

Schooling for dissent? Education, autocratic regime instability and transitions to democracy

Sirianne Dahlum

s.a.dahlum@stv.uio.no



Department of Political Science

Faculty of Social Sciences

University of Oslo

February 2017

Contents

List of Figures	vii
List of Tables	xi
1 Linking education to political instability in dictatorships	1
1.1 Introduction	1
1.1.1 The contribution	1
1.1.2 Education as a plausible determinant of democratization	4
1.2 Revisiting modernization theory: In search for micro-foundations	7
1.2.1 The micro-foundations of democratization	7
1.2.2 Existing literature	11
1.2.3 Moving beyond average education levels: Three examples	18
1.3 Concepts	23
1.3.1 Democracy and autocracy	23
1.3.2 Variations in autocratic regimes	26
1.3.3 Democratic transitions and autocratic breakdown	28
1.3.4 Contentious collective action	30
1.3.5 Education	32
1.4 Opportunities and preferences for anti-regime mobilization	33
1.4.1 The actors	34
1.4.2 Obstacles to collective action	35
1.4.3 Mechanism I: Opportunities for anti-regime protest	36
1.4.4 Mechanism II: Anti-authoritarian grievances	37
1.5 Data and units of analysis	41
1.6 Methodological considerations	42
1.6.1 Methodological foundations	43
1.6.2 Methodological challenges	48
1.7 Structure of the thesis	55

1.7.1	Part I: The reciprocal relationship between education quality and democratization	55
1.7.2	Part II: Disaggregating the relationship between education and authoritarian regime instability	56
1.8	Findings	57
1.9	Implications	59
1.9.1	Implications for future research	59
1.9.2	Policy implications	61
I	The reciprocal relationship between education quality and democracy	63
2	What do they learn? Education quality and democratization	65
2.1	Introduction	66
2.2	Moving from education quantity to quality	68
2.3	Why education quality and content should matter	72
2.3.1	The capacity-enhancing effect of education	72
2.3.2	The democratic-preferences effect of education	73
2.4	Data	75
2.4.1	Education quality	75
2.4.2	Education autonomy and type	78
2.4.3	Dealing with missing values	79
2.4.4	Regime change and democratization	79
2.4.5	Controls	80
2.5	Empirical analysis	80
2.5.1	Education quality and democracy	80
2.5.2	Education type and education autonomy	85
2.5.3	Education quality and autocratic breakdown	87
2.6	Implications: Why quantity rather than quality matters	90
2.7	Conclusion	91
3	Do democracies provide better education? Revisiting the democracy–human capital link	93
3.1	Introduction	95
3.2	Democracy, education, and education quality	98
3.3	Why democracy may not enhance education quality	99
3.3.1	Democracy and incentives to improve education quality	100
3.3.2	Democracy and obstacles to improving education quality	102
3.3.3	Some suggestive evidence and observations	104
3.4	Data, measures and validity issues	104
3.4.1	Measuring education quality	104
3.4.2	Data and measures for democracy and control variables	107

3.5	Empirical analysis	108
3.5.1	Democracy and measures of “education quantity”	109
3.5.2	Main analysis: Democracy and education quality	111
3.5.3	Robustness tests on democracy and education quality	115
3.5.4	Extensions and some nuances	118
3.6	Conclusion	120
II	Disaggregating the relationship between education and authoritarian regime instability	123
4	Schooling as a “dangerous good”? Education, mass protest and authoritarian regime stability	125
4.1	Introduction	126
4.2	Education and democratization	128
4.3	Education and authoritarian regime stability	131
4.4	Educated protesters and prospects for autocratic breakdown	133
4.5	Data and empirical model	135
4.5.1	Dependent variables	135
4.5.2	Independent variables	136
4.5.3	Confounders	137
4.6	Empirical analysis	138
4.6.1	Country-level variation in education	138
4.6.2	Variation in education levels between protest campaigns	142
4.7	Exploring the mechanisms	146
4.8	Conclusion	150
5	Educating Demonstrators: Education and Mass Protest in Africa	151
5.1	Introduction	152
5.2	Education, participation and protest	154
5.3	Linking education to popular protests	157
5.3.1	Motivations	157
5.3.2	Opportunities	158
5.3.3	Expectations	161
5.4	Data	162
5.4.1	Covariates	165
5.5	Empirical strategy	166
5.5.1	Baseline estimates	168
5.5.2	Exploring mechanisms	172
5.5.3	Robustness and additional tests	174
5.6	Conclusion	175

6 Students in the streets: Education and nonviolent conflict	179
6.1 Introduction	180
6.2 Education, violence and collective action	182
6.2.1 Violent vs. nonviolent mobilization	182
6.2.2 Education, collective action and regime change	184
6.3 How education affects political protest	184
6.3.1 Education and capacity for nonviolent protest	185
6.3.2 Education and preferences for nonviolent protest	187
6.4 Data	189
6.4.1 Campaign participants' education profile	189
6.4.2 Campaign outcomes	194
6.4.3 Control variables	194
6.5 Results	199
6.5.1 Education and nonviolent protest	199
6.5.2 Education and successful outcomes	201
6.6 Empirical extensions	205
6.7 Conclusion	207
III Appendices and Bibliography	209
7 Appendix to Chapter 2	211
7.1 Introduction	211
7.2 Distinguishing between democratic transitions and survival	212
7.3 Checking the robustness of the quality-democracy link: Dealing with endogeneity	216
7.4 Extending the analysis of interaction effects between education quality and enrollment	220
7.5 Details about the multiple imputation procedure	227
7.5.1 Imputation model and diagnostics	227
7.5.2 Assessing the sensitivity of the models to the multiple imputation procedure	232
7.6 Additional robustness tests	237
7.6.1 Alternative lag structures	237
7.6.2 Additional controls	240
7.6.3 Subcomponents of education quality	242
7.6.4 Alternative dependent variable	244
7.6.5 5-year panel	246
7.6.6 Education quality and democracy without a lagged dependent variable	247

8 Appendix to Chapter 3	249
8.1 Introduction	249
8.2 The education quality data	251
8.3 Investigating mechanisms	253
8.4 Imputation model and diagnostics	259
8.5 Robustness tests: Democracy and measures of “education quantity”	265
8.6 2SLS models: Discussion, first-stage regressions for core specifications, and alternative 2SLS models using different instruments	268
8.7 Other robustness tests: Democracy and education quality	273
8.8 Additional robustness tests on education quality: Models run on imputed data	279
8.9 Mediation analysis on democracy, education, and growth	282
8.10 Extension: Democracy and education quality by subject and level of education	285
8.11 Extensions: Tests on variance between and within countries, by regime type, and tests on within-population variability in PISA 2012 test scores	287
9 Appendix to Chapter 4	291
9.1 Introduction	291
9.2 Replicating the education-democracy relationship	292
9.3 Country-level education in authoritarian regimes	294
9.4 Educated protest campaigns in authoritarian regimes	298
9.5 Empirical extensions	310
10 Appendix to Chapter 5	313
10.1 Overview	313
10.2 Evaluating the relative importance of education	313
10.3 Descriptives	315
10.4 The data	317
10.5 Additional controls	321
10.5.1 Alternative controls for local wealth	321
10.5.2 Literacy and media consumption	323
10.5.3 Additional proxies for urban settings	324
10.5.4 Different types of land cover	325
10.5.5 Models with spatial lags	327
10.6 Dropping cells with few respondents (< 30 respondents)	329
10.7 Education in Anglophone versus Francophone countries	330
11 Appendix to Chapter 6	333
11.1 Introduction	333
11.2 Codebook for indicators of education	334
11.3 Summary statistics for main independent and dependent variables	337
11.4 Sensitivity tests leaving out observations coded as uncertain	338

11.5 Alternating control variables	340
11.6 Alternating the operationalization of the education index	344
11.7 Alternative specifications of general sensitivity analysis	353
11.8 Robustness tests of the relationship between education and successful outcome	357
11.9 Robustness tests of mediation analysis	368
11.10 Education profile of campaign leadership	368
11.11 Multiple imputation of missing values	375
11.12 List of campaigns included in the analysis	377
12 Bibliography	387

List of Figures

1.1	Developments in global education attainments (share of population with no schooling) and democracy, 1900-2010	13
1.2	Electoral democracy and average years of schooling by period, 1900-2010. Kernel density estimates.	14
2.1	Schooling achievements (from Angrist (2013)) and mean years of education (from Barro-Lee (2013))	77
2.2	Schooling achievements (from Angrist (2013)) and the Polity Democracy index	78
3.1	Average years of schooling (from Barro and Lee (2010)), and “schooling achievements” (from Angrist (2013)).	106
3.2	Democracy and residualized mean student achievement test scores.	112
4.1	Predicted probabilities of autocratic breakdown	146
5.1	Grid cells with DHS education data and protest events overlaid, 1990-2012	165
5.2	The Location of mission stations in colonial Africa (Based on data from Nunn, 2011)	168
5.3	Simulated probability of protest events	171
6.1	Involvement by the educated in protest campaigns.	196
6.2	Involvement by students in protest campaigns.	197
6.3	Involvement by the educated in violent and nonviolent campaigns.	198
6.4	Involvement by students in violent and nonviolent campaigns.	198
6.5	The predicted probability of nonviolence	202
6.6	The mediation effect of education on successful outcome, mediated through nonviolence	204
6.7	Assessing the threat of omitted variable bias	207

7.1	Distributions of observed and imputed observations for aggregated measure of student achievement test scores	228
7.2	Overimputation test plot, aggregated measure of student achievement test scores	230
7.3	Overimputation test plot, student achievement test measure for primary school students	230
7.4	Overimputation test plot, student achievement test measure for secondary school students	231
8.1	The mediation effect of democracy on primary education enrollment, mediated through education spending. Parsimonious model specification.	256
8.2	The mediation effect of democracy on education quality, mediated through education spending. Parsimonious model specification.	258
8.3	Distributions of observed and imputed observations for aggregated measure of student achievement test scores	260
8.4	Distributions of observed and imputed observations for student achievement test measure for primary school students	261
8.5	Distributions of observed and imputed observations for student achievement test measure for secondary school students	261
8.6	Overimputation test plot, aggregated measure of student achievement test scores	263
8.7	Overimputation test plot, student achievement test measure for primary school students	263
8.8	Overimputation test plot, student achievement test measure for secondary school students	264
8.9	Mediation effects of democracy on economic growth, mediated through education quality	283
8.10	Mediation effects of democracy on economic growth, mediated through education quantity	284
9.1	Predicted probabilities for uneducated campaigns vs highly educated campaigns	303
11.1	Assessing the threat of omitted variable bias. Coefficient estimate = 0	354
11.2	Assessing the threat of omitted variable bias. Coefficient estimate = 0. Estimation using logit model.	354
11.3	Assessing the threat of omitted variable bias. Coefficient estimate is halved. Estimation using logit model.	355
11.4	Assessing the threat of omitted variable bias. Coefficient estimate is halved. Simple model.	355
11.5	Assessing the threat of omitted variable bias. Coefficient estimate is halved. With region dummies.	356
11.6	Sensitivity analysis of mediation effect	369

11.7 The mediation effect of education on successful outcome, mediated through nonviolence. Simple model.	370
11.8 Sensitivity analysis of mediation effect. Simple model	371
11.9 The mediation effect of education on successful outcome, mediated through nonviolence. Categorical education index.	372
11.10 Sensitivity analysis of mediation effect. Categorical education index.	373

List of Tables

2.1	Democracy and education quality	83
2.2	Democracy, education type and education autonomy	86
2.3	Education quality and authoritarian breakdown.	89
3.1	Democracy and average years of schooling	110
3.2	Democracy and mean student achievements test score	113
3.3	Democracy and mean student achievements test score: Imputed data . .	118
4.1	Changes in electoral democracy and school enrollment.	139
4.2	Changes in electoral democracy and regime breakdown in authoritarian regimes	140
4.3	Changes in electoral democracy and regime breakdown in authoritarian regimes. Campaigns by educated protesters	143
4.4	Predicted probabilities of autocratic breakdown, given an ongoing protest campaign.	145
4.5	Participation by educated citizens in protest movements	147
5.1	Baseline probit and OLS models of protest events	169
5.2	IV-probit and 2SLS models of protest events	170
5.3	Models exploring mechanisms	173
6.1	Non-violent outcome on campaign involvement by students and graduates (logit models)	200
6.2	Successful outcome on campaign involvement by students and graduates (logit models)	203
7.1	Education quality. Transitions to democracy and autocratic survival . . .	213
7.2	Education type and education autonomy. Transitions to democracy and democratic survival	215

7.3	Democracy and education quality. 2SLS models where quality is instrumented using historical levels of numeracy skills	217
7.4	First-stage regressions with primary education quality as DV	218
7.5	First-stage regressions with secondary education quality as DV	219
7.6	Democracy and education quality. Adding interaction effects between quality and enrollment.	221
7.7	Democracy and education quality. Adding interaction effects between quality and years of schooling. Combined education quality indicator.	222
7.8	Distinguishing between democratic transitions and democratic survival. Interacting education quality with enrollment. Combined education quality indicator.	223
7.9	Education quality and authoritarian breakdown. Combined average education quality indicator.	224
7.10	Primary education quality and authoritarian breakdown. Interacting education quality with primary enrollment.	225
7.11	Secondary education quality and authoritarian breakdown. Interacting education quality with secondary enrollment	226
7.12	Democracy and education quality. 5-year panel.	233
7.13	Democracy and education type. Unimputed dataset.	233
7.14	Democracy and autonomy. Unimputed dataset.	234
7.15	Democracy and education quality. Estimations across all 5 imputed datasets, with MI-corrected standard errors	234
7.16	Democracy and education type. Estimations across all 5 imputed datasets, with MI-corrected standard errors	235
7.17	Democracy, education type and education autonomy. Full imputed time series to deal with selection bias.	236
7.18	Democracy and education quality (5-year lag)	237
7.19	Democracy and education quality (10-year lag)	237
7.20	Democracy, education type and education autonomy (5-year lag)	238
7.21	Democracy, education type and education autonomy (10-year lag)	239
7.22	Democracy and education quality. Additional controls.	240
7.23	Democracy, education type and education autonomy. Additional controls	241
7.24	Democracy and primary education quality	242
7.25	Democracy and secondary education quality	242
7.26	Democracy and subcomponents of primary education quality.	243
7.27	Democracy and different subcomponents of secondary education quality	243
7.28	Democracy (Freedom House) and education quality	244
7.29	Democracy, education type and education autonomy. Freedom House as DV.	245
7.30	Democracy and education quality. 5-year panels	246
7.31	Democracy and education quality. Without lagged democracy.	247

8.1	Cross-country regressions on democracy and education spending	254
8.2	Cross-country regressions on education spending and education quality .	257
8.3	Democracy, as measured by the Freedom House Index (average of the Political Rights and Civil Liberties indices), and average years of schooling .	265
8.4	Democracy and primary school enrollment rates	266
8.5	Democracy and secondary school enrollment rates	266
8.6	Democracy and mean years of schooling. Alternative estimators	267
8.7	First-stage regressions (with Polity Index as dependent variable) for Fixed Effects 2SLS models presented in Table 3.1 of the paper.	268
8.8	First-stage regressions (with Polity Index as dependent variable) for Fixed Effects 2SLS models presented in Table 3.2 of the paper.	269
8.9	Democracy and mean student achievements test score: Second-stage regressions from alternative FE 2SLS specifications using different instruments	272
8.10	Democracy and mean student achievements test score: Alternative estimators	273
8.11	Democracy, as measured by the Freedom House Index(average of political and civil liberties indices), and mean student achievement test scores. .	274
8.12	Democracy, as measured by the Democracy–Dictatorship (DD) measure, and mean student achievement test scores.	274
8.13	Democracy and mean student achievements test score. 10-year lag.	275
8.14	Democracy and mean student achievement test scores, checking for a non-linear relationship by including a squared term for Polity.	275
8.15	Democracy and mean student achievement test scores, checking for a non-linear relationship by using autocracy, partial democracy, full democracy categorization from Epstein et al. (2006).	276
8.16	Democracy and mean student achievement test scores, checking for an interaction between democracy and income level.	276
8.17	Democracy and mean student achievement test scores, checking for an interaction between democracy and income inequality.	277
8.18	Democracy and mean student achievement test scores, controlling for quality of government.	277
8.19	Democracy and mean student achievement test scores, checking for an interaction between democracy and quality of government.	278
8.20	Democracy (reference category), autocracy types and mean student achievement test scores.	278
8.21	Democracy and mean student achievements test score	279
8.22	Democracy and first differences in mean student achievements test score	280
8.23	Democracy and mean student achievements test score. 5-year lag	280
8.24	Democracy and mean student achievements test score. 10 year lag	281
8.25	Democracy and subcomponents of student test scores. Random effects models.	286

8.26 Democracy and subcomponents of student test scores. OLS (Fixed effects) models.	286
8.27 Goldfeld-Quandt test results, based on residuals from cross-section OLS (between-country variance, 2005–2009) and OLS fixed effects regression (within-country variance, all time periods).	287
8.28 Democracy and standard error in PISA 2012 test scores within countries' student populations	288
8.29 Democracy and absolute gender difference in PISA 2012 test scores	289
8.30 Democracy and mean PISA 2012 score	289
 9.1 Changes in democracy score and education. Average year of education as explanatory variable.	292
9.2 Changes in democracy score and education. Polity as dependent variable	292
9.3 Changes in democracy score and education. Polity index as dependent variable and average year of education as explanatory variable	293
9.4 Democratic upturns and education.	293
9.5 Democratic upturns in authoritarian regimes.	294
9.6 Democratization and regime breakdown in authoritarian regimes. Sample consisting of autocracies according to Geddes et al.	295
9.7 Democratization in authoritarian regimes. Democratic transitions according to Boix, Rosato and Miller (2013)	295
9.8 Democratization and regime breakdown in authoritarian regimes. Education measured using average years of education from Barro-Lee.	296
9.9 Democratization and regime breakdown in authoritarian regimes. Additional control variables	296
9.10 Democratization and regime breakdown in authoritarian regimes. Controlling for authoritarian regime type, from Geddes, Wright and Frantz (2014).	297
9.11 Democratic upturns in authoritarian regimes.	298
9.12 Democratization and regime breakdown in authoritarian regimes. Sample consisting of autocracies according to Geddes et al.	299
9.13 Democratization in authoritarian regimes. Democratic transitions according to Boix, Miller and Rosato (2013)	300
9.14 Democratization and regime breakdown in authoritarian regimes. Controlling for years of education.	301
9.15 Democratization and regime breakdown in authoritarian regimes. Including additional control variables.	302
9.16 Democratization and regime breakdown in authoritarian regimes. Controlling for authoritarian regime type	304
9.17 Democratization and regime breakdown in authoritarian regimes. Independent variables lagged by five years.	305

9.18	Democratization and regime breakdown in authoritarian regimes. Protest campaigns with educated participants described as having participated.	306
9.19	Democratization and regime breakdown in authoritarian regimes. Protest campaigns dominated by educated participants.	308
9.20	Democratization and regime breakdown in authoritarian regimes. Alternative education index.	309
9.21	Democratization and regime breakdown in authoritarian regimes. Controlling for governance index.	311
9.22	Democratization and regime breakdown in authoritarian regimes. Controlling for governance index.	312
10.1	Substantive effects of different predictors	314
10.2	Descriptives for variables used in baseline models	315
10.3	Countries included and democracy scores (Polity)	316
10.4	Countries and time-periods included in the DHS	318
10.5	SCAD protest events in our sample (column 2) and all countries (in our sample) over 1990-2011 period (column 3)	320
10.6	Probit and OLS models of protest events	321
10.7	IV Probit and 2SLS models of protest events	322
10.8	Probit, OLS, IVProbit and 2SLS models of protest events	323
10.9	Probit, OLS, IVProbit and 2SLS models of protest events	324
10.10	Probit and OLS models of protest events	325
10.11	IV Probit and 2SLS models of protest events	326
10.12	Probit and OLS models of protest events	327
10.13	IV Probit and 2SLS models of protest events	328
10.14	Probit and OLS models of protest events	329
10.15	Probit models of protest events	331
11.1	Summary statistics	337
11.2	Non-violent outcome on campaign involvement by educated protesters (logit models). Leaving out cases for which coding was considered uncertain.	338
11.3	Successful outcome on campaign involvement by educated protesters (logit models). Leaving out cases for which coding was considered uncertain.	339
11.4	Non-violent outcome on campaign involvement by educated protesters (logit models). Leaving out one control at a time.	341
11.5	Non-violent outcome on campaign involvement by educated protesters (logit models). Without controlling for membership size.	342
11.6	Non-violent outcome on campaign involvement by educated protesters (logit models). Additional controls.	343
11.7	Non-violent outcome and whether campaign originated among educated protesters (logit models)	345
11.8	Non-violent outcome and whether campaigns had at least some educated protesters (logit models)	346

11.9 Non-violent outcome and whether campaign was dominated by educated protesters (logit models)	347
11.10 Non-violent outcome on campaign involvement by graduates (logit models)	348
11.11 Non-violent outcome on campaign involvement by student protesters (logit models)	349
11.12 Non-violent outcome on campaign involvement by educated protesters (logit models). Separate indicators for students and graduates, included in the same model	350
11.13 Non-violent outcome on campaign involvement by educated protesters (logit models). Dichotomous education index.	351
11.14 Non-violent outcome on campaign involvement by students and graduates (logit models). Multiplicative education index.	352
11.15 Successful outcome and whether campaigns had some educated protesters (logit models)	358
11.16 Successful outcome and whether campaigns were dominated by educated protesters (logit models)	359
11.17 Successful outcome and whether campaigns originated among educated protesters (logit models)	360
11.18 Successful outcome on campaign involvement by graduates (logit models)	361
11.19 Successful outcome on campaign involvement by students (logit models) .	362
11.20 Successful outcome on campaign involvement by educated protesters (logit models). Multiplicative education index.	363
11.21 Successful outcome on campaign involvement by educated protesters (logit models). Dichotomous education index.	364
11.22 Successful outcome on campaign involvement by educated protesters. Multi- nomial logit models (base category = failure)	365
11.23 Successful outcome on campaign involvement by educated protesters (logit models). Omitting one control at a time.	366
11.24 Successful outcome on campaign involvement by educated protesters (logit models). Additional controls	367
11.25 Non-violent outcome on campaign involvement by students and graduates (logit models)	374
11.26 Non-violent outcome on campaign involvement by students and graduates (logit models)	375
11.27 Successful outcome on campaign involvement by graduates (logit models)	376

Acknowledgement

Pursuing a Phd has been a wonderful experience. I was warned that Phd life could be tough and discouraging, but my verdict is that it was surprisingly fun. Although I am glad to have finally arrived, I am also slightly sad to leave behind such an ideal job and way of living - with the freedom to read, write and learn about (almost) anything, the possibility of stretching my analytical abilities every day, and the opportunity to interact with brilliant people from around the world. I am very grateful for having had this opportunity, and for everyone who have contributed along the way. There are several people I would like to thank.

First and foremost, I would like to thank my main supervisor, Carl Henrik Knutsen, who has taught me how to be a scholar. He has shaped my academic development ever since giving an undergraduate lecture on the relationship between economic development and democratization in 2008. The lecture sparked my initial interest in this topic, that has later materialized into one bachelor thesis, two master dissertations and finally a doctoral degree. As a Phd supervisor, Carl Henrik has never been short of invaluable advice on how to do research, how to survive in academia, and how to generally make good choices. He has from the very beginning treated me not only as a student but also as a trusted collaborator, introducing me to the high standards and methods of world class political science. Together with the rest of our small ?research family? at the University of Oslo we have embarked on numerous research projects, op-eds, academic debates, data collections, travels (both work and holiday) and late night discussions. I am looking forward to continue our collaboration! Thanks also to my co-supervisor, Håvard Hegre, who joined later in the process, but who has given valuable input.

I am also grateful to my academic home throughout the past 4 years, the Department of Political Science at the University of Oslo, for offering the institutional backing and safety net that has allowed me to freely pursue the project I wanted. I would also

like to thank my colleagues at the University of Oslo, including senior members of the Comparative Politics section, who have taken an active interest in my work at seminars and other occasions. I am particularly grateful to have been surrounded by a great group of fellow Phd students and post docs from the very first day, who have created a welcoming atmosphere at the department.

There is no doubt that the quality of this dissertation would have been substantively poorer if it had not been for those who have commented at various workshops and conferences, helping to expose flaws that I have subsequently tried my best to fix. Thank you to Sven-Erik Skaaning, Agnes Cornell, Jørgen Møller, Milan Svolik, Silje Lyder Hermannsen, Tatjana Stankovic, Kacper Szulecki, Henning Finseraas, Zoltan Fazekas, Andreas Kot-sadam, Andreas Hvidsten, Solveig Hillestad, Cristina Bucur, Helge Holtermann, Magnus Rasmussen, Kalle Moene, Raino Malnes, Fenja Søndergaard Møller, Jacob Tolstrup, Lasse Lykke Rørbaek, Bjørn Høyland, Håvard Strand, Elin Allern, Kristian Gleditsch, Scott Gates, Håvard Nygård, Marianne Dahl, Charles Butcher, James Hollyer, Suthan Krishnarajan, David Delfs Erbo Andersen and Casper Sakstrup for commenting on one or more of the different articles comprising this thesis on various occasion, in addition to other participants at APSA, EPSA, MPSA, and NOPSA. I would particularly like to thank Tore Wig for offering important advice and feedback at every stage of the dissertation.

I am also grateful to the editors and anonymous reviewers from World Development, Journal of Conflict Resolution and Comparative Political Studies for providing detailed comments, that are either already incorporated or currently being incorporated in a revised version.

I would also like to thank a group of fellow political scientists who first crossed my path during our undergraduate years and who all have continued to pursue a Phd at some point, for offering a stimulating environment filled with intellectual curiosity that always inspires me to learn more and think harder. I would particularly like to thank Aksel, for always sharing new ideas, and my fellow Phd traveler Magnus, for teaching me that trade unions are the key to understanding everything and that all Phd challenges can be solved with a bottle of fruit beer.

Although UiO has been my base throughout these four years, a substantial share of the next pages were written in other places. I would like to thank Columbia University (where I spent a semester), Metochi Study Center at Lesvos (where I had two writing retreats), Albergo del Vino in Alba, my parents cabin at Gålå, and my neighborhood cafe Liebling for hosting me and offering valuable breathing space and time to write. Finally, I never would have been this motivated to study mass mobilization and authoritarian regimes if it had not been for my one year adventure in Teheran, in which I got to observe the (failed) Green Revolution up close. The strong disappointment among students and

intellectuals in Teheran when the green movement was crushed is a vivid memory, and I hope that liberal Iranians will be able to successfully mobilize to promote the desired political change in the future.

I would also like to thank my friends for making Oslo the best city in the world. Finally, I would like to thank my closest family - my parents and my brother, Jens Martin - for supporting me in all my pursuits. Thank you to my father, for encouraging me to work hard, and my mother, for urging me to work less.

Last but not least, I am deeply grateful to Tore, travel companion, colleague, collaborator, (almost) office neighbour, best friend and boyfriend. Thank you for your endless support at every part of the journey, for doing the dishes, and for always reminding me why we are doing research. You are the major reason why the past four years were so wonderful.

Oslo, July 2017

Sirianne Dahlum

1 Linking education to political instability in dictatorships

1.1 Introduction

1.1.1 The contribution

In 1900, the overwhelming majority of the world’s population had never set foot in a school. Education was a benefit reserved for an elite minority, primarily residing in the most developed parts of the world. Today, the large majority spends at least parts of its childhood in school. While almost 8 in 10 of the world’s population had not received *any* formal education in 1900, this number was reduced to 1 in 10 by 2010 (Barro and Lee, 2013). Hence, it is not unwarranted that the 20th century has been referred to as the “century of education” (Goldin and Katz, 2009; Morrisson and Murtin, 2009). The boom in schooling access has altered people’s lives and opportunities, through knowledge, cognitive resources and more autonomous jobs, but it has also had some profound impacts on society and politics. In particular, it is commonly presumed that an increasingly educated population will eventually lead to the demise of authoritarian regimes, as educated individuals will be more averse to dictatorship and gradually more capable of carrying out organized resistance to secure political inclusion and civil liberties (e.g., Diamond, 2015; Glaeser, Ponzetto and Shleifer, 2007).

This dissertation investigates whether mass education actually induces popular mobilization, authoritarian regime breakdown and democratic transitions. The notion that education is a threat to autocratic regimes aligns with probably the most influential account of democratization, referred to as modernization theory. A prominent advocate of

this explanation, Seymour Martin Lipset (1959), argued that socio-economic development promotes transitions from autocracy to democracy, and that education is one of the key components in this development. He argued that "If we cannot say that a high level of education is a sufficient condition for democracy, (...) it comes close to being a necessary condition" (Lipset, 1959, p.80). Prompted by this thesis, several studies investigate the relationship between indicators of education and transitions to democracy (see, e.g., Benavot, 1996; Barro, 1999; Acemoglu et al., 2005; Murtin and Wacziarg, 2014; Sanborn and Thyne, 2014). Although the findings are somewhat mixed, most of these contributions argue and present corroborating evidence that education promotes democratization. In line with these findings, there seems to be widespread agreement among policy-makers and in the development community that promoting education is an important step to fostering democratization.

Against this backdrop, I take issue with current findings on the link between education and democratization. In a nutshell, this dissertation indicates that education may be a force for political instability in authoritarian regimes and democratic transitions, but *not* in the way assumed and tested in existing literature: Contrary to many previous findings, I do not find much evidence that highly educated autocracies are more likely to experience regime downfall and democratization. Rather, there is some evidence that these regimes are in fact *more stable* than autocracies with low education levels. Yet, *when* educated citizens first mobilize and protest, they are more likely to promote authoritarian regime breakdown and transitions to democracy.

More specifically, I take issue with what I argue are too simplified arguments and highly aggregated tests of how education influences prospects of regime change in authoritarian regimes. Close to all existing studies rely on empirical analyses of the macro relationship between one or more country-level indicator of access to education on the one hand, and democratization on the other hand. Yet, this approach fails to capture the processes and dynamics that lead to political instability and regime change in autocracies: First, with a few exceptions, existing studies do not offer comprehensive theoretical accounts for why education should promote democratization.¹. This includes accounting for the mechanisms taking place at the *micro or meso level*, that tie education to the choices and behavior of individuals or groups. As I argue below, the most plausible mechanisms relating education to regime change are concerned precisely with the role of agents, but this is not sufficiently incorporated in existing theoretical frameworks or research designs. Second, this highly aggregated empirical approach not only fails to shed light on the micro-mechanism, it may also produce biased results, as it relies on macro-level tests to draw inferences regarding mechanisms unfolding at the level of groups of individuals.

¹For exceptions, see Glaeser, Ponzetto and Shleifer (2007) and Campante and Chor (2012b)

This is problematic for several reasons. For instance, an aggregate education indicator hides substantial within-country variation in education, both geographically and between groups. And, an aggregate education indicator may capture country-level processes or characteristics (that are hard to measure) other than merely the education-level of the population, for instance related to processes of reverse causality or the capacity or inherent stability of authoritarian regimes.

With this research gap in mind, my thesis is a plea for a closer attention to the micro-foundations of the proposed relationship between education and regime change, both theoretically and empirically. *Theoretically*, I detail the mechanisms explaining why education should induce democratization. This involves shifting the focus towards events and processes occurring *within authoritarian regimes*, which may promote authoritarian regime instability and breakdown. Here, I build on recent scholarship on authoritarian regime survival, suggesting how aspects of socio-economic development may serve to the benefit of dictators, by suggesting how schooling provision may also boost popular support and promote autocratic survival (e.g., Kennedy, 2010; Bueno de Mesquita and Smith, 2010; Miller, 2012; Bueno de Mesquita and Smith, 2003). It also involves detailing the processes through which education should influence the behavior of relevant individuals or group that may conduct anti-regime mobilization in authoritarian regimes. Here, I build on literature on violent uprisings and collective mobilization, by focusing on how education influences the *preferences* and *opportunities* for anti-regime mobilization (Gurr, 1970; Lichbach, 1995; Collier and Hoeffler, 2004; Cederman, Gleditsch and Buhaug, 2013). I argue that education should increase the capacity for collective action, through enhanced skills and knowledge, and raise anti-authoritarian grievances and thereby preferences for regime change. This should raise the likelihood of anti-regime mobilization, as well as the likelihood that such mobilization succeeds at overthrowing dictators and promoting democracy.

Empirically, I develop research designs that can accommodate this more nuanced theoretical framework, drawing on different types of data, models and levels of aggregation. Three important steps are taken to offer empirical tests that are more aligned with the proposed (micro-)mechanisms:

- The traditional empirical focus on democratic transitions is complemented with studies of contentious politics in authoritarian regimes and authoritarian regime breakdown – regardless of whether this breakdown leads to democratization.
- Macro-level tests are supplied with more disaggregated research designs, using new data from different levels of analysis to study the role of education at the level of actors and sub-national areas. For instance, I offer the first test of whether protest

movements consisting of educated groups or individuals are more likely to succeed at overthrowing authoritarian regimes. And, I am the first to explore whether sub-national areas (such as districts) with a high concentration of educated people are more likely to experience political protests. This approach is important because it allows investigating the micro-mechanisms at the level of analysis where they are expected to play out, thereby mitigating ecological fallacies.

- More nuanced tests are offered also at the national level, through the first investigation of whether education quality and variations in schooling contents matter for the likelihood of democratization, and vice versa.

This set of studies produces several novel insights regarding the relationship between education and prospects for regime change in autocracies. To highlight some of the key findings, there is little evidence that increases in schooling access or education quality at the national level promote autocratic breakdown or democratization in authoritarian regimes.² Rather, the evidence suggests that states with high education levels on average have a lower likelihood of authoritarian breakdown. I suggest several reasons for this, including that high education levels are symptomatic of authoritarian regimes' capacity – both for offering public services for staying in power, and that schooling provision is a “signal” by the dictator of his commitment to citizen's welfare, which may strengthen popular support. At the same time, I find that when educated groups and individuals first initiate and organize anti-regime protest campaigns, this increases the likelihood of authoritarian breakdown. Further, I find that this may be due to educated protesters being more likely than other protesters to choose nonviolent resistance, which again is more likely to produce successful protest outcomes. Hence, variation in education levels between anti-regime actors matters for the likelihood of autocratic breakdown and transitions to democracy, but not variation in education between countries. Finally, more disaggregated indicators of education at the sub-national level suggest that areas with high concentration of educated people are more likely to experience protest. I return to and elaborate upon these and additional findings at the end of this chapter.

1.1.2 Education as a plausible determinant of democratization

The research gap pointed to above, relating to the lack of attention to relevant micro-mechanisms, applies to most extant empirical literature on modernization theory. This

²I find some indications that average years of achieved education and school enrollment are positively related to changes in democracy in all regimes, but this is not very robust and disappears when including country fixed effects and/or year dummies. Meanwhile, when looking exclusively at autocracies in chapter 4, there is evidence that education is negatively related to changes in democracy.

includes numerous studies investigating the relationship between income level and democratization. Yet, the effect of education on regime change is particularly understudied, all the while both theoretical expectations and policy-considerations suggest that the role of education deserves particular attention.

First, theoretical considerations point to education as perhaps the most plausible driving force behind the strong correlation that exists between socio-economic modernization and democracy. The large majority of extant studies investigating modernization theory have analysed the relationship between indicators of income level to measure socio-economic development, and democracy, yielding mixed results (see, e.g. Londregan and Poole, 1996; Boix, 2003; Acemoglu et al., 2008; Przeworski and Limongi, 1997; Moral-Benito and Bartolucci, 2012; Heid, Langer and Larch, 2012; Treisman, 2011; Kennedy, 2010; Miller, 2012). Yet, in line with Lipset, it has been noted that a high income level per se is not the most plausible explanation for why socio-economic modernization should promote democratization. Importantly, GDP per capita may be a poor proxy of real levels of economic modernization. A high income level is found in a great variety of societies, with very different economic and social structures, including those that do not even resemble what we associate with modernized economies. For instance, the oil-producing Arab countries now have very high income levels, but have not necessarily experienced the changes in society or the economy that we associate with economic modernization. In addition, a high income level could also be a source of strength for authoritarian regimes, which can utilize access to increasing state revenues to boost regime survival (see, e.g. Bueno de Mesquita and Smith, 2010).

Education, on the other hand, arguably captures to what extent economic development actually empowers and benefits citizens. Education offers knowledge and skills that not only the dictator but also the recipients of education will benefit from. It can in many ways be considered a “coordination good”, which can “determine the ability of citizens to coordinate and organize” (Bueno de Mesquita and Smith, 2010, p.938). At the same time, it is less likely to strengthen authoritarian survival prospects than high GDP levels, the latter being a resource that dictators can tap directly into as a means of buying off support or repressing potential opposition. In this sense, education is intuitively a much more plausible candidate as a driving force of democratization than GDP per capita, which captures so much more than citizen empowerment and does not necessarily indicate that the economy is genuinely modern and developed.³

Second, this thesis is motivated by the fact that the assumption that education breeds democracy has been deeply influential in the work of policy-makers, international organi-

³Other aspects of socio-economic modernization, such as level of industrialization and urbanization, are also not as easily rewarding to the dictator as a high income level. At the same time, these developments can not as directly be considered a resource to potential protesters as education.

zations and development agencies. Efforts to expand access to education in less developed societies have been regarded as a crucial step in democracy promotion by international organizations such as the UN and the World Bank, and this has been manifested in key governing documents such as the UN Resolution on Education for Democracy 2012. This assumption is sustained by scholars and journalists as well, who frequently point to education as a likely trigger for political change and democratization around the world. For instance, the head of the US National Endowment for Democracy, Larry Diamond (2015, p.314), predicts that “China cannot keep moving forward to the per capita income, educational, and informational levels of a middle-income country without experiencing (...) pressures for democratic change”. Moreover, discussing Tunisia, Filipe Campante (2014) argues that “With its relatively high levels of education, its secular traditions, and its constitutional progress since the fall of Zine el-Abidine ben Ali’s dictatorship in 2011, Tunisia has a real chance to consolidate democracy”.

Third, as opposed to other proposed determinants of democracy, such as income level, urbanization and industrialization, education is a variable that, at least theoretically, can be directly manipulated by policy-makers and social engineers. Not only can policy-makers influence the levels of investments in education and thereby influence access to schooling, they can also influence the contents and design of education. Hence, knowledge about the effects of education access and content on regime stability or change is directly applicable to policy-making.

Finally, access to education is still increasing rapidly in all parts of the world – including in the most authoritarian regimes – and there is nothing indicating that this development will stop or slow down. The question of whether education promotes collective action, authoritarian instability and democratic transitions, at least in the long-run, is therefore also a question about the future and viability of today’s authoritarian regimes.

These considerations suggest that it is crucial to identify exactly how and in what ways education impacts on autocratic regime endurance and democratization. Moreover, in order to respond to the strong assumptions held among policy-makers and researchers, it is crucial to go beyond simply detecting correlations between education and democracy, but also establishing how education impacts on the relevant agents and processes of political instability in dictatorships.

The remainder of the chapter proceeds as follows: Section 1.2 situates the thesis within extant literature, and elaborates on how it adds to previous theoretical and empirical studies. Section 1.3 continues by discussing the main concepts and assumptions that the thesis builds upon, and section 1.4 presents the overarching theoretical framework that all the subsequent chapters elaborate on. Some considerations regarding data and units of analysis used in the different chapters are presented briefly in section 1.5, and section 1.6

discusses the methodological foundations and issues related to choice of research design and model specifications. In section 1.7, I lay out the structure of the rest of the thesis and briefly summarize the following chapters. Finally, section 1.8 discusses some implications of the findings for future research and policy-making.

1.2 Revisiting modernization theory: In search for micro-foundations

This section situates my main contribution within existing literature. I start by elaborating on one of the overarching themes of the dissertation: The plea for a more disaggregated empirical test of modernization theory, that is better aligned with the supporting micro-mechanisms. Second, I discuss existing literature and its shortcomings. In brief, I show that many studies seem to build on explicit assumptions about micro-mechanisms, but these assumptions are rarely elaborated on explicitly or tested systematically. Finally, I discuss three cases that illustrate the need for a more disaggregated approach than the reliance on state-level indicators of education levels in existing literature allows.

1.2.1 The micro-foundations of democratization

As I elaborate on below, current empirical scholarship on the education - democratization link mainly relies on evidence in the form of investigations of the relationship between indicators of average education levels on the hand and indicators of democracy or regime transitions on the other hand. At the same time, this literature generally pays little attention to how education is thought to produce the events and mobilization processes that induce regime change. Hence, extant studies largely fail to test and investigate the *micro-mechanisms* underpinning the proposed relationship, which can be understood as the causal process through which agents bring about the social change. Indeed, this highly aggregated approach also characterizes current empirical literature on other socio-economic determinants of democracy such as income levels (see, e.g. Acemoglu et al., 2008; Przeworski and Limongi, 1997; Moral-Benito and Bartolucci, 2012; Heid, Langer and Larch, 2012; Faria, Montesinos-Yufa and Morales, 2014; Gundlach and Paldam, 2009). This has also been pointed out by proponents of so-called agent-centered explanations of democratization (e.g. O'Donnell and Schmitter, 1986). For instance Rustow (1970, p.340) argued that the structural literature investigating the economic correlates of democracy neglects the “genetic question of how a democracy comes into being”.

At the same time, and as I elaborate on below, the most plausible reasons for why

democracy should emerge from education are not related to societies' *education levels*. Rather, it involves educated individuals or groups carrying out behavior that is conducive to regime change, such as forms of collective action. This assumption seems to be shared by most scholars studying the relationship between education and regime change. Those that do outline a theoretical account almost always invoke mechanisms focusing on the role of educated individuals or groups. For instance, in their account of why education should promote democratization, Sanborn and Thyne (2014, p.777) point to how schooling alters individuals values, teaching students "how to work together and engage in democratic practices", and how education is "empowering a growing middle class". Hence, a complete study of this relationship should specify and test how education influences relevant actors or groups, including how it alters their preferences, opinions, choices and behavior.

Positing and investigating micro-explanations is already a widely recognized goal in social science, as noted by Gerring (2011, p.164):

"These days, interpretivists, case-study researchers, experimentalists, neo-behavioralists, formal modelers and epistemologists seem to agree on at least one thing (and perhaps only one thing): the traditional focus of social science has been too large, too highly aggregated and too influenced by – often spurious – associations between X and Y. Greater leverage on causal questions can be gained only by scaling down. We need to get inside the box of causation".

On a similar note, Little (2005) argues that social science should be concerned with how social outcomes are brought about by the actions and dispositions of the relevant agents.

Yet, the quantitative empirical literature on democratization and regime change has not sufficiently emphasized this perspective. In this regard, the empirical democratization literature is out of touch with developments in other areas in political science.⁴ For instance, the related field of civil conflict research presents theoretical mechanisms tying structural factors such as economic development and poverty to agents' preferences and choices. For example, the "greed theory" details how relatively poor people are more likely to join civil war, due to lower opportunity costs, while the "grievance theory" emphasizes how relatively deprived people are more motivated for violent conflict due to frustration with current conditions (Gurr, 1970; Collier and Hoeffler, 2004; Fearon and Laitin, 2003).

Moreover, there is by now an extensive empirical (and quantitative) research agenda that integrates determinants of civil war, such as socio-economic and geographic factors, with their micro-foundations (Kalyvas, 2006; Cederman and Gleditsch, 2009; Weidmann,

⁴The qualitative and case study literature, however, pays much more attention to the role of agents in bringing about democratization (see prominent contributions such as Moore, 1966; Rueschemeyer, Stephens and Stephens, 1992; Collier, 1999). Yet, this literature is rarely concerned with factors such as education or socio-economic modernization. And, it rests heavily on the selection of certain cases.

2009). This literature criticizes the large body of studies on the effects of structural indicators such as GDP per capita or urbanization on civil war for looking at “national aggregates and averages that are only loosely linked to their rationale for conflict and the postulated micro-level mechanisms” (p.488 Cederman, Gleditsch and Hug, 2013). Following up on such criticisms, numerous studies have analyzed the relationship between socio-economic factors such as poverty (or relative poverty) in a disaggregated manner, either relying on sub-national units of analysis such as districts, regions or “grid-cells”, (see, e.g. Hegre, Østby and Raleigh, 2009; Buhaug et al., 2011; Holtermann, 2012; Fjelde and Østby, 2014; Berman et al., 2011), or looking at the individual or group level (e.g. Cederman, Wimmer and Min, 2010; Humphreys and Weinstein, 2008; Arjona and Kalyvas, 2006).

Although the *empirical* democratization literature has rarely studied micro-mechanisms, prominent contributions in political economy have developed more nuanced *theoretical* models of democratization, based on rational choice models where the preferences and utility calculations of individuals are invoked. These studies have been particularly concerned with modeling how economic inequality will influence the prospects of democratic transitions, by shaping agents’ preferences for redistribution and democracy (Acemoglu and Robinson, 2006; Boix, 2003; Ansell and Samuels, 2014). Yet, although these mechanisms are specified at the agent-level, the accompanying empirical tests mainly rely on macro-level tests of the relationship between various indicators of economic inequality and democratic transitions (Boix, 2003; Ansell and Samuels, 2010, 2014). Moreover, there is actually little empirical evidence in favor of the main macro-implications of such theories, for instance that economic inequality is systematically related to democratization (Houle, 2009; Teorell, 2010; Knutsen, 2015). Finally, these theoretical models have not incorporated education as an explanatory factor.

Of course, the actual changes in institutions that make up a transition from autocracy to democracy happen at the nation-state level, and this may explain why the focus in existing literature has been on the nation-state as unit of analysis. Although there may be sub-national variations in local institutions (see, e.g. Wig and Tollefsen, 2016), democracy as defined and operationalized in this thesis is indeed a state-level phenomena. Yet, most of the numerous processes and developments that *precede* and explain this change unfold at a more disaggregated level, through for instance groups of people planning a coup, individuals realizing that it is no longer rational to support the autocratic incumbent, individuals becoming more favorable towards democracy, or groups of people voicing their opposition through strikes or demonstrations. This suggests a theoretical and empirical turn towards elaborating and testing micro-mechanisms of theories of democratization.

In sum, the thesis takes issue with what I argue is a too vague and highly aggregated

approach to the education-democratization link in existing empirical literature. On this background, I elaborate and test an argument linking education to authoritarian breakdown and democratization through the effect it has on the *behavior and preferences of citizens and groups*. Different implications of this argument is explored across five empirical studies, with the outcome of yielding a richer and more nuanced picture of the education-democratization nexus. To exemplify, this involves asking questions such as: Are educated protesters more likely to promote regime change? Are (sub-national) areas with a high concentration of educated individuals more likely to experience anti-regime protests?

The theoretical focus is on mechanisms taking place in authoritarian regimes, as this is where the events and process thought to produce regime breakdown and transitions to democracy should unfold. In brief, I discuss how education raises the *capacity* of individuals to organize and carry out anti-regime resistance, as well as nurtures *preferences* for more democracy and civil liberties that stimulates grievance against authoritarian regimes. Due to these developments, educated citizens should be more likely to challenge regimes collectively, and more likely to successfully overthrow regimes. I also consider possible feedback effects from regime type to education, including how dictatorships may use education policies as strategic tools to strengthen regime survival.

Empirically, all contributions in the thesis investigate different implications of this theoretical framework. *First*, I shift the empirical focus from only studying democratic transitions as outcome variable, as has been common practice in existing literature, to also studying outcomes such as authoritarian breakdown and contentious politics in authoritarian regimes (chapters 4-6). This follows recent literature noting that in order to understand the roots of democratization, it is necessary to also study processes of authoritarian instability and regime breakdown independently of its outcome – whether it is a new democracy or a new autocracy (Kennedy, 2010; Miller, 2012).

Second – and this is one of the main contributions of the thesis – I supply macro-level tests with empirical investigations using new data from different levels of aggregation, including at the protest movement level and at the sub-national level. This is the main uniting theme of chapters 4-6. For instance, I offer the first test of whether protest movements dominated by educated groups or individuals are more likely to succeed at overthrowing authoritarian regimes than other protest movements (chapter 6). Moreover, I offer the first test of whether sub-national areas with higher education levels are more likely to experience onsets of political protest (chapter 5).

This disaggregated approach has several advantages. As the proposed mechanisms are expected to play out at the micro or meso level, it allows empirical tests at a more appropriate level of analysis, closer to where the actual mechanism is unfolding. Further-

more, an empirical test restricted to the macro level may produce ecological fallacies. For instance, an aggregate measure of education such as the average years of schooling in the population hides substantial within-country variation, both in terms of who receives the education and how it is distributed. It does not tell us whether the actors promoting authoritarian instability are educated or not, and these actors may very well not be representative for the social profile of the population at large. For instance, we might expect that the prospects for regime change is likely to depend on the strongest groups, or the most educated actors, rather than on those with the average education profile.

An average state-level indicator of education can also capture other characteristics of societies and regimes than simply variations in how many people are educated. For instance, a very high level of school enrollment can signal that the regime has a very high capacity, or that it is invested in providing services and welfare benefits to its citizens. Furthermore, a high enrollment level may signal that the dictator wants to quell grievances in the population, and that education has been offered as a means to mitigate social unrest and potential political instability. Hence, education levels are likely endogenous to both political instability and regime type, at the same time as it may be a symptom of many other state-level processes or characteristics.

Third, I offer a more nuanced empirical test at the state level, using indicators of education that are more closely aligned with the micro-mechanisms. A measure of the average amount of schooling offered to the population does not capture variations in schooling quality and contents. Yet, as I discuss at length in subsequent chapters, the proposed micro-mechanism assume that education should influences the preferences and opportunities of individuals through *what they learn* in schools. To account for this, I analyse the relationship between indicators of education quality and variations in learning outcomes on the one hand and regime change on the other hand. This is done in chapters 2-3.

1.2.2 Existing literature

The argument

The notion that an educated population is essential to well-functioning democracy can be traced back to early political theorists, such as Montesquieu, Locke, Rousseau, Mill and Aristotle, who argued that education empowers individuals and societies, and promotes a culture of enlightened public debate and effective citizenship. After the emergence of modern democratic regimes, this notion was carried forward by scholars tying education to the functioning of existing democracies. For instance, Dewey (1938) considered an educated population as a prerequisite for democracy, as schooling should stimulate the

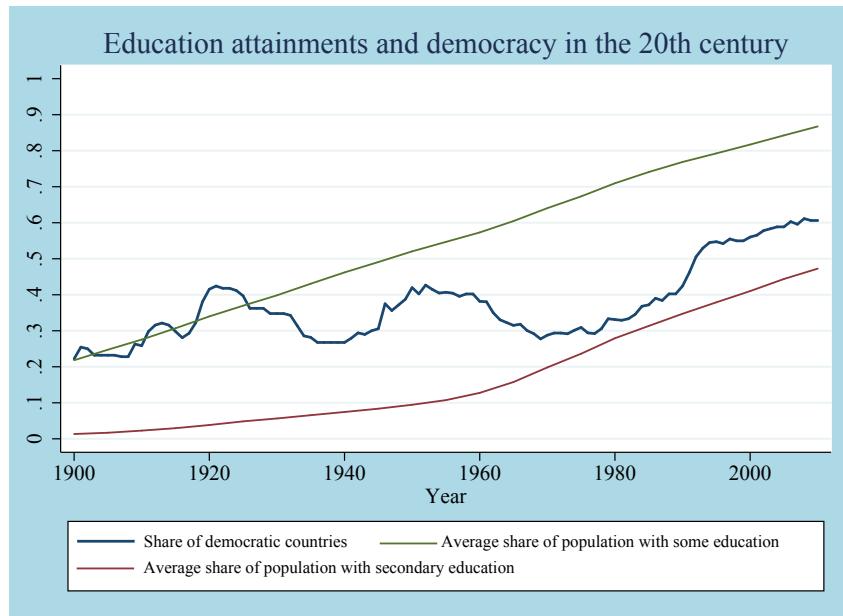
individual growth and reflection necessary to function as democratic citizens. Lerner (1958) also pointed to education as one of the core factors in the process of individual modernization and political participation.

Building on this, Lipset's (1959) famous version of modernization theory argued that socio-economic development is the key explanation for the emergence and survival of democracies. This proposition has been argued to be "undoubtedly one of the most well-known propositions in comparative politics" (Clark and Linzer, 2015). Lipset considered modernization as a major restructuring of society, consisting of developments such as urbanization, industrialization and work differentiation. In his argument, education was considered as perhaps the most crucial component of this development. To support the theory, Lipset presented empirical evidence suggesting that GDP per capita, used as a proxy for socio-economic modernization, is strongly related to democracy. This finding has inspired a massive body of literature investigating empirically the relationship between GDP per capita and democracy, drawing on a wide range of sophisticated estimation methods (see, e.g. Londregan and Poole, 1996; Acemoglu et al., 2008; Przeworski and Limongi, 1997; Moral-Benito and Bartolucci, 2012; Heid, Langer and Larch, 2012; Faria, Montesinos-Yufa and Morales, 2014; Gundlach and Paldam, 2009; Treisman, 2011; Kennedy, 2010; Miller, 2012).

Yet, as noted, Lipset's main concern was not with changes in income levels per se, and there are several reasons why changes in GDP per capita should not be expected to lead to democratization. First, an indicator of GDP may be a poor proxy of the socio-economic modernization process that is suggested to influence regime type. As pointed out in recent literature on democratization and authoritarian regimes, many countries today that experience very high GDP per capita levels, such as oil and gas exporting countries, have not necessarily gone through the socio-economic developments associated with modernization (Ross, 2001, 2012; Clark, Golder and Golder, 2012). Second, an increase in GDP per capita is a development that may also strengthen the authoritarian regime, and help to prolong regime survival due to increased access to state revenues (Kennedy, 2010; Bueno de Mesquita and Smith, 2010; Haggard and Kaufman, 2008; Treisman, 2011). Democratization should be expected to happen through socio-economic changes that strengthen civil society and potential protesters on the expense of the dictator. An increasing education level represents such a development: Education should empower citizens and civil society in ways conducive to authoritarian regime instability and democratization, through promoting democratic preferences and raising the capacity to mobilize against dictators.⁵

⁵Other proposed determinants of democracy according to modernization theory, that are investigated empirically, are, e.g., urbanization (Wallace, 2013), media proliferation (Teorell, 2010) and so-called "youth bulges" (Cincotta, 2008).

Figure 1.1: Developments in global education attainments (share of population with no schooling) and democracy, 1900-2010



The share of democracies in the world is calculated based on the dichotomous democracy measure by Boix, Miller and Rosato (2013). The share of citizens with no schooling and secondary schooling is calculated based on data from Barro and Lee (2013).

Empirical findings

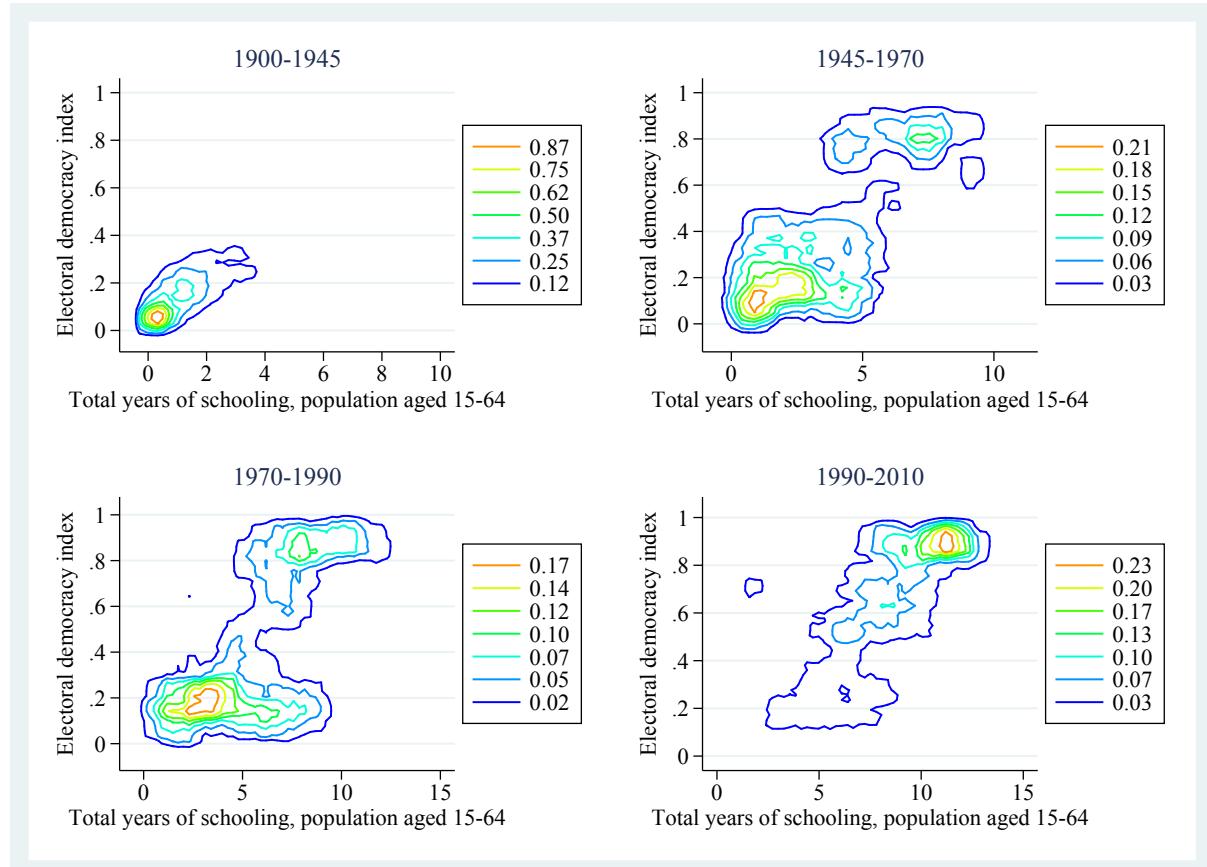
Despite the theoretical expectations, the hypothesis that education promotes democratization currently has a mixed empirical record.

On the one hand, it seems clear that education levels are correlated with (changes in) democracy. This is illustrated in figure 1.1, which shows the developments in global education attainments throughout the 20th century. It clearly illustrates how the share of population with at least *some education* has increased dramatically throughout the century. Moreover, although the share of the population with secondary schooling is at a much lower level, it has increased substantially as well. In the same period, the share of countries in the world that are democratic (rather than autocratic) has more than tripled.

Figure 1.2 illustrates joint movements in democracy and schooling over time. Here, the bivariate distribution between mean years of education and democracy levels from 1900-2010 is presented using Kernel densities.⁶ The color scheme illustrates the density of data points in regions of the graph. The data is broken down into four different sub-

⁶Kernel density estimation is a non-parametric method for measuring the *density* of observations on one or several variables, and can be considered a smoothed version of a histogram, the latter obtained by counting the observations in different bins. Kernel density estimates are easier to interpret than histograms, since they do not contain artifices induced by a binning grid.

Figure 1.2: Electoral democracy and average years of schooling by period, 1900-2010. Kernel density estimates.



The electoral democracy index is from the Varieties of Democracy dataset (Coppedge et al., 2015), and data on years of education is from Barro and Lee (2013). The colored contours represent degrees of density. The kernel density estimates are calculated using the “bivariate” package in Stata 14, using an Epanchenikov kernel with bandwidth adjusted to the finite sample size.

periods, and the sequence of graphs illustrates the joint developments in schooling levels and democracy over time. In brief, it shows that most societies were autocratic and highly uneducated until the end of World War II, indicated by the fact that one small area at the bottom left-hand side of the graph has the highest density. Over the next sub-period (1945-1970), a group of countries became democratic and their education levels jumped, but most countries remained in the lower left-hand side of the distribution. After 1970 and up until 1990, there was a large increase in variation between non-autocracies when it comes to education levels, as many autocracies became much more educated while others did not experience much development. During the last sub-period (1990-2010), most countries joined the club of democratic countries with high education levels, and the relationship between education levels and democracy is quite linear.

In line with this bivariate pattern, numerous studies present results from regressions

showing indicators of education are positively associated with both democracy levels and democratization (Benavot, 1996; Barro, 1999; Papaioannou and Siourounis, 2008; Sanborn and Thyne, 2014). For example, Benavot (1996) presents evidence that elite education is positively related to democracy, while Barro (1999), on the other hand, finds that primary education rather than higher education is the explaining factor. Another example is Murtin and Wacziarg (2014), who in a thorough study based on time-series data extending back to 1870 shows that education is a much stronger and more robust predictor of democracy than income.⁷

Yet, critics of modernization theory have questioned whether the education-democracy correlation is due to a causal effect from the former to the latter. Acemoglu et al. (2005) challenges the positive findings, using fixed effects models to argue that the positive correlation from previous studies is most probably due to a spurious relationship driven by omitted variables. The perhaps strongest disagreement in the recent education-democracy debate has to do with causal direction. There is ample evidence that democratization may increase access to education, and numerous studies find that democracy enhances educational enrollment and years of schooling (see, e.g. Lake and Baum, 2001; Lindert, 2005; Ansell and Samuels, 2010; Engerman and Sokoloff, 2005; Harding and Stasavage, 2014; Acemoglu, Gallego and Robinson, 2014). Moreover, there is evidence that education is among the key mediators through which democracy enhances economic growth (Baum and Lake, 2003; Tavares and Wacziarg, 2001; Doucouliagos and Ulubasoglu, 2008)⁸.

In line with these findings, several scholars argue that both the introduction of mass schooling and economic growth were originally induced by democratic or so-called “inclusive institutions”, rather than the other way around (Acemoglu and Robinson, 2001; Acemoglu et al., 2005; Stasavage, 2005; Acemoglu, Gallego and Robinson, 2014). To support this argument, Acemoglu and Robinson (2001) utilize exogenous sources of persistent variations in political institutions that emerged due to European colonizers in the modern world setting up different form of institutions, depending on the feasibility of setting up settlements. They argue that in countries where settlements were infeasible, colonizers set up “extractive institutions” which involved, for instance, labor coercion or lack of property rights. Colonizers in more feasible areas set up “inclusive institutions”, that were much more conducive to economic growth and human capital investments. Using potential settler mortality rates to instrument for institutions, Acemoglu et al find evidence in line with the claim that institutions influence economic development and education. Acemoglu, Gallego and Robinson (2014) also use the historical presence of

⁷These positive findings are corroborated by a recent sensitivity analysis indicating that education is one of the most robust predictors of democratization measured as changes in polity score (Hegre et al 2012).

⁸But, see also Knutsen (2011b)

Christian mission stations to instrument for human capital formation, assuming that protest missionaries played an important role in setting up schools (see also, Woodberry, 2012). Relying on this strategy, they demonstrate that human capital does not have an effect on democratic institutions.

Acemoglu et al's rejection of modernization theory has again been disputed, however. For instance, Castelló-Climent (2008) challenges Acemoglu et al's use of fixed effects models to demonstrate that education does not promote democracy. The former argue that these estimators are inappropriate when variables are highly persistent, such as education and democratic institutions. Using arguably more appropriate estimators such as system gmm models, and accounting for the distribution of education in the population, Castello-Climent finds that the positive relationship between education and democracy is robust (see also Bobba and Coviello, 2007).

Prominently, Acemoglu et al's claim that education is mainly a product of political institutions is also criticized by Glaeser and colleagues in several studies (Glaeser et al., 2004; Glaeser, Ponzetto and Shleifer, 2007; Campante and Glaeser, 2009) (see also, Murtin and Wacziarg, 2014), arguing that Acemoglu and Robinson (2001) is putting the "institutional cart before the human capital horse". For instance, Glaeser et al. (2004) criticizes the identification strategy of Acemoglu and Robinson (2001), arguing that crucial differences between European colonies were not the type of institutions introduced, but rather variations in human capital, which again promoted democratic institutions. Thus, there is an ongoing debate regarding the empirical merits of the education-democratization link.

Education is, of course, likely to be strongly related to other aspects of socio-economic development such as urbanization, industrialization and youth cohorts. A timely question is whether education is just a symptom of other aspects of socio-economic developments, that are really driving democratization, or whether education is the fundamental cause of modernization. In response to this, studies indicate that human capital development is indeed "trigger" that sent countries off to different political and economic development paths (Easterlin, 1981). Hence, according to this view, education is an important determinant of other components of economic development.

Mechanisms

Most large-N studies documenting the education-democracy link do not explicitly specify a causal mechanism, but many seem to hinge on implicit (and untested) assumptions about such a mechanism. Glaeser, Ponzetto and Shleifer (2007) offers the perhaps most comprehensive account of why education should produce regime transitions. They argue that education should raise the benefits (or reduce the costs) of political protest, which

again should promote democratization. Moreover, Sanborn and Thyne (2014) propose that education may produce transitions to democracy because “the highly educated are apt to take to the streets”. Finally, Campante and Chor (2012b, p.169) suggest that education should promote authoritarian instability through promoting “political activities of a protest nature”. Hence, these studies all point to anti-regime mobilization or collective action as the mechanism through which education should induce authoritarian instability and regime change, assuming that education will promote such events.

This mechanism has rarely been tested in a large-N framework, however. A large body of literature has documented that education is strongly associated with *conventional forms of political participation in democracies*, such as voting, contacting politicians, joining political parties and participating in community meetings (see e.g. Almond and Verba, 1963; Wolfinger, 1980; Putnam, 1993; Hanushek, 2002; Milligan, Moretti and Oreopoulos, 2004; Dee, 2004; Hillygus, 2005; Bratton, Mattes and Gyimah-Boadi, 2005). Yet, we know less about how and whether the education-participation link plays out in authoritarian regimes. And, it is rarely studied how education influences “contentious collective action”, which includes behavior seeking to challenge existing institutions rather than channeling demands within the established political framework. A couple of exceptions are studies by Campante and Chor (2012b) and Campante and Chor (2012a), that argue and demonstrate empirically – both at the individual level and the country level – that education combined with macroeconomic weakness is likely to spur protest. For instance, they show that individuals whose income falls below the predicted value based on their achieved education are more likely to engage in protest activities and that education coupled with unemployment is likely to spur uprisings.

There is an extensive literature devoted to explaining social movements and their outcomes, offering theoretically rigorous contributions discussing the internal life and dynamic of social movements, focusing on factors such as identity formation, organization and leaderships (e.g. Tarrow, 1994; Tilly and Tarrow, 2006). Yet, this strand of research has not paid much attention to whether socio-economic or structural factors (such as education) influence the dynamic or outcome of social movements. Moreover, there is a recent quantitative literature that utilizes new data on various characteristics of protest movements to explain their outcomes and prospects for success. But, this literature mainly focuses on strategic aspects – such as the interaction between protesters and the target, the size of the protest and competition between different factions – rather than structural explanations (e.g., Chenoweth and Stephan, 2011; Celestino and Gleditsch, 2013; Cunningham, 2013b).⁹

⁹Two exceptions are Butcher and Svensson (2014) and Chenoweth and Ulfelder (2017). The former studies the relationship between industrialization and nonviolent resistance, while the latter investigates several state-level structural variables as determinants of nonviolent protest. Yet, none of these investigate

In sum, existing literature on the education–regime change relationship offers little more than highly aggregated studies of the effects of average education indicators on democratic transitions. While collective action is commonly assumed to be the most prominent mechanism, this is rarely tested.

It has often been proclaimed that it is time to disregard modernization theory, especially by those arguing forcefully that political institutions promote economic development rather than the other way around. For instance, Acemoglu and Robinson (2013) argue that “despite its great intuitive appeal (especially to the well-educated middle classes already inclined to see themselves as the harbinger of all good things including democracy), Lipset’s modernization theory just doesn’t have empirical support”. Reviewing existing studies of modernization theory, Munck (2016, p.48) concludes that “despite repeated testing of some key hypotheses, results are not positive. Instead of converging in support of modernization theory, they are decidedly mixed”. He goes on to argue that this lack of unified support suggests that it is time to put the theory to rest. However, as I have pointed out, there are potentially crucial shortcomings in the literature, suggesting that the relationship between socio-economic development and democratization has not yet been put to its best test. The lack of findings could be due to a true null-hypothesis, but it could also be due to empirical tests that are too simplified to capture this relationship. This suggests that before being put to rest, modernization theory should be given a new chance – through the kind of comprehensive and nuanced testing required to pick up mechanisms occurring at the micro- or meso level.

1.2.3 Moving beyond average education levels: Three examples

In this section I discuss some cases that illustrate one of the main overarching themes in this dissertation: That empirical tests of the education–democratization link using aggregate indicators of access to education may hide substantial variation and relevant nuances. And, these missing nuances may in fact be the key to understanding how education is related to mass mobilization and authoritarian regime instability. This is illustrated with three examples of events taking place in Indonesia, Malawi and Qatar.

Revolution in Indonesia (1998)

In the mid-1990s, Indonesia was a seemingly stable dictatorship, headed by dictator and president Suharto who had been in charge for 31 years, after coming to power in a military coup in 1967. At that point, several other large Asian autocracies had experienced regime breakdown and transition to democracy during the past decade, including South Korea in

education as an explanatory variable, neither at the state or at the protest level.

1987 and the Philippines in 1986. Yet, these countries were more economically developed than Indonesia, who had suffered from bad economic policies and political hesitance to invest in education policies (Boudreau, 2009). For instance, in 1995 the Indonesian enrollment rate in tertiary education was only 2.2 percent, and the average years of education in the population was 4.9. In South Korea and the Philippines, 23 and 24 percent were enrolled in tertiary education, and the average citizen had attended school for 11 and 8 years respectively (Barro and Lee, 2013). Hence, a true believer in modernization theory would probably not predict that Indonesia would experience authoritarian instability or transition to democracy any time soon. Yet, in February 1998 mass protests broke out in Jakarta, which soon spread throughout the country and culminated in Suharto's exit from power.

At first glance, these state-level education statistics suggest that the Indonesian case does not comply with modernization story, and that the 1998 revolution must have been triggered by other factors than education or socio-economic development. However, when considering more closely the events and protests that led up to Suharto's choice to step down, it seems that *educated protesters* were crucial to both initiating and coordinating the revolution. As noted by most records of the events that took place in Indonesia in 1998, the first protests that kicked off the mass movement took place at the country's universities, initiated by university students who campaigned for democracy. Throughout the spring of 1998 these protests erupted. Not only did they spread to universities all over Indonesia, they were also joined by other social groups such as artists, journalists and other professionals. Yet, the students evidently still played an instrumental role when it came to orchestrating and coordinating the protests: "Recognizing the stagnancy the opposition found itself in, the students took decisive action to bring protest to a national, confrontational level, demanding Suharto's removal" (Swarthmore, 2015).

There are several potential mechanisms through which the student profile of the protesters may have influenced the outcome of the 1998 events, as suggested by the theoretical framework in this dissertation (detailed in section 1.4). First, the fact that many of these protesters were students or educated may explain why they chose to protest in the first place, as educated individuals arguably have stronger anti-authoritarian attitudes. Second, the high education levels among many of the protesters may explain why the protest movement succeeded in promoting regime change, due to the skills and resources of educated citizens. If these mechanisms played out, education did influence the downfall of the Indonesian military regime, despite the low average education level in the country.

In chapter 6, I follow up on this point by studying variations in education levels between protesters, presenting evidence that protest campaigns consisting of students

and educated groups are more likely to use nonviolent resistance and to accomplish their goals.

Political instability in Malawi (2002)

The Indonesian example illustrates the importance of looking to the actual groups and individuals participating in anti-regime mobilization. At the same time, there may be huge *geographical variations* within a certain country – both when it comes to education levels and the occurrence of political protest events – that are disguised by average state-level education indicators.

This can be illustrated by developments in Malawi in 2002. At that time, the country had undergone a series of democratic reforms since the dictator of 30 years, Kamuzu Banda, had stepped down almost a decade earlier. Yet, the country was still far from a full-fledged democracy, experiencing violations such as election “irregularities” and alleged fraud, imprisonment of journalists and police brutality – including e.g. arbitrary arrests and violence against opposition sympathisers (House, 2003). As a result, Malawi in 2002 was considered only partly free by Freedom House. Meanwhile, the country was highly uneducated. As many as 49 percent of citizens had never been to school at all, and the enrollment rate in higher education institutions was as low as 0.4 percent (Barro and Lee, 2013). Hence, the country did not have the type of highly educated population that – in line with modernization theory – should be expected to carry out pro-democracy protests to push for further democratization.

Despite this, a number of anti-regime protests broke out in 2002, partly directed against perceived lack of democracy and respect for civil liberties. In addition, protesters mobilized against a proposition by President Bakili Muluzi and his supporters to alter the constitution to allow the president to run for a third term in the elections scheduled for 2003. Yet, the protests were not spread evenly across the country, but were contained to a few areas or cities. According to the Social Conflict in Africa Database, protests broke out in June in the city of Zomba, in October in Mzuzu and in November in the city of Blantyre (Salehyan et al., 2012). According to survey data, all of these three areas or cities have education levels that are much higher than the national average levels: In Zomba, the enrollment in tertiary education was 25 percent, while the same number was 18 percent in Mzuzu and 17 percent in Blantyre (Afrobarometer, 2008).

Hence, if Malawi in 2002 had been included in a country-year regression analysis of the relationship between education and political protests, this data point would have pulled the coefficient estimate down, downplaying the effect of education on protest. Yet, a closer look at the relevant events suggest that educated individuals likely played an important role in the 2002 protest: For instance, the protests in Zomba consisted of

students and academics, who organized rallies at the university campuses and across the city. Moreover, the protests against the suggested constitutional amendments in Blantyre was organized under the The Forum for the Defence of the Constitution (FDC), consisting of e.g. doctors, journalists and pastors from the Presbyterian Church.

The Indonesian and Malawian cases illustrate how low average education levels may hide crucial within-country variation, which may be relevant to explaining regime opposition. Hence, tests using state-level education indicators may downplay the role of education in promoting authoritarian instability. Sub-national variation in education and protest is returned to in chapter 5, which demonstrates that sub-national areas in Sub-Saharan Africa with high education levels are more likely to experience anti-regime protest. Below, I round off this section by discussing how a *high* level of access to education at the state level may *overplay* the potential for democratic change due to an educated population.

Authoritarian stability in Qatar (2010s)

With its constant inflow of state revenues from natural resources and its enduring authoritarian government, Qatar is the typical gulf state. The al-Thani clan has ruled Qatar for nearly 200 years, currently through the Emir Tamim bin Hamad al-Thani as head of state. On most accounts, the Qatari kingdom is considered highly resilient. This can be illustrated by the fact that the country did not experience any tumults at all throughout the course of the 2011 Arab spring, that shook many of its regional neighbors. At the same time, the country has invested huge amounts in the education sector, and the Qatari population is by now highly educated.¹⁰

There are several potential plausible reasons for why the Qatari dictatorship remains so resilient, despite the highly educated citizenry. Most accounts have appropriately pointed to the country's seemingly inexhaustible pool of natural resource rents and the fact that it has the highest GDP per capita in the world (Ross, 2012). Yet, another feature may also be relevant, especially for the purpose of understanding the relationship between education and democracy; namely the surprisingly low quality of Qatari education. Despite being the richest country in the world and spending vast amounts of resources on the education system, Qatari students perform far below average in terms of student achievement tests. For instance, in a recent report by OECD, Qatar came in 68th place out of 75 countries when it comes to average scores in student achievement tests in reading and maths (Hanushek and Woessmann, 2015). Teachers lacking formal qualifications and high levels of student absenteeism were pointed to as potential expla-

¹⁰In 2015, the average years of schooling in the country was 9.3. Only 2.8 percent of the population was fully uneducated (Barro and Lee, 2013)

nations for the poor learning achievements. This means that while education access is high, Qatar performs worse in terms of student learning outcomes than countries such as the Dominican Republic, Swaziland and Mozambique, that are much poorer and have much lower education levels. This is in line with trends in other Gulf countries, according to Bahgat (1999), arguing that “the number of individuals with formal schooling is rising, but the quality of the education they receive does not correspond to the need of Gulf societies”. During the recent decade, the Qatari government has placed a lot of emphasis on improving the quality of education, in line with international campaigns and emphasis on learning outcomes. Yet, the country is still underachieving majorly in terms of student achievements (Angrist, 2013).

This suggests that the real human capital of Qatar is overplayed when looking exclusively at indicators of education *quantity*. Although Qatari students spend many years in schools, they simply do not learn much more than severely poorer countries with considerably lower enrollment rates. The theoretical framework in this dissertation indicates that education may promote anti-regime mobilization authoritarian instability by strengthening skills and resources and inducing preferences for democracy. Yet, if the schooling offered is of a very poor quality, these resources and preferences may not develop.

Of course, this is far from the only explanation for the endurance of the Qatari monarchy. Indeed, this example may actually serve to illustrate another point made in this dissertation, related to state-level indicators of education and state capacity. In brief, the high education enrollment rates may capture the high capacity of the Qatari regime. The Emir is in position where he can buy off the population through social spending, inflated earnings for public employees, investments in culture and entertainment and public services – including education. The regime arguably also has high capacity when it comes to controlling information in order to discourage political opposition, through government ownership and censorship of the mass media, surveillance of citizens and the threat of sanctions against government criticism, which again promotes high levels of self-censorship (House, 2013). If high education enrollments are simply a proxy for the regime’s capacity - both when it comes to offering public services, buying off citizens and controlling information, it should not necessarily be expected to promote authoritarian instability.¹¹

Variation in schooling quality and contents is returned to in chapters 2 and 3, which study whether learning outcomes matter for democratization, and vice versa. Discussions related to state-level indicators of education access, state capacity and authoritarian

¹¹Another feature that is not captured by the mean education level statistics is the fact that both education access and quality are probably distributed highly unequally in Qatar, with the rich elite (consisting of local Qataris) likely to receive much more and better education than for instance immigrant workers.

survival are included in chapter 4, which presents evidence that high nationwide education levels are associated with a lower likelihood of authoritarian breakdown.

1.3 Concepts

In linking education to political instability and regime transitions, I make a number of conceptual distinctions. This section presents and discusses some fundamental conceptual distinctions that the contributions in the thesis build upon. I start by discussing key aspects related to the *outcomes* of interest, and how they relate to each other: 1) Democracy and autocracy, 2) Variations in autocratic institutions, 3) Democratic transitions and autocratic breakdown and 4) Contentious collective action. Finally, I discuss the main explanatory variable – education.

1.3.1 Democracy and autocracy

This dissertation is concerned with changes in political institutions along one specific dimension, namely variations in the extent to which political institutions are democratic rather than autocratic. The term democracy is used interchangeably as a label for a wide range of political institutions and ideals. For instance, it is commonly used to describe societies with low socio-economic inequality, citizen participation, elections, the realization of the public interest, popular sovereignty, political rights or constitutional checks on executive or legislative powers (Held, 2006; Coppedge, 2012). This illustrates the importance of pinning down a more exact definition of democracy. In order to define and operationalize democracy it is insufficient to simply point to a list of political institutions that characterize or are associated with countries considered democratic. We need to offer an answer to the question of why some institutions are *considered* democratic, based on the essence of democracy is and what core principles it aims to realize.

In this study, democracy is conceptualized as a way of *establishing popular control over political decision-making*. This is in line with the definition offered by Beetham (1992, p.40):

“Democracy I take to be a mode of decision-making about collectively binding rules and policies over which the people exercise control, and the most democratic arrangement to be that where all members of the collectivity enjoy effective equal rights to take part in such decision-making directly – one, that is to say, which realizes to the greatest conceivable degree the principles of popular control and equality in its exercise”.

According to this definition, democracy is a political system that realizes both popular control over decision-making *and* political equality. Autocracy, by definition, is a political system where either popular control over decision-making or political equality is lacking. Hence, autocracy is defined as a “residual category”, with reference to the definition of democracy (see, e.g., Svolik, 2012).

Moreover, democracy is conceptualized as lying at one end of a spectrum, with a system where people are totally excluded from decision-making lying at the other end of the spectrum (Beetham 1992). This is in line with Dahl (1971), who described democracy in its ideal form as *polyarchy*. In reality, all regimes will usually fall short of this ideal, but I assume that a regime can still be considered democratic if the extent of popular control over decision-making reaches a certain level. In practice, the ideal of popular control does not require direct democracy in the Greek sense – requiring all decisions to be made by the electorate. In large, complex societies it is sufficient that the populace is indirectly in control through elections of their preferred representatives (see, Dahl, 1971).

The literature offers plenty of thorough discussions of what institutional components are necessary to fulfill popular control over decision-making (for thorough discussions, see Munck and Verkuilen, 2002; Knutsen, 2011b; Coppedge, 2012). Here, I will only briefly suggest some relevant institutional components, well aware that the various democracy measures used throughout the thesis vary somewhat in which components they emphasize. In line with common descriptions of democratic institutions, I take *contested elections* to be at the core of democracy, as this is crucial to realizing popular control over decision-making in a representative democracy. As noted by Dahl (1971) genuine “contestation” requires not only free and fair elections, but also that citizens are free to organize themselves into fractions, present their desired political program and compete for political power (see Teorell et al., 2016).

Political contestation is not, however, sufficient to ensure popular control over decision-making. As noted by Dahl, and recognized by most democracy scholars, we should also look to the *inclusiveness* of the political system. This refers to the extent to which all citizens are allowed to participate in the political process and whether contestation is restricted to only a few or the majority. Other important requirements, also discussed by Dahl (1971), are that the elected are the ones that can actually influence decision-making and that citizens can access diverse sources of information about politics.

This democracy conception requires more than simply elections, it is restrictive in the sense that control over decision-making is the defining factor. It is a definition that focuses on the “electoral democracy” dimension, in the sense that it includes aspects that “ascertain that elections work as intended” (Teorell et al., 2016, p.6). At the same time, it excludes potential *outcomes* or *explanations* of democracy, such as socio-economic

components or equal distribution of wealth, income and social goods. This conceptual distinction is particularly important for this dissertation, that studies education as a determinant of democratization. My preferred democracy definition also excludes goals such as realizing the “common good” or the “general will” of the people, which is considered the ultimate goal of democracy in the classical definition of Schumpeter (1976).¹²

All the institutions listed above – and perhaps also additional institutional dimensions, such as more civil liberties and judicial independence – are necessary to ensure democracy in its most ideal form. Yet, the question of what institutions are necessary for an ideal democracy should be distinguished from the question of what institutions are required to qualify as a *minimal* democracy. A regime can be considered democratic (at a minimum) even if some of the institutions discussed above are lacking.

Hence, although I consider democracy in this thesis as fundamentally being a gradual concept, I do not take a specific stance in the widely debated question of whether democracy should be *operationalized* using a continuous measure, ranging from the most autocratic to the most democratic regimes, or as a dichotomous measure, distinguishing between either autocracy or democracy (see, e.g. Cheibub, Gandhi and Vreeland, 2010; Knutsen, 2011b). It has been argued that the choice of gradual versus dichotomous democracy measure should depend on the research question (Clark, Golder and Golder, 2012). I take a pragmatic stance when it comes to choice of democracy measure, utilizing both gradual and dichotomous democracy measures for various purposes.

For instance, when studying the breakdown of authoritarian regimes and their replacement by democracy, it is necessary to determine a threshold at which a country fulfills the *minimum criteria* for democracy. This *could* have been based on a chosen threshold on a continuous democracy scale, but the choice of such a threshold is likely to be arbitrary. This suggests distinguishing the difference between democracy and autocracy not simply in quantitative terms but also in qualitative terms, by specifying a substantial criteria for qualifying as a minimum democracy. For example, I use the Boix, Miller and Rosato (2013) dichotomous democracy index, which considers regimes as democratic if political leaders are chosen through free and fair elections and it satisfies a threshold value of suffrage. When using this measure, I still acknowledge that an advanced democracy requires more than contested elections with universal participation, and, that a regime lacking some of the institutions discussed above may be a relatively poorly functioning democracy

¹²This view is rendered problematic due to the inherent difficulty of aggregating individual preferences to a collective preference (Arrow 1951). Not only is the outcome of such an aggregation process sensitive to the preferences of dominating individuals. It will also fall short due to individuals “intransitive preferences”, implying that it is meaningless to rank all preferences from most desirable to less desirable (Shepsle and Boncheck, 1997). While the “common good” is the outcome of political decision-making, democracy is about the procedures giving people the opportunity to control politics.

At the same time, this thesis is also concerned with changes in degree of political contestation *within* democracies and autocracies. For instance, the altered preferences and resources due to education may prompt citizens to push for more accountability or political rights in poorly performing democracies (often referred to as semi-democracies), or it may promote more political contestation in autocracies – although not enough to push the country over a minimal democracy threshold. Moreover, democracy may emerge due to more gradual processes such as a series of political liberalization reforms introduced gradually by a dictator, rather than from a clear-cut transition from autocracy to democracy. Such variations and processes are not necessarily picked up by a dichotomous democracy measure. Hence, this dissertation also employs gradual democracy measures to study processes of democratization and institutional change within autocracies and democracies. The exact choice of democracy measure is described in each chapter, as well as discussions regarding the robustness of results to various democracy measures.

1.3.2 Variations in autocratic regimes

Above, autocracy is defined with reference to democracy, as the lack of institutions that realize popular control and political equality. Yet, as has been appropriately noted in the comparative literature on authoritarianism, there is vast diversity within regimes classified as autocracy. Moreover, these differences may be important to explaining prospects for authoritarian regime breakdown and democratization.

First, autocratic regimes have been classified in various ways based on differences in key actors and political institutions. For instance, Geddes, Wright and Frantz (2014) offer a classification of authoritarian regimes based on the type of actors that are in control – both when it comes to appointing key political positions and controlling policy-making. This classification broadly consists of four categories: 1) personalistic dictatorships, 2) military regimes 3) one-party regimes and 4) monarchies. One other commonly used classification is offered by Hadenius and Teorell (2007). This classification, distinguishing between 18 regime types, shares some similarities with that of Geddes, Wright and Frantz (2014), but focuses on the institutions that are in place rather than the actors. For instance, regimes are classified based on whether only one or several political parties are operating, whether elections are held and the use of military force.

Drawing on such classifications, several studies have demonstrated that authoritarian regime type matters for prospects of regime breakdown and democratization (Geddes, Wright and Franz, 2012; Magaloni, 2008), civil war (Fjelde, 2010) as well as economic and social outcomes such as investment and growth (Wright, 2008) and property rights (Knutsen and Fjelde, 2013). Notably, there is evidence that military regimes and multiparty regimes have lower survival prospects and higher likelihood of democratization

than e.g. monarchies or one-party regimes (Magaloni and Kricheli, 2010; Geddes, 1999; Geddes, Wright and Franz, 2012; Hadenius and Teorell, 2006; Wright and Escriba-Folch, 2012). There is also evidence that one-party regimes are associated with greater investment and stronger growth performance than other types of autocracies (Gandhi, 2008; Wright, 2008).¹³

Along this line, regime type could influence both education levels and prospects of regime instability and democratization. For instance one-party regimes, in line with literature indicating their stability and advantage at promoting economic development, may offer more education and be less likely to break down than other authoritarian regimes. This perspective is incorporated, for instance, through robustness tests of whether the identified effects of education on regime instability or democratization is sensitive to controlling for authoritarian regime type (in chapters 2, 3 and 4).

Second, autocracies display a lot of variation when it comes to where potential threats to regime survival are likely to emerge from. In particular, it is common to distinguish between the threat from the masses and the threat from the ruling coalition (Svolik, 2012; Gandhi, 2008; Magaloni, 2008). The latter source of opposition is defined as “those individuals who support the government and, jointly with the dictator, hold enough power to be both necessary and sufficient for its survival” (Svolik, 2009, p.478), and are often also referred to as the “winning coalition” (Bueno de Mesquita and Smith, 2003) or the “launching organization” (Haber, 2006). Regimes that are mainly concerned with the potential of a revolution may choose prioritize building up a military and repressive apparatus that can quell potential mass resistance, or carry out redistribution policies to satisfy the larger public (Svolik, 2012; Bueno de Mesquita and Smith, 2003). Dictators whose greatest fear is a coup d’etat by the elites, or who desires to acquire more power at the expense of the ruling coalition, may choose to emphasize public policies that will satisfy the elites or even build down traditional security forces that may serve as stepping stones for coups (Svolik, 2012; Albertus and Menaldo, 2012). This dissertation is mainly focused on threats to survival emerging from the masses, although it does also offer discussions of how education can be used to co-opt the elite, or the ruling coalition, for instance through restricting access to high-quality education to a selected few.

Finally, autocracies vary when it comes to the strength and capacity of the dictator to carry out policies that promote regime survival. An important aspect of this is so-called *state capacity*, commonly defined as the ability of state institutions to effectively imple-

¹³Other important institutional variations between autocracies, having to do with the presence of “democratic elements”, may also matter for regime survival. Studies suggest that introducing legislatures may reduce the likelihood of authoritarian breakdown (Gandhi, 2008). Although the findings are somewhat mixed, there is also evidence that elections stabilize dictatorships – at least in the long run (Gandhi and Lust-Okar, 2009; Knutsen, Nygård and Wig, 2017).

ment official goals (Sikkink 1991), or as the “degree of control that state agents exercise over persons, activities, and resources within their government’s territorial jurisdiction” (McAdam, Tarrow and Tilly, 2001, p.78). It is commonly distinguished between different types of state capacity, such as “extractive capacity” – the ability to extract revenues¹⁴, or “bureaucratic quality” – related to bureaucracy’s efficiency when it comes to implementing decisions (see, e.g. Hendrix, 2010). Yet, the distinction between such aspects is not clear-cut, as they are likely to be related – for instance, access to resources may improve the ability to effectively implement policies, and an effective bureaucracy should improve the ability to collect resources.

State capacity could very well influence education levels. For instance, a regime with available resources and bureaucratic capacity to implement education reforms should be more capable of boosting enrollment rates or expand the duration of public schooling. At the same time, high-capacity regimes with necessary resources and implementing power should also be able offer a range of other policies that strengthen regime stability. This includes, for instance, the implementation of public services and benefits that boosts popular support, or building a strong security apparatus that discourages and in worst case suppresses anti-regime mobilization.

Such discussions are returned to in chapter 4, which also tries to test whether the identified (negative) relationship between education and authoritarian breakdown is driven by available indicators of state capacity. As noted in the literature, aspects such as state capacity are difficult to measure. At the same time, existing suggestions for operationalization of state capacity suffer from lack of available data. As a result, it can not be ruled out that country-level analyses of the education-regime change relationship can potentially be driven by such unobserved factors. Indeed, this is one important motivation for also conducting more disaggregated analyses, by operating with indicators of education at the group or sub-national level, that are (more) disentangled from state-level characteristics such as regime capacity or inherent sources of regime stability.

1.3.3 Democratic transitions and autocratic breakdown

Building loosely on the above definitions of democracy and autocracy, *democratization* can be defined as institutional change towards more democratic institutions. This can happen gradually, through small incremental changes from year to year, or more abruptly, through for instance a revolution. In line with the notion that we can distinguish between

¹⁴In particular, the literature on state capacity has been concerned with the ability to extract revenues through taxation rather than through primary commodity exports (Hendrix, 2010, see, e.g.) Yet, this focus is less relevant to my dissertation, as access to resources may enhance both education levels and regime survival regardless of the origin of the resource. Hence, natural resources may also strengthen dictator’s “degree of control” over society.

an autocracy and a minimal democracy, a *democratic transition* should be understood as a regime change from autocracy to (minimal) democracy. Gradual changes towards democracy within autocracies or existing democracies do not qualify as a democratic transitions. Although democratization can be distinguished conceptually from a democratic transition, this difference is not crucial to the research conducted in the dissertation. My main argument applies to both of these instances - I suggest that education should make people more likely both to push for democratic transitions and for promoting democratization. In most chapters, I focus on a measure capturing either democratic transitions or gradual democratization in the main analysis. Yet, I always also run robustness tests with a measure capturing the alternative concept.

Although democratic transitions are the most well-researched, the most common form of regime change is actually transitions from one authoritarian regime to another. The latter definition also requires clarifying what a regime is. In line with Geddes, Wright and Frantz (2014), a regime can be defined as “a set of formal and/or informal rules for choosing leaders and policies”. According to this definition, a regime transition takes place if an autocracy is replaced by an autocracy with different “rules of the game”. Yet, a leadership change such as the installation of a new president or a new king is not a regime transition, unless the “rules of the game” are also altered at the same time. For instance, the ruling period by the Somoza family in Nicaragua is considered as one regime despite several changes in leadership (Geddes, Wright and Frantz, 2014).

The literature review presented above shows that existing studies on regime change nearly always focus on democratic transitions or democratization as the outcome. Yet, as noted by recent studies, this approach may offer limited insight into the dynamics and nature of regime change processes. An important conceptual building block in this thesis is the distinction between the *onset* of regime change and the *outcome* of regime change. (see e.g. Kennedy, 2010; Miller, 2012). In order for a successful democratic transition to take place, two distinct events need to happen: First, the authoritarian regime needs to experience some kind of instability or even breakdown that initiates the regime change. Second, the unstable or failed authoritarian regime must be replaced by a democratic regime, and this is the outcome of the regime change.¹⁵

Importantly, authoritarian instability or breakdown does not necessarily lead to the emergence of democratic institutions. In fact, most authoritarian breakdowns culminate in the emergence of new authoritarian or non-democratic regimes (see e.g. Geddes, Wright and Frantz, 2014). And, once a regime change is initiated, the outcome of that regime change is not given. By focusing exclusively on transitions from autocracy to democracy,

¹⁵This was also noted by Huntington (1991), who argued that democratization involves three stages: the end of an authoritarian regime, the installation of a democratic regime, and the consolidation of the democratic regime.

existing literature only compares cases of transitions from autocracy to democracy to cases that do not experience such transitions. In other words, by excluding all instances of authoritarian breakdown that do not produce democracy from the analysis, only one specific instance of authoritarian instability is studied. Yet, as authoritarian breakdown or instability is a precondition for democratization, it is necessary to understand *how and why authoritarian instability arises*, regardless of whether this instability or breakdown leads to a democratic outcome (Kennedy, 2010; Ulfelder, 2005).

The importance of distinguishing between regime change onset and outcome is illustrated by Kennedy (2010), showing that these two processes may have different and even contradictory causes. Re-investigating the common claim that economic development breeds democratization he finds – contradictory to this popular assumption – that income level *reduces* the likelihood of authoritarian breakdown (or regime change onset). Through boosting state capacity and increasing the opportunity costs of potential opposition members, a high income level may actually stabilize authoritarian regimes. Kennedy does, moreover, find evidence that income level raises the likelihood of a democratic outcome *given that* regime breakdown has taken place. Yet, the fact that income simultaneously stabilizes dictatorships explains why many empirical studies do not identify an effect of income level on democratization. Similarly, a complete account of how education affects regime change and democratization should pay attention to how education influences the likelihood that citizens will oppose a dictator and promote democratization, be it through an elite coup, a popular uprising or by pressuring the dictator to political liberalization. This insight is incorporated in my thesis through a focus on how education influences the dynamic and stability of authoritarian regimes. For instance, chapter 4 investigates how education influences the likelihood of authoritarian regime breakdown, demonstrating that state-level indicators of education access is actually associated with lower likelihood of authoritarian breakdown, in line with Kennedy's (2010) finding that income reduces authoritarian instability. Moreover, chapters 5 and 6 look at how education influences processes of political instability in dictatorships due to mass mobilization – regardless of whether these processes lead to democracy.

1.3.4 Contentious collective action

Collective action can be defined as behavior or action carried out by a group of individuals working towards a common goal. This dissertation is concerned with a particular form of collective action, namely *contentious* collective action. Collective action becomes contentious when it is “used by people who lack regular access to representative institutions, who act in the name of new or unaccepted claims, and behave in ways that fundamentally challenge others or authorities” (Tarrow, 1994). Hence, I am interested in behavior that

seeks to challenge regimes and governments, and to “transform existing patterns of political authority” (Ulfelder, 2005). This form of collective action should be distinguished from “routine” political behavior that challenges demands through accepted and established political institutions in democracies, such as voting, participation in community meetings, joining political parties or influencing elected politicians through lobbying.¹⁶ As noted, there is ample evidence in the literature that education raises the frequency of such conventional forms of political participation in democracies (see, e.g. Wolfinger, 1980; Milligan, Moretti and Oreopoulos, 2004). Yet, this does not easily generalize to “contentious” collective action, and particularly not to contentious collective action in dictatorships, where the costs of anti-regime protest are potentially severe. Hence, anti-government protests is quite a different activity from e.g. voting - described as a “political act requiring very little initiative” (Nie, 1987, p.77).

Contentious movements can take the form of brief, spontaneous events or more continuous and sustained form. The latter is often referred to as protest “campaigns”, which are institutionalized, contentious protest movements that challenge a target over a period of time. Contentious movements can employ a variety of strategies, both violent and non-violent. This dissertation is primarily concerned with *anti-regime* collective action, i.e. actions seeking to promote regime change, and especially democratization. This means that contentious protest directed against other targets than the government fall outside of the scope of this investigation. Contentious protest against the government but aiming for other goals than regime change, such as riots aiming to lower food prices, do not technically fall into the dissertation’s main scope either. Yet, the latter distinction is far from clear-cut. Many protests that do not explicitly work to overthrow regimes, such as food riots, may play an important role when it comes to undermining the regime’s legitimacy, which again may promote regime breakdown. Hence, all forms of contentious collective action directed against the government are usually included in my analyses. More details regarding operationalization and measurements of protest is discussed in section 1.6 and in chapters 5-6.

The focus in this thesis is on political instability due to mass opposition. This is in line with what has been referred to as the “bottom-up approach”, which includes efforts to theorize and explain regime change from autocracy to democracy in terms of variations in mass mobilization (Clark, Golder and Golder, 2012). This approach has been criticized by proponents of “top-down” explanations of democratization, who argue that the preferences, strategies and behavior of the elites are far more important (see, e.g. Rustow, 1970; Di Palma, 1990; O’Donnell and Schmitter, 1986; Linz and Stepan, 1996)

¹⁶In authoritarian regimes, there are fewer legitimate channels for influencing decision-making in a non-contentious matter.

There are indeed many instances of democratic transitions where the elites have played important roles, through initiating processes of gradual liberalizations or through securing so-called “pacted transitions”. Yet, my focus on mass mobilization is backed by empirical studies indicating that collective contentious action is an important determinant of democratization (Ulfelder, 2005; Celestino and Gleditsch, 2013; Teorell, 2010). Moreover, even classical instances of elite-driven regime change, such as pacted transitions, combine elements “from above” with elements “from below” (see, Wood, 2003). Studying typical instances of pacted transitions – Spain and Peru – Collier (1999) finds that popular mobilizations played a greater role than had been hitherto understood. Moreover, Bratton and Van de Walle (1997) finds that collective action was often a trigger behind the liberalization reforms introduced by several autocratic rulers in sub-Saharan Africa in the 1980s and 90s. Hence, mass mobilization in these instances is not only a *symptom* of regime instability, but also a facilitating factor.

Throughout the rest of the dissertation I commonly refer to contentious collective action as protest. Hence, protest does not simply refer to the familiar image of people demonstrating in the streets, although this is a common instance, but also other forms of protest activities such as e.g. boycotts, sit-ins, petitions, strikes and uprisings.

1.3.5 Education

This thesis is concerned with all forms of education offered by accredited educational institutions. In line with the OECD, I consider educational institutions as all entities that offer instructional services or organized learning to individuals, regardless of which authority is responsible for it, and whether it is private or public (OECD, UNESCO and Eurostat, 2006). As a starting point, I do not distinguish between types of education or education contents. I exclude, however, learning that takes place outside of accredited educational institutions and offers instruction in particular skills which are not part of the recognized educational structure of the country (for example, in-service training courses in factories). In line with previous literature, I consider both primary, secondary and tertiary education¹⁷, and the relevant theoretical mechanisms are assumed to apply to

¹⁷According to OECD, UNESCO and Eurostat (2006) *primary education* usually begins at age five, six or seven and lasts for four to six years (the mode of the OECD countries being six years). According to OECD, “programmes at the primary level generally require no previous formal education, although it is becoming increasingly common for children to have attended a pre-primary programme before entering primary education”. *Secondary education* consists of lower and higher secondary education. This former usually consists of two to six years of schooling (the mode of OECD countries is three years) and “generally continues the basic programmes of the primary level” (OECD, UNESCO and Eurostat, 2006). The latter, typically ranging from two to five years of schooling, “is often more organised along subject-matter lines, and teachers typically need to have a higher level of qualifications (OECD, UNESCO and Eurostat, 2006). Finally, tertiary education is either programmes designed to provide sufficient qualifications for entry to professions with high skill requirements (minimum three years) or programmes of minimum

all of them. For instance, each education stage is assumed to strengthen people's skills and capacities and each stage may stimulate anti-authoritarian grievances. Yet, the empirical focus differs somewhat between the different chapters, due to research question and research design. For instance, the study of education quality and democratization in chapter 2 is restricted to primary and secondary schooling, due to data availability on education quality. When collecting data on the educational background of protest movements in chapter 6, I restrict the analysis to variations in tertiary education, as this is the only information listed in the sources.

The above definition of education is concerned with education *inputs*, or the *amount of education* offered to students. It does not consider variations in education *outputs*, including what students actually learn in schools – the skills and learning outcomes. Yet, this is of course also relevant to how education should influence prospects for regime change, especially according to mechanisms specifying how education influences skills and preferences. There are indeed large variations in what is being taught in schools, and this thesis also considers how this variation may affect regime type and change. For instance, schools may vary in the extent to which they actually offer students the skills and resources that make them more efficient political actors that can threaten dictators. Moreover, some educational institutions may offer education with very specific content, for instance a curriculum designed to install particular ideologies. This may include religious schools, army academies or schools with a high level of indoctrination in regime loyalty. Chapter 2 considers how variation in schooling quality and contents influences regime change. Moreover, chapter 3 considers the “reverse effect”, namely how regime type influences education quality.

1.4 Opportunities and preferences for anti-regime mobilization

All the empirical contributions in chapters 2-6 various implications of the same theoretical framework, which I briefly outline here. As noted, the democratization literature offers few detailed attempts at tying education to particular agents and their choices. Hence, to develop a theoretical “micro-argument” I draw on a framework commonly applied in studies on civil conflict and violent mobilization, explaining the behavior of individuals or groups as a function of two factors: their *preferences* and *opportunities* (Gurr and Duvall, 1973; Tarrow, 1994; Cederman, Gleditsch and Buhaug, 2013).

two years focusing on practical, technical or occupational skills for direct entry into the labour market (OECD, UNESCO and Eurostat, 2006).

The choice to participate in protest can be considered a function of the expected utility to be gained from protesting, minus the costs associated with the participation. I consider, first, how education reduces the costs of carrying out regime-challenging contentious actions, by raising *opportunities* for successful collective action. In addition, education raises the expected utility from promoting regime change, due to emerging *preferences* for alternative regimes than the incumbent, notably democracy. Below, I start by describing the main actors. Next, I briefly discuss some important obstacles that may prevent collective action from emerging. Finally, I outline the two main mechanisms through which education promotes and influences collective action, namely the “opportunity-mechanism” and the “preferences-mechanism”.

1.4.1 The actors

The theoretical framework in this thesis considers how education influences anti-regime opposition in dictatorships. Hence, I am concerned with mechanisms expected to play out within autocratic regimes. I build on a simple model with two main actors, the (potential) political opposition and the dictator. The potential political opposition consists of citizens who, under certain circumstances, will challenge the power of the regime through actions such as mass protests or attempted coup d'états. I assume that there is variation both in terms of whether citizens prefer to overthrow the regime and in whether they actually do so. Although many authoritarian regimes center heavily around one person, what I refer to as the dictator is in practice usually more than one person, such as an family or a leadership group (Svolik, 2012; Geddes, Wright and Frantz, 2014).

I assume that the main priority of the dictator is political survival, in line with Bueno de Mesquita and Smith (2003), who argue that political leaders choose those policies that they consider will maximize their chances of remaining in office. This assumption of office-seeking rulers underlies much classical work on both democracy and dictatorship (Downs, 1957; Wintrobe, 1998) While holding office yields power and prestige it also enables the ruler to accumulate rents, whereas losing office has direct consequences, including the prospects of exile, imprisonment, or even death (Goemans, 2000).

My argument emphasizes how education influences the behavior and preferences of *the political opposition*, in ways conducive to contentious collective action and authoritarian instability. At the same time, the opportunities and choices of the opposition are likely influenced by the regime's policies and behavior, and hence the two actors cannot be studied completely in isolation. This is incorporated through discussions of, e.g., the use of education policies as a potential co-optation mechanism in autocracies in chapter 2. Moreover, chapter 3 is devoted entirely to studying how regime type may influence education access and quality. I also return to such discussions when considering endogeneity

problems in section 1.6.¹⁸ In the final section of this chapter, I suggest some avenues for future research, highlighting the need for more attention to how the dictator may adopt certain strategies and policies in response to potential threats from an empowered, educated population.

1.4.2 Obstacles to collective action

Participating in political protest as opposed to being inactive or staying at home always comes with a cost, that any rational individual will only accept if the expected utility to be gained from it exceeds the cost. In an authoritarian regime, where political opposition is considered illegitimate by the very nature of the political system, the costs of protesting will be especially high if the protests fail. An individual will not only have wasted valuable time that could have been spent differently, she will also face the potential of severe, even life-threatening sanctions in the form of repressions, purges or imprisonment.

At the same time, even if the expected utilities are considered to exceed the costs, the nature of collective action problems may still make it rational for each single individual to refrain from protesting. For instance, every individual lacks information about the likelihood of a protest succeeding, because she cannot know or trust that others will also participate in the protest, and thereby ensure successful mobilization (Downs, 1957; Tarrow, 1994; Kuran, 1995). This will discourage collective action. Participation in collective action is also discouraged by the dilemma known as “free-riding”, describing how individuals can reap the benefits of collective action without taking on the costs of participating in it. Such collective action problems raise the costs of protesting. I consider collective action problems both as an obstacle to the *emergence* of anti-regime opposition and a factor that may influence the outcome of the opposition movement, including whether it is successful at achieving its aims or not (Taylor, 1988). In chapter 5, I discuss and investigate whether education promotes the *emergence* of protest, and chapters 4 and 6 investigates whether education strengthens protest movements’ potential for success.

The costs and utilities associated with protest are not, however, identical for all individuals: Some will have relatively lower costs associated with protesting, and some will have a higher expected utility. The first part of my theoretical framework (mechanism I) discusses how education can be considered a resource that *increases the capacity* of potential protesters for organizing collective action, thereby reducing the expected costs of protest. The second part (mechanism II) considers how education raises the expected

¹⁸Regarding more detailed discussions concerning the strategic considerations made by the dictator and authoritarian politics, including use of repression and co-optation, the thesis relies on existing literature such as Svolik (2012) and Gandhi (2008).

utility of contentious collective action, due to increasing grievances against dictatorship.

1.4.3 Mechanism I: Opportunities for anti-regime protest

The focus on opportunities and capacities for collective action is in line with the emphasis in the civil war literature on the role of “opportunity structures” in facilitating armed conflict. For instance, Collier, Hoeffer and Rohner (2009) claim that the most important determinant of armed rebellion is “feasibility”, arguing that when rebellion is possible and beneficial for potential rebels, it will happen. Collier et al. focus in particular on the availability of financial and military resources as important factors raising the feasibility of rebellion, and argue that this is much more important than motivations or grievances. Other civil war scholars emphasize the *political* opportunity structures that potential rebels are faced with, showing for instance that strong states yield higher costs of rebellion (see, e.g. Fearon and Laitin, 2003; Collier and Hoeffer, 2005; Gleditsch and Ruggeri, 2010).¹⁹

Economic models of democratic transitions, such as that offered by Acemoglu and Robinson (2006, p.31), also acknowledge that the resources of citizens may matter for the prospects of regime change, arguing that “when the citizens are not well organized, the system will not be challenged and transition to democracy will be delayed indefinitely”. Yet, their focus is mainly on the *preferences* of the relevant actors and how these are influenced by structural factors such as economic inequality (see, e.g. Boix, 2003; Acemoglu and Robinson, 2006; Ansell and Samuels, 2014).²⁰ The theoretical framework in this dissertation places much more emphasis on opportunities, by applying the “opportunity structure” perspective to authoritarian instability and democratization, focusing especially on citizen’s *resources and capacities* as factors lowering the costs of anti-regime mobilization.

The consecutive chapters consider several ways in which education strengthens resources and raises capacities, thereby making contentious action more feasible and effective. First, by enhancing skills and knowledge levels, education creates individuals that are more efficient political actors. For instance, literacy skills and more advanced communication skills induced by several years of schooling enables efficient use of communication technology to share information, opinions and even concrete plans about the timing and

¹⁹This focus on opportunities is integrated in much of the literature on social movements as well, although most of theses studies study political opportunity structures (McCarthy and Zald, 1977; Tilly, 2003; Tarrow, 1994).

²⁰For instance, in the model of Acemoglu and Robinson (2006) democratic transition is expected to occur under medium levels of economic inequality, because this offers an incentive structure in which the poor will prefer democratization and the elite will be able to agree on this, expecting the economic redistribution resulting from democratization to be moderate.

location of protest activities. Moreover, schooling is a source of general knowledge about the world as well as more specific information about society, political institutions, and political opposition movements. This should make citizens better able to understand what their interests are as individuals and groups, and how these interests can be maintained (Galston, 2001). A certain core of knowledge also makes it easier to acquire more knowledge, and make sense of that knowledge in a systematic manner. Due to such skills and knowledge, educated citizens should have a comparative advantage when it comes to developing an efficient campaign strategy.

Second, education helps to strengthen social networks, that are of great importance when it comes to facilitating cooperation and political protest (Siegel, 2009). An individual who goes to school will have to interact and socialize daily with a large number of fellow students from outside of family and kin. In this sense, schooling strengthens social networks and creates ties across different segments of society that otherwise would interact much less frequently (Alesina and Giuliano, 2011). Given that education will make employment later in life more likely, this effect will be extended by continued contact with a large number of individuals at the workplace. Several studies indicate that social networks are crucial to the recruitment of new protesters, and important to overcoming collective action problems (e.g Goldstone, 1994; Kuran, 1991; Lichbach, 1995; McAdam, 1986; Siegel, 2009).²¹ Social networks may also reduce collective action problems: When people interact more frequently with other people in their social networks, they will also have more information about what others are doing and what they want, and should be more likely to trust that others will also participate in collective action. Siegel (2009, p.12) writes that one benefit of social networks is that “information exchange allows people to update their beliefs about the costs and the benefits inherent in participation, and so change their decisions”.

Both the skills and knowledge acquired from education and the social networks that are formed in schools suggest that educated individuals will face lower costs of opposing authoritarian regimes. I return to these discussions throughout the rest of the dissertation.

1.4.4 Mechanism II: Anti-authoritarian grievances

While the opportunity-mechanism is concerned with how education lowers the costs of anti-regime mobilization, education may also increase the *expected utility* from protesting. In this thesis, I consider two ways in which people develop preferences for anti-regime mobilization in autocracies: Due to ideological or economic grievances. Here I build on extant literature considering grievances as a fundamental source of collective mobilization

²¹Moreover, the size of one’s network affects the probability of political participation (McClurg, 2003), and this finding is probably also applicable to regime-challenging protest.

(Davies, 1962; Gurr, 1970, 2000). Grievances and preferences are also at the core of economic models of democratization such as those proposed by Acemoglu and Robinson (2006) and Boix (2003). Yet, my theoretical framework differs from theirs in the sense that economic calculations are not the only mechanism considered to influence preferences and behavior. In addition, I assume that education will influence individuals' behavior by shaping ideological and moral dispositions. In particular, we can expect educated individuals to develop a moral aversion against authoritarian government and a preference for democratization and political liberalization. This should again increase the likelihood of anti-regime collective action in dictatorships.

Ideological grievances

One potential mechanism, tying education to anti-authoritarian grievances, is through certain shifts in preferences and values referred to as "human development" (Inglehart and Welzel, 2005; Welzel, 2013). Inglehart and Welzel (2005) discusses how increasing material and cognitive resources as a result of economic modernization shift people's preferences towards demands for freedom, individual autonomy, civil liberties and democracy. This logic builds on Maslow's "hierarchy of needs", according to which "primary" needs such as material security and cognitive resources must be satisfied before demands for "higher" needs such as self-fulfillment will emerge (Maslow, 1943). When people are safe, free and have the resources "to devote energy to concerns that do not immediately threaten them" (Inglehart and Welzel, 2005, p.33), they start demanding more freedoms.²²

One component in this process of "human development" is material resources: Young people grow up taking material security and survival for granted, promoting "existential security" (Inglehart and Welzel, 2005). This make people less inclined to emphasize money and material resources, as more material benefits is no longer a necessity of survival, and less likely to value group conformity, state authority, and strong leadership – orientations motivated by self-preservation and fear (Inglehart and Welzel, 2005).

Another important force for changes in preferences according to this framework, which is particularly relevant to emerging preferences for democracy, is increases in cognitive resources. Education is at the core of this development. It strengthens people's intellectual abilities and cognitive autonomy, thereby pushing people's grievances higher up on Maslow's hierarchy of needs Inglehart and Welzel (2005).²³ Relatedly, through both

²²It has been questioned empirically, however, whether such liberal values are really associated with transitions to (and the survival of) democracy (see, Dahlum and Knutsen, 2016; Hadenius and Teorell, 2005)

²³According to Inglehart and Welzel (2005) is it not only the actual schooling that explains why educated individuals are more likely to hold self-expression values, but also the fact that educated individuals usually grow up under high levels of material security.

the education system and potentially more autonomous tasks in the workplace, educated individuals grow used to exercising initiative and independent thinking. As a result, they will also increasingly emphasize goals that previously were given lower priority, including self-realization, more individual autonomy and the pursuit of freedom.²⁴ Given these emerging preferences, democracy naturally becomes the preferred form of government, while the aversion against dictatorships should increase. This preference should raise the expected utility of anti-government protest in dictatorships, as the utility of the expected outcome of protest (democracy) increases. According to this mechanism, education may also raise the *direct* utility of collective action: Participation in protest is a way of realizing self-expression values, suggesting that it may produce utility even when the movement does not successfully induce regime change (Inglehart and Welzel, 2005).

Education may also promote the ability to use abstract, moral reasoning. Pinker (2011) argues that this ability stimulates preferences for policies and political regimes that best serves the interests of most people, rather than the political system that serves each individual's self-interest. As I discuss below, educated people may not always have economic incentives to prefer democracy, as they may fear that democracy leads to radical redistribution. Yet, a strengthened level of moral reflection may promote the ability to downplay one's self-interest in favor of virtues such as political equality and civil liberties, that should yield preferences for liberal democracy.

Economic grievances

Of course, individual preferences are not only driven by ideological or moral considerations, and economic incentives should also have a strong impact. There is, however, no unequivocal theoretical expectations relating to how education should impact on preferences for democracy based on economic calculations:

In economic models of democratic transitions such as Acemoglu and Robinson (2006) and Boix (2003), democratization is preferred by those who will benefit from the economic redistribution expected to take place in a democracy. Meanwhile, those who fear redistribution will work to prevent transition to democracy. Educated individuals and groups are likely to be relatively affluent and may therefore lack the type of grievances that may spur anti-regime mobilization. This is in line with Acemoglu and Robinson's (2006) argument about the role and preferences of the middle class, which is considered as a third group, distinguishable from the poor and the elite (the regime). Although the middle class is considered as having the capacity for carrying out collective action, it is not assumed to have strong preferences for democracy in this model, due to their (relative) economic advantage and fear of redistribution under democracy. There is no perfect

²⁴These preferences have been referred to as "self-expression values" (Inglehart and Welzel, 2005).

overlap between educated groups and what is commonly referred to as the middle class, but it may be useful to consider highly educated individuals as part of the middle class when considering their economic preferences for redistribution and democratization.²⁵

Relatedly, regime change may not take place if people are sufficiently satisfied under current institutions. In the framework of Acemoglu and Robinson (2006), this situation emerges when society is relatively egalitarian and prosperous, eliminating economic grievances. Education could be such a source of satisfaction. First, education is itself a benefit that recipients will value on its own terms. At the same time, education opens up other opportunities such as more attractive job opportunities, power and material rewards (Card, 1999; Oreopoulos and Salvanes, 2011; Clark and Roayer, 2013).

On this point, it is worth considering the strategic calculations made by the dictator and the regime when developing education policies. The provision of public policies such as schooling is often used deliberately as a tool to reduce grievances and create loyal followers (see, e.g., Bueno de Mesquita and Smith, 2003; Bueno de Mesquita et al., 2002; Haggard and Kaufman, 2008; Mares and Carnes, 2009; Knutzen, 2015). In particular, the dictator may choose to offer education to his supporters, or “ruling coalition”, in exchange for loyalty and continued support. Hence, recipients of education may have received this benefit precisely as a means of inducing political stability. If this policy is successful, education may actually reduce the likelihood of contentious politics. It is, however, hard to imagine that most dictators have the leverage to use education policies *only* to buy off supporters. After all, investing in human capital may be necessary to strengthen the economy and boost government revenues. The dictator may therefore also sometimes have to offer education to potential opposition members in order to secure prosperity.

At the same time, neither schooling nor the economic rewards from schooling may be sufficient to quell emerging political grievances among educated individuals in the long run. Indeed, if we are to follow the psychological “human development theories” discussed above, pressure for political freedom and liberalization may start to grow *especially strong* if individuals have satisfied their material needs and feel physically secure. Moreover, educated individuals may prefer transitions to democracy precisely *because* they are relatively affluent and want to protect their own property against arbitrary expropriation by the dictator (Ansell and Samuels, 2014). On this account, the middle class may prefer democracy as a tool to restrict autocratic power-holders from grabbing a share of their wealth, in line with Ansell and Samuel’s (2014) theoretical framework.

²⁵Most individuals and groups with tertiary education will be considered part of the middle class, but not everyone in the middle class will have tertiary education: For instance, merchants and the urban business elite are not necessarily educated, but are usually still considered part of the (urban) middle class. Moreover, social groups that fall within the “petty bourgeoisie” category, such as white-collar workers and functionaries, will not necessarily have tertiary education.

Hence, the economic considerations of educated citizens may play out in different ways when it comes to influencing prospects of authoritarian instability. Implications of these mechanisms are explored further in this thesis. For instance, chapters 2 and 3 consider the possibility that schooling as a public good may reduce anti-authoritarian grievances and boost loyalty. These chapters discuss, for instance, how high-quality education offered only to a few citizens may serve as a means of co-opting the elites (chapter 2), and how authoritarian regimes may face the same incentives as democracies to offer education to its citizenry to ensure continued support (chapter 3). Chapter 5 discusses and investigates the possibility of an interaction effect between education and unemployment, according to which education may boost anti-regime protest *only if* combined with a high level of unemployment, as this situation creates both moral and economic grievances against the regime.

In sum, education should raise the likelihood of collective action through enhancing citizens' skills, capacities and anti-authoritarian grievances. These mechanisms also suggest that education should promote more *successful* collective movements, more likely to overthrow authoritarian regimes and install democracies. The moral grievances against autocracy could, however, be neutralized by potential stronger economic preferences *against* democracy, but it is far from clear that economic calculations may lead educated citizens to fear democracy. Moreover, such economic incentives may in any case be offset by stronger ideological commitments to democracy that become increasingly more prominent with more education. These expectations are elaborated on and tested in chapters 2-6.

1.5 Data and units of analysis

The five consecutive chapters rely on different types of data at different levels of analysis to investigate how education impacts on authoritarian instability and prospects for democratization. Details about the data and operationalization of the particular variables are discussed, where relevant, in each chapter, but here I would like to highlight the overall logic when it comes to use of data and research design. In brief, I operate at three levels of analysis:

First, I present new tests of the relationship between education and regime instability and change at the *country level*, relying on country-years as the units of analysis. For instance, chapter 2 offers the first empirical analysis of how increases in education quality affects the likelihood of democratization. Chapter 3 investigates the possibility that education quality is endogenous to regime type (and change), by studying whether

democracy promotes education quality. To do this, I utilize a recent data by (Angrist, 2013), that compares student achievement tests from 117 countries from 1970 to 2013. Chapter 4 also uses country-years as units of analysis, to investigate how average education indicators impacts on autocratic survival. In these models, I operate with democracy indicators such as the Electoral Democracy index (the “Polyarchy index”) from Varieties of Democracy (V-Dem) (Coppedge et al., 2015) or the Polity index (Marshall, 2014) to measure changes in democracy score.²⁶ To measure autocratic breakdown I mainly use the indicator of authoritarian regime failure from Geddes, Wright and Frantz (2014).

Second, this dissertation investigates the relationship between education and political instability at the *sub-national level*. In this approach, the units of analysis are so-called “grid cells”, which divide the world into 50X50 kilometers areas that are completely independent of country-borders and other administrative units (Tollefson, Strand and Buhaug, 2012). More specifically, chapter 5 investigates the relationship between education and political protest onset at the sub-national level in Africa. To measure education, the study draws on geo-referenced survey data recording the education levels of respondents, taken from the Demographic and Health Survey (DHS). The average level of education in each grid cell is then matched to geo-referenced protest event data from the Social Conflict in Africa (SCAD) dataset Salehyan et al. (2012).

Finally, chapter 6 investigates the relationship between education and mass protest and prospects for authoritarian breakdown at the *level of the relevant actors* – political protest campaigns. In particular, I investigate how education impacts on political instability and regime change through investigating whether protest campaigns consisting of educated individuals and groups are more likely to succeed in terms of achieving their goals, and more likely to use non-violent methods. To do this, I use original data on the education profile of the anti-regime protest campaigns that are listed in the Nonviolent and violent conflict and outcomes 1.0 (NAVCO) database (Chenoweth and Stephan, 2011). Details about data collection and the different variables measuring education profile are described in chapter 6.

1.6 Methodological considerations

The first part of this section discusses some fundamental assumptions regarding research methods and philosophy of science that the thesis builds upon – the methodological foundations. The second part considers specific methodological challenges to drawing

²⁶The recently released democracy indicators from the V-Dem dataset are generally preferable due to the transparency, detail and reliability of these data. Yet, as chapters 2 and 3 were written before the V-Dem data was released, these chapters rely on the Polity index.

inferences about the effects of education on regime change and contentious collective action. Here, I focus on two issues that reoccur throughout the dissertation: Trade-offs between biased and inefficient estimators and selection problems challenging the extent to which the empirical findings can be interpreted in a causal manner.

1.6.1 Methodological foundations

Mechanisms and micro-foundations

The proposed focus on micro-mechanisms in this thesis involves two different things: Investigating causal mechanisms and specifying and testing micro-foundations. Although these two concerns are often conflated in the literature (see, e.g., Hedström, 2008), and often overlap, the two should be distinguished conceptually.

First, several of the next chapters are concerned with the *causal mechanisms* tying education to regime change, and in particular the role of mass mobilization and contentious collective action. Following Gerring (2005), this thesis adheres to a minimal definition of a causal mechanism, as the “pathway or process by which an effect is produced”. This can be one or a long chain of mediating variables. The concern with causal mechanisms in political science is often associated with certain positions in philosophy or science or preferences for certain methods. Hence, some clarifications are necessary.

First, my concern with mechanisms should not be equated with a position stressing the context-dependent or bounded nature of social relations (e.g., George and Bennett, 2005). When specifying and studying the pathway between an explanatory variable and an outcome, I assume that the goal is still to identify generalizable effects.

Second, it does not entail a distinct approach to *proving causation*. Hedström (2008) argues that covariation between two variables, X and Y, does not give us insight about the “black box of causation” tying the two variables together. Rather, he argues, causation can only be identified through distinct methods such as process tracing or “causal process observation”, that can offer insights into the “nuts and bolts” of causal relations (see also Brady, 2008a; Coleman, 1986). However, this view is problematic in its rejection of covariational evidence. Having identified a series of causal relationships that tie an explanatory variable to an outcome, the “mechanistic” understanding of causation would still require us to search for the causal pathways that link these mediating factors, and this approach would quickly lead to infinite regress (King, Keohane and Verba, 1995). Hence, both a quantitative analysis of the (covariational) relationship between X and Y and an analysis that seeks to identify the pathway or process tying the two variables together, rest on the same goal: The attempt to evaluate a causal hypothesis by presenting convincing evidence that X influences Y.

Finally, and partly following the previous argument, my approach does not imply that causal arguments or inferences about causal relations are *impossible to make* without a supporting causal mechanism. A sound research design, for instance drawing on a quasi-experimental set-up, could allow convincing inferences regarding the causal effect of X and Y, assuming that this effect is supported by a causal mechanism. Yet, ideally we should do more than simply offer a minimal support for a causal statement, by presenting the most *convincing* causal argument as possible. An explanation identifying a causal mechanism tying X to Y is arguably more convincing than an explanation that does not invoke any causal mechanism. The former yields a more comprehensive explanation, and it is more precise and intelligible. Moreover, as this study relies on observational data, the findings will inevitably be vulnerable to omitted confounders, selection bias and other estimation problems. Identifying a causal mechanism may therefore help to bolster the credibility of an identified effect. As Gerring (2011, p.128) notes, “*Ceteris paribus*, a causal argument is stronger (i.e. more convincing) if a causal mechanisms can be specified and a set of covariational patterns consistent with that mechanism can be identified”.

Second, I advocate for the need to tie structural explanations of democratization to their *micro-foundations*. Micro-foundations of macro-level phenomena can be understood as hypotheses about the underlying circumstances of purposive agents whose choices bring about the macro-level outcome. This builds on an assumption that structural conditions produce causal effects *when* they affect the behavior of individuals. Little (2005) argues that the search for micro-mechanisms implies that “The causal capacities of social entities are to be explained in terms of the structuring of incentives and opportunities for agents”. Identifying a causal mechanism will often involve accounting for the micro-foundations. Yet, this is not a requirement if we accept the minimal definition of a causal mechanism, as the specified pathway through which X causes Y, which could also be at the macro level (Gerring, 2011). For instance, a theory specifying how education causes democratization through producing anti-regime protests accounts for the causal mechanism, but not the micro-foundations. When this theory is expanded to also including hypotheses about how education influences the incentives and opportunities of individuals in ways conducive to anti-regime protest, it also invokes the micro-mechanism.

Ideally, empirical tests should be conducted at the same level of analysis that the theories or hypotheses are specified at (Gerring, 2011; Cho et al., 2008). For instance, if the theory proposes that protest movements consisting of educated individuals are more likely to choose non-violent resistance than movements with workers or unskilled labor, this proposition should be tested at the meso level (the level of the protest movements), rather than at the state level. Yet, in practice this recommendation is commonly ignored,

for instance due to lack of data at the appropriate level of analysis. There are numerous examples in the literature of studies using macro-level evidence to support statements at the meso or micro level. And, in many cases this is not warranted, leading to so-called “ecological fallacies” - emerging from incorrect use of aggregate data to make inferences about individuals (or groups) (King, 2013).²⁷

Yet, it is not *always* unwarranted to make inferences regarding micro-explanations at higher levels of analysis. For instance, social science methodologists have proposed various methods of unbiased “ecological inferences” (King, 2013). Moreover, many micro-level theories have *implications* that apply to different levels of analysis. King, Keohane and Verba (1994) stress the need to specify and analyse as many implications of a theory as possible, to “maximize leverage”. One way of improving leverage is to specify and test implications of a theory at different levels of analysis. In this thesis, I combine evidence collected at different levels of analysis, thereby utilizing various forms of evidence to improve leverage. If different types of evidence point in the same direction, this should strengthen the confidence in the main findings.

Assumptions about individuals and groups

As noted, this dissertation develops and discusses (micro-)theories that specify how education will influence the preferences, choices and behavior of individuals and groups. Theories about the behavior of individuals and groups rest on certain ontological assumptions, that require a brief discussion.

At the individual level, all five chapters discuss how education may alter behavior due to changes in both preferences and resources. This account builds on the assumption that we can consider individuals as rational actors. For instance, if the decision to participate in regime-challenging opposition is based on a whim or a roll of the dice, theories about motivations and resources will be of little use for explaining protest participation, collective action and regime breakdown. Hence, I assume that actors will make a rational choice, which involves choosing the course of action with the highest expected utility.

This requires, first of all, that the actor knows which actions he can possibly take. Second, it requires the actor to be able to form a belief regarding which actions will lead to which outcome. Finally, the actor should be able to rank the possible actions in a coherent way. This is a minimal definition of rationality, which is much less strict than definitions such as that proposed by Elster (2009, p.191). He argues that the beliefs that

²⁷For instance, in the 2012 US presidential election the republican candidate Mitt Romney was less likely to get support from the richer states. Yet, this does not mean that richer people are more likely to support the democrats: At the individual level, there is a clear positive relationship between income and the likelihood of voting republican (Gelman, 2009). If we had inferred from the state level to the individual level in this instance, it would have been an ecological fallacy.

the chosen action is based upon must also be “as well supported as possible”, and that “the evidence must from an optimal investment in information gathering”. There are many reasons why most individuals will not live up to this ideal, for instance due to limited cognitive capacities, lack of knowledge or cognitive biases (see Kahneman, 2011). This is in line with Simon’s (1961) notion of “bounded rationality”, assuming that individuals are “intendently rational, but only limited so”.²⁸

An emphasis on micro-foundations is often associated with *methodological individualism*, the doctrine that all social phenomena must be reducible to the behavior of individuals. Yet, in practice, accounts of the micro-foundations of social outcomes at the macro level often assume *groups of individuals* as agents rather than single individuals, as groups are required to produce the hypothesized outcomes. This is also the case here. The proposed theories suggest how education influences the preferences and resources of individuals, but assumes that groups of educated individuals will be the agents of collective action and regime change. This application of individual-level theories to groups rests on the assumption that a group to a certain extent can be considered a unitary actor. This is a problematic premise in the sense that even individuals within a group united by a common identity will usually have many divergent opinions and preferences. When aggregating these preferences to the group level, irrational outcomes may emerge - that is, outcomes that clash with the interests of even a majority of the individuals in the group. Such problems are discussed in the literature on collective action problems. Free-riding problems and information problems may make it difficult to know how other individuals will behave, and this prevents commitment to the action and outcome that otherwise should be preferred by most individuals (Olson, 1965).

Despite these problems, it is common practice in political science to take a pragmatic stance and treat groups of individuals “as if” they were unitary actor, despite recognizing the problems behind this assumption.²⁹

Collective action problems are dealt with explicitly in this dissertation, for instance in chapter 5, which considers education as a resource that may help individuals to overcome collective action problems. In other parts, such as in chapter 6, analysing the strate-

²⁸Yet, even the minimal understanding of rationality may not always apply in practice. For instance, ranking the utility of possible actions is an exercise that is often done in an incoherent way. In reality, it is also often hard to assess what the outcome of a certain action will be. Each individual therefore has to rely on estimates of the likelihood that certain outcomes will follow an action, and these estimates will be used to rank and subsequently decide on the best action. Handling and weighing such probabilities is an endeavor filled with potential pitfalls that makes it hard to live up to the minimalist view of rationality described above. Despite these problems, “bounded rationality” is assumed in this dissertation.

²⁹One factor that can make the unitary actor assumption less problematic is a minimum level of *common identity* among the individuals in a group, based on which individuals recognize that they share some characteristics that makes it feasible to seek to promote a common interest (Little, 2012). As noted by Little (2012), “persons sharing an identity have some level of preparation for cohesive action and they have a common set of assumptions about the world that encourages similar behavior”.

gies and choices made by educated protest movements, I follow the common practice of treating groups “as if” they were unitary actors.

Causation

This dissertation ultimately seeks to investigate whether education has a *causal impact* on authoritarian instability and regime change. There are different accounts of causation in the social sciences. For instance, Brady (2008b) distinguishes between four different approaches to causation: a regularity theory, a manipulation theory, a counterfactual theory and a mechanisms theory, the latter discussed above.

This dissertation is based on a counter-factual understanding of causation, which states that a factor, X, is a cause of an outcome, Y, if X generates a change in Y relative to what Y *would otherwise be* (Gerring, 2011; Rosenberg, 2012). The causal effect of X on Y is conditioned on certain background factors (*ceteris paribus* assumptions).

One of the benefits of this conception of causation is that it distinguishes causal relationships from mere accidental regularities. Many accounts of the counterfactual notion of causation builds on the premise that causal relationships, as opposed to coincidental sequences, are instances of law-like propositions (see, e.g., Rosenberg, 2012, p.88). This does not require that we already know the law that brings about the causal effect, but it assumes that there is a law (or laws), already known or not yet discovered, which does so (see also Salmon 1966). While actual laws support their counterfactuals, accidental generalizations do not. As Rosenberg (2012, p.64) points out, laws support their counterfactuals because “laws express some real, necessary connection, some glue, that holds instances of their antecedents and their consequences together and that is missing between the antecedent and the consequent of an accidental generalization” (see also, Okasha, 2016).

The laws we appeal to in the social sciences are different, however, from those in for instance physics or chemistry. The covering law model presented by Carl Hempel equaled scientific explanation with a general law of a natural kind, but this view has rendered too strict at least for the behavioral and social sciences (see, Rosenberg, 2012; Woodward, 2002). As Rosenberg points out, at the very least we have to qualify Hempel’s covering law model. A cause in social science studies is rarely sufficient. What we identify as the cause of an event is almost always merely one among many conditions that could bring it about, but by no means guarantees that it will happen. Hence, at best we can turn to *ceteris paribus* laws or *probabilistic* laws.

Based on this counter-factual notion of causation, a causal effect can understood according to the *potential outcomes framework* (King, Keohane and Verba, 1994; Rubin, 2005). According to this definition, a causal effect of a treatment, X, on an outcome, Y, is

the difference between the value of Y given the treatment and Y given the counterfactual condition, which is the absence of the treatment. That is, for a unit of observation at time t , the causal effect can be described as $Y_{it}(1) - Y_{it}(0)$.

Since each unit has either experienced the treatment or not experienced the treatment, this effect can of course never actually be observed. We can not rerun the world and assign units to their counterfactual states. This is referred as the “fundamental problem of causal inference” (Holland, 1986). Simply comparing treated units with units that have not been treated cannot solve this problem in the type of observational data used in this study, as treated units will likely differ in many ways from the untreated. Hence, in non-randomized studies it is a fundamental challenge to find an appropriate control group, that shares the same characteristics as the treatment group *except from the treatment*.³⁰ Below (in section 1.6.2), I discuss concrete challenges (and potential solutions) to estimating the causal effect of education on outcomes such as collective action or regime change.

1.6.2 Methodological challenges

The next five papers in this dissertation make numerous choices when it comes to model specifications and assumptions, including choice of estimation techniques, lag structures, included covariates, use of instrumental variables, and diagnostic tests to assess the robustness and validity of the chosen models. Here, I will not go into all the details of the different research designs, as these are described in each paper, but I will single out two important and general methodological discussions that are important throughout the entire thesis.

Inefficiency versus bias

When evaluating the performance of an estimator, two criteria are commonly used – unbiasedness and efficiency (see, e.g. Wooldridge, 2010; Kennedy, 2008). Unbiasedness refers to whether the coefficient estimate is on *average* true. This means that if we were to estimate the coefficient across repeated samples, the average of the sampling distribution should equal the true but unknown value of the parameter.

In addition, we should be interested in choosing the most efficient estimator, which is a measure of the sampling variance of an estimator. The smaller the variance of the estimator, the more efficient it is, leading to higher certainty that the estimated

³⁰Randomized experiments, which are usually not feasible when studying regime change and democratization, have been suggested as a way of solving this problem: By dividing the units randomly into two groups, whereby one group receives the treatment and the other group does not receive the treatment, we can assume that the units do not differ systematically between groups in ways that may bias the estimated effect of the treatment. Hence, we can estimate the causal effect by estimating the average difference in an outcome between the treatment group and the control group.

coefficient is close to the true value of the parameter. The unbiased estimator with the lowest variance is often referred to as the “best unbiased” estimator (Wooldridge, 2010; Kennedy, 2008). In practice, there is often a trade-off between these two criteria, in cases where we only have the choice between an unbiased but inefficient estimator and a biased and efficient one. It will often be preferable to accept higher variance to get at a less biased coefficient (see, e.g. Plümper, Troeger and Manow, 2005). Conversely, although unbiasedness is an appealing ideal, choosing the least biased estimator is not necessarily the best solution to this trade-off (Kennedy, 2008). Sometimes it is preferable to choose an estimator that is only slightly biased but more efficient rather than an unbiased estimator with considerably higher variance. Below, I discuss how the trade-off between efficiency and bias guides the choice of estimators throughout the thesis.

In chapters 2-4, I use time series cross-sectional (TSCS) data, and a natural starting point for model estimation with this data structure is the ordinary least squares regression (OLS). As long a set of specific assumptions are met, OLS produces the most unbiased and efficient estimate.³¹ Unfortunately, the OLS assumptions are commonly violated in social science research designs, and they are particularly problematic when using TSCS data (see, e.g. Beck 2001; Beck and Katz 1995; Green, Kim, and Yoon 2001; Plümper, Troeger, and Manow 2005).³²

One widely accepted econometric technique for TSCS data, suggested by Beck and Katz (1995), is to use OLS with panel corrected standard errors (PCSE) to avoid getting inefficient coefficient estimates due to the pooled data structure. This procedure corrects the standard errors of the OLS estimates by dealing with heteroskedasticity, thereby improving the efficiency of the estimator. Yet, it does not make the coefficient estimate less biased, as autocorrelation may still have an impact (see, e.g. Plümper, Troeger and Manow, 2005). In response to this, Beck and Katz (1995) suggests also adding a lagged dependent variable to the regression model. The combination of panel corrected standard errors and lagged dependent variable – referred to as the “de facto Beck-Katz standard” (Plümper, Troeger and Manow, 2005) – is used as a baseline model in the TSCS analyses in this dissertation. This model reduces inefficiency and bias, but is far from a panacea against the estimation challenges emerging from the TSCS data structure.

In particular, it has been pointed out that the “Beck-Katz standard” will often produce biased estimates if models do not account for unmeasured factors that cause each country to have its own base rate, or intercept (Green, Kim and Yoon, 2001; Kennedy, 2008). At the same time, it can rarely be assumed that the included covariates eliminate these persistent cross-sectional differences. Throughout the thesis I account for this by following

³¹In addition, the OLS estimates are relatively easy to interpret.

³²Estimating OLS regression on TSCS data may lead to problems stemming from ‘interdependence’ between observations over time and between cross-sectional observations in space.

the standard recommendation of including country fixed effects, which is the same thing as including a dummy variable for each cross-sectional unit. In brief, this approach removes time-invariant country-specific variation. In line with common practice, I combine this with the inclusion of year dummies, to remove variation that is invariant across cross-sectional units but varies over time.

This model combining unit and year fixed effects with the Beck-Katz recommendation of panel corrected standard errors and lagged dependent variable will usually perform very well in terms of unbiasedness (Plümper, Troeger and Manow, 2005; Kennedy, 2008).³³ Yet, it may not always be an appropriate remedy, as the fixed effects estimator is a highly inefficient estimator (Beck et al., 2001; Plümper, Troeger and Manow, 2005; Plümper and Troeger, 2007). Since regression with fixed effects only uses the within-variation in each cross-sectional unit and eliminates the between-variance, it does not allow estimation of time-invariant variables.³⁴ It may also be too inefficient when it comes to estimating effects of variables that have very little within-variance, for instance due to little change over time. This problem is highly relevant for the slowly-moving variables used throughout this study. For instance, political institutions such as democracy rarely change, even in long time series. Despite education having increased rapidly in recent years, as discussed in the introduction, indicators of education also show much more variation across units than over time. As fixed effects will soak up most of the explanatory power of these slowly changing variables, it will be hard to detect statistically significant results (Beck et al., 2001).³⁵

In response to the problem of inefficient fixed effects-estimation, several methodologists recommend using more efficient estimators such as random effects models or the system GMM estimator, which has been shown to perform much better when variables are highly persistent over time (Blundell and Bond, 2000; Clark and Linzer, 2015). The latter method is used in chapters 2, 3 and 4. By doing so, a higher degree of bias is accepted, but this is traded off against higher efficiency that allows sharper inferences. This methodological point is illustrated by a number of empirical studies demonstrating that the system GMM estimator identifies effects that are soaked up by fixed effects

³³One potential challenge to this model is Nickell bias, which may emerge if the time-series is very short (Nickell, 1981).

³⁴Such variables are constant over time either by definition or for the particular selection of cases and period under analysis.

³⁵Moreover, the fixed effects model may actually also increase the likelihood of bias. Researchers often tend to emphasize that the coefficient estimate holds up to fixed effects estimation as evidence that the finding is robust (Plümper, Troeger and Manow, 2005). Yet, although inefficiency due to fixed effects estimation is commonly associated with a lower likelihood of type-I errors, it may in fact also produce biased estimates, since the influence of the error on the estimated coefficients becomes larger as inefficiency increases (Clark and Linzer, 2015). Hence, fixed effects may in worst case produce *both* inefficiency and bias.

models. Notably, Castelló-Clement (2008) revisits the finding that there is no significant relationship between education and democracy, arguing that Acemoglu et al's use of fixed effects models is inappropriate for these variables. Using the system GMM model, they identify a positive and statistically significant effect of education on democracy.³⁶

To sum up, I take a pragmatic approach in response to the potential trade-off between efficient and unbiased results, by running different estimators performing differently in terms of bias and efficiency, to check for coefficient stability. This point touches upon a final methodological discussion important to the empirical analyses in this thesis. In line with Leamer's (1985) recommendation, I always test and show that my results hold up to a variety of robustness tests using different estimators and model specifications. The value of running numerous sensitivity and robustness tests to check for coefficient estimate stability has been questioned by those that emphasize the importance of identifying the “true model” (Angrist and Pischke, 2010). For instance, proponents of randomized studies suggest a sound randomized research design resolves any concerns regarding bias or efficiency, as it will yield the true causal effects of an independent variable. Yet, randomized studies are often infeasible, and without knowing the data-generating process one should make sure not to rely overly on one specification (Leamer, 1985; Kennedy, 2008; Stock, 2010; Sala-i-Martin, 1997). Assessing coefficient stability across numerous model specifications is a pragmatic solution to this challenge.

Threats to estimating causal effects

As discussed above, all the contributions in this thesis rely on non-randomized observational data. Drawing on the “potential outcomes framework” discussed above, this means that being exposed to the treatment – education – is not random. The “treated” units (e.g., the highly educated countries) may therefore be systematically different from the “untreated” units (e.g., countries with low education). This is often referred to as the “selection problem” (Angrist and Pischke, 2008).

There are three important sources of systematic differences between the treatment group and untreated group (the control group). If unsolved, these situations are commonly refereed to as endogeneity problems. First, systematic differences can emerge due to *observable* factors that influence both the likelihood of being treated and the outcome variable. For instance, several socio-economic factors such as income level, urbanization, degree of industrialization and age composition may influence both education levels and authoritarian instability or democratization.

³⁶Moreover, as opposed to Acemoglu et al. (2005), both Heid, Langer and Larch (2012) and Che et al. (2013) identify significant effects of income levels on democracy when using the system GMM estimator rather than a fixed effects model.

Second, endogeneity problems can emerge due to *unobserved* factors that influence both the likelihood of being treated and the outcome (omitted variable bias). For instance, there are likely to be cultural factors and historical events that have been very important to the divergent political-economic development paths of different countries, and that are (close to) impossible to measure systematically, such as colonial experience, cultural heritage or historical institutional patterns (Acemoglu, Johnson and Robinson, 2001; Acemoglu and Robinson, 2006).

Third, they can be due to reverse causality, emerging when the outcome variable influences the treatment. Reverse causation is a highly likely feature when studying the effects of education on democratization. Indeed, there is a vast literature indicating that regime type and democratization influence access to education (see, e.g. Lake and Baum, 2001; Lindert, 2005; Ansell and Samuels, 2010; Acemoglu, Gallego and Robinson, 2014). Instances of authoritarian instability and political protest are also likely to influence education policies, for instance through dictators using education investments to appease potential protesters. This is also studied directly in chapter 3, which analyses the effects of regime type on education quality.

The first type of situations can be solved simply by including the relevant observable factors as control variables in the model, as done throughout the study. The choice of controls must be theoretically justified, however, and the inclusion of inappropriate control variables may induce bias (and inefficiency). First, it is important to avoid including controls influenced by the treatment variable, as these may induce “post-treatment bias”, reducing the estimated effect of the treatment variable (King, 2010). To complicate things, a potential omitted variable is often also influenced by the explanatory variable. For instance, when studying whether the involvement of educated individuals in protest movements promotes non-violent tactics (such as in chapter 6), the size of the protest movement is a plausible omitted variable: Larger movements should make it more likely that educated individuals will be among the participants, and there is also evidence that larger movements are more likely to become non-violent (Chenoweth and Stephan, 2011). At the same time, and as argued in chapter 6, educated individuals may also be especially skilled at building up movements with high potential for success, and this may attract more participants. Hence, campaign size may also be influenced by the education level of the protesters. In such cases, there are no clear methodological answers: While controlling for campaign size may lead to post-treatment bias, excluding it from the models may lead to omitted variable bias. Without knowing what the appropriate model is, I remain agnostic in these cases and present models both with and without the potential confounder (and potential post-treatment variable), to check for coefficient stability.

In addition, one should be aware of controls that are unnecessary, even if they are not

post-treatment controls. Achen (2002) has criticized the practice of throwing all possible controls into a model, labeling it as “garbage can regressions”. He argues that this practice is problematic, because “estimated coefficients can be very sensitive to the inclusion of even slightly correlated controls”. In line with this recommendation, I usually start by presenting simple models in the main empirical analysis, and include additional controls as robustness tests.

As I am often dealing with TSCS data, certain instances of the second problem of omitted variable bias can be solved with the help of country fixed effects or year dummies, both used throughout the thesis. As discussed above, fixed effects estimation may create inefficiency problems, but it will generally reduce problems of omitted country-specific time-invariant variables. In other words, country fixed effects may eliminate heterogeneity between the treatment and the control group that is due to time-invariant effects. And, year dummies will correct for systematic differences between observations due to time trends. Yet, country and year dummies do not deal with omitted variables that vary both over time and across units. Nor do these models account for the third selection problem, reverse causality.

A recommended way of dealing with selection problems in observational data due the second problem of omitted variable bias, or the third problem of reverse causation, is instrumental variable regression (Angrist and Pischke, 2010). This estimation procedure rests on the identification of a valid *instrument*, which is exogenous in the sense that it is not affected by any of the variables included in the model. The instrument must only affect the outcome variable through the treatment variable, conditioned on the included confounders (the exclusion restriction). At the same time, the instrument must be a sufficiently strong predictor of the relevant explanatory variable (Stock and Yogo, 2005a). If these requirements are fulfilled, the instrumental variable regression utilizes the variation in the treatment variable that is explained by the instrument to estimate the effect of the treatment on the outcome.

This procedure is used in several chapters. For instance, chapter 2 utilizes numeracy skills in the 1820s, which are estimated using age-heaping techniques, to instrument for current levels of education quality³⁷. This builds on several studies documenting that numeracy skills tend to be highly persistent over time, and that early numeracy skills affect current levels of education quality (A’Hearn, Baten and Crayen, 2009; Crayen and Baten, 2010). I argue that we can assume that early numeracy skills only affect current levels of democracy through education or any of the other included confounders such as

³⁷Age-heaping techniques relies on the fact that people lacking numeracy skills are not able to state their exact age and therefore reported a rounded age in official documents, rather than the exact age. The share of people who were unable to state an exact age is used to proxy for historical levels of human capital (A’Hearn, Baten and Crayen, 2009; Crayen and Baten, 2010).

gdp per capita, level of industrialization or early experience with democracy.

Another example is the use of variation in historical presence of christian mission stations to instrument for current sub-national education levels in Africa, which is done in chapter 5. Christian mission stations have had a strong persistent impact on education levels through their provision of schools in the areas they settled in (Nunn, 2010). In addition, chapter 5 argues that the assumption can be made that christian missions only affect current protest activities through education levels or included confounders such as the present presence of Christians, the consumption of media and communication technology, infant mortality rate and other proxies for local economic development.

In lack of a robust identification strategy such as a strong, plausibly exogenous instrument, it is not really warranted to interpret the estimates as a causal effect. Yet, measures can be taken to *raise the confidence* that an estimate is picking up a causal relationship. One suggested procedure, used in this thesis, utilizes the coefficient estimate's sensitivity to the *included* confounders to draw inferences about the size of potential bias due to *omitted* confounders (Imbens, 2003; Harada, 2015). More specifically, this procedure, which is described in more detail in chapter 6, generates a sequence of pseudo-random covariates that reduce the coefficient estimate by a certain magnitude. This can be used to asses how large the omitted variable bias has to be to reduce the coefficient estimate by a certain amount, such as to half its original size or to 0. Chapter 6 applies this strategy to assess the threat of potential omitted variables to the identified effects of education on non-violent protest.

Another more pragmatic strategy that may reduce endogeneity problems due to reciprocal effects is to simply choose education indicators less likely to be influenced by the dependent variable(s) – including democratization or mass protest. In chapter 2, I argue in favor of using indicators of education quality rather than quantity. One rationale behind this choice is that indicators of education quality arguably are more appropriate operationalizations of the theory, but this choice is also beneficial because education quality is arguably *less endogenous* to democracy than education quantity. In chapter 3, I argue and present evidence that education quality is indeed not affected by regime type, while there is generally a strong and positive effect of democracy on education quantity. In light of this finding, operating with education quality as independent variable should offer an empirical test less plagued by endogeneity issues.

1.7 Structure of the thesis

Each of the five subsequent chapters investigate the relationship between education and authoritarian regime instability or transitions to democracy. More specifically, the thesis is divided into two parts, each offering different types of empirical tests of the theoretical framework discussed above – highlighting the resources and preferences of educated citizens. Part I revisits the reciprocal relationship between education and democracy at the state level, and takes issue with the common practice of operationalizing education using indicators of education quantity such as school enrollment or average years of achieved education. Part II investigates empirically how education is related to authoritarian instability or mass mobilization, through three different chapters at three levels of disaggregation. Three out of five chapters are single-authored, and two chapters are first-authored (each with one other co-author).

1.7.1 Part I: The reciprocal relationship between education quality and democratization

Before I move on to a more disaggregated research design, Part I gives state-level education indicators a new chance, but takes issue with the way education has been used and operationalized in previous literature.

Chapter 2 revisits the relationship between education and democratization. It starts by detailing and discussing two plausible micro-mechanisms that are often hinted at in the literature: the “preferences-mechanism” and the “opportunities-mechanism”. It shows that these theories yield implications about what the education received by citizens consists of, indicating that high-quality and free education should be more likely to lead to democratization than low-quality and unfree education. Yet, all existing studies rely on indicators of the *amount of education* offered to the population, such as average years of schooling. Taking the above implications seriously, this chapter is the first study to investigate the effects of education quality on democratization. The empirical analysis draws on a recent panel dataset by Angrist (2013) on education quality, measured using student achievement test results. In brief, this investigation yields no evidence that variation in education quality or content is systematically related to education, despite the theoretical expectations. This finding questions our current understanding of the education-democracy relationship, by suggesting that implications of plausible underlying mechanism do not hold up.

When investigating the effects of education levels on regime type and change, education is likely to be endogenous, particularly due to the potential effects of regime type on

education levels. Chapter 3 (co-authored with Carl-Henrik Knutsen) investigates the argument that democracy promotes education, by exploring whether democracies also offer *better education*. We argue that while democracies should be expected to increase access to education (through, e.g., higher school enrollment rates), democracies may actually *not* do better than autocracies when it comes to boosting education quality. Not only is education quality inherently difficult to promote, it is also a less transparent policy outcome than education access, in the sense that it is harder to monitor and less visible to voters. Hence, a democratic leader may not face stronger incentives than a dictator to invest in education quality. This argument is supported using data on education quality from Angrist (2013). This finding also adds to the analysis conducted in chapter 2, by indicating that education quality, as opposed to education quantity is likely not endogenous to democracy. Hence, the finding in chapter 2 may depart from previous literature precisely because indicators of education quality are less endogenous to democracy, while the finding from extant studies that education levels are related to democracy may really be driven by reciprocal effects. In sum, by accounting for variations in the quality of education, Part I offers evidence suggesting that we should reconsider prominent theories about the relationship between education and democracy at the country level.

1.7.2 Part II: Disaggregating the relationship between education and authoritarian regime instability

The second part of the dissertation investigates how education affects authoritarian regime instability and collective action in authoritarian regimes. Here, I respond to the plea for more disaggregated empirical tests – in order to avoid ecological fallacies and move closer to the level of analysis where the proposed micro-mechanisms should play out.

Chapter 4 adds to existing literature by shifting the focus to studying the relationship between education and regime breakdown in dictatorships, rather than only focusing on democratic transitions. The chapter starts by investigating this relationship at the macro level. I find no evidence that high levels of nationwide education enrollment is positively related to authoritarian breakdown. Rather, there is evidence that high education levels actually *reduce* the likelihood of authoritarian breakdown, and I suggest that this could be due to factors such as the strength and capacity of highly educated regimes. However, utilizing new data on the education profile of protest campaigns, I show that ongoing protest movements with educated protesters *increase* the likelihood of authoritarian breakdown. Moreover, it shows that the (non-robust) evidence for the hypothesis that high education levels are associated with democratization identified in previous studies disappears completely when controlling for educated protest movements.

This indicates that any effect of education on democratization goes *through* protest by educated individuals.

Drawing on the findings in chapter 4, displaying the benefits of studying the relationship between education and authoritarian instability at a more disaggregated level, chapters 5 and 6 continue to offer more disaggregated tests of the relationship between education and authoritarian instability. In chapter 5 (co-authored with Tore Wig), we investigate the relationship between education and contentious protest at the sub-national level. Drawing on the theoretical framework discussed above, the chapter argues that education should increase the frequency of mass protest, by alleviating collective action problems and motivating mass opposition, particularly in autocracies. Combining geo-referenced survey data on education from the DHS with geo-referenced protest event data from the SCAD dataset (Salehyan et al., 2012), the empirical analysis demonstrates that areas with more educated people have higher levels of protest activity, and that this effect is particularly strong in authoritarian regimes. The latter finding supports the grievance mechanism, which suggests that schooling breeds collective action because it stimulates anti-authoritarian grievances.

Finally, chapter 6 operates at an even more disaggregated level of analysis, by studying the relevant actors, which in this case are political protest campaigns. Building on arguments from the literature on non-violent and violent protest, it asks whether protest campaigns consisting of or dominated by educated individuals are more likely to 1) turn nonviolent and 2) succeed in terms of achieving their goals. Utilizing a novel data set on the education profile of the anti-regime protest campaigns listed NAVCO 1.0, I find strong and robust evidence that educated campaigns are more likely to turn nonviolent. Moreover, I find some evidence that these campaigns are more likely to succeed, but this is *only* due to their nonviolent dispositions. Hence, the non-violent dispositions of educated protesters, and the fact that non-violent movements are more likely to promote democratization, may explain why they are potent agents of democratization.

1.8 Findings

This section rounds off by highlighting five key insights derived from the combination of the empirical contributions described above.

First, based on a number of results, this dissertation throws cold water on the proposition that dictatorships with high education levels are more likely to break down and experience transitions to democracy. Despite the implications of convincing theoretical mechanisms tying education to regime change, I find no evidence that high-quality school-

ing is positively related to democratization. The only evidence identified in this thesis that supports the proposed education-democratization link at the country level, is some (but weak) evidence that indicators of education *quantity* are related to democratization, but such indicators are highly likely to be endogenous to democracy and democratization. Moreover, the findings in the thesis indicate that *all* of the (weak) evidence suggesting a positive effect of education on democratization, identified in line with previous literature, disappears completely when controlling for ongoing protest campaigns conducted by educated protest campaigns. Hence, this suggests that if there is any effect of education levels on the likelihood of democratization, it goes through educated protesters being more likely than other protesters to induce democratic transitions. In other words, educated groups of individuals may matter for democratization, but not average education levels.

Second, this thesis is the first to show that high education levels may actually be associated with a lower likelihood of authoritarian regime breakdown. I suggest that this may be due to the fact that highly educated dictatorships are also likely to have high state capacity, and can therefore offer public policies and services to its population to boost loyalty. At the same time, these regimes may have the capacity and resources to be able to discourage opposition through (the threat of) repression. This interpretation finds some support in results showing that the negative effect of education levels on authoritarian regime instability is reduced when controlling for proxies for state capacity (in chapter 4). This suggests that state-level indicators of education levels may be too imprecise to capture the proposed mechanisms, as education levels may proxy for and be related to numerous other processes and dynamics happening at the state level. Hence, those predicting the downfall of current dictatorships with reference to their high average education levels, may be missing important nuances.

Third, those that believe in the democratizing potential of education may still be on the right track after all, although not in ways captured by the simple aggregated test of the link between education levels and democratic transition. This thesis finds evidence that ongoing protest campaigns carried out by educated protesters (as opposed to non-educated protesters) increases the likelihood of authoritarian breakdown, as well as of democratic upturns in autocracies. Indeed, the analysis presented in chapter 4 suggests that predicted probability of autocratic breakdown increases from 0.06 if there is an ongoing protest campaign largely consisting of uneducated protesters to 0.26 if a protest campaign that originated among and is dominated by students or educated groups is on-going. This finding supports the micro-mechanisms that I discussed above, indicating how education is a resource and source of capacity that may increase the likelihood of successful collective action.

Fourth, and consistent with mechanisms specifying how education should increase opportunities and strengthen preferences for anti-regime protests in non-democracies, the thesis presents evidence that educated (sub-national) areas are associated with a higher likelihood of protest. This finding reinforces the previous point: When moving closer to the level of analysis where the mechanisms are expected to play out, and using measures of education that come closer to capturing the education levels of the relevant groups, education is positively related to protest. This finding is also robust to instrumental variable regression.

Finally, I argue and find evidence that protest campaigns consisting of educated protesters are more likely to turn non-violent. Due to this non-violent disposition, educated protest movements are also more likely to succeed in terms of achieving their main goals – usually regime change. This finding validates implications from the common proposition that education is negatively related to violent conflict, and does so at the protest group level. At the same time, it suggests that the reason why protest campaigns by educated protesters are more likely to produce authoritarian breakdown and transitions to democracy than other protest movements, is because they are more likely to use non-violence, which according to existing studies is again likely to produce success (Chenoweth and Stephan, 2011).

1.9 Implications

1.9.1 Implications for future research

The findings in this thesis have important implications for future research on democratization and political instability. An important take-home lesson is that existing research has suffered from a limited understanding of the substantive content of average education indicators. Even if it was the case that education generally makes citizens more likely to promote democratization, the various findings in this dissertation suggest that a high average education level may capture much more than simply the population's capacity and propensity for anti-regime resistance: For instance, it can signal that the regime enjoys a high level of state capacity, or that the dictator enjoys a high degree of safety that allow him to prioritize education policies rather than for instance security measures (as discussed in chapter 4). Or, it may hide substantial within-country variation, for instance due to geographical clustering or because supporters of the regime receive many years of education, while the rest of the population has little or no education (discussed in chapter 5). Given this lack of certainty regarding what a high education level cap-

tures, coefficient estimates from regression models investigating the relationship between average education levels and regime change should be interpreted with caution.

On this background, future studies should incorporate the role of actors in explanations of mass protest and regime change – both theoretically and empirically. This point applies not only to the education-democratization literature, but also to the general literature on socio-economic determinants of democracy. For instance, when assessing the relationship between urbanization and democratization, existing studies rely on an aggregate indicator of how urban the country is on average (Wallace, 2013). Yet, this test may produce invalid inferences if the supporting mechanism is political opposition being more likely to emerge in urban settings. Even if a country is largely urban, the anti-regime resistance may still arise in rural areas, and vice versa: Even in predominantly rural countries, the political opposition may emerge in cities. Hence, it would be more appropriate to investigate the urbanization-mass protest link at a more disaggregated level, for instance by exploring whether urban areas are more prone to protest, or whether individuals living in cities are more likely to join anti-regime resistance.

In extension of the arguments made and findings presented in the five consecutive chapters, there are many fruitful ways in which future research can dig deeper into the relationship between socio-economic development and regime change in a more disaggregated manner. I highlight three avenues for future research. First, this dissertation illustrates the importance of collecting (and analyzing) more data on *characteristics of relevant actors* in authoritarian regimes, including on aspects such as on profession, location, income level and social background. These actors include the members of the political opposition, manifested through for instance participants in mass protest movements and coup d'etats, as well as members of the regime's support base, or "ruling coalition". For instance, in order to properly test theories arguing that high levels of inequality promote democratization because poor people demand redistribution (e.g., Acemoglu and Robinson, 2006), aggregate indicators of economic inequality are highly imperfect proxies. Rather, we want to determine whether those who work to promote regime change are poorer than the rest of society, leading them to prefer economic redistribution.

Second, I have hinted at how the behavior of the political opposition is influenced by the behavior of the dictator, and vice versa. Hence, it is necessary to explore further how aspects of socio-economic modernization impact on the choices and behavior of the authoritarian regime or the dictator. For instance, dictators can influence education levels (suggesting that education is likely endogenous), but dictators may also adjust their behavior to prevent potential life-threatening mass opposition by educated citizens. If the dictator buys the finding in this dissertation that educated protesters are more efficient at overthrowing regimes, he may carry out targeted repression against highly educated

individuals that may act as agents of mass resistance, such as university professors or journalists. Or, he may carry out deliberate efforts at co-opting educated groups through e.g. higher salaries or targeted social services. This suggests that we should also study how education impacts on authoritarian strategies for survival, including how it impacts on the use of repression and co-option.

Third, this dissertation has discussed two main theoretical mechanisms – related to opportunities and preferences. These mechanisms yield similar implications regarding the main outcomes of interest discussed here – such as more protest, and higher likelihood of authoritarian breakdown. Future research should draw additional implications from the relevant theoretical framework that can be used to further tease out if both, or any, of these two mechanisms are at play. One potential way to do this is suggested in chapter 5, which shows that the effect of education on protest is stronger in autocratic regimes, indicating that preferences against autocracy may partly be driving the result. Similar interactions between education and relevant contextual conditions such as regime type could also be tested at the regime level or the protest level, to explore whether education has a stronger impact on regime change or protest when anti-authoritarian preferences should be especially vital.

1.9.2 Policy implications

The belief that investing in mass schooling and education quality will promote democratization in authoritarian regimes has been deeply influential in policy circles and the development community. The emphasis on education policies by international organizations such as the UN, the World Bank, and donor countries has partly been explicitly justified with reference to goals such as strengthening civil society and promoting the prospects of democratization.³⁸ Yet, despite the plausibility of the education-democratization link, this recommendation is currently based on an unsettled empirical foundation, as indicated by the literature review at the beginning of this chapter, showing that there is currently no clear agreement on whether education promotes democratization. Moreover, those studies that find supporting evidence for this link, do not offer an explanation of exactly how and why education should be related to institutional transitions. This includes whether it is due to a causal effect or how education influences citizens and their potential for promoting regime change.

This dissertation adds important insights that should be valuable to policy-makers, both those working to promote democracy and those concerned with education policies

³⁸Of course, this is not the only reason for why education investments should be prioritized - education is unquestionably a major benefit and source of human empowerment in itself, at the same as it commonly considered a major determinant of economic development (see, e.g. Cohen and Soto, 2007)

as a tool for economic development. First, it has important implications for so-called third-parties in events of political unrest in dictatorships, such as foreign governments, international organizations and international security alliances, considering whether to throw their support behind anti-regime protest movement or not. In a nutshell, it suggests that such actors should consider the social background of protesters when deciding whether to offer their support. In many cases mass protest campaigns fail, or even worse, culminate in violent clashes with the government or (enduring) civil war, such as in Libya or in Syria in recent years. Other times, mass opposition leads to the installation of (stable) democracy, such as in Tunisia after 2011. This thesis indicates that protest movements consisting of educated protesters should be more likely to choose non-violent methods, and because of this, more likely to succeed. This finding suggests that third parties committed to avoiding supporting movements that may spur civil war should be well advised to consider who the participants in mass protest movements are when evaluating their prospects for success.

Finally, my findings suggest that average education levels should not be used as a guiding principle to assess the prospects of anti-regime opposition and democratization. Rather, if the goal is to promote democratization through education, policy-makers should focus on *who* the education is offered to. Conversely, the findings may imply that a low education level is not necessarily an impediment to the emergence of democracy. Several studies, drawing on modernization theory, have argued that democracy should not be promoted nor installed in countries that do not score sufficiently high on indicators of socio-economic development such as education. Yet, my findings suggest that the relevant characteristic is not education at the country level per se, but rather the educational background of the relevant actors, and these actors need not be representative of the population as such. Even if a country scores very low on indicators such as enrollment and average years of education, educated citizens may still be able to conduct successful protest movements culminating in democratization.

Part I

**The reciprocal relationship between
education quality and democracy**

2 What do they learn? Education quality and democratization

Sirianne Dahlum

Abstract:

Does an educated population spur democratization? Numerous studies investigate this proposition using indicators of time spent in schools to measure education, but none have accounted for *what students learn*. Yet, the widely held argument, drawn from modernization theory, that educated citizens are inclined to oppose dictatorships and initiate democratization assumes that the schooling received is of a certain quality and exposes students to reflection rather than indoctrination. This study investigates the relationship between education and democratization by accounting for variations in *education quality*, measured using a recent panel data set of student achievement tests from 118 countries. It also looks at variations in the autonomy of school curriculums and types of tertiary education. Contrary to expectations, no robust evidence is identified that high-quality schooling promotes democracy or democratic transitions. This finding challenges plausible mechanisms in modernization theory, by suggesting that it is not the learning outcomes of education that matters for democracy.

Publication status:

Submitted

2.1 Introduction

All current regimes provide some kind of public education to their citizens, including at the university level. The share of citizens who receive schooling is also in rapid expansion, especially in the developing world, where primary school enrollment has almost doubled in the past 30 years (Unesco, 2015). Yet, this expansion in schooling access hides dramatic variations in education quality and contents. This was stressed, for instance, in a recent Unesco (2014) report, claiming there is a “learning crisis” based on the finding that one third of primary school age children are not learning basic skills including reading and writing, even if they have been to school.

Against this backdrop, I analyze how the prospects of democratic transitions are affected by what students learn in schools and universities. Modernization theory postulates that more economically developed societies will experience the downfall of dictatorships and the rise of democracy, and it is commonly argued that the trigger is a more educated population, more willing and able to overthrow dictators. Several empirical studies find evidence consistent with this claim, demonstrating that indicators of education are systematically related to democratic transitions (see e.g. Glaeser et al., 2004; Glaeser, Ponzetto and Shleifer, 2007; Sanborn and Thyne, 2014; Murtin and Wacziarg, 2014), but these findings are challenged by studies maintaining that the education-democracy correlation is rather due to democracies being more prone to invest in education (see e.g. Acemoglu et al., 2005; Acemoglu, Gallego and Robinson, 2014). Yet, all these contributions rely on indicators of *how much time students spend in school*, such as enrollment rates or the average years of schooling, that do not adequately capture what students learn.

This paper is the first to empirically take into account variations in schooling quality and content in the study of the education-democracy link. It discusses two plausible mechanisms relating education to democratization, showing how they assume that education affects political behavior and regime change through *what student learn in classrooms*, by producing more assertive and demanding citizens. Hence, indicators of education quality and contents should arguably be better at capturing the potential for regime change due to education than a more distant proxy for learning such as years of education, which previous studies have relied upon. Looking at what students learn should therefore offer a more appropriate empirical test of the education-democratization link, that may help to resolve the ongoing debate between modernization theorists and their critics.

Based on extant literature, I outline two theoretical mechanisms tying education to democratization. First, a *capacity-enhancing* mechanism, according to which the skills and knowledge attained through education bolster citizens’ capacity for carrying out acts

of political resistance and push for democratization. Second, a *preferences-enhancing* mechanism, through which educated citizens develop preferences for civil liberties and democracy. As I show, both mechanisms make certain assumptions about what is being taught in schools, that clearly are not always met in schooling systems around the world. For instance, these arguments assume that people are offered education of a certain quality, that stimulates skills and reflection. Yet, there is often a gap between education quantity and education quality. For instance, it has been estimated that out of Africa's nearly 128 million school-aged children, 111 attend school, but 37 million of these learn so little in school that they will "not be much better off than those kids who never attend school" (Van Fleet, 2012). Furthermore, these mechanisms assume that educational institutions have a certain degree of autonomy from governments. Yet, many regimes intervene heavily in the curriculum of both schools and universities, such as the Ahmadinejad administration in Iran, that set out to "purify" university curricula from secularist and individualistic thought, and replace it by more mandatory courses on Qu'ran and Islamic family morality.

To capture variations in learning outcomes, the paper relies primarily on an indicator of *education quality*, operationalized using international student achievement tests from 118 countries made comparable by Angrist (2013). In addition to capturing learning outcomes that should promote democratization, operating with education quality has another advantage: As I argue below (and show in chapter 3), education quality is less likely to be affected by democracy than indicators of "education quantity" (such as years of schooling), suggesting that this model offers a test of the education-democracy relationship that is less vulnerable to endogeneity issues. Yet, as endogeneity can not be ruled out completely, I also expand on models assuming education quality is exogenous using an instrumental variable set-up with numeracy skills in the 1820s to instrument for current levels of education quality. While high-quality education is likely to boost capacity for collective action and preferences for democracy, it does not guarantee a free and critical curriculum. To account for this I also use two additional indicators of variations in learning outcomes. First, I use a proxy for *education autonomy*, capturing the extent to which the education is free from government intervention. Second, I use a measure of *education type*, indicating to what extent certain academic disciplines are prioritized, measured using data on share of graduates from social sciences, humanities, sciences or engineering.¹

Contrary to implications of the presented mechanisms trying education to democratization, I find very little evidence that variations in education quality and content

¹This paper is concerned with all education levels (including primary, secondary and tertiary), but some of these indicators naturally only capture variations in one of these, such as education type, concerned with tertiary education contents.

matter. More specifically, I find little evidence that high-quality education promotes higher democracy levels, transitions to democracy or democratic survival. Nor do I find that education autonomy or type matters for democracy. To unpack this null-finding further, alternative mechanisms are also investigated, such as whether the effect of education quality on democracy is conditioned on enrollment levels. It can be hypothesized that high-quality education only promotes democratization when offered to a large share of citizens, but there is no evidence for this. There is, on the other hand, some evidence that education quality combined with low enrollment rates *reduces* the likelihood of authoritarian breakdown, and I suggest that this may be due to dictators restricting high-quality education to a selected few to boost loyalty and regime survival.

I discuss two potential interpretations of this null-finding. First, the hypothesized effect of education, which is supported by several studies, could be due to *other mechanisms* than the ones focused on here. For instance, education and schools may foster regime change by bringing together people from different social segments in schooling institutions, thereby stimulating "social networks" that are conducive to anti-regime mobilization. Alternatively, I find no support for this hypothesis simply because education *does not promote democratization*. This is in line with the view that previous findings identifying an effect of education quantity on democracy are driven by, for instance, an effect of democracy on schooling expansions.

The study proceeds as follows. In section 2.2, I discuss the literature on education and democratization, including numerous studies relying on indicators of education quantity. Against this background, section 2.3 lays out a more detailed account of the mechanisms underpinning the education-democracy link in accounts of modernization theory, and show why and in what ways education quality and contents should matter for democratization. Section 2.4 discusses the research design and data, while section 2.5 presents the findings from the main empirical analysis and discusses some potential explanations for the null-findings from the main analysis. Finally, section 2.6 discusses and interprets the findings from previous sections, and section 2.7 concludes.

2.2 Moving from education quantity to quality

The notion that education is one of the main ingredients in a well-functioning democracy goes all the way back to Aristotle. More recently, Dahl (1971) argued that education and literacy stimulates political participation, which is crucial to maintaining representative institutions (see also Dewey, 1938; Lerner, 1958; Huntington, 1991). The education-democracy relationship was elaborated on in Lipset's (1959) modernization theory, maintaining that affluent and economically developed societies are more likely to experience

democracy, and that education is a crucial component of this development.

Numerous studies have investigated this proposition empirically, yielding surprisingly mixed results. On the one hand, there is evidence that an educated population is associated with higher probabilities of democratic transitions, in line with modernization theory (Papaioannou and Siourounis, 2008; Sanborn and Thyne, 2014; Murtin and Wacziarg, 2014; Glaeser et al., 2004). For instance, Glaeser et al. (2004) argues that schooling is a causal explanation of differences in democracy levels, and also of other variations in political institutions. A thorough study by Murtin and Wacziarg (2014) corroborates this finding, relying on a wide range of model specifications and a time-series extending back to the 19th century.

On the other hand, these positive findings have been challenged by those arguing that both education and economic growth are induced by democracy, rather than the other way around (Acemoglu and Robinson, 2001; Acemoglu et al., 2005; Stasavage, 2005; Acemoglu, Gallego and Robinson, 2014). Acemoglu and Robinson (2001) use the potential mortality of European settlers to instrument for institutions, assuming that the degree of settler mortality affected the colonization strategy pursued by the Europeans, which again created “inclusive institutions”. They find that institutions are the fundamental cause of long-run development including human capital formation. In line with this, Acemoglu, Gallego and Robinson (2014) find similar evidence using the historical presence of Christian mission stations to instrument for early levels of human capital, showing that education does not have an effect on democratic institutions.

The latter findings have, in turn, been criticized on methodological grounds. For instance, using estimators that are more appropriate for analysing slowly-moving variables, such as education, and considering how education is distributed, Castelló-Climent (2008) identifies a positive relationship between education and democracy (see also Bobba and Coviello, 2007). Moreover Glaeser et al. (2004) criticizes the identification strategy of Acemoglu and Robinson (2001), arguing that what Europeans primarily brought to their colonies was not inclusive institutions but rather human capital, which again promoted democratic institutions. Hence, there is an ongoing debate regarding the empirical merits of the education-democratization link.

All the empirical contributions discussed above rely on empirical tests of the effects of the *amount of schooling* offered to a population, either measured as the average years of education achieved, enrollment rates, or share of citizens having completed primary, secondary, or tertiary education. In other words, they do not account for the content and quality of schooling. Meanwhile, much of current scholarship also assumes that it is *what* students learn that influences their behavior and consequently the prospects for regime change, although few have offered detailed accounts of the causal mechanisms.

For example, among classic proponents of modernization theory, education is often considered to induce democratization through changes in political culture (see e.g. Almond and Verba, 1963; Inglehart and Welzel, 2005; Glaeser, Ponzetto and Shleifer, 2007). This again builds on studies showing that education stimulates political participation by creating more effective political actors (Galston, 2001; Hillygus, 2005). Inglehart and Welzel (2005) maintain a recent version of modernization theory, arguing that socio-economic modernization, and in particular education, fosters individual autonomy, empowerment and liberal inclinations. They argue that such attitudes will induce demands for more political participation and freedom, thereby promoting democracy. Hence, the proposed effect of education on regime change according to this theory is expected to run through changes in attitudes brought about by education, and it seems plausible that the content of schooling will influence whether this mechanism takes place. Meanwhile, Glaeser, Ponzetto and Shleifer (2007) is one of few attempts at specifying a comprehensive theoretical model and unified account of how education fosters democracy, arguing that education raises the potential for citizens to participate in collective action, while also making democracy more attractive. On their account, democracy is clearly influenced by certain learning outputs, as education “teach their students that political participation is good” and because “educated people are better able to express what they know, to inform and to persuade” (Glaeser, Ponzetto and Shleifer, 2007, p. 81). Yet, variations in learning outcomes are not accounted for empirically in any of these contributions.

In contrast, studies exploring determinants of economic growth have recently focused on variations in education quality, drawing on recent data sets comparing international students performance tests in cognitive skills such as mathematics and reading across a large number of countries (see e.g. Angrist, 2013). For instance, Hanushek et al. (2008) identify a positive effect of education quality on economic growth that dwarfs the association between indicators of education quantity and economic growth (see also Hanushek and Kimko, 2000). This finding is reiterated using instrumental variable regression analysis by Jamison, Jamison and Hanushek (2007) and Hanushek and Woessmann (2012a). Studies utilizing data on comparative education quality also indicate that the human capital stock in developing studies has been exaggerated by indicators of education quantity. For instance, Hanushek and Woessmann (2012b) show that the Latin American “growth puzzle” – referring to the discrepancy between relatively high education levels and low economic growth – is solved when accounting for the poor quality of Latin American schooling, that predict growth levels quite accurately. Meanwhile, Hanushek and Woessmann (2011) demonstrate that cognitive skills, as opposed to education quantity, can account for the very large differences in long-run growth performance across OECD countries.

Hence, the growth literature suggests that education quality outweighs the importance of education quantity when it comes to explaining variations in economic growth. In this study, I apply this insight to the study of education and political regime type (and change). Looking at schooling contents also has an additional advantage: As opposed to indicators of education quantity, measures of quality are less likely to be endogenous to democracy and democratic transitions. Studies suggest that democracies are more likely to offer widespread school enrollment or many years of education (Acemoglu, Gallego and Robinson, 2014; Stasavage, 2005), but that they do not necessarily boost education quality (Harding and Stasavage, 2014).² In general, it can be argued that for a (young) democratic government, expanding education enrollment is a very powerful way of satisfying constituencies, at the same time as expanding school-enrollment is an available and relatively straight-forward policy for governments with sufficient financial leverage. Improving education quality, on the other hand, is less visible to voters and hence less likely to be prioritized by elected officials. It is also much more difficult to achieve, especially when schooling expansion puts pressure on teaching facilities. That democracy does not promote education quality is indeed documented in detail in chapter 3.³ In short, an indicator of education quality will likely offer an empirical test that is a) more theoretically relevant and b) less vulnerable to endogeneity bias than indicators of education quantity.

To summarize, there are two diverging accounts of how education relates to democracy, with proponents of modernization theory such as Glaeser, Ponzetto and Shleifer (2007) maintaining that education promotes democracy, while critics, such as Acemoglu, Gallego and Robinson (2014), arguing that education rather is a product of democratic institutions. The rest of this paper adds to the literature by specifying (and subsequently testing) how education quality and content may impact on democracy. As this framework should offer a more appropriate test of the relevant theory, it may help to settle questions regarding modernization theory's empirical merits.

²This is illustrated by Stasavage's (2005) account of how shifts to multiparty competition in many African countries resulted in increased education spending but not improvements in quality. For instance, he shows that although the Ugandan school reform of 1996 secured universal primary education, the massive expansion in education enrollment was followed by a drop in education quality, due to subsequent shortage of basic materials and a dramatic increase in pupil-teacher ratios (Stasavage, 2005).

³Even some of today's rich, advanced democracies are suffering to improve education quality, and studies indicate no easy relationship between education spending and learning outcomes (Hanushek and Kimko, 2000).

2.3 Why education quality and content should matter

As discussed, few extant studies explicitly specify the causal mechanism tying education to regime change. This section takes modernization theory at face value, and details two crucial theoretical mechanisms underpinning the education-democratization link. From these mechanisms I derive hypotheses specifying why and how we would expect variations in education quality and content to impact on democratization – *if* modernization theory holds.

2.3.1 The capacity-enhancing effect of education

Carrying out acts of political resistance – especially in authoritarian regimes where genuine political opposition is deemed illegitimate *per se* – is a difficult and potentially very risky enterprise. Hence, it requires a certain amount of capacity. Education, at the primary, secondary and tertiary level, can be considered a resource enhancing people's skills, knowledge and thereby capacity for organizing efficient resistance, for instance through mass protests or elite coups. I call this the *capacity-enhancing effect* of education.

First of all, schooling should strengthen language skills. Not only does it yield basic literacy skills, it should also improve people's ability to communicate effectively and articulate complex ideas, through essay-writing, discussions and presentations. Communication skills enable individuals to take advantage of the internet as a communication channel, and to articulate experiences and opinions on political affairs. These capacities are crucial when working to build a successful opposition movement, including mobilizing people and communicating effectively with other opposition members and allies. Language-competent individuals will also be more apt to obtain information from abroad, including e.g. information about foreign anti-dictatorship movements and their strategies, and uncensored information about home governments and their policies.

Second, schooling should offer knowledge about the world and society, including knowledge about the political system, the current regime and their policies. This is supported by evidence that education is correlated with knowledge about politics (Almond and Verba 1963, Verba and Nie 1972, Hanushek 2002, Bratton et al. 2005, Mattes and Bratton 2007) and interest in acquiring more information (Dee 2004). Some level of knowledge about the current state of affairs is an absolute minimum for participating in anti-government opposition. Schooling should also expose students to analytical and systematic thinking, enabling them to process and make sense of large quantities of information, such as that acquired from the media.

Finally, education is likely to strengthen the ability to organize collective action through enhancing organizational skills. Schooling helps to teach cooperation through the means of collective learning tasks, joint activities such as games or school plays and participation in the school community (Newton, 1997). Such activities should have a socializing function through teaching people to successfully interact to solve complex tasks. This includes understanding and appreciating others' points of view, tolerating diverging opinions, as well as working with others to obtain certain goals.

However, the extent to which the capacity-enhancing mechanism will materialize should depend on the schooling quality. The skills and abilities described above will *not* simply emerge from spending many years in a classroom. A significant share of students finish many years of schooling lacking even the most basic skills such as reading and simple mathematics (Unesco, 2014). An illiterate population, lacking knowledge about the political system and unable to access or make sense of modern media channels, will arguably be less likely to overcome collective action problems and create successful opposition in non-democratic regimes. Moreover, if education is to induce a better understanding of society and political affairs, it has to offer genuine knowledge about the world, rather than e.g. religious or political doctrines.

2.3.2 The democratic-preferences effect of education

In addition to strengthening capacities for promoting democratization, education may affect citizens' motivations for participating in political activity and opposing dictatorships, by nurturing stronger preferences for democratic government. I call this the *democratic-preferences effect* of education.

One version of this argument maintains that educated people, through both the schooling system and potentially more autonomous jobs, grow used to exercising initiative and independent thinking. As a result of this, they will also demand more autonomy and freedom in the political system (Inglehart and Welzel, 2005). Relatedly, a number of empirical studies suggest that educated individuals are more supportive of democracy as a political system and more likely to emphasize civil liberties, individual autonomy and empowerment (Welzel, 2013; Friedman et al., 2011). This implies that educated citizens should have stronger preferences for democracy, but also that they should be more prone to joining regime-challenging activities against authoritarian rule, in order to promote the desired changes. Given that open opposition against dictatorships comes with a certain risk, an individual strongly committed to democratic ideals will be more likely to accept that risk, as she will place a higher value on the potential outcome (regime change) of successful opposition.

This mechanism rests on certain assumptions about the content of schooling. In

particular, it assumes that students are exposed to critical thinking, expanded knowledge about the surrounding world and society, and new perspectives and opinions. This should develop students' ability to think independently and critically assess information.

Simply attending a school for several years – even if the school enhances sophisticated reading and maths skills – does not necessarily train students to think for themselves. For instance, an education system whose purpose is to indoctrinate students in regime ideology (or theology), or induce regime loyalty, may not stimulate the type of critical thinking and liberal inclinations that nurture preferences for democracy. Indeed, it may even have the opposite effect, as it presents predefined doctrines and encourages acquiescence rather than questioning official ideology. Josef Stalin reportedly grasped this, arguing that “Education is a weapon, whose effects depend on who holds it in his hand and at whom it is aimed” (Wells, 1937). That education may be used as a tool to strengthen regime loyalty and shape the beliefs of citizens has been documented in several studies (see e.g. Lott, 1999; Pritchett, 2002).

There is much anecdotal evidence of education being used by political leaders in manipulative manners. For instance, it has been argued that historical primary school expansion in Western Europe was carried out to strengthen national-building (Alesina and Reich, 2013) and to “inculcate habits of conformity, discipline and morality that would counter the widespread problems of social disorder” (Green, 2013, p.59). Education is commonly used for indoctrination purposes in modern dictatorships, too. For instance, in China, where schooling quality is generally very high according to reading and mathematics tests, school curriculums are tightly directed by the government. This may have intensified under the current leadership, who recently announced a ban on university textbooks containing “Western values” - except from “Western-imported marxism” (Economist, 2015). Schooling in these forms should not train people to think independently.

Furthermore, certain types of education, even if high-quality and not distorted by governments, may not stimulate the same levels of social awareness and critical thinking as others. For instance, social sciences, humanities and verbal skills are shown to be particularly conducive to political engagement at the individual level (see Hillygus, 2005; Paterson, 2009). Educations in other fields such as the natural sciences and engineering may very well stimulate systematic thinking and problem-solving skills, but may not necessarily encourage political awareness, which is identified as an important determinant of political opposition. Social sciences and humanities on the other hand, usually encourage reflection and independent reasoning about society. There is also anecdotal evidence that certain types of education are more conducive to political opposition in dictatorships. For instance, in a Foreign Affairs article called “Did liberal studies enable Hong Kong’s

youth awakening?” Tsoi (2014) suggests that the introduction of mandatory liberal studies in Hong Kong’s secondary schools in 2009 is to be blamed for the outbreak of (mainly student) protests in September 2014.

Hence, the above noted mechanisms both assume that what students learn in schools should matter for democratization. The primary implication is the expectation that variations in education quality should influence the prospects for democratization. Students may not gain the capacity to carry out collective action if the quality of schooling is low, nor will they develop preferences for democracy. This yields the following hypothesis:

Hypothesis 1. *High-quality education raises the probability of democratization*

To this hypothesis, two nuances can be added. First, although education quality is likely to correlate with a free curriculum⁴, the two will not always coincide. Several non-democratic countries have been able to produce advanced skills in reading, maths and sciences while at the same time restricting free discussion and access to ideas antithetical to official doctrine – China being a prime example (Zhao, 2014). This suggests that an indicator of education quality is perhaps not fully sufficient to capture the variation in education relevant to explaining democratization, and by implication, that a certain degree of autonomy from the regime is also required. This yields the following hypothesis:

Hypothesis 2. *Academic autonomy raises the likelihood of democratization*

Finally, certain types of tertiary education have a content that is more likely to stimulate critical thinking and preferences for democracy. This yields a third hypothesis:

Hypothesis 3. *Countries with a large share of students in social sciences and humanities should be more likely to experience democratization than countries with an emphasis on natural sciences and engineering*

To summarize, this section has considered two mechanisms tying education to democratization, and shown how they imply that three aspects of schooling – quality, autonomy and type – should matter for democratization. The next section discusses the data used to test the stipulated hypotheses.

2.4 Data

2.4.1 Education quality

As noted, this paper considers variations learning outcomes to explain democracy, and therefore relies on the availability of cross-country indicators of schooling contents. The

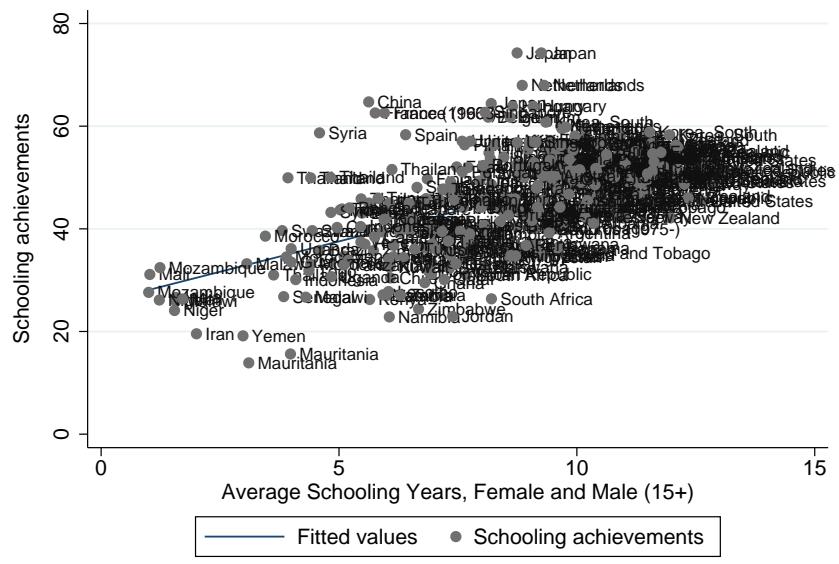
⁴For instance, a heavy curricular influence from regime ideology may reduce the amount of time devoted to developing skills and knowledge

mechanisms discussed above should apply to both primary, secondary and tertiary level education. However, due to data availability, hypothesis 1 will be tested using data from primary and secondary schooling, while hypothesis 2 and 3 will be tested on data from tertiary education.

To measure *education quality*, an indicator of the output of schooling in terms of skills and knowledge is required, and I therefore draw on student achievement tests measuring students' knowledge and skills. The recent dataset by Angrist (2013) compares international student achievement tests across 128 countries around the world from 1965 to 2010 in 5-year intervals, allowing for studying variations in student performance in a panel data set-up. In addition to an aggregated test score for each country-interval, this dataset contains scores on three different subjects (mathematics, reading and science) and at two different education levels (primary and secondary). In terms of cross-sectional coverage, this data is a clear improvement compared to previous data sets such as that developed by Hanushek and Woessmann (2012a), covering 50 countries across 40 years. A problem for the comparative study of education quality is the relatively limited coverage of all international achievement tests. For instance, the widely known Program for International Student Assessment (PISA) test only covers 54 countries, and few of these are from developing countries. Building on Hanushek and Woessmann (2012a), the Angrist dataset solves this by developing a method for linking regional assessments such as "South and Eastern African consortium for Monitoring Educational Quality" (SACMEQ) and the "Latin American UNESCO Laboratorio Latinoamericano para la Evaluacion de la Calidad de la Educacion" to international assessments such as PISA and Trends in International Mathematics and Science Study (TIMSS). Regional tests are linked to international tests using countries participating in both as reference points. Furthermore, different international assessment tests are linked using the United States – having participated in almost all tests since their inception – as reference point. The Angrist dataset provides indicators of student achievements in reading, mathematics and science, but in the main analysis I combine these measures into two composite indicators – one measuring student achievements in primary education and the other measuring achievements in secondary education. More thorough discussions of reliability and the potential of measurement errors in the education quality indicators are included in section 3.4.1 in this dissertation, and more details about the construction of the quality data are included in the appendix (section 8.2).

Figure 2.1 plots country-year observations of education quantity, measured as average years of education (from Barro-Lee 2013), against education quality, measured using the composite measure of student achievement tests across both primary and secondary education Angrist (2013). Not surprisingly, the two are clearly related (correlated by

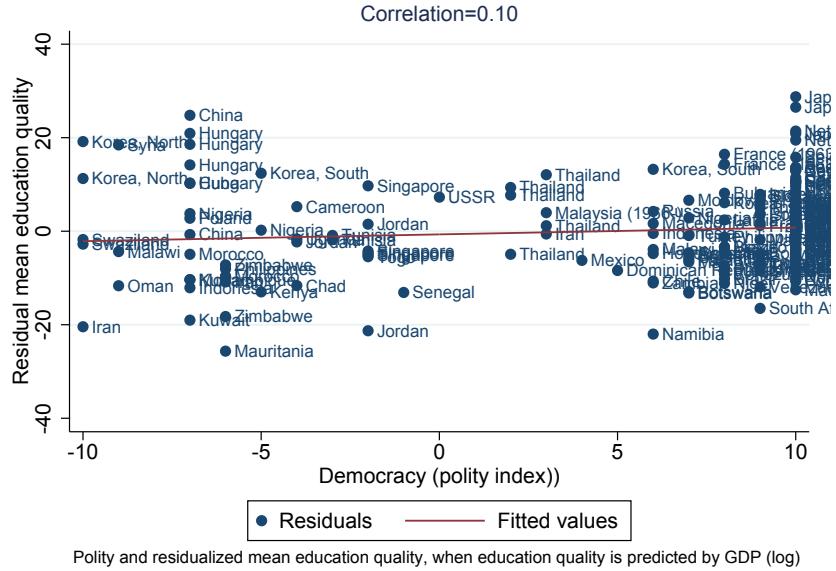
Figure 2.1: Schooling achievements (from Angrist (2013)) and mean years of education (from Barro-Lee (2013))



0.61), and countries with high education quantity are on average more likely to perform well in terms of student achievement tests. However, the country-years with the highest test scores are far from the best on education quantity. For instance, while China in 1990 scored among the top five percent in student achievements, its mean years of schooling was 5.6 years – well below average. Japan gains the highest student achievements score in this period, but does not even fall within the top 25 percent on quantity. Countries such as Namibia and Zimbabwe on the other hand, have a higher education quantity than e.g. China in 1990, but their education quality is dramatically lower.

Hence, education quality clearly diverges from education quantity, in ways that may matter for its relation to democracy. Figure 2.2 shows the bivariate distribution between democracy and schooling achievements when log GDP per capita levels are accounted for. The deviation from the predicted average student test score – based on OLS regression where log GDP is the only explanatory variable - is plotted along the Y-axis and the Polity index is plotted along the X-axis. The figure indicates no relationship between education quality and democracy (when accounting for income), contrary to what we should expect if the version of modernization theory discussed above is correct. This motivates a more systematic study in a multivariate set-up in section 2.5.

Figure 2.2: Schooling achievements (from Angrist (2013)) and the Polity Democracy index



2.4.2 Education autonomy and type

The perhaps most intuitive way of measuring *education autonomy*, especially in terms of education curriculums, would be an indicator of violations of academic freedom, i.e. to what extent governments dictate the curriculum and operation of schools. However, such an indicator could potentially be problematic in models with democracy on the left-hand side of the equation, as academic freedom is likely strongly related to levels of democracy. Indeed, it is even one sub-component in the Freedom House democracy index.⁵ There are, however, ways of capturing degrees of education autonomy less connected to degrees of democracy, that mitigates problems emerging when including aspects related the same phenomena on both sides of the regression equation. I therefore rely on an available proxy variable, measuring how many students in a country that are educated abroad, from the UNESCO Institute for Statistics. This data also contains information on bilateral student flows at the tertiary level, allowing for distinguishing between education acquired abroad in democracies and in autocracies, based on the expectation that education acquired in democracies should be more likely to spur democratization. This is, of course, not a direct measure of the independence of domestic schools, but likely captures the extent to which a population is exposed to education independent of the home government, includ-

⁵These subcomponents are, in any case, not currently released.

ing education that is acquired in a democratic system.⁶ I operate with two indicators measuring the number of students in a given year studying abroad in either a democracy or an autocracy respectively, and these variables are lagged by five years to capture the number of graduates from foreign universities in year t.⁷

To measure variations in emphasis on different academic disciplines, I use UNESCO's indicators of share of graduates from different fields. I operate with four indicators measuring the from 1) the humanities and arts, 2) social sciences, business and law, 3) engineering, manufacturing and construction and 4) sciences.

2.4.3 Dealing with missing values

A problem for this study is the relatively short time-series with many gaps. For instance, the education quality data only comes in five-year panels and many countries have even larger gaps in the time-series. This may create selection bias, as the data available may be systematically better for highly educated countries with very high democracy scores. To avoid selection bias due to this “missingness”, I conduct multiple imputation, employing the Amelia II software developed by Honaker et al. (2011), which accounts for the time-series cross-section structure of the data. The imputation model produced 5 different data sets, that were subsequently used in the empirical analysis of the paper.⁸

2.4.4 Regime change and democratization

My main dependent variable is the **level of democracy** measured using the Polity index (Marshall, 2014), building on numerous previous studies (see e.g. Glaeser, Ponzetto and Shleifer, 2007; Acemoglu et al., 2005; Castelló-Climent, 2008; Bobba and Coviello, 2007; Murtin and Wacziarg, 2014).⁹ By including a lagged dependent variable, these models estimate the effect of the independent variables on *changes* in democracy levels from one year to another. Additionally, I present results from models distinguishing between democratic transitions and democratic breakdown, building on the by now conventional

⁶Indeed, Spilimbergo (2009) finds evidence that individuals promote democracy in their home country if their education is acquired abroad in democratic countries.

⁷A potential problem with this indicator is that many students may never return to their home-country. Yet, this indicator is likely (strongly) correlated with the level of foreign-educated in each country, as a certain share of students always return. Furthermore, it has even been argued that foreign-educated diaspora may influence political developments in their home country through e.g. money transfers and transmission of ideas to family and friends at home (Pérez-Armendáriz and Crow, 2010).

⁸I only carry out imputation for the 118 countries that are in the Angrist (2010) data on student achievement tests. The imputation model allows for country-specific, second-order polynomial time trends (Honaker et al., 2011, see). Furthermore, it specifies that indexes, fractions and other restricted variables have their theoretical minimum and maximum values as bounds. For more details see the appendix.

⁹As a robustness check I also estimate the main models with the Freedom House index.

insight that these two processes may have different explanations (Przeworski and Limongi, 1997). In these models I operate with a dichotomous measure of democracy as dependent variable, studying simultaneously the likelihood of a country transitioning from autocracy to democracy on the one hand and remaining democratic on the other hand with the help of Markov transition models. More specifically, I use the Boix, Miller and Rosato (2013) minimalist and dichotomous measure of democracy to model democratic transitions.

2.4.5 Controls

To separate the effect of education quality from the effect of education quantity, I include the latter as a control, measured using an indicator of the average years of education from the Barro-Lee (2010) dataset.

In the main models, I also add control variables that the literature points to as key determinants of democratization, that may also influence education quality and contents. This includes indicators of urbanization, GDP per capita (log-transformed), the gini index and oil dependency, all from the World Development Indicators dataset. In several model specifications I also include country fixed effects to account for country-omitted time-invariant variables and year fixed effects to account for time trends.

As robustness test, I also include additional control variables that could potentially be confounding factors, including an indicator of ethnic fractionalization from Alesina et al. (2003), an indicator of youth bulges from Urdal (2006) as well as unemployment, economic growth and indicators of share of muslims and protestants, the latter from the World Development Indicators dataset.

2.5 Empirical analysis

I here start by presenting results from models investigating how *education quality* affects democracy, including models accounting for the potential endogeneity of education quality. Second, I present results from models investigating if democracy is influenced by variations in *academic discipline* and *education autonomy*.

2.5.1 Education quality and democracy

This section investigates whether education quality impacts on democracy and democratization, and is a test of the main implication of the arguments discussed above (summarized in hypothesis 1). Table 2.1 summarizes the main results from regressions of democracy, operationalized using the Polity democracy index, on education quality, measured using

two indicators of international student test scores at the primary and secondary level.¹⁰ The multiple imputation procedure employed to deal with missing values produced 5 country-year datasets, where the imputed values vary across these datasets. All models presented in the following are estimated on the average-values across the five imputed datasets.¹¹ All independent variables are lagged by one year, but the results are also robust to alternative lag specifications, such as 5 and 10-year lags (see appendix for results).

Model A1 shows results from OLS regression (with panel-corrected standard errors (PCSE)) of democracy level on indicators of primary and secondary education quality, only controlling for log GDP and lagged democracy.¹² The coefficient estimate for primary education quality is low and insignificant at conventional levels with a t-value of 1.56. The coefficient estimate for secondary quality is negatively signed and even lower (close to zero), and far from significant with a t-value of -0.21.¹³ In model A2 the two indicators of education quality are replaced by an indicator of education *quantity* (average years of education), and this coefficient estimate is positive with a t-value of 1.70, in line with previous studies documenting this effect. In model A3, the two indicators of education quality are again included, and when controlling for education quantity, in addition to log GDP and lagged democracy, the two are still weak and statistically insignificant (and secondary quality has a negative sign).¹⁴

No effect of primary education quality is identified when adding country fixed effects to account for omitted country-specific factors (in model 4), as well as additional control

¹⁰The main findings are also robust to using the Freedom House index as dependent variable

¹¹The recommended procedure for imputed data is to use imputation-corrected standard errors to account for uncertainty associated with the imputation, and in the appendix I also present models based on this procedure and models estimated on the original (unimputed) dataset. However, due to the complexity of the former, it was not feasible combined with complex models such as fixed effects probit models, and the main analysis therefore relies on estimations on the average across the 5 data sets. This is arguably less problematic here, as my main finding supports the 0-hypothesis. If anything, this procedure reduces the threat of type 2 errors (and raises the likelihood of type 1), as the standard errors will not incorporate uncertainty from the imputation procedure, and may therefore be biased downwards. Hence, my main findings are based on estimations that make it *harder* to find 0-results.

¹²All models in the paper use an average indicator of primary schooling test scores and an average score for secondary schooling (across three different subjects). I also run separate models for test scores in maths, science and reading at both schooling levels, and the results are reported in the appendix. In sum, there is some heterogeneity between different sub-components. There is, for instance, some evidence that primary math skills and secondary reading skills are positively associated with democracy. However, these results are not robust to looking at democratic transitions, or to estimations using multiple imputation-corrected standard errors. I also run models with a combined average test score across primary and secondary level and all three subjects, revealing similar results.

¹³In the main models, the indicators of secondary and primary education quality are included together. However, in the appendix I also present models where only each of the two are included at the same time (see appendix, section 7.6.3).

¹⁴The Variance Inflation (VIF) test suggests that collinearity is not a major concern in these models, as the main independent variables both have quite low VIF values (2.09 for primary quality and 1.50 for secondary quality). Most control variables have VIF values lower than or slightly above 2.

variables, including urbanization, the gini coefficient and oil exports. When replacing the country fixed effects with year dummies (in model A5) to account for time trends, the primary quality coefficient increases slightly and now has a t-value of 1.70. This effect disappears, however, when adding country fixed effects in addition to year dummies (and controls) in model A6, suggesting that the (weak) positive effects from model A5 is likely only driven by omitted country-specific time-invariant factors. When it comes to secondary quality, no positive effect on democracy is identified in models with fixed effects or time dummies (models A4-A6). The coefficient estimate for secondary quality is actually *negative* with a t-value of -2.12 in the fixed effects model (A4), and negative with a t-value of -1.90 in model A6 with both country fixed effects and year dummies. If anything then, secondary education quality seems to be *negatively* associated with democracy in these models.

Finally, I run system gmm models to account for the possibility that the results in models A1-A6 reflect the slow-moving nature of education quality (and democracy), making it hard to identify effects using e.g. fixed effects models (Blundell and Bond, 2000). In brief, I find no evidence that primary or secondary education quality are related to democracy in these models (A7-A8).¹⁵ The t-values of the education quality indicators never exceed 0.41, and the secondary education quality coefficient is still negative in model A7, while primary quality is negative in model A8. Hence, there is overall little evidence that education quality boosts democracy, even when employing arguably more efficient estimators.

The coefficient estimate for *years of education* is low and insignificant in models A3-A5 and negative with a t-value above 2 in model A6. Yet, there is evidence that years of education is associated with democracy when using the system gmm estimator in models A7-A8. The coefficient estimate is very high with a t-value of 6.82 in model A7, and the coefficient estimate and t-value only drop slightly when adding year dummies in model A8. Hence, although no robust effect is identified across model specifications, there is slightly more evidence suggesting that mean years of education impacts on democracy than indicators of education quality.

In addition to the system gmm models, I run two additional tests to account for the fact that cross-sectional time-series models on annual data and with a lagged dependent variable is a strict test, that may produce type II errors when dealing with slowly-moving variables such as education quality and democracy. First, I estimate all models on 5-year rather than 1-year panels, and these yield similar results (see section 7.6.5 in the appendix). Second, I also estimate the models in Table 2.1 without controlling for lagged

¹⁵Ideally, country dummies should have been included in the system gmm model, but this combination failed to produce estimates.

Table 2.1: Democracy and education quality

	A1 OLS PCSE	A2 OLS PCSE	A3 OLS PCSE	A4 OLS FE	A5 OLS Year dummies	A6 OLS (FE) Year dummies	A7 System GMM	A8 System GMM
Education quality								
Primary quality	0.0613 (1.56)		0.0511 (1.29)	0.0250 (0.60)	0.0730+ (1.79)	0.0137 (0.33)	0.0187 (0.41)	-0.00678 (-0.15)
Secondary quality	-0.00738 (-0.21)		-0.0140 (-0.38)	-0.0946* (-2.12)	0.00359 (0.09)	-0.0845+ (-1.90)	-0.0152 (-0.34)	0.0168 (0.37)
Controls								
Democracy (lag)	0.945*** (207.32)	0.943*** (184.42)	0.943*** (184.77)	0.849*** (72.80)	0.933*** (152.61)	0.833*** (68.40)	0.805*** (98.76)	0.794*** (94.24)
GDP (log)	0.395** (2.63)	0.320+ (1.81)	0.294+ (1.66)	1.081*** (3.53)	0.517* (1.97)	-0.266 (-0.65)	1.146*** (3.82)	0.487 (1.43)
Years of schooling		0.156+ (1.70)	0.103 (0.97)	-0.0294 (-0.24)	0.118 (1.03)	-0.272* (-2.18)	0.936*** (6.82)	0.736*** (5.10)
Urbanization				0.0775* (2.14)	-0.0134 (-1.31)	0.0672+ (1.86)	0.0410+ (1.68)	0.0532* (2.18)
Gini index					0.0326 (1.02)	0.0566* (2.21)	0.0557+ (1.74)	0.102** (3.23)
Fuel					0.0102 (0.86)	-0.0189** (-3.08)	0.00910 (0.76)	-0.0803*** (-6.35)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	4876	4876	4876	4876	4876	4876	4873	4873
R ²	0.923	0.923	0.923	0.927	0.925	0.930		
chi2							16852.4	17272.4

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are clustered on country. Dependent variable is the Polity democracy score. Constant, country dummies and year dummies are omitted from the table.

democracy, and these tests yield mixed results (see appendix section 6.6). This is indeed a very liberal test, and it does not capture *changes in democracy* but rather variations in democracy levels. Yet, the fact that even these generous models do not identify a consistent effect of education quality on democracy, suggests that the null-finding from Table 7.28 is not simply a type II error driven by inefficient estimation.

By controlling for lagged democracy, all models above can be interpreted as *changes* in democracy from one year to another. These models do not, however, tell us whether education quality impacts on the likelihood of transitions to democracy or on the likelihood of democratic survival. In addition, I also run dynamic probit models separating between transitions from autocracy to democracy and the survival of existing democracies using the Boix, Miller and Rosato (2013) dichotomous democracy index. The results and details about the models are included in the appendix (section 7.2). In short, I find no evidence that education quality at the primary or secondary level has a positive impact on neither transitions to democracy nor democratic survival.

As discussed above, there are theoretical reasons to expect education quality to be less endogenous to regime type (and change) than education quantity, but we still can not rule out endogeneity bias. Both democracy levels and transitions to democracy may impact on levels of education quality, and there may be other unobservable characteristics influencing both education quality and democracy that are not included in the above

estimated models. The system GMM models in Table 2.1 account for endogeneity to some extent, but a better test requires an explicit identification strategy.

In response to this, I run two-stage-least-square (2SLS) instrumental variable models, treating education quality as endogenous, and using the levels of numeracy skills in the 1820s as instrument for current levels of education quality. It has been demonstrated that cognitive skills such as numeracy abilities are highly persistent over long periods of time, due to countries with early reading and numeracy abilities entering a path-dependent process whereby education increases the prevalence of human capital-intensive industries, that in turn help to foster future skills (Baten and Juif, 2014). Unique information on early numeracy skills is collected by Crayen and Baten (2010) (see also A'Hearn, Baten and Crayen (2009)) for 165 countries covering the period 1820 to 1945. To estimate numeracy skills Crayen et al use an age-heaping technique, utilizing irregularities in people's ability to report their age to estimate societies' level of education. The identifying assumption is that the instrument should only impact on current levels of democracy (from 1970-2010) through current levels of education quality - when controlling for alternative channels such as GDP per capita, education quantity, unemployment and oil dependency. This assumption is strengthened by the fact that the instrument directly taps into early skill levels rather than relying on proxies such as enrollment rates. The results from the instrumental variable regression is presented in the appendix (section 7.3). In brief, I find no evidence that neither primary or secondary education quality are positively related to democracy when using numeracy skills as instrument. Moreover, tests for weak instruments indicate that numeracy skills is a very strong instrument for education quality.¹⁶ Hence, these models should likely identify an effect of education quality if there is one.

Finally, in defense of modernization theory, it could be argued that the tests presented above are unfair since they do not acknowledge that the effect of education quality on democracy may be conditioned on schooling *quantity*. More specifically, high-quality education could be expected to induce democracy *if* combined with high enrollment rates, covering broader segments of the population. First, if high-quality education is only offered to a small section of society, the mechanisms discussed above may be prevented from materializing at the national level, as the number of people experiencing raised capacities and altered preferences may not reach the critical mass necessary to stage large-scale mobilization for democracy. Second, high-quality education may sometimes

¹⁶In the first-stage regressions, numeracy skills in the 1820s is a statistically significant determinant of both primary and secondary education quality (at the 0.001 level). The instrument is also strong, with F-tests ranging from 256.1 to 11.9. The Wu-Hausmann test rejects the null that education is exogenous, suggesting that instrumenting for education quality is sensible (see appendix, section 7.3, for results from first-stage regressions).

be used by autocrats for purposes of co-optation. The literature on authoritarian survival strategies indicates that autocrats will often focus on ensuring the loyalty and support of a small "ruling coalition" (e.g. Bueno de Mesquita and Smith, 2003; Svolik, 2012). Consistent with this, one would expect dictators (with a given amount of resources) to offer high-quality schooling to an exclusive segment of society, to ensure loyalty from the "ruling coalition", while restricting access to schooling for the remaining part of society. If this strategy works, high-quality education will not induce democracy when combined with low enrollment rates. To test this proposition, I also investigate a version of the baseline model including interaction terms between primary and secondary education quality and enrollment rates. Contrary to this line of reasoning, however, I find no evidence that education quality matters for democracy even when enrollment levels are high (see online appendix, section 7.4, for results).¹⁷ Hence, high-quality education does not promote democracy even if offered to a large share of the population, a finding casting further doubt on the proposed mechanisms tying education to democratization.

2.5.2 Education type and education autonomy

Even if high-quality education does not influence democratization, other aspects of education may still matter, as discussed above. In Table 2.2, education quality is replaced by indicators of education type, measured as the share of tertiary graduates from different academic disciplines (models B1-B5) and education autonomy (models B6-B10), operationalized using a proxy variable of the number of students educated abroad (in democracies and autocracies).

As discussed, countries with a large share of graduates from the humanities or social sciences should be expected to be more likely to experience democracy or democratic transitions than countries with a high share of graduates from engineering or sciences. In brief, models B1-B5 offer no evidence that emphasis on the humanities or social sciences is positively associated with democracy. Contrary to expectations, the coefficient estimates for humanities graduates are negative in these models, with t-values ranging from -0.49 to -4.41. The coefficient estimates for social sciences are also negative in four out of five models, but never reach conventional levels of significance.¹⁸ When it comes to the share

¹⁷In the appendix I also present results from models distinguishing between democratic transitions and survival, where the interaction term between education quality and enrollment levels is included (see appendix, section 7.4). These models investigate whether the effect of education quality on democracy or democratic transitions is also conditional on enrollment levels. I find little evidence that this is the case.

¹⁸Table 7.2 in the appendix presents results from dynamic probit models distinguishing between transitions to democracy and democratic survival. In brief, there is no evidence in these models that a large share of graduates from humanities or social sciences are related to transitions to democracy or

Table 2.2: Democracy, education type and education autonomy

	B1 OLS	B2 OLS	B3 OLS FE	B4 OLS FE	B5 System GMM	B6 OLS	B7 OLS	B8 OLS FE	B9 OLS FE	B10 System GMM
Education type										
Humanities	-0.0769+ (-1.71)	-0.0886+ (-1.93)	-0.0619 (-1.13)	-0.0275 (-0.49)	-0.194*** (-4.41)					
Social science	-0.0303 (-1.17)	-0.0278 (-1.08)	-0.00928 (-0.25)	-0.00340 (-0.09)	0.0402 (1.55)					
Science	0.00113 (0.02)	0.0137 (0.25)	0.0792 (1.06)	0.0549 (0.75)	0.0302 (0.50)					
Engineering	0.0209 (0.81)	0.0777 (0.30)	0.0484 (0.98)	0.0642 (1.19)	0.00709 (0.18)					
Education autonomy										
Students abroad(aut.)						-1.419 (-0.63)	-1.507 (-0.74)	2.768 (1.22)	2.366 (1.04)	-0.0748 (-0.02)
Students abroad(dem.)						-1.235 (-0.69)	-1.167 (-0.84)	-1.276 (-0.86)	-1.162 (-0.80)	-3.285+ (-1.87)
Control										
Democracy (lag)	0.965*** (131.57)	0.959*** (113.64)	0.677*** (12.00)	0.675*** (12.01)	0.823*** (68.55)	0.949*** (103.87)	0.947*** (152.47)	0.837*** (52.59)	0.814*** (47.18)	0.802*** (75.98)
GDP(log)	-0.171 (-0.69)	-0.601* (-2.01)	-0.315 (-0.77)	-0.441 (-1.08)	-1.001** (-2.78)	0.566* (2.38)	0.379 (1.55)	1.605** (2.70)	-0.654 (-0.78)	2.119*** (4.21)
Tertiary graduates	9.070** (2.79)	12.00** (3.27)	8.719* (2.99)	15.52*** (2.38)	5.789 (1.17)	-3.124 (-0.49)	-11.56+ (-1.76)	18.49* (2.51)		
Urbanization	-0.305* (-2.45)	-0.221 (-1.47)	0.0266 (1.10)	0.0250 (1.07)	0.0250 (-0.62)	0.116* (2.21)	0.0856 (1.58)	0.0731* (2.16)		
Gini index	0.0175 (0.45)	0.0250 (0.65)	-0.0197 (-0.62)	0.0471 (1.11)	0.0750+ (1.78)	0.121** (1.11)	0.0557 (3.05)			
Fuel	-0.0345 (-1.26)	-0.0302 (-1.01)	-0.0540*** (-4.11)	0.00557 (0.34)	0.0124 (0.77)	-0.103*** (-6.12)				
Country dummies?										
no	no	yes	yes	no	no	yes	yes	yes	no	
Year dummies?	1232	1232	1232	1317	2990	2990	2990	2990	no	3323
N	0.962	0.962	0.970	0.971	0.924	0.924	0.926	0.932		
R ²										

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year, except from indicators of students abroad, which are lagged by 5 years. Standard errors are clustered on country. Dependent variable is the Polity democracy score. Constant, country dummies and year dummies are omitted from the table.

of graduates from engineering and natural sciences, no effect on democracy is identified.¹⁹

Finally, there is no evidence that education autonomy, operationalized using a proxy indicator of students abroad in foreign democracies, is positively related to democracy. The two indicators of numbers of students studying abroad (in democracies and autocracies) are lagged by five years, in order to capture the number of citizens in year t with an education from a foreign democracy. The coefficient estimate for students studying in foreign democracies in $t-5$ is actually *negative* in all five models in Table 7.29, but the t -values never exceed 1 except from in model B10, where the t -value is -1.87.²⁰ The coefficient estimate for students studying in autocracies, on the other hand, is actually positive in five out of six models, but never significant at conventional levels.

Hence, this section suggests no systematic effects of variations in education type or a proxy of education autonomy, on levels of democracy. Combined with the previous section, these results indicate that the skills and knowledge gained from education do not matter for the likelihood of democratization. Nor does an autonomous curriculum or an emphasis on social sciences and humanities promote changes in democracy, contrary to expectations.²¹ Rather, there is some evidence that the relationship between education and democracy really does seem to be driven by average years of education, in line with previous studies. However, as this indicator is likely to be endogenous to democracy, this effect could very well be driven by other dynamics than a causal effect of schooling access on democratization.

The next section extends this discussion by investigating a potential explanation for this apparent null-finding.

2.5.3 Education quality and autocratic breakdown

No relationship was identified in section 2.5.1 between education quality and democracy or democratic transitions, even under high enrollment rates. To probe this finding further, I round off by investigating how education quality impacts on *authoritarian breakdown*, regardless of whether it leads to democracy or not. The downfall of dictatorship is a precondition for democratic regime change. But, the determinants of autocratic breakdown democratic survival.

¹⁹But, there is some evidence that engineering is *positively* related to democratic transitions. Moreover, science seems to be positively related to democratic transitions in some model specifications, although not in others (see appendix, Table 7.2).

²⁰Nor is there any evidence that a large amount of students having studied abroad in democracies (in $t-5$) is positively related to democratic transitions (see appendix, Table 7.2)

²¹It should be noted that it cannot be ruled out that the latter findings are affected by endogeneity bias, since education type and autonomy have not been instrumented for. Yet, the most plausible endogeneity scenarios suggest that democracy should produce more emphasis on humanities and social sciences and more students studying in foreign democracies, and would therefore increase the likelihood of type 1 errors.

– the *onset* of regime change – may differ from the determinants of the *outcome* of regime change (see e.g. Kennedy, 2010; Miller, 2012). When a dictator has stepped down, there is a point in time with great uncertainty regarding whether a democracy will emerge, or a new autocracy, as is very often the case. The factors that promote installation of democracy at this stage, may be different from the ones that initially brought down the dictatorship (Kennedy, 2010).

This suggests that studying the explanations of autocratic breakdown, in addition to studying democracy levels, could improve our understanding of why education quality does not induce democracy. A couple of studies suggest that indicators of modernization, such as income level, actually reduce the likelihood of authoritarian breakdown, as higher government revenues raises the dictator's ability to buy support (see e.g. Kennedy, 2010; Miller, 2012). Income may raise the likelihood of democracy being installed, once the dictator has stepped down, but the stabilizing effect of income on dictatorship cancels out the overall effect of income on democratic transitions (Kennedy, 2010). A similar dynamic may explain why no effect of education quality on democracy is identified: High-quality education may very well make citizens more eager for democracy, but this effect could be canceled out by a negative effect of education quality on authoritarian breakdown, due to the mechanism discussed above where recipients of high-quality education become more loyal to the dictator. On this account, education quality, by reducing the likelihood of authoritarian breakdown, prevents the onset of regime change from happening. This may erase any overall effect of education quality in models with democracy (or democratic transitions) as dependent variable. Pushing this story even further, we could expect education quality to be particularly stabilizing for authoritarian breakdown if coupled with low enrollment rates – as high-quality education targeted towards an exclusive group of supporters may be a particularly potent way of ensuring regime survival.

To probe this hypothesis, I operate with a dichotomous indicator of authoritarian breakdown from the Geddes, Wright and Frantz (2014) dataset as my dependent variable, capturing regime breakdown (regardless of outcome). The results are presented in Table 2.3. All models in Table 2.3 control for school enrollment, log GDP, and three variables accounting for regime age (regime duration, regime duration squared and regime duration cubed) (Carter and Signorino, 2010). Model C1 includes the primary quality indicator and the controls described above, to investigate whether primary quality has an unconditional effect on authoritarian breakdown, but no such effect is identified. When adding an interaction term between primary quality and primary enrollment (in models C2-4), the coefficient estimate for primary quality turns negative, but the z-scores are quite low (ranging from -1.05 to -1.44). Hence, we can not conclude that primary quality has a negative effect on the likelihood of autocratic breakdown when enrollment is low, either.

The coefficient estimates for the interaction terms between primary quality and enrollment are positive in all models, but never significant at conventional levels.

Table 2.3: Education quality and authoritarian breakdown.

	C1 logit	C2 logit	C3 logit FE	C4 logit FE	C5 logit	C6 logit	C7 logit FE	C8 logit FE
Primary quality	0.0168 (0.78)	-0.0926 (-1.05)	-0.126 (-1.34)	-0.206 (-1.44)				
Primary enrollment	-0.00114 (-0.19)	-0.0367 (-1.23)	-0.0909* (-2.28)	-0.0897 (-1.46)				
Primary quality*enrollment		0.00111 (1.21)	0.00151 (1.40)	0.00145 (0.88)				
Secondary quality					-0.0343 (-1.50)	-0.0692+ (-1.69)	-0.105* (-1.98)	-0.331*** (-4.35)
Secondary enrollment					0.00464 (0.84)	-0.0234 (-0.80)	0.0176 (2.66)	-0.133 (0.36)
Secondary quality*enrollment						0.000636 (0.98)	0.000759 (0.70)	0.00419* (1.98)
GDP (log)	0.0348 (0.23)	0.0431 (0.29)	0.399 (0.52)	-0.298 (-0.27)	0.0321 (0.19)	0.0467 (0.29)	-0.691 (-0.97)	-0.0698 (-0.07)
Regime duration	0.164* (2.07)	0.161* (2.03)	0.364* (2.31)	0.280 (1.47)	0.159* (2.04)	0.157* (2.02)	0.378* (2.50)	0.386 (1.63)
Regime duration ²	-0.0129* (-2.32)	-0.0127* (-2.27)	-0.0241* (-2.18)	-0.0135 (-0.84)	-0.0130* (-2.35)	-0.0129* (-2.34)	-0.0257* (-2.23)	-0.0235 (-1.29)
Regime duration ³	0.000235* (2.34)	0.000230* (2.28)	0.000891** (3.08)	0.000636 (1.29)	0.000238* (2.40)	0.000236* (2.39)	0.000882** (3.25)	0.000858 (1.63)
Country dummies?	no	no	yes	yes	no	no	yes	yes
Year dummies?	no	no	no	yes	no	no	no	yes
N	1895	1895	1415	933	1895	1895	1415	933
ll	-460.1	-459.3	-293.2	-193.7	-459.1	-458.5	-293.0	-185.5

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Geddes, Wright and Franz authoritarian regime breakdown indicator.

Constant, country dummies and year dummies are omitted from the table.

Looking at secondary education quality, however, there is more evidence in favor of the “elite co-optation” explanation. Model C5 suggests no unconditional effect of secondary quality on autocratic breakdown, as the coefficient estimate is negative, but low and with a z-score under 1.60. However, when also including the interaction term between secondary quality and enrollment, the negative coefficient for secondary quality, which can now be interpreted as the effect of education quality when enrollment is low, is strengthened and the z-score becomes -1.69 (in model C6). The secondary quality coefficient jumps in size and the z-score increases further when adding country fixed effects (in model C7). When also adding year dummies, the coefficient estimate is tripled in size and is now significant at the 0.001 level (in model C8). Hence, there is some evidence that secondary education quality reduces the likelihood of authoritarian breakdown when secondary schooling enrollment is low. In line with the argument discussed above, this could explain why no effect of education quality on democracy is identified in models that only look at the overall likelihood of democratization. If high-quality education is offered to only a few, this may stabilize authoritarian regimes, and enhanced ruling coalition

loyalty is a plausible explanation for this. Furthermore, if authoritarian regimes become more stable, the overall likelihood of democratization should decrease.

Meanwhile, the coefficient estimate for the interaction term between secondary quality and enrollment is positive in all models. Moreover, the z-score increases from lower than 1 to 1.98 when including year dummies and country fixed effects in model C8. Hence there is also some, although weak, evidence that secondary quality has a positive impact on authoritarian breakdown if enrollment is high.

This suggests that even if high-quality education does have an impact on democratization, for instance due to more democratic preferences and higher mobilization potential, this could be canceled out by the fact that dictators, by restricting the education to a selected few, may *also* reduce the overall likelihood of authoritarian regime breakdown. Hence, at best there seems to be no effect of education quality on democratization. At worst, a more complex picture emerges, where education quality targeted towards elites may also boost authoritarian regime stability.

2.6 Implications: Why quantity rather than quality matters

Above, I tested a number of empirical implications derived from two plausible causal mechanisms tying education to democratization. I identified no evidence that education quality matters directly for democratization. This suggests that it is not the skills and knowledge gained from education that matters for democratization, nor is it a free curriculum stimulating reflection or critical thinking. Rather, there is some evidence that the relationship between education and democracy really does seem to be driven by average years of education, in line with previous studies.

One potential explanation for this finding is that the relationship between education and democracy is driven by other mechanisms than those discussed in this paper (and those focused on in the literature). The shifts in people's behavior that could ultimately spur regime change may result not from the content of education, but rather from simply attending schools or universities. For instance, education may bring together people from different social segments in the same classrooms and institutions, and this may strengthen social networks. Moreover, education increases work differentiation and specialization, and this creates new patterns of interactions that partly replace family and kin as the dominant units of social interaction. The role of social networks in overcoming collective action problems has been discussed extensively in the literature (e.g Goldstone, 1994; Kurian, 1991), and education may potentially have an effect on democratic transitions

through this path.

Another interpretation is that we find no effect of education quality simply because education does not promote democratization. If the most plausible implications of a given theory are refuted by data, the most conventional response would be to question that theory. So why then do I find some evidence of a positive effect of education *quantity* on democracy levels? I suggested above that education quantity is much more likely to be affected by the introduction of democratic institutions, as schooling expansion is more manipulable by political regimes than schooling quality. Since I have not instrumented for education quantity in this paper, I can not exclude that the positive association between education quantity and democracy levels is driven by endogeneity bias. Indeed, the fact that the coefficient estimate for years of education usually drops in level of statistical significance (usually becoming insignificant) when estimating system gmm models may indicate that this is the case (see, e.g., Table 2.1). Hence, the identified effect of education quantity on democracy could simply be driven by reverse causation. In sum, I only find evidence that an arguable quite poor proxy of education matters for democratization, and this proxy is in addition likely to be very endogenous to regime type and change.

2.7 Conclusion

Modernization theory postulates that economically developed societies will experience democratization, and it is commonly argued that the trigger is a more educated population. This paper re-investigates the education-democracy link by accounting for variations in learning outcomes. I present detailed accounts of prominent mechanisms tying education to regime change, showing that they are really concerned with the effects of certain learning *outcomes*. Hence, I provide a more appropriate test of the education-democracy link than numerous recent studies who have rejected it, thereby offering modernization theory a “second chance”. Still, it does not hold up to empirical scrutiny. I find no robust evidence that high-quality, free education – expected to stimulate critical thinking and anti-authoritarian sentiments – is positively related to democracy. In other words, the hypothesized effect of education on democracy does not run through learning outcomes. Rather, there is some (although not robust) evidence that it is really driven by education quantity, which again is likely endogenous to democracy. Nor do I find that certain types of education such as social sciences and humanities are more conducive to democracy than other types. This finding contributes to the empirical assessment of modernization theory, by suggesting that some of its key implications receive very little empirical support. A citizenry having achieved free, high-quality schooling should not be more prone to overthrow dictators, suggesting that education may not, after all, be a panacea against

resilient authoritarian government.

3 Do democracies provide better education? Revisiting the democracy–human capital link

Sirianne Dahlum and Carl Henrik Knutsen¹

Abstract:

We investigate whether democracy enhances the skills and knowledge of citizens through improving education quality. This, in turn, could have ramifications for other development outcomes such as economic growth. We offer the first systematic cross-national study on democracy and education quality. Democracy is widely regarded as superior to autocracy in terms of providing access to education, and several studies find that democracy enhances educational enrollment and years of schooling. Yet, we do not know whether democracies provide *better education*. We argue that democracies should not too readily be expected to outperform autocracies on education quality. First, it is inherently difficult to implement quality-enhancing education reforms, even for well-intentioned (democratic *and* autocratic) governments with ample resources. Second, education quality is less visible to voters than, e.g., expanding education enrollment, making quality-enhancing policies a less attractive option for office-seeking democratic politicians. We employ a recent dataset comparing international student tests for 128 countries, from 1965 onwards. While democracies typically provide “more” education than autocracies, we find no systematic evidence that democracies offer better education. The result is very robust and holds in both cross-section and panel specifications. The null-relationship is not explained simply by democracies providing education access to more (and different types of) children than autocracies, and it appears both in rich and poor and in low- and high-capacity states. We also present relevant nuances: For instance, autocracies display more variation in

¹Professor at the Department of Political Science, University of Oslo

education quality outcomes than democracies, and we find *some* evidence that democracy may be associated, more specifically, with better reading skills. In sum, this study provides new insights to the democracy and education literature, where extant studies often report strong links between democracy and various education outcomes not directly related to education quality, and informs literatures linking democracy to development outcomes such as growth via effects on human capital.

Publication status:

Published in *World Development*

3.1 Introduction

We investigate the relationship between political institutions and education quality, which is closely linked to the knowledge and skills (i.e., “human capital”) obtained by citizens during their years in school. More specifically, we ask: Do democracies provide citizens with an education that raises their knowledge and skill levels relative to the education provided in autocracies? To foreshadow the conclusion, our empirical analysis suggests that democracy does *not* systematically relate to citizens’ knowledge and skills.

While education quality and the resulting human capital can be considered important development outcomes in their own right (e.g., Sen, 1999), any effect of democracy would arguably have broader developmental ramifications since human capital is anticipated to affect a range of social and economic phenomena. Human capital is, for example, typically considered among the key ingredients for generating economic growth (e.g., Mankiw, Romer and Weil, 1992; Lucas, 1988; Benos and Zottou, 2014). Early empirical analysis of human capital and growth often employed measures such as years of schooling or school enrollment ratios to proxy for human capital (e.g., Mankiw, Romer and Weil, 1992). Yet, economic growth *theory* is not merely concerned with “the quantity of schooling” offered to students, but rather the skills and capabilities that (prospective) workers acquire, in school or otherwise. Years of schooling or education spending are, at best, only distant proxies of human capital. Acknowledging this, the growth literature has recently turned its focus towards more valid proxies of human capital. More specifically, it has started taking education content into account, revealing – as theoretically expected – that education quality is a much stronger determinant of growth than indicators of “education quantity”. For instance, Hanushek et al. (2008) find that cognitive skills are positively related to growth, and that any positive effect of education quantity disappears when controlling for skills (see also Hanushek and Kimko, 2000; Hanushek and Woessmann, 2008).² Further, the human capital stock in developing countries has likely been exaggerated when using indicators such as enrollment rates and average years of education (see, e.g. Glewwe, Maiga and Zheng, 2014).³ Also differences in long-run growth performances across OECD countries can be explained by variations in education quality, as opposed to education quantity (Hanushek and Woessmann, 2011). Hence, understanding the determinants of education quality is important for understanding economic development.

Democratic political institutions have long figured among the proposed determinants

²Jamison, Jamison and Hanushek (2007) and Hanushek and Woessmann (2012a) find that education quality remains positively related to growth when using instrumental variables regression to account for endogeneity problems.

³For example, studies of Latin America have noted that the relatively low growth rates in the region cannot be explained by the high education levels, but when accounting for the poor quality of Latin American schools this “growth puzzle” is solved (Hanushek and Woessmann, 2012b).

of human capital. A vast literature, drawing on contemporary and historical data from different regions of the world, suggests that democracies are more likely than autocracies to provide *more* education to their citizens (Lake and Baum, 2001; Bueno de Mesquita and Smith, 2003; Lindert, 2005; Brown, 1999; Brown and Hunter, 2004; Stasavage, 2005; Engerman and Sokoloff, 2005; Engerman, Mariscal and Sokoloff, 2009; Ansell and Samuels, 2010; Gallego, 2010; Huber and Stephens, 2012; Acemoglu et al., 2015; Harding and Stasavage, 2014). Countries experiencing successful democratic transitions can thus, over the coming years, expect lower schooling fees, a larger share of their children enrolled into primary and secondary school, and that their youth will, on average, spend more years in school (but, see Murtin and Wacziarg, 2014). Indeed, different statistical studies have suggested that education is among the key mediators through which democracy enhances economic growth (Baum and Lake, 2003; Tavares and Wacziarg, 2001; Doucouliagos and Ulubasoglu, 2008). However, also these studies have drawn on measures of “education quantity”, such as secondary school enrollment rates, to proxy for human capital. In order to more directly evaluate the effect of democracy on human capital, and thereby also the potential indirect effect of democracy on growth via human capital, we need to study how regime type affects education quality.

One might intuitively expect the clear answer to the question of whether democracy promotes education quality to be “yes”. Why would voters not pressure politicians to give their kids a high-quality education, and not only increase the number of years they stay in school? In autocracies, truly competitive elections are lacking, leaving citizens with one less tool available for pushing for an education system that properly teaches their children skills such as reading and basic mathematics. Instead, autocratic regimes might expend effort and resources to provide high-quality education only for the children of their (often fairly narrow) supporting groups, or use the national education system to indoctrinate government ideology. A plausible hypothesis is thus that democracy enhances not only education quantity, but also education quality. Yet, we retain that this is an open empirical question: First, previous empirical studies suggest that education quality is intrinsically hard to increase by legislation or by simply increasing the education budget (see, e.g., Hanushek et al., 2008). If political choices have little bearing on education quality, the link between regime type and such quality should be weak. Second, democratic politicians may not have strong incentives to even *try* to increase education quality after all. Following Harding and Stasavage (2014), we highlight that the quality of education (and the policies that may improve it) is often difficult to verify and monitor by voters. Hence, in a world of budget constraints, democratic politicians may be wise to rather channel resources to measures increasing (more visible) features associated with “education quantity”.

With one notable recent exception (Harding and Stasavage, 2014), there is a lack of systematic empirical studies on regime types and education quality. The main reason is presumably that cross-country data on education quality have previously been scarce. We employ the recent dataset by Angrist (2013) containing comparable measures on international students achievements tests in mathematics, reading, and science from 128 countries. We thus provide the first systematic analysis drawing on cross-country data to assess the relationship between democracy and education quality.

To quickly summarize our results, we first replicate the finding that democracy is positively related to standard measures of education quantity, namely average years of schooling and primary and secondary enrollment ratios. But, the core finding – which is a novel, important, and perhaps (to many) surprising null finding – is that democracy does *not* systematically relate to education quality. While we acknowledge the possibility that this null-result could partly stem from measurement errors and other aspects with the data, and should thus not be viewed as entirely conclusive, the result does hold across different cross-section and panel specifications, different samples, and for different measures. One might hypothesize that the null-relationship between democracy and nation-wide measures of education quality is simply due to expansion of enrollment rates under democracy; more students entering school puts pressure on teaching facilities and new groups with worse initial prospects for learning (such as kids living in poor rural areas) are channeled into the schooling system. However, we find no clear support for this alternative explanation. Democracy is not associated with education quality, on average, even when accounting for the expanded access to education. Neither do we, when looking beyond the average performance of countries, find that democracies have smaller differences between good- and bad-performing students, though this result is based on far more limited data material. That being said, we do find that autocracies clearly display more cross-national variability in terms of aggregate education quality outcomes. While democracy does not relate to better education quality on average, democratic countries are less prone to observing extremely poor (and extremely good) outcomes.

In sum, democracy may increase the number of kids in school, but it does not, in general, clearly improve the measurable capacities, skills and knowledge of its young citizens. While democracy may increase economic growth through other channels, our findings thereby also cast doubt on the anticipated indirect effect of democracy on growth via improving human capital, since the first link in this chain is not robust.

The paper proceeds as follows: In Section 2, we review relevant literature on democracy and education. Thereafter, in Section 3, we elaborate on the argument for why democracy might not lead to higher-quality education outcomes. In Section 4, we present the data, focusing on our preferred measure of education quality. We present the empir-

ical analysis in Section 5, before concluding and pointing to avenues for future research in Section 6.

3.2 Democracy, education, and education quality

The literature suggests very plausible reasons for why democracy improves outcomes associated with “education quantity”. In brief, democratic politicians accountable to a wide constituency – many of whom have kids in schooling age – are incentivized to respond to calls for lower school fees, or expand education access to large population groups in order to be re-elected. Extant empirical studies use education measures that reflect these arguments such as enrollment rates, share of the population having completed primary, secondary and tertiary education, or the abolition of school fees. We will not survey this vast literature in its entirety, but mention some prominent arguments and results:

Lake and Baum (2001) describe how political competition within democracies generates political dynamics that increase the provision of public services at the expense of rents extracted by leaders. In contrast, autocratic politicians can extract substantial rents by limiting such services without facing grave consequences. A slightly different argument for why democracy widens access to and increases funding of education is provided by Bueno de Mesquita and Smith (2003). They highlight that broad-based education systems can be characterized as public goods, and that investing in such systems is a cost-effective way to maintain political support for leaders with large “winning coalitions” (as in democracies), and relatively expensive for leaders with small coalitions (as in many autocracies). Acemoglu and Robinson (2006) theorize that the relatively poor majority will push for universal schooling under democracy, where the poor hold more political power, whereas autocratic elites will shun such expensive schooling systems, which also benefit the non-elites (see also Meltzer and Richards, 1981; Saint-Paul and Verdier, 1993; Boix, 2003; Ansell, 2008; North, Wallis and Weingast, 2009). Finally, Stasavage (2005), focusing on the African context, develops a model explaining why democratically elected leaders will invest more in primary education. Populous rural groups in African countries have clear preferences for spending limited public resources on primary education rather than policies benefiting urban elites, such as expensive university systems, and these rural groups face far lower costs of organizing collective action under democracy than autocracy (see also Bates, 1981).⁴

⁴Nuancing this picture, Kosack (2014) argues that autocratic leaders may also face strong incentives to expand mass education when they rely on mobilizing the disorganized poor to stay in power. Moreover, Kumah and Brazys (2016) argues that accountability (which can also exist in autocracies), and not democracy, is the relevant feature for promoting investments such as education spending in Africa. But, there are also other arguments suggesting that democratic politics could increase education spend-

Concerning empirical analysis, Lake and Baum (2001) report a strong positive cross-country correlation between democracy and various proxies of human capital, including measures of persistence of students to the fourth grade, different school enrollment ratios, pupil-teacher ratio, and literacy rates (the latter two can, admittedly, be considered as crude proxies of education quality). But also several other studies – including case studies, small-n comparative studies and statistical studies employing time series variation – suggest that democracy expands access to and improves funding of both primary and secondary schooling (Lindert, 2005; Stasavage, 2005; Engerman, Mariscal and Sokoloff, 2009; Huber and Stephens, 2012). For instance, the historical expansion of access to lower-level education in the US, UK and Scandinavian countries seemingly followed from the expansion of political rights, and suffrage extensions in particular, during the 19th and early 20th century (see Lindert, 2005). Studying post-colonial Africa, Stasavage (2005) shows that democracy relates to spending on primary education in particular, and Harding and Stasavage (2014) detail how elections lead to subsequent abolition of school fees. They also report evidence from surveys of Kenyan voters suggesting that they condition their voting behavior on school fees being abolished, providing more direct evidence for the type of theoretical argument reviewed above.

One concern is that the above findings reflect a causal effect running from education to democracy – such an effect is among the central propositions of “modernization theory” (e.g., Lipset, 1959), and the empirical study by Murtin and Wacziarg (2014) suggests that the correlation mainly reflects a “reverse effect”. Nonetheless, other stringent studies (e.g., Acemoglu et al., 2005) do not find that education systematically enhances democracy, casting doubt on the reverse relationship, and yet other identify a causal effect of democracy on measures of education quantity even when accounting for the potential endogeneity of democracy (e.g., Gallego, 2010; Acemoglu et al., 2015). Hence, while contested, a substantial effect of democracy on, for instance, lower-level school enrollment or years of schooling seems more plausible than not.

3.3 Why democracy may not enhance education quality

Thus, democracy likely improves access to education. However, we can not automatically expect democracies to provide *better* schooling than autocracies. Following Harding (2015), we contend that vote-seeking democratic politicians are incentivized to provide

ing: Jennings (2015) argues that even voters without material incentives to further increase education spending could vote “expressively” for it (knowing they are unlikely to affect policy individually) due to ethical concerns.

better public policies only when it comes to *certain public services*. More specifically, responsibility for the quality of the particular public service should be easy to trace back to the elected officials and the service provision should be transparent, in the sense that voters can verify whether improvements have been made. Further, the public service in question must be linked to an available policy that is within the reach and capacity of elected official to effectuate. We discuss below how education quality does not have these characteristics, making it less feasible for politicians to pursue as a way of boosting popularity.

There are two parts to our argument. First, education quality is hard to verify and is often not attributed to the central government. This reduces the incentives of vote-seeking politicians to pursue legislation or other measures anticipated to improve education quality. Second, democratically elected politicians (as well as autocratic) may simply be incapable of increasing education quality even if they desire to do so. There is fairly strong evidence that the quality of education systems is (highly) path dependent, and does not respond clearly, and especially not immediately, to pieces of legislation – which may also be hard to implement effectively in the education area – or to increased education spending. We elaborate below.

3.3.1 Democracy and incentives to improve education quality

First, there are reasons to expect that improvements in education quality may not be a very effective way of gathering public support, and this reduces the incentives of democratic incumbents (and opposition leaders) to pursue policies that may enhance education quality. Whereas a democratic system creates strong incentives for office-seeking politicians to expand education enrollment, it does not necessarily incentivize them to pursue quality-improving programs.

One reason is that improvements in education quality are not easily *verifiable* for the average voter (see also Harding and Stasavage, 2014), whereas, e.g., expanding access to schooling (or abolishing school fees) *is* fairly visible to voters: Experiencing that your child or grandchild finally gets to go to school will likely make a clear impression. In contrast, few citizens have exhaustive knowledge about the quality of an education system. Investments in measures that improve student achievements, such as developing new methods of class-room teaching, are often less transparent (Harding and Stasavage, 2014). Thus, they are difficult to promote and explain at campaign rallies or televised debates.⁵ The opaque relationship between specific policy measures and improvements in

⁵The public release of student achievement tests, such as PISA, have received attention in the public debate, especially in some OECD democracies. This could contribute to increasing transparency for attentive voters, for example by providing measuring sticks for cross-country comparisons. Yet, this is

education quality means that incumbents are less likely to convince retrospective voters, even those concerned with education outcomes, in the next election.

Another reason why democratic incumbents may not be punished for neglecting education quality is that it is often not clearly *attributable* to the central government (see Harding, 2015). While policy reforms such as abolishing school fees, implementing mandatory schooling laws, or removing formal barriers to entering schools can be clearly traced back to government policies, poor education quality can often not. Low education quality can be caused by lack of school equipment or poor facilities, and these factors are commonly managed at the local level. Meanwhile, poor teaching is often attributed to characteristics of the teacher or the school, rather than the central government.

If education quality does not attract much popular support, democratically elected leaders may even accept a depreciation quality in order to, for instance, raise enrollment rates (Michaelowa, 2001). This is particularly pertinent for governments presiding over limited financial resources. Addressing the Ugandan abolition of primary school fees in 1996, Stasavage (2005) argues that it was most likely the introduction of democratic politics that prompted President Museveni to promise and implement this reform. Stasavage points out that while the Ugandan education reform secured universal primary education, the massive expansion in education enrollment was followed by a drop in education quality – possibly due to subsequent shortage of basic materials and a dramatic increase in pupil-teacher ratios. Likewise, Colclough and Al-Samarrai (2000) suggest that increased enrollment ratios in Sub-Saharan Africa and South Asia has often been achieved by lowering the “unit costs” of schooling. Studying education policy in current England, Jennings (2015) argues that also other education policies (such as banning private schools) may be supported by many voters on ethical (and sometimes instrumental) grounds, leading to “expressive voting” for such policies even if they induce sub-optimal education quality outcomes. Hence, there may even be negative effects of democracy on education quality due to the types of education policies that are electorally attractive. For a democratic government, expanding education enrollment is a powerful tool for satisfying constituencies, whereas improving education quality is less visible to voters. The latter is thus less likely to be prioritized by elected officials, particularly if budget constraints force a trade-off between quality-enhancing measures and expanding education access.

Survey evidence from Kenya illustrates that education quality does not strongly affect popular support, as intentions to vote for incumbent President Kibaki was not affected by poor and even deteriorating quality (Harding and Stasavage, 2014). Meanwhile, these authors also find that a reform abolishing school fees (and raising enrollment) clearly did increase support for the Kenyan president. Further, studying Ghana, Harding (2015) finds

mainly a recent phenomena (occurring towards the end of our time-series).

evidence supporting the “attribution argument” discussed above, showing that services that are easily attributable to the government, including roads and number of primary schools, increase incumbent vote share. In contrast, factors such as classrooms per school and school facilities – which are provided at the district level – does not affect voting behavior. These results constitute fairly direct evidence supporting of the first (electoral incentive) mechanism in our argument.

While democratic governments are not necessarily incentivized by electoral competition to improve education quality, they may still want to improve it for other reasons such as growing the economy. However, dictators often have similar such incentives. In dictatorships where incumbents rely on patronage and economic co-optation of their “ruling coalitions”, having extra financial resources serves as a buffer against political instability (e.g., Geddes, 1999; Bueno de Mesquita and Smith, 2009). Dictators may also invest in human capital, and increase production, in order to strengthen the country’s capacity to withstand external security threats (e.g., Thyne, 2006; Knutsen, 2011c). Hence, dictators and democratic leaders may face resembling incentives for promoting education quality.

3.3.2 Democracy and obstacles to improving education quality

Second, politically engineering improvements in education quality is inherently difficult. Even when democratically elected politicians are genuinely motivated to bolster the quality of schooling, the intended outcome is not ensured. Indicatively, different studies find no straightforward relationship between education spending and educational outputs (e.g., Hanushek and Kimko, 2000). One reason may simply be that it is hard to determine what exact policies are effective at enhancing student learning, as is reflected in the field of education research. It is, for example, still unclear which types of (and even whether) homework boosts learning (Cooper, Civey Robinson and A. Patall, 2006), what manipulable factors determine teacher quality (Harris and Sass, 2011), and whether number of students in the classroom affects learning (Hoxby, 2000). Given this lack of established knowledge about ways to improve education quality, there is little reason to expect that democratic leaders have profound insights into effective learning enhancers. Even some of today’s rich, advanced democracies such as Norway and the United States, where politicians at the very least pay lip-service to raising education quality, struggle to improve it.

Such uncertainty not only makes it difficult to identify exactly what a quality-enhancing education reform looks like in practice, but it also provides fodder for those opposed to resource-demanding education reforms. More generally, even if democratically elected politicians happen to know the “cure” against poor student achievements, they often face well-organized interest groups that make legislating and effectively implementing

education reforms difficult. Organized special-interest groups may, for various reasons, be particularly strong in established democracies (Olson, 1982), and markedly influence policy-making in “their” respective areas. One interest group that has received much attention in the literature is organized teachers’ unions. Although teachers’ unions should prefer more investments in the education sector, their members will often bear different costs associated with reforms, inducing unions to work against them (Loveless, 2011; Moe, 2011). For example, there is evidence documenting the efficiency of policies that link salaries to teachers’ performance, but such reforms are commonly impeded by organized interests. Teachers’ unions in the U.S. have in practice blocked any merit pay reforms (Hanushek, 2016). Meanwhile, the strong Mexican teachers’ union has worked to promote merit pay, but has exploited it to increase salaries while working to minimize requirements for moving up the payment ladder, thereby preventing the reform from boosting quality (Hancock, 2014). Hence, *even if* democratically elected governments were blessed with knowledge about the policies that boost education quality *and* would want to invest in and prioritize them, these policies may be difficult to implement.⁶

In sum, the uncertainty concerning what quality-enhancing reforms actually look like, and the difficulty of implementing education reforms, may help explain why previous studies have shown that a society’s overall level of skills and knowledge is strongly path-dependent (see, e.g., Crayen and Baten, 2010; Baten and Juif, 2014). Enhancing education quality is difficult to achieve by the stroke of a pen, even if the pen is held by an education minister, prime minister or president. This suggests that politics more generally, and the political regime type more specifically, may not strongly affect education quality.

Now, presiding over a well-functioning, non-corrupt state with high capacity might help politicians overcome opposing interest groups in passing different reforms, including education reforms. A high-capacity bureaucracy *might* also enable politicians and bureaucrats to overcome the difficult issue of discerning and evaluating information on which reforms actually work. As such, institutional features that enhance the “quality of government” (Rothstein and Teorell, 2008) could be important for achieving the desired policy outcome (Rothstein, 2011). Insofar as democracy enhances “quality of government” (see Charron and Lapuente, 2010), it could therefore indirectly enhance education quality. However, the link between such institutional features and education quality may be weaker than the link with other outcomes, simply given the extent of the informational

⁶The impact of reforms on learning may also take time to materialize. While the time lag likely differs with the type of reform and context, Fullan proposes that “it takes about three years to achieve successful change in student performance in an elementary school. Depending on size, it takes about six years to do so in a secondary school” (Fullan, 2000, 1). Below, we experiment with different lag lengths when estimating effects on education quality, and our baselines employ 5-year lags on the independent variables.

problems on what constitutes a quality-enhancing education reform. And, as discussed, having the capacity to implement the appropriate education reforms does not suffice; politicians need to be motivated to pass such reforms in the first place, and democratic accountability mechanisms may not work particularly well in this issue area.

3.3.3 Some suggestive evidence and observations

The relationship between democracy and education quality has previously not been systematically evaluated using cross-country data. But, the proposition that democracy does not, in general, enhance quality seems to be supported by some suggestive evidence. For instance, PISA test score studies reveal large variations in student achievements – which, as we argue below, is a strong indicator of education quality – among OECD democracies (see Fuchs and Woessmann, 2008). Further, there seems to be large variation in education quality also across autocratic regimes. In some autocracies the education system focuses on indoctrinating government ideology, and less on basic writing and math skills or knowledge of science. One example is North Korea, where students spend more than one-third of their time on “political education” in the *Juche* ideology (Martin, 2004, 167). Other autocracies have historically had high-quality education systems. 19th century Prussia had a high-quality system that was widely emulated by other countries (see, e.g., Lindert, 2005; Clarke, 2006). The quality of the education in mathematics and the natural sciences in the Soviet Union and other Eastern European Communist regimes was considered high (Boesman, 1993; Balzer, 1993; Chengze, Overland and Spagat, 1999). South Korea and Taiwan are two more recent examples, with students achieving top scores on cross-national skills tests before democratization (Stiglitz, 1997, 884), but also after. We proceed by investigating the relationship between regime type and education quality more systematically.

3.4 Data, measures and validity issues

3.4.1 Measuring education quality

We assume that education quality is strongly related to the actual skills and knowledge (i.e., human capital) that students gain from schooling. Indeed, our interest in education quality is, in large part, driven by a concern with the knowledge and skills that students actually acquire through their education. The latter is the outcome often hypothesized to, in turn, produce various other beneficial outcomes such as economic growth.

Thus, we draw on measures of students’ knowledge and skills to operationalize education quality, and more specifically a measure using international student achievement

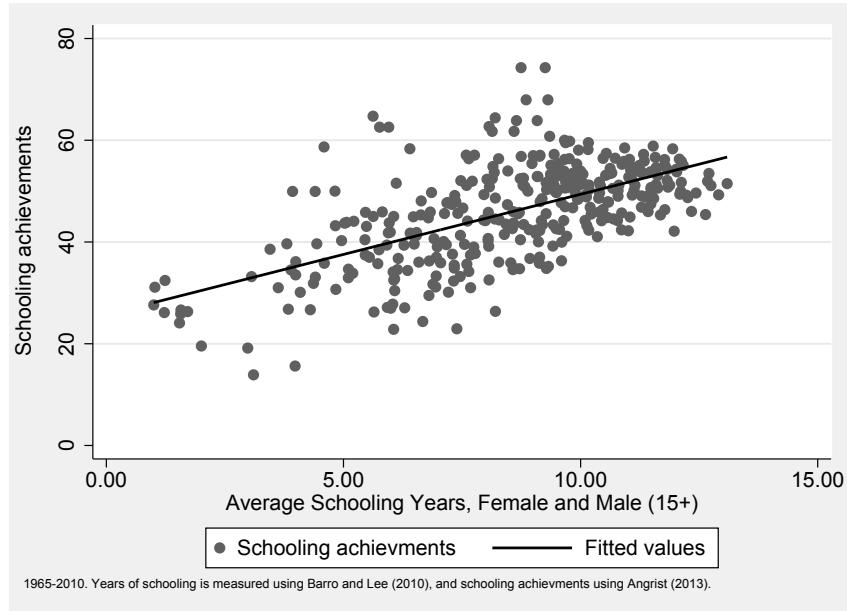
test scores. Although there are several challenges related to measuring education quality (and also skills and knowledge more directly), we submit that our preferred indicator comes close to capturing the outcome of interest, and far closer than more distal proxy indicators such as education spending or even the ratio of teachers to students. The latter measure *education inputs* rather than *education outputs*. While we acknowledge and discuss validity issues and limitations below, we first describe and highlight the beneficial aspects of the data, for example the coverage:

Angrist (2013) provides the first dataset on student achievement tests across a majority of countries globally, covering 128 countries around the world measured in 5-year intervals from 1965–2010. This is a clear extension of the cross-sectional coverage of previous datasets (up to about 50 countries), which directly resulted from the limited coverage of any single international achievement test such as PISA. Building on Hanushek and Woessmann (2012a), Angrist (2013) solves this by linking regional tests to international tests using countries that participated in both as reference points. Further, different international assessment tests are linked by using the United States – which has participated in almost all tests since their inception – as a reference point. Thus, regional assessments such as “South and Eastern African consortium for Monitoring Educational Quality” (SACMEQ) and the “Latin American UNESCO Laboratorio Latinoamericano para la Evaluacion de la Calidad de la Educacion” (LLCE) are converted to the same metric as international assessments such as PISA and Trends in International Mathematics and Science Study (TIMSS). All surveys included are conducted among students enrolled in primary or secondary schools, and we further detail the different surveys included and how the Angrist et al. dataset and our baseline measure of education quality is constructed in Appendix Section 8.2.

Our main measure of education quality is the composite measure from Angrist et al. that aggregates both primary and secondary level test-score performances for mathematics, reading and science. The measure ranges from 0 to 100 theoretically, and from 13.9 (Mauritania-1995) to 74.3 (Japan-1980) empirically, with a mean of 45.8 and standard deviation of 9.7. The Angrist et al. dataset also contains disaggregated test score measures on the three different subjects for each of the two education levels. We employ these disaggregated measures in extensions of our main analysis.

Figure 3.1 shows that education quantity, measured as average years of education (for 15-year olds+; the pattern is very similar when using data on 25-year olds+) from the Barro and Lee (2013) dataset, is correlated with our preferred measure of education quality by .60. Hence, countries with higher education quantity are, on average, more likely to have higher quality, but many observations also deviate from this pattern. Japan is the country with the highest observed score on student achievements, but on average

Figure 3.1: Average years of schooling (from Barro and Lee (2010)), and “schooling achievements” (from Angrist (2013)).



years of schooling Japan does not even fall within the top 25 percentile in our sample. Countries such as Jordan and South Africa, on the other hand, perform better on quantity than, say, China in 1990 (only 5.6 years of average schooling), but their education quality is dramatically lower than China’s. Indeed, when accounting for income level – which presumably affects both the quality and quantity of education provided – the partial correlation between average years of schooling and mean student achievement test score is only .20.

Despite capturing central learning outcomes that should be strongly linked to education quality, and despite the thorough measurement modelling and extensive information used to generate the Angrist et al. measure, there are potential validity issues. *First*, even though different test scores in different subjects and at different levels of education are included, there may still be noise in the measure. Unsystematic measurement error in our dependent variable would not systematically bias regression results, but would increase standard errors. Still, given that the coefficient on democracy flips sign between models, and t-values are very low, it is implausible that this is driving our results; the magnitude of such errors would have to be (unrealistically) high to explain our results.

Second, and potentially more problematic, there could be systematic errors in our education quality measure. One well-known example is that China only conducts PISA tests in Shanghai (plus in Macau and Hong-Kong), and students in this relatively wealthy urban area presumably score better on such tests. China alone does not drive our results

– we re-ran our regression omitting China, and results are almost identical. But, if autocracies, in general, only allow areas with expected high-performing students to participate, or otherwise manipulate test scores more blatantly than democracies do, we will underestimate the relationship between democracy and education quality. Hence, we cannot completely exclude the possibility that our results are produced by several authoritarian regimes manipulating their test scores. However, such a measurement-induced bias must be very large in order to explain our very clear null-results. Further, if “manipulation”, such as testing only particular schools or areas, is fairly stable within countries – persisting also after democratization – then controlling for country-specific effects should mitigate biases. The country-specific effects should also mitigate biases potentially related to PISA and other tests capturing particular topics and skills (within mathematics, reading and science) that are emphasized more in some countries than in others due, e.g., to cultural or historical differences.

Third, our main measure is based on test results for mathematics, reading and science. If one operates with a fairly broad concept of education quality, and democracies tend to put more resources into, and enhance learning outcomes in, subjects such as social sciences and civics education, we could potentially underreport differences in education quality. This is a plausible hypothesis, which we unfortunately cannot test given the lack of measures with comprehensive coverage. If data become available, future research could test for regime differences in education outcomes outside the core areas of mathematics, reading and science. With these caveats in mind, we proceed to our empirical analysis.

3.4.2 Data and measures for democracy and control variables

We employ the Polity Index from the Polity IV Project (Marshall, Gurr and Jaggers, 2013) as our main measure of democracy, and report results for other democracy measures in the appendix.⁷ Alongside the Freedom House Index, Polity is the most widely used democracy measure in the literature addressing relationships between democracy and education variables. Polity ranges from -10 to +10 (most democratic), and its dimensions are competitiveness and openness of executive recruitment, constraints on the chief executive, and competitiveness and regulation of political participation.

Regarding control variables we opt for a parsimonious baseline model controlling only for income level (ln GDP per capita; mainly from the Maddison project, but we also test models using data from Gleditsch (2002)). Income level, which also correlates with other aspects such as class structure and urbanization that could affect education quality, is a

⁷The relationships between education (both quantity and quality) and democracy are stable for quite different democracy measures, including the minimalist measure from Cheibub, Gandhi and Vreeland (2010) and the maximalist from Freedom House.

particularly important control. Income is strongly correlated with democracy, and richer societies are presumably more likely to have the capacity to build and maintain high-quality education systems. Further, children of wealthier citizens – for various reasons such as better access to information technology (Jackson et al., 2006) and increased parental involvement in children’s education (Desimone, 1999) – systematically perform better in school.

In many of our panel data models, we control for country-fixed effects since different country-specific factors, e.g. related to the historical characteristics of the education system or particular cultural traits, may affect education quality. Likewise time-fixed effects account for possible systematic time trends. For example, education quality might generally increase over time as we learn more about the measures improving it. While we consider the possibility of country-specific factors affecting education quality as highly plausible, and therefore prefer models including country-fixed effects, we also report cross-section and panel models without them for two reasons: First, given the moderately long time series and the slow-moving nature of democracy and education quality, fixed effects models could yield large standard errors. Fixed effects results could also be downward biased, due to the Nickell (1981) effect. A reported null-relationship in models including country-fixed effects *could* thus lead to a “Type II” error.

Also other factors can potentially affect education and correlate with democracy. We therefore report models including measures of income inequality (market income Gini, from Solt, 2009), natural resource income (ln oil and gas income, from Ross, 2012), and country size (ln population, from World Bank, 2011) in addition to income and the country and year dummies. Models controlling also for the latter three factors should thus mitigate omitted variable bias; if democracy is related to education outcomes in such models this would be a fairly strong result. Yet, these models risk “over-controlling” by taking out relevant channels that transmit indirect effects of democracy on education. Features such as income inequality (Acemoglu et al., 2015) or even population size (Przeworski et al., 2000) may be affected by democracy, and controlling for such covariates could introduce post-treatment bias. This is why we test both extensive and parsimonious specifications.

3.5 Empirical analysis

We start out by replicating the finding that democracy is positively related to measures of “education quantity”, before moving on to test the relationship with education quality. In addition to our baseline analysis, we present models checking, for example, for non-linearities in the relationship between democracy and education quality and whether the relationship depends on countries’ level of income or quality of government. Finally, we

present extensions on how democracy relates to different types of variability in education performance – both between different countries and among citizens in a given country – and how democracy relates to education performance in particular subjects, such as mathematics or reading.

3.5.1 Democracy and measures of “education quantity”

Table 3.1 reports our core specifications on education quantity, and alternative specifications are reported in Appendix Section 8.5. In accordance with much of the literature we find positive, and mainly statistically significant, relationships between democracy and measures of education quantity. The dependent variable in Table 3.1 is average years of schooling for 25-year olds from the Barro and Lee dataset. Model A1 is a parsimonious OLS cross section regression, run on 120 countries. The dependent variable (averaged over 2005–2009) is regressed on the Polity Index and ln GDP p.c. (both averaged over 2000–2004). The coefficient for Polity is positive with a t-value of 2.37, suggesting that an increase from minimum to maximum Polity score is associated with 1.3 more average years of schooling. The result holds up well when adding other plausible controls. Model A2 adds the Gini measure of market income inequality, ln oil and gas income, and ln population, and shows very similar results as A1.

Model A3 is run on the entire 5-year period panel dataset, where the first time period is 1965–69 and the last is 2005–09, to incorporate within-country variation over time. A3 controls for income and country-fixed effects. Both the Polity coefficient and its t-value increase relative to the cross section models, suggesting that the democracy–schooling relationship is *not* driven by country-specific omitted variables related to, e.g., culture, geography or political history. The coefficient is weakened, but remains positive and with a p-value of 0.11, when adding time-period dummies, oil and gas income, income inequality and population controls to the fixed effects model (A4).

Models trying to account for democracy being endogenous to years of schooling also show a positive, and often statistically significant, effect of democracy. For example, Models A5 and A6 are fixed effects 2SLS models that instrument for democracy by using WAVE from Knutsen (2011a). This variable records whether or not the last regime change (as identified by Polity’s duration variable) was within or outside one of Huntington’s “reverse waves of democratization” (Huntington, 1991) – time periods that were less and more conducive to democratization worldwide. WAVE has previously been used to estimate the effects of democracy on, e.g., property rights protection (Knutsen, 2011a) and civil conflict (Hegre and Nygård, 2014). The logic behind this instrument – which resembles the international “constitutional fashions” logic behind the instruments for electoral system and form of government created by Persson and Tabellini (2003) – is that WAVE

taps exogenous, international-political sources of variation in countries' democracy levels. As Huntington (1991) details, the international climate for democratization has varied over time due to differing preferences by regional and global hegemonic powers (see also Boix, 2011) and to various learning- or other spill-over effects (see also Weyland, 2005) from neighboring countries undergoing democratization or de-democratization processes. These international trends should provide us with a source of variation in democracy that is presumably exogenous to the domestic education system. Concerning the exclusion restriction, we cannot think of any obvious theoretical reason for why the international-political environment under the last regime change should directly relate to the domestic education system, conditional on the other covariates.

WAVE turns out to be a strong instrument, as absolute t-values in the first stage are 9.9 and 7.2 for Models A5 and A6, respectively, and weak-instrument tests show F-statistics far beyond conventional critical values. Despite the theoretical rationale, one can, of course, never *guarantee* that the exclusion restriction is satisfied, although empirical tests sooth such concerns.⁸ In any case, the parsimonious Model A5 reports a very large and statistically significant ($t=10.6$) Polity coefficient, whereas the extensive Model A6 reports a smaller, positive estimate with a p-value of 0.11.

Table 3.1: Democracy and average years of schooling

	A1 OLS Cross-sec. b/t	A2 OLS Cross-sec. b/t	A3 OLS (FE) 5-yr panel b/t	A4 OLS (FE) 5-yr panel b/t	A5 FE 2SLS 5-yr panel b/t	A6* FE 2SLS 5-yr panel b/t
Polity Index	0.065* (2.37)	0.076* (2.04)	0.118** (8.52)	0.014 (1.61)	0.240** (10.62)	0.023 (1.59)
Ln GDP p.c.	1.879** (12.80)	1.856** (9.69)	1.963** (6.51)	0.559** (2.72)	1.681** (9.73)	0.574** (4.64)
Ln oil+gas income p.c.	0.092 (1.21)			-0.087* (-2.28)		-0.088** (-3.80)
Gini (market income; reversed)	0.044 (1.58)			0.009 (1.42)		0.009* (2.09)
Ln population	-0.244+ (-1.75)			0.699* (2.00)		0.717** (3.75)
Period dummies				Y	Y	Y
Country dummies				Y	Y	Y
N	120	104	1027	707	1019	698

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is average years of schooling for 25 year olds. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for Polity in the FE 2SLS models (see Appendix Section 8.6 for first stage regressions). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

In Appendix Section 8.4 we report regressions testing alternative measures of edu-

⁸We cannot run standard overidentification tests with only one instrument. As an alternative (albeit an imperfect one, see Sovey and Green, 2011), we probed for a direct link between the instrument and schooling by including WAVE directly in fixed effects models otherwise similar to the second-stage regression. WAVE only obtains a t-value of 0.16 in the parsimonious model and -0.45 in the extensive model.

cation quantity (primary and secondary enrollment ratios) and democracy, alternative estimators, and alternative control sets. We consistently find a positive relationship between democracy and education quantity, and the relationship is most often statistically significant at conventional levels.⁹

3.5.2 Main analysis: Democracy and education quality

Whereas democracy is strongly related to measures of how many citizens obtain education, and of how long they stay in school, we find no systematic relationship between democracy and education quality. For starters, the raw correlation between Polity and the aggregated student achievement test score measure from Angrist (2013) for the 2005–2009 period is .34, but the partial correlation is only .14 when accounting for differences in ln GDP per capita.¹⁰ In contrast, *the partial correlation between the student test score measure and ln GDP per capita is close to .7*, suggesting that the measure does not only pick up noise.

Figure 3.2 shows the deviation from the predicted mean test score – where the prediction is based on an OLS regression with ln GDP p.c. as the independent variable – along the y-axis, and Polity along the x-axis, for 95 countries measured in 2005–2009. There is basically no net relationship between democracy and education quality, when accounting for income level. Strongly overachieving countries can be found both among autocracies (e.g., Cuba) and democracies (Finland), and the same goes for underachievers (Kuwait and Mauritius, respectively). This is an initial indication that democracy does not explain variation in the knowledge and skills of students.

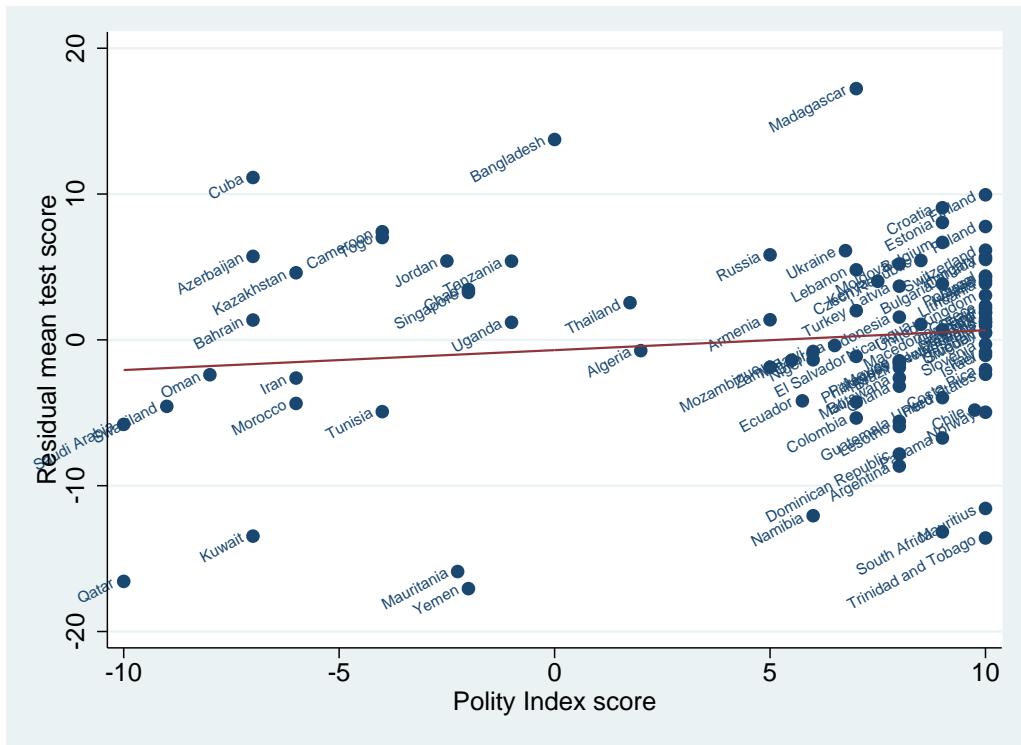
This null-relationship is corroborated in different regression models. Table 3.2 reports our core models, which are similar to those in Table 3.1 but substitute average years of schooling with the test score measure. The cross section OLS regressions in Models B1 and B2 corroborate the lack of any systematic cross-country correlation between democracy and education quality, with t-values of 1.1 and 0.8, respectively.

Given the lack of any relationship when drawing on cross-country comparisons, there is

⁹One notable issue is possible sample selection bias, as many countries – especially autocratic – do not report education data. To deal with this we employed a multiple imputation model using Amelia II software from Honaker et al. (2011) to predict missing values at the country-year level. The exact specifications of the imputation model and tests evaluating its performance (which is quite good) are described when discussing education quality regressions including imputed data and in Appendix Section 8.4. We constructed 5 imputed datasets and ran our regression models on these datasets using country-year as unit and employing imputation-corrected standard errors. The positive relationships between democracy and measures of education quantity turn out even clearer when including imputed data.

¹⁰The partial correlation is identical (.14) when lagging Polity with 5-years, and even lower when rather employing Freedom House's Political Rights Index (.04) or Civil Liberties Index (.11) as democracy measures.

Figure 3.2: Democracy and residualized mean student achievement test scores.



Note: The figure plots the residual (i.e. predicted - observed value from a regression with Ln GDP per capita as independent variable) of the mean student achievement test score measure from Angrist (2013) (y-axis), and the Polity Index (x-axis)

Table 3.2: Democracy and mean student achievements test score

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t	B5 FE 2SLS 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
Polity Index	0.165 (1.08)	0.146 (0.79)	0.052 (0.22)	-0.029 (-0.19)	-0.186 (-0.67)	-0.017 (-0.07)
Ln GDP p.c.	5.765** (8.19)	6.776** (8.34)	0.620 (0.36)	1.915 (0.65)	1.360 (0.92)	1.924 (0.77)
Ln oil+gas p.c.		-0.420 (-1.40)		-0.322 (-0.67)		-0.323 (-0.64)
Gini (market income; reversed)		0.429** (5.88)		0.013 (0.15)		0.013 (0.16)
Ln population		1.007+ (1.76)		7.918+ (1.86)		7.879* (2.07)
Period dummies				Y		Y
Country dummies		Y	Y	Y	Y	Y
N	91	84	341	321	317	297

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for Polity in the FE 2SLS models (see Appendix Section 8.6 for first stage regressions). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

little reason to expect panel models controlling for country-fixed effects or accounting for the endogeneity of democracy to identify a relationship. If certain cultural factors (e.g., related to customs and norms highlighting individualism and strenuous work, Weber, 2002) or political-historical factors (e.g., related to colonial experiences and transmission of institutions enhancing political and economic development, Acemoglu et al., 2005) affect both educational outcomes and democracy, the extant literature suggests that these factors should affect them in the same direction. To take one example, Woodberry (2012) argues that the historic prevalence of Protestant missionaries, and the resulting conversionary Protestants, was a key factor behind reading skills and other education outcomes, historically, in Africa, Asia, Latin America and Oceania, and that such developments, in the end, helped induce democratization. If anything, cross-section results should thus be upwards biased, and accounting for country-fixed effects will alleviate common historical roots underlying both an effective education system and democracy.

Related, if democracy is endogenous to education quality, various contributions to modernization theory, for instance, indicate that high-quality education should enhance democracy (e.g., Lipset, 1959). Well-educated individuals are more politically active, also in autocracy-challenging protest movements (e.g., Glaeser, Ponzetto and Shleifer, 2007), and highly educated citizens are more capable in overcoming collective action problems and challenging autocratic regimes more effectively (e.g., Welzel, 2013). In contrast, we are unaware of plausible arguments suggesting that higher education quality should negatively affect democracy in a systematic manner.

Still, we run models accounting for country-fixed effects and the endogeneity of democ-

racy to mitigate the *possibility* that any unforeseen “downward biases” are masking a positive effect of democracy in the cross country regressions – these models should, in any case, yield more consistent estimates. For the 2SLS models, which should also mitigate any attenuation bias stemming from random measurement errors in Polity, we again draw on the WAVE instrument capturing exogenous, international impulses on democratization. WAVE is a strong instrument, with absolute t-values around 8 in the first-stage regressions and F-test values surpassing any conventional threshold (see Appendix Section 8.6). We see no clear theoretical reason to suspect that the exclusion restriction is violated; when conditioning on, e.g., income level and time period dummies, there is little reason to expect the international-political environment during the latest regime change to directly affect education quality except via affecting regime type domestically.¹¹ As expected, neither the fixed effects OLS (B3–B4) nor fixed effects 2SLS models (B5–B6) identify any systematic relationship, with the t-values of the Polity coefficients ranging from -0.7 to 0.2.

Our theoretical argument highlighted two potential reasons for why democracy may not enhance education quality, a hypothesis that finds clear support in the analysis above. The first relates to the lack of clear incentives for democratically elected politicians to pursue policies that may boost the quality of education, as politicians facing re-election in competitive races gain from rather focusing on more visible measures such as expanding the access to education. Second, even if democratic politicians may want to improve education quality, we highlighted the tenuous link between available and manipulable policy measures for national politicians and education quality. To further check the validity of our argument and gain insight into whether one or both of these mechanisms seem to explain the null-finding, we run additional tests using various indicators of education spending.

Regarding the first part of the argument – on the lack of incentives for democratic politicians to pursue policies that could improve education quality – we investigate whether democracy is associated with both total education spending and spending per student. Results are reported and more thoroughly discussed in Appendix Section 8.3. We highlight that the available data and simple cross-section designs that we must use should caution against drawing clear inferences. Still, these analyses provide some suggestive evidence supporting the electoral-incentive part of our argument. In brief, democracy is not correlated with higher levels of education spending per student, nor with total education spending. The latter result might be surprising since democracy enhances education

¹¹This is corroborated by further empirical tests, which do not show any link between WAVE and education quality once conditioning on democracy level and the control variables. These tests are discussed in Appendix Section 8.6, which also reports and discusses alternative 2SLS specifications employing other instruments than WAVE.

enrollment. Yet, higher enrollment rates may not necessarily increase total spending levels by much if unit costs can be lowered by, e.g., increasing the number of students per teacher. (We do, indeed, find indications that democracies have more students per teacher in primary schools). In line with this, we also conduct mediation analysis suggesting that the effect of democracy on education quantity does not run through increased education spending.

The second part of the argument – on the difficulty of achieving education quality through political engineering – also receives support from cross-sectional regressions. Specifically, education spending is *not* clearly associated with higher levels of education quality, suggesting that politicians can not easily enhance education quality simply by channeling more funds to educational purposes.

3.5.3 Robustness tests on democracy and education quality

Different specifications and estimators: The null-relationship between democracy and education quality is robust in the sense that no OLS or 2SLS model with country fixed effects that we have tested give a positive Polity coefficient that is statistically significant at conventional levels.¹² Still, the absence of any relationship in these models could be due to the relatively few observations and small number of time series units for each panel. The longest panel is only 9 time periods (each covering 5 years), and the average is 3.2. With so few time series units, fixed effects estimates are both inefficient and downward biased (Nickell, 1981). But, the null-results remain for less restrictive pooled OLS and Random Effects (RE) models. The null-result is also retained in System GMM models, which are designed for short panels and for estimating effects of slow-moving variables (see Blundell and Bond, 1998). In contrast with the null findings for democracy, several of the control variables yield substantively large coefficients, and are often statistically significant. For instance, pooled OLS and RE models suggest that richer, more egalitarian and more populous countries have better education quality. We also find that natural resource rich countries produce worse education results.

Non-linear relationship: We probed different specifications to investigate whether our core models fail to pick up a potential non-linear relationship between education quality and democracy, or whether particular types of autocracies systematically fare better or worse than democratic regimes (see Appendix Section 8.7). We included a

¹²See Appendix Sections 8.6–8.7. For example, we tested different error specifications than our baseline robust errors, e.g. clustering by country to account for panel-specific autocorrelation. In our baseline models the independent variables are lagged with one 5-year period, but results are robust to, e.g., lagging the independent variables by 10 years. We note that some FE 2SLS specifications report a significant negative Polity coefficient, although we discuss in Appendix Section 8.6 why we put less trust in these specifications.

squared polity-term to investigate the possibility that “semi-democratic” regimes fare worse in terms of education quality than both democracies and autocracies. However, only Fixed Effects OLS models suggest such a relationship, while other panel models and cross section regressions do not show signs of any non-linear relationship between democracy and education quality. We further tested specifications including dummy variables for both “full” and “partial” democracies (calculated from cut-offs on the Polity scale, drawing on Epstein et al., 2006), but find no systematic differences between the two regime categories, nor between these two regime categories and autocratic regimes. In sum, we do not find robust evidence of any systematic non-linearities in the relationship between democracy and education quality.

We also tested models including the autocratic regime dummies from Hadenius and Teorell (2007), and do find some evidence that autocratic monarchies perform systematically worse than democracies, and some indications that multi-party autocratic regimes perform better than democracies (but only when we omit the Polity variable from the regression). However, neither military regimes nor one-party autocracies differ systematically from democracies.

Context-dependent relationship: We further probed whether our null-results could mask that democracy matters for education quality in some contexts, but not in others. These results are also reported in Appendix Section 8.7. We checked for an interaction between income and democracy, as democracy might matter for education quality only in rich-country contexts. Since primary school enrollment is close to universal in many rich countries, it could be that the political debate on education turns more towards improving quality (whereas democratic accountability mechanisms center on education expansion in poor countries). However, we find no evidence of any interaction between democracy and income on education quality.¹³

We discussed how institutions pertaining to the “quality of government” might enhance education quality, and we find some evidence of this in our cross-country regressions when using the composite QoG index, taken from Teorell et al. (2011). This measure is based on indicators of bureaucratic quality, rule of law, and corruption from the ICRG dataset (Political Risk Group, 2010). However, this result does not hold up in panel regressions. What is more relevant here, adding this measure does *not* affect the null-result for democracy. We also tested whether democracy matters more for education quality in countries with high quality of government, but do not find any clear evidence of a systematic interaction here either.

Multiple imputation: We identify fairly clear (and sensible) results for several con-

¹³Likewise, we tested whether the relationship depends on whether societies are egalitarian or not, but fail to find any systematic interaction between Polity and the market income Gini coefficient.

trol variables, meaning that the relatively low number of observations is not an impenetrable barrier to identifying effects on education quality, if one exists. Still, the inefficiencies and potential selection biases resulting from the few data points might still account for the null-result on democracy. To alleviate this, we tested different models on the five imputed datasets constructed by using Amelia II (Honaker and King, 2010; Honaker et al., 2011). In brief, the imputation model takes into account the cross-section time-series data structure, using second-order polynomial country-specific time trends. Diagnostics and overimputation tests suggest that the imputation model performs quite well, although it predicts the most extreme education quality scores less well (see Appendix Section 8.4). In order to provide as strong a test as possible on our main hypothesis, the imputation model was run so that it predicted observations for every country-year.

When using the imputed datasets, some parsimonious pooled OLS models report a statistically significant Polity coefficient (see Appendix Section 8.8). However, this is likely due to omitted variable bias; when running Random Effects (RE) or Fixed Effects (FE) models, Polity turns statistically insignificant. Table 3.3 reports six RE and FE models that we tested on imputed data. These are parsimonious models with a 1-year lag on the independent variables. Coefficients are averaged over the 5 imputed datasets, and standard errors are imputation corrected. Models C1 (RE) and C2 (FE) includes income level, whereas C3 (RE) and C4 (FE) also include year dummies. The Polity t-values vary between -0.6 and 1.4 in these four models.¹⁴

Controlling for differences in education access: We highlighted in our theoretical discussion that the relationship between democracy and education quality might be attenuated by democracies expanding access of education to more citizens. This may generate capacity issues such as fewer teachers, textbooks, and square meters of class room per pupil. Still, our results on nation-wide averages of student achievement test scores may also be attenuated for a more subtle reason that should be considered more of a measurement issue: The background characteristics of the children that are typically given access to education under democracy but not autocracy (e.g., children of poor, rural voters; see Stasavage, 2005) *might* make them less likely to perform well in school and on student tests, *ceteris paribus*, than the children that typically have access both under autocracy and democracy (e.g., children of rich, urban elites). The lack of any positive relationship between democracy and aggregated education quality might therefore simply reflect expanded education access driving down democratic countries' *average* test scores.

This conjecture is admittedly difficult to test absent more nuanced measures or individual-

¹⁴As displayed in Appendix Section 8.8, the null-result is (mostly) retained when employing alternate democracy indices, lag structures, sets of controls, and when first-differencing the dependent variable to investigate how democracy affects changes in education quality.

Table 3.3: Democracy and mean student achievements test score: Imputed data

	C1 RE 1-yr panel b/t	C2 OLS (FE) 1-yr panel b/t	C3 RE 1-yr panel b/t	C4 OLS (FE) 1-yr panel b/t	C5 RE 1-yr panel b/t	C6 OLS (FE) 1-yr panel b/t
Polity index	0.094 (1.38)	0.024 (0.34)	0.083 (1.29)	-0.037 (-0.57)	0.073 (1.10)	-0.041 (-0.62)
Ln GDP per capita	5.074** (5.12)	3.414** (3.51)	4.701** (3.92)	1.478 (1.33)	2.824* (2.56)	0.786 (0.76)
Primary school enrollment ratio					-0.014 (-0.65)	-0.019 (-0.95)
Secondary school enrollment ratio					0.078** (5.08)	0.025+ (1.99)
Year dummies			Y	Y	Y	Y
Country dummies		Y		Y		Y
N (per imputed dataset)	4914	4914	4914	4914	4914	4914
Countries	117	117	117	117	117	117

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. Country-year is unit of analysis, and all independent variables are lagged by 1 year. The coefficients are averaged over 5 imputed datasets, and standard errors are robust and imputation-corrected. The time series run from 1970–2011 (balanced panel). The log-transformed GDP p.c. variable is constructed as $\ln(\text{GDP p.c.} + 18000)$, due to the multiple imputation being conducted without bounds and the resulting occurrence of negative GDP p.c. values in the imputed datasets (see Appendix Section 8.4 for details on the imputation model). Constant, country dummies and period dummies are omitted from the table.

level data, although we note that our results in the next subsection, conducted on PISA 2012 results, suggest that democracy is not associated with higher variability in test scores within student populations, further alleviating concerns that this alternative mechanism is driving our results. Another implication is that the relationship between democracy and education quality should become stronger positive once controlling for measures of education access. *This turns out not to be the case*. We tested several models incorporating controls for primary and secondary enrollment ratios, and there is no evidence that the Polity coefficient subsequently increases (or is statistically significant, for that matter). Models C5 (RE) and C6 (FE) in Table 3.3 represent two such models, where both primary and secondary enrollment ratios are added to C3 and C4, respectively. The Polity coefficient actually decreases slightly in both models, suggesting that the null-relationship between democracy and the nation-wide average on student test scores does not come from democracies expanding education access to more kids.¹⁵ In sum, our analysis leaves little ground for optimism on the part of democracy in improving education quality.

3.5.4 Extensions and some nuances

Education quality in different subjects: The results presented so far rely on the aggregate indicator of student achievements tests that are based on (up to) six subcomponent scores, namely math, reading, and science test scores at the primary level *and* at the secondary level. Although we find no effect of democracy on this aggregate measure,

¹⁵These results hold up when adding enrollment ratios to models not including imputed data.

democracies could still be systematically better at promoting *certain types* of education. For example, democracies may be better at promoting reading skills, but not science and math skills. Various East Asian and former Communist dictatorships in Eastern Europe have emphasized education in technical subjects, engineering, or natural sciences as opposed to social sciences and humanities (see, e.g., Hayhoe, 1995). One could hypothesize that a dictator looking to increase state revenues would find improving the natural sciences skills of citizens as one attractive option. Doing so increases the tax base by enhancing the productivity of the labor force, while such education may be less likely to serve the needs of oppositional forces wishing to build up protest movements or campaigns (see Stasavage, 2005).

To investigate this we re-ran selected models from Table 3.3, replacing our comprehensive education quality measure with the disaggregated measures available from Angrist (2013). We note, however, that we put less trust in these tests, as the disaggregated measures have lower coverage across countries and over time than the composite measure. This also means that we can only rely on the imputed datasets for conducting our tests. The results, reported and discussed more carefully in Appendix Section 8.10, are thus only suggestive. In brief, Random Effects models – which are more efficient than corresponding Fixed Effects models, but potentially biased – provide no evidence that democracy is positively related to math skills. There are indications that democracy is positively related to reading skills, both at the primary and secondary levels, whereas democracy also correlates with science achievements at the primary (but not the secondary) level. But, when including country dummies, democracy is not systematically related to knowledge and skills in any of the three subjects, neither at the primary nor at the secondary level. Hence, while there are some indications that democracy is good for stimulating reading skills, more specifically, the evidence is far from robust and disappears when we account for country-fixed effects.

Variation in national-level outcomes: Our second extension relates to the *variability* in education quality in democracies and in autocracies, between countries, within countries over time, and within the student population in a given country. Although there is no difference in “average quality” between democracies and autocracies, autocracies could produce more outcome variation – as they have been shown to do for many other outcomes (such as economic growth; e.g., Rodrik, 2008). The higher concentration of political power within autocracies could make them more susceptible to the priorities and preferences of a small ruling elite, some of which may be highly concerned with education for various reasons, and others not.

To test this, we first run cross-section and panel models on education quality and next check for systematic differences in variation by employing Goldfeld-Quandt tests on the

residuals, separating relatively democratic and relatively autocratic regimes using various operationalizations. We do, indeed, find that autocratic countries display more variability in education quality (see Appendix Section 8.11). This result holds also when looking at variation within countries over time. Thus, even if predicted (average) education quality is not clearly related to regime characteristics, autocracies display more variation around their predicted level. In other words, democracies are less prone than autocracies to observe extremely bad, but also extremely good, education quality outcomes, also when accounting for income level and country-fixed effects.

Variation within student populations: Finally, one could expect that either autocracies or democracies display more variation in achievement tests also *within their populations*. That is, regime type could matter for the *spread* in skills and performance between “good” and “bad” students, even if the mean national levels are similar. However, the direction of such a relationship is hard to anticipate. Autocratic regimes might, for example, heavily favor children from their narrower supporting coalitions, suggesting larger spread within student populations in autocracies. Yet, democratization expectedly extends enrollment to student groups (such as sons and daughter of poor farmers) that may not have all the pre-requisites for performing well in school, and these students are often not in school (and are therefore not tested) in autocracies.

While the Angrist et al. dataset only includes country-aggregated scores, we can conduct more limited tests on regime type and within-population variation using data from the 2012 PISA test covering 56 countries. PISA mostly covers democracies, but 10 countries score lower than 6 on the Polity scale (and 8 score below 0, including as different countries as China, Jordan, Kazakhstan and Qatar). The results from regressions on the within-population standard error in test performance – we separately tested this for mathematics, reading and science, at the primary *and* at the secondary level – show no signs of any systematic relationship, independent of whether we control for the mean test result in the population or not. Further analysis shows that regime type is not related to absolute gender differences in test scores either. Insofar as these cross-country tests on a modest number of countries are informative, there is no indication that democracy relates to the variability in performance within student populations.

3.6 Conclusion

In this paper, we have replicated the result that democracy relates positively to features of the education system such as the share of young citizens that attend school and the number of years they stay in school. This is congruent with extant findings from the literature, and with theoretical arguments highlighting how democratically elected politi-

cians have incentives to address aspects of the education systems that are fairly visible to voters. However, we argued that democracy should not readily be expected to enhance the quality of the education that young citizens are exposed to, and thus the skills and knowledge that they obtain. Altering education quality is no easy matter for politicians, and democratic politicians may not even have strong incentives to *try* to alter them in the first place. Employing a panel dataset of international student test achievements, we find no clear evidence of any relationship between democracy and education quality, on average, although we do find indications that democracies display less cross-country variability in education quality outcomes than autocracies. Future data collection for more countries and preferably also for other subjects than mathematics, reading and science, will allow for even more appropriate testing of this relationship. Yet, our best guess, at current, is that democracy does not enhance the skills and knowledge of students.

Our findings have wider ramifications for venerable literatures on other subjects. It, for instance, bears on the suggested indirect effect of democracy on economic growth via increasing human capital. While several studies have identified a clear and positive indirect effect using measures of what we have termed “education quantity” (Tavares and Wacziarg, 2001; Baum and Lake, 2003; Doucouliagos and Ulubasoglu, 2008), analysis employing more appropriate measures of the actual knowledge and skills of the future work force would likely not identify any indirect effect, due to the non-existence of the first link in the causal chain. To illustrate this, we present the results of simple mediation analyses – with democracy as independent variable, the student performance measure as mediator, and economic growth as dependent variable – in Appendix Section 8.9. The same might be the case for the proposed indirect effects of democracy via education on other outcomes of interest such as higher political participation, social trust and reduced political violence. This is because education is often theorized to stimulate political participation (Hillygus, 2005), as well as higher trust and fewer incidences of political violence (Thyne, 2006), mostly through enhancing knowledge and skills among citizens, and not simply through kids being in school.

Part II

**Disaggregating the relationship
between education and authoritarian
regime instability**

4 Schooling as a “dangerous good”? Education, mass protest and authoritarian regime stability

Sirianne Dahlum

Abstract:

The argument that dictatorships with educated populations are more likely to democratize is popular among political scientists, and numerous empirical studies investigate the relationship between education and democratic transitions. Yet, this paper argues that extant literature offers too highly aggregated tests to capture the processes through which education influences prospects of authoritarian regime instability. First, I argue and show that a high education level is actually associated with a lower likelihood of authoritarian breakdown. I suggest that this could be driven by several plausible factors, for example that education levels reflect high state capacity. Second, I argue that when educated citizens mobilize against dictatorships, they are more likely to promote authoritarian breakdown and democratization than other citizens. This is supported using novel data on the education profile of 200 protest campaigns aiming for regime change. These findings indicate that education matters for the prospects of regime change in autocracies, but the relevant feature is not nationwide education levels, but rather the education level of the mass protest campaigns seeking regime change.

Publication status:
Submitted

4.1 Introduction

Is an educated population a threat to a dictator's survival? The notion that increases in education will induce political instability and democratization in authoritarian regimes is popular, both among scholars and policy-makers. For example, Bueno de Mesquita and Smith (2012, p.110) notes in his "Handbook for dictators" that education services are "dangerous public goods that should be doled out carefully". Diamond (2012, p.43) claims that China is likely to democratize within a generation, mainly due to the rise of an educated middle class. Meanwhile, the uprisings and downfalls of authoritarian leaders in the Arab Spring are commonly explained by an expanded access to education in the region, proposed by observations that the demonstrators in Tahrir square and elsewhere were often students or college-educated youth expressing their discontent with authoritarian regimes and preferences for more accountable government (see e.g. Campante and Chor, 2012b). If we are to believe these claims, dictator's all over the world should shiver from the rapid improvements in access to education that have taken place in the last decades.

The notion that education enhances prospects for democratization is in line with modernization theory, as elaborated on by Seymour Martin Lipset (1959). Several studies have investigated this claim empirically, fueled by the strong positive correlation – both at the country level and individual level – between schooling on the one hand and both political participation and democracy on the other. Some of these argue and present corroborating evidence that education breeds democracy (Bobba and Covello, 2007; Castelló-Climent, 2008; Sanborn and Thyne, 2014; Murtin and Wacziarg, 2014), while others claim that the correlation between education and democratization is not due to a causal effect but rather driven by omitted variables or reverse causality (Acemoglu et al., 2005; Acemoglu, Gallego and Robinson, 2014). Hence, the literature offers no unequivocal answer to the question of whether educated dictatorships are more likely to experience democratization.

At the same time, extant literature suffers from two important limitations that may explain these inconclusive findings. First, most existing studies look exclusively at the general relationship between education and various indicators of transitions from autocracy to democracy. Yet, these tests may be too highly aggregated to capture the process through which regime change occurs in dictatorships. In order to fully understand and explain democratization, one should study not only the correlates of *democracy* but also the determinants and processes that influence *authoritarian failure* – which is a precondition for any regime change (see, e.g. Przeworski and Limongi, 1997; Kennedy, 2010; Miller, 2012). Relatedly, several studies demonstrate how aspects of socio-economic development, such as economic growth, high GDP per capita and generous public services, often

strengthen the dictator's grip on power – in contrast to expectations from modernization theory (see, e.g. Bueno de Mesquita and Smith, 2010). Rich countries may very well have a higher likelihood of a new democracy emerging *once the dictator has stepped down*, and a higher likelihood of democratic stability, but they are not more likely to experience authoritarian failure (see, e.g. Fearon and Laitin, 2003; Kennedy, 2010). Education might have similar stabilizing effect on dictatorships, in line with anecdotal evidence suggesting that many of the most highly educated dictatorships, such as Russia, Bahrain and Singapore, also have strong record of regime endurance and political stability. Yet, the effects of a more educated population on authoritarian stability and breakdown have yet to be investigated empirically in a systematic manner.

Second, extant studies rely on highly aggregated indicators of education that measure the average level of schooling offered to a population. This may be problematic, insofar as the most plausible arguments for why democracy should emerge from education are not related to the education level of a society per se, but focus on the role of educated citizens and their mobilization for regime change. This mirrors anecdotal evidence of students or educated groups being at the forefront of mass movements against authoritarian regimes, such as the student protests in Hong Kong 2013, students protesting in the Ukrainian orange revolution or university professors, and lawyers taking part in the 1986 revolution in the Philippines. As I discuss below, the most plausible accounts for why education should promote transitions from autocracy to democracy involve individuals or groups of people altering their preferences, opinions and behavior as a result of education. An average education indicator at the state level may hide substantial within-country variation and capture a variety of other aspects of the relevant regime. At the same time, it does not necessarily tell us whether the actors that carry out political opposition culminating in regime change are educated or not. Hence, existing findings relying on average education indicators may suffer from “ecological fallacies” – emerging due to country-level data being used to infer about mechanisms occurring at the meso level (Gerring, 2011).

Accounting for these shortcomings, I offer new evidence that may help to explain the inconclusive findings in extant literature as well as the paradox of numerous stable highly educated dictatorships and the common occurrence of educated protesters. The paper makes two main points. First, when shifting the focus to explaining authoritarian regime breakdown and mechanisms occurring in dictatorships, I argue that we should not necessarily expect educated dictatorships to be more vulnerable. Authoritarian regimes offering many years of schooling to a large share of its citizens are likely to have high levels of capacity, both for offering goods and public services to its citizens and for staying in power. In line with this argument, I present evidence that a high education level is actually associated with a *lower* likelihood of authoritarian breakdown.

Second, I argue and show that *when educated citizens protest*, they are more likely to overthrow authoritarian regimes than less educated protesters. In brief, I argue that educated citizens have the capacities and resources that enable more efficient anti-regime resistance, more capable of overthrowing dictatorships, such as communication abilities, knowledge and leadership skills. This argument is supported using original data on the education profile of 203 protest campaigns aiming for regime change, recording whether educated groups or individuals a) participated in the protest movement and b) dominated the protest movement or c) initiated the movement. I present results suggesting that ongoing protest campaigns with a high degree of involvement by educated (rather than uneducated) protesters increase the likelihood of authoritarian breakdown in the preceding year. There is also evidence that ongoing protest campaigns with educated participants increase the likelihood of democratization in authoritarian regimes in the preceding year.

In sum, these findings suggest that existing studies have been too aggregated and insufficiently concerned with developing tests that capture the causal mechanisms underlying the education-democracy argument, and hence do not offer a complete picture of how education influences the prospects of regime change in authoritarian regimes. Not only do they not account for the processes promoting authoritarian failure, they may also yield biased conclusions, as we can not infer from state-level indicators of education to the education profile of those who bring down dictators. I find that education does promote authoritarian breakdown and democratization, but not in the ways assumed in existing literature: While high nationwide education levels are associated with lower likelihood of authoritarian breakdown, highly educated protesters should promote authoritarian failure. Hence, rather than looking to variation in education level between countries to explain prospects for democratization, we should look to the educational profile of mass protest movements to explain prospects for authoritarian breakdown.

By distinguishing between nationwide education levels and the education level of the protesters, I contribute to solving the apparent tension between expectations that dictatorships investing in education should also be more stable – mirroring cases such as the Bahraini and Singaporean regimes – and expectations that education should be a force for regime change through successful mass protest – reflecting instances such as the Ukrainian orange revolution and the 1986 revolution in the Philippines.

4.2 Education and democratization

The proposition that an educated population is conducive to democracy is deeply engrained in the political science literature. For instance, Aristotle argued that “The best

laws, though sanctioned by every citizen of the state, will be of no avail unless the young are trained by habit and education in the spirit of the constitution" (Aristotle, 2000). Aristotle argued that education induces moderation and tolerance that is necessary to avoid political extremism, and keep the state in balance from corruption and oppression. Later, the education-democracy relationship was maintained by Dahl (1971) who argued that education brings about democracy because literacy stimulates the types of political participation necessary to maintain representative institutions. Most of the early studies paid attention to education as a key requisite for democracy more generally, and did not specify further whether education is a determinant of democratic transitions rather than simply a condition for democratic survival. Huntington (1991), however, argued that education was an important explanation of the third wave of democratization.

The education-democracy relationship was elaborated on in Lipset's (1959) modernization theory, maintaining that affluent and economically developed societies are more likely to experience democracy. In Lipset's argument, education was considered a crucial component of economic modernization: "If we cannot say that high level of education is a sufficient condition for democracy, the available evidence does suggest that it comes close to being a necessary condition", argued Lipset (1959, p.80). Lipset assumed that education would lead already existing democracies to flourish and persevere, but his argument also implied that education would bring about democratization in previously autocratic regimes.

Lipset's thesis is consistent with the fact that indicators of education are highly correlated with democracy levels in global cross-sections.¹ Yet, there are mixed findings when it comes to the question of whether education promotes transitions to democracy. As of today, most empirical contributions study the effects of an indicator of average education levels on changes in democracy levels or transitions from autocracy to democracy. Many of these studies argue and present corroborating evidence that education does indeed induce democratization (Barro, 1999; Glaeser et al., 2004; Papaioannou and Siourounis, 2008; Murtin and Wacziarg, 2014; Sanborn and Thyne, 2014).²

Acemoglu et al. (2005) challenge these findings, arguing that the positive correlations from previous studies mainly capture cross-sectional variation. Using fixed effects models they find no positive relationship between changes in education and changes in democracy (see, also, Acemoglu, Gallego and Robinson, 2014). Hence, they conclude that the positive correlation is most probably due to a spurious relationship driven by omitted

¹For instance, the correlation between school enrollment and the Polyarchy index from V-Dem is 0.69 in the sample used in this study.

²A recent sensitivity analysis of the determinants of democracy also indicate that education is one of the most robust predictors of democratization measured as changes in polity score, although no such effect is found when using the ACLP dichotomous measure of democracy (Hegre, Knutsen and Rød, 2012)

variables that are fairly constant at the country level, such as institutional history or cultural characteristics. In line with their own empirical studies looking at income level and democracy (Acemoglu et al., 2008) they interpret this finding as evidence of how education and democracy have developed jointly and that historical factors, such as the mortality rates faced by the European settlers and the density of indigenous population, are important determinants of both outcomes (but, see Castelló-Climent, 2008). However, this finding has again been challenged by Castelló-Climent (2008), arguing that the estimators used by Acemoglu et al are not appropriate when variables are highly persistent such as education and democratic institutions (see also Bobba and Coviello, 2007). Hence, although the findings are mixed, several studies indicate that education promotes democratization.

The literature has much less to offer when it comes to the causal mechanisms tying education to democracy. One exception is Campante and Chor (2012*b*, p.169), suggesting that education should induce authoritarian instability through promoting “political activities of a protest nature” (see, also Sanborn and Thyne, 2014). Glaeser, Ponzetto and Shleifer (2007) offer another attempt at specifying a theoretical model, although they do not test it empirically. They argue that education may promote democracy because it raises the benefits (or reduces the costs) of political participation, and that this will make democracy a more feasible option the larger the share of educated citizens. Most other contributions also seem to hinge on implicit assumption that education influences democratization through political protest or anti-regime mobilization, but this is rarely tested empirically.

As a result of this lack of attention to theoretical mechanisms, as well as the highly aggregated approach of existing empirical contributions, we know very little about how and in what ways education should impact on prospects of authoritarian instability or democratic transitions. This includes an incomplete understanding of how education influences processes and dynamics within authoritarian regimes, such as how education influences anti-authoritarian mobilization. In addition, we know little about how education impacts on the preferences and resources of the relevant actors that are potential participants in anti-regime opposition – the micro-mechanisms of regime change.

The next two sections elaborate on these two shortcomings. In doing so, I suggest that the expected effects of education on authoritarian regimes and prospects of democratic transitions is likely much more nuanced than what previous contributions have indicated. Implications of these discussions are tested in section 6.

4.3 Education and authoritarian regime stability

As noted, existing literature has relied exclusively on indicators of transitions to democracy or changes in democracy levels to analyze how education influences prospects of regime change in authoritarian regimes. This section discusses how restricting the focus to instances of democratization or changes in democracy level yields incomplete insights into the processes that promote authoritarian instability and regime change.

As pointed out in recent contributions, students of democratization should distinguish between the *onset* and *outcome* of institutional change (see e.g. Kennedy, 2010; Miller, 2012). In order for a successful democratic transition to take place, two distinct events need to happen: First, the authoritarian regime must break down, due to for instance a coup d'etat or popular uprising – initiating the onset of regime change. Second, the failed authoritarian regime must be replaced by a democracy - which is the outcome of the regime change. Although these two processes often go together, they should not be conflated: First, authoritarian regime breakdown does not necessarily entail democratization. In fact, most authoritarian breakdowns culminate in the emergence of new authoritarian or non-democratic regimes (see e.g. Geddes, Wright and Frantz, 2014). Immediately after an authoritarian regime collapses there is a lot of uncertainty with regard to the outcome, which can be a new authoritarian regime such as in Egypt, a semi-democratic government as in Tunisia, or a civil war breakout as in Libya in 2011. Second, the outcome and onset of the regime change may have different and even contradictory explanations. For example, instances of popular protest in neighboring countries may induce mass mobilization and regime downfall, but not the institutional outcome. At the same time, socio-economic conditions may determine why the regime downfall leads to the emergence of a new democracy rather than a new dictatorship, but this may not matter for regime breakdown. Yet, by restricting the theoretical and empirical focus to only studying democratic transitions, one fails to account for regime change onset, or why the dictatorship collapses – regardless of its outcome. This yields an incomplete picture of the impact of education on dictatorships.

The importance of identifying the determinants of authoritarian regime instability is illustrated by a couple of recent studies re-investigating the effect of income levels on democratization. For instance, contrary to modernization theory, Kennedy (2010) finds that economic development (measured by GDP per capita) lowers the overall likelihood of authoritarian breakdown. He argues that as a country develops economically, state capacity increases and the regime gains access to a vast amount of resources to maintain stability. At the same time, Kennedy argues that income increases the likelihood that the authoritarian breakdown culminates in a new democracy rather than democracy,

and this explains why income is correlated with democracy. This makes it difficult for individuals and groups demanding political inclusion to overthrow existing institutions (Kennedy, 2010).³ Similarly, Miller (2012) shows that economic development strengthens autocratic regimes in the sense that it reduces the likelihood of violent leader removal. Simultaneously, greater development increases the probability of democratization, but only if a violent leader removal has occurred in the recent past. Hence, democratization is arguably best promoted by a combination of regime vulnerability and pro-democratic citizens.

A similar logic may apply to the effects of education on authoritarian regime stability. First, the ability to enroll a large amount of citizens could be a testament to the regime's strength or capacity. A regime with high levels of school enrollment may also be able to offer high-quality governance more generally, through offering goods such as other public services, economic productivity or a high level of security and order. Such government achievements may help to neutralize popular discontent due to lack of democratic freedoms and reduce political instability. One important aspect of the ability to offer high-quality governance is “state capacity”, defined as the ability of the political system to implement decisions and policies (see, e.g. Hendrix, 2010). If education is a proxy for high state capacity, it may also be associated with a higher ability to successfully conduct repression to prevent mass mobilization, or the ability to carry out systematic censorship or surveillance to strengthen regime control over potential opposition groups. Second, schooling access – both at the primary, secondary and tertiary level – is a public service that in itself may bolster popular support. High school enrollment rates is a tangible, easily quantifiable signal of public-goods provision by the regime. For instance, the Cuban communist regime has received a lot of appraisal – both domestically and also internationally, for its education system, being the only country in Latin America and the Caribbean region to have reached the United Nation’s “Education for All” objectives (Unesco, 2015). This is a strong signal to Cuban citizens that the government cares for the well-being of its citizens, that may mitigate popular discontent. Finally, education may also be endogenous to regime stability in the sense that regimes that have experienced a high degree of political stability are the ones that invest in schooling. If the dictator faces a low potential of opposition from below, offering education to its citizens is less risky, as the likelihood that education will be converted into anti-regime mobilization is less precarious. At the same time, regimes that feel existentially secure can spend less on the military and security apparatus, which allows for more resources on public services –

³This builds on the empirical finding that economic development is generally stabilizing for all political regimes, both democratic and authoritarian, while lower levels of economic development tend to destabilize all regimes (Collier et al., 2003; Gates et al., 2006; Thyne, 2006; Fearon and Laitin, 2003; Bueno de Mesquita and Smith, 2003).

including education. If these mechanisms are at play, we could expect that high education levels to be associated with lower likelihood of regime breakdown.

The hypothesis that high education levels are associated with authoritarian survival, as well as the findings that income stabilizes dictatorships, underscores a more general point. Factors considered as the classical determinants of democracy, such as socio-economic development and a “democratic culture”, are often variables that really should be expected to stabilize *all political regimes*. While economic performance and investments in public services may reduce revolutionary grievances and anti-authoritarian sentiments, education and a democratic culture involves a dedication to moderation and compromise – traits that may discourage contentious collective action to overthrow existing regimes (MacCulloch, 2004). Hence, we should not necessarily expect such factors to promote authoritarian breakdown. This is consistent with the argument that economic modernization is related to democracy because it stimulates democratic survival, not because it induces democratic transitions.⁴

4.4 Educated protesters and prospects for autocratic breakdown

Although we should not necessarily expect authoritarian regimes with high levels of school enrollment to be more prone to political instability, the presence of educated anti-regime forces may still influence regime stability. In other words, highly educated people may be more likely to promote authoritarian regime breakdown, giving rise to the familiar image of students and highly educated protesters mobilizing against autocracy. More specifically, I surmise that political protest movements consisting of educated protesters should be more likely to promote authoritarian regime breakdown than movements consisting of less educated or uneducated groups or individuals, such as peasants or industrial workers.

If education is to destabilize authoritarian regimes, this should be expected to happen through actions carried out by educated individuals and groups. Hence, it is necessary to specify exactly how education may alter people’s choices and behavior in ways conducive to authoritarian breakdown. Here, I focus on two mechanisms, which both describe how education raises people’s resources and capacity for carrying out collective action,

⁴Indeed, a democratic transition often requires a certain amount of conflict; either in the form of non-violent conflict of interest or violent conflict (see Miller, 2012; Kennedy, 2010; Bermeo, 1997; Wood, 2003). There are many examples of democracy arising out of very different climates than the peaceful, consensual social order which is fundamental to democratic survival. For instance, Wantchekon (2004) rejects the popular claim that democracy cannot arise out of the chaos of civil war, showing how democracy can emerge in instances of anarchy, such as in Mozambique and El Salvador. Studying transitions in South Africa and El Salvador, Wood (2003) launches what she calls the insurgent path to democracy, characterized by sustained mobilization by poor and working-class people.

implying that protest campaigns consisting of educated protesters should be more efficient at overthrowing authoritarian regimes.⁵

First, educated individuals should have skills and knowledge that promotes effective protest campaigns. This includes language and communication abilities, that facilitate effective planning and organizational capacities. The most effective protest campaigns (that are most threatening to the dictator) are those that are able to persuade a large number of fellow citizens, and thereby successfully undermine the dictators legitimacy (Chenoweth and Stephan, 2011). Effective communication strategies are crucial to mobilizing and persuading the larger citizenry. Indeed, many protest campaigns have conducted highly professionalized communication campaigns, drawing on networks of academics and journalists and using tools such as professional mass media advertisements, use of polls to develop effective strategies, and in more recent time, professional social media advisors (see, e.g. Chenoweth and Stephan, 2011; Nepstad, 2011). Such campaigns have arguably drawn heavily on educated campaign members and their expertise, such as journalists and university professors. Protest campaigns where individuals lack these resources may not, for instance, be equally efficient when it comes to developing effective communication strategies to attract more protesters, both from the same social segment as existing participants and from the outside.⁶ Educated citizens are also usually more knowledgeable about politics and society, and hence more able to maneuver on the political stage and make sense of both the domestic and international political context.⁷

Second, education should stimulate social networks and cooperation abilities, that increases group efficiency Putnam:1993, Newton:1997. Schools and educational institutions bring together students from different social and geographical segments of society that otherwise would interact much less frequently. Every day, students cooperate with and socialize with fellow students from outside of her family or neighborhood, and this may strengthen social networks - which are crucial to efficient collective action (McAdam, 1986; Siegel, 2009). This helps to build a coherent organization that stays unified and sticks to the same strategy rather than becoming fractionalized, thereby loosing leverage. At the same time, exercising social interaction and cooperation is an important aim of education, which teaches people to work and study together through classroom discussions, joint learning tasks and group projects. Educated individuals are also more likely to end

⁵Education may also raise grievances against authoritarian regimes, due to for instance preferences for more democracy and protection of civil liberties. Here, such preferences are considered as being constant, as I am seeking to explain variations in efficiency between on-going protest campaigns with an explicit goal to overthrow authoritarian regimes. As these protest campaigns emerged in the first place, I assume that their members have some kind of grievance against the autocratic regime.

⁶The size of the protest movement has been identified as an important predictor of success protest outcome (in terms of achieving their goals) (Chenoweth and Stephan, 2011).

⁷For more thorough discussions of educated citizen's particular skills, see chapters 1 and 6

up in jobs where social interaction and cooperation is important, for instance through leadership positions that require social abilities.

Altogether, the combination of resources such as language and communication skills, political knowledge, social networks and cooperation abilities should increase the likelihood that protest campaigns with more educated participants are successful at overthrowing authoritarian regimes.

Summing up sections 3 and 4, we should expect high nationwide education levels to have a different effect on authoritarian breakdown and regime change than high education levels among participants in anti-regime campaigns. While high education levels should not necessarily promote regime change – rather, it may promote regime survival – highly educated protesters could be a potent threat to authoritarian stability.

4.5 Data and empirical model

4.5.1 Dependent variables

To measure autocratic breakdown, I utilize the dichotomous indicator of regime failure from the Geddes, Wright and Frantz (2014) data set, with the value 1 if the authoritarian regime collapsed in a given year, and 0 otherwise. One advantage of this indicator is that it captures autocratic breakdown regardless of whether the regime is replaced by a democracy or any other form of regime, reflecting the approach in this paper. At the same time, it codes the regime as surviving if a leader succession takes place within the same set of rules and institutions, for instance, if the son of the monarch inherits the throne.

In addition, I also analyse the effects of education levels and educated protest campaigns on democratization in authoritarian regimes. Democratization is commonly operationalized using a continuous democracy index as dependent variable and controlling for lagged levels of that index. In order to make my results easily comparable to previous studies, I operate with such a specification in the main analysis. By controlling for lagged democracy, these models capture changes in democracy from one year to another. More specifically, I use the Polyarchy index from the Varieties of Democracy project (V-Dem), which is a continuous indicator of electoral democracy, measuring whether the chief executive is elected in free and fair elections, the level of suffrage and freedom of speech and association (Coppedge et al., 2015). A problem with this model specification is that it does not identify whether the results are driven by positive changes in democracy score (democratic upturns) or negative changes in democracy score (democratic downturns). As this paper seeks to explain democratization (and not democratic backlashes), I

therefore also run robustness tests with an indicator of democratic upturns as dependent variable, to check if the results are similar when excluding backlashes, and the results are included in the appendix. Democratic upturns are measured using positive changes in the Polyarchy index from V-Dem.

In addition, it can be argued that positive changes in electoral democracy score can capture different processes, including both small incremental changes from year to year, or more large-scale transitions from autocracy to democracy that happen from one year to another. To investigate whether the identified effect of education in the main analysis is not only driven by small incremental changes, but also applies to more rapturous changes, I also run robustness tests using the Boix, Miller and Rosato (2013) dichotomous democracy indicator as dependent variable (see appendix for results). Reflecting the focus on processes and mechanisms occurring within authoritarian regimes, the main empirical analysis excludes all countries that were coded as democratic by Boix, Miller and Rosato (2013) in year $t-1$ (when the outcome variable is measured in year t).

4.5.2 Independent variables

To measure average education levels, I use indicators of primary, secondary and tertiary enrollment which I combine into an average education enrollment measure. The three enrollment indicators, taken from the Varieties of Democracy dataset Coppedge et al. (2015) are created by combining UNESCO's adjusted net enrollment rates and the long-run education enrollment data from Barro and Lee (2013). The advantage of using enrollment rates is that it captures mass education, by measuring the share of the population that receives schooling.⁸ Commonly used indicators of mean years of education, on the other hand, do not account for the distribution of schooling, and a country may receive a high score if most citizens only attend schooling for a few years, receive, while a small minority receives extensive education. Yet, I also run robustness tests using an indicator of years of schooling from Barro and Lee (2013), and the results generally hold up (see appendix).

To measure the presence of educated protest campaigns, I rely on novel data recording the educational background of all protest campaigns aiming for regime change that are listed in the NAVCO 1.0 dataset (203 in total). In this data set, a protest campaign is defined as a series of "observable, continuous, purposive mass tactics or events in pursuit of a political objective (Chenoweth and Lewis, 2013). Details about the coding scheme and discussions regarding reliability and validity of the data are described in

⁸It does not, however, account for the quality of the schooling. As this paper aims to study the long-term effects of education on authoritarian regime breakdown, it is not feasible to incorporate indicators of education quality in this study, as they have quite short time-span and relatively poor coverage for autocracies.

chapter 6 in this dissertation. In brief, the data includes a variable recording whether the sources explicitly report that students or people with tertiary education at a minimum *participated* in the protest campaign, or whether these groups were not reported to have participated at all. In addition, I coded a variable recording whether students or educated protesters are reported to have *dominated* the protest movement, and a variable recording whether the protest campaign *originated among* students or educated citizens. The three variables were coded based on sources such as news reports, books and databases and encyclopedias such as “The international encyclopedia of revolution and protest” (Ness, 2015) and the “Global Non-Violent Action Database” (Swarthmore, 2015). In the main analysis, I combine these three indicators into one lexical index, capturing the degree of involvement by educated protesters. This index ranges from 0 to 3, with the value 0 if no educated or student groups are reported to have been involved and 3 if the movement originated among and was dominated by student or educated groups. To assess the robustness of my results, I also run models using each of the three composite indicators to measure the education profile of protest campaigns (see appendix Table 9.18-9.19).

4.5.3 Confounders

To avoid problems of collinearity or “garbage can regressions” (Achen, 2005), I present relatively parsimonious models in the main analysis, but several key confounders are included in most models – both when looking at changes in democracy score and when looking at autocratic breakdown. First, I control for log GDP per capita, which is expected to both lead to higher schooling levels (Bils and Klenow, 2000), higher democracy levels (e.g., Moral-Benito and Bartolucci, 2012; Heid, Langer and Larch, 2012), and a lower likelihood of regime breakdown (Miller, 2012). Second, I always control for share of GDP per capita that is due to oil and gas incomes, and this variable is also log-transformed. A high level of revenues from natural resources have been shown to reduce the likelihood of authoritarian regime breakdown and democratic transitions (Wright, Frantz and Geddes, 2015), and may also influence education levels, for instance through reducing the need to invest in human capital to promote economic growth (Gylfason, 2001). Third, I control for levels of urbanization, which is expected to lead to a higher education level, as well as higher levels of democracy (Barro, 1999). As a robustness check, I also include additional controls related to country-level socio-economic characteristics or variables related to variation in autocratic regime types from Geddes, Wright and Frantz (2014) (for results, see Table 9.15 and 9.16 in the appendix).

In models with autocratic breakdown as dependent variable, problems of time dependency are likely to emerge – the probability of regime breakdown is likely to be related to the age of the regime. To deal with this, I follow the advice of (Carter and Signorino,

2010) and include time since last breakdown in the equations, as well as time squared and time cubed, to form a smooth function of time.

I also include country fixed effects, to account for omitted time-invariant country-specific factors, for instance related to a country's history, cultural background or colonial heritage that may both have promoted long-run schooling investments and high levels of democracy. Finally, year dummies are also included, to account for country-invariant year-specific trends in both education levels and authoritarian instability or democratization.

4.6 Empirical analysis

I start off the empirical analysis by replicating models commonly used in existing literature to investigate the relationship between education and democratization - using changes in electoral democracy (measured using V-Dem's Polyarchy index) as dependent variable and an average education indicator as independent variable. Next, I proceed to my own empirical contribution, in which I exclude all existing democracies from the analysis in order to study the effects of education on prospects of regime instability and change in dictatorships. The analysis is conducted in two parts. In the first part, I investigate how average education indicators at the state-level impact on 1) democratization in authoritarian regimes and 2) authoritarian regime breakdown. In the second part of the analysis, I add an indicator of ongoing protest movements by educated citizens to the models analysed in step one. Hence, these latter models investigate the relationship between *educated campaigns* and regime instability and change in dictatorships.

4.6.1 Country-level variation in education

Results from OLS models replicating tests of the relationship between school enrollment and changes in electoral democracy are summarized in Table 4.1. In line with extant findings, there is some evidence that education is positively related to changes in democracy in the preceding year, but the results are far from robust. The coefficient estimate is positive and significant at the 0.001 level in the bivariate model (A1). When controlling for GDP per capita, an indicator of urbanization and income from oil and gas (in model A2), the t-value is reduced by more than half of its size, and when also including country fixed effects in models A3 the t-value is reduced to closer to 1. This suggests that positive coefficient estimate is driven by omitted time-invariant factors. Finally, when including year dummies in addition to country fixed effects in model A4, the coefficient estimate actually becomes negative, but it is far from reaching conventional levels of statistical

significance. Hence, according to standard norms of inference we can not conclude from the results that education is related to positive changes in democracy.⁹ This is in line with findings such as Acemoglu (2005), who identifies no effect of education on democracy when including country fixed effects.

Table 4.1: Changes in electoral democracy and school enrollment.

DV: Model:	A1 Electoral democracy logit	A2 Electoral democracy logit	A3 Electoral democracy logit	A4 Electoral democracy logit
School enrollment	0.0196*** (6.31)	0.0137** (3.12)	0.0141 (1.19)	-0.0207 (-1.10)
Ln GDP p.c.		0.000591 (0.44)	-0.00306 (-0.90)	0.000900 (0.26)
Urbanization		0.0132** (2.76)	0.0676*** (4.59)	0.0440** (3.20)
Ln Oil and gas income		-0.00489+ (-1.92)	-0.00247 (-0.42)	-0.00327 (-0.50)
Electoral democracy _{t-1}	0.980*** (368.82)	0.976*** (278.69)	0.945*** (170.62)	0.940*** (154.30)
Country dummies?	no	no	yes	yes
Year dummies?	no	no	no	yes
N	10372	4905	4905	4905
N (countries)	158	103	103	103
R ²	0.008	0.013	0.0337	0.103

Notes: + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$. T-values in parentheses. Electoral democracy is measured using the Polyarchy index from V-Dem (Coppedge et al., 2015). Lagged democracy is the Polyarchy score in year t-1. School enrollment is the average of primary, secondary and tertiary enrollment, from Barro and Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from the table.

As noted, if we are to understand how education affects democratization we should also study how education influences processes of instability and regime change within authoritarian regimes, and particularly how it influences the likelihood of regime failure. I now turn to authoritarian regimes and authoritarian breakdown, by excluding all regimes classified as democracy according the Boix, Miller and Rosato (2013) democracy indicator. The results are presented in Table 4.2.

To assess how the estimated coefficient for education enrollment changes when restricting the sample to only autocracies rather than all regimes, I start by presenting models similar to those presented in Table 4.1, investigating the effects of education on changes in democracy levels. Hence, these models investigate the effects of education on democratization in authoritarian regimes. As seen from models B1-B4, there is some evidence that education is associated with positive changes in democracy score. The education enrollment coefficient is positive and significant at the 0.001 level in the bivariate model B1, and significant at the 0.01 level when adding key control variables, including

⁹Nor is there any robust evidence that education is positively related to democratization when operating with average years of education as independent variable or changes in Polity score or democratic upturns as dependent variable (see section 9.2 in the appendix).

log GDP per capita, urbanization and log oil and gas income. Yet, this positive result is far from robust. When adding country fixed effects in model B3 the coefficient estimate for school enrollment changes from positive to negative (but is insignificant), and when including both country and year fixed effects in model B5, the coefficient estimate is still negative and the size of the t-values increases to -2.6. As I am interested in whether increases in school enrollment are associated with positive changes in electoral democracy (rather than simply cross-sectional covariation), models B3 and B4 are more appropriate model specifications. Hence, the results suggest that there is no positive association between education and democratization in authoritarian regimes. Rather, there is more convincing evidence that the this relationship is negative. The latter result is in line with the above discussions suggesting that high enrollment rates may proxy for authoritarian regime capacity and strength, which should be expected to be negatively associated with democratization.

Table 4.2: Changes in electoral democracy and regime breakdown in authoritarian regimes

DV: Model:	B1 Electoral democracy OLS	B2 Electoral democracy OLS	B3 Electoral democracy OLS	B4 Electoral democracy OLS	B5 Autocratic breakdown logit	B6 Autocratic breakdown logit	B7 Autocratic breakdown logit	B8 Autocratic breakdown logit
School enrollment	0.0320*** (6.03)	0.0290** (2.98)	0.0151 (0.83)	-0.0590** (-2.58)	-1.810** (-2.62)	-3.065** (-3.28)	-6.920** (-3.14)	-7.484** (-2.85)
Ln GDP p.c.		0.000388 (0.14)	0.000390 (0.10)	0.0116** (2.90)		0.388 (1.63)	0.471 (1.14)	0.374 (0.78)
Urbanization		0.0216* (2.01)	0.0617** (2.96)	0.000457 (0.02)		0.151 (0.15)	3.821 (1.47)	2.562 (0.76)
Ln oil and gas income		-0.0113+ (-1.94)	-0.00159 (-0.18)	-0.000282 (-0.03)		-0.553 (-1.13)	-1.134 (-1.14)	-1.343 (-1.27)
Electoral democracy _{t-1}	0.941*** (121.94)	0.944*** (107.31)	0.917*** (87.94)	0.887*** (77.02)		3.339*** (5.10)	3.490*** (4.36)	4.365*** (4.46)
Country dummies?			yes	yes			yes	yes
Year dummies?				yes			yes	yes
Time, Time ² , Time ³ ?					yes	yes	yes	yes
N	4331	3575	3575	3575	2856	2743	2161	2133
N (countries)	122	117	117	117	106	105	75	75
R ²	0.005	0.008	0.0212	0.103		-665.7	-623.6	-525.5
ll								-486.3

Notes: + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$. T-values in parentheses in models B1-B4, and z-values in parentheses in models B5-B8. Dependent variable in models B1-B4 is the Polyarchy index from V-Dem (Coppedge et al., 2015). Lagged electoral democracy is the Polyarchy score in year t-1. Dependent variable in models B5-B8 is authoritarian regime failure from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro and Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from the table. The sample is restricted to regimes coded as autocracies by Boix, Miller and Rosato (2013).

As noted, an indicator of change in democracy score does not distinguish between democratic upturns and democratic downturns. As this paper focuses on democratization, I would like to tease out whether the findings are similar when only looking at democratic upturns (or whether the above results are driven by democratic downturns). The results, presented in Table 9.5 in the appendix, indicate that the estimated effect does not change substantially when looking at democratic upturns: There is no robust evidence that national education enrollment is related to democratic upturns. Rather, there is some evidence that the association is negative. The main findings also hold up when operating with a dichotomous democracy measure from Boix, Miller and Rosato (2013) as dependent variable (see appendix Table 9.6).¹⁰

In models B5-B8 I investigate the relationship between schooling enrollment and autocratic breakdown, to investigate how education influences authoritarian instability *regardless* of what the outcome of that instability is. Here, I run logit models with the dichotomous indicator of autocratic breakdown from Geddes, Wright and Frantz (2014) as dependent variable. All models control for time (since last breakdown), time squared and time cubed. The coefficient estimate for school enrollment is negative in all four models, with z-values ranging from at least -2.6 to -3.3. In the simple bivariate model (B5), the coefficient estimate is negative and significant at the 0.01 level. Both the z-value and the negative coefficient estimate increase in size when adding key controls in model B6 – including electoral democracy, log GDP, urbanization and log oil and gas income. Finally, the negative coefficient estimate is more than doubled in size when adding fixed effects in model B7 as well as both country fixed effects and year dummies in model B8, and the coefficient estimate is still significant at the 0.01 level in both models. This result adds further support to the argument that a high nationwide education level is not necessarily a force for authoritarian breakdown or democratization, contrary to modernization theory.¹¹ Rather, high enrollment rates seem to be associated with a lower likelihood that the authoritarian regime will fail. This is not necessarily due to a causal effect of school enrollment on regime survival, however, but could be due to other factors that coincide with high education levels, such as high state capacity. Moreover, it could be

¹⁰The main result is also robust to operating with average years of schooling as education indicator, to including additional controls (such as authoritarian regime type) and to restricting the sample to countries defined as autocracies according to Geddes, Wright and Frantz (2014). The results are included in section 9.3 in the appendix).

¹¹The main result is robust to including additional control variables (including indicators of authoritarian regime type), to restricting the sample to autocracies according to Geddes, Wright and Frantz (2014). When replacing school enrollment with an indicator of years of education, the negative association between education and autocratic breakdown is somewhat less robust, however. The coefficient estimate for years of schooling is always negative, but when including country fixed effects and year fixed effects the coefficient estimates is not statistical significant at conventional levels. Results from these robustness tests are included in section 9.4 in the appendix.

due to reverse causation, if more stable autocracies offer more education. In section 7, return to this finding and the suggested mechanisms.

4.6.2 Variation in education levels between protest campaigns

Having investigated the relationship between education levels at the state level and the likelihood of autocratic breakdown or changes in democracy score in dictatorships, the final step of the empirical analysis involves exploring the impact of ongoing protest movements with educated participants. To investigate this, I utilize the index describing the extent to which educated individuals are involved in a protest campaign, ranging from 0 to 3, with the value 0 if no educated or student groups are reported to have been involved and 3 if the movement originated among and was dominated by student or educated groups. This variable is added to the models presented in Table 4.2. In addition, I control for whether there was an on-going protest movement at all in each country-year. These models, investigating either changes in democracy level or authoritarian breakdown, are presented in Table 4.3. In this model set-up, the coefficient estimate for the indicator of educated protest campaigns can be interpreted as increases in either democracy score or the likelihood of autocratic breakdown when the involvement of educated protesters increases – given that there is campaign. The coefficient estimate for the protest campaign variable can be interpreted as the effect of a protest campaign with no educated protesters. All models also control for school enrollment.

Table 4.3: Changes in electoral democracy and regime breakdown in authoritarian regimes. Campaigns by educated protesters

DV:	C1 Electoral democracy	C2 Electoral democracy	C3 Electoral democracy	C4 Electoral democracy	C5 Autocratic breakdown	C6 Autocratic breakdown	C7 Autocratic breakdown	C8 Autocratic breakdown
Educated protesters (index)	0.0125*** (5.15)	0.0128*** (4.78)	0.0140*** (4.85)	0.0115*** (4.03)	0.505*** (3.49)	0.502** (3.28)	0.585** (2.66)	0.613** (2.60)
Protest campaign	0.00421 (1.23)	0.00363 (0.92)	0.00308 (0.71)	-0.0000170 (-0.00)	0.394 (1.55)	0.502+ (1.85)	0.701+ (1.80)	0.766+ (1.85)
School enrollment	0.0270*** (5.03)	0.0219* (2.23)	0.00499 (0.28)	-0.0582* (-2.56)	-1.791** (-3.24)	-3.119*** (-4.28)	-6.990** (-3.13)	-6.587* (-2.44)
Ln GDP p.c.	0.00110 (0.39)	0.00167 (0.44)	0.0121** (3.03)			0.488* (2.34)	0.462 (1.09)	0.317 (0.65)
Urbanization	0.0222* (2.04)	0.0594** (2.87)	0.00591 (0.24)			-0.0658 (-0.08)	1.585 (0.59)	2.767 (0.80)
Ln oil and gas income	-0.0110+ (-1.88)	-0.00177 (-0.20)	0.000492 (0.05)			-0.529 (-1.36)	-0.980 (-0.99)	-1.167 (-1.09)
Electoral democracy_t-1	0.940*** (122.03)	0.943*** (107.42)	0.916*** (88.61)	0.887*** (77.33)		2.873*** (5.15)	3.319*** (4.17)	4.071*** (4.09)
Country dummies?			yes	yes			yes	yes
Year dummies?				yes			yes	yes
Time, Time ² , Time ³ ?					yes	yes	yes	yes
N	4331	3575	3575	3575	2864	2743	2161	2133
N (countries)	122	117	117	117	106	105	75	
R ²	0.832	0.825	0.813	0.833				
ll					-642.9	-602.0	-506.2	-468.4

Notes: + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$. T-values in parentheses in models B1-B4, and z-values in parentheses in models B5-B8. Dependent variable in models B1-B4 is the Polyarchy index from V-Dem (Coppedge et al., 2015). Lagged electoral democracy is the Polyarchy score in year t-1. Dependent variable in models B5-B8 is regime failure from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro and Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from the table. The sample is restricted to regimes coded as autocracies by Boix, Miller and Rosato (2013).

Investigating changes in democracy level in authoritarian regimes in models C1-C4, there is clear evidence that higher involvement by educated protesters is associated with positive changes in democracy, although the coefficient size is moderate, ranging from 0.05 to 0.06. The coefficient estimate for educated protest campaigns is positive in all four models, and always significant at the 0.001 level. This result holds when including key control variables in model C2 – including log GDP per capita, urbanization and log oil and gas income – and when also adding country fixed effects in model C3. The t-value drops somewhat when including year dummies in addition to country fixed effects in model C4, but the coefficient estimate is still positive and significant at the 0.01 level. This finding is also robust to operating with democratic upturns as dependent variable, or to using the indicator of democratic transitions from Boix, Miller and Rosato (2013) (see appendix, Table 9.11 and 9.13).¹² Meanwhile, there is still some evidence that school enrollment is negatively associated with changes in democracy. For instance, the coefficient estimate for school enrollment is negative and significant at the 0.05 level when including key controls, country dummies and year dummies in model C4.

In models C5-C8, autocratic breakdown is the dependent variable, and here I also include the indicator of educated protest campaigns. All models control for ongoing protest campaigns, school enrollment and time since last breakdown. In line with the findings from models C1-C4, the results indicate a positive association between involvement by educated protesters and the likelihood of autocratic breakdown. In the most sparse model (C5), the coefficient estimate for educated protest campaigns is positive with a z-value of 2.9. The coefficient size and the z-value only drop slightly when adding key controls in model C6. When adding country fixed effects in model C7 and both country fixed effects and year dummies in model C8, the coefficient estimate for involvement by educated protesters increases, but the z-value drops to around 2.3 in both models. Hence, there is evidence that educated protesters promote autocratic breakdown even in the most complex models. There is also quite robust evidence that high school enrollment levels are inversely related to autocratic regime breakdown, especially when looking at autocratic breakdown in models C5-C8.

The results above indicate that there is a statistically significant effect of involvement in protest campaigns by students and educated groups on both democratization and autocratic breakdown. Yet, it may be hard to interpret the size of this effect substantially, especially for the logit models with autocratic breakdown as dependent variable. To sim-

¹²The main results are also robust to restricting the sample to autocracies according to Geddes, Wright and Frantz (2014), to controlling for authoritarian regime type, and to alternative measures of educated protest campaigns, including both the dichotomous indicator of whether educated protesters participated in the movement or not and the dichotomous indicator of whether educated protesters dominated the campaign (the results are presented in section 9.4 in the appendix)

Table 4.4: Predicted probabilities of autocratic breakdown, given an ongoing protest campaign.

Education index	Predicted probability
0: No students or educated groups reported to have participated	0.06***
1: Students or educated groups reported to have participated	0.12**
2: Students or educated groups participated, and the campaign either originated among or was dominated by these groups	0.12**
3: Students or educated groups participated, and the campaign both originated among and was dominated by these groups	0.26***

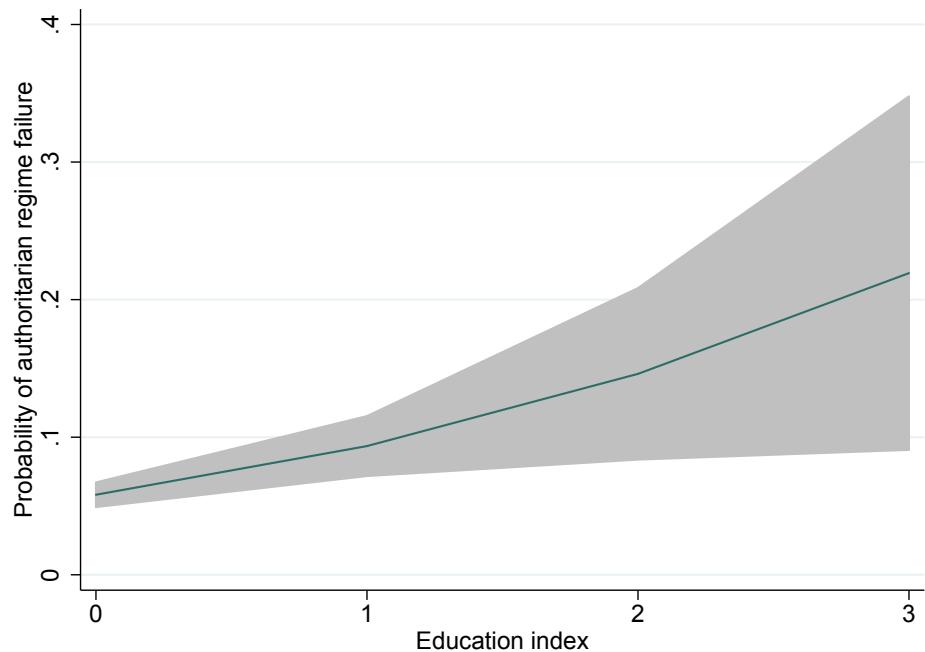
Notes: + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$. Predicted probabilities are calculated using the margin command in Stata 14, and are based on logit regression where included controls are education enrollment lagged democracy and regime duration. All control variables are set to the mean.

plify the interpretation of this result, I also calculate how the predicted probabilities of autocratic breakdown changes with increases in the education index, that varies from 0 to 3. Table 4.4 lists the predicted probability of autocratic breakdown for all levels of the education index, together with an explanation of what each value of the index corresponds to. The predicted probabilities are calculated based on a simple logit model where all control variables are set to their mean. The table indicates that the likelihood of regime breakdown is 0.06 if there is an on-going campaign that does not have students or educated participants. The predicted risk of breakdown is doubled in size if there is an on-going campaign where students or graduates are among the participants, or if there is an on-going campaign either initiated by or dominated by students or educated protesters. Finally, the predicted probability of authoritarian regime failure jumps to 0.26 if there is an on-going campaign with students or educated protesters who both initiated and dominated the protests. Hence, the likelihood of regime breakdown increases by 0.20 when comparing “un-educated” campaigns to the familiar image of “student protests” or mass opposition mainly orchestrated by the highly educated. This means that while only about 1 out of 20 ongoing protest campaigns conducted by uneducated citizens should promote autocratic breakdown, 1 in 3 campaigns that were initiated and dominated by students or educated citizens should culminate in regime breakdown.

Figure 4.1 also illustrates how the likelihood of autocratic breakdown increases gradually as students and educated protesters become more involved in an on-going protest movement.

This evidence supports the argument that educated protest campaigns, due to their skills, knowledge and social networks, should be more likely to induce autocratic breakdown. In addition, it shows that educated protest movements are more likely to promote democratization. This insight is important to understanding the relationship between

Figure 4.1: Predicted probabilities of autocratic breakdown



Predicted probabilities are calculated using the margin command in Stata 14, and are based on logit regression where included controls are education enrollment lagged democracy and regime duration. All control variables are set to the mean.

education and prospects for regime change in authoritarian regimes.

4.7 Exploring the mechanisms

The results presented above indicate that high average school enrollment rates do not promote democratization in authoritarian regimes. Yet, *when* educated citizens protest, they are more likely to promote authoritarian breakdown and democratization in authoritarian regimes. Why this mismatch between country-level education and the role of education at the meso level?

One potential explanation is that the mismatch between effects of nationwide education and the effects of education at the protest campaign level, is due to the fact that high education levels per se do not make it more likely that educated citizens will protest. As noted above, the most plausible mechanism tying indicators of education at the state level to authoritarian instability is through educated citizens carrying out political protest, which is more likely to result in democratization due to the resources of educated citizens. This hypothesized link assumes that high education levels should make it more likely that educated citizens participate in protest.

Table 4.5: Participation by educated citizens in protest movements

DV: DV:	E1 Educated protesters participated in campaign	E2 Educated protesters participated in campaign	E3 Educated protesters participated in campaign	E4 Educated protesters dominated campaign	E5 Educated protesters dominated campaign	E6 Educated protesters dominated campaign
School enrollment	0.919 (0.81)	-1.219 (-0.74)	-8.018** (-2.78)	3.572* (2.04)	4.864+ (1.66)	-5.301 (-1.01)
Democracy _{t-1}	-1.881+ (-1.95)	-2.216* (-2.04)	-1.685 (-1.19)	-2.391 (-1.44)	-2.417 (-1.33)	-2.657 (-1.10)
Campaign	4.042*** (20.01)	3.784*** (19.55)	4.421*** (15.54)	2.631*** (7.98)	2.379*** (7.15)	2.750*** (6.02)
GDP p.c.(log)	0.391 (1.19)	0.227 (0.57)	0.624 (1.24)	0.593 (1.13)	0.180 (0.27)	-0.297 (-0.33)
Urbanization	0.485 (0.36)	3.422+ (1.71)	0.737 (0.25)	-3.326 (-1.55)	-4.027 (-1.16)	-6.049 (-1.08)
oilgas_log	-0.540 (-0.78)	0.362 (0.37)	0.592 (0.49)	0.448 (0.40)	1.659 (0.96)	7.573* (2.44)
Country dummies?	yes	yes	yes	yes	yes	yes
Year dummies?						yes
N	3575	2269	2269	3575	1052	1052
N_g	117	53	53	117	23	23
ll	-536.7	-364.3	-285.7	-227.3	-138.1	-89.52

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is dichotomous variable indicating whether educated citizens or students were reported to have participated in the campaign. School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from the table.

To assess whether this mechanism is at play, I investigate whether high education levels increase the likelihood of protest carried out by educated citizens. The results from this investigation are presented in Table 4.5. Here, the first three models operate with the dichotomous variable indicating whether the protest campaign consisted at least partly of educated protesters or not as dependent variable. As I also control for whether there was an ongoing protest campaign, the main explanatory variable of interest, education enrollment, can here be interpreted as the effect of enrollment on the likelihood that a campaign has some educated protesters versus no educated protesters (given that there is a campaign). In brief, the results indicate that higher enrollment levels do not make it more likely that ongoing protest campaigns consist of educated protesters. In model E1, which only controls for ongoing campaign and log GDP per capita, the coefficient estimate for education enrollment is positive but with a low t-value (0.8). When also including country fixed effects in model E2, the coefficient estimate becomes negative, but the t-value is still quite low. In model E3, which includes year dummies in addition to country fixed effects, the negative coefficient estimate for school enrollment increases substantially in size and becomes significant at the 0.01 level. Hence, there is no evidence that high enrollment rates are associated with a higher likelihood that ongoing protest campaigns consist of educated protesters.

Models E4-E6 investigate whether high enrollment levels increases the likelihood that campaigns are dominated by educated protesters. When only including key control vari-

ables in model E4, the coefficient estimate for education enrollment is positive with a t-value just above 2. The coefficient estimate is still positive but the t-value drops somewhat when also including country fixed effects in model E5. Finally, when including year dummies in addition to country fixed effects, the coefficient estimate becomes negative (but with a quite low t-value of just above 1). Hence, there is no robust evidence that education enrollment increases the likelihood of campaigns dominated by educated protesters, either. Apparently, a high nationwide education level does not equal collective action by the educated. This finding may also help to explain why there is no positive link between education enrollment and autocratic breakdown and democratization in dictatorships, despite the fact that educated protesters seem to be more likely to promote regime change. In sum, despite what seems to be a common assumption in the modernization theory literature, extensive school enrollment does not make it more likely that educated individuals participate in protest. Nor does it make it more likely that the typical “student protests” or the campaign dominated by professionals or the “intelligentsia” emerges. Hence, this finding is a blow to a central premise in modernization theory.

While the null relationship between education level and the presence of educated protest campaigns may seem unexpected, it is not without potential explanations. First, it could result from the endogeneity of education levels to unobservable but possibly quite important confounders, as discussed above. Second, it could be that education has different effects on campaigns, pulling in opposite directions. For example, in highly educated societies, campaigns should be more likely to include educated participants, at least if they reflect the population at large. At the same time however, highly educated citizens in low-education societies should represent the “elite”, which may be the most likely agents of opposition campaigns in *all societies*. Throughout history, there are numerous examples of this: student- or graduate-led protest movements in rural societies with low education. It could also very well be that the educated have fewer grievances in societies where they are in the majority, dampening their inclinations to take part in protests.

Probing further into this finding, a second, and somewhat related explanation for the identified negative effect of education enrollment on autocratic breakdown is related to the argument discussed above that education could be a proxy for the capacity and strength of the authoritarian regime. A regime capable of enrolling a large share of citizens in primary, secondary and tertiary schooling should also be likely to possess the resources and abilities to carry out policies that promote regime survival, such as offering high-quality governance or maintaining a strong security apparatus. This may explain why high enrollment is not associated with a lower likelihood of breakdown. It may also explain why school enrollment does not seem to induce protest by educated

groups. As noted, there is currently little data available on state capacity, and the few indicators that exist have quite poor coverage, especially over time. Yet, to investigate whether this mechanism may account for the evidence that school enrollment is associated with autocratic stability, I run models where indicators of state capacity are added to the models from Table 4.2 that identified a negative effect of education enrollment on autocratic breakdown. Several indicators have been used to measure state capacity in the literature, and here I use two relevant indicators from Hegre and Nygård (2014), that are imputed to mitigate selection problems stemming from missing values. First, I use an index of “bureaucratic quality” – measured using indicators of perceptions of the quality of public services and bureaucratic effectiveness. Second, I use an index measuring performance in economic policies, including factors such as economic management, social inclusion and economic freedom (see, Hegre and Nygård, 2014). These two aspects of state capacity are selected because they should be particularly likely to both lead to higher education levels and promote autocratic survival.

These variables are added to the models presented in Table 4.2, investigating the effect of education enrollment on autocratic breakdown. The results are presented in Table 9.21 in the appendix, where models 1-4 replicate the models from Table 4.2 on the sample for which state capacity data is available, and the two indicators of state capacity are added in models 5-8. In brief, although these models are a lot less efficient due to a lower number of observations, there is evidence that the negative coefficient estimate for education enrollment decreases in size when controlling for the two indicators of state capacity. For instance, in a model specification with key control variables and country fixed effects, the coefficient estimate for education enrollment decreases by 20 percent (from -2.5 in model 3 to 1.9 in model 7) when controlling for the two indicators of state capacity (see appendix Table 9.21)¹³ Hence, this adds some support to the argument that the negative effect of education enrollment on democratization and autocratic breakdown is driven by state capacity. Yet, this only seems to be part of the story, indicated by the fact the negative coefficient estimate for school enrollment only decreases moderately when controlling for the state capacity indicators. Other, additional explanations could be that education in itself is a source of popular content and a signal to the population of the government's dedication to its well-being. Or, that other aspects related to the regime's strength, that are not properly captured by indicators of state capacity, are both promoting higher education levels and regime survival. Moreover, as the results from these models are vulnerable to endogeneity problems, it can not be ruled out that

¹³In Table 9.22 I present results where the more general governance index from Hegre and Nygård (2014) is used to measure state capacity. When including this variable in models investigating the association between school enrollment and autocratic breakdown, the negative coefficient estimate for education enrollment is also reduced substantially.

parts of the identified negative association between education and regime breakdown is due to reverse causation.

4.8 Conclusion

This chapter has presented support for the proposition that education may help to destabilize authoritarian regimes and promote democratization, but not in the way commonly assumed by existing literature. First, I find little evidence that average school enrollment rates are positively related to regime change or authoritarian breakdown. Rather the evidence suggests that high education enrollments are negatively related to the likelihood of autocratic breakdown and changes in democracy. Exploring this finding further, I propose and find some suggestive evidence that this negative association is partly driven by state capacity, indicating that the high capacity of educated countries will also promote authoritarian survival. Second, I do find that when educated citizens participate in political protest campaigns, this promotes both autocratic breakdown as well as democratization in dictatorships. This comports with well known examples of stable dictatorships with highly educated populations, such as Singapore and Bahrain, as well as with prominent examples of (successful) anti-regime movements led by students and educated citizens. Hence, education may play a role in promoting authoritarian instability and regime change at the meso level, but education at the macro level does not. This evidence contrasts with modernization theory by defying expectations of education as a "dangerous good" that promotes authoritarian failure. It suggests that authoritarian regimes can safely expand education access without fearing subsequent threats to regime survival. Yet, if educated protesters take to streets, the dictator should be well advised to fear for his future. The next two chapters will delve further into the role of education in anti-regime protest and regime breakdown.

5 Educating Demonstrators: Education and Mass Protest in Africa

Sirianne Dahlum and Tore Wig¹

Abstract:

Does a more educated population spur regime-challenging mass protest? It is commonly argued that educated individuals are more likely to collectively challenge governments through protests, and that this may explain why education is associated with democratization. While many studies have investigated education's effect on conventional political participation (voting, petitioning etc.), we do not know whether education levels affect contentious mass protest. We argue that education increases the frequency of mass protest, by alleviating collective action problems and motivating mass opposition, particularly in autocracies. We investigate these links at the sub-national level in Africa, by mapping over 650 000 survey respondents to spatialized protest event data. We find that areas with more educated people have higher levels of protest activity. We proceed to pinpoint the causal effect of education by using the location of colonial-era christian missions to instrument for local education levels. We find mixed evidence consistent with both opportunity- and grievance-related mechanisms driving this relationship.

Publication status:

Forthcoming in *Journal of Conflict Resolution*

¹Postdoctoral research fellow at the Department of Political Science, University of Oslo

5.1 Introduction

The notion that education increases a population’s tendency to initiate mass uprisings against existing institutions is popular. University campuses have often served as locus points for protests and riots, and students and young graduates spearheaded the revolts against the crumbling Soviet empire in 1989, the Tiananmen Square uprisings in the same year, and the Arab Spring of 2011, to mention some notable examples. Meanwhile, there is evidence that educated individuals are more politically aware, more critical of government and more politically active (see, e.g., Dee, 2004; Hillygus, 2005; Solis, 2013). Based on this, a growing number of studies point to mass mobilization as the most plausible causal mechanism explaining why education is correlated with democratization (Glaeser, Ponzetto and Shleifer, 2007; Murtin and Wacziarg, 2014; Sanborn and Thyne, 2014). Yet, this education-protest link is rarely investigated systematically. While a comprehensive literature indicates a strong relationship between education and “conventional” political activities within democracies – such as voting, party membership and community meetings – few studies probe how education levels affect collective action aimed at transforming existing patterns of political authority (for exceptions, see e.g., Campante and Chor, 2012*b,a*). Such studies are essential to answer the question of whether education leads to more social instability, and, consequently, whether education is a force for regime change, as suggested by proponents of modernization theory (see e.g., Lipset, 1959; Glaeser, Ponzetto and Shleifer, 2007).

This paper provides new evidence that education can indeed increase the frequency of protests. It offers the first systematic analysis linking education to protest events, drawing on local-level evidence from 42 countries in Africa in the period 1990-2012. We present a simple framework for relating education to protest, assuming that protest emerges when people have both sufficient *motivation* to protest and the *opportunity* to do so. Education should affect protest motivation in two ways. First, educated citizens will arguably be more politically aware, knowledgeable and more likely to demand civil liberties, individual autonomy and democracy. Second, educated individuals are more critical of government and should be strongly motivated to protest when their education is not converted into jobs. Hence, they should overall be more motivated to raise their voice and protest, and especially against non-democratic, corrupt and poorly run governments. Education may also enhance protest *opportunities*, by serving as a resource that helps solve collective-action problems of potential protesters, by *enhancing skills, stimulating social networks* and creating *common understandings*. For these reasons, we expect education to raise the likelihood of protest overall, and to be particularly conducive to anti-government protest in (poorly governed) non-democracies.

To investigate these expectations, we create a dataset that captures fine-grained variation in education and protest, by mapping local education levels to protest event data. We rely on georeferenced data from the Demographic and Health Survey (DHS) project combined with georeferenced information on the location of regime-challenging protest events in Africa from the Social Conflict Analysis Dataset (SCAD) (Salehyan et al., 2012). This maps over 650 000 survey respondents to protest events, allowing us to investigate whether areas within countries with high education levels are more likely to experience protest events. Studying Africa has several advantages - in addition to the availability of fine-grained, georeferenced data that this study draws upon. A majority of the countries included in our sample are authoritarian or semi-democracies (according to the Polity index), and all have in the relevant period suffered from problems such as unemployment, inflation, corruption or poor governance.² Hence, we consider the countries in our sample to represent a particularly fertile ground for testing the proposed mechanisms tying education to protest, such as anti-government grievances. At the same time, the region exhibits enough variation in both democracy, governance and economic conditions to assess empirically whether the effect of education is conditioned on economic and political context. Moving beyond the country-year setup analyzed in previous contributions (Campante and Chor, 2012^{b,a}) also has several benefits: First, dis-aggregating this information plausibly brings us closer to the actual *events* and processes through which education affects collective action, reducing the danger of ecological fallacies. Second, it allows us to exploit variation in the sub-national spread of education that is plausibly exogenous to modern protest dynamics, such as the location of colonial-era christian missions.³

Empirically, we find that education generally does increase protest. This result holds across a wide range of specification choices, and also obtains when we endogenize education by using patterns in the settlements of Christian missionaries in colonial Africa to instrument for contemporary education levels, following (Nunn, 2010). In fact, the result is strengthened in models that account for endogeneity. To be able to probe the opportunity- and motivation-mechanisms discussed above, we also test some additional implications, utilizing information in SCAD on whether the protests were directed at the government or not. To evaluate the motivation mechanism, we first test whether education is particularly conducive to anti-regime protest when combined with more unfavorable economic opportunities, in line with Campante and Chor (2012^{b,a}). Second, we look at whether education is more strongly linked to anti-government protest in more authori-

²When using an average democracy-score for coding a democracy of > 5 on the Polity scale (27 anocracies/dictatorships and 15 democracies)

³An additional benefit is the fact that it allows us to control for omitted variables at the country-level without loosing to much variation in our education measures. Since education is a slowly-moving variable over time, this is often not feasible with country-level data.

tarian regimes. These tests are intended to distinguish between economic and ideological aspects of education-induced motivations for protest. We find no support for the first expectation, but strong and consistent support for the second, leading us to conclude that education levels are particularly protest-inducing in dictatorships. This suggests that education could affect protest through enhancing preferences for democracy and more accountable governance. To gauge the opportunity-mechanism, we investigate whether education also affects protest events that should not be motivated by education-induced preferences, captured by looking at protests that are *not aimed at the government*. In this analysis, we also find a link between education and protest, suggesting that education also impacts on capacities to stage protests, holding preferences constant.

We contribute to the literature by being the first study to demonstrate that education is linked to a higher likelihood of actual protest events, and that this effect is particularly strong in more authoritarian settings, suggesting that education may spur protest through anti-authoritarian grievances. We utilize an approach that allows us to study (aggregated) protest events and at the same time benefit from a more fine-grained analysis than a country-level study would permit. The study offers new insights on the causal nature of the education-protest link, by leveraging (plausibly) exogenous spatial variation. Finally, it adds to a growing literature on the determinants of sub-national political instability (e.g., Hendrix and Salehyan, 2012; Pierskalla and Hollenbach, 2013; Garcia and Wimpy, 2016) by introducing an important new predictor of local protest.

5.2 Education, participation and protest

It is commonly held that educated citizens are more politically active. For instance, Dahl (1971) conjectured that socioeconomic development creates an educated populace likely to nurture the kind of participatory culture and political awareness necessary to sustain democracy. The claim that education is related to more political participation in democracies is supported by several studies, and has been referred to as “one of the most thoroughly-explored and best-established facts in the social sciences” (Campante and Chor, 2012b), and “one of the major contributions of political science to the general body of knowledge” (Solis, 2013). In particular, numerous studies report that educated individuals are more likely to vote (see e.g., Wolfinger, 1980; Dee, 2004; Freeman, 2004). Relatedly, there is a strong association between education and individual political knowledge and democratic awareness (see e.g., Almond and Verba, 1963; Verba and Nie, 1987; Hanushek, 2002; Bratton, Mattes and Gyimah-Boadi, 2005), interest in obtaining political information (Dee, 2004), party membership (see e.g., El-Said and Rauch, 2015), and greater dissatisfaction with existing institutions (Weakliem, 2002).

While most of the above studies indicate a positive association between education and participation, they mainly draw on purely correlational evidence that does not support causal interpretations (without highly implausible assumptions). Against this background, several authors have utilized inference techniques such as quasi-experiments and instrumental variable regression in order to more properly assess whether education causes political participation, yielding mixed evidence (see e.g. Solis, 2013; Tenn, 2007; Kam and Palmer, 2008; Persson, 2012). While a number of studies – chiefly drawing on evidence from the US – identify causal effects (Sondheimer and Green, 2010; Milligan, Moretti and Oreopoulos, 2004), others do not (e.g., Solis, 2013; Friedman et al., 2011). Moreover, probing this question in a non-democratic setting, Croke et al. (2016) find no effects of education on political participation in Zimbabwe.

While one might surmise from this literature that there is suggestive – yet mixed – evidence for a positive effect of education on political participation, all the studies mentioned above deal with “normal” political participation, aiming to channel demands and preferences *within the framework of established political institutions*, mainly in western democracies. The literature is much less comprehensive when it comes to the question of how education affects regime-challenging political action, or “contentious politics”; defined by Tarrow (1994, 7) as “used by people who lack regular access to representative institutions, who act in the name of new or unaccepted claims, and who behave in ways that fundamentally challenge others or authorities”.

A prominent literature, associated with modernization theory, suggests that an educated population is more likely to induce political change, especially change towards more democracy and political liberalization (Lipset, 1959; Glaeser, Ponzetto and Shleifer, 2007). Yet, if education is to create a citizenry that will ultimately oppose and push through changes in existing political regimes, it must do more than just strengthen the extent to which people take part in political activities *within* the framework of current institutions. Indeed, one of the central mechanisms proposed to explain the education-democracy relationship in influential studies on democratization is a suggested effect of education on *mass protest* (e.g., Glaeser, Ponzetto and Shleifer, 2007); a quintessentially unconventional form of participation. While an individual who is likely to oppose a dictatorship would perhaps also be likely to vote and participate in a functioning democracy, this does not mean that participation in democracies generalizes to dictatorships: Taking part in anti-government protest is much more costly than political participation in conventional democratic channels. While casting a vote in a democracy comes with the mere cost of traveling to the polling station, protests against a dictator may lead to prison, strict sanctions or even death. While this is especially the case in repressive regimes, even protesting in (semi-)democracies can be fraught with peril, as many protesters face teargas, jail and

police brutality.

There are very few empirical investigations of the link between education and “contentious” protest. A couple of studies have looked at education and various forms of protest within *single countries*. For example, Miller (1977) finds that schooling levels could predict individual participation in the U.S. urban social unrest in the 1960s and 70s, and that the relationship is most accurately captured by Davies (1962) “J-curve theory”, positing that “rapid expansions in opportunities build up expectations which, if unmet, can lead to frustration and violence”.⁴ Studying Kenya, Friedman et al. (2011) finds that education does *not* make individuals more likely to challenge authority. When it comes to cross-country studies, a few contributions investigate the link between education and self-reported participation in protest at the *individual level*. Campante and Chor (2012b,a) argue and demonstrate empirically that education combined with economic grievances is likely to spur protest. For instance, they show that individuals whose income falls below the predicted value based on their achieved education are more likely to report to have participated in protest, and that education coupled with unemployment is associated with higher levels of reported protest, in line with the “J-curve” tested in Miller (1977).

While studies of *self-reported* protest activities yield valuable insights; they do not tell us whether education increases *levels* of protest as such. It could for example be that education raises individual-level inclinations to join already existing protests and demonstrations, while not affecting the coordination-processes that are necessary for protests to get underway in the first place. Hence, using individual-level findings to draw expectations about overall protest incidence is perilous. In light of this, the lack of studies investigating the link between education and actual *protest events* represents a considerable gap. Furthermore, with the exception of Friedman et al. (2011), the above-mentioned studies pay little attention to causal inferences, which means that we can not rule out that the correlations between education levels and reported protest activities are driven by endogeneity problems. Our study sets out to address these issues, through investigating the effects of education on regime-challenging protests at the sub-national level in Africa and developing a research design that should identify causal effects and allow for distinguishing between potential mechanisms.

In addition to informing the debate relating to the general relationship between education and regime-challenging contentious collective action, our study also illuminates ongoing work on the impacts of education on democracy, collective action and political participation in Africa specifically (e.g., Albaugh, 2014; Mattes and Mughogho, 2009;

⁴In a different context, Krueger and Maleckova (2003) found that education predicts participation in Hezbollah activities in Lebanon.

Bleck, 2015).

5.3 Linking education to popular protests

What are the mechanisms by which education should generate protest? To structure this discussion, we draw on a distinction from the literature between two determinants of successful collective action: *opportunity* and *motivation* (see e.g. Kuran, 1991; Tarrow, 1994). For protest activities to materialize, potential dissenters must first be motivated to oppose a regime, for instance due to frustrations with lack of democratic freedoms or economic deprivation. But, as Kuran (1991, p.21) notes: “Anti-government feelings can certainly bring a revolution within the realm of possibility, but other conditions must come together to set it off”. Prospective dissidents must gain the opportunity to organize, coordinate and carry out opposition such as demonstrations or general strikes. Hence, a minimum of both motivation and opportunity are necessary conditions for collective action to materialize, but the two can to a certain extent compensate for each other. That is, collective action can materialize despite slender opportunities as long as motivations are sufficiently strong, and people may choose to protest even based on moderate grievances if conditions are sufficiently favorable. In the following we relate education to both motivations and opportunities for contentious collective action. In brief, we argue that education should motivate people to protest, especially in dictatorships and if economic expectations are unmet. Moreover, education will probably improve *capacities* for collective action - through boosting skills, social networks, and coordination abilities.

5.3.1 Motivations

Education should make people more motivated to engage in regime-challenging protest – at least when the political regime and its policy performance is considered sub-optimal. We here specify two branches of the causal channel tying education to protest through motivations; an *ideological branch* that involves citizens becoming more challenging of existing institutions due to changing preferences for more accountability and better governance and an *economic branch* where educated citizens protest only when education is unmet by economic opportunities.

First, education should alter preferences for political institutions. Schooling yields greater knowledge of political affairs, increased political awareness and a more critical view on government performance. This suggests that educated individuals should be increasingly eager to influence government decision-making, and raise their voice if the government does not satisfy expectations in terms of e.g. public policy provisions or

macro-economic performance. Relatedly, it has been argued that education makes people more endorsing of democracy and more likely to emphasize civil liberties, individual autonomy and citizen empowerment (Welzel, 2013). This builds on a proposed psychological mechanism according to which citizens who become used to exercising critical thinking in schools, universities and in the work-place, start demanding the right to influence political affairs as well (Inglehart and Welzel, 2005). This is pertinent in the context of Africa, where education is often found to make citizens more demanding towards political institutions. This is for example documented in cross-national surveys (Mattes and Mughogho, 2009) and case-based work, such as Bleck (2015)'s study of education and political engagement in Mali. Based on such altered preferences, educated individuals should generally be more likely to express political grievances through protest activities. This should strongly apply in the African context. As Bleck (2015, 3), studying Mali, notes "in environments of fragile or imperfect institutions, educated, empowered citizens might be more critical of regimes than their uneducated peers are". In line with this, we expect education to be particularly protest-conducive in political regimes that fail to live up to standards of democracy, civil liberties and well-functioning institutions, given that education induces sympathies with such ideals. This implies that more educated citizens in authoritarian regimes and semi-democracies should be especially likely to express their political grievances through mass protest.

Second, education should affect economic grievances. For example, increased political awareness and a more critical approach to politics may be especially protest-inducing if combined with economic hardship and unemployment. Conversely, if combined with a healthy economy, education is likely to get converted into better jobs and better material conditions for the educated. Schooling is, after all, also a public good that creates satisfaction and other material benefits for its recipients. Even if education generally makes people more critical of the government, it may create incentives in favor of status quo that trumps such dissenting views. Hence, the protest-enhancing effect may be less pronounced for educated people who have converted their acquired human capital into employment and material benefits.

5.3.2 Opportunities

The window of opportunity to form regime-challenging movements is a function of two factors: The capacity of potential protesters to organize, and the capacity of the regime to prevent protest (see e.g Gurr, 1970). There are some reasons to suppose that education could strengthen the regime, since educated labor may boost economic productivity and government revenue, increasing state capacity. However, most arguments point in favor of education strengthening the capacity of protesters, thereby tipping the balance in favor

of the protesters' side of the equation. Crucially, education can enhance the capacity of potential protesters, since successful mobilization requires human capital and overcoming collective action problems. Below, we briefly elaborate on this mechanism.

The inherent difficulty of organizing collective action has been described in depth (see e.g Olson, 1965). In brief, a collective action problem occurs when a certain outcome that would be desirable for most members of society could be reached through coordinated collective action, but when individual participation is stifled by prohibitively high costs (Olson, 1965). When deciding whether to support an opposition movement or not, an individual will consider the potential benefits she would receive if the movement succeeded, relative to the costs if the movement failed, and benefits under the status quo.⁵ At the same time, she needs to consider the likelihood that the movement will succeed. Even if an individual will benefit greatly from a successful protest outcome such as regime change, she might still refrain from protest participation, either due to the participation costs or due to doubts that the movement will succeed - a source of even higher costs.⁶

Against this background, protest success is to a large extent a function of the number of dissidents who join: the more people, the more likely a successful outcome. In such settings, people are usually faced with a severe information deficit regarding other people's intentions or preferences and lack the trust that others will participate (e.g. Kuran, 1995). Hence, individuals may end up choosing to refrain from protest based on fear that others will also do so, implying that collective action commonly fails to materialize.

There are several ways in which education can mediate collective action problems. First, education should improve *communication skills*, which are essential for collective action. For instance, well-developed reading and writing skills are crucial for circulating critical information about the government, to spread calls for protest, and to utilize the internet and social media for collective action purposes. More generally, enhanced literacy should improve abilities to effectively communicate and articulate opinions on complex policy issues.

Second, schooling should enhance *cooperation skills* in important ways. Through making students participate in collective learning tasks, as well as joint activities such as games, school plays, research projects, and general socializing with classmates, schools teach pupils how to successfully interact with others (Newton, 1997). An important side product of such activities is a greater understanding and appreciation of other points of

⁵The costs of participating in collective movements may become particularly high in non-democratic regimes, as individual participation will be strongly discouraged through high punishments for political dissent if the popular movement fails to challenge an authoritarian regime

⁶The prospects of collective action is made even more difficult by free-rider problems, since it is generally in a person's interest to let others make the sacrifices required to secure regime downfall. Popular protest can be seen as a "public good" that individuals can benefit from without contributing themselves.

view. Hence, class environments foster toleration of diverging opinions, as well as effective communication of one's own. These are qualities that are vital to overcoming collective action problems among potential anti-regime protestors.

Third, education should create and strengthen *social networks*. The role of such networks in overcoming collective-action problems has been extensively discussed (e.g Goldstone, 1994; Kuran, 1991). Networks strengthen group-mentality and make people more likely to trust that others will also participate in collective action. Indeed, there is evidence that the size of one's network affects the probability of engaging in political activity (McClurg, 2003), and this is probably also applicable to regime-challenging protest. Interestingly, since schools bring people together that might not otherwise have interacted, education should be expected to enhance the kinds of "weak ties" that are particularly conducive to collective action. As argued by (Granovetter, 1973, 1983), weak ties connect previously unconnected parts of a network and create relationships between "friends of friends". As Granovetter highlights, these ties will often be conducive to collective action, since they create network integration rather than fragmentation. Larger social networks, composed of weak ties, have smaller information costs: When people interact more frequently with other network members, they will get access to more information about the intentions and behavior of others, easing collective action. As Siegel (2009, 12) writes, "information exchange allows people to update their beliefs about the costs and the benefits inherent in participation, and so change their decisions".

Relatedly, education should increase work differentiation and specialization, immersing citizens in more integrated and specialized networks. These dynamics create patterns of interactions that can replace and complement family and kin as the dominant social units. This kind of expansive non-familial interaction could be crucial to social mobilization, since strong family ties have been identified as an obstacle to political participation and civic engagement (Alesina and Giuliano, 2011). In sum, education should create a more networked population, which in turn can stimulate collective action.

Fourth, education should enhance inter-group *solidarity*. People who have attended the same schooling or university system will share at least a minimal common framework of reference, easing communication and coordination. According to Tarrow (1994, p.11): "Education provides us with a common knowledge of a set of dates, places, names, events, concepts, references and quotations that help the social interaction of otherwise disparate individuals". He shows how collective action problems can be solved when leaders or key agents are able to "tap into and expand deep-rooted feelings of solidarity or identity" (Tarrow, 1994, p.11), arguing that the crucial ingredients for a successful mass movement are mobilization around a collective challenge, with a common purpose that is recognized by participants through a sense of social solidarity.

The communication and cooperation skills, enhanced networks, and solidaristic communities described above as stimulated by education, will be familiar to scholars of of *social capital* (see e.g. Putnam, 1993) - argued to “improve the efficiency of society by facilitating co-ordinated actions” (Putnam, 1993, 176). In short, we contend that the social capital generated by education could help dissidents overcome collective action problems and enable them to effectively challenge authority.

At this point, we should note that education *in Africa* - the empirical context we investigate - exhibits great variety. For example, the continent hosts a wide variety of religious schools (such as Madrassas), schools run by development actors, as well as the more standard public and privately run schools adhering to baseline curricular standards. Education systems also vary greatly, as illustrated by the differences between Francophone and Anglophone countries when it comes to supporting multilingual versus unilingual language policies in education (see e.g., Albaugh, 2014). Despite this heterogeneity in schooling in Africa, we still think investigating the impact of education *as such* on protest in Africa is warranted. While school types in Africa will undoubtedly differ slightly in content, there is no strong theoretical reason to suspect that the theoretical mechanisms outlined above should apply strongly to some but not to others. Furthermore, while our data does not allow us to distinguish empirically between forms of education (such as public vs. private) it does allow us to investigate heterogeneity across country-level *education systems* such as is found between Francophone and Anglophone countries (discussed in Albaugh (2014)). We will therefore return to the question of heterogeneity in education systems in our empirical analysis.

5.3.3 Expectations

Several expectations are derived from the above discussion. Most importantly, we contend that most of the arguments point to education being on net conducive to mass protest. Education should enhance the type of critical sentiments and self-expression values that should motivate protest overall, at least in the African context, where there are many sources of grievances such as poor governance and lack of democracy. When we combine these arguments with the purported effect of education in boosting the protest-capacities of potential dissidents, we get the following expectation about the unconditional effect of education on protest: *Increases in education should increase the risk of protest overall.*

The discussion above illuminates two causal channels through which education should affect protest; a motivational channel and a capacity channel. We will here present some conditional expectations attempting to tease these different channels apart. First, the motivational channel, which can be broken down to an ideology-component and an economic-component, is concerned with how education motivates individuals to oppose

and protest *against the government*. Hence, this mechanism yields several expectations regarding the effect of education on anti-government protest. The ideological component of the motivational channel implies that education generates the type of liberal democratic sentiments that stimulate protest against dictatorships. For this reason, we expect an interaction effect between education and regime-type on protest, whereby *education should have a stronger positive effect on anti-regime protests in more autocratic regimes*. Additionally, we discussed the economic component of the motivational channel, whereby education should interact with economic opportunities, such that individuals with high education but poor job prospects should be more likely to protest against the government. This yields the second conditional expectation, namely that *education should be more conducive to anti-government protest when combined with high unemployment*.

Finally, we discussed the opportunity channel which stipulates that education should also increase capacities for protest, holding preferences constant. To test whether this operates, we perform tests where we circumscribe our sample to look at protests that are not directed at governments. If the opportunity mechanism operates, we would expect education to increase non-government protests also, through affecting protest-opportunities generally. Hence, we expect that: *education should increase protests not aimed at the government*. We now proceed to test each of these expectations.

5.4 Data

To investigate the expectations presented above, we need fine grained data on both sides of the education-protest equation. To this end, we combine georeferenced survey data from the Demographic and Health Surveys (DHS) with georeferenced data on protests in Africa taken from the Social Conflict in Africa Dataset (SCAD) (Salehyan et al., 2012).⁷ Africa is particularly ideal for our purposes. The continent has experienced a surge in education over the past three decades, and protests and riots are quite frequent, occurring in different regime contexts.

Our data on education comes from the DHS survey, and is at a high level of spatial resolution. DHS provides GPS coordinates for each survey “cluster” where a given respondent was surveyed, and by aggregating survey responses in these areas we get a local picture of education levels in the areas where clusters are located.⁸ The DHS covers 42 countries in Africa. The countries in this sample are mostly either semi-democracies or authoritarian, and most of them are either poorly governed and suffer from (relatively)

⁷The DHS data are available at <http://www.dhsprogram.com/Data/>

⁸The DHS surveys are household surveys, often covering more than one respondent in a household. In these cases we aggregate values for all respondents in a household to survey clusters.

poor economic performance. The average democracy score in the sample (measured using the Polity index) is 1.78, indicating that the average country in our sample is a semi-democracy. This suggests that there are many potential sources of grievances, that may interact with education-induced resources and greater political awareness to induce mass protest. At the same time, there is enough variation in terms of both degree of democracy, quality of governance and economic conditions to actually be able to assess whether the grievance mechanisms are at play, by testing whether the effect of education on protest is conditioned on the political and economic context. Hence, Africa is a particularly vital context for investigating how education impacts on political protests.

Our data on protest events, from SCAD, is based on news reports, and covers all countries in Africa in the 1990-2012 period. A highly beneficial property of the SCAD dataset is that it contains geocodes for each protest event, allowing us to map protest events to the mentioned survey clusters. To do this, we need a common spatial frame of reference that can give us units of analysis comprising information on protest events, and the education levels in the areas where these events took place. For such a template, we use the PRIO-GRID spatial template (TollefSEN, Strand and Buhaug, 2012), version 2.0, which partitions our map of Africa into 50X50KM grid-cells. The PRIO-GRID is now a standard template for sub-national analyses of spatial data in conflict research, commonly used by scholars doing disaggregated conflict analysis (e.g., Pierskalla and Hollenbach, 2013; Wischnath and Buhaug, 2014). The grid-cell units are completely independent of country-borders and other administrative units.

To construct our dataset, we proceed as follows. First, we aggregate respondent information (from households) to the level of the survey-clusters, and remove all clusters with no GPS coordinates, leaving us with over 650000 respondents situated in roughly 26000 clusters in the period 1990-2012. These clusters are then merged into the grid cells that encompass them, leaving us with just above 6800 grid cells containing survey clusters with education data and respondents. The DHS contains data for every year in the 1990-2012 period covered by the SCAD data, but only some grid-cells are surveyed in more than one year. This essentially gives us an unbalanced panel of grid-cells with survey and conflict data for the 1990-2012 period. Figure 5.1 shows a map of Africa with all the grid cells in our dataset and SCAD events overlaid.

In our interpretations we treat the grid-cell values as more or less representative for the population in the area that the given grid-cell covers. In relation to this, one could argue that the information contained in this dataset is compromised by the aggregation of survey-respondents to grid-cells. There are about 98 survey respondents on average in each grid-cell, with a minimum of 1 and a maximum of 2972. Most cells have acceptable numbers of respondents. The median number of survey respondents is 47 and more than

80% of cells have more than 29 respondents. Nevertheless, some of our units-of-analysis will contain too few respondents, and will thus not be representative.⁹ We believe this should generate attenuation bias, pulling our estimates towards zero. This makes our tests slightly more conservative and should bias *against* finding a relationship between education and protest. However, to further probe this, we conduct robustness checks dropping all cells with less than 30 respondents, with no qualitative change in results.

To measure education we take the grid-cell level mean of the main education item in the DHS surveys. At the respondent level, the item takes the values; 0 =no education, 1 =incomplete primary education, 2 =completed primary education, 3 =incomplete secondary education, 4 =completed secondary education, 5 =higher education. Taking the grid-cell mean of this survey item, which we use in our analysis, we get a measure that is normally distributed, with an empirical mean of 1.1, and an empirical minimum and maximum of 0 and 4.41. The countries with the lowest scores on this education variable are Algeria (.07), Somalia (.15) and Niger (.24), while the countries with the highest scores are Swaziland (2.46), South Africa (2.45) and Botswana (2.37).

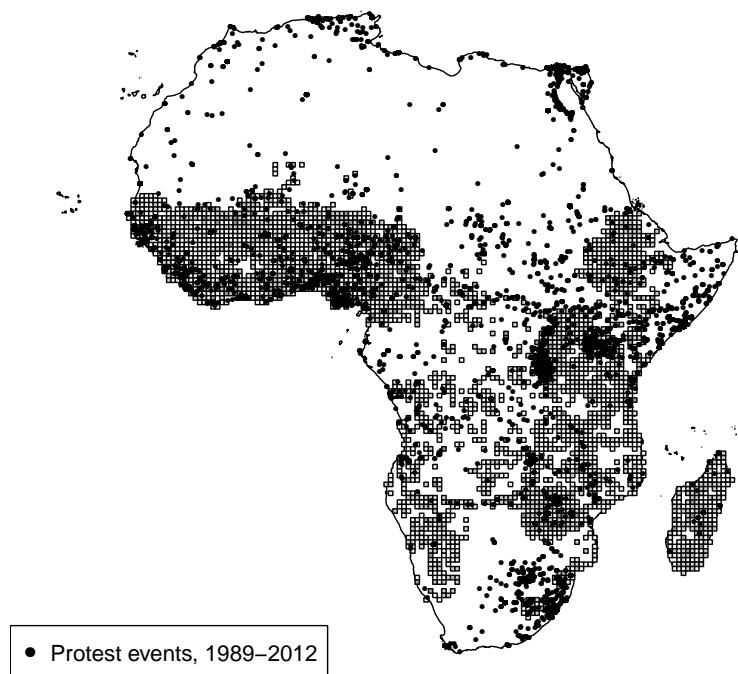
As our measure of protest events, we create a measure based on the Social conflict in Africa (SCAD) data (Salehyan et al., 2012). This counts the number of “Organized” and “unorganized” demonstrations and riots, defined as: “Distinct, continuous”, and “directed toward members of a distinct “other” group or government authorities.” These include peaceful anti-regime demonstrations, as well as violent riots. It also includes protest events that are part of general or limited strikes. In the SCAD data, each such event includes a brief description of the event. Typical examples of events that we focus on are pro-democracy rallies, student demonstrations, strikes among public employees and so on. Our protest measure counts the absolute number of protest events in a grid-cell in a 3-year time-window after each grid-cell has a survey.¹⁰ In our baseline model we include all demonstrations and riots, but the data also details whether protest was directed towards the government or not, and we also present results from models distinguishing between the two. This allows us to tease out whether the anti-government grievance mechanism is at play, by investigating whether education impacts on anti-government protest as well as assessing whether this effect is condition on democratic deficit and government performance.

In our sample where we have DHS information on education, 11% of grid-cells experience at least one protest event, while the average frequency of protest events for all

⁹DHS clusters are randomly sampled within regions of a country, and respondents are randomly sampled within clusters.

¹⁰While our results are robust to using a 1-year window, we use the 3-year time-window to get more information on conflict events in our dataset. Since most grid-cells are measured in the DHS only once, or with more than 3-year intervals, and since education levels are expected to be relatively static from year to year, using a 3-year window is a sensible solution.

Figure 5.1: Grid cells with DHS education data and protest events overlaid, 1990-2012



grid-cells (where we have DHS data) is .07. Of the countries in our dataset, in the 1991-2012 period, the countries with the most protest events are Nigeria (327 events), South Africa (168) and Kenya (106), while the least protest-prone (as measured by SCAD) countries in our sample in the relevant period are Eritrea (8 events), Lesotho (5) and Gambia (4).

5.4.1 Covariates

In our baseline models, we include a sparse set of potential confounders, while including a large list of additional controls in robustness tests. As controls, we include the log distance to the capital, since the capital expectedly has higher education levels and probably also more protests because of the location of central national symbolic spaces, government buildings, important infrastructure and similar focal points. We also include the average travel time to the nearest city in the grid cell, using data from the PRIO-GRID

database, capturing the development of infrastructure and urbanization in the grid-cell. Furthermore, we include a list of mean-aggregated items from the DHS surveys; the mean age among respondents in the grid-cell, the mean number of respondents in the grid-cell living in an urban area (as defined by DHS), and the average unemployment level in the grid cell. It should be noted that the latter covariate is potentially a bad control, since it could induce post-treatment bias (i.e. be an intermediary variable between education and protest, since education can reduce unemployment). We therefore display models both with and without this post-treatment variable. Since protest events cluster in space, a spatial lag could be a candidate for the baseline model, but we have opted to include it only in additional tests; both because it is also potentially a post-treatment variable, and because it is insignificant when conditioning on observables and its inclusion does not alter our results. A number of other variables are plausible candidates for confounders, and will be discussed in the robustness section.

5.5 Empirical strategy

In our main analysis, we estimate a set of probit models capturing the probability of a grid-cell experiencing at least one protest event in a given year, along with OLS linear models capturing the log of the count of protest events in a grid cell.¹¹ These models are estimated with and without controls, and country- and year dummies.

After this exercise, we probe whether our results can support a causal interpretation, by comparing the OLS and the probit models to Two Stage Least Squares (2SLS) and Instrumental Variable probit (IVprobit) models for the same outcomes. This is done to address the (quite plausible) eventuality that education is endogenous; either due to reverse causality or to unobserved confounders. Reverse causality would occur if observed or latent protests affect education levels. This will be the case if conflict or political instability drives out educated people indirectly, through affecting the economy, or directly; if educated are more resourceful and thus more able to migrate in the face of conflict. Unobserved confounders are plentiful. It could, for example, be that certain areas are not serviced with public goods such as education precisely because they are hostile to the government. If this is the case, then we would see a downwardly biased correlation between education and protests. One could also imagine the opposite pattern; that places with high levels of protest received higher levels of education as a concession to stave off

¹¹An arguably more appropriate choice of models would be negative binomial count models, which would be more fitting for our count outcome. However, these models are infeasible for a number of reasons. First, they are computationally intensive, and the models do not converge when estimated on our data (with the full set of controls). Second, there is no easy-to-use two-stage estimator for count models, meaning that we can not use these to estimate causal effects.

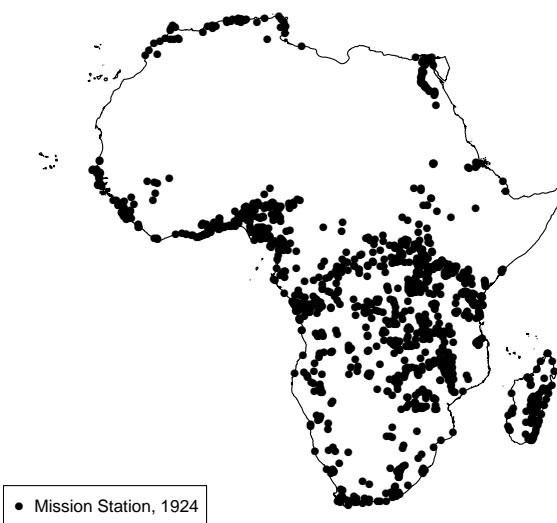
further unrest.

In short, we have good reasons to suspect endogeneity, and need a strategy for addressing this issue. To do this, we employ models that endogenize education by using the historical presence of Christian missions as an instrument for contemporary education levels. The logic behind using the historical mission-station instrument is that Christian missionaries provided education to the localities they settled in, and, given the persistence of education levels through social reproduction and other channels of persistence, increased contemporary education levels. This builds on recent contributions showing that the presence of Christian missionaries at the local level in Africa substantially affected historical education levels locally (Nunn, 2010), whereby Christian missionaries established local schools that boosted the level of education in localities close to missions. Building on a large historical qualitative literature documenting the activities of colonial missionaries, Nunn (2010) finds that mission stations increase local levels of civic participation, and persuasively demonstrates that this effect is due to increases in education.

To create the historical mission-stations instrument, we use the map created by Roome (1925), and digitized in Nunn (2010), containing coordinates of all the Protestant and Catholic missions and bible depots in Africa in 1924. We combine this information with the DHS clusters that we have georeferenced in Grid-cells, and count the number of historical Christian mission stations in a grid-cell, taking the log of the number of historical Christian mission in a cell as our instrument for contemporary grid-cell level education today. Around 20% of the cells we have survey-clusters for has at least one Christian mission in 1924. Figure 5.2 shows the digitized map of mission stations in colonial Africa. The map shows that there is substantial spatial variation in the location of mission stations, making this a useful variable in our subnational research design. The exclusion restriction assumed in models using this instrument is that the presence of missions is unrelated to protests when conditioning on observables, and the IV-strategy requires that missions strongly predict higher education levels in our grid-cells (the expectation yielded by the argument linking missionaries to the spread of education). We return to these considerations in our discussion of the IV models.

In our baseline estimates and in the instrumental variable models, we use a protest measure including both anti-government protest and protest aimed at other targets, but in section 5.2 we also distinguish between these two forms of protest, in order to evaluate the proposed grievance and opportunity- mechanisms.

Figure 5.2: The Location of mission stations in colonial Africa (Based on data from Nunn, 2011)



5.5.1 Baseline estimates

First, we estimate a panel of probit and OLS models, where we assume that education levels are exogenous conditional on the observables. Table 5.1 presents results from the probit models (models 1.1-1.4) and OLS estimates (models 1.5-1.8).

Model 1.1 is most parsimonious, including only lagged protest and education level. As expected, the coefficient is positive (.252) and precisely estimated (Z-value= 10.12). This result is strengthened when country- and year-fixed effects are added (model 1.2), and remains positive and strong in the models with additional grid-cell level controls (1.3 and 1.4), also when the unemployment control (potentially post-treatment, as discussed above) is included. Based on a simulation of model 1.4, depicted in Figure 5.3 below, we find that when increasing average education levels from no education to completed secondary education, protest probability in a given year rises from roughly 2% to 8%. This is a *substantial* increase, and suggests that education could be a powerful driver of unrest. Education is a powerful predictor of protest also when compared to other variables tapping crucial aspects of socioeconomic development. Indeed, as we discuss in the appendix, it compares favorably to variables such as unemployment, literacy, urbanization and the wealth-index provided in the DHS surveys. This indicates that education is a powerful driver of mass protest activity in Africa.

Table 5.1: Baseline probit and OLS models of protest events

	Probit (1.1)	Probit (1.2)	Probit (1.3)	Probit (1.4)	OLS (1.5)	OLS (1.6)	OLS (1.7)	OLS (1.8)
Education	0.252*** (10.12)	0.469*** (8.24)	0.206*** (3.50)	0.225*** (3.76)	0.142*** (6.27)	0.246*** (7.68)	0.124*** (3.86)	0.078** (3.02)
Protest _{t-1}	0.628*** (7.44)	0.730*** (7.19)	0.648*** (6.33)	0.643*** (6.05)	1.160*** (8.03)	1.001*** (6.69)	0.952*** (6.43)	1.065*** (7.19)
Age			0.034 (0.11)	0.176 (0.45)			-0.045 (-0.52)	-0.032 (-0.35)
Age ²			-0.001 (-0.23)	-0.004 (-0.57)			0.001 (0.41)	0.000 (0.21)
Urban			0.591*** (5.87)	0.562*** (5.40)			0.327*** (5.12)	0.382*** (6.01)
L(Capital distance)			-0.169*** (-3.62)	-0.175*** (-3.58)			-0.109*** (-3.52)	-0.111*** (-3.85)
Time to city			-0.001** (-3.02)	-0.001*** (-3.54)			-0.000* (-2.42)	-0.000*** (-4.56)
Unemployment			0.150 (0.80)					0.100* (2.13)
β	-1.553***	-2.405***	-1.134	-3.003	-4.544***	-4.369***	-3.023*	-2.812
Country-FE	NO	YES	YES	YES	NO	YES	NO	YES
Year-FE	NO	YES	YES	YES	NO	YES	NO	YES
N	6699.000	6097.000	6097.000	5719.000	6272.000	6272.000	6272.000	5905.000
Log likelihood	-2247.764	-1031.161	-980.606	-896.914				
R ²					0.067	0.112	0.126	0.111

Notes: *p<0.05; **p<0.01; ***p<0.001, standard errors are clustered at the level of the grid-cell.
 Z-scores (probit) and T-values (OLS) in parentheses.

A similar pattern obtains when we estimate the OLS models of the log number of protest events, in models 1.5-1.8. All the education coefficients are positively signed, and precisely estimated, with T-values ranging from 7.68 (model 1.6) to 3.02 (model 1.8). These results indicate that protests are more common in highly educated areas, even when possible confounders such as urbanity, travel time to nearest city, age composition, capital distance and unemployment, as well as unobserved country- and year-level confounders are included.

As noted above, the assumption that education is exogenous is heroic. Hence, to investigate whether the patterns identified in Table 5.1 warrant a causal interpretation, we estimate a series of IV-probit (of the binary outcome) and 2SLS models (for the log-linear outcome) where we endogenize education levels in grid-cells using the mission instrument. This comparison will take us some way towards a causal interpretation of the estimated associations.

Table 5.2 conducts this exercise, where education is instrumented using the presence of historical Christian mission stations (described above). In this set of models we control

Table 5.2: IV-probit and 2SLS models of protest events

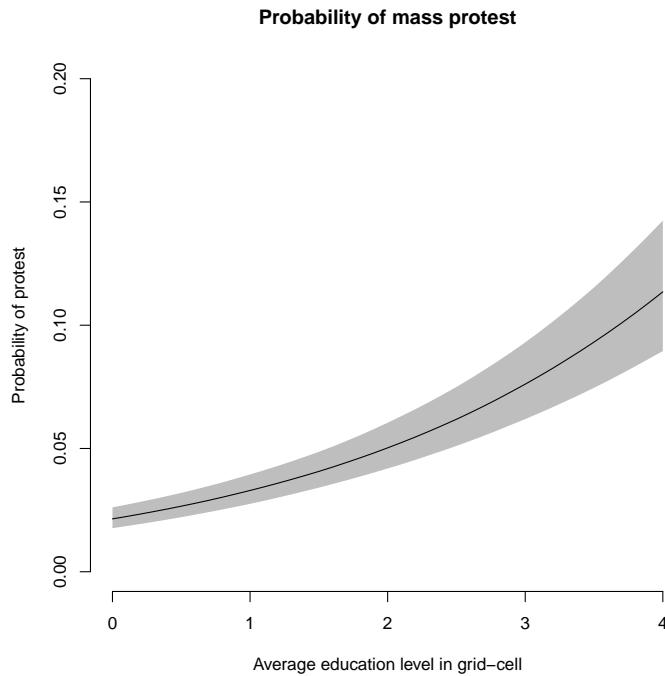
	IV-Probit (2.1)	IV-Probit (2.2)	IV-Probit (2.3)	IV-Probit (2.4)	2SLS (2.5)	2SLS (2.6)	2SLS (2.7)	2SLS (2.8)
Education	0.430*** (3.73)	1.103*** (6.81)	0.251 (1.19)	0.900** (2.69)	0.695*** (5.38)	0.703*** (4.92)	0.655** (3.18)	0.738** (2.86)
Christians	-0.139 (-0.81)	-1.103*** (-4.46)	0.109 (0.37)	-0.805* (-2.01)	-0.904*** (-5.22)	-0.664*** (-3.66)	-0.862** (-3.13)	-0.715* (-2.49)
Protest _{t-1}	0.487*** (4.97)	0.444*** (3.69)	0.491*** (5.12)	0.477*** (3.73)	0.949*** (6.26)	0.842*** (5.62)	0.945*** (6.30)	0.846*** (5.66)
Age			-0.339* (-2.40)	0.006 (0.02)			-0.012 (-0.10)	-0.025 (-0.22)
Age ²			0.005* (2.10)	-0.000 (-0.05)			0.000 (0.09)	0.001 (0.27)
Urban			0.095 (0.50)	-0.021 (-0.06)			-0.094 (-0.53)	-0.162 (-0.77)
L(Capital distance)				-0.086** (-2.84)	-0.076 (-1.07)		-0.097*** (-3.36)	-0.038 (-0.88)
Travel time				-0.000 (-1.48)	-0.001 (-1.91)		-0.000 (-0.34)	-0.000 (-0.63)
β	-1.592*** (-32.43)	-2.336*** (-6.14)	4.515* (2.16)	-1.782 (-0.44)	-4.639*** (-93.28)	-5.189*** (-17.00)	-3.862* (-2.07)	-4.619* (-2.52)
Country-FE	NO	YES	NO	YES	NO	YES	NO	YES
Year-FE	NO	YES	NO	YES	NO	YES	NO	YES
N	5747.000	5157.000	5747.000	5157.000	5320.000	5320.000	5320.000	5320.000
ll	-7968.774	-4973.652	-7355.571	-4167.407				
R ²					.	0.068	0.003	0.063

Notes: *p<0.05; **p<0.01; ***p<0.001, standard errors are clustered at the level of the grid-cell. Z (probit) and T (OLS) values in parentheses.

for the presence of (current day) Christians in each specification, since this is the most obvious threat to the exclusion restriction (we report several additional tests in the appendix, where we further probe the exclusion restriction). This is because the historical presence of Christian missionaries should affect religiosity in these areas today, through the missionaries' spread of their faith. Furthermore, since Christian areas might see more protests (especially where Christians are national minorities, such as in Sierra Leone, Burkina Faso, and Mauritius), this could be a threat to the exclusion restriction.

Columns 2.1-2.4 in Table 5.2 display the IV-probit models. These models show substantively similar results to the probit models reported in Table 5.1. The coefficients are positive and substantive, ranging from 1.103 (model 2.2) to .251 (model 2.3). With the exception of model 2.3, all coefficients are precisely estimated, with fairly large Z-values. The pattern is even stronger for the 2SLS models in columns 2.5-2.8, showing very precise coefficients across all specifications. For most IVprobit and 2SLS models, the education coefficients are stronger than their counterparts in Table 5.1, with no first-stage. This is the pattern we would expect if there was a negative endogeneity bias, which would

Figure 5.3: Simulated probability of protest events



happen if, for example, governments provide less education to protest-prone areas. Such a relationship is certainly plausible. IV-diagnostics show that in each model, there is a need for a first-stage, since education shows signs of endogeneity (all Wu-Hausmann p-values are below .05). Furthermore, the mission-stations instrument is clearly relevant, with F-values well above values that are usually considered critical (never dropping below $F = 50$) (see e.g., Stock and Yogo, 2005*b*).

It is important to stress the limitations of these IV models. While they buy us some added internal validity, this comes at the price of a loss in generalizability. This is because the treatment effects estimated are local average treatment effects (LATEs) that are less generalizable than the associations estimated in the OLS models. Crucially, the coefficients for the LATEs we estimate describes the effect of specifically “missionary-induced” education, i.e. the effect of education on protest in areas where education levels are predicted by the presence of mission stations. This might significantly reduce the scope for generalization, especially if this local treatment effect is inherently different from the global average treatment effect of education. While there is no specific reason to suspect this, it might very well be the case. In spite of this caveat, these results warrant greater confidence in claiming that education has a causal effect on increasing protest incidence.

5.5.2 Exploring mechanisms

As a second step, we probe the specific mechanisms proposed above, starting with the two branches of the motivational mechanism; ideological and economic grievances. First, we test an implication of the ideological grievances explanation; that education should generate protest more strongly in dictatorships and especially protests that are anti-government. To test this, we first circumscribe our dependent variable such that it only registers protests aimed at the government. Then we introduce an interaction term to our baseline models, between education and a country-level variable measuring democracy, using the POLITY IV standard index of democracy (Marshall, Gurr and Jagers, 2013). Since these interactions are all hard to implement in the IV models, we enter them into our baseline specifications with no first stage, estimating both probit and OLS versions. Second, we investigate whether the effect of education is particularly salient where the educated have fewer economic opportunities, to capture the “unmet-expectations” mechanism proposed by Campante and Chor (2012b) and Campante and Chor (2012a). To model this, we create an individual-level variable registering the individual level multiplicative term between education (0-3) and whether the respondent is unemployed (1=unemployed, 0 otherwise). We then aggregate this term to the grid-cell and enter this into our baseline, along with each constitutive term. If the economic-opportunities aspect of the motivational mechanism is operating, we would expect to see more protest in areas with high scores on this variable.

Third, we investigate the opportunity-channel. We do this by circumscribing our dependent variable to exclude all protests that are aimed at (*any*) level of government, and include those that are aimed at some other entity or group. This will often be other ethnic groups, international companies or other actors that are not specifically tied to the government. We believe this probes the opportunity-channel since this mechanism would imply that education would increase *all kinds of protest* holding the level of anti-government grievances constant, since education simply makes it easier for people to organize protests. Hence, if we find that education also increases non-governmental protests, then this is at least suggestive evidence that this effect is about more than just motivations.

To identify protests that are specifically aimed at the government (and those that are not), we use the SCAD variables registering whether the government (central or regional) was the target of the given protest event, and separate between protest events (as operationalized above) that are aimed at either central or regional governments and those aimed at some other entity or group.

Table 5.3 shows the results of this exercise. Models 3.1-3.4 shows our baseline (Probit

Table 5.3: Models exploring mechanisms

	(3.1) (Anti-gov.) (Probit)	(3.2) (Anti-gov.) (OLS)	(3.3) (Anti-gov.) (Probit)	(3.4) (Anti-gov.) (OLS)	(3.5) (Non-gov.) (Probit)	(3.6) (Non-gov.) (OLS)
Education	0.516*** (2.65)	0.318*** (3.40)	0.405**** (3.81)	0.095*** (3.04)	0.224*** (3.07)	0.089**** (3.38)
Education·Democracy	-0.024* (-1.71)	-0.020*** (-3.09)				
Democracy	0.029 (1.10)	0.022*** (2.69)				
Education·Unemployment			-0.459** (-2.14)	-0.098 (-1.44)		
Unemployment	-0.248 (-0.95)	-0.037 (-0.63)	0.403 (1.37)	0.091 (1.52)	0.342 (1.42)	0.040 (0.94)
Protest _{t-1}	0.587**** (5.23)	0.627**** (4.29)	0.599**** (5.55)	0.655**** (4.52)	0.329*** (2.81)	0.330*** (2.92)
Age	0.748* (1.74)	0.027 (0.58)	0.698 (1.59)	0.003 (0.07)	-0.042 (-0.10)	-0.019 (-0.21)
Age ²	-0.014* (-1.86)	-0.001 (-0.77)	-0.013* (-1.72)	-0.000 (-0.28)	0.000 (0.02)	0.000 (0.16)
Urban	0.576**** (4.42)	0.227**** (4.48)	0.578**** (4.43)	0.249**** (4.82)	0.383*** (2.92)	0.079* (1.85)
L(capital distance)	-0.114** (-2.08)	-0.062** (-2.46)	-0.126** (-2.36)	-0.071*** (-2.78)	-0.236**** (-3.50)	-0.056** (-2.29)
Time to nearest city	-0.001** (-2.29)	-0.000* (-1.80)	-0.001*** (-2.94)	-0.000** (-2.45)	-0.001** (-2.30)	-0.000 (-1.36)
β	-11.582* (-1.89)	-4.589**** (-6.15)	-10.606* (-1.70)	-3.909**** (-5.41)	2.783 (0.44)	-4.034*** (-2.89)
Country-FE	YES	YES	YES	YES	YES	YES
Year-FE	YES	YES	YES	YES	YES	YES
N	5310	5547	5668	5905	5223	5905

Notes: *p<0.05; **p<0.01; ***p<0.001, standard errors are clustered at the level of the grid-cell. Z (probit) and T (OLS) values in parentheses.

and OLS) models estimated on anti-government protests rather than all protests, since we expect anti-government protests to be particularly strongly affected by the motivational set of mechanisms. Models 3.1-3.2 investigates whether the education-protest link is stronger in non-democracies, which is implied by the ideological component of the motivational channel. This indeed seems to be the case, as the interaction term is significant and negative in both the probit (3.1) and OLS (3.2) models. Models 3.3 and 3.4 turn to the question of whether areas with high levels of education *and* unemployment are more likely to experience education-induced protest, in line with the arguments highlighted above. These models indicate patterns that go against this hypothesis. The probit model (3.3) actually displays an interaction term which is negative (i.e areas with high-education-high-unemployment are less likely to experience protest than contrasting areas) and weakly significant, while the OLS models yield a similarly signed coefficient,

but with low T-values. We believe this combination of results provides some, although no more than suggestive, support for the argument that education leads to protest through affecting preferences for political institutions rather than economic aspirations.

Finally, models 3.5-3.6 evaluate whether there is (at least suggestive) evidence that the opportunity mechanism is operating, by assessing whether education also affects protest aimed at some other entity than (local or national) government. Here, we find a slightly weaker coefficient for education for this outcome, although it is both positively signed and precisely estimated in both models. This suggests that the capacity/opportunity mechanism is probably also driving the relationship between education levels and protest incidence.

5.5.3 Robustness and additional tests

Our key results – both the democracy-interactions and our unconditional result – are retained across a wide variety of specification changes. Most are detailed in the appendix, but we will highlight some of them here.

First, we investigate whether our results remain stable when introducing a number of plausible additional controls, that can both be seen as tapping potential pre-treatment confounders, but that also represent potential threats to the exclusion restriction. For instance, we investigate coefficient stability across a number of controls representing alternative ways to measure economic development. We include measures of the average infant mortality in the grid-cell, child malnutrition, and estimated local GDP (all taken from the PRIO-GRID database). We also include a measure of light-emission density at night as measured by satellites, taken from PRIO-GRID. Satellite-measured light density has been found to be a strong proxy for local economic development (Weidmann and Schutte, 2016), and has been used in several recent studies of political conflict (e.g. Cederman, Weidmann and Bormann, 2015). We also run models where we include controls for local child malnutrition and local infant mortality rates, taken from the PRIO-GRID database.

Second, we control for two factors that are known to have been affected by the presence of Christian missions. This includes a control for consumption of media and communication technology in the grid-cell, using a simple additive index registering the share of respondents in a grid cell that have a telephone, a TV set, a radio, and newspapers. This is partly because the presence of Christian missionaries have also been claimed to affect the propagation of early print media (Cagé and Rueda, 2016), which should impact on media consumption patterns and (potentially) protest patterns today. We also factor in a control for average literacy in the grid-cell, since Christian missions (and their printing presses) affected the spread of literacy, through books and reading and through education,

and it has been conjectured that reading and literacy should affect societal propensities for violence and political instability (Pinker, 2011).

Third, we control for a number of additional proxies for urban development, since urbanization is a particularly salient alternative explanation that could account for our results; education is higher in urban areas, and protests should be much more likely in cities. We therefore include a measure of the share of grid-cell land covered by urban area, and a measure registering the log of population in the grid-cell, relying on data from PRIO-GRID.

Fourth, we include a number of controls for different types of geographic land cover, since these could both represent omitted variables and potential threats to the exclusion restriction; due to the non-random nature of mission-station placement. In this respect, we include variables for the share of the grid-cell that is covered by forest, mountains and agricultural land respectively, all found in the PRIO-GRID dataset.

Additionally, we also present tests probing causal heterogeneity, such as seeing if we get similar results when we split the sample into Francophone vs. non-Francophone countries, yielding no qualitative change in the results.

The results from these specifications are shown in the appendix. Our main results are stable to these additional controls (for further discussion, see appendix). In addition to these tests, we control for a number of additional potential controls, and estimate models with different functional forms. We also estimate models including a spatial lag, as well as models dropping grid-cells with low numbers of respondents. All of these tests show that education is robustly linked to protest incidence.

5.6 Conclusion

The spread of education is one of the most consequential social developments in the modern era. Yet, the effects of education on regime-challenging protest have been curiously missing from the literature looking at the political effects of education and the literature on political instability in developing countries more generally. This paper adds to the discussion surrounding this topic by presenting a geographically focused, but still cautiously generalizable, study of this link. In a sample of over 6000 grid-cells spread across 42 African countries in the period 1991-2012, we find that local education levels are positively associated with protest events. This result holds across a wide range of specifications, and it is also retained when we endogenize education by using the variety in missionary settlements in colonial Africa as an instrument for contemporary education levels.

We outline two potential channels that can account for a link between education and

protest. A motivational channel, where education induces protest-motivations through changing ideological preferences for democracy and economic aspirations, and an opportunity channel, where education increases the capacity of people to protest through solving collective action problems. While the empirical analysis has been mainly focused on the unconditional link between education and protest incidence, results from additional tests yield some suggestive evidence on each of these mechanisms. First, we find that education is particularly conflict-inducing when it comes to anti-government protest in autocracies, and that education combined with high unemployment is not particularly conducive to protest. This suggests that while education might increase motivations for protest it primarily does so through making populations increasingly critical of political institutions. Furthermore, we find that education also impacts on protests not aimed at governments, suggesting that education might also increase protest incidence through enhancing capacities for protesting as such.

Our overall interpretation of these results is that education increases both the capacity and motivations for regime-challenging collective action, and that these dynamics are particularly strong in authoritarian settings. Capacities are enhanced through the strengthening of social networks, creating common understandings, and through enhancing the skills and organizational capacities needed to organize collective action against a regime. Motivations (for regime change) should be strengthened by education in dictatorships, since education has been shown to stimulate democratic sentiments. Our results shed systematic light on the conventional wisdom that education in dictatorships can generate uprisings among the young, educated and disaffected.

Our study makes several novel contributions to the existing literature. Ours is the first study to use disaggregated event data to look at the link between education and protest. Second, we present a novel way of instrumenting for education in the context of regime-challenging collective action. Finally, ours is the first study to investigate the link put forward in Glaeser, Ponzetto and Shleifer (2007), where education is claimed to lead to democratization through stimulating mass uprisings. We show that in our sample of African localities, education does indeed have this effect, and more strongly so for anti-government protest in dictatorships.

This corresponds with a number of famous historical examples from African political history. For example, in Zimbabwe, students from higher education institutions have been one of the most vocal groups in opposing the Zanu-PF regime (see, e.g., Zeilig, 2008). In South Africa, students and the educated were vital in the anti-apartheid movement, and similar examples proliferate (Nkinyangi, 1991; Meredith, 2006). These examples, combined with the systematic evidence uncovered above, provide a somewhat cautionary tale for governments wanting to expand education while not improving other governance

outcomes and institutions that citizens care about. For example, our results suggest that if African autocrats continue to expand education to their populations, they might find themselves more frequently challenged by riots and protests than they would like.

6 Students in the streets: Education and nonviolent conflict

Sirianne Dahlum

Abstract:

This study investigates whether protest campaigns consisting of students and educated protesters are more likely to 1) choose nonviolent methods and 2) successfully reach their goals. Extant literature suggests that education reduces the prospects of violent conflict, and the commonly assumed mechanism is that educated groups are less likely to resort to violence. Moreover, many argue that education is a force for regime change, by creating successful protest campaigns. This paper is the first to systematically test implications of these mechanisms at the protest level. The empirical analysis builds on original data on the educational background of participants in all protest campaigns aiming for regime change from 1900 to 2006 identified in the NAVCO 1.0 dataset. I find very robust evidence that protest movements with a high degree of involvement by students and graduates are more likely to turn nonviolent. Moreover, there is some (although weaker) evidence that these movements are more likely to achieve their goals, but *only due to* their nonviolent dispositions. This adds to the literature explaining why some movements resort to nonviolence (and succeed), by establishing that the identity and socio-economic background of protesters matters.

Publication status:

Revise and resubmit in Comparative Political Studies

6.1 Introduction

Students and university graduates have been at the forefront of protests and opposition movements across the world, pushing through government concessions and even overthrowing political regimes. For instance, university students played a prominent role in protests against Milosevic in Yugoslavia in 2000, in the 2003 Georgian “rose revolution” and in protests against the Suharto regime in Indonesia in 1998. Meanwhile, university graduates, including professors, lawyers and doctors were crucial, for instance, in the 1986 protests against Marcos in the Philippines and in the 1986 Chilean pro-democracy campaign. These examples comport with the commonly held view that an educated population poses a threat to authoritarian regimes. Yet, we have no *systematic* knowledge about so-called “student protests”, and we do not know if movements consisting of educated and student protesters stand out, e.g. when it comes to choice of strategies or effectiveness. While recent literature has yielded several insights on protest campaigns and their outcomes (e.g., Chenoweth and Stephan, 2011), including why some movements turn nonviolent and successful, we do not know whether the social background of protesters matters.

The literature offers two clues to the nature of protest campaigns consisting of educated and student protesters. First, studies of domestic conflict suggest that education has a pacifying effect, as educated countries are associated with fewer instances of civil war (see e.g. Collier and Hoeffler, 2004; Thyne, 2006). The implied mechanism is usually that educated groups or populations are less likely to take up arms. Second, modernization theory postulates that education induces transitions from authoritarian regimes to democracy, supported by several empirical studies (see e.g. Glaeser, Ponzetto and Shleifer, 2007; Sanborn and Thyne, 2014; Murtin and Wacziarg, 2014). An implicit and untested assumption behind the democratization-education link is the notion that an educated population is more capable of organizing efficient mass resistance. Applying these findings and proposed mechanisms to protest movements, one should expect campaigns consisting of students and university graduates to be more likely to 1) choose nonviolent methods and 2) accomplish their aims.

This study is the first to test these propositions, by compiling original data on the educational background of participants in all protest movements aiming for regime change included in the Nonviolent and Violent Campaigns and Outcomes Dataset 1.1 (NAVCO) (Chenoweth and Stephan, 2011). This departs from the aggregated approach of previous studies investigating whether peace or regime change are more likely in educated countries. By offering macro-level tests of mechanism that usually operate at the meso-level, such as educated groups being prone to instigate successful regime change and less likely

to participate in civil war, these studies are vulnerable to ecological fallacies. Meanwhile, there is a burgeoning literature seeking to explain protest campaigns and their outcome, especially why some movements turn nonviolent and successful (see e.g. Chenoweth and Stephan, 2011; Celestino and Gleditsch, 2013; Cunningham, 2013b). Although a couple of recent studies have investigated structural or socio-economic explanations at the country-level to explain protest outcomes (see e.g. Butcher and Svensson, 2014; Chenoweth and Ulfelder, 2017), no studies have yet looked at the socio-economic correlates and background of actual protesters.

The compiled data contains information on whether and to what extent students and university graduates were involved in the 203 protest movements aiming for regime change between 1900 and 2006 from the NAVCO 1.0 dataset. For each campaign, I have coded whether students or university graduates at a minimum *participated* in the movement, whether they *dominated* the movement and whether the movement *originated* among these groups. Utilizing this information, I conduct tests showing that protest campaigns consisting of students and university graduates are indeed more likely to resort to non-violent methods. This is a very robust finding, that holds across numerous sensitivity tests. Furthermore, I find some evidence, although weaker, that these movements are more likely to succeed in terms of achieving their main goals. This is only, however, *due to their pacifist inclinations*, and there is no evidence that campaigns with educated and student protesters are more successful when accounting for nonviolence.

I suggest two mechanisms explaining why educated protest campaigns should be more likely to adopt nonviolence: First, the *preferences-effect* suggests that educated people will have developed intrinsic preferences for nonviolent rather than violent protest. Second, the *capacity-enhancing effect* suggests that educated people have the capacities and resources facilitating the type of mobilization necessary to create peaceful (and successful) movements, and hence will be more likely to choose nonviolent opposition.

The finding in this paper adds to previous literature in several ways. First, it adds to the burgeoning literature on political protest campaigns and their outcome (see e.g. Chenoweth and Stephan, 2011), by being the first study to show that the identity and socio-economic background of protest participants matter for whether protest movements turn nonviolent. Second, it contributes to the literature on education and violent conflict. It corroborates a plausible mechanism implicit in the argument that education breeds peace, by indicating that groups of students or educated individuals are more likely to choose nonviolent resistance. Finally, it adds to the study of democratization by suggesting *why* it is that education is correlated with regime change (and democratization) (see e.g. Murtin and Wacziarg, 2014): By promoting nonviolent rather than violent campaigns, thereby potentially raising the probability of succeeding at promoting regime

change – in line with literature showing that nonviolent resistance is more successful

The paper proceeds as follows. Section 6.2 discusses the literature that motivates this study and section 6.3 lays out the theoretical considerations specifying how and why educated protest movements should be expected to turn nonviolent (and thereby more likely to succeed). The research design and data is discussed in section 6.4, with an emphasis on the coding of data on the educational background of protesters. Results from the main analysis are presented in section 6.5, indicating that protest campaigns with educated members have a higher likelihood of nonviolence (and a higher success rate, due to this). Section 6.6 contains an extended discussion of whether these findings are driven by my proposed mechanisms, suggesting a causal effect of educated protesters on nonviolence, or rather, for instance, by nonviolent campaigns recruiting educated members. Various tests suggest the former, implying that educated protest members promote a peaceful outcome, and that this again helps to ensure success.

6.2 Education, violence and collective action

6.2.1 Violent vs. nonviolent mobilization

Protest campaigns crucially vary according to whether they (primarily) employ peaceful or violent forms of resistance. There is a growing research agenda investigating the nature of political protest movements and their outcomes (see Ackerman and DuVall, 2000; Schock, 2005; Chenoweth and Stephan, 2011; Nepstad, 2011; Chenoweth and Cunningham, 2013; Butcher and Svensson, 2014; Chenoweth and Ulfelder, 2017). These studies find that nonviolent campaigns are more likely to succeed in terms of reaching their goals (Chenoweth and Stephan, 2011), and thus more likely to affect outcomes such as transitions from autocracy to democracy (Celestino and Gleditsch, 2013). Hence, whether protest movements turn violent or not is an important concern in itself, given the amount of human suffering produced by violent clashes, at the same time as nonviolent resistance have other beneficial consequences.

When it comes to explaining the adoption of nonviolent methods, most studies consider agent-centered factors such as the strategies adopted by protesters (Cunningham, 2013*b*), the interaction between protesters and the target (in most cases the government) (Chenoweth and Stephan, 2011) and the competition between different factions of the protest movements (Cunningham, 2013*a*). We know less about the structural factors that facilitate nonviolent methods. Two exceptions are Chenoweth and Ulfelder (2017), using aggregate state-level indicators of economic development to explain nonviolent campaign onsets, and Butcher and Svensson (2014) investigating whether nonviolent campaigns

can be explained by variations in manufacturing as a share of the economy.¹ Both of these examine *structural factors at the state-level*, and no studies have looked at the socio-economic correlates of *protest movements and their members*. Dahl et al. (2013) advocate for the need to link the choices of agents to their resources and constraints, by studying group characteristics and profiles, but relies empirically on country-level proxies such as size of the urban population, GDP and population size. This study fills this gap, by studying how the socio-economic correlates of protesters matter for strategic choices and outcomes, with a focus on education.²

While no studies look specifically at the educational background of groups and their conflict behavior, many have studied the education- armed conflict link at the country level.³ As noted in a recent review article, “there seems to be an emerging consensus in the literature that education has a general pacifying effect on conflict” (Østby and Urdal, 2011). In the perhaps most comprehensive analysis of education and domestic conflict, Thyne (2006) finds that various education indicators such as higher levels of primary enrollment and literacy levels are associated with a lower risk of civil conflict. This finding is corroborated by several studies showing that countries with high education levels have less armed conflict (see e.g. Barakat and Urdal, 2009; Collier, Hoeffer and Söderbom, 2004).

Regarding causal mechanisms, Thyne (2006) suggests that education pacifies by reducing anti-government grievances and offering people tools to resolve disputes peacefully. Education also plays a prominent part in Pinker’s (2011) account of why violence is declining globally. He argues that education and literacy enable people to interact with and access the ideas of a wide range of fellow citizens, and that this stimulates empathy, moral concerns and a reduced willingness to accept that violence and suffering is inflicted on other people.

In short, there is evidence that an educated population is inversely related to civil war, and this should lead us to expect that *those that engage in violence are uneducated*. Yet, this can not be inferred from highly aggregated studies at the country level. This is the first paper to test this link at the group level, thereby investigating empirically a key assumption in the literature on education and political violence.

¹Butcher and Svensson (2014) also include education as a control, and the results indicate that education significantly raises the likelihood of nonviolent campaign onsets (but not violent campaign onsets)

²Hence, this study links structural conditions to the agents of protest.

³The civil war literature also offers studies at the local (district-) level. These studies may bring us somewhat closer to the mechanism, but do not tell us whether those participating in violence have education or not.

6.2.2 Education, collective action and regime change

The second literature motivating this study is concerned with the relationship between socio-economic development and democratization. The strong correlation between education and democracy levels is well established (see e.g. Lipset, 1959; Benavot, 1996), and has led many to assume that educated societies are more likely to democratize. This argument is in line with modernization theory, positing that socio-economic development, and particularly education, is conducive to the emergence and survival of democracies (Lipset, 1959). Numerous studies investigate this proposition empirically, and many report a positive effect of education on democratic transitions (Papaioannou and Siourounis, 2008; Glaeser, Ponzetto and Shleifer, 2007; Sanborn and Thyne, 2014; Campante and Chor, 2012b; Murtin and Wacziarg, 2014). Yet, it has been questioned if this is really due to an effect of education on democratization, or rather, for instance, due to democracies investing in more education (see e.g. Acemoglu et al., 2005; Acemoglu, Gallego and Robinson, 2014).⁴

Few have gone into depth to specify the causal mechanism tying education to democracy. Those that do, usually consider education to promote successful collective action, due to the skills and resources obtained by educated individuals. For instance, Glaeser, Ponzetto and Shleifer (2007) argue that education will induce more people to participate in collective action, thereby boosting the prospects of successful mass mobilization. Yet, the proposed education-collective action link has rarely been tested systematically at the country-level (but, see Campante and Chor, 2012a), and there is, to my knowledge, no examinations of whether educated protest movements are more likely to successfully promote regime change.⁵ Moreover, it is rarely specified *why* we should expect educated movements to succeed at inducing regime change. I contribute by showing that education may induce successful collective action *because* it promotes peaceful rather than violent resistance.

6.3 How education affects political protest

The dynamic and outcome of a political protest campaign should depend on the resources and preferences of the protesters. This section outlines two mechanisms predicting that educated protest movements should stand out from other campaigns: First, the *capacity-mechanism*, implying that educated protesters have the capacities – and are in the position – to put pressure on the regime without resorting to violence. Second, the

⁴But, see also Bobba and Coviello (2007) and Castelló-Climent (2008).

⁵A number of studies investigate whether educated individuals are more likely to *take up* protest, but none of these examine protest outcomes (see e.g. Friedman et al. 2011)

preferences-mechanism, suggesting that educated protesters will have intrinsic preferences for nonviolent rather than violent opposition,

6.3.1 Education and capacity for nonviolent protest

Nonviolent opposition is a costly form of resistance, that requires certain skills and resources. In their study of nonviolent resistance, Chenoweth and Stephan (2011) argue that nonviolent campaigns in fact are more costly than violent campaigns. The former require a wide repertoire of contentious actions, a large organizational capacity and the knowledge and strategies necessary to impose costs on the regime without the use of violence. I argue in this section that campaigns with educated participants have a higher potential for pressuring the regime in a nonviolent manner. Due to this comparative advantage, they should be more likely to choose a nonviolent tactic. I discuss two modes of non-violent resistance – noncooperation and symbolic protest – and show how they both have requirements that match the resources and skills of educated protesters, making it a more feasible option for educated protest campaigns.

Noncooperation refers to people refusing to carry out their regular tasks or continue cooperation with the regime (Sharp, 2005; Nepstad, 2011). The aim of noncooperation is to put pressure on the ruler by disrupting the existing system. This can happen through economic noncooperation, including actions such as economic boycotts, strikes or withdrawal of bank deposits. It can also happen through social or political noncooperation, such as election boycotts, civil disobedience, reluctant and slow compliance, deliberate inefficiency by enforcement agents, student strikes, boycotts or stay-at-home (Sharp, 2005). Resistance through noncooperation is an attractive method if the state depends on the networks that a protest campaign builds upon, as this dependence makes the regime vulnerable to even peaceful forms of resistance (Nepstad, 2011; Butcher and Svensson, 2014). Educated groups, including professionals such as doctors, lawyers, teachers and university professors, possess generic knowledge and skills that are indispensable for the functioning of the economy. Hence, the state can not simply disregard these groups and replace them with people lacking the necessary skills. For instance, strikes or boycotts by professionals and educated groups could be very disruptive both to the economy and to the functioning of society, disruptions that could produce even more discontent and political opposition. Moreover, educated individuals have leverage in the sense that they often hold key roles in government agencies, the bureaucracy and the judicial system. For instance, if the bureaucracy simply refuses to carry out or deliberately delay instructions, the ruler is deprived of his implementing capacity. For this same reason, groups of educated people are also more likely to be part of the regime's support base - or the "winning coalition" (Bueno de Mesquita and Smith, 2003), as people who are indispensable for the

regime are more likely co-opted. Hence, the regime is especially vulnerable to the acts of opposition of educated individuals, especially when it is prolonged, and even nonviolent resistance will impose costs on the regime.⁶

Nonviolent *symbolic protest* refers to attempts to undermine the regime's legitimacy through expressing disapproval and mobilizing widespread discontent. This approach includes many different methods such as demonstrations, petitions, singing, leafleting or protest meetings (Sharp, 2005; Nepstad, 2011). The rationale behind this form of protest is that the ruler depends on legitimacy and popular obedience. If people start challenging the power of the dictator and manage to win the hearts and minds of large portions of the population, this undermines the ruler's legitimacy and thereby its power (Sharp, 2005). Waging such a "symbolic war" against the regime is very demanding because it requires shaping public opinion, and persuading key agents that the regime lacks legitimacy, including for instance the military, foreign actors such as governments or international organizations and the regime itself. Education, especially at the tertiary level, is a resource that should enhance people's capacity for organizing and coordinating this form of nonviolent resistance. I call this the *capacity-enhancing effect* of education.

First, educated individuals usually have more advanced language and communication skills, and this is a great advantage when articulating the movements' message and communicating with fellow protesters. Educated individuals are also more likely to use social and mass media, and more competent when it comes to utilizing it in an efficient manner. The 1986 pro-democracy campaign in Chile illustrates the advantage of communication and media competence. As part of this campaign, a group of media professionals worked voluntarily to develop a very successful TV advertisement campaign that arguably helped to persuade many Chileans (Puryear, 1994). Foreign language abilities also encourage collaboration and mobilization directed towards foreign governments, non-governmental organizations and like-minded opposition movements abroad. For instance, the student movement that organized the Tiananmen square protests in Beijing in 1989 had been built up under close supervision from student members of Poland's Solidarity movement (Cunningham, 2014). Moreover, the students and intellectuals that helped to organize the color revolutions had strong transnational ties to other color revolution campaigns (see, e.g. Wilson, 2005).

Second, educated individuals should be more knowledgeable of politics, society and complex social mechanisms, and better at gathering new information. Hence, they should be better equipped to develop a suitable and well-informed campaign strategy. For instance, as part of the Chilean anti-Pinochet movement, a group of social scientists were

⁶Moreover, these factors also suggest that the costs of violent repression of educated protesters are higher.

responsible for developing the campaign strategy prior to a plebiscite. This group utilized focus groups and surveys to assess how the campaign message should be developed to appeal to and mobilize as many Chileans as possible (Puryear, 1994). Students and educated protesters will also likely be familiar with effective protest methods and strategies from having read about earlier experiences of protest in their home-country or abroad. For instance, educated protesters of the color revolutions were reportedly influenced by detailed readings of political scientist Gene Sharp, offering guidance for planning symbolic protest and a list of suggested protest activities. This includes the student organization Pora – which was crucial in the Ukrainian orange revolution – that utilized Sharp’s readings to develop a creative nonviolent campaign strategy. Among other things, they used entertainment and music to mobilize and keep up the spirit. When soldiers surrounded the central square in Kiev, the Pora protesters handed out flower and played music as a symbolic move. One of the Pora leaders have described the Orange revolution as partly an ”information war” against the government (see e.g. Wilson, 2005).

This is not to say that violent protest does not require resources and capacities, but it requires other skills than nonviolent protest, such as e.g. battle training, physical force and military strategy. These are not areas of strength for students and educated protesters. Of course, other prominent social groups, such as e.g. industrial workers, may also have some qualities that encourage nonviolent resistance. For instance, workers may benefit from economic leverage through the possibility of labor strikes, and this may encourage large-scale, organized non-violent resistance. However, educated groups have a much wider repertoire both when it comes to conducting symbolic resistance and non-cooperation. Their generic skills and formal positions suggests that they can disrupt the system in a much more fundamental way than other groups, at the same time as they should be supreme when it comes to available resources and tools for undermining regime legitimacy, such as inventing powerful symbolic moves and persuasive communication strategies. Due to this comparative advantage, educated protesters should be more likely to choose nonviolent opposition than e.g. farmers or unskilled workers.

6.3.2 Education and preferences for nonviolent protest

The capacity-mechanism discussed above considers how the attractiveness of nonviolent resistance depends on the resources and positions of (potential) protesters. This section discusses how the choice between nonviolent and violent resistance can also be affected by protesters’ ideology or moral considerations. I argue that educated individuals are more likely to be ideologically or morally averse to the use of violence. Due to this, educated protest campaigns may choose to stick to nonviolent resistance even in instances where violence would have been a more feasible option given the strength and capacity of the

campaign vis-a-vis the target.

Humans generally have a strong disposition to want to prevent harm and suffering from being inflicted on family and kin. Yet, this sense of empathy does not necessarily apply to groups of people that a person does not identify with, such as other ethnic groups, political opponents or other social classes (see, e.g. Tajfel and Turner, 1979; Abrams and Hogg, 2006). Indeed, a major share of political or organized violence has been conducted against groups that in some ways are considered an opponent or out-group by the perpetrators (e.g., Gat, 2008). Yet, education should reduce willingness to accept that suffering and death is inflicted on other people through violence, by increasing empathy for people outside of one's closest circle.

First, educated people are generally more knowledgeable about other people, cultures or societies, and this generally reduces ignorance and intolerance (see, e.g. Galston, 2001; Dee, 2004). The knowledge may emerge from the education in it itself, or due to educated individuals reading more books. Reading and learning about the experiences of strangers arguably creates a stronger sense of identification with other people, and stimulates empathy with their sufferings. For instance, Pinker (2011) describes how the flourishing of books and reading after the introduction of mass schooling opened up a new world of insights into other people's thoughts and lives. He goes on to argue that these insights were crucial to the subsequent reductions in violent behavior. By being able to imagine other people's pain, one grows less inclined to accept that violence is inflicted upon humans outside of ones closest circle. In this sense, education should increase levels of empathy.

Second, education should strengthen individuals' ability to carry out abstract moral reasoning (Kohlberg, 1976; Kurtines, Gewirtz and Lamb, 2014), and this may discourage use of violence. Abstract moral reasoning stimulates "the ability to detach oneself from parochial knowledge of one's own little world and explore the implications postulated in purely hypothetical worlds" (Pinker, 2011, p.791). This includes the ability to reason about moral issues from a neutral standpoint and adopt the perspectives of other people. In this sense, education may help to expand the "circle of moral considerations", by strengthening the extent to which moral principles are applied to humans in general and not only family and kin. Violence and suffering inflicted on others is than no longer considered acceptable from a moral standpoint. Abstract reasoning also makes people able to rationalize violence as a "Prisoner's dilemma", inferring that everyone would benefit from abstaining from violence, rather than taking up violence as a short-term preemptive solution to avoid that violence is inflicted upon oneself.

Inherent in this ability to apply empathy and moral principles to strangers is the ability to execute self-restraint and avoid spontaneous decisions based solely on emotions. Tertiary education requires self-discipline and patience, and nurtures the ability to refrain

from jumping the gun for short-term gains. If nonviolent protest leads to temporary failure in the face of repressive regimes, persistence and willpower is required to continue sticking to nonviolence. And, self-restraint is needed to overcome the impulse to use violence when faced with provocative or unjust actions being carried out by a perceived enemy. Education should also make protesters more informed of the strategic benefits of nonviolent opposition – in line with studies showing that nonviolent strategies are associated with a higher success rate (e.g. Chenoweth and Stephan, 2011).

The notion that education promotes a political culture of tolerance and empathy conducive to nonviolent behavior, resonates well with classic arguments such as that presented by Lipset (1959). He asserts that education creates a “culture of peace” because it “presumably broadens men’s outlook, enables them to understand the needs for norms of tolerance, restraining them from adhering to extremist and monistic doctrines”. It is also in line with individual-level studies establishing that educated individual are more tolerant, trustful and less willing to accept violence (see e.g. Golebiowska, 1995; Inglehart and Welzel, 2005; Inglehart, Puranen and Welzel, 2015), or Shayo’s (2007) finding that educated individuals are less likely to hold “militaristic attitudes”. There is also evidence that educated individuals are less likely to *participate in violence*. For instance, Humphreys and Weinstein (2008) finds that educated individuals in Sierra Leone are less likely to join insurgencies, and Oyefusi (2008) finds that educated youth are hesitant to participate in forms of civil unrest in the Niger Delta.⁷

These considerations suggest that campaigns led by, or made up of, students and educated protesters are more likely to opt for nonviolent tactics, due to their intrinsic preferences for nonviolence and capacity for organizing the type of coordinated, large-scale resistance necessary to sustain a peaceful strategy. The capacity-mechanism may also imply that educated campaigns should have an overall higher probability of succeeding, due to their skills and resources. Section 6.5 demonstrates that educated campaigns are indeed more likely to become nonviolent. However, it finds no evidence of a direct effect of education on success – only of an indirect effect on success operating through nonviolence.

6.4 Data

6.4.1 Campaign participants’ education profile

There is no previous systematic data on socio-economic characteristics of participants in protest campaigns. Hence, to test the arguments presented above, new data was compiled

⁷The results are somewhat mixed, however. For instance, there is some evidence that educated individuals are more likely to participate in terrorist activity (Krueger and Maleckova, 2003).

on the extent to which students and educated individuals participated in the campaigns included in the NAVCO 1.0 dataset. This section describes the coding procedure and the steps taken to ensure reliability, such as tests of inter-coder reliability and incorporating uncertainty into the coding scheme and empirical analysis.

In NAVCO, campaigns are defined by Chenoweth and Lewis (2013) as “a series of observable, continuous, purposive mass tactics or events in pursuit of a political objective.” Campaigns should also have a discernible leadership, and have at least 1000 observed participants and a coherent organization.⁸ I include in the sample all campaigns except from self-determination and secession campaigns. This yields a sample of 203 campaigns reported to aim for either regime change or “other goals”, the latter usually referring to goals such as policy concessions.⁹ There are several reasons for restricting the sample to regime change campaigns. First, as noted, this paper partly seeks to contribute to the regime change literature, by exploring the argument that education promotes regime change through inducing non-violent opposition. Hence, the scope of this paper suggests testing the education-nonviolence link in *the context of anti-regime campaigns*. Second, national self-determination and secessionist campaigns are primarily carried out by ethnic groups as such, ranging across many or all social segments. They therefore do not necessarily exhibit systematic variation in social composition, making it less relevant to determine their social composition. There are also reasons to expect these campaigns to have a different dynamic and different mobilizing potential than campaigns working for regime change or liberalization (see e.g. Denny and Walter, 2014).

The coding scheme distinguishes between students and university graduates. While students are defined as individuals enrolled in and actively studying in a tertiary education institution, including universities and colleges, university graduates are defined as individuals with a completed degree from such institutions. The latter includes professionals such as lawyers, doctor and engineers. I include all types of education from universities and colleges, including for instance religious or military education.¹⁰ To gather the relevant information I partly relied on databases and encyclopedias such as “The international encyclopedia of revolution and protest” (Ness, 2015) and the “Global Non-Violent Action Database” (Swarthmore, 2015) and partly on news stories, books, journal articles and various reports by international or non-governmental organizations such as UNHCR

⁸As only campaigns of a certain size are included, sample selection bias can not be ruled out. For instance, education may influence to what extent campaigns become big enough to enter the dataset, and future research should examine this more closely. Note, however, that no students or graduates were reported to have participated in around 40 percent of the campaigns.

⁹Out of 203 campaigns, there are 8 with missing values on the education and student variables, as it was not possible to obtain information about the educational profile of their participants. Hence, the number of missing values is quite low.

¹⁰An area for further research would be to distinguish between types of higher education.

or Freedom House, were the relevant events were described. The main variables focus on the profile of the campaign's mass membership, but I also code variables recording the educational background of the campaign leader(ship).

Determining the educational profile of large, complex movements is associated with some challenges, as many movements are broad-based, consisting of coalitions of a wide range of groups and actors. For instance, the movement that staged the Iranian revolution comprised a large number of groups including students, teachers, doctors, communists, clerics, university professors and merchants. Meanwhile, there is a lot of variation in how crucial students and graduates are to a campaign's destiny, and we should expect this variation in *degrees of involvement* to matter for campaign outcomes, and not only whether educated protesters participate or not. To distinguish between degrees of involvement by students and educated individuals I collected data on several indicators. The codebook and details about the coding scheme are described in the appendix (section 11.2).

First, I coded a dichotomous variable with the value 1 if educated participants were explicitly reported by the sources to have *participated* in the movement and 0 if no educated participants were reported to be involved. Second, a variable was created indicating whether the movement *originated among* educated participants. Finally, I created a variable recording whether the movement was *dominated* by educated participants. For students, the equivalent to the three variables described above was coded.¹¹

To briefly illustrate the coding of the first two variables, the Swarthmore database lists the groups participating in the Bangladeshi protests against military rule from 1989-1990 as students, doctors lawyers, intellectuals, including members of the opposition party, but also other groups including workers and trade unionists. These groups are also listed by International Encyclopedia of Revolution and Freedom House (see also Immigration and Refugee Board of Canada, 1991). Thus, the campaign is coded as consisting of students and graduates. Another example is the Senderista Insurgency (Sendero Luminoso) against the Peruvian government from 1980-1999. As noted by several sources (e.g. McClintock, 1984; Swarthmore, 2015), this movement was founded among a group of students in San Cristobal de Huamanga, implying that the movement both consists of and originated among students.

To determine if a movement was *dominated* by students or educated individuals, I rely on a twin criteria requiring at least that students and educated groups were in majority, or, if not, they should have had a critical impact on the campaign, in the sense that they were the most important force behind the strategies and outcomes of the campaign. The majority criteria is based on descriptions in the sources of who the movement was made

¹¹ Again, see codebook for details about the coding criteria.

up of. The critical impact criteria is admittedly a more subjective criteria, but it also usually builds on explicit statements in the sources describing certain groups as the most crucial to the outcome.

To illustrate using the example of the Bangladeshi protests, the campaign clearly was a broad-based movement comprising many different groups, and so it is hard to determine if students or graduates were in majority. But, as noted by several sources, students and graduates seemed to play a critical role, as noted by UNHCR stating that "the student movement emerged as the driving force of protest, forcing the traditionally divided opposition parties to maintain a united front against Ershad" (UNHCR, 2005) (see also Freedom House 1991). Hence, the movement is coded as dominated by students. Regarding the Peruvian Senderista Insurgency, Global Britannica states that throughout its lifespan the movement's main recruiting base was among indigenous peasants and from poorer urban districts, and these groups dominated the campaign (and were in majority) during its final ten years (see also McClintock, 1984). Based on this information it can be inferred that, although the movement originated among students, it was dominated by neither students or educated protesters.

The coding procedure requires a certain amount of subjective judgment, especially when it come to the campaign domination variable, that may open up for measurement errors. To strengthen the confidence in the data reliability, several steps were taken. First, to be able to take the uncertainty more explicitly into account a variable was coded with the value 1 if the coding is fairly certain based on the available data, and 0 if considered quite uncertain. This uncertainty may arise from e.g. lack of data, making it hard to settle questions like whether students only participated in or dominated the movement, or from strong disagreement between sources regarding the campaign membership. I conduct robustness tests in which I exclude all the insecure observations from the analysis, and the main findings are robust to this (see appendix, Table 11.2 and 11.3).

Second, to assess the coding reliability I conducted inter-coder reliability tests across three different coders¹², by comparing the 10 percent of the cases that were double-coded by two researchers. This comparison indicates that inter-coder reliability is quite good: Out of 160 double-coded data points (20 cases were double-coded, with 8 indicators for each case), 81 percent of the double-coded observations were coherent across the two coders whose codings were compared. Finally, the fact that the coding was always based on multiple sources helps to mitigate measurement errors. In most cases, several sources pointed in the same direction, raising the confidence in the coding.

Another challenge to the data reliability is the possibility that the extent to which

¹²Around 50 percent of the cases were coded by the author, and the rest were coded by two research assistants

sources cover campaigns exhaustively, and thereby offer sufficient accounts of their social composition, is correlated with the outcome variable. For instance, I want to ensure that data availability and quality is not systematically poorer for violent campaigns, which could lead to an undercount of participation by educated groups because they are not mentioned in the sources. This could emerge either due to the fact that violent movement's are often smaller and carried out in poor, remote areas where news coverage is limited, or due to educated groups deliberately covering up or underplaying their own role in violent movements. Although such a bias can not be ruled out definitely, it does not seem to a major problem in this study. The descriptions of the relevant campaigns in the sources are usually exhaustive enough – also for violent campaigns – to be able to determine who the movement consisted of. This is supported by the fact that there is a very low correlation ($r=0.0051$) between "uncertainty status", which is assigned to cases where it was harder to obtain sufficient or high-quality information about the campaign's social profile, and violence. Meanwhile, the correlation between uncertainty and participation by educated protesters is also very low ($r=0.0852$). It should also be noted that my sample only include campaigns with at least 1000 participants, that are "observable", "continuous" and have a "discernible leadership" (Chenoweth and Lewis, 2013). Hence, we are only dealing with campaigns that are manifested in the literature to a certain extent, and the most small-scale or low-profile campaigns are excluded.

Based on the three indicators of participation by graduates and students I create two indexes, capturing degrees of involvement. The *graduate index* was given the value 0 if the movement neither consisted of, originated among or was dominated by graduates. If the movement scores the value 1 on *either* the involvement variable or the origin variable, and 0 on the dominate variable, the index was assigned the value 1. If the movement has the value 1 on both the involvement variable and the originate variable, but 0 on the dominate variable, the index was given the value 2. Finally, it was coded as 3 if the movement consisted of, originated among and was dominated by graduates. The *student index* was created according to the same rule. In the main analysis, I combine the student and graduate indexes into one additive *combined education index*. As the results may be influenced by the way this index is composed, I also run robustness tests with several alternative constructions of the education variable, including a multiplicative education index, a dichotomous index with the value 1 if students or graduates at least participated in the movement, separate indexes for students and graduates, and each of the dichotomous indicators described above, capturing various degrees of involvement. In brief, all these operationalizations yield similar results.¹³

The variables discussed above were coded based on campaigns' mass participation

¹³See appendix, section 11.6

profile. It could be argued, however, that to account for the involvement of educated individuals and students one should also study the campaign leader. For instance, it could be argued that a movement is dominated by educated individuals if the leader was a university graduate who exerted his power as a leader to inflict a certain outcome. One example is the Senderista insurgency discussed above, that consisted primarily of indigenous peasants and the urban poor, but the leader of the movement was Abimael Guzman, a former philosophy professor who arguably was crucial to the campaign’s destiny. To account for this, I also coded whether the campaign leader was a student or educate individual (see appendix, section 11.2 for details and coding rules). As a robustness check, I added this leadership variable to the education index. Replicating the main models with this revised indicator of protest participation does not, however, yield substantially different results (see appendix, section 11.10 for results).

6.4.2 Campaign outcomes

To measure whether the campaign opted for nonviolent tactics, I use the dichotomous indicator from the NAVCO 1.0 dataset, with the value 1 if the campaign was primarily nonviolent and 0 if otherwise. To measure if the campaign was successful, I use another dichotomous variable from NAVCO, with the value 1 if the campaign achieved all of its stated goals within a year of the “peak” of its activities and the value 0 if not. In a few instances, the campaigns’ goals were achieved years after the peak of the struggle, but if the success was considered a consequence of campaign activities these campaigns are coded as successful (Chenoweth and Stephan, 2011). This paper considers nonviolence and success as outcome variables, in line with Chenoweth and Stephan (e.g. 2011), and this approach departs somewhat from studies such as Celestino and Gleditsch (2013) and Dahl et al. (2013), that study nonviolent campaign onsets. Both approaches, however, assume that choice of nonviolent tactic is a relatively static value of interest, and this is supported empirically by indications that tactical choices are very stable in sustained movements (Dahl et al., 2013).

6.4.3 Control variables

There are several confounding factors that may affect *both* the extent to which students or educated were involved in the movement *and* the likelihood of nonviolence. One potential confounder is campaign size. We could expect large campaigns with many protest participants to be more likely to contain some students or educated participants (the larger the movement, the more likely that *any* group will be a part of it). At the same time, studies suggest that campaign size is correlated with nonviolence, for

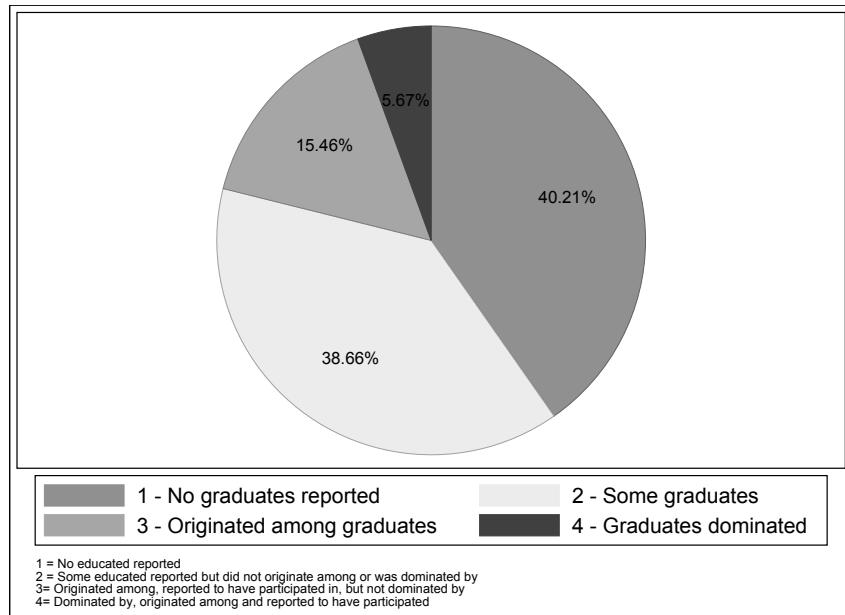
instance due to larger campaigns being more able to pressure the regime in a nonviolent manner or through nonviolent campaigns attracting more protesters (Chenoweth and Lewis, 2013). I therefore control for campaign size, using an indicator of the log number of participants in the campaign, from NAVCO. Since this control may also induce post-treatment bias, however, due to the possibility of educated protesters being able to build up larger movements. I therefore present models both with and without it, to check for coefficient stability.

I also include additional controls from NAVCO on characteristics of the campaign and the campaign target (most often the regime). We could expect factors determining the strength of the campaign vis-a-vis the target, to influence educated individuals decisions on whether to join a campaign or not. Educated individuals may face high opportunity-costs, and will be more hesitant to join campaigns where the expected likelihood of campaign success is low, or where protest participation can induce high costs, for instance due to regime repression. At the same time, these factors are likely to influence choice of non-violent versus violent strategies. I therefore control for the following characteristics: of the campaign and the target: Target capacity, target support from foreign powers, foreign support for campaign and whether the regime uses violence. In addition, key characteristics of the location country could affect both who participates and use of violence, in particular socio-economic variables that both have been shown to influence the likelihood of violence and that may influence the social composition of campaigns, in particular by influencing the social composition of a population. To account for this I control for GDP per capita, university enrollments, an indicator of manufacturing as proportion of GDP and the regime's democracy level (using the Polity index) in the main analysis. Additional controls are added in robustness tests presented in the appendix (section 11.5).

Finally, region dummies are included to control for context-specific confounding factors, for instance related to culture, history or geo-political factors, as well as decade dummies to account for time trends. Due to the low number of observations in each country and low number of campaigns in each year, there is not enough variation to include country fixed effects or year dummies.

A challenge emerging from the structure of the data – that e.g. Chenoweth and Stephan (2011) are also facing – is that it generally can not be ruled out that the causal sequence of the campaign characteristic variables is different than what is hypothesized, especially since the data is purely cross-sectional. Several of the discussed controls could potentially both induce omitted variable bias if excluded and post-treatment bias if included. These issues call for some agnosticism when choosing model specifications, and this is the reason why e.g. all models are also presented with each control left out at a

Figure 6.1: Involvement by the educated in protest campaigns.

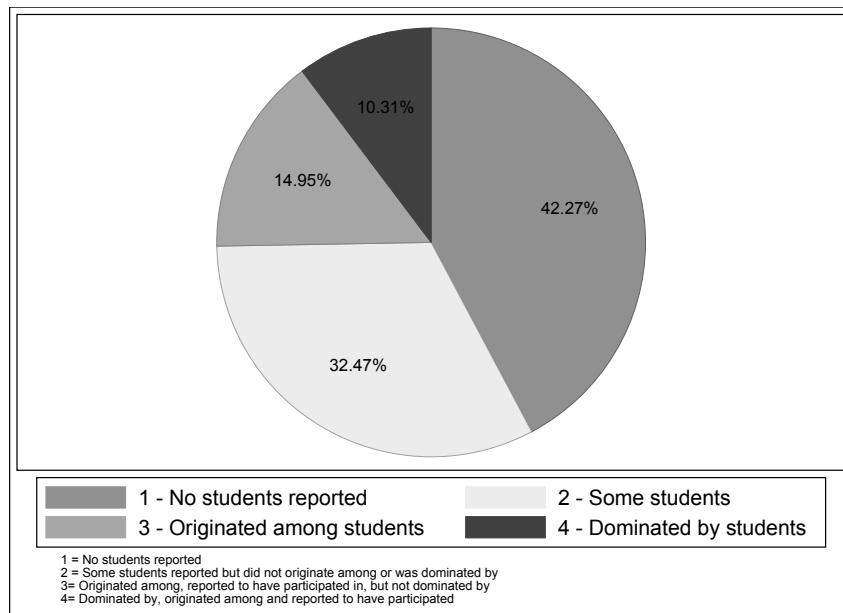


time (see appendix, Tables 11.4 and 11.5). The issue of endogeneity in the education-nonviolence relationship, as well as the causal sequence of nonviolence and success, is returned to in section 11.5 and 11.6.

The extent to which the protest campaigns in the sample consist of educated individuals and students is summarized in Figure 6.1 and Figure 6.2. Figure 6.1 shows the distribution of campaigns consisting of no graduates, some graduates, campaigns that originated among graduates and dominated by graduates. It shows that slightly more than half of the campaigns had at least some involvement by educated individuals. Meanwhile, there are only a few campaigns that were dominated by educated protesters. Figure 6.2 shows the distribution of campaigns with different degrees of involvement by students, showing a similar distribution to the educated campaigns.

Figures 3 and 4 offer some preliminary insights suggesting that campaign involvement by graduates or students is strongly related to use of nonviolence. The bar charts in Figure 3 shows the number of campaigns within each category (representing different degrees of involvement by university graduates), as well as the share of campaigns within each category that was nonviolent rather than violent. For instance, it shows that among campaigns with no reported graduates, a very large share (close to all) were nonviolent. Considering campaigns with at least some university graduates, the share of violent campaigns drops dramatically, and when considering campaigns dominated by university

Figure 6.2: Involvement by students in protest campaigns.



graduates, *all* of the campaigns were nonviolent (but the total number of campaigns in this group is low). The pattern is similar when looking at student involvement in Figure 4.

Figure 6.3: Involvement by the educated in violent and nonviolent campaigns.

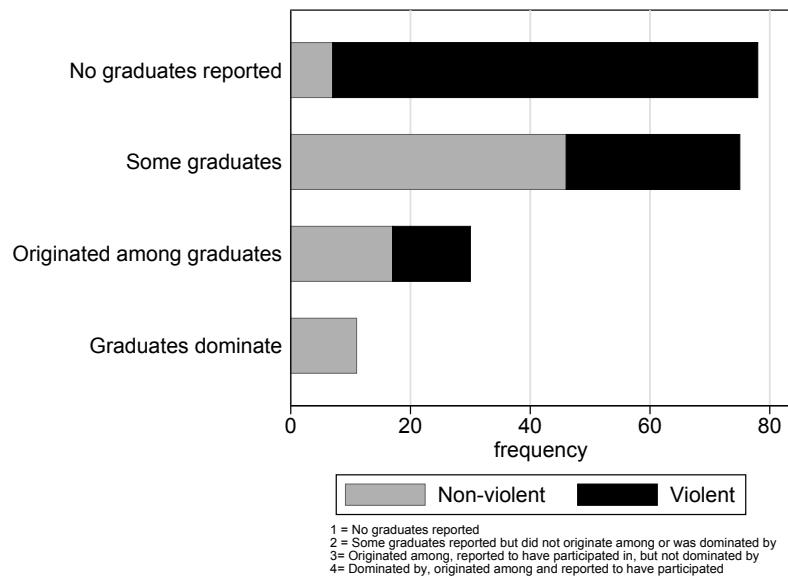
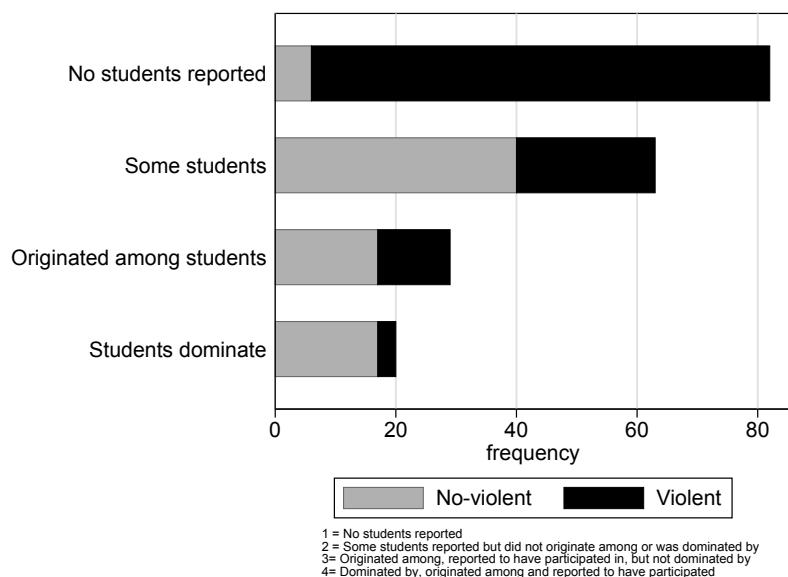


Figure 6.4: Involvement by students in violent and nonviolent campaigns.



6.5 Results

In this section I first present results from models investigating whether campaigns consisting of educated protesters are associated with less violence. In brief, there is strong evidence supporting this hypothesis. This motivates the second part of this analysis, examining whether campaigns with educated protesters are more likely to succeed. I find some, although not very robust, evidence that campaigns with a high degree of involvement by educated protesters are more prone to success, but, as I show, *only* due to their nonviolent dispositions.

6.5.1 Education and nonviolent protest

Table 6.2 contains results from logit models where the outcome is whether the campaign was nonviolent. The main explanatory variable is the education index, measuring the degree of involvement by both current students and graduates.¹⁴

Model A1 is a simple bivariate regression between the degree of involvement by educated protesters on the one hand and nonviolence on the other hand, revealing that the education index has a positive coefficient estimate of 0.947, which is highly significant (with a z-value of 5.9). When including region dummies in model A2, to account for cultural or geographic (time-invariant) variation between campaigns from different regions, both the coefficient estimate for education index and its z-value actually increase in size. Moreover the coefficient estimate increases further and is still strongly significant when introducing decade dummies in model A3 to account for campaign-invariant omitted variables that vary over time, and when including both region dummies and decade dummies in model A4.

In models A5-A7 I include several control variables capturing the strength of the campaign target and the dynamic between the campaign and the target. For instance, I control for a plausible potential confounder, namely campaign size.¹⁵ In addition, I add log GDP per capita, university enrollment, manufacturing as share of GDP, and the Polity score for the country where the campaign is located. The coefficient estimate for the education index jumps further (to 1.3) and is still precisely estimated (with a z-value of 4.4) when adding these controls in model A5. I also present results in the appendix (Table 11.4) from models where each of these controls are excluded one by one, and this

¹⁴I also run models where the education index is disaggregated by separating between involvement by students and graduates, and I find that both are positively and significantly associated with nonviolence (see appendix, Tables 11.10-11.12).

¹⁵However, all the models are also estimated without campaign size, and the results are not significantly altered (see appendix, Table 11.5).

Table 6.1: Non-violent outcome on campaign involvement by students and graduates (logit models)

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Education index	0.947*** (5.85)	1.029*** (5.79)	1.061*** (4.94)	1.146*** (4.68)	1.314*** (4.37)	1.354*** (4.51)	1.729** (3.13)
Membership size (log)					0.697*** (3.32)	0.846** (3.27)	0.663+ (1.83)
Target capacity (log)					-0.557* (-2.07)	-0.999* (-2.52)	-0.410 (-1.43)
Target support					1.965** (2.74)	2.225** (2.70)	5.238+ (1.80)
Campaign support					-3.578*** (-3.98)	-4.311*** (-4.13)	-9.301* (-2.13)
Regime violence					-0.961 (-0.69)	-0.939 (-0.66)	0.964 (0.66)
Polity index					-0.114* (-2.00)	-0.154* (-2.30)	-0.215+ (-1.78)
GDP p.c.(log)					0.677 (1.32)	1.575* (2.43)	1.378 (1.39)
University enrollment					-0.00319 (-0.74)	-0.00496 (-0.87)	-0.0337+ (-1.96)
Industrial activity					-0.000646 (-0.73)	-0.000819 (-0.59)	0.00237 (0.94)
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes
R2 (pseudo)	0.281	0.402	0.464	0.558	0.588	0.690	0.711
ll	-93.76	-61.94	-55.60	-31.45	-31.48	-23.71	-17.18
N	192	153	153	107	111	111	86

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-scores in parentheses. Standard errors are clustered on country. Constant term, region and decade dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily nonviolent.

does not alter the main results. When adding region dummies in model A6 to account for cultural or regional variation between campaigns from different regions, the coefficient estimate for the education index increases further and is still precisely estimated (with a z-score of 4.5), and when replacing region dummies by decade dummies the coefficient estimate jumps to 1.7 (with a z-score of 3.1).¹⁶

The number of observations drops when adding region dummies and decade dummies, due to the fact that regions and decades without sufficient variation on the dependent variable are dropped. Due to missing values on the control variables, the number of

¹⁶Due to the low number of observations, the logit models unfortunately did not converge when including both region and decade dummies simultaneously

observations also drops quite a bit when adding these (from 192 observations in the most simple model A1 to 111 in model A5 with full set of controls). To make sure that listwise deletion is not inducing selection bias - due to the possibility that missing values are correlated with non-violence and campaigns' social profile, I also conduct multiple imputation using the Amelia II software (Honaker et al., 2011). Yet, when replicating Table 6.2 on the full imputed sample, the coefficient estimate(s) for the education index actually increases in size, and is now always significant at the 0.001 level.¹⁷ This suggests that the results are not systematically biased due to missing values.

Finally, all models in Table 6.2 are also estimated with numerous additional controls (see appendix, Table 11.6), to account for other sources of heterogeneity between campaigns. First, I include additional controls capturing socio-economic characteristics of the host country, including e.g. urbanization, population size and GDP growth. Moreover, I include dummy variables for whether the campaign took place in the cold war or in the post- cold war, an indicator of whether other violent campaigns took place simultaneously and whether the government was involved in intrastate war. Adding these controls does not significantly alter the results. Hence, it seems that the identified relationship is not driven by the political, socio-economic and campaign characteristics included in these models, nor by regional differences.

In sum, despite the relatively low number of observations, there is robust evidence that the level of involvement by educated protesters is positively associated with a nonviolent strategy.

To assess the substantial importance of education, Figure 6.5 displays predicted probabilities of campaigns being nonviolent for different values of the education index, based on model A5 from Table 6.2, with all controls set to their mean value, and created using the Clarify package in Stata. It illustrates how the likelihood of nonviolence increases with more involvement by educated protesters, from close to 0 when considering campaigns with no reported educated participants to 1 when considering campaigns dominated by educated protesters.¹⁸

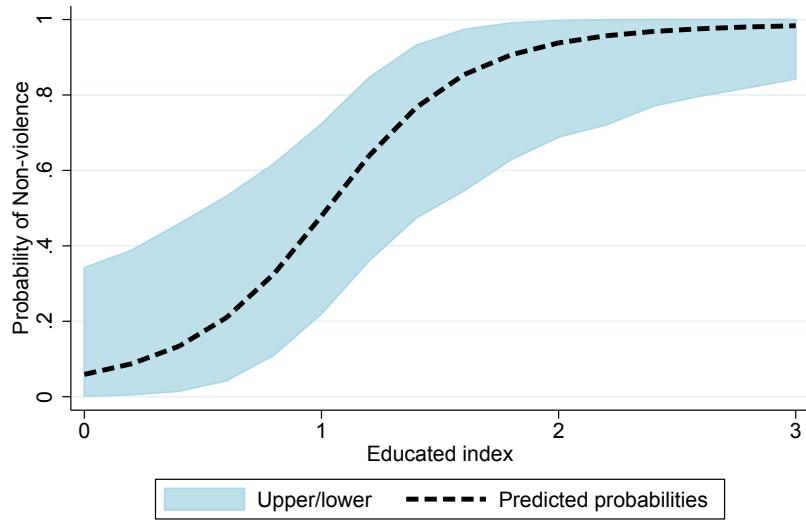
6.5.2 Education and successful outcomes

The finding that protest campaigns with considerable involvement by educated protesters are more likely to be nonviolent begs the question of whether educated movements, due to their pacifist inclinations, also are more likely to succeed. This is a plausible implication,

¹⁷See appendix, section 11.11 for results.

¹⁸Note that the number of observations with 3 on the education index is very low, most of which are nonviolent.

Figure 6.5: The predicted probability of nonviolence



Based on model A5 in table 1 with several controls. Logit regression.
All control variables are set to mean.

especially in light of studies indicating that nonviolent movements are more successful (Chenoweth and Stephan, 2011). Also, the discussed mechanisms imply that the presence of educated participants may both have an indirect effect on the likelihood of success through nonviolence *or* a direct effect on success.

Table 6.2 contains results from logit models of the likelihood of a protest campaign succeeding (in terms of obtaining their main goals). I start by investigating the relationship between participation by educated protesters and success, without controlling for potential mechanisms. In brief, the education index coefficient is positive with z-values ranging from 2.2 to 1.9 in the simple bivariate regression in model B1, when controlling for log GDP, democracy level and region dummies in model B2 and when replacing region with decade dummies in model B3. Yet, the z-score drops to 1.1 in the more complex model B4. Hence, there is some, although not very strong evidence that campaigns with educated protesters are more likely to succeed in terms of achieving their goals.

In model B5 I include the indicator of nonviolent campaigns, to probe the mechanism through which education induces success. In line with previous literature, this model indicates that nonviolent campaigns are associated with more success. Meanwhile, the coefficient estimate for the education index drops dramatically and is now close to zero with a negative sign and a z-value of only -0.04. Hence, the effect of education on success is almost completely soaked up when controlling for nonviolence, suggesting that educated participants *only* affect success rates indirectly, through inducing peaceful resistance. In models B6-B8 I add the indicator of nonviolence to models with additional controls, region

Table 6.2: Successful outcome on campaign involvement by students and graduates (logit models)

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
Education index	0.228* (2.17)	0.301* (1.99)	0.248+ (1.87)	0.191 (1.10)	-0.00497 (-0.04)	0.110 (0.66)	0.0586 (0.38)	0.0126 (0.07)
Nonviolent					1.369*** (3.51)	1.361* (2.51)	1.386* (2.25)	2.209** (2.71)
Polity index		0.0269 (0.87)	0.0375 (1.24)	0.00882 (0.21)		0.0456 (1.50)	0.0596+ (1.85)	0.0527 (1.20)
GDP p.c. (log)		0.269 (1.22)	0.319+ (1.81)	0.445 (1.47)		0.103 (0.42)	0.228 (1.27)	0.330 (1.07)
Target capacity				8.558 (1.47)			9.227 (1.52)	
Duration					-0.227 (-1.50)			-0.185 (-1.14)
Target support					0.588 (1.03)			0.199 (0.33)
Campaign support					0.661 (0.94)			1.602+ (1.67)
Military defection					1.650** (2.80)			1.816** (2.96)
Regime sanctioned					0.418 (0.58)			0.142 (0.21)
Regime violence					-1.476* (-2.28)			-1.331+ (-1.93)
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes
r2 (pseudo)	0.0252	0.162	0.114	0.318	0.0773	0.200	0.147	0.364
ll	-126.0	-87.00	-91.08	-65.56	-119.3	-83.08	-87.65	-61.16
N	192	154	152	143	192	154	152	143

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-scores in parentheses. Standard errors are clustered on country. Region and decade dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing its main goals.

dummies and decade dummies, and there is still no significant effects for the education index, and in models B7 and B8 the coefficient estimate is again approaching zero.

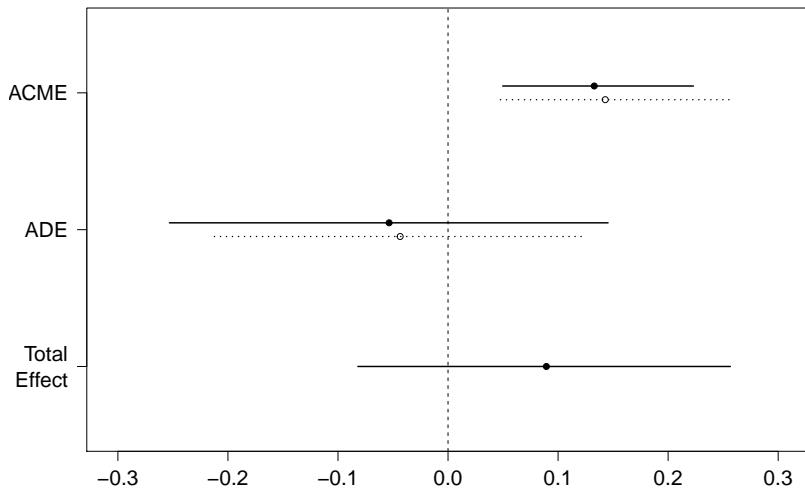
However, as has been pointed out by e.g. Imai et al. (2011) (see also Imai et al. 2010), comparing the coefficient estimate of the main independent variable before and after adding the proposed mediating variable, is often a poor way of estimating mediation effects, especially for non-linear (logit) models as used in Table 6.2. Hence, I also apply the causal mediation analysis procedure proposed by Imai et al. (2011).¹⁹ Drawing on the

¹⁹I use the R mediation package developed by Tingley et al. (2014)

potential outcomes framework for causal analysis, this estimation procedure decomposes the average treatment effect of an explanatory variable into the average causal mediation effect (ACME) and the average direct effect (ADE). Figure 6.6 plots results from the causal mediation analysis with 95 percent confidence intervals. The average causal mediation effect of education through nonviolence on the likelihood of success is 0.13, and this is statistically significant at the 0.00 level. In total, the model estimates that 0.28 percent of the total effect of education on success is mediated through this channel. Moreover, there is no statistically significant *direct* effect of education on the likelihood of success. This procedure rests on several assumptions, including the absence of omitted variables influencing both the mediator and the outcome, that could potentially be violated here. To assess to what extent the findings are sensitive to such potential violations, I conduct numerous sensitivity test checking how large a potential omitted variable bias must be to push the mediation effect to zero (or even negative). The results are presented in the appendix, suggesting this is unlikely (see section 11.9 in the appendix).

Hence, this result further corroborates the idea that the level of education among protesters does not have a direct effect on success, but only an indirect effect, through nonviolence. The next section discusses the robustness of this finding, as well as the main finding that education is associated with nonviolence.

Figure 6.6: The mediation effect of education on successful outcome, mediated through nonviolence



Notes: The total effect of education on the likelihood of success is decomposed into the average direct effect (ADE) and the average causal mediation effect (ACME). The effects are plotted with their 95 percent confidence intervals. Education is measured using a binary version of the combined education index. The pre-treatment covariates are log number of participants, log GDP, the Polity index, target capacity, campaign duration, regime violence, government capacity, foreign aid to the regime and military defection. See appendix for alternative model specifications.

6.6 Empirical extensions

The findings that campaigns with students and educated protesters are clearly associated with a lower likelihood of violence is interesting on its own term, given that it is the first systematic evidence mapping the educational background of participants in protest campaigns, and the first evidence linking socio-economic characteristics of protest campaigns to their nature and outcome. Ideally, however, we also want to determine whether this is due to an *effect of* involvement by educated protesters on the likelihood of nonviolence. This section explores this.

First, campaigns consisting of a large fraction of educated protesters will most likely differ in many other ways from campaigns without educated protesters than only their nonviolent disposition. If the findings in this paper are to support a causal interpretation this potential heterogeneity can not be correlated with the outcome, but it is highly likely that certain campaign characteristics may induce nonviolent outcome *and* attract educated protest participants. The models presented above control for many important potential confounders related to the economic and political context of the campaign location or characteristics of the campaign or its target. As noted, I have also excluded a major source of campaign heterogeneity by restricting the sample to campaigns aiming for regime change or "other goals" and national self-determination campaigns, that may have characteristics that could correlate with both the education index and the likelihood of nonviolence.

Yet, there may still be *unobserved* variables correlated with both the education index and nonviolent outcomes. To assess to what extent this is a threat, I use the strategy proposed by Imbens (2003) (see also Clarke, 2009), offering a benchmark to assess how large omitted variable bias has to be to reduce the "treatment" coefficient substantially - in this instance the coefficient estimate for the education index. If this method shows that potential unobserved confounders need to be unreasonably larger than included confounders, we can strengthen our confidence that the results are not due to omitted variable bias. More specifically, I use the Generalized Sensitivity Analysis developed by (Harada, 2015), which is a refinement of the original setup proposed by Imbens (2003).²⁰ After identifying the treatment variable in the regression equation, the procedure generates a sequence of pseudo-random confounders that, once added to the regression equation, reduce the coefficient of interest by a chosen magnitude.

The curve in Figure 6.7 describes how much of the variance in nonviolence *and* in the education index an unobserved covariate would have to explain to reduce the education index coefficient by half. It shows that it would have to explain, for example, around

²⁰The plots are created using the gsa package in Stata

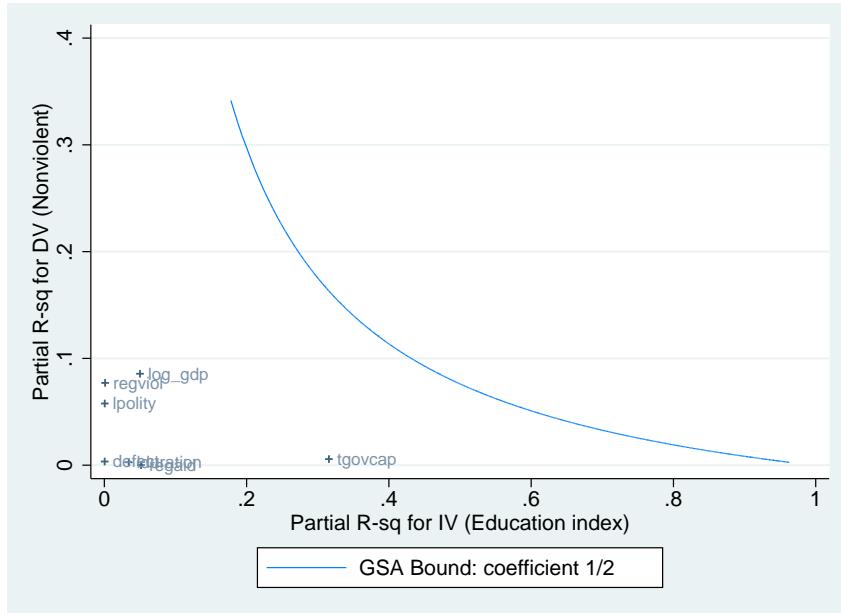
30 percent of nonviolent protest and 20 percent of participation by educated protesters. When compared to the variance explained by the included covariates, that lie far below the curve, this is a lot. For instance, the level of socio-economic development, measured using log GDP per capita, is among the variables that explain the most of nonviolence (less than 0.1), but it only explains less than around 0.03 of the variation in the education index. Meanwhile, government capacity, which explains more of the variation in the education index than any other variable, around 0.3, explains close to zero of the variation in nonviolence. In order for an unobserved covariate to wipe out the effect of the education index on the likelihood of nonviolence it would have to explain more of the variation the two variables than the combined variance explained by several plausible variables including log GDP, the polity level, government capacity, campaign duration, external support and whether the regime uses violence. Given that the existence of an unobserved variable of that importance is unlikely, the education index coefficient should be considered robust to potential omitted variable bias. When running the generalized sensitivity analysis to investigate how large the omitted variable bias would be to push the coefficient estimate for the education index down to zero, it produces a curve that, as expected, is even further away from zero and the other included controls (see appendix, Figure 11.1).²¹

Another finding that soothes some concerns about omitted confounders is the fact that I also identify effects of the dichotomous indicator of whether the campaign *originated among* educated protesters on the likelihood of nonviolence (see appendix, Table 11.7). I have discussed how campaign characteristics suggesting it will turn nonviolent may attract educated protesters, inducing endogeneity, but whether the campaign originated among educated protesters is more plausibly exogenous to other characteristics of the campaign that evolves as the campaign develops, such as responses by the regime, choices made by the leaders and participants, other groups entering or leaving the campaign and the choice of foreign or external actors to throw in their support. Hence, this indicator is less vulnerable to endogeneity bias.

So far I have discussed threats to inference arising from omitted variables, but another source of endogeneity is the possibility of reverse causation, due to a direct effect of nonviolence on campaign participation by educated protesters. It should be noted, however, that many of the factors that may plausibly signal to potential protesters that the movement will turn violent, and thereby induce participation, are really factors *correlated* with a nonviolent outcome rather than the actual nonviolent campaign nature,

²¹I also run this test with various other model specifications, including using the logit estimator rather than the ylpm model, adding region dummies and excluding all controls but log GDP and polity level (see appendix, section 11.7).

Figure 6.7: Assessing the threat of omitted variable bias



Notes: The plotted line shows where the coefficient estimate for the education index (in models explaining nonviolent outcome) would be half of its original size. Included control variables are logged GDP, the use of violence by the regime (regviol), the polity level (lpolity), government capacity (tgovcap), foreign aid to the regime (regaid) and the duration of the campaign (lduration). The model is estimated using a linear probability model.

such as for instance statements, personality and background of campaign leaders and use of violence in previous protests in that country. Moreover, although reverse causality can not be completely ruled out, it is also made less plausible by the fact that I identify effects of campaigns originating among educated protesters on nonviolence. The indicator of nonviolence used in this paper measures whether the movement is mainly nonviolent throughout its lifespan, and it is unlikely that this is already determined when the movement originates. In sum, this section strengthens the confidence in the findings in this paper, indicating that the most plausible interpretation is that educated protesters induce nonviolence.

6.7 Conclusion

Utilizing a novel data set on the educational background of participants in major protest movements, this study finds that protest campaigns with a high degree of involvement by educated participants are more likely to be nonviolent rather than violent. After discussing and assessing possible explanations for this finding, I contend that it is plausibly due to an effect of participation by educated protesters on nonviolence. Meanwhile, I identify some evidence that campaigns consisting of educated participants and students

are more likely to succeed, but this is likely *because* they are also more nonviolent. This finding has implications for the engrained claim in political science that education is a force for peace, by demonstrating that groups of educated individuals will be more inclined than non-educated groups to choose peaceful protest. Moreover, it adds to the democratization literature by suggesting a potential mechanism explaining the correlation between education and democratic transition, namely by promoting peaceful resistance. First and foremost, however, it is a contribution to the literature on the nature of protest campaigns (e.g. Chenoweth and Stephan, 2011), suggesting that who the protesters are matters for campaign outcomes, and not only the context they operate in and the goals they are working for.

Part III

Appendices and Bibliography

7 Appendix to Chapter 2

7.1 Introduction

This appendix offers extensions of the main discussions from chapter 2, as well as several robustness tests. In the next section, I present results from models distinguishing between democratic transitions and democratic breakdown that were discussed in the main text. In section 7.3, I present and discuss results from instrumental variable regressions of democracy on education quality, including both the first stage and second stage regressions. In section 7.4, I present and discuss various models, referred to in the main text, investigating whether there is an interaction effect of education quality and quantity on democracy or democratic transitions. In section 7.5 I present additional details about the multiple imputation procedure and assessments of its performance. In this section, I also present results from estimations based on the original, non-imputed dataset, as well as estimations using multiple imputation-corrected standard errors.

Finally, in section 7.6 I present additional robustness tests based on numerous alternative model specifications. For instance, I present models with alternative lag structures of the independent variables, different operationalization of the main dependent variable, additional control variables and models distinguishing between different subcomponents of the primary and secondary quality indicators.

7.2 Distinguishing between democratic transitions and survival

This section discusses results from dynamic probit models distinguishing between democratic transitions and democratic survival. The dependent variable is the dichotomous democracy indicator by Boix, Miller and Rosato (2013). All independent variables are lagged by one year, and interacted with lagged democracy. This means that each constitutive term can be interpreted as the effect of the independent variable on the likelihood of transitioning from autocracy to democracy (more specific, the likelihood of being a democracy in t if the country was an autocracy in $t-1$), and the interaction terms can be interpreted as the effect of each independent variable on democratic survival (the likelihood of being a democracy in year t if the country was a democracy in $t-1$)

Table 7.1 contains results from models with education quality as independent variable. The constitutive terms for primary and secondary quality can now be interpreted as the effect of each variable on the likelihood of democratic transitions. In brief, there is no robust evidence that neither primary or secondary quality are positively related to democracy in models 1-6 in table 7.1. The primary education quality is positive and significant at the 0.05 level in model 1, but there is no significant effects in models 2-6 when e.g. controlling for education quantity and adding fixed and year dummies. Moreover, there is no evidence at all that secondary education quality is positively related to democratic transitions, and the coefficient estimate is actually negative in all models, but model 3 is the only model where the z-values exceeds -1.50. Nor is there any evidence that primary or secondary quality are positively related to democratic survival.

In Table 7.2, indicators of education quality are replaced by indicators of education type (in models 1-5) and education autonomy (in models 6-10). In brief, and contrary to the hypotheses, there is no evidence that emphasis on social sciences or humanities is associated with democratic transitions. The coefficient estimate for humanities, which can now be interpreted as the effect of a large share of humanities graduates on democratic transitions, is actually negative in models 1-5 (but with low z-values) and the coefficient estimate for social science graduates is positive but never reaches conventional levels of statistical significance. Nor is there any evidence that emphasis on humanities or social sciences is positively related to democratic survival. The coefficient estimates for the interaction term between the humanities and democracy is actually negative in models 1-3, with z-values exceeding -2 in all three models. The interaction term between social sciences and democracy is very low and never statistical significant.

Finally, looking at education autonomy in models 6-10, there is no evidence that

Table 7.1: Education quality. Transitions to democracy and autocratic survival

	1 Probit	2 Probit	3 Probit	4 Probit	5 Probit	6 Probit
Education quality						
Primary quality	0.0504*		0.0257	0.0226	0.0143	-0.0604
	(2.42)		(1.35)	(0.99)	(0.55)	(-1.55)
Secondary quality	-0.0275		-0.0431*	-0.0183	-0.0260	-0.0426
	(-1.34)		(-2.36)	(-0.80)	(-1.02)	(-1.27)
Primary quality*Democracy	-0.0780		-0.0436	-0.0216	-0.0653	0.0582
	(-0.77)		(-0.52)	(-0.30)	(-0.79)	(0.89)
Secondary quality*Democracy	0.00966		0.0357	-0.0509	-0.0452	-0.0349
	(0.09)		(0.33)	(-0.53)	(-0.37)	(-0.40)
Controls						
Democracy	4.587	2.406	2.092	16.67*	4.397	3.362
	(1.29)	(0.84)	(0.53)	(2.37)	(0.53)	(0.51)
GDP(log)	0.768***	0.636***	0.627***	2.067***	0.763***	1.702***
	(5.14)	(7.37)	(7.27)	(9.10)	(4.63)	(4.93)
Democracy*GDP	0.661					
	(1.53)					
Years of schooling	0.176***	0.202***	0.167*	0.231**	0.0118	
	(4.11)	(3.84)	(2.23)	(3.23)	(0.10)	
Democracy*Schooling(years)	-0.306	-0.306	-0.0170	-0.449	-0.255	
	(-1.10)	(-1.06)	(-0.06)	(-1.26)	(-1.21)	
Democracy*GDP	0.864+	0.895+	-1.174*	1.977**	0.487	
	(1.77)	(1.77)	(-1.96)	(2.67)	(0.71)	
Urbanization			0.0761***	-0.0240***	0.103***	
			(3.58)	(-4.28)	(4.16)	
Democracy*Urban			0.0422	0.00534	0.0155	
			(1.45)	(0.12)	(0.53)	
Gini index			0.0139	0.00620	0.0415+	
			(0.93)	(0.39)	(1.93)	
Democracy*Gini			-0.0160	-0.110	-0.00874	
			(-0.26)	(-1.53)	(-0.18)	
Fuel			-0.0103*	-0.0265***	-0.0261**	
			(-2.04)	(-5.63)	(-2.70)	
Democracy*Fuel			0.000633	0.0113	0.000141	
			(0.05)	(0.67)	(0.01)	
Country dummies?	no	no	no	yes	no	yes
Year dummies?	no	no	no	no	yes	yes
N	4876	4876	4876	3365	4312	2026
ll	-818.6	-813.7	-810.8	-592.0	-478.8	-301.3

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are clustered on country. Dependent variable is the dichotomous democracy indicator by Boix, Miller and Rosato (2013). Constant, country dummies and year dummies are omitted from the table.

a large share of students studying abroad in a democracy in t-5 is positively related to democratic transitions in t. The coefficient estimate for the constitutive term for students abroad in democracies is positive, but the z-value is only barely higher than 1 in one out of six models. Nor is there any robust evidence that foreign education (from democracies) is positively related to democratic survival in the home country, as the coefficient estimate for the interaction term between students abroad (in democracies) and democracy is only weakly significant at the 0.10 level in model 2, but insignificant in all other models.

In sum, the models distinguishing between democratic transitions and democratic survival yield no more evidence suggesting that education quality is associated with democratic transitions, nor is it associated with democratic survival. Nor is a country with an emphasis on certain types of education (social sciences and humanities) or a large share of students having studied abroad in democracies, more likely to experience democratic transition or democratic survival.

Table 7.2: Education type and education autonomy. Transitions to democracy and democratic survival

	1 Probit	2 Probit FE	3 Probit FE	4 Probit	5 Probit FE	6 Probit	7 Probit FE	8 Probit FE	9 Probit FE	10 Probit FE
Education type										
Engineering	0.0604** (2.58)	0.0574* (2.37)	0.233*** (5.56)	0.220*** (4.58)	-0.619 (-0.77)					
Science	0.0884** (2.19)	0.0940* (2.34)	0.121+ (1.92)	-0.0749 (-0.58)	-1.994* (-2.10)					
Humanities	-0.00750 (-0.27)	-0.00990 (-0.36)	-0.0358 (-1.04)	-0.218 (-1.22)	-2.245+ (-1.93)					
Social science	0.0177 (0.85)	0.0183 (0.90)	0.0123 (0.41)	0.0364 (0.47)	-0.456 (-1.37)					
Democracy*engineering	-0.104 (-1.21)	-0.180* (-2.13)	-0.258* (-2.19)	-0.286*** (-4.69)	-0.263 (-0.34)					
Democracy*humanities	-0.181** (-2.87)	-0.178** (-2.94)	-0.191* (-2.37)	-0.151 (0.86)	0.878 (1.47)					
Democracy*soc.science	-0.00290 (-0.04)	0.0122 (0.18)	-0.0103 (-0.10)	-0.0445 (-0.56)	-0.777 (-1.51)					
Democracy*GDP	-0.0000538 (-0.85)	-0.000186** (-2.62)	-0.000457* (-2.16)	0.000248** (2.67)	0.000826 (0.83)	0.000551** (2.99)	0.000674** (2.99)	0.0000484 (0.39)	0.000706** (2.99)	0.000129 (6.95)
Education autonomy										
Students abroad(dem.)						1.216 (1.08)	1.182 (0.96)	0.888 (0.53)	1.324 (0.99)	1.222 (0.66)
Students abroad(aut.)						-0.0652 (-0.05)	-0.123 (-0.11)	1.674 (0.71)	-0.213 (-0.18)	1.101 (0.49)
Abroad(dem.)*Democracy						6.809 (6.760+)	3.676 (1.79)	7.141 (0.86)	4.751 (1.59)	4.751 (0.96)
Abroad(aut.)*Democracy						0.587 (0.09)	0.502 (0.09)	-0.632 (-0.09)	0.208 (0.04)	0.862 (0.15)
Controls										
Democracy	10.96** (2.74)	9.592** (2.69)	12.38* (2.37)	8.709** (2.89)	42.48 (1.44)	5.913*** (8.88)	4.670*** (7.43)	6.464*** (8.09)	5.372*** (7.13)	
GDP(log)	0.836*** (4.04)	0.794*** (3.35)	2.786*** (4.47)	-0.304 (-0.88)	1.124 (0.23)	0.299* (2.05)	0.130 (1.01)	1.547*** (4.05)	0.168 (1.27)	1.679*** (3.66)
Democracy*GDP	-0.0000538 (-0.85)	-0.000186** (-2.62)	-0.000457* (-2.16)	0.000248** (2.67)	0.000826 (0.83)	0.000551** (2.99)	0.000674** (2.99)	0.0000484 (0.39)	0.000706** (2.99)	0.000129 (0.95)
Tertiary graduates		0.887 (0.21)	12.28** (2.66)	-12.47 (-1.04)	228.4+ (1.87)	6.978+ (1.83)	1.792 (0.32)	6.939+ (0.32)	0.657 (1.70)	0.657 (0.11)
Democracy*tertiary		28.66* (1.97)	37.45 (1.57)	19.32 (1.30)	-0.133 (-0.00)	-16.19 (-1.18)	4.877 (0.46)	-16.03 (-1.16)	5.505 (0.46)	
Country dummies?	no	no	yes	no	yes	no	no	yes	no	yes
Year dummies?	no	no	no	yes	yes	no	no	yes	yes	yes
N	1321	1321	483	887	119	2879	2879	1219	2879	1219
ll	-208.8	-206.0	-155.3	-46.34	-14.34	-354.3	-352.3	-239.7	-330.6	-215.2

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. All explanatory variables are lagged by 1 year. Standard errors are clustered on country. Dependent variable is the dichotomous democracy index from Boix, Miller and Rosato (2013). Constant, country dummies and year dummies are omitted from the table.

7.3 Checking the robustness of the quality-democracy link: Dealing with endogeneity

This section presents results from models attempting to deal with the fact that education quality may be endogenous to democracy. Table 7.3 shows results from instrumental variable regression (2SLS) models, where primary and secondary education quality are in turn instrumented for using and indicator of numeracy skills in 1820. The exclusion restriction requires that early numeracy skills should only affect current levels of democracy via current levels of education quality. For this to hold, the models need to control for a number of potential alternative channels through which early numeracy skills may affect current levels of democracy. Hence, all models control for democracy and log gdp, lagged by one year. Numeracy skills in the 1820s may also have affected early democracy levels and economic development levels, which may have pushed countries on to certain paths that make it more likely to have higher democracy levels today. Hence, in all models I also control for democracy in the 1850s, measured using Polity, and log gdp in the 1850s.¹ When adding only this set of controls, there is no evidence that neither primary quality (model 1) and secondary quality (model 2) affects democracy. Moreover, this result holds when adding region dummies (as the instrument is time-invariant, country fixed effects are not possible), year dummies, as well as both region and year dummies combined with numerous additional controls: Years of schooling, urbanization, the gini index, fuel dependency, gdp growth, youth bulges, ethnic fractionalization and unemployment.

First-stage regressions of primary quality on early numeracy skills are presented in table 7.4, and first-stage regressions of secondary quality on early numeracy skills are presented in 7.5. The results suggest that the instrument is relevant. The coefficient estimate for early numeracy skills is significant at the 0.001 level in all models explaining primary quality, and in all models explaining secondary quality. Moreover, the F-test suggests that the instrument is strong. For models explaining primary quality, the F-test values ranges from the 102.9 to 11.9, and for models explaining secondary quality the F-tests range from 256.1 to 16.5.

¹The VIF test suggests that collinearity is a not a major problem in these models

Table 7.3: Democracy and education quality. 2SLS models where quality is instrumented using historical levels of numeracy skills

	1 2SLS	2 2SLS	3 2SLS	4 2SLS	5 2SLS	6 2SLS	7 2SLS	8 2SLS
Primary quality	0.297 (1.50)	0.213 (1.12)	0.217 (0.87)	0.276 (0.43)				
Secondary quality					0.0947 (0.86)	0.148 (1.12)	0.211 (0.87)	0.214 (0.43)
Democracy	0.873*** (54.44)	0.871*** (54.83)	0.874*** (56.49)	0.860*** (47.82)	0.886*** (62.05)	0.870*** (54.38)	0.872*** (53.14)	0.857*** (39.78)
GDP(log)	0.663 (0.87)	0.992 (1.49)	0.897 (0.62)	1.659 (1.27)	0.761 (1.40)	1.271* (2.43)	0.662 (0.40)	1.283 (0.98)
Polity 1850s	0.0811 (1.57)	0.0507 (0.78)	0.0489 (0.78)	0.0818 (1.06)	0.0500 (1.06)	0.0489 (0.75)	0.0505 (0.80)	0.0905 (1.25)
GDP (log) 1850s	-1.911* (-2.46)	-2.167* (-2.46)	-2.140* (-2.51)	-2.896*** (-3.32)		-2.114* (-2.45)	-2.149* (-2.50)	-2.934*** (-3.34)
Years of schooling				-0.303 (-0.86)				-0.270 (-0.88)
Urbanization					0.0752 (1.59)			0.0812+ (1.92)
Gini index					0.0285 (0.24)			0.0239 (0.22)
Fuel					0.000656 (0.03)			0.00575 (0.27)
GDP growth					0.119 (1.07)			0.115 (1.08)
Youth bulge					0.140 (0.61)			0.105 (0.61)
Ethnic fractionalization					-2.501 (-0.87)			-2.656 (-1.01)
Unemployment					0.158 (1.45)			0.170 (1.33)
Region dummies?	yes	yes	yes	yes		yes	yes	yes
Year dummies?		yes	yes	yes		yes	yes	yes
N	799	799	799	799	841	799	799	799
r2	0.883	0.885	0.897	0.898	0.877	0.885	0.896	0.898

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index (Marshall, 2014). Primary quality and secondary quality are instrumented using the level of numeracy skills in 1820. Constant, country dummies and year dummies are omitted from the table.

Table 7.4: First-stage regressions with primary education quality as DV

	1	2	3	4
Numeracy skills 1820s	0.110*** (3.55)	0.108*** (10.36)	0.125*** (7.70)	0.0612*** (3.90)
Democracy	0.0134 (0.98)	0.0125 (0.97)	0.0116 (0.94)	0.0102 (1.50)
GDP(log)	1.568*** (6.27)	1.604*** (6.60)	0.720 (0.93)	-0.533 (-0.75)
Polity 1850s	-0.0844 (-0.96)	-0.0231 (-0.54)	-0.0177 (-0.37)	0.0305 (0.72)
GDP p.c (log) 1850s	1.473+ (1.67)	1.858*** (3.39)	1.992** (3.28)	0 (.)
Years of schooling				0.114 (0.74)
Urbanization				0.0167 (0.55)
Gini index				-0.0724* (-2.37)
Fuel				0.0135 (1.06)
GDP growth				-0.158*** (-4.27)
Youth bulge				-0.230*** (-3.38)
Ethnic fractionalization				-3.386*** (-3.92)
Unemployment				-0.0590 (-1.41)
Region dummies?	no	yes	yes	yes
Year dummies?	no	no	yes	yes
N	799	799	799	799
F-test	102.90	129.48	68.86	11.92

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Primary quality is instrumented for using the level of numeracy skills in 1820. Constant, country dummies and year dummies are omitted from the table.

Table 7.5: First-stage regressions with secondary education quality as DV

	1	2	3	4
Numeracy skills 1820s	0.172*** (3.57)	0.157*** (17.25)	0.139*** (10.02)	0.0750*** (6.27)
Democracy	0.0133 (1.16)	0.0130 (1.14)	0.0135 (1.20)	0.0171* (2.21)
GDP(log)	0.329 (1.33)	0.373 (1.53)	1.436+ (1.88)	0.361 (0.48)
Polity 1850s	-0.126 (-1.11)	-0.0101 (-0.15)	-0.0159 (-0.24)	0.0168 (0.32)
GDP p.c. (log) 1850s	1.325 (1.28)	2.121** (3.26)	1.925*** (3.37)	0 (.)
Years of schooling				0.0194 (0.15)
Urbanization				0.0114 (0.42)
Gini index				-0.131*** (-4.87)
Fuel				0.000827 (0.07)
GDP growth				-0.0293 (-0.54)
Youth bulge				-0.166* (-2.47)
Ethnic fractionalization				-3.530* (-2.09)
Unemployment				-0.103** (-2.98)
Region dummies?	no	yes	yes	yes
Year dummies?	no	no	yes	yes
N	799	799	799	799
F-test	256.13	225.92	62.71	16.50

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Secondary quality is instrumented for using the level of numeracy skills in 1820. Constant, country dummies and year dummies are omitted from the table.

7.4 Extending the analysis of interaction effects between education quality and enrollment

This section presents results from models investigating interaction effects between education quality and quantity on democracy. In brief, it finds no evidence of this. There are no statistically significant negative effects of primary or secondary education quality on democracy when school enrollment is low (see table 7.6). Nor is there evidence that primary or secondary schooling are positively related to democracy when school enrollment is high. This finding is robust to looking at the combined education quality indicator (see table 7.7). Moreover, there is no evidence that education quality is conditioned by school enrollment when looking at democratic transitions or democratic breakdown (see table 7.8).

Table 7.6 replicates Table 2.1 from the paper, but also includes interaction effects between A) primary quality and primary enrollment and B) secondary quality and secondary enrollment.²

²Unfortunately, it was not possible to include these interaction terms in system gmm models, as it would have given too many endogenous regressors for this estimator to perform well.

Table 7.6: Democracy and education quality. Adding interaction effects between quality and enrollment.

	1 OLS	2 OLS	3 OSL	4 OLS	5 OLS	6 OLS
Education quality						
Primary quality	0.0613 (1.56)		-0.0697 (-0.37)	-0.0867 (-0.42)	-0.122 (-0.65)	-0.215 (-1.05)
Primary enrollment		0.00419 (0.48)	-0.0367 (-0.59)	-0.0480 (-0.73)	-0.0726 (-1.18)	-0.0918 (-1.39)
Primary quality*enrollment			0.00132 (0.72)	0.00112 (0.56)	0.00204 (1.11)	0.00233 (1.17)
Secondary quality	-0.00738 (-0.21)		-0.0440 (-0.54)	-0.0550 (-0.58)	-0.0193 (-0.23)	0.00645 (0.07)
Secondary enrollment		0.00331 (0.39)	-0.0272 (-0.64)	0.0332 (0.60)	-0.00915 (-0.22)	0.0207 (0.38)
Secondary quality*enrollment			0.000546 (0.62)	-0.000644 (-0.58)	0.000356 (0.41)	-0.00137 (-1.22)
Controls						
Democracy	0.945*** (207.32)	0.946*** (215.24)	0.945*** (203.40)	0.849*** (73.23)	0.934*** (164.34)	0.831*** (68.05)
GDP(log)	0.395** (2.63)	0.511*** (3.61)	0.408+ (1.93)	1.084*** (3.38)	0.580* (2.13)	-0.141 (-0.35)
Urbanization				0.0753* (1.98)	-0.0136 (-1.28)	0.0788* (2.12)
Gini index				0.0384 (1.18)	0.0667* (2.47)	0.0615+ (1.89)
Fuel				0.0104 (0.88)	-0.0197** (-3.17)	0.00931 (0.78)
Country dummies?	no	no	no	yes	no	yes
Year dummies?	no	no	no	no	yes	yes
N	4876	4876	4876	4876	4876	4876
r2	0.923	0.923	0.923	0.927	0.925	0.930

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index (Marshall, 2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.7: Democracy and education quality. Adding interaction effects between quality and years of schooling. Combined education quality indicator.

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 system gmm	8 system gmm
Education quality								
Education quality	0.0382+ (1.66)		0.0854 (1.27)	0.0430 (0.54)	0.0614 (0.90)	0.0512 (0.64)	-0.0924 (-1.07)	-0.0302 (-0.35)
Years of schooling		0.156+ (1.70)	0.469 (1.30)	0.513 (1.18)	0.263 (0.72)	0.332 (0.75)	0.384 (0.78)	0.517 (1.06)
Quality*Quantity			-0.00830 (-1.08)	-0.0123 (-1.33)	-0.00282 (-0.37)	-0.0139 (-1.44)	0.0120 (1.12)	0.00439 (0.41)
Controls								
Democracy	0.945*** (206.87)	0.943*** (184.42)	0.943*** (186.15)	0.849*** (73.00)	0.934*** (155.55)	0.833*** (68.34)	0.800*** (97.62)	0.791*** (93.13)
GDP (log)	0.427** (2.97)	0.320+ (1.81)	0.290+ (1.65)	1.113*** (3.63)	0.522* (2.01)	-0.256 (-0.62)	1.276*** (4.27)	0.667+ (1.95)
Urbanization				0.0732* (1.99)	-0.0126 (-1.22)	0.0630+ (1.72)	0.0664** (2.70)	0.0826*** (3.34)
Gini index				0.0337 (1.07)	0.0463+ (1.94)	0.0560+ (1.79)	0.0952** (3.12)	0.0997** (3.27)
Fuel				0.0127 (1.06)	-0.0194** (-3.08)	0.0116 (0.96)	-0.0826*** (-6.71)	-0.0794*** (-6.39)
Country dummies?	no	no	no	yes	yes	no	no	
Year dummies?	no	no	no	no	yes	no	yes	
N	4876	4876	4876	4876	4876	4876	4873	4873
r2	0.923	0.923	0.923	0.927	0.925	0.930		
chi2							17227.5	17653.6

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index (Marshall, 2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.8: Distinguishing between democratic transitions and democratic survival. Interacting education quality with enrollment. Combined education quality indicator.

	1 logit	2 logit	3 logit	4 logit	5 logit	6 logit
Education quality						
Education quality	0.0202 (1.27)		-0.0893 (-1.39)	-0.132 (-1.42)	-0.00635 (-0.07)	-0.136 (-1.05)
Democracy*Quality	-0.0757 (-1.01)		0.371* (2.23)	0.372+ (1.76)	0.270 (1.23)	0.351+ (1.71)
School enrollment		0.00818+ (1.76)	-0.0410 (-1.37)	-0.0342 (-0.75)	-0.0121 (-0.27)	-0.0586 (-0.90)
Democracy*enrollment	-0.00372 (-0.14)		0.205*** (3.46)	0.243** (2.96)	0.209* (2.55)	0.191* (2.32)
Quality*enrollment			0.00121 (1.64)	0.00170 (1.59)	0.000244 (0.22)	0.000473 (0.30)
Democracy*quality*enrollment			-0.00533*** (-3.35)	-0.00580* (-2.56)	-0.00513* (-2.25)	-0.00474* (-2.10)
Controls						
Democracy	4.995 (1.42)	3.217 (1.28)	-12.10* (-1.99)	0.526 (0.06)	-8.230 (-0.79)	-7.651 (-0.89)
GDP(log)	0.779*** (5.23)	0.722*** (8.35)	0.726*** (8.17)	1.982*** (8.61)	0.931*** (5.67)	1.947*** (5.49)
Democracy*GDP	0.709 (1.60)					
Democracy*GDP		0.576 (1.35)	0.750 (1.63)	-1.141* (-2.01)	1.557+ (1.90)	0.0926 (0.15)
Urbanization				0.0726*** (3.39)	-0.0186** (-3.21)	0.127*** (4.96)
Democracy*Urban				0.0317 (1.02)	0.0000716 (0.00)	0.0179 (0.60)
Gini index				0.0123 (0.82)	0.00727 (0.47)	0.0449* (2.04)
Democracy*Gini				-0.0559 (-1.15)	-0.138* (-2.30)	-0.0247 (-0.57)
Fuel				-0.0104* (-2.02)	-0.0292*** (-6.38)	-0.0296** (-3.09)
Democracy*Fuel				0.00229 (0.17)	0.0160 (0.98)	0.00558 (0.31)
Country dummies?	no	no	no	yes	no	yes
Year dummies?	yes	yes	yes	yes	yes	yes
N	4876	4876	4876	3365	4312	2026
ll	-820.3	-821.1	-813.3	-583.5	-478.8	-293.2

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the dichotomous democracy indicator from Boix, Miller and Rosato (2013).

Constant, country dummies and year dummies are omitted from the table.

This section also presents results from additional models looking at authoritarian breakdown as dependent variable, to check the robustness of the results discussed in section 2.5.3 in the paper. All models use the Geddes, Wright and Frantz (2014) measure as dependent variable, and I include indicators of education quality as well as interaction terms between education quality and quantity as independent variables. I present models where the two primary and secondary education quality indicators are replaced by the average education quality indicator (see table 7.9), and several additional model specifications including the primary and secondary education quality indicators (see tables 7.10 and 7.11). In brief, the results are are in line with the results presented in the paper. Primary education quality is negative (but not significant) when combined with low enrollment rates in table 7.10, but secondary education quality is negative and statistical significance when combined with low enrollment rates in table 7.11. Moreover, the combined education quality index is also negatively associated with authoritarian breakdown when education enrollment is low (see table 7.9).

Table 7.9: Education quality and authoritarian breakdown. Combined average education quality indicator.

	1 logit	2 logit	3 logit	4 logit	5 logit	6 logit	7 logit
Education quality							
Education quality	-0.00371 (-0.23)		-0.102* (-1.97)	-0.107* (-2.04)	-0.461** (-3.21)	-0.0872 (-1.55)	-0.122 (-1.30)
School enrollment		0.00503 (0.76)	-0.0367 (-1.37)	-0.0342 (-1.17)	-0.194* (-2.24)	-0.0279 (-0.89)	-0.0753 (-1.13)
Quality*enrollment			0.00116+ (1.74)	0.00111 (1.52)	0.00428* (2.07)	0.000797 (1.01)	0.000949 (0.70)
Controls							
GDP	-0.0000361 (-1.00)	-0.0000515 (-1.35)	-0.0000537 (-1.40)	-0.0000150 (-0.31)	0.0000702 (0.31)	-0.0000343 (-0.72)	0.000242 (1.15)
Regime duration	0.168* (2.14)	0.165* (2.12)	0.163* (2.09)	0.173* (2.28)	0.289+ (1.83)	0.139 (1.34)	0.339* (2.17)
Regime duration ²	-0.0132* (-2.39)	-0.0131* (-2.38)	-0.0131* (-2.38)	-0.0134* (-2.48)	-0.0200 (-1.26)	-0.0115+ (-1.70)	-0.0207+ (-1.75)
Regime duration ³	0.000249* (2.50)	0.000251* (2.50)	0.000251* (2.49)	0.000245* (2.44)	0.000692 (1.40)	0.000187 (1.61)	0.000639* (2.38)
Unemployment				-0.00637 (-0.23)	0.0394 (0.82)	-0.00973 (-0.31)	0.0569 (1.55)
Urbanization				-0.000865 (-0.11)	0.120 (1.52)	0.00651 (0.83)	-0.0242 (-0.22)
Gini index				-0.00112 (-0.06)	-0.00590 (-0.20)	-0.0100 (-0.56)	-0.0121 (-0.44)
Fuel				-0.0136* (-2.04)	-0.0210 (-1.60)	-0.0179* (-2.40)	-0.0137 (-0.64)
GDP growth				-0.0479** (-3.24)	-0.0940*** (-4.44)	-0.0455** (-2.84)	-0.0505+ (-1.95)
Country dummies?	no	no	no	no	yes	no	yes
Year dummies?	no	no	no	no	no	yes	yes
N	1911	1911	1911	1911	1067	1804	1358
ll	-466.5	-466.0	-464.5	-453.4	-203.4	-402.3	-258.7

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the indicator of autocratic regime breakdown from Geddes, Wright and Frantz (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.10: Primary education quality and authoritarian breakdown. Interacting education quality with primary enrollment.

	1 logit	2 logit	3 logit	4 logit	5 logit	6 logit	7 logit
Education quality							
Primary quality	0.0251 (1.17)		-0.0780 (-0.86)	-0.0738 (-0.81)	-0.101 (-1.06)	-0.0528 (-0.51)	-0.0625 (-0.53)
Primary enrollment		0.00140 (0.25)	-0.0332 (-1.05)	-0.0355 (-1.08)	-0.0820* (-2.12)	-0.0265 (-0.74)	-0.0812+ (-1.86)
Primary quality*enrollment			0.00105 (1.08)	0.00106 (1.07)	0.00115 (1.06)	0.000704 (0.65)	0.000754 (0.57)
Controls							
GDP	-0.0000503 (-1.38)	-0.0000396 (-1.14)	-0.0000501 (-1.34)	-0.00000682 (-0.14)	0.000142 (1.05)	-0.0000308 (-0.65)	0.000153 (0.74)
Regime duration	0.171* (2.16)	0.167* (2.13)	0.169* (2.13)	0.177* (2.28)	0.333** (2.63)	0.141 (1.37)	0.315* (2.00)
Regime duration ²	-0.0133* (-2.39)	-0.0131* (-2.38)	-0.0132* (-2.35)	-0.0133* (-2.42)	-0.0227* (-2.31)	-0.0114+ (-1.70)	-0.0167 (-1.44)
Regime duration ³	0.000253* (2.49)	0.000248* (2.49)	0.000249* (2.44)	0.000242* (2.36)	0.000814** (3.12)	0.000185 (1.60)	0.000558* (2.18)
Unemployment				0.00768 (0.29)	0.0391 (1.11)	0.000725 (0.02)	0.0738+ (1.95)
Urbanization				-0.00237 (-0.31)	0.105 (1.55)	0.00361 (0.49)	-0.0508 (-0.51)
Gini index				0.0157 (0.89)	0.0191 (0.78)	0.00572 (0.33)	0.00692 (0.26)
Fuel				-0.0117+ (-1.86)	-0.0176 (-1.00)	-0.0163* (-2.22)	-0.0152 (-0.65)
GDP growth				-0.0453** (-3.10)	-0.0395* (-2.09)	-0.0432** (-2.77)	-0.0473* (-2.03)
Country dummies?	no	no	no	no	yes	no	yes
Year dummies?	no	no	no	no	no	yes	yes
N	1911	1911	1911	1911	1425	1804	1358
ll	-465.5	-466.5	-464.8	-454.1	-295.1	-402.9	-254.8

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the indicator of autocratic regime breakdown from Geddes, Wright and Frantz (2014). country dummies and year dummies are omitted from the table.

Table 7.11: Secondary education quality and authoritarian breakdown. Interacting education quality with secondary enrollment

	1 logit	2 logit	3 logit	4 logit	5 logit	6 logit	7 logit
Education quality							
Secondary quality	-0.0185 (-0.89)		-0.0750+ (-1.84)	-0.0822* (-2.15)	-0.287*** (-3.82)	-0.0810+ (-1.91)	-0.168** (-3.06)
Secondary enrollment		0.00648 (1.35)	-0.0199 (-0.70)	-0.0144 (-0.48)	-0.158+ (-1.65)	-0.0252 (-0.81)	-0.0615 (-1.03)
Secondary quality*enrollment			0.000697 (1.10)	0.000610 (0.93)	0.00403+ (1.91)	0.000708 (1.02)	0.00231+ (1.91)
Controls							
GDP	-0.0000308 (-0.86)	-0.0000657 (-1.64)	-0.0000683+ (-1.73)	-0.0000296 (-0.61)	0.0000377 (0.16)	-0.0000454 (-0.95)	0.000182 (0.82)
Regime duration	0.167* (2.14)	0.167* (2.13)	0.163* (2.11)	0.171* (2.29)	0.302* (2.36)	0.137 (1.32)	0.383* (2.32)
Regime duration ²	-0.0131* (-2.40)	-0.0133* (-2.40)	-0.0134* (-2.44)	-0.0135* (-2.51)	-0.0210 (-1.49)	-0.0114+ (-1.69)	-0.0251+ (-1.90)
Regime duration ³	0.000249* (2.51)	0.000257* (2.52)	0.000260** (2.58)	0.000249* (2.48)	0.000657 (1.58)	0.000189 (1.63)	0.000763* (2.44)
Unemployment				-0.0172 (-0.61)	0.0432 (0.73)	-0.0159 (-0.52)	0.0404 (1.03)
Urbanization				-0.00181 (-0.22)	0.0525 (0.70)	0.00546 (0.66)	-0.103 (-0.94)
Gini index				-0.00426 (-0.23)	-0.00536 (-0.17)	-0.0129 (-0.68)	-0.0253 (-0.95)
Fuel				-0.0144* (-2.17)	-0.0267* (-2.02)	-0.0180* (-2.45)	-0.0135 (-0.63)
GDP growth				-0.0445** (-2.94)	-0.0853*** (-4.20)	-0.0432** (-2.61)	-0.0443+ (-1.72)
Constant	-2.082* (-2.16)	-3.051*** (-8.27)	0.0263 (0.02)	0.870 (0.45)	2.142 (0.48)	1.258 (0.60)	0.0850 (0.02)
Country dummies?	no	no	no	no	yes	no	yes
Year dummies?	no	no	no	no	no	yes	yes
N	1911	1911	1911	1911	1067	1804	1358
ll	-466.0	-465.4	-463.0	-451.7	-216.0	-401.5	-257.1

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the indicator of autocratic regime breakdown from Geddes, Wright and Frantz (2014). Constant, country dummies and year dummies are omitted from the table.

7.5 Details about the multiple imputation procedure

7.5.1 Imputation model and diagnostics

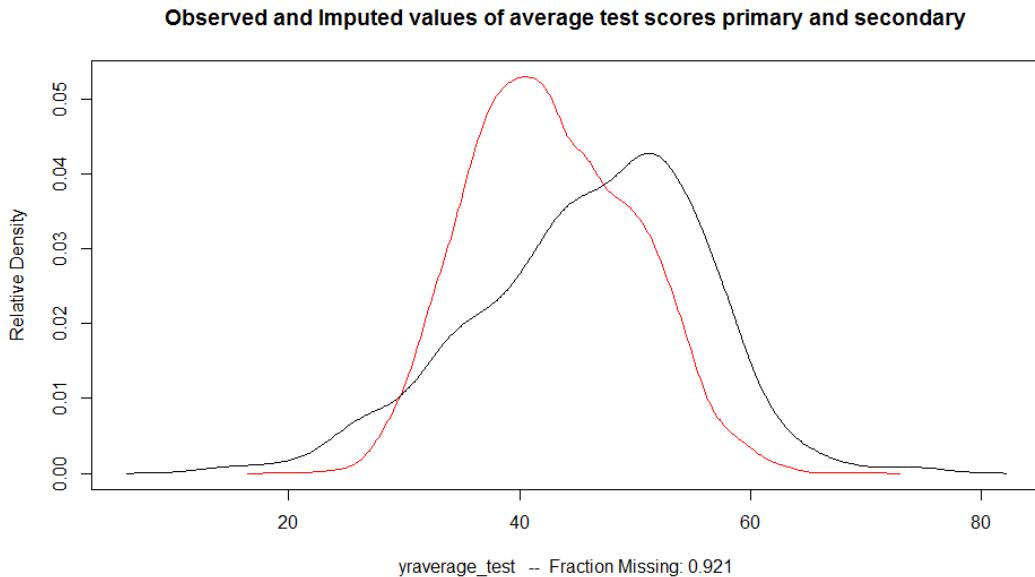
I employed the Amelia II software (Honaker and King, 2010; Honaker et al., 2011) to construct the imputation model. 5 different data sets were produced by the imputation, and all 5 were used in the empirical analysis of the paper. The imputation model allowed for country-specific (second-order polynomial) time trends. Following standard recommendations, I included a wide range of relevant variables (see below) in order to improve the predictive power of the multiple imputation model as much as possible,

Multiple imputation builds on an assumption that the data are missing at random (MAR). This means that the pattern of “missingness” only depends on the observed data included in the imputation model, and not on unobserved variables or features with the data (Honaker and King, 2010). The pattern of missingness on many of the key variables used in the analysis will inevitably depend on a number of measurable factors. For instance, the extent to which governments will allow student achievement tests to be carried out in their country could depend on how democratic it is, and poorer countries are less likely to have participated in student achievement test due to organizational and capacity issues. To account for this, and to attempt to reduce the threat of the MAR-assumption being grossly violated, I included a wide range of variables in the imputation model. For instance, I included various indicators of democracy, indicators of human rights, as well as numerous socio-economic indicators such as economic growth, GDP per capita, and urbanization.

To evaluate the outputs from the imputation model I conducted different imputation diagnostic tests to assess its quality. If the MAR-assumption is still not fully met, despite the measures described above, the imputation model may perform poorly. There are several ways in which the outputs from the imputation model can be described in a condensed manner. One way, which can potentially also be used as a check on the plausibility of the imputation model specification, is to map the distribution of imputed values and the distribution of observed values for particular variables. One should not necessarily expect the distributions of the missing values to be completely identical to the distributions of the observed values, but very large discrepancies between the two suggest that the imputation model is problematic, for instance due to violations of the assumptions discussed above. However, the main reason I imputed to begin with is based on the assumption that observed and missing values could differ systematically; correcting for this alleviates the issue of selection biases affecting our regression results

Figure 7.1 shows the distributions of imputed and observed values for the average

Figure 7.1: Distributions of observed and imputed observations for aggregated measure of student achievement test scores



student achievement test measure from Angrist (2013). The red line represents the density of the mean of each imputed observation across all 5 datasets, and the black line represents the density of the observed values. The shape of the two distributions follow a roughly similar pattern, but still deviate especially when it comes to the lower scores on student achievement tests, where there are higher relative frequencies for the imputed data. As noted, it is plausible to expect that poorer countries will have more missing values on the student achievement test measure than rich countries, suggesting that this plot is plausible. Indeed, the distribution of the imputed values for the aggregated student achievement test measure may actually indicate some of the selection biases discussed above is alleviated by the imputation.

Second, I present results from the so-called overimputation procedure (Honaker et al., 2011), that may help to assess the performance of the imputation procedure. Since missing data are, by definition, unobservable, it is impossible to tell whether each imputed value is close to the unobserved value that we would like the imputation model to predict accurately. An overimputation test is a diagnostics test that attempts to solve this problem by (sequentially) treating each observed value *as if* it is actually missing, and then generate a large number of imputed values for it. Based on this large set of imputed values, I can then construct a confidence interval, and the final step is then assessing whether the actually observed value falls within this interval. If it does, the imputation model performs well in the sense that it is able to accurately predict this “fictive missing

“data” point (that we *know* the actual score on).

This diagnostics test is offered by the Amelia II software, as well as a simple way of graphically inspecting whether our observed data tends to fall within the region where it is imputed by the mode, Figure 7.2 plots the results for the overimputation test for the average student achievement test score. The observed test scores are plotted against the mean of the imputed scores for the very same observation on this index (when the observed test score is treated as if it is missing). The 90 percent confidence intervals that are plotted for each imputed value thus allow for visually inspecting the performance of the imputation model. As noted by (?), checking how many of the confidence intervals cover the $y = x$ line lets us tell how often the imputation model can confidently predict the true value of the observation. The line colors on the confidence intervals tell the fraction of missing observations for all covariates entered in the imputation model for that particular observation.

Figure 7.2 indicates that the imputation model performs well when it comes to predicting the observed values. Every single confidence interval, except from two, covers the 45 degree line. The predictions are especially good for the medium to higher values of the student achievement index. However, the imputation model is worse at predicting the observed values at the lowest level of the student achievement test score measure, but it should be noted that there are very few observations in this area.

Figure 7.2: Overimputation test plot, aggregated measure of student achievement test scores

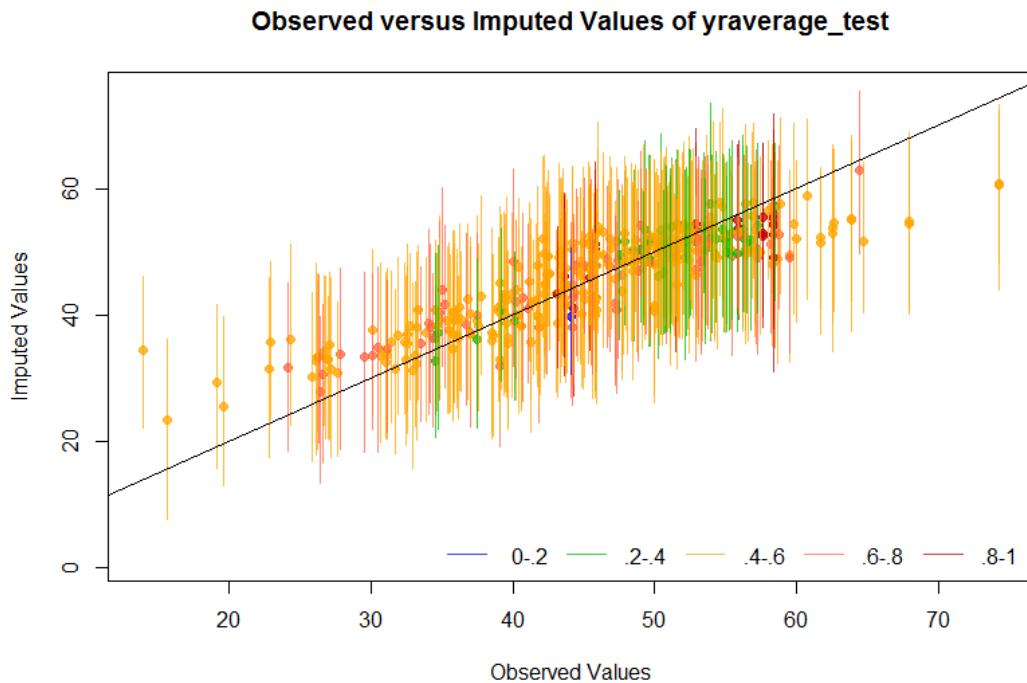


Figure 7.3: Overimputation test plot, student achievement test measure for primary school students

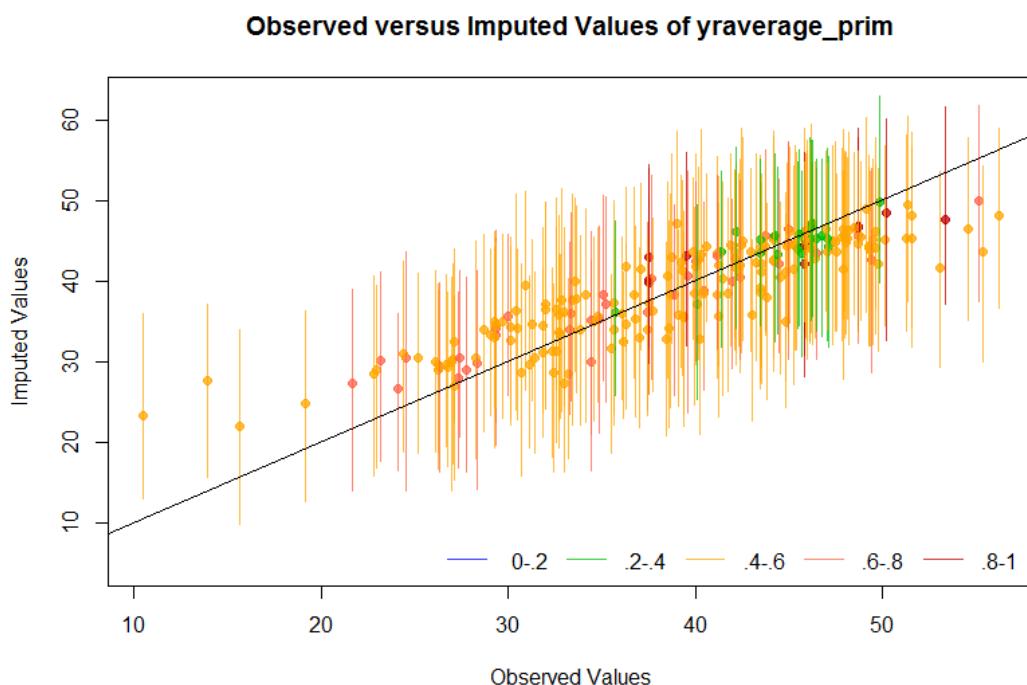
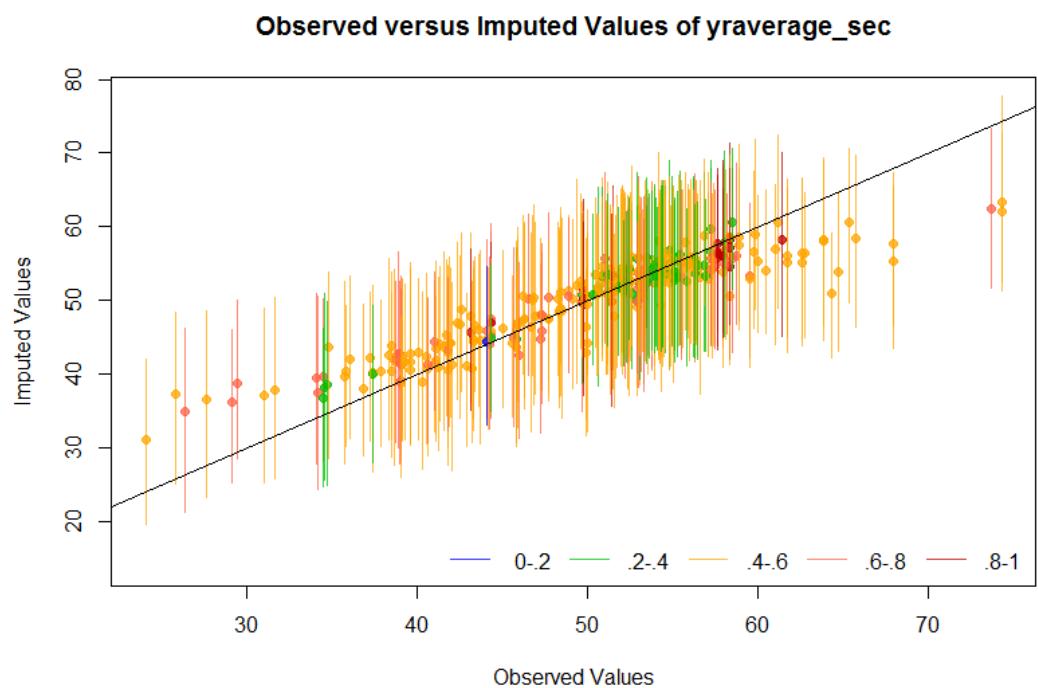


Figure 7.4: Overimputation test plot, student achievement test measure for secondary school students



7.5.2 Assessing the sensitivity of the models to the multiple imputation procedure

This section starts by presenting results from models estimated on the unimputed datasets. As the education quality data only comes in 5-year panels, I estimate this in a 5-year set up. Moreover, as I have far fewer degrees of freedom when estimating on the unimputed data, I only include the most essential controls – lagged polity, years of education and log gdp, in addition to year and country dummies. Second, I present models based on standard errors corrected for the multiple imputation procedure. Finally, I also present results from models investigating the relationship between education type and autonomy on democracy, estimated on the entire time-series that is also used to estimate models for the relationship between education quality and democracy. As discussed in the paper, I restricted my main analysis to looking at the years and countries which had at least one observation on the original indicators for education type and autonomy. However, to be able to assess whether the results in the paper are driven by selection bias, I also run estimations on an expanded sample, despite the fact that the multiple imputation is much more insecure for observations where little original data is present.

Table 7.12: Democracy and education quality. 5-year panel.

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS (FE)
Polity Index	0.902*** (21.31)	0.868*** (49.62)	0.913*** (19.06)	0.891*** (16.59)	0.922*** (22.58)	0.909*** (21.23)
Education quality	0.00480 (0.24)		0.00373 (0.21)	-0.00973 (-0.59)	-0.000600 (-0.03)	-0.0103 (-0.45)
Ln GDP p.c.	-0.117 (-0.42)	-0.203+ (-1.71)	-0.442* (-2.05)	-0.142 (-0.51)	-0.487* (-2.29)	-0.158 (-0.55)
Years of education		0.266*** (4.98)	0.0937 (1.00)	0.0141 (0.17)	0.0920 (1.01)	-0.0103 (-0.12)
Country dummies?	no	no	no	yes	no	yes
Year dummies?	no	no	no	no	yes	yes
N	197	981	190	190	190	190
r2	0.862	0.851	0.878	0.886	0.892	0.899

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 5 years. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.13: Democracy and education type. Unimputed dataset.

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6
Democracy	0.979*** (69.22)	0.986*** (63.85)	1.032*** (59.31)	1.048 (.)	1.032*** (59.31)	1.048 (.)
Engineering	0.00291 (0.38)	0.00892 (1.39)	-0.0189+ (-1.93)	-0.0437 (.)	-0.0189+ (-1.93)	-0.0437 (.)
Humanities	-0.0101 (-1.00)	-0.00438 (-0.76)	0.0286* (2.05)	0.0577 (.)	0.0286* (2.05)	0.0577 (.)
Social science	-0.00501 (-0.74)	0.0121 (1.41)	-0.00904 (-0.97)	-0.0278 (.)	-0.00904 (-0.97)	-0.0278 (.)
Science	-0.00136 (-0.09)	-0.0227+ (-1.81)	-0.0212 (-1.58)	-0.0382 (.)	-0.0212 (-1.58)	-0.0382 (.)
GDP (log)	-0.0818 (-1.21)	0.0561 (1.38)	-0.119 (-1.64)	-0.175 (.)	-0.119 (-1.64)	-0.175 (.)
Tertiary graduates		0.828 (1.04)				
Years of schooling			-0.0301 (-0.69)	-0.000979 (.)	-0.0301 (-0.69)	-0.000979 (.)
Country dummies?	no	no	no	yes	no	yes
Year dummies?	no	no	no	no	yes	yes
N	297	39	44	44	44	44
r2	0.975	0.999	0.997	1	0.997	1

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 years. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.14: Democracy and autonomy. Unimputed dataset.

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS
Democracy	0.963*** (97.53)	0.941*** (62.29)	0.912*** (39.89)	0.960*** (82.73)	0.909*** (40.42)
Foreign edu.(autocracy)	-28.30 (-0.87)	-47.61 (-1.32)	-109.8+ (-1.73)	-9.958 (-0.34)	-82.62 (-1.42)
Foreign edu.(democracy)	-34.76 (-1.08)	-41.26 (-1.58)	-140.7+ (-1.73)	-40.11 (-1.38)	-146.7+ (-1.82)
GDP (log)	0.133* (2.13)	0.293** (2.67)	0.269 (1.54)	-0.0440 (-0.74)	-0.0975 (-0.44)
Tertiary graduates		-0.401 (-0.24)			
Years of schooling			0.0657 (0.87)	0.0639 (1.60)	-0.0807 (-0.80)
Country dummies?	no	no	yes	no	yes
Year dummies?	no	no	no	yes	yes
N	905	666	772	772	772
r2	0.950	0.944	0.961	0.956	0.963

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 5 year. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.15: Democracy and education quality. Estimations across all 5 imputed datasets, with MI-corrected standard errors

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 System gmm	8 System gmm
Democracy	0.898*** (81.57)	0.896*** (86.25)	0.895*** (81.48)	0.741*** (36.07)	0.885*** (73.97)	0.710*** (35.18)		
Education quality	0.00529 (0.64)		0.00370 (0.44)	-0.00521 (-0.59)	0.00462 (0.47)	-0.00601 (-0.68)	0.00449 (0.42)	0.00479 (0.53)
GDP(log)	0.191** (3.36)	0.163* (2.49)	0.160* (2.47)	0.348* (2.78)	0.170 (1.51)	-0.102 (-0.54)	0.653*** (6.99)	0.320 (1.81)
Years of education	0.0302 (1.45)	0.0268 (1.31)	0.0107 (0.48)	0.0250 (1.09)	-0.0229 (-1.03)	0.158** (4.33)		0.122* (3.07)
Urbanization			0.0126 (0.98)	-0.000676 (-0.18)	0.00797 (0.61)	0.0266* (2.75)		0.0248* (2.84)
Gini index				0.00680 (0.58)	0.00662 (0.77)	0.0101 (0.97)	0.0127 (1.31)	0.0137 (1.30)
Oil				0.00511 (1.00)	-0.00520+ (-1.90)	0.00496 (1.00)	-0.0236** (-4.47)	-0.0225** (-4.74)
Democracy						0.608*** (31.32)		0.580*** (37.06)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	4803	4803	4803	4803	4803	4803	4803	4803

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.16: Democracy and education type. Estimations across all 5 imputed datasets, with MI-corrected standard errors

	1 OLS	2 OLS	3 OLS	4 OLS
Democracy	0.951*** (79.20)	0.948*** (75.37)	0.930*** (60.72)	0.559*** (7.21)
Engineering	0.00451 (0.42)	0.00364 (0.34)	0.00703 (0.67)	0.00217 (0.13)
Humanities	-0.0140 (-1.32)	-0.0147 (-1.37)	-0.0150 (-1.46)	-0.0136 (-0.99)
Social science	-0.00295 (-0.38)	-0.00273 (-0.35)	-0.00196 (-0.26)	0.00164 (0.17)
Science	-0.00997 (-0.61)	-0.00935 (-0.58)	-0.00772 (-0.47)	-0.00166 (-0.06)
GDP(log)	-0.00118 (-0.02)	-0.0353 (-0.41)	-0.00279 (-0.03)	-0.0466 (-0.30)
Tertiary education		0.803 (0.84)	0.992 (0.98)	0.482 (0.41)
Urbanization			-0.00223 (-0.56)	-0.0473 (-0.96)
Gini index				-0.00361 (-0.60) -0.000327 (-0.04)
Oil				-0.00764* (-2.35) 0.000475 (0.03)
Country dummies?	no	no	yes	yes
Year dummies?	no	no	no	yes
N	1213	1213	1213	1213

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.17: Democracy, education type and education autonomy. Full imputed time series to deal with selection bias.

	B1 OLS	B2 OLS	B3 OLS	B4 OLS	B5 OLS	B6 S.gmm	B7 OLS	B8 OLS	B9 OLS	B10 OLS	B11 OLS	B12 S.gmm
Education type and autonomy												
Engineering	0.0400 (1.24)	0.0354 (1.09)	0.0297 (0.68)	0.0577 (1.61)	0.0404 (0.92)	0.0592 (1.36)						
Humanities	-0.121** (-2.96)	-0.122** (-2.83)	-0.148** (-2.37)	-0.106* (-2.25)	-0.117* (-10.49)	-0.498*** (-10.49)						
Social science	0.0396 (1.63)	0.0415+ (1.71)	0.0479+ (1.66)	0.0528* (2.44)	0.0714* (1.09)	0.0315 (1.09)						
Science	-0.121* (-2.12)	-0.118* (-2.05)	-0.104+ (-1.25)	-0.0626 (-0.94)	-0.219** (-3.06)							
Students abroad(dem.)							-1.653 (-1.24)	-1.841 (-1.56)	0.605 (0.45)	-1.830 (-1.46)	-0.644 (-0.47)	-1.810 (-1.31)
Students abroad(aut.)							0.482	0.601	0.543	0.750	0.613	-2.075
Controls												
Democracy	0.932*** (156.21)	0.931*** (148.12)	0.837*** (66.84)	0.924*** (135.02)	0.821*** (62.67)	0.789*** (103.23)	0.951*** (124.01)	0.948*** (205.08)	0.851*** (68.99)	0.936*** (156.49)	0.831*** (63.68)	0.815*** (102.24)
GDP (log)	0.628*** (4.14)	0.502* (2.54)	1.231*** (3.64)	0.657* (2.43)	-0.0101 (-0.47)	1.372*** (4.47)	0.289 (1.58)	0.0790 (0.39)	0.608+ (1.78)	0.273 (1.06)	-0.475 (-1.17)	1.206*** (3.73)
Tertiary graduates	3.597 (1.04)	-2.244 (-0.52)	5.332 (1.41)	-10.69* (-2.34)	15.26*** (3.35)	5.720+ (1.68)	-0.0282 (-0.01)	7.978* (2.06)	-9.056+ (-1.95)	21.90*** (4.48)	0.0884*** (1.39)	
Urbanization		0.0632+ (1.79)	-0.00991 (-0.97)	0.0506 (1.42)	0.0590** (2.72)	0.0756+ (1.94)	-0.00317 (-0.32)	0.0555 (1.39)	0.0555 (3.60)			
Gini index	0.0251 (0.75)	0.00751 (0.30)	0.0514 (1.53)	0.0129 (0.44)	0.0440 (1.35)	0.0367+ (1.68)	0.0680* (2.10)	0.0706* (2.37)				
Fuel	0.0149 (1.25)	-0.0209*** (-3.51)	0.0157 (1.31)	-0.0770*** (-6.80)	-0.00654 (-0.52)	-0.0280*** (-4.43)	-0.00277 (-0.22)	-0.0874*** (-6.99)				
Country dummies?	no	no	yes	no	no	no	no	no	yes	no	no	
Year dummies?	no	no	yes	yes	no	no	yes	yes	yes	no	no	
N	4876	4876	4876	4876	4873	4410	4410	4410	4410	4410	4870	
r2	0.923	0.923	0.927	0.925	0.930	0.928	0.929	0.933	0.931	0.935		
chisq												17635.2
sargan												3385.4

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year, except from indicators of students abroad, which are lagged by 5 years. Standard errors are clustered on country. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

7.6 Additional robustness tests

7.6.1 Alternative lag structures

Table 7.18: Democracy and education quality (5-year lag)

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS (FE)	7 System	8 System
Education quality	0.241*** (5.63)		0.202*** (4.31)	-0.00648 (-0.12)	0.194*** (3.53)	0.00556 (0.10)	-0.0402 (-0.98)	-0.0196 (-0.48)
GDP (log)	0.898*** (3.32)	0.625+ (1.86)	0.535 (1.60)	4.496*** (9.01)	1.547*** (3.70)	0.456 (0.79)	1.447*** (4.06)	0.801+ (1.94)
Years of schooling		0.698*** (3.86)	0.348+ (1.75)	0.00235 (0.01)	0.348+ (1.67)	-0.471* (-2.18)	0.0939 (0.62)	-0.0351 (-0.22)
Urbanization				0.193*** (3.66)	-0.0562** (-3.09)	0.144** (3.00)	0.0222 (0.79)	0.0238 (0.84)
Gini index				-0.146** (-2.86)	0.0484 (1.11)	-0.0838+ (-1.67)	-0.0783* (-2.43)	-0.0587+ (-1.83)
Fuel				-0.0425* (-2.26)	-0.0967*** (-7.08)	-0.0445* (-2.35)	-0.0455*** (-3.32)	-0.0461*** (-3.35)
Democracy	0.813*** (83.19)	0.808*** (76.68)	0.807*** (76.72)	0.503*** (25.51)	0.776*** (64.13)	0.473*** (23.96)	0.825*** (92.01)	0.807*** (85.83)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	4423	4423	4423	4423	4423	4423	4537	4537
r2	0.760	0.759	0.760	0.816	0.774	0.828		
chi2							12707.2	13538.2

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 5 years. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.19: Democracy and education quality (10-year lag)

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS (FE)	7 System	8 System
Education quality	0.413*** (7.01)		0.360*** (5.56)	0.0922 (1.46)	0.380*** (5.16)	0.0812 (1.34)	0.0223 (0.51)	0.0274 (0.62)
GDP (log)	0.563 (1.48)	0.290 (0.63)	0.103 (0.22)	7.628*** (11.57)	2.062*** (3.79)	-0.240 (-0.30)	1.881*** (4.43)	0.371 (0.72)
Years of schooling		1.081*** (4.44)	0.462+ (1.72)	-0.192 (-0.76)	0.497+ (1.84)	-0.779** (-3.14)	-0.0323 (-0.19)	-0.0519 (-0.30)
Urbanization				0.126* (2.06)	-0.104*** (-4.31)	0.103* (2.00)	0.00722 (0.24)	0.0370 (1.21)
Gini index				-0.0589 (-0.98)	0.200*** (3.69)	-0.00598 (-0.10)	0.0277 (0.82)	0.0326 (0.98)
Fuel				-0.0911*** (-4.08)	-0.180*** (-10.53)	-0.114*** (-5.12)	-0.0126 (-0.89)	-0.0219 (-1.53)
Democracy	0.694*** (54.22)	0.689*** (49.78)	0.687*** (49.88)	0.232*** (11.86)	0.629*** (41.46)	0.192*** (9.82)	0.826*** (84.09)	0.809*** (78.17)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	3851	3851	3851	3851	3851	3851	3968	3968
r2	0.620	0.617	0.620	0.777	0.653	0.799		
chi2							9895.2	10626.6

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 10 years. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.20: Democracy, education type and education autonomy (5-year lag)

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	S.gmm	6 OLS	7 OLS	8 OLS	9 OLS	10 OLS	11 OLS	12 S.gmm
Democracy	0.790*** (67.54)	0.788*** (64.44)	0.491*** (24.11)	0.756*** (56.96)	0.458*** (22.33)	0.789*** (103.23)	0.827*** (28.99)	0.823*** (80.06)	0.495*** (23.25)	0.783*** (63.23)	0.458*** (21.28)	0.830*** (95.53)	
Engineering	0.188** (2.80)	0.183** (2.70)	0.127+ (1.71)	0.257*** (3.66)	0.147* (2.03)	0.0592 (1.36)							
Humanities	-0.186* (-2.40)	-0.187* (-2.41)	-0.142+ (-1.74)	-0.213*** (-2.60)	-0.0970 (-1.20)	-0.498*** (-10.49)							
Social science	0.0807 (1.60)	0.0829 (1.63)	0.0246 (0.49)	0.0956+ (1.90)	0.0786 (1.61)	0.0315 (1.09)							
Science	-0.556*** (-4.61)	-0.554*** (-4.59)	-0.382** (-3.29)	-0.461*** (-3.98)	-0.307** (-2.71)	-0.219** (-3.06)							
GDP (log)	1.854*** (6.31)	1.715*** (4.63)	5.666*** (10.19)	2.313*** (5.21)	1.593*** (2.61)	1.372*** (4.47)	1.077 (1.50)	0.709+ (1.73)	4.629*** (7.65)	1.741*** (3.75)	1.697*** (2.65)	1.406*** (3.70)	
Tertiary graduates	4.122 (0.64)	-25.41*** (-3.47)	4.459 (-0.62)	-40.62*** (-5.45)	15.26*** (3.35)	-19.39** (-2.64)	10.34 (1.59)	-19.39** (1.61)	-34.28*** (-4.50)	9.098+ (1.69)	-0.377* (-2.64)	-0.106+ (0.597)	
Urbanization	0.193*** (3.69)	-0.0446* (-2.49)	0.0390*** (2.91)	0.137*** (2.72)	0.178*** (2.99)	0.0377* (-2.06)							
Gini index	-0.184*** (-3.53)	-0.0683 (-1.56)	-0.1114* (-2.23)	0.0129 (0.44)	-0.227*** (-4.25)	-0.0260 (-0.63)							
Fuel	-0.0418* (-2.24)	-0.107*** (-8.17)	-0.0413* (-2.20)	-0.0770*** (-6.80)	-0.0754*** (-3.60)	-0.123*** (-8.72)	-0.123*** (-3.69)	-0.0601** (-2.86)	-0.0601** (-3.69)	-0.0487*** (-2.25)			
Foreign edu (autocracy)				-730.1* (-2.39)	-758.9*** (-2.75)	-132.9 (-0.59)	-766.2** (-2.66)	-191.1 (-0.83)	-94.48 (-0.49)	-427.6* (-0.83)			
Foreign edu (democracy)				-370.4 (-1.52)	-3 51.0+ (-1.87)	-403.0* (-2.17)	-468.6* (-2.49)	-468.6* (-2.30)	-439.9** (-3.07)				
Country dummies?	no	no	yes	no	no	no	no	no	no	no	no	no	
Year dummies?	no	no	no	yes	no	yes	no	yes	yes	no	yes	no	
N	4423	4423	4423	4423	4873	3957	3957	3957	3957	3957	4419		
r ²	0.761	0.761	0.817	0.776	0.830	0.762	0.762	0.825	0.777	0.836			
chi2				2089.4							13496.7		

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 5 years. Standard errors are robust. Dependent variable is the Polity

Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.21: Democracy, education type and education autonomy (10-year lag)

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 System	7 OLS	8 OLS	9 OLS	10 OLS	11 OLS	12 System
Democracy	0.654*** (43.25)	0.632*** (41.50)	0.224*** (10.95)	0.600*** (36.85)	0.182*** (8.85)	0.789*** (103.23)	0.710*** (14.92)	0.706*** (51.20)	0.209*** (11.00)	0.624*** (40.01)	0.154*** (8.10)	0.830*** (88.56)
Engineering	0.221* (2.44)	0.215* (2.35)	0.175+ (1.96)	0.394*** (4.25)	0.152+ (1.78)	0.0592 (1.36)						
Humanities	-0.326** (-3.10)	-0.326** (-3.10)	-0.112 (-1.18)	-0.299** (-2.83)	-0.0568 (-0.62)	-0.498*** (-10.49)						
Social science	0.209*** (3.03)	0.212*** (3.05)	0.0112 (0.19)	0.173** (2.59)	0.0533 (0.93)	0.0315 (1.09)						
Science	-0.965*** (-5.66)	-0.961*** (-5.63)	-0.317* (-2.19)	-0.801*** (-4.95)	-0.212 (-1.55)	-0.219*** (-3.06)	1.360 (1.01)	1.108y*** (2.00)	7.635*** (9.24)	2.923*** (4.75)	0.366 (0.39)	1.747*** (3.94)
GDP (log)	2.178*** (5.70)	2.050*** (4.12)	9.036*** (12.31)	3.239*** (5.60)	1.046 (1.22)	1.372*** (4.47)						
Tertiary graduates	4.091 (0.45)	-39.16*** (-4.30)	9.646 (0.99)	-53.09*** (0.99)	15.26*** (5.89)	15.26*** (3.35)	7.510 (0.83)	-35.68*** (4.11)	17.87+ (1.78)	-49.91*** (5.83)	0.540 (0.09)	
Urbanization	0.134* (2.19)	-0.0826*** (-3.48)	0.0979+ (1.91)	0.0590*** (2.72)			0.178** (2.63)	-0.0642** (2.60)	0.123* (2.12)	-0.0206 (-0.73)		
Gini index	-0.108+ (-1.77)	-0.00466 (-0.08)	-0.0405 (-0.70)	0.0129 (0.44)	0.0129 (-0.70)	0.0129 (0.44)	-0.0778 (-1.31)	-0.0778 (1.31)	0.0679 (-0.27)	-0.0151 (0.27)	0.0341 (1.07)	
Fuel	-0.0945*** (-4.26)	-0.198*** (-12.26)	-0.112*** (-5.05)	-0.0770*** (-6.80)	-0.112*** (-5.05)	-0.112*** (-6.80)	-1.769.0*** (-2.89)	-1.788.1*** (-3.86)	-0.122*** (-4.85)	-0.245*** (-14.49)	-0.125*** (-5.17)	-0.0204 (-1.50)
Foreign edu (autocracy)									16.45 (0.04)	-2142.8*** (-4.59)	-80.10 (-0.23)	-58.38 (-0.23)
Foreign edu (democracy)									-335.5+ (-1.77)	-634.5+ (-2.66)	-385.7* (-2.12)	-64.83 (-0.44)
Country dummies?	no	no	yes	no	yes	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	yes	yes	no	no	no	yes	yes	yes	no
N	3851	3851	3851	3851	4873	3385	3385	3385	3385	3385	3385	3847
r ²	0.624	0.624	0.778	0.656	0.800	0.619	0.620	0.808	0.659	0.829		
chi ²												10174.2

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 10 years. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

7.6.2 Additional controls

Table 7.22: Democracy and education quality. Additional controls.

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 S.gmm	8 S.gmm
Democracy	0.758*** (58.16)	0.760*** (58.30)	0.761*** (58.49)	0.493*** (24.20)	0.753*** (56.90)	0.461*** (22.68)	0.821*** (90.38)	0.801*** (83.75)
Education quality	0.150** (2.77)		0.174** (3.21)	0.0654 (1.19)	0.133* (2.32)	0.0710 (1.31)	-0.00441 (-0.10)	0.0105 (0.25)
GDP (log)	1.798*** (5.06)	2.118*** (5.51)	2.235*** (5.74)	5.130*** (9.36)	2.187*** (4.89)	1.304* (2.18)	1.810*** (4.82)	1.043* (2.44)
BNP growth	-0.0141 (-0.25)	-0.00829 (-0.15)	-0.00636 (-0.11)	0.00448 (0.09)	0.00641 (0.11)	-0.00377 (-0.07)	0.0157 (0.52)	0.00384 (0.12)
Youth bulges	-0.00395 (-0.06)	-0.146* (-2.13)	-0.0510 (-0.68)	0.364*** (3.87)	0.00485 (0.05)	0.577*** (6.05)	0.158* (2.35)	0.215** (3.10)
Ethnic fractionalization	2.224+ (1.73)	1.269 (1.02)	2.126+ (1.65)	17.43*** (3.41)	2.409+ (1.94)	5.592 (1.06)	2.496 (1.02)	-0.461 (-0.18)
Muslim	-0.113*** (-9.73)	-0.122*** (-10.31)	-0.120*** (-10.13)	-0.0373 (-0.93)	-0.111*** (-8.74)	-0.0784* (-2.04)	-0.0301 (-1.59)	-0.0456* (-2.38)
Protestant	-0.0159+ (-1.84)	-0.0160+ (-1.86)	-0.0142 (-1.64)	-0.0322 (-0.69)	-0.0112 (-1.26)	0.00701 (0.16)	-0.0171 (-0.61)	-0.00875 (-0.31)
Unemployment	0.122* (2.32)	0.107* (2.07)	0.143** (2.73)	0.266*** (4.22)	0.122* (2.35)	0.130* (2.14)	0.0961* (2.09)	0.0498 (1.08)
Years of schooling	-0.364+ (-1.67)	-0.485* (-2.20)	-0.485* (-2.20)	0.0401 (0.18)	-0.497* (-2.16)	-0.398+ (-1.80)	0.0828 (0.52)	-0.0614 (-0.37)
Urbanization				0.257*** (4.62)	-0.00772 (-0.41)	0.184*** (3.66)	0.0408 (1.37)	0.0386 (1.31)
Gini index				-0.207*** (-3.78)	-0.119* (-2.19)	-0.171** (-3.22)	-0.103** (-3.03)	-0.0916** (-2.72)
Fuel				-0.0550** (-2.65)	-0.0664*** (-5.01)	-0.0480* (-2.34)	-0.0465** (-3.15)	-0.0412** (-2.78)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	4423	4423	4423	4423	4423	4423	4537	4537
r ²	0.768	0.768	0.768	0.818	0.780	0.830	12764.3	13635.4

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.23: Democracy, education type and education autonomy. Additional controls

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 System gmm	8 System gmm
Democracy	0.922*** (124.22)	0.922*** (123.81)	0.836*** (63.98)	0.917*** (121.21)	0.822*** (60.52)	0.773*** (100.68)	0.932*** (100.86)	0.932*** (137.22)
Engineering	0.0789* (2.32)	0.0784* (2.31)	0.0524 (1.19)	0.0925* (2.51)	0.0682 (1.53)	0.0721+ (1.67)		
Humanities	-0.0816+ (-1.94)	-0.0816+ (-1.94)	-0.147** (-2.81)	-0.0760+ (-1.72)	-0.111* (-2.13)	-0.496*** (-10.54)		
Social science	0.0560* (2.19)	0.0562* (2.19)	0.0541+ (1.90)	0.0692** (2.64)	0.0713* (2.46)	0.0628* (2.19)		
science	-0.0223 (-0.37)	-0.0224 (-0.37)	-0.0817 (-0.22)	-0.0318 (-0.53)	-0.0678 (-1.02)	-0.143* (-2.01)		
GDP (log)	0.777*** (3.63)	0.761*** (3.20)	1.261*** (3.47)	0.790** (2.66)	1.559*** (0.57)	1.559*** (4.98)	0.377 (1.57)	0.336 (1.45)
BNP growth	-0.0727* (-2.48)	-0.0727* (-2.48)	-0.0662* (-2.48)	-0.0653* (-2.15)	-0.0687* (-2.07)	-0.0613* (-2.07)	-0.0694* (-2.27)	-0.0695* (-2.24)
Youth bulges	0.005339 (0.17)	0.00746 (0.22)	0.167** (2.87)	0.00878 (0.20)	0.257*** (4.26)	-0.135* (-2.37)	-0.0157 (-0.42)	-0.00987 (-0.31)
Ethnic frac.	-0.212 (-0.34)	-0.217 (-0.35)	10.62* (2.20)	-0.208 (-0.33)	6.535 (1.34)	-3.396+ (-1.73)	0.141 (0.18)	0.135 (0.24)
Muslim share	-0.0236*** (-3.67)	-0.0238*** (-3.64)	0.0644+ (1.87)	-0.0199** (-2.81)	0.0446 (1.32)	-0.0761*** (-5.38)	-0.0311*** (-4.44)	-0.0309*** (-4.90)
Protestant shar	0.00705 (1.48)	0.00694 (1.46)	0.0001633 (0.00)	0.00923+ (1.82)	0.0138 (0.34)	0.0387+ (1.77)	-0.000408 (-0.07)	-0.000691 (-0.15)
Unemployment	0.0713* (2.12)	0.0716* (2.13)	0.123** (2.58)	0.0671* (1.99)	0.0788+ (1.69)	0.255*** (6.57)	0.0437 (1.40)	0.0447 (1.36)
Tertiary graduates	0.687 (0.18)	1.687 (0.18)	1.804 (0.40)	1.350 (0.32)	-5.185 (-1.12)	11.61* (2.49)	1.798 (0.46)	
Urbanization		0.1117** (3.15)	-0.00547 (-0.53)	0.0981** (2.60)	0.0179 (0.78)			
Gini index		-0.0152 (-0.43)	-0.00343 (-0.11)	-0.000518 (-0.01)	0.0317 (1.03)			
Fuel		-0.00934 (-0.77)	-0.0148* (-2.31)	-0.00523 (-0.43)	-0.0457*** (-3.80)			
Foreign edu (autocracy)					-105.8 (-0.85)	-111.2 (-0.95)		
Foreign edu (democracy)					86.28 (0.63)	89.84 (0.80)		
Country dummies?	no	no	yes	no	yes	no	no	no
Year dummies?	no	no	no	yes	yes	no	yes	no
N	4876	4876	4876	4876	4873	4410	4410	
r2	0.924	0.924	0.928	0.926	0.931	0.929	0.929	0.929

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

7.6.3 Subcomponents of education quality

Table 7.24: Democracy and primary education quality

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 S.gmm	8 S.gmm
Democracy	0.945*** (207.16)	0.943*** (184.42)	0.943*** (184.44)	0.849*** (72.95)	0.933*** (151.99)	0.834*** (68.42)		
Primary quality	0.0562* (2.06)		0.0425 (1.39)	-0.00435 (-0.11)	0.0745* (2.05)	-0.0123 (-0.32)	0.0173 (0.38)	0.00290 (0.06)
GDP (log)	0.391** (2.59)	0.320+ (1.81)	0.295+ (1.67)	1.122*** (3.67)	0.517* (1.97)	-0.237 (-0.58)	1.079*** (3.47)	0.444 (1.28)
Years of schooling	0.156+ (1.70)	0.0955 (0.92)	-0.0392 (-0.32)	0.120 (1.08)	-0.285* (-2.29)	0.936*** (6.49)		0.700*** (4.60)
Urbanization			0.0705+ (1.94)	-0.0132 (-1.28)	0.0608+ (1.68)	0.0544* (2.12)		0.0701** (2.73)
Gini index			0.0469 (1.46)	0.0559* (2.26)	0.0687* (2.13)	0.0885** (2.81)		0.0957** (3.05)
Fuel			0.0133 (1.12)	-0.0190** (-3.14)	0.0120 (1.00)	-0.0842*** (-6.45)		-0.0804*** (-6.10)
Democracy						0.801*** (94.82)		0.788*** (90.20)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	4876	4876	4876	4876	4876	4876	4873	4873
r2	0.923	0.923	0.923	0.927	0.925	0.930		
chi2						15797.5		16230.9

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.25: Democracy and secondary education quality

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 S.gmm	8 S.gmm
Democracy	0.946*** (210.67)	0.943*** (184.42)	0.943*** (184.57)	0.849*** (72.82)	0.934*** (153.63)	0.833*** (68.38)		
Secondary quality	0.0286 (1.15)		0.0116 (0.40)	-0.0872* (-2.07)	0.0305 (0.84)	-0.0804+ (-1.91)	0.0181 (0.40)	0.0397 (0.88)
GDP (log)	0.468** (3.24)	0.320+ (1.81)	0.316+ (1.80)	1.092*** (3.57)	0.531* (2.03)	-0.264 (-0.64)	1.178*** (3.81)	0.488 (1.40)
Years of schooling	0.156+ (1.70)	0.140 (1.32)	-0.0201 (-0.16)	0.162 (1.41)	-0.268* (-2.13)	0.914*** (6.44)		0.698*** (4.63)
Urbanization			0.0787* (2.18)	-0.0112 (-1.11)	0.0677+ (1.88)	0.0632* (2.40)		0.0750** (2.85)
Gini index			0.0290 (0.93)	0.0447+ (1.81)	0.0538+ (1.72)	0.0914** (2.90)		0.101** (3.20)
Fuel			0.00971 (0.82)	-0.0202*** (-3.30)	0.00884 (0.74)	-0.0790*** (-5.96)		-0.0736*** (-5.51)
Democracy						0.801*** (94.55)		0.791*** (90.21)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	4876	4876	4876	4876	4876	4876	4873	4873
r2	0.923	0.923	0.923	0.927	0.925	0.930		
chi2						15858.3		16295.6

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.26: Democracy and subcomponents of primary education quality.

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 S.gmm	8 S.gmm
Democracy	0.942*** (193.39)	0.943*** (184.42)	0.942*** (178.31)	0.850*** (72.51)	0.931*** (147.20)	0.834*** (68.17)		
Primary math skills	0.0895** (2.65)		0.0884** (2.62)	0.0652+ (1.78)	0.113** (2.97)	0.0761* (2.03)	0.0602 (1.54)	0.0667+ (1.69)
Primary reading skills	-0.0459 (-1.07)		-0.0505 (-1.14)	-0.0491 (-1.07)	-0.0146 (-0.32)	-0.0626 (-1.35)	0.122* (2.51)	0.0952+ (1.95)
Primary science skills	0.0563 (1.38)		0.0521 (1.23)	0.00895 (0.21)	0.0673 (1.55)	-0.00515 (-0.12)	0.117* (2.57)	0.132** (2.90)
GDP(log)	0.389* (2.45)	0.320+ (1.81)	0.355+ (1.95)	1.167*** (3.76)	0.547* (2.08)	-0.198 (-0.48)	1.110*** (3.75)	0.534 (1.59)
Years of schooling	0.156+ (1.70)	0.0419 (0.39)		-0.0437 (-0.35)	0.0569 (0.50)	-0.292* (-2.33)	0.779*** (5.78)	0.633*** (4.49)
Urbanization				0.0695+ (1.93)	-0.0150 (-1.45)	0.0609+ (1.70)	0.0260 (1.11)	0.0397+ (1.68)
Gini index				0.0568+ (1.74)	0.0782** (3.03)	0.0794* (2.43)	0.137*** (4.51)	0.139*** (4.59)
Fuel				0.0147 (1.23)	-0.0156** (-2.59)	0.0132 (1.11)	-0.0726*** (-6.03)	-0.0676*** (-5.57)
Democracy							0.800*** (100.17)	0.793*** (97.21)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	4876	4876	4876	4876	4876	4876	4873	4873
r2	0.923	0.923	0.923	0.927	0.925	0.930		
chi2							18111.5	18587.0

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

Table 7.27: Democracy and different subcomponents of secondary education quality

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 S.gmm	8 S.gmm
Democracy	0.942*** (193.39)	0.943*** (184.42)	0.941*** (178.31)	0.847*** (72.51)	0.931*** (147.20)	0.832*** (68.17)		
Secondary math skills	-0.00913 (-0.28)		-0.0109 (-0.33)	-0.0580 (-1.57)	0.0115 (0.32)	-0.0499 (-1.36)	0.00241 (0.06)	0.0150 (0.36)
Secondary reading skills	0.122* (2.28)		0.115* (2.13)	0.0783 (1.36)	0.131* (2.32)	0.0571 (0.99)	0.133* (2.12)	0.110+ (1.75)
Secondary science skills	0.00322 (0.08)		0.00000673 (0.00)	-0.0458 (-1.04)	0.0151 (0.35)	-0.0383 (-0.87)	-0.0941+ (-1.94)	-0.0734 (-1.50)
GDP(log)	0.287+ (1.80)	0.320+ (1.81)	0.237 (1.30)	1.068*** (3.44)	0.452+ (1.69)	-0.267 (-0.65)	1.272*** (4.30)	0.673* (2.01)
Years of schooling	0.156+ (1.70)	0.0557 (0.52)		-0.0378 (-0.30)	0.0816 (0.70)	-0.278* (-2.22)	0.835*** (6.32)	0.645*** (4.63)
Urbanization				0.0754* (2.09)	-0.0159 (-1.59)	0.0647+ (1.79)	0.0533* (2.19)	0.0664** (2.72)
Gini index				0.0329 (1.00)	0.0666* (2.43)	0.0572+ (1.72)	0.103*** (3.31)	0.111*** (3.55)
Fuel				0.0111 (0.93)	-0.0164** (-2.63)	0.0101 (0.84)	-0.0729*** (-6.04)	-0.0674*** (-5.56)
Democracy							0.800*** (99.77)	0.791*** (95.66)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	4876	4876	4876	4876	4876	4876	4873	4873
r2	0.923	0.923	0.923	0.927	0.925	0.930		
chi2							17615.9	18014.0

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

7.6.4 Alternative dependent variable

Table 7.28: Democracy (Freedom House) and education quality

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 S.gmm	8 S.gmm
Freedom House Democracy	0.931*** (174.21)	0.925*** (159.61)	0.925*** (159.85)	0.787*** (62.19)	0.914*** (131.71)	0.785*** (61.44)		
Education quality	0.00182 (1.32)		-0.000654 (-0.41)	-0.00552** (-2.66)	-0.000330 (-0.17)	-0.00517* (-2.44)	-0.00107 (-0.46)	-0.000333 (-0.14)
GDP (log)	0.0459*** (4.73)	0.0237* (2.11)	0.0239* (2.12)	0.0843*** (4.32)	0.0647*** (3.62)	0.0467+ (1.82)	0.0972*** (5.21)	0.106*** (4.95)
Years of schooling		0.0202*** (3.88)	0.0213*** (3.58)	0.0208** (2.71)	0.0245*** (3.81)	0.00903 (1.15)	0.0394*** (4.71)	0.0377*** (4.16)
Urbanization				-0.00203 (-0.85)	-0.00149* (-2.36)	-0.00237 (-0.97)	0.00547*** (3.48)	0.00569*** (3.57)
Gini index					0.00385* (2.05)	0.00254+ (1.83)	0.00423* (2.21)	0.00389* (2.10)
Fuel					0.000711 (0.93)	-0.00117** (-3.27)	0.000913 (1.18)	-0.00573*** (-7.48)
Freedom House Democracy							0.707*** (72.00)	0.709*** (71.72)
Country dummies?	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	no	yes	yes	no	yes
N	4876	4876	4876	4876	4876	4876	4873	4873
r2	0.904	0.904	0.904	0.912	0.906	0.914		
chi2							8466.1	8527.2

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Freedom House democracy score. Constant, country dummies and year dummies are omitted from the table.

Table 7.29: Democracy, education type and education autonomy. Freedom House as DV.

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 S.gmm	7 OLS	8 OLS	9 OLS	10 OLS	11 OLS	12 S.gmm
Freedom House Democracy	0.920*** (150.12)	0.919*** (143.03)	0.777*** (58.67)	0.908*** (121.65)	0.775*** (58.20)	0.720*** (79.94)	0.943*** (96.06)	0.940*** (175.12)	0.806*** (6 1.21)	0.926*** (137.96)	0.801*** (60.12)	0.725*** (77.73)
Engineering	0.0000404 (0.12)	-0.0000404 (0.02)	-0.00164 (-0.63)	0.000465 (0.22)	-0.00104 (-0.40)	0.000824 (0.31)						
Humanities	-0.00656** (-3.13)	-0.00659** (-3.14)	-0.0121*** (-4.17)	-0.00736** (-3.05)	-0.0104** (-3.59)	-0.0257*** (-9.00)						
Social science	0.00104 (0.77)	0.00109 (0.81)	0.00097 (0.61)	0.00129 (0.95)	0.00162 (0.98)	0.000480 (0.28)						
Science	-0.00966** (-2.96)	-0.00950** (-2.90)	-0.00931* (-2.43)	-0.00754* (-2.33)	-0.00861* (-2.27)	-0.00613 (-1.42)						
GDP(log)	0.0552*** (5.88)	0.0507*** (4.11)	0.106*** (0.66)	0.0774*** (5.00)	0.0689*** (4.32)	0.0351** (3.67)	0.0210+ (2.92)	0.0707*** (1.71)	0.0429* (2.05)	0.0205 (2.53)	0.0719*** (0.82)	
Tertiary graduates (Urbanization)												
	-0.002124 (-0.92)	-0.001033+ (-1.66)	-0.00269 (-1.11)	0.00513*** (3.89)	0.00630 (0.25)							
Gini index	0.00181 (0.89)	-0.000392 (-0.27)	0.00241 (1.18)	-0.000552 (-0.31)	0.00166 (2.16)							
Fuel	0.000995 (1.31)	-0.00119*** (-3.52)	0.00128+ (1.66)	-0.00544*** (-7.95)	0.00529 (0.65)							
Foreign edu (autocracy)												
Foreign edu (democracy)												
Country dummies?	no	no	yes	no	no	no	no	yes	no	yes	no	no
Year dummies?	no	no	no	yes	no	no	no	yes	no	yes	yes	no
N	4876	4876	4876	4876	4873	4410	4410	> 4410	4410	4410	4410	4870
r ²	0.905	0.905	0.913	0.906	0.914	0.918	0.918	0.925	0.920	0.927	0.927	9877.8

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Freedom House democracy score. Constant, country dummies and year dummies are omitted from the table.

7.6.5 5-year panel

In the table below, the models from Table 2.1 in the paper are estimated on a 5-year panel, rather than on a 1-year panel. The main findings from the paper are robust to these estimations: I identify no effect of education quality on democracy.

Table 7.30: Democracy and education quality. 5-year panels

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS	7 S.gmm	8 S.gmm
Polity Index	0.902*** (21.31)	0.868*** (49.62)	0.913*** (19.06)	0.743*** (6.27)	0.892*** (18.34)	0.760*** (7.30)		
Education quality	0.00480 (0.24)		0.00373 (0.21)	0.00247 (0.10)	0.00293 (0.11)	0.0182 (0.57)	0.0106 (0.71)	0.0277 (1.10)
GDP (log)	-0.117 (-0.42)	-0.203+ (-1.71)	-0.442* (-2.05)	1.138 (0.62)	-0.439+ (-1.81)	1.970 (0.98)	0.340 (1.07)	0.211 (0.54)
Years of education		0.266*** (4.98)	0.0937 (1.00)	-0.135 (-0.34)	0.139 (1.55)	-0.509 (-1.12)	-0.164 (-1.35)	-0.212 (-1.19)
Oil				-0.593 (-1.46)	-0.0259 (-0.45)	-0.865+ (-1.97)		0.178+ (1.71)
Gini index					-0.0124 (-0.27)	0.00815 (0.36)	-0.0582 (-1.17)	-0.0240 (-0.78)
L.Polity Index							0.833*** (21.38)	0.821*** (19.73)
Country dummies?	no	no	no	yes	no	yes	no	no
Period dummies?	no	no	no	no	yes	yes	no	yes
N	197	981	190	182	182	182	279	262
r2	0.860	0.850	0.876	0.877	0.868	0.894		

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 5 years. Standard errors are robust. Dependent variable is the Polity Democracy index Marshall (2014). Constant, country dummies and year dummies are omitted from the table.

7.6.6 Education quality and democracy without a lagged dependent variable

Below, Table 2.1 from the paper is replicated *without* controlling for lagged democracy. It should be noted that this is *not* really a test of the hypothesis that education quality should promote democratization, since these models operate with levels rather than changes democracy as dependent variable. Yet, the fact that even these more generous models do not identify a consistent effect of education quality on democracy, indicates that the null-findings from the paper are not merely driven by too strict assumptions about annual effects of education quality on democracy. In brief, the coefficient estimate for education quality is positive and significant in only two out of five models (1 and 5), the first of these models only control for log gdp, and the other does not include country fixed effects. Meanwhile, no effect of education quality on democracy is identified when controlling for years of education in model 3, and the coefficient estimate for education quality is actually *negative* and significant in two out of five models (4 and 6).

Table 7.31: Democracy and education quality. Without lagged democracy.

	1 OLS	2 OLS	3 OLS	4 OLS	5 OLS	6 OLS
Education quality	1.007*** (12.64)		0.0901 (1.12)	-0.115+ (-1.88)	0.503*** (5.72)	-0.116+ (-1.95)
GDP(log)	11.34*** (22.90)	2.789*** (4.32)	2.754*** (4.24)	6.334*** (11.91)	9.585*** (14.81)	1.485** (2.59)
Years of schooling		7.094*** (26.25)	6.932*** (23.69)	3.132*** (13.73)	5.826*** (19.88)	1.673*** (7.01)
Urbanization				0.443*** (8.39)	-0.327*** (-10.63)	0.369*** (7.50)
Gini index				0.604*** (10.69)	1.078*** (17.64)	0.654*** (11.98)
Fuel				-0.160*** (-7.00)	-0.320*** (-15.97)	-0.149*** (-6.59)
Country dummies?	no	no	no	yes	no	yes
Year dummies?	no	no	no	no	yes	yes
N	4876	4876	4876	4876	4876	4876
r2	0.266	0.347	0.347	0.753	0.465	0.775

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All explanatory variables are lagged by 1 year. Standard errors are robust. Dependent variable is the Polity democracy score. Constant, country dummies and year dummies are omitted from the table.

8 Appendix to Chapter 3

8.1 Introduction

In this appendix we expand on information that is relevant for our empirical analysis, but that was presented in a condensed manner in the paper due to space constraints. Examples are discussions on the data material used for our main dependent variable and information about our multiple imputation model. Further, we present and discuss additional tests assess the robustness of the main findings from the paper. We also report and discuss more nuanced empirical evidence pertaining to mechanisms suggested by our theoretical argument and various extensions of the core results, e.g. considering more disaggregated education outcomes.

The more specific sequence of the sections in this Appendix is as follows: *First*, we provide a more detailed description and discussion of how the education quality data are constructed. *Second*, we present and discuss tests pertaining to the mechanisms highlighted by our argument. *Third*, we present additional details about our imputation model. We also discuss and present a number of tests to assess the performance of our imputation model.

Fourth, we include tables with robustness tests of the core models presented in the paper, which indicated that democracy is systematically related to education quantity. *Fifth*, we discuss and evaluate our instrumental variable(s) and present first-stage regressions from the core 2SLS specification as well as second-stage regressions from alternative 2SLS specifications. *Sixth*, we present a number of other robustness tests on democracy and education quality. *Seventh*, we report various tests on democracy and education quality including imputed data.

Eighth, we present results from a mediation analysis probing whether education quality

has an effect on economic growth *through education quality*. *Ninth*, we report and discuss panel regression models on democracy and more disaggregated student achievement test measures pertaining to particular subjects and levels of education. *Tenth*, and finally, we present our extensions assessing different kinds of variability in education quality in democracies and in autocracies.

8.2 The education quality data

This section provides more details on the data used for constructing our main dependent variable, education quality. The measure that we use to capture this concept comes from the dataset on student achievement tests from Angrist (2013). This dataset covers, in total, 128 countries, including more than 40 countries from the developing world. The data are reported in 5-year intervals. The first observation is from 1965 and the final from 2010, but there is no full coverage over the whole period for any country, and some countries have better coverage than others (we note here that we deal with the issue of missing data in analyses where we conduct multiple imputation; see section A.3 for specifics). We note that no countries have test scores available from 1975 in the Agrist dataset, due to the paucity of international students achievements tests.

Since many countries, especially in the developing world, are not included in international tests such as PISA and TIMSS, regional student achievement test scores are linked to the international tests by Angrist (2013) through using the method proposed by Altinok and Murseli (2007): More specifically, this means that each regional test is linked to the international tests by using countries that participated in both regional and international tests as reference points. Moreover, different international tests are linked together using the United States as reference point. The reason for this is that the United States has participated in every international achievement test conducted during the past half century. In practical terms, this means that each country's test score is expressed relative to the test score result for the United States result in a given year. If countries participated in several different achievement tests in the same 5-year interval, the average between these (normalized) test scores is included in the data set.

All the international and regional achievement tests are conducted on students enrolled in either primary or secondary school, although the specific age/grade levels that are included vary from survey to survey (see the list provided directly below). The age/grade level ranges from second grade students at the lowest (surveyed in the African Programme d'Analyses des Systemes Educatifs de la Confemena) to 15 and 16-year olds at the highest (the latter are surveyed in the PISA assessment). This means that although all the test scores are taken from primary and secondary school students, the different regional and international tests may not measure the same grade in a given year. According to Angrist (2013) this should not bias the results, however, as all the different test scores are transformed with the help on an index that accounts for varying scales of different tests as well as varying difficulty among different tests. We refer to the paper by Angrist (2013) for the specific formulae used for aggregation.

As discussed in the paper, the dataset includes specific measures on test scores for

maths, natural sciences, and reading, and such disaggregated measures are provided for both primary and secondary level students. We mainly use the composite measure covering all these three subjects for both levels of education. The theoretical minimum and maximum values on this measure is 0 and 100, whereas the empirical minimum is 13.9 (Mauritania in 1995) and the empirical maximum is 74.3 (Japan in 1980). The mean score on the measure is 45.8, and the standard deviation is 9.7.

The following international and regional tests are listed by Angrist (2013) as having been included in the dataset:

- International tests:
 - Programme for International Student Assessment (PISA) – conducted for 15 and 16-year old students
 - Trends in International Mathematics and Science Study (TIMSS) – conducted for students in fourth and eighth grade
 - Progress in International Reading Literacy Study (PIRLS) – conducted for students in fourth grade
- Regional tests:
 - South and Eastern African Consortium for Monitoring of Educational Quality (SAQMEQ) – conducted for students in third and fourth grade in a set of South and Eastern African countries
 - Programme d'Analyses des Systèmes Educatifs de la Conférence (PASEC) – conducted for students in second and fifth grade in a set of countries in Francophone Africa
 - Three tests conducted by the UNESCO Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLCE)
 - * Primer Estudio Regional Comparativo y Explicativo (PERCE) – conducted for students in third and fourth grades in a set of Latin American and Caribbean countries
 - * Segundo Estudio Regional Comparativo y Explicativo (SERCE) – conducted for students in third and sixth grades in a set of Latin American and Caribbean countries
 - * Tercer Estudio Regional Comparativo y Explicativo (TERCE) – conducted for students in third and sixth grades in a set of Latin American and Caribbean countries

8.3 Investigating mechanisms

In our theoretical argument we pointed to two features that may explain why having a democratic regime may not enhance education quality, despite the widely held notion that democratic politicians are more attentive than autocratic leaders when it comes to providing public goods and services that are beneficial for larger parts of the country's population. The first feature is that democracy may not – via the “standard” electoral accountability channel – lead democratic politicians to pursue policy measures that boost education quality. Rather, democratic politicians may want to focus their policy efforts in the education area on measures that are more visible to voters, such as expanding access to education. Second, we highlighted that *even if* (democratic or other) politicians would have wanted to pursue policies enhancing education quality, it is often unclear what these specific policies are, and politicians may lack the capacity to identify and implement the most appropriate measures.

Finding specific measures and designs that allow us to directly test these mechanisms in a general manner, using cross-national data, is hard (in this regard, we note that alternative designs drawing on data from particular country contexts may allow for more direct tests, as evidenced by the excellent study on the electoral incentive mechanism conducted by Harding and Stasavage, 2014). Yet, there are data that allow us to come somewhat closer to the relevant mechanisms, also at the cross-national level. Thus, to shed further light on whether any, or both, of the theorized mechanisms are at play, we conducted additional tests using indicators of education spending. More specifically, we test whether: 1) Democracy correlates with education spending (per student); 2) Education spending correlates with our measure of education quality.

Our argument suggests that democracies have incentives to pursue observable and presumably popular education policy reforms, such as boosting enrollment rates. While it is unclear by how much boosting enrollment rates increases total education spending – new classrooms may need to be built and new textbooks bought, but many classrooms may simply become more densely populated and textbooks may be shared – it is natural to assume that total education spending increases, *ceteris paribus*. Yet, given that democratic incumbents might prefer to promote enrollment rates *at the expense of education quality*, we might not expect to see that democracy strongly correlates with total education spending, and, especially, we should assume no, or even a negative, correlation with *education spending per student*. Nor should we expect to see that democracy correlates negatively with the number of students per teacher, rather to the contrary.

To assess these implications, we run models with various indicators of education spending from Unesco (2015) as dependent variable. The available data on education spending

that is comparable in a global sample of countries has limited coverage, unfortunately, especially over time. Hence, these models rely on cross-sectional data on education spending taken from the period 2005-2009 with all independent variables measured in 2000-2004, thus conforming with the core cross-sectional specifications that we use in the paper.¹ We underline that we should not draw very strong conclusions based on these results, since they are based on cross-sectional comparisons. Yet, the correlations that we obtain can offer some suggestive evidence as to why democracy does not enhance education quality. The main results are included in Table 8.1.

In brief, our tests provide no clear evidence that democracy increases education spending. The point estimate of Model 1, which draws on data from 120 countries and is a parsimonious specification only controlling for income level, suggests that going from -10 to +10 on Polity increases the percentage share of GDP going to education spending by about 0.5. Yet, we cannot reject the hypothesis that this could simply be due to noise, as the t-value of the Polity coefficient is only 0.7. The finding turns even weaker when adding controls for natural resources income, income inequality and population size in Model 2. Further, Models 3 (parsimonious) and 4 (extensive) show very low and insignificant correlations between democracy and spending per primary school student. Further, Models 5 and 6 even display negative, though statistically insignificant, Polity coefficients when spending per secondary school student is the dependent variable.

Table 8.1: Cross-country regressions on democracy and education spending

DV:	1 OLS Education spending (share of GDP)	2 OLS Education spending (share of GDP)	3 OLS Spending per student (primary)	4 OLS Spending per student (primary)	5 OLS Spending per student (secondary)	6 OLS Spending per student (secondary)
Polity Index	0.026 (0.71)	0.017 (0.41)	0.028 (0.18)	0.120 (0.75)	-0.156 (-0.57)	-0.359 (-0.92)
Ln GDP p.c.	0.321* (2.12)	0.259 (1.38)	1.807** (2.98)	2.069** (2.75)	-0.803 (-0.65)	0.263 (0.17)
Ln oil + gas inc. p.c.		0.062 (0.96)		-0.437 (-1.57)		-0.519 (-0.96)
Gini (market income; reversed)		-0.030 (-1.24)		0.094 (1.04)		0.321+ (1.90)
Ln population		-0.428** (-2.93)		-0.462 (-0.99)		-1.543+ (-1.75)
Constant	1.659 (1.36)	10.737** (2.79)	0.181 (0.04)	0.829 (0.08)	29.846** (2.90)	31.602 (1.44)
N	120	108	100	93	95	88

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All independent variables are lagged by one 5-year period, and standard errors are robust.

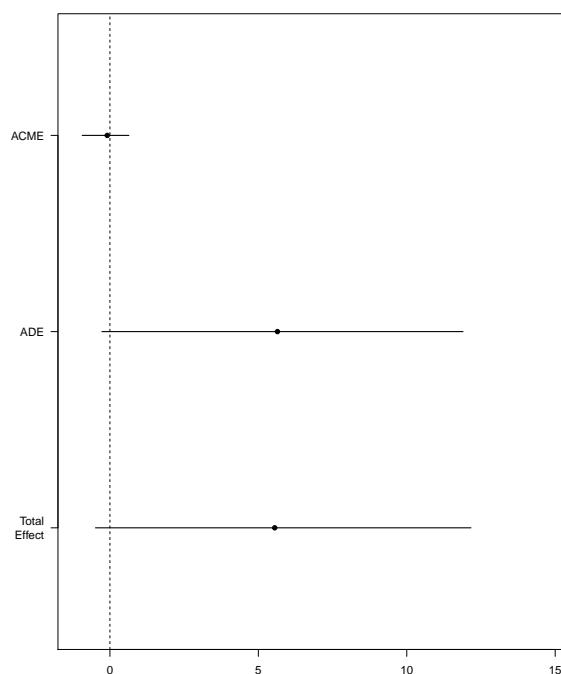
¹Given that coverage is spotty for the spending variables, we can not take averages over the entire time interval and retain a decent number of observations. Instead, we use 2005 as the default year, and fill in with data from the closest available year within the time interval for those countries that have missing data on a particular spending variable in 2005.

The patterns in Models 1 and 2 might, at first glance, seem somewhat surprising, given that democracies have incentives to invest in education in order to boost education enrollment. At the same time, we note that widening the access to education can often be done *without raising education spending by too much*. For example, countries may fill up classrooms with more students without spending on hiring new teachers, and otherwise cut costs per students.² When further investigating this, we do find evidence that the relationship between democracy and education quantity – as measured by enrollment rates or average years of schooling – that we replicated, also holds up fairly well when controlling for education spending, despite the reduced number of observations.

This is illustrated in Figure 8.1, which displays results from a mediation analysis, following Imai et al. (2011). This estimation procedure decomposes the average treatment effect of an explanatory variable into the average causal mediation effect (ACME) and the average direct effect (ADE). According to Imai et al. (2011), this is a more robust way of estimating mediation effects than simply adding the proposed mediator to the regression, and assessing the change in the coefficient estimates of the variable of interest when this is done. The analysis is conducted on the parsimonious model with primary enrollment as dependent variable, education spending as share of GDP as mediator, and the Polity Index as independent variable. The estimated direct “effect” of democracy on primary enrollment is substantial in size and (weakly) significant, whereas the estimated indirect effect of democracy via education spending is minuscule and statistically indistinguishable from zero. Hence, the relationship between democracy and education quantity does not seem to be mediated by increased education spending.

²Another piece of evidence suggesting that democracy may increase access to education without diverting too much resources comes from employing pupil to teacher ratios as the dependent variable, within the same cross-section set-up. While the finding is not robust to adding control variables or considering secondary schooling – although the Polity coefficient is always positively signed – the Polity coefficient is statistically significant at 5 percent in a parsimonious controlling for income level on primary education pupil to teacher ratios. The point estimate suggests that going from -10 to +10 on Polity increases the number of pupils per teacher by about 6 students, holding income level constant.

Figure 8.1: The mediation effect of democracy on primary education enrollment, mediated through education spending. Parsimonious model specification.



Notes: The total effect of democracy on education quantity is decomposed into the average direct effect (ADE) and the average causal mediation effect (ACME). The effects are plotted with their 95 percent confidence intervals. The pre-treatment covariate is Ln GDP per capita.

We now turn to the second part of our argument, concerning the difficulty of implementing education policies that boost the quality of education. Also here we test cross-section regressions, but use education spending as the independent variable and our main student achievements test score measure as the dependent variable. While spending constitutes a crude proxy, one would expect that if education quality is a fairly easily manipulable feature, it should react to politicians spending more resources on education. Conversely, if our assumption about education quality being hard to manipulate through policy is correct, it should not correlate strongly with education spending. Again, we warn that the reported results are cross-country regressions, and results may be biased by unobserved features affecting both education spending and education quality. Thus, this constitutes only suggestive evidence. Still, the available evidence largely points in the same direction as our argument. As Table 8.2 shows, education spending is not systematically correlated with our measure of education quality, neither when omitting (Models 1 and 2) or including (Models 3 and 4) a control for education quantity (average years of education). This does not mean that education quality is a randomly distributed property, as it correlates strongly with various structural features such as a high income level and low income inequality.

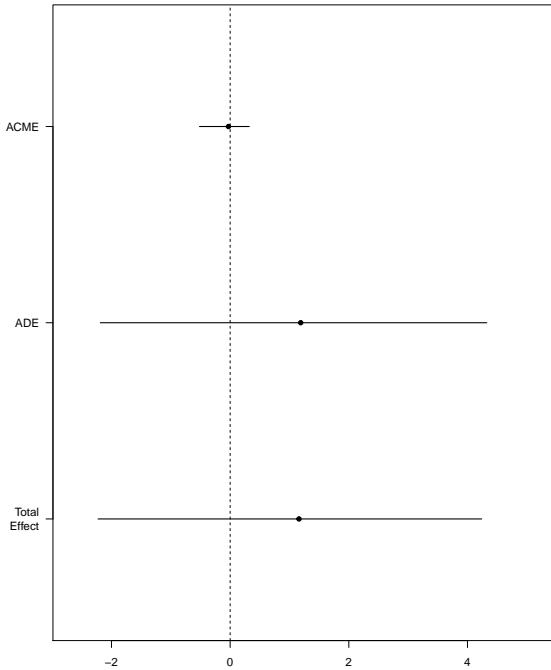
Table 8.2: Cross-country regressions on education spending and education quality

DV:	OLS Education quality	OLS Education quality	OLS Education quality	OLS Education quality
Total education spending (share of GDP)	-0.067 (-0.17)	-0.011 (-0.05)	-0.115 (-0.37)	0.052 (0.22)
Polity Index	0.197 (1.27)	0.256 (1.40)	0.132 (0.99)	0.144 (0.89)
Ln GDP p.c.	5.821** (8.27)	6.431** (7.95)	3.663** (4.05)	5.651** (6.37)
Ln oil + gas income p.c.		-0.182 (-0.77)		-0.387 (-1.66)
Gini (market income; reversed)		0.469** (6.87)		0.389** (5.01)
Ln population		0.528 (1.11)		0.915+ (1.68)
Average years of schooling 15 yr olds			1.533** (4.69)	0.891** (3.01)
Constant	-8.855 (-1.50)	-48.598** (-4.73)	-2.126 (-0.34)	-50.270** (-4.85)
N	89	82	84	78

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. All independent variables are lagged by 5 years, and standard errors are robust.

In sum, we find little evidence that democracy correlates with education spending in our sample, but also little evidence that education spending systematically relates to education quality. This is neatly summed up in Figure 8.3, which presents the results of a mediation analysis (again following Imai et al., 2011) with Polity as independent variable, total education spending as share of GDP as mediator, and the test-score measure of education quality as dependent variable. Again, we employ the parsimonious model specification controlling for Ln GDP per capita and employ the same cross-country design as above. The estimated indirect effect running from democracy on education quality via education spending is virtually zero, but also the estimated direct effect of democracy on education quality is far from statistically significant at conventional levels.

Figure 8.2: The mediation effect of democracy on education quality, mediated through education spending. Parsimonious model specification.



Notes: The total effect of democracy on education quality is decomposed into the average direct effect (ADE) and the average causal mediation effect (ACME). The effects are plotted with their 95 percent confidence intervals. The pre-treatment covariate is Ln GDP per capita.

8.4 Imputation model and diagnostics

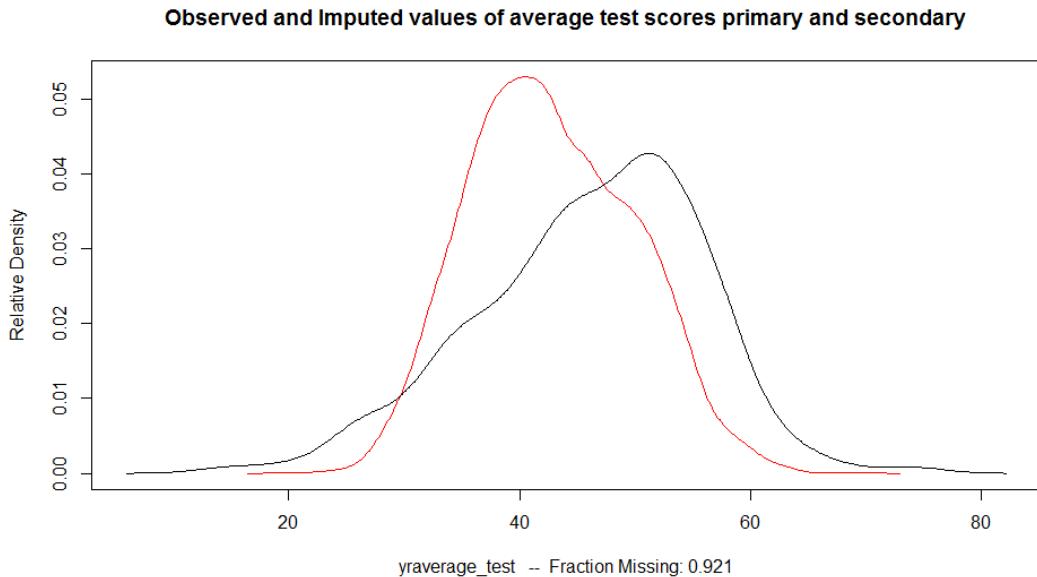
In order to construct our imputation model we employed the Amelia II software (Honaker and King, 2010; Honaker et al., 2011), accounting for the time-series cross-section data structure. We ran the imputation model and produced 5 different data sets, which were subsequently used in the empirical analysis of the paper. Our imputation model allowed for country-specific second-order polynomial time trends (see Honaker et al. 2012, 20-21). Following standard recommendations, we included several relevant variables in order to improve the predictive power of our multiple imputation model.

Multiple imputation builds on the assumption that the data are missing at random (MAR). This means that the pattern of “missingness” should only depend on the observed data included in the imputation model, and not on unobserved variables or features with the data (see Honaker and King 2010). The pattern of missingness on many of the key variables used in our analysis, such as the student achievement test measures, likely depends on a number of factors. For instance, the extent to which governments will allow student achievement tests to be carried out in their country could depend on the level of democracy. Moreover, poorer countries are less likely to have participated in student achievement tests, for instance due to the difficulties of organizing the required capacity for conducting tests. To account for such issues we included a wide range of variables in the imputation model that may influence the missingness, including level of democracy, indicators of human rights, economic growth, GDP per capita, and urbanization. This likely reduces the threat of the MAR-assumption being grossly violated.

Yet, we can not exclude the possibility that the MAR-assumption is still not fully met, for instance because the missingness depends on unobserved factors, and that this causes the imputation model to perform poorly. To evaluate the outputs from the imputation model we therefore conducted different imputation diagnostic tests that help us to assess its quality. One way to describe the outputs from the imputation model in a condensed manner, which can potentially also be used as a check on the plausibility of the imputation model specification, is mapping the distribution of imputed values and the distribution of observed values for particular variables. While very large discrepancies between the observed and imputed distributions may provide warning signs that something may be wrong with the imputation model, one should not necessarily expect the distributions of the missing values to be completely identical to the distributions of the observed values. In fact, the main reason why we impute to begin with is the assumption that observed and missing values could differ systematically; correcting for this alleviates the issue of selection biases affecting our regression results (e.g., Honaker and King, 2010).

Plots showing the distributions of imputed and observed values for the average stu-

Figure 8.3: Distributions of observed and imputed observations for aggregated measure of student achievement test scores



dent achievement test measure from Angrist (2013) are presented in Figure 8.3. The red line represents the density of the mean of each imputed observation across all 5 datasets, while the black line represents the density of the observed values. The figure shows that the shape of the two distributions follow a similar pattern, very roughly, but still deviate especially when it comes to the lower scores on student achievement tests, where there are higher relative frequencies for the imputed data. This is, however, a very plausible result: As noted, there are reasons to expect that poorer countries will have more missing values on the student achievement test measure than rich countries, and income is strongly correlated with education quality. Hence, the distribution of the imputed values for the aggregated student achievement test measure may suggest that the imputation procedure alleviates some of the selection biases discussed above related to the under-representation of developing countries. In Figure 8.4 and Figure 8.5 we present similar plots for student achievements test measures from Angrist (2013), which separately measure performance for students at the primary and secondary levels of schooling, respectively, and the patterns are similar to that for the aggregated index.

Another way of assessing the imputation model is by employing the so-called overimputation procedure (Honaker et al., 2011). Since missing data are, by definition, unobservable, it is impossible to tell whether each imputed value is close to the unobserved

Figure 8.4: Distributions of observed and imputed observations for student achievement test measure for primary school students

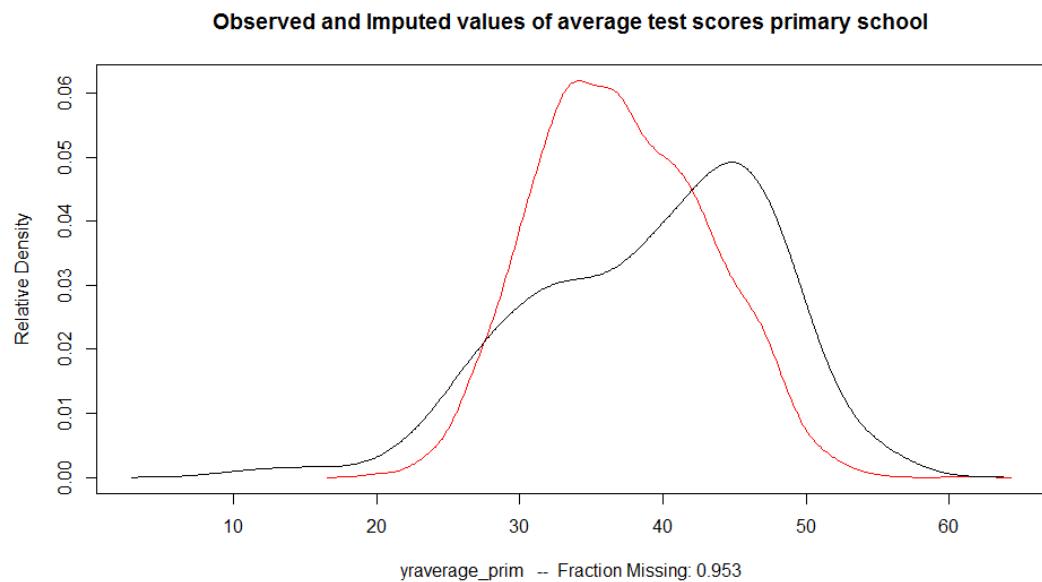
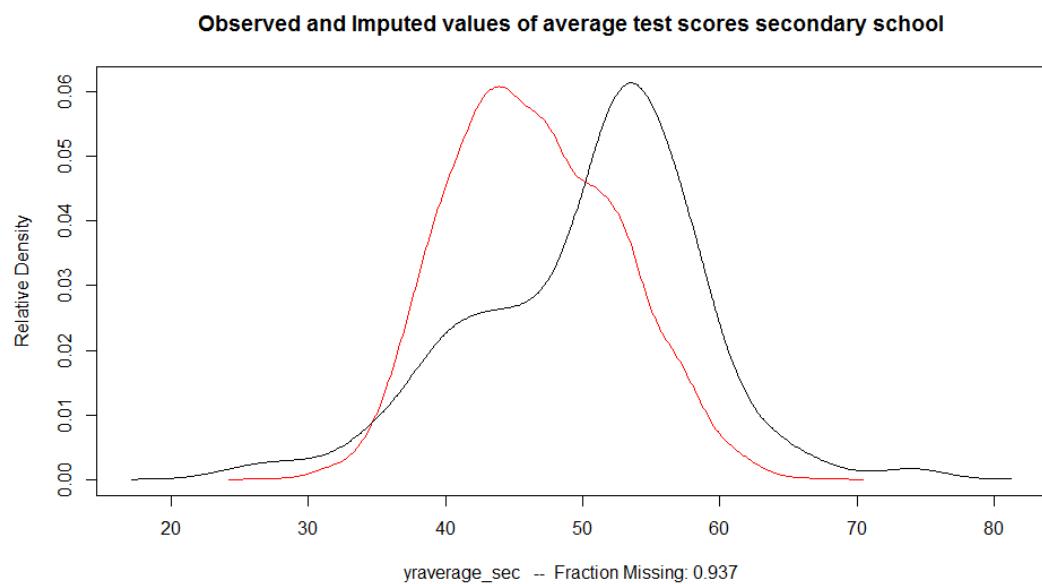


Figure 8.5: Distributions of observed and imputed observations for student achievement test measure for secondary school students



value that we would like the imputation model to predict accurately. An overimputation test is a diagnostics test that attempts to solve this problem by (sequentially) treating each observed value *as if* it is actually missing, and then generate a large number of imputed values for it. After we have this large set of imputed values, we can then construct a confidence interval, and the final step is then assessing whether the actually observed value falls within this interval. If it does, our imputation model performs well in the sense that it is able to accurately predict this “fictive missing data” point (which we *know* the actual score on).

The Amelia II software provides this diagnostics test as well as a simple way of graphically inspecting whether our observed data tends to fall within the region where it is imputed by the model. Figure 8.6 plots the results for the overimputation test for the average student achievement test score. The observed test scores are plotted against the mean of the imputed scores for the very same observation on this index (when the observed test score is treated as if it is missing). The 90 percent confidence intervals that are plotted for each imputed value thus allow us to visually inspect the performance of the imputation model. As noted by Honaker et al. (2012, 30), checking how many of the confidence intervals cover the $y = x$ line lets us tell how often the imputation model can confidently predict the true value of the observation. The line colors on the confidence intervals tell the fraction of missing observations for all covariates entered in the imputation model for that particular observation.

As we can see from Figure 8.6, our imputation model performs well when it comes to predicting the observed values. Every single confidence interval, except from two, covers the 45 degree line. The predictions are especially good for the medium to higher values of the student achievement index. However, the imputation model is worse at predicting the observed values at the lowest level of the student achievement test score measure, but we note that there are very few observations in this area. In Figures 8.7 and 8.8 we plot the results for the overimputation tests on the primary level and secondary level test score measures, respectively, and they show a similar pattern: The imputation model, overall, seems to predict well, even though it performs slightly worse when it comes to predicting the very lowest values on the student achievement test measures.

Figure 8.6: Overimputation test plot, aggregated measure of student achievement test scores

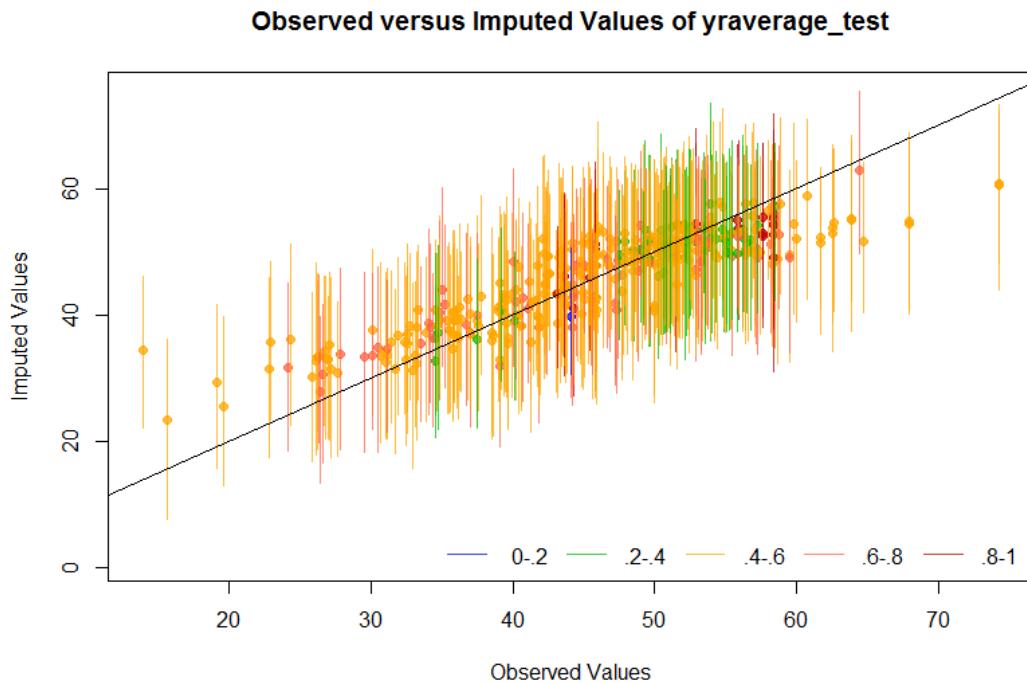


Figure 8.7: Overimputation test plot, student achievement test measure for primary school students

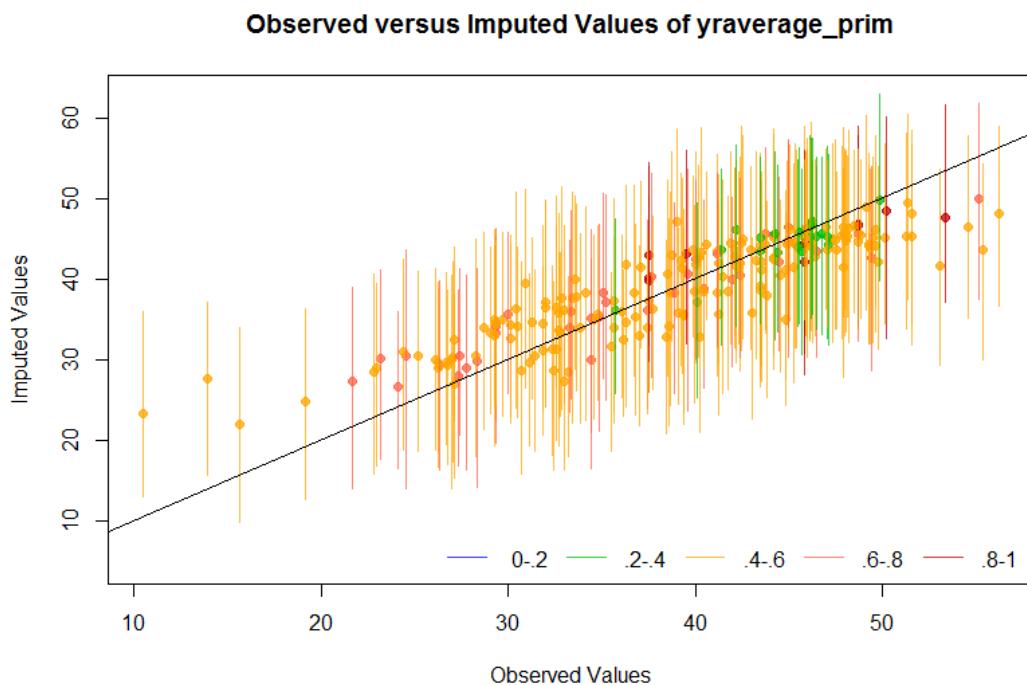
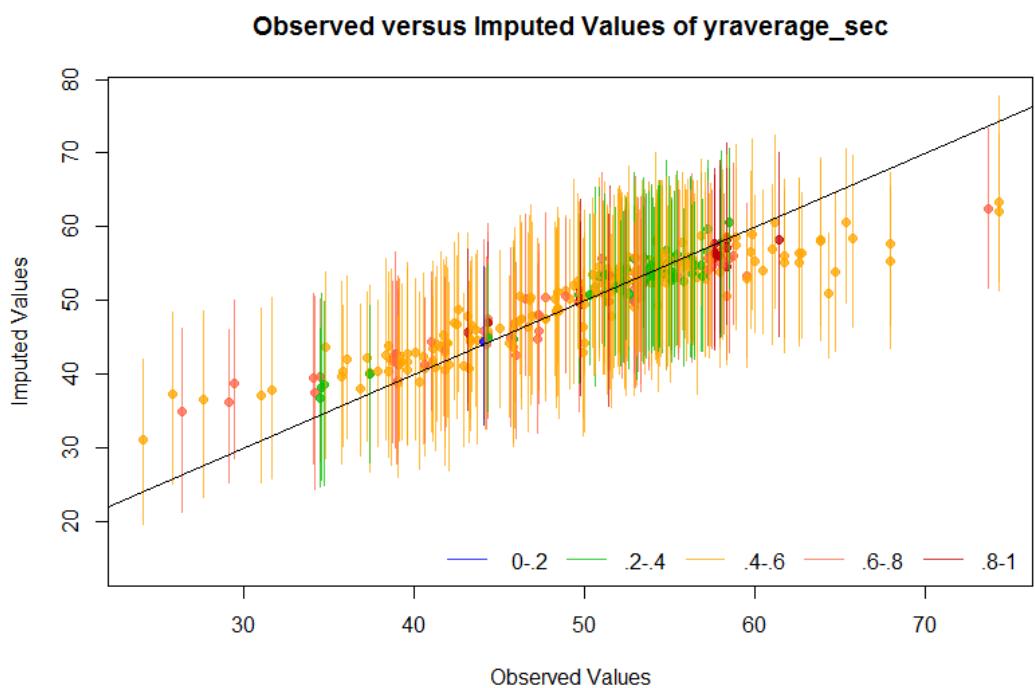


Figure 8.8: Overimputation test plot, student achievement test measure for secondary school students



8.5 Robustness tests: Democracy and measures of “education quantity”

In this section we present robustness tests on the relationship between democracy and *education quantity*. First, we replicate Table 3.1 in the paper, but substitute the Polity Index with the Freedom House Index. The two subsequent tables show results for models substituting the average years of schooling measure from Table 3.1 with measures of enrollment rates. The final table shows results from models based on alternative estimators.

In short, the models below consistently display a positive point estimate for democracy on the different measures of “education quantity”, and while several models do not yield coefficients that are significant at conventional levels, many others do. Albeit not robust, we thus find evidence that democracy relates positively to educational aspects such as enrollment rates and average years spent in school.

Table 8.3: Democracy, as measured by the Freedom House Index (average of the Political Rights and Civil Liberties indices), and average years of schooling

	A1 OLS Cross-sec. b/t	A2 OLS Cross-sec. b/t	A3 OLS (FE) 5-yr panel b/t	A4 OLS (FE) 5-yr panel b/t	A5 FE 2SLS 5-yr panel b/t	A6 FE 2SLS 5-yr panel b/t
Freedom House Index	0.200+ (1.74)	0.230 (1.47)	0.372** (6.90)	0.018 (0.51)	1.203** (8.36)	0.112 (1.39)
Ln GDP p.c.	1.840** (10.18)	1.807** (7.57)	1.177** (2.71)	0.525* (2.57)	0.708** (2.78)	0.577** (4.33)
Ln oil+gas income p.c.		0.107 (1.42)		-0.083+ (-1.67)		-0.082* (-2.21)
Gini (market income;reversed		0.040 (1.44)		0.006 (0.84)		0.007 (1.35)
Ln population		-0.240+ (-1.77)		1.072* (2.31)		1.163** (4.41)
Period dummies				Y	Y	Y
Country dummies				Y	Y	Y
N	120	104	828	599	818	591

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is average years of schooling for 25 year olds. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a)

is used as instrument for Freedom House in the FE 2SLS models (see Appendix for first stage regressions). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.4: Democracy and primary school enrollment rates

	A1 OLS Cross-sec. b/t	A2 OLS Cross-sec. b/t	A3 OLS (FE) 5-yr panel b/t	A4 OLS (FE) 5-yr panel b/t	A5 FE 2SLS 5-yr panel b/t	A6 FE 2SLS 5-yr panel b/t
Polity Index	0.040 (0.24)	-0.179 (-0.69)	1.030** (4.56)	0.220 (1.42)	2.041** (6.33)	0.638 (1.46)
Ln GDP p.c.	0.975 (0.68)	1.953 (1.14)	3.504 (1.30)	-5.334 (-1.62)	1.558 (0.79)	-4.261+ (-1.65)
Ln oil+gas income p.c.		-0.441 (-1.11)		-0.022 (-0.04)		0.005 (0.01)
Gini (market income;reversed)		-0.319* (-2.24)		0.142 (1.09)		0.142 (1.23)
Ln population		1.894* (2.20)		29.807** (3.83)		30.053** (6.62)
Period dummies				Y		Y
Country dummies			Y	Y	Y	Y
N	133	119	1101	788	1094	777

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is primary school enrollment rates. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for Freedom House in the FE 2SLS models (see Appendix for first stage regressions). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.5: Democracy and secondary school enrollment rates

	A1 OLS Cross-sec. b/t	A2 OLS Cross-sec. b/t	A3 OLS (FE) 5-yr panel b/t	A4 OLS (FE) 5-yr panel b/t	A5 FE 2SLS 5-yr panel b/t	A6 FE 2SLS 5-yr panel b/t
Polity Index	-0.006 (-0.02)	0.351 (0.96)	1.131** (7.04)	0.123 (0.81)	2.193** (8.43)	0.355 (1.15)
Ln GDP p.c.	22.010** (13.54)	21.899** (13.03)	22.448** (7.16)	11.465** (4.98)	19.916** (10.24)	11.848** (6.13)
Ln oil+gas income p.c.		0.997+ (1.71)		0.481 (1.06)		0.493 (1.30)
Gini (market income;reversed)		0.506* (2.08)		0.067 (0.79)		0.072 (1.00)
Ln population		-2.398* (-2.29)		14.258** (2.78)		14.318** (4.49)
Period dummies				Y		Y
Country dummies			Y	Y	Y	Y
N	127	115	1047	747	1044	740

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is secondary school enrollment rates. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for Freedom House in the FE 2SLS models (see Appendix for first stage regressions). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.6: Democracy and mean years of schooling. Alternative estimators

	A1 OLS PCSE 5-year panel b/t	A2 OLS PCSE 5-year panel b/t	A3 OLS (RE) 5-yr panel b/t	A4 OLS (RE) 5-yr panel b/t	A5 System gmm 5-yr panel b/t	A6 System GMM 5-yr panel b/t
Polity Index	0.129** (11.80)	0.111** (8.56)	0.121** (8.82)	0.014 (1.64)	0.002 (0.36)	-0.000 (-0.01)
Ln GDP p.c.	1.740** (23.95)	1.746** (18.68)	1.912** (8.68)	0.639** (3.97)	0.212** (3.86)	0.013 (0.14)
Ln oil+gas income p.c.		-0.029 (-0.89)				0.012 (0.60)
Gini (market income;reversed)		0.069** (8.74)				0.001 (0.36)
Ln population		-0.126* (-2.24)				0.379** (3.44)
Lagged DV (Avg yrs schooling)					1.014** (68.84)	0.963** (33.19)
Period dummies				Y	Y	Y
Country dummies				Y	Y	Y
N	1027	707	1027	1027	982	688

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean years of schooling. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for Freedom House in the FE 2SLS models (see Appendix for first stage regressions). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

8.6 2SLS models: Discussion, first-stage regressions for core specifications, and alternative 2SLS models using different instruments

In this section we include two tables displaying the first-stage regression estimates from four instrumental variable regression models (Fixed Effects 2SLS), for which the second-stage regressions were presented in Tables 3.1 and 3.2 in the paper. In these first-stage regressions, the dependent variable is the Polity Index, and the WAVE instrument from Knutsen (2011a) is included as an independent variable alongside all the covariates included also in the second-stage regressions. In brief, the first-stage regressions show that WAVE has a strong and highly significant negative relationship with democracy, both when adding country fixed effects and country fixed effects combined with period dummies. Also F-tests show that this instrument is very strong, with F-values (see bottom row of the two tables) far exceeding those considered threshold values for strong instruments (e.g., $F = 10$). Also other weak identification tests, such as comparing the Stock-Yogo critical values against the Cragg-Donald Wald F-statistic, leave no doubt that WAVE is a very strong instrument, and the t-values of WAVE range between -4.6 and -13.7.

Table 8.7: First-stage regressions (with Polity Index as dependent variable) for Fixed Effects 2SLS models presented in Table 3.1 of the paper.

	F1 5-year panel b/t	F2 5-year panel b/t t
WAVE	-6.597** (-13.64)	-5.078** (-7.25)
Ln GDP p.c	1.838** (4.85)	-1.400+ (-1.83)
Ln oil+gas income p.c.		0.131 (0.93)
Gini (market income;reversed)		0.009 (0.28)
Ln population		-3.303** (-2.57)
Period dummies		Y
Country dummies	Y	Y
N	1019	698
F-test value excl. instrument	98.36	52.50

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is democracy, measured using the Polity Index. Standard errors are robust. First time period is 1960–64, and final time period is 2000–2004. WAVE from Knutsen (2011a), recording whether or not the last regime change was within or outside one of Huntington's "reverse waves of democratization", is used as instrument for democracy in the first-stage regression. Constant, country dummies and period dummies are omitted.

Table 8.8: First-stage regressions (with Polity Index as dependent variable) for Fixed Effects 2SLS models presented in Table 3.2 of the paper.

	F1 5-year panel b/t	F2 5-year panel b/t t
WAVE	-7.893** (-4.76)	-7.853** (-4.59)
Ln GDP p.c	2.341** (4.22)	-0.860 (-0.56)
Ln oil+gas income p.c.		0.109 (0.59)
Gini (market income;reversed)		-0.035 (-0.67)
Ln population		-2.687 (-1.40)
Period dummies		Y
Country dummies	Y	Y
N	317	297
F-test value excl. instrument	22.69	21.09

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is democracy, measured using the Polity Index. Standard errors are robust. First time period is 1960–64, and final time period is 2000–2004. WAVE from Knutsen (2011a), recording whether or not the last regime change was within or outside one of Huntington's "reverse waves of democratization", is used as instrument for democracy in the first-stage regression. Constant, country dummies and period dummies are omitted.

Table 8.9 presents second-stage results on education quality from the alternative fixed effects 2SLS specifications that we tested. The two rightmost columns (IV1 and IV2) report the models also reported in the paper, namely the parsimonious and extensive models employing WAVE as instrumental variable. As we also discuss in the paper (for a more extensive discussion, see Knutsen, 2011a), this instrument aims to capture exogenous influence on regime type, notably coming from regional and global spill-over effects and features of the international-political climate, through leveraging the notion that democratic regime changes have tended to cluster temporally in “waves” (Huntington, 1991). As we will return to, standard over-identification tests of the exclusion restriction can only be conducted in the presence of two different instruments. While this is not an accurate test of the exclusion restriction (Sovey and Green, 2011), we do, however, take some solace in the fact that when we re-run the second-stage regressions (as standard fixed effects OLS models), but including WAVE directly in the second stage, WAVE is not directly linked to education quality. More specifically, the t-value of WAVE is 0.16 in the parsimonious model and -0.45 in the extensive model. This is not bullet-proof evidence that the exclusion restriction is valid in Models IV1 and IV2, but they do mitigate concerns that these estimates are strongly biased by a clear, direct link between WAVE and education quality.

An alternative, and more direct, way of capturing spill-over effects on regime type from neighbors at the regional level is to calculate the “democracy environment” in the geographic region and use that as an instrument. When doing so, we naturally exclude the country in question whenever scoring the regional democracy environment variable. More specifically, we take the average Polity score within the country’s geographic region, omitting the relevant country from the calculation. We rely on the geographic region coding in Hadenius and Teorell (2007), but collapse some of the smaller regions so that we are left with the following major world regions: Eastern Europe and the (post-)Soviet space; Western Europe plus Australia, Canada, New Zealand and USA; Latin America and the Caribbean; Middle East and North Africa; Sub-Saharan Africa; Asia and the Pacific.

The exogeneity of this instrument might be somewhat more problematic than for WAVE, in particular for relatively larger countries. While WAVE draws on a broad (and global) temporal categorization to identify exogenous variation in political regime type, which are arguably less influenced by events in most countries, regime changes and other political events in a regional power (such as Brazil in Latin America) might influence the regional democracy score. Nonetheless, for most countries we consider that treating the democracy environment in the wider regions, at any given point in time, as an exogenous feature should be a decent approximation. Regarding the exclusion restriction we are

– as we are for the models using WAVE – making the (plausible) assumption that the wider international political regime environment should not have any direct effect on education quality in the country in question (other than through affecting regime change domestically), conditional on the period dummies, income level and other covariates in the model. Regarding strength, the regional average polity score instrument turns out highly significant in the first-stage, and the F-statistics (see bottom row Table 8.9) show that the instrument passes any conventional thresholds for being considered a strong instrument.

Model IV3 presents results from the parsimonious model, controlling only for GDP per capita, country dummies and year dummies, that uses the regional average instrument. As for the models using WAVE, the Polity coefficient is negatively signed in the second stage, but statistically insignificant at all conventional levels. In Model IV4, which adds control for resource income, income inequality and population, Polity remains negatively signed and actually turns significant at the 5 percent level.

We also tested FE 2SLS model simultaneously including WAVE and the regional average instrument, and again the Polity coefficient turns out statistically insignificant in the parsimonious model (IV5) and negative and significant at 5 percent in the extensive model (IV6). A benefit of using two instruments is that we can run standard overidentification tests on the validity of the exclusion restriction. The Hansen j-test p-value for Model IV5 is .19, suggesting that we cannot reject the hypothesis that the exclusion restriction holds at conventional levels, whereas the p-value is .08 for Model IV6.

Still, the WAVE and regional average instruments draw on a similar underlying logic related to geographic spill-over effects, and overidentification tests may not give accurate results under such conditions (Murray, 2006). Therefore, we also tested models combining our preferred instrument, WAVE, with a quite different instrument. More specifically, we follow Helliwell (1994) in using the 15-year lagged value of Polity as an instrument for current Polity (which, in turn, is measured five years before the dependent variable). This instrument draws on the notion that democratic institutions are often persistent, suggesting that lagged democracy is a strong instrument. Further, we need to assume that there is no direct effect of historical political regime type on education quality other than its influence through affecting the current regime. This is a very strong assumption, for instance due to the potential long time lag between education policy decisions and improvements in education quality discussed in the paper. This is why we do not include the lagged Polity variable as an instrument in our main specifications reported in the paper. Still, we tested it to gauge the results of Hansen j-tests based on two very different instruments, and they actually suggest that the parsimonious model (IV7) might yield consistent estimates ($p=.61$) whereas the extensive models (IV8) might not ($p=.05$). The

results for Polity, in both models, suggest a negative effect of democracy on education quality. Despite the outcome of the Hansen j-test for Model IV7, we note that we put less trust in these models given our theoretical concerns with exclusion restriction in these specifications. Nonetheless, we note that neither these nor any of the other 2SLS models that we have tested show any indications of a positive effect of democracy on education quality.

Table 8.9: Democracy and mean student achievements test score: Second-stage regressions from alternative FE 2SLS specifications using different instruments

	IV1 b/t	IV2 b/t	IV3 b/t	IV4 b/t	IV5 b/t	IV6 b/t	IV7 b/t	IV8 b/t
Polity Index	-0.186 (-0.67)	-0.017 (-0.07)	-0.343 (-1.22)	-0.564* (-2.25)	-0.273 (-1.22)	-0.335* (-2.14)	-0.434* (-2.56)	-0.394* (-2.49)
Ln GDP p.c.	1.360 (0.92)	1.924 (0.77)	1.771 (0.73)	1.560 (0.56)	1.490 (0.69)	1.712 (0.65)	-0.518 (-0.31)	-0.559 (-0.23)
Ln oil+gas income p.c.	-0.323 (-0.64)			-0.265 (-0.48)		-0.289 (-0.55)		0.514 (1.20)
Gini (market inc.; reversed)	0.013 (0.16)			-0.003 (-0.03)		0.004 (0.05)		-0.007 (-0.07)
Ln population	7.879* (2.07)			9.626* (2.31)		8.896* (2.30)		4.165 (1.03)
Period dummies	Y	Y	Y	Y	Y	Y	Y	Y
N	317	297	262	297	262	297	286	265
Instruments	WAVE		Regional average Polity		WAVE + Reg. avg. Polity		WAVE + Polity 15 yrs lag	
Hansen j-test p-val.	—	—	—	—	.186	.078	.609	.053
Cragg-D. Wald F-stat.	80.8	69.9	86.4	98.8	63.5	87.9	72.2	55.1

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean years of schooling. Time series units are 5-year panels, and all independent variables are lagged by 5 years, and standard errors are robust. Polity Index is the endogenous independent variable, and instrument(s) used in the first stage are reported towards bottom of the table. First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

8.7 Other robustness tests: Democracy and education quality

The following tables contain robustness tests of our core result on democracy and education quality. The first table employs alternative estimation techniques (OLS PCSE, RE, System GMM) to the core tests reported in Table 3.2. The following tables reports result where Table 3.2 from the paper is replicated using other widely used measures of democracy than the Polity index. Then follows a table where the independent variables from Table 3.2 are lagged 10, rather than 5, years behind the outcome. Thereafter, we present results for similar robustness tests of Table 3.3 of the paper. At the end, we report tables with models including a squared Polity term or dummies capturing different ranges on the Polity scale, in order to check for non-linearities, models checking for interaction between democracy and income level/income inequality/quality of government on education quality, as well as models using the autocratic regime type dummies from Hadenius and Teorell (2007).

Table 8.10: Democracy and mean student achievements test score: Alternative estimators

	C1 OLS PCSE 5-yr panel b/t	C2 OLS PCSE 5-yr panel b/t	C3 RE 5-yr panel b/t	C4 RE 5-yr panel b/t	C5 System GMM 5-yr panel b/t	C6 System GMM 5-yr panel b/t
Polity Index	0.039 (0.43)	-0.023 (-0.24)	0.025 (0.16)	0.008 (0.07)	0.243 (0.71)	0.000 (0.00)
Ln GDP p.c.	6.566** (15.35)	6.956** (14.60)	5.440** (8.29)	6.088** (9.94)	2.628 (1.30)	5.376 (1.56)
Ln oil+gas income p.c.		-0.433** (-3.06)		-0.371* (-2.10)		-1.691** (-3.43)
Gini (market income; reversed)		0.330** (5.87)		0.261** (4.40)		0.248 (1.47)
Ln population		1.320** (3.48)		1.072* (2.51)		0.793 (0.39)
Lagged dep. var.					-0.123 (-1.34)	0.047 (0.20)
Period dummies		Y		Y		Y
N	341	321	341	321	190	185

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust for random effects and system GMM models, and account for panel-specific AR(1) autocorrelation and panel-level heteroskedasticity in OLS PCSE models. The System GMM models treat the Polity Index as endogenous, and there are no restrictions on the number of lags used for instrumentation. First time period is 1965–69, and final time period is 2005–2009. Constant and period dummies are omitted from the table.

Table 8.11: Democracy, as measured by the Freedom House Index (average of political and civil liberties indices), and mean student achievement test scores.

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t	B5 FE 2SLS 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
Freedom House Index	0.244 (0.44)	0.133 (0.25)	-0.228 (-0.25)	-0.326 (-0.54)	-1.604+ (-1.95)	-0.477 (-0.65)
Ln GDP p.c.	5.891** (7.07)	7.013** (8.05)	-1.432 (-0.69)	0.720 (0.26)	-0.324 (-0.20)	0.719 (0.33)
Ln oil+gas income p.c.		-0.464+ (-1.68)		0.401 (1.04)		0.405 (1.05)
Gini (market income; reversed)		0.409** (5.93)		-0.039 (-0.43)		-0.042 (-0.47)
Ln population		1.043+ (1.88)		3.030 (0.63)		2.950 (0.74)
Period dummies				Y		Y
Country dummies		Y	Y	Y	Y	Y
N	94	86	328	307	302	282

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for the Freedom House Index in the FE 2SLS models (see Appendix for first stage regressions). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.12: Democracy, as measured by the Democracy–Dictatorship (DD) measure, and mean student achievement test scores.

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t	B5 FE 2SLS 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
DD	2.745 (1.61)	0.870 (0.56)	-0.039 (-0.01)	-0.314 (-0.15)	-2.936 (-0.66)	-0.234 (-0.07)
Ln GDP p.c.	5.676** (8.30)	6.980** (9.44)	0.789 (0.48)	1.931 (0.65)	1.324 (0.90)	1.932 (0.78)
Ln oil+gas income p.c.		-0.454 (-1.62)		-0.321 (-0.67)		-0.322 (-0.64)
Gini (market income; reversed)		0.402** (6.02)		0.014 (0.17)		0.014 (0.18)
Ln population		0.996+ (1.75)		7.878+ (1.84)		7.864* (2.09)
Period dummies				Y		Y
Country dummies		Y	Y	Y	Y	Y
N	94	86	344	323	317	297

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for DD (from Cheibub, Gandhi and Vreeland, 2010) in the FE 2SLS models (see Appendix for first stage regressions). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.13: Democracy and mean student achievements test score. 10-year lag.

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t	B5 FE 2SLS 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
Polity Index	0.145 (0.95)	0.063 (0.33)	0.084 (0.71)	-0.115 (-1.00)	-0.042 (-0.32)	-0.166 (-1.16)
Ln GDP p.c.	5.652** (7.49)	6.489** (7.64)	-0.517 (-0.29)	0.525 (0.18)	-0.099 (-0.07)	0.451 (0.19)
Ln oil+gas income p.c.	-0.340 (-0.95)	-0.340 (-0.95)	-0.330 (-0.68)	-0.330 (-0.68)	-0.334 (-0.73)	-0.334 (-0.73)
Gini (market income;reversed)	0.416** (4.55)		-0.132 (-1.55)		-0.130+ (-1.55)	-0.130+ (-1.81)
Ln population	0.754 (1.32)		11.337* (2.34)		11.633** (2.68)	
Period dummies				Y		Y
Country dummies		Y	Y	Y	Y	Y
N	91	83	339	313	313	288

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 10 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for Freedom House in the FE 2SLS models (see Appendix for first stage regressions). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.14: Democracy and mean student achievement test scores, checking for a non-linear relationship by including a squared term for Polity.

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t	B5 FE 2SLS 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
Polity Index	1.092 (1.58)	0.639 (0.78)	1.683** (2.65)	1.091+ (1.71)	-2.215 (-0.54)	1.180 (0.35)
Polity Index squared	-0.043 (-1.42)	-0.021 (-0.63)	-0.076** (-2.71)	-0.051+ (-1.85)	0.096 (0.54)	-0.055 (-0.38)
Ln GDP per capita	6.595** (6.85)	7.109** (6.61)	1.065 (0.62)	2.542 (0.81)	0.758 (0.46)	2.592 (0.79)
Ln oil+gas income p.c.	-0.434 (-1.45)	-0.434 (-1.45)	-0.205 (-0.44)	-0.205 (-0.44)	-0.197 (-0.32)	-0.197 (-0.32)
Gini (market income;reversed)	0.430** (5.88)		0.047 (0.54)	0.047 (0.54)	0.050 (0.34)	0.050 (0.34)
Ln population	0.956+ (1.70)		5.907 (1.55)	5.907 (1.55)	5.743 (0.83)	5.743 (0.83)
Period dummies				Y		Y
Country dummies		Y	Y	Y	Y	Y
N	91	84	341	321	317	297

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for the linear term of the Polity Index in the FE 2SLS models (squared Polity term is not instrumented). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.15: Democracy and mean student achievement test scores, checking for a non-linear relationship by using autocracy, partial democracy, full democracy categorization from Epstein et al. (2006).

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t	B5 FE 2SLS 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
Full Democracy	0.956 (0.49)	0.506 (0.30)	-1.270 (-0.58)	-1.045 (-0.61)	-3.520 (-0.73)	-0.783 (-0.22)
Partial Democracy	0.795 (0.43)	1.068 (0.60)	-0.399 (-0.25)	0.545 (0.32)	-1.554 (-0.50)	0.684 (0.27)
Ln GDP per capita	6.010** (7.35)	7.224** (9.46)	1.035 (0.61)	1.867 (0.62)	1.498 (0.99)	1.859 (0.75)
Ln oil+gas income p.c.		-0.455 (-1.58)		-0.325 (-0.67)		-0.328 (-0.65)
Gini (market income;reversed)		0.412** (5.68)		0.012 (0.14)		0.012 (0.14)
Ln population		1.056+ (1.91)		7.752+ (1.86)		7.730* (2.06)
Period dummies				Y		Y
Country dummies		Y	Y	Y	Y	Y
N	91	84	341	321	317	297

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust. Partial democracy ($1 \leq Polityscore \leq 7$) and full democracy ($Polityscore \geq 8$) are categorized using the criteria from Epstein et al. (2006). WAVE from Knutsen (2011a) is used as instrument for Full Democracy in the FE 2SLS models (squared Polity term is not instrumented). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.16: Democracy and mean student achievement test scores, checking for an interaction between democracy and income level.

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t	B5 FE 2SLS 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
Polity Index	-1.757 (-1.25)	0.882 (0.63)	2.625+ (1.83)	-0.275 (-0.17)	-8.271 (-0.54)	0.453 (0.05)
Ln GDP per capita	4.618** (4.22)	7.356** (6.14)	2.796 (1.37)	1.809 (0.57)	-5.828 (-0.46)	2.124 (0.41)
PolityX Ln GDP p.c.	0.224 (1.40)	-0.090 (-0.57)	-0.318+ (-1.77)	0.030 (0.15)	1.011 (0.55)	-0.058 (-0.05)
Ln oil+gas income p.c.		-0.421 (-1.41)		-0.320 (-0.66)		-0.327 (-0.65)
Gini (market income;reversed)		0.436** (5.75)		0.014 (0.18)		0.009 (0.10)
Ln population		1.016+ (1.77)		8.305+ (1.68)		7.150 (0.44)
Period dummies				Y		Y
Country dummies		Y	Y	Y	Y	Y
N	91	84	341	321	317	297

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for the Polity Index in the FE 2SLS models (squared Polity term is not instrumented). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.17: Democracy and mean student achievement test scores, checking for an interaction between democracy and income inequality.

	B2 OLS Cross-sec. b/t	B4 OLS (FE) 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
Polity Index	0.430 (0.41)	1.381+ (1.73)	0.624 (0.29)
Gini (market income;reversed)	0.460** (3.49)	0.098 (0.79)	0.053 (0.32)
PolityXGini	-0.005 (-0.26)	-0.023+ (-1.88)	-0.011 (-0.32)
Ln GDP per capita	6.777** (8.25)	2.198 (0.80)	2.034 (0.81)
Ln oil+gas income p.c.	-0.434 (-1.35)	-0.444 (-0.93)	-0.378 (-0.73)
Ln population	1.006+ (1.76)	5.435 (1.16)	6.806 (1.21)
Period dummies		Y	Y
Country dummies		Y	Y
N	84	321	297

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for the Polity Index in the FE 2SLS models (squared Polity term is not instrumented). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.18: Democracy and mean student achievement test scores, controlling for quality of government.

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t	B5 FE 2SLS 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
Polity Index	0.106 (0.68)	0.058 (0.29)	0.021 (0.14)	-0.048 (-0.27)	-0.217 (-0.89)	-0.523* (-1.99)
Quality of Government Index	12.577* (2.26)	8.500+ (1.74)	-1.145 (-0.25)	-6.428 (-1.18)	1.231 (0.23)	0.230 (0.03)
Ln GDP per capita	3.824** (3.18)	5.536** (4.49)	3.122+ (1.71)	1.302 (0.39)	3.909* (2.18)	-0.370 (-0.12)
Ln oil+gas income p.c.		-0.470 (-1.66)		0.078 (0.15)		0.213 (0.44)
Gini (market income;reversed)		0.439** (5.96)		-0.030 (-0.27)		-0.011 (-0.11)
Ln population		0.666 (1.14)		1.078 (0.15)		0.845 (0.16)
Period dummies				Y		Y
Country dummies		Y	Y	Y	Y	Y
N	84	77	282	265	260	244

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for the Polity Index in the FE 2SLS models (squared Polity term is not instrumented). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.19: Democracy and mean student achievement test scores, checking for an interaction between democracy and quality of government.

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t	B5 FE 2SLS 5-yr panel b/t	B6 FE 2SLS 5-yr panel b/t
Polity Index	-1.057+ (-1.81)	-0.735 (-1.17)	0.551* (2.46)	0.473 (1.20)	5.835 (0.29)	10.099 (0.39)
Quality of Government Index	-4.921 (-0.44)	-3.349 (-0.30)	2.349 (0.48)	-3.643 (-0.63)	20.376 (0.30)	9.330 (0.23)
PolityXQoG	2.054+ (1.98)	1.365 (1.27)	-1.005* (-2.52)	-0.941 (-1.60)	-9.102 (-0.30)	-14.739 (-0.39)
Ln GDP per capita	4.264** (3.50)	5.934** (4.39)	2.747 (1.60)	1.604 (0.49)	-3.619 (-0.15)	13.030 (0.38)
Ln oil+gas income p.c.		-0.602* (-2.05)		0.026 (0.05)		-1.304 (-0.35)
Gini (market income;reversed)		0.409** (5.14)		-0.052 (-0.46)		-0.459 (-0.37)
Ln population		0.699 (1.24)		-0.311 (-0.04)		-19.711 (-0.34)
Period dummies				Y		Y
Country dummies		Y	Y	Y	Y	Y
N	84	77	282	265	260	244

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. All independent variables are lagged by 5 years, and standard errors are robust. WAVE from Knutsen (2011a) is used as instrument for the Polity Index in the FE 2SLS models (squared Polity term is not instrumented). First time period is 1965–69, and final time period is 2005–2009. Constant, country dummies and period dummies are omitted from the table.

Table 8.20: Democracy (reference category), autocracy types and mean student achievement test scores.

	B1 OLS Cross-sec. b/t	B2 OLS Cross-sec. b/t	B3 OLS (FE) 5-yr panel b/t	B4 OLS (FE) 5-yr panel b/t
Monarchy	-4.646+ (-1.78)	-0.866 (-0.26)	-46.432** (-10.78)	-37.976** (-5.09)
Military	-1.918 (-0.47)	-3.818 (-0.89)	-3.986 (-1.08)	-3.787 (-1.15)
One party	12.761** (11.50)	0.000 (.)	3.855 (0.93)	2.116 (0.79)
Multi-party	3.814* (2.20)	3.662** (2.66)	2.775+ (1.78)	3.769* (2.19)
Other autocracy	0.057 (0.03)	-1.803 (-1.43)	-25.736** (-6.97)	-19.829** (-3.53)
Ln GDP per capita	6.833** (11.03)	7.701** (12.28)	-1.556 (-0.71)	-0.375 (-0.09)
Ln oil+gas income p.c.		-0.507+ (-1.88)		-0.141 (-0.25)
Gini (market income;reversed)		0.383** (5.77)		0.032 (0.30)
Ln population		1.120* (2.03)		1.813 (0.38)
N	94	86	295	276
Period dummies				Y
Country dummies		Y	Y	Y
N	94	86	295	276

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. Standard errors are robust. First time period is 1970–74, and final time period is 2005–2009. Regime dummies are taken from Hadenius and Teorell (2007), and democracy is the reference category. Since time series for the regime dummies start in 1972, these are not lagged by 5 years to preserve the number of observations. Other independent variables are lagged by 5 years. Constant, country dummies and period dummies are omitted from the table.

8.8 Additional robustness tests on education quality: Models run on imputed data

In this section we present different Random Effects and Fixed Effects models on the imputed data sets. The coefficients reported are averaged over regressions run on our five imputed datasets, and the errors are imputation-corrected. We report fairly parsimonious specifications controlling for income level (both with and without year-fixed effects). We also present models controlling for enrollment ratios to account for the alternative explanation of our core result discussed in the paper (related to democracies expanding schooling to students that expectedly perform less well in tests). Tests are conducted both on (levels of) mean student achievement test scores, and on first-differences of such mean test scores (i.e., changes in mean test scores between t-1 and t). Further, we also report models employing alternative lag structures (1, 5 and 10 year lags). Again, the models show no systematic relationship, although one fixed-effects first-difference model suggests that democracy reduces education quality, whereas two random effects models (on levels) suggest a positive effect of democracy when independent variables are lagged by 10 years.

Table 8.21: Democracy and mean student achievements test score

	D1 RE 1-yr panel b/t	D2 OLS (FE) 1-yr panel b/t	D3 RE 1-yr panel b/t	D4 OLS (FE) 1-yr panel b/t	D5 RE 1-yr panel b/t	D6 OLS (FE) 1-yr panel b/t
Polity index	0.094 (1.38)	0.024 (0.34)	0.083 (1.29)	-0.037 (-0.57)	0.073 (1.10)	-0.041 (-0.62)
Ln GDP per capita	5.074** (5.12)	3.414** (3.51)	4.701** (3.92)	1.478 (1.33)	2.824* (2.56)	0.786 (0.76)
Primary school enrollment ratio					-0.014 (-0.65)	-0.019 (-0.95)
Secondary school enrollment ratio					0.078** (5.08)	0.025+ (1.99)
Year dummies			Y	Y	Y	Y
Country dummies		Y		Y		Y
N (per imputed dataset)	4914	4914	4914	4914	4914	4914
Countries	117	117	117	117	117	117

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. Country-year is unit of analysis, and all independent variables are lagged by 1 year. The coefficients are averaged over 5 imputed data sets, and standard errors are robust and imputation-corrected. The time series run from 1970–2011 (balanced panel). The log-transformed GDP p.c. variable is constructed as $\ln(\text{GDP p.c.} + 18000)$, due to the multiple imputation being conducted without bounds and the resulting occurrence of negative GDP p.c. values in the imputed data sets (see Appendix for details on the imputation model). Constant, country dummies and period dummies are omitted from the table.

Table 8.22: Democracy and first differences in mean student achievements test score

	D1 RE 1-yr panel b/t	D2 OLS (FE) 1-yr panel b/t	D3 RE 1-yr panel b/t	D4 OLS (FE) 1-yr panel b/t	D5 RE 1-yr panel b/t	D6 OLS (FE) 1-yr panel b/t
Polity index	-0.009 (-0.63)	-0.036 (-1.13)	-0.011 (-0.79)	-0.062 (-1.56)	-0.034* (-2.03)	-0.066 (-1.64)
Ln GDP per capita	-0.625 (-0.96)	-1.144 (-0.96)	-0.707 (-0.94)	-1.926 (-1.11)	-2.309+ (-1.92)	-2.374 (-1.35)
Primary enrollment					-0.005 (-0.66)	-0.001 (-0.09)
Secondary enrollment					0.026* (2.63)	0.042* (2.39)
Year dummies			Y	Y	Y	Y
Country dummies		Y	Y	Y	Y	Y
N (per imputed dataset)	4914	4914	4914	4914	4914	4914
Countries	117	117	117	117	117	117

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is first differences in mean student achievement test scores. Country-year is unit of analysis, and all independent variables are lagged by 1 year. The coefficients are averaged over 5 imputed data sets, and standard errors are robust and imputation-corrected. The time series run from 1970–2011 (balanced panel). The log-transformed GDP p.c. variable is constructed as $\ln(\text{GDP p.c.} + 18000)$, due to the multiple imputation being conducted without bounds and the resulting occurrence of negative GDP p.c. values in the imputed data sets (see Appendix for details on the imputation model). Constant, country dummies and period dummies are omitted from the table.

Table 8.23: Democracy and mean student achievements test score. 5-year lag

	D1 RE 1-yr panel b/t	D2 OLS (FE) 1-yr panel b/t	D3 RE 1-yr panel b/t	D4 OLS (FE) 1-yr panel b/t	D5 RE 1-yr panel b/t	D6 OLS (FE) 1-yr panel b/t
Polity index	0.094 (1.71)	0.024 (0.41)	0.087 (1.63)	-0.025 (-0.46)	0.082 (1.63)	-0.026 (-0.48)
Ln GDP per capita	5.693** (5.31)	3.908** (3.38)	5.584** (4.30)	2.244 (1.58)	4.014* (2.48)	1.899 (1.04)
Primary enrollment					-0.006 (-0.40)	-0.016 (-1.02)
Secondary enrollment					0.066* (2.57)	0.011 (0.56)
Year dummies			Y	Y	Y	Y
Country dummies		Y	Y	Y	Y	Y
N (per imputed dataset)	4914	4914	4914	4914	4914	4914
Countries	117	117	117	117	117	117

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. Country-year is unit of analysis, and all independent variables are lagged by 5 years. The coefficients are averaged over 5 imputed data sets, and standard errors are robust and imputation-corrected. The time series run from 1970–2011 (balanced panel). The log-transformed GDP p.c. variable is constructed as $\ln(\text{GDP p.c.} + 18000)$, due to the multiple imputation being conducted without bounds and the resulting occurrence of negative GDP p.c. values in the imputed data sets (see Appendix for details on the imputation model). Constant, country dummies and period dummies are omitted from the table.

Table 8.24: Democracy and mean student achievements test score. 10 year lag

	D1 RE 1-yr panel b/t	D2 OLS (FE) 1-yr panel b/t	D3 RE 1-yr panel b/t	D4 OLS (FE) 1-yr panel b/t	D5 RE 1-yr panel b/t	D6 OLS (FE) 1-yr panel b/t
Polity index	0.083* (2.10)	0.006 (0.14)	0.082* (1.97)	-0.032 (-0.77)	0.074 (1.66)	-0.037 (-0.90)
Ln GDP per capita	5.698** (3.65)	3.592* (2.29)	5.409* (2.86)	1.831 (1.03)	3.768* (2.17)	1.662 (0.97)
Primary enrollment					0.019 (0.92)	0.009 (0.47)
Secondary enrollment					0.062* (3.00)	0.006 (0.25)
Year dummies			Y	Y	Y	Y
Country dummies		Y		Y		Y
N (per imputed dataset)	4914	4914	4914	4914	4914	4914
Countries	117	117	117	117	117	117

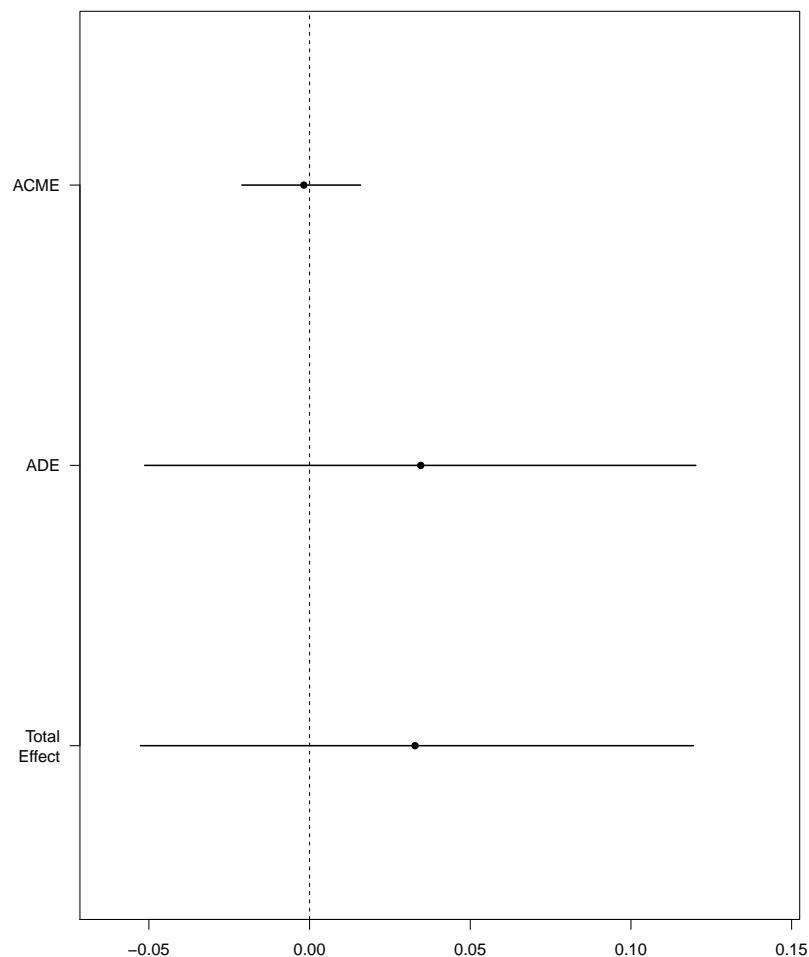
Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. Country-year is unit of analysis, and all independent variables are lagged by 10 years. The coefficients are averaged over 5 imputed data sets, and standard errors are robust and imputation-corrected. The time series run from 1970–2011 (balanced panel). The log-transformed GDP p.c. variable is constructed as $\ln(\text{GDP p.c.} + 18000)$, due to the multiple imputation being conducted without bounds and the resulting occurrence of negative GDP p.c. values in the imputed data sets (see Appendix for details on the imputation model). Constant, country dummies and period dummies are omitted from the table.

8.9 Mediation analysis on democracy, education, and growth

This section presents results from simple mediation analysis, as proposed by Imai et al. (2011). The mediation analysis is estimated using the mediation package in R, where the total effect of democracy on economic growth is decomposed into the average direct effect of democracy (ADE) on growth and the mediation effect of democracy on growth through education quality (Tingley et al., 2014). The latter is referred to as the average causal mediation effect (ACME). Figure 8.9 shows results from mediation analysis estimated with log gdp per capita, log oil and gas income, the gini index and log population as included controls. In brief, no indirect effects of democracy on growth through education quality is identified. Nor is there any direct effect of democracy on growth.

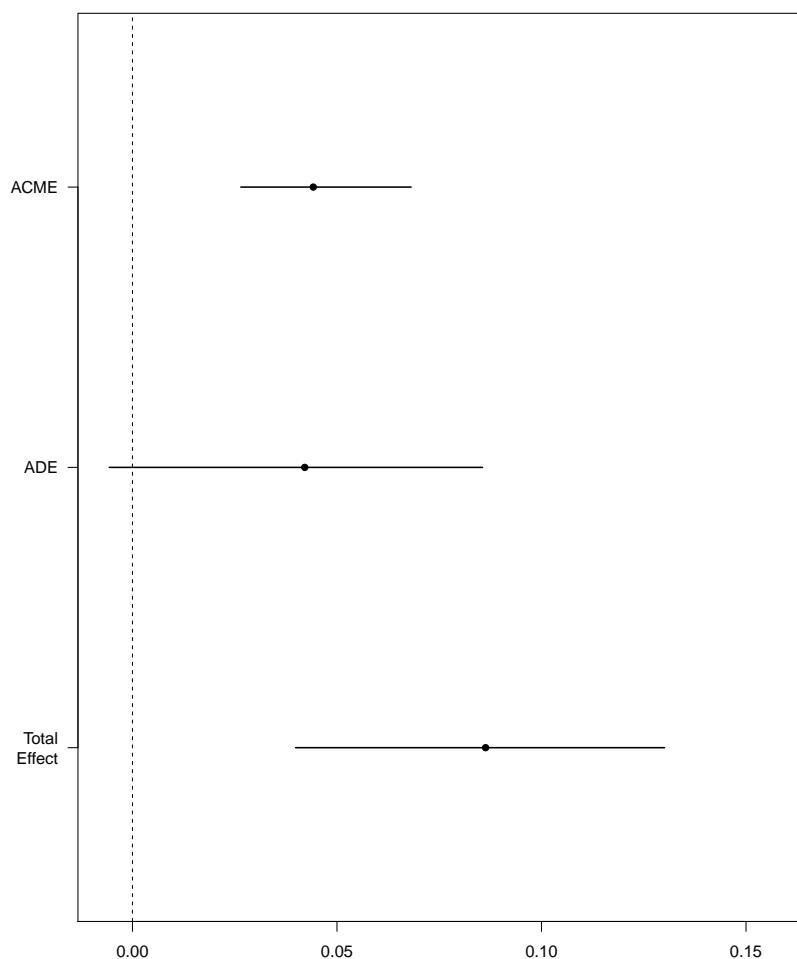
In contrast, there is evidence that democracy impacts on economic growth through education quantity. In figure 8.10, the indicator of education quality is replaced by an indicator of education quantity, years of education. Hence, this mediation analysis model decomposes the total effect of democracy on economic growth into the average direct effect of democracy (ADE) on growth and the mediation effect of democracy on growth through education *quantity*. In brief, this model identifies a mediation effect of democracy on economic growth through education quantity. The average causal mediation effect is estimated to 0.05, and is significantly different from zero. Moreover, it estimates that 51 percent of the total effect of democracy on economic growth is through education quantity as a mediating variable.

Figure 8.9: Mediation effects of democracy on economic growth, mediated through education quality



Notes: The total effect of democracy on economic growth is decomposed into the average direct effect (ADE) and the average causal mediation effect (ACME). The effects are plotted with their 95 percent confidence intervals. Education quality is measured using the indicator of mean student achievement test scores from Angrist et al, and democracy is measured using the Polity index. The pre-treatment covariates are log gdp per capita, log oil and gas income, the gini index and log population.

Figure 8.10: Mediation effects of democracy on economic growth, mediated through education quantity



Notes: The total effect of democracy on economic growth is decomposed into the average direct effect (ADE) and the average causal mediation effect (ACME). The effects are plotted with their 95 percent confidence intervals. Education quantity is measured using mean years of education from the Barro-Lee dataset, and democracy is measured using the Polity index. The pre-treatment covariates are log gdp per capita, log oil and gas income, the gini index and log population.

8.10 Extension: Democracy and education quality by subject and level of education

In this section we present the first extension discussed in the paper, which considers how democracy relates to measures of education quality from the Angrist (2013) dataset that are restricted to particular subjects. These subjects are mathematics, natural sciences and reading skills, and these measures are disaggregated to cover the primary and secondary level of schooling separately.

In Table 8.25 we replicate a Random Effects model – which is more efficient than a corresponding Fixed Effects model, but potentially biased – from Table 3.3 (Model C3) for the six disaggregated components. They provide no evidence that democracy is positively related to math skills. However, we do find indications that democracy is positively related to reading skills, both at the primary (Model D2) and secondary level (D5); the coefficient estimates for Polity are positive and significant at 1%. Meanwhile, there is no effect of democracy on science achievements at the secondary level (D6), but the coefficient estimate for democracy is positive and significant for primary-level science achievements (D3). Nonetheless, these results are far from robust.

In Table 8.26 we replicate Model C4 from Table 3.3, which is a Fixed Effects model, for the various sub-component measures. When including country dummies, democracy is no longer systematically related to knowledge and skills in any of the three subjects, neither at the primary nor at the secondary level, with Polity t-values ranging from -1.2 to +1.0. Hence, while there are some indications that democracy is good for stimulating reading skills, the evidence is far from robust and disappears when we account for country-fixed effects.

Table 8.25: Democracy and subcomponents of student test scores. Random effects models.

	D1 Maths primary 1-yr panel b/t	D2 Reading primary 1-yr panel b/t	D3 Science primary 1-yr panel b/t	D4 Maths secondary 1-yr panel b/t	D5 Reading secondary 1-yr panel b/t	D6 Science secondary 1-yr panel b/t
Polity index	0.0531 (1.04)	0.118** (2.73)	0.173*** (3.84)	0.0585 (0.94)	0.132** (3.24)	0.0480 (1.05)
Ln GDP per capita	3.900* (2.42)	5.175*** (4.10)	4.268*** (4.01)	4.958*** (3.75)	4.898*** (4.27)	4.300** (3.11)
Year dummies	Y	Y	Y	Y	Y	Y
N (per imputed dataset)	4914	4914	4914	4914	4914	4914
Countries	117	117	117	117	117	117

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. Country-year is unit of analysis, and all independent variables are lagged by 1 year. The coefficients are averaged over 5 imputed datasets, and standard errors are robust and imputation-corrected. The time series run from 1970–2011 (balanced panel). The log-transformed GDP p.c. variable is constructed as $\ln(\text{GDP p.c.} + 18000)$, due to the multiple imputation being conducted without bounds and the resulting occurrence of negative GDP p.c. values in the imputed datasets (see appendix for details on the imputation model). Constant and year dummies are omitted from the table.

Table 8.26: Democracy and subcomponents of student test scores. OLS (Fixed effects) models.

	E1 Maths primary 1-yr panel b/t	E2 Reading primary 1-yr panel b/t	E3 Science primary 1-yr panel b/t	E4 Maths secondary 1-yr panel b/t	E5 Reading secondary 1-yr panel b/t	E6 Science secondary 1-yr panel b/t
Polity index	-0.0563 (-0.97)	0.0230 (0.55)	0.0480 (1.02)	-0.0539 (-0.85)	0.0398 (0.97)	-0.0483 (-1.18)
Ln GDP per capita	1.290 (1.00)	2.240+ (1.82)	1.082 (1.05)	2.013+ (1.71)	2.356* (2.48)	1.625 (1.23)
Constant	27.05* (2.02)	13.97 (1.11)	29.71** (2.81)	30.29* (2.47)	22.39* (2.17)	29.47* (2.19)
Year dummies	Y	Y	Y	Y	Y	Y
Country dummies	Y	Y	Y	Y	Y	Y
N (per imputed dataset)	4914	4914	4914	4914	4914	4914
Countries	117	117	117	117	117	117

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is mean student achievement test scores. Country-year is unit of analysis, and all independent variables are lagged by 1 year. The coefficients are averaged over 5 imputed datasets, and standard errors are robust and imputation-corrected. The time series run from 1970–2011 (balanced panel). The log-transformed GDP p.c. variable is constructed as $\ln(\text{GDP p.c.} + 18000)$, due to the multiple imputation being conducted without bounds and the resulting occurrence of negative GDP p.c. values in the imputed datasets (see appendix for details on the imputation model). Constant, country dummies and period dummies are omitted from the table.

8.11 Extensions: Tests on variance between and within countries, by regime type, and tests on within-population variability in PISA 2012 test scores

This section pertains to the various extensions investigating relationships between democracy and the variability of education quality, both at the macro-level as well as for variation within populations.

Table 8.27 reports F-statistics from so-called Goldfeld-Quandt tests. These tests assesses whether there is a systematic variation in regression residuals between relatively democratic and relatively autocratic countries. The cut-off on the Polity Index used here is 6 – observations scoring higher than or equal to 6 are considered relatively democratic regimes. The upper rows report a test where the residuals come from a cross-section OLS regression on the final time-period (2005–2009) with Ln GDP per capita as control variable, thus yielding a test on differential (between-country) variation between the group of autocratic countries and the group of democratic countries. The lower rows report a test where residuals are drawn from a fixed-effects OLS regression, again with Ln GDP per capita as a control, over all time periods. Hence, this provides a test on differential within-country variation (over time) in autocracies and in democracies. While the differences in variation are not huge, they are statistically significant. This suggests that education quality displays more variation in autocracies than in democracies.

Table 8.27: Goldfeld-Quandt test results, based on residuals from cross-section OLS (between-country variance, 2005–2009) and OLS fixed effects regression (within-country variance, all time periods).

	Sum square residuals	Number of obs.	F-value (aut./dem.)	Critical F-value (5 percent)
Between-country variance test				
Autocracies	1367.1	26		
Democracies	2147.6	65	1.8	1.7
Within-country variance test				
Autocracies	10840.4	79		
Democracies	17067.7	213	1.8	1.3

Notes: The regressions used for generating the residuals control for Ln GDP per capita. The regimes are grouped so that countries with Polity Index score lower than 6 are considered autocracies, and equal to or higher than 6 are considered democracies. The critical F-values are approximate (interpolated from F-statistic tables), and correspond to the 5 percent significance level

Tables 8.28–8.30 provide correlations between regime type and different measures

based on data from the PISA 2012 test. Table 8.28 display cross-section regression results on variability in performance within a country's population (measured as the standard error of individual PISA tests scores in the country). We ran models separately for test score results in mathematics, science, and reading. Both regressions controlling for and not controlling for the mean test score of the country are reported. Table 8.29 displays regression results on absolute gender differences (point differential between average male student and average female student) for the 2012 PISA tests. Table 8.30 takes the mean PISA 2012 test score as dependent variable, hence testing our main hypotheses on a smaller sample using a different measure. Regime type is not found to matter, neither for individual variability in tests scores, gender differences, nor mean PISA test scores.

Table 8.28: Democracy and standard error in PISA 2012 test scores within countries' student populations

Test subject	D1 Math OLS b/(t)	D2 Science OLS b/(t)	D3 Reading OLS b/(t)	D4 Math OLS b/(t)	D5 Science OLS b/(t)	D6 Reading OLS b/(t)
Polity Index	-0.241 (-0.92)	-0.241 (-0.96)	-0.014 (-0.04)	-0.161 (-0.57)	-0.240 (-0.93)	-0.006 (-0.02)
Ln GDP per capita	7.218** (4.22)	9.835** (4.95)	5.214+ (1.92)	2.855 (1.41)	9.693** (4.62)	7.217* (2.64)
PISA math mean				0.090** (3.90)		
PISA science mean					0.003 (0.14)	
PISA reading mean						-0.047 (-1.53)
Constant	23.693 (1.54)	-0.612 (-0.03)	44.014+ (1.80)	21.357 (1.51)	-0.808 (-0.04)	47.581+ (1.95)
N	56	56	56	56	56	56

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Cross-section OLS regressions (with robust errors). T-values are reported in parentheses. Polity and GDP per capita are lagged by 5 years. Dependent variable is standard error in on PISA 2012 test scores for the country's student population for PISA 2012, for particular subject (see top row).

Table 8.29: Democracy and absolute gender difference in PISA 2012 test scores

Test subject	D1 Math OLS b/(t)	D2 Science OLS b/(t)	D3 Reading OLS b/(t)	D4 Math OLS b/(t)	D5 Science OLS b/(t)	D6 Reading OLS b/(t)
Polity Index	0.130 (0.72)	-0.590 (-1.65)	-0.494 (-1.10)	0.083 (0.50)	-0.605* (-2.14)	-0.488 (-1.19)
Ln GDP p.c.	-0.216 (-0.14)	0.002 (0.00)	2.449 (0.80)	2.382 (1.39)	2.813 (1.54)	3.930 (1.16)
PISA math mean					-0.054** (-2.95)	
PISA science mean						-0.063* (-2.56)
PISA reading mean						-0.035 (-0.74)
Constant	11.568 (0.81)	12.115 (0.72)	19.599 (0.70)	12.960 (0.93)	16.005 (1.03)	22.238 (0.77)
N	56	56	56	56	56	56

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Cross-section OLS regressions (with robust errors). T-values are reported in parentheses. Polity and GDP per capita are lagged by 5 years. Dependent variable is absolute gender difference in PISA 2012 test scores for particular subject (see top row).

Table 8.30: Democracy and mean PISA 2012 score

Test subject	D1 Math OLS b/(t)	D2 Science OLS b/(t)	D3 Reading OLS b/(t)
Polity Index	-0.885 (-0.37)	-0.247 (-0.12)	0.166 (0.09)
Ln GDP per capita	48.536** (4.51)	44.945** (4.45)	42.595** (4.87)
Constant	26.000 (0.26)	62.202 (0.65)	75.845 (0.91)
N	56	56	56

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Cross-section OLS regressions (with robust errors). T-values are reported in parentheses. Polity and GDP per capita are lagged by 5 years. Dependent variable is mean PISA 2012 test scores for particular subject (see top row).

9 Appendix to Chapter 4

9.1 Introduction

This appendix presents numerous additional tests, both to assess the robustness of the main findings to alternative model specifications and to explore further what mechanisms are driving the main results.

Below, I start by presenting a few robustness tests of the results from table 4.1 in the paper, which replicates the relationship between education and democratization. Second, I present robustness tests of the findings from table 4.2 in the paper, investigating the relationship between education and democratization in authoritarian regimes or autocratic breakdown. This includes tests with alternative operationalizations of both education and democratization (such as measures of democratic upturns) and models with additional control variables. Third, I present robustness tests for models investigating the relationship between educated protesters and autocratic breakdown and democratization in dictatorships, based on the main results that are presented in table 4.3 in the paper. In addition to presenting models with different operationalization of democratization, alternative lag structures and alternative samples, I here present results from models with different measures of involvement by educated protesters in campaigns. Finally, I present additional tests seeking to explore some proposed mechanisms that may account for the finding that education enrollment is negatively related to autocratic breakdown.

Table 9.1: Changes in democracy score and education. Average year of education as explanatory variable.

	1 Electoral democracy	2 Electoral democracy	3 Electoral democracy	4 Electoral democracy
Electoral democracy index	-0.0168*** (-6.21)	-0.0426*** (-7.46)	-0.0466*** (-7.66)	-0.0480*** (-7.85)
Years of education	0.00125*** (5.48)	0.00102 (1.32)	0.000819 (0.88)	-0.000785 (-0.82)
4 GDP p.c.(log)		-0.00164 (-0.59)	-0.00130 (-0.43)	0.000195 (0.06)
Urbanization		0.0389*** (3.34)	0.0453** (3.26)	0.0351* (2.37)
Ln oil and gas income p.c.		-0.000171 (-0.35)	-0.0000841 (-0.15)	-0.000124 (-0.19)
Country dummies?	no	no	yes	yes
Year dummies?	no	no	no	yes
N	12939	5522	5522	5522

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is the Electoral democracy index from V-Dem (Coppedge et al., 2015). Education is measured using the average years of education indicator (for citizens between 15–64) from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.2: Changes in democracy score and education. Polity as dependent variable

	1 Polity	2 Polity	3 Polity	4 Polity
Polity (t-1)	0.000118 (1.23)	0.0000991 (0.68)	0.0000726 (0.47)	0.0000985 (0.66)
School enrollment	-0.00305 (-1.26)	-0.0241** (-2.61)	-0.0324** (-2.82)	-0.0269+ (-1.75)
GDP p.c.(log)		-0.00215 (-0.62)	-0.000845 (-0.21)	-0.000458 (-0.12)
Urbanization		0.0332* (2.47)	0.0398* (2.50)	0.0299* (2.15)
Ln oil and gas income p.c.		0.000799* (2.00)	0.000881+ (1.95)	0.000582 (1.11)
Country dummies?	no	no	yes	yes
Year dummies?	no	no	no	yes
N	8509	4821	4821	4821
R ²			0.00474	0.0781

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is the Polity Democracy index (Marshall, 2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro and Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

9.2 Replicating the education-democracy relationship

Table 9.3: Changes in democracy score and education. Polity index as dependent variable and average year of education as explanatory variable

	1 Polity index	2 Polity index	3 Polity index	4 Polity index
Polity index (t-1)	0.000128+ (1.74)	0.0000632 (0.46)	0.0000282 (0.19)	0.000107 (0.76)
Years of education	-0.000245 (-1.38)	-0.000160 (-0.27)	-0.000226 (-0.32)	0.00105 (1.33)
GDP p.c.(log)		-0.00521+ (-1.88)	-0.00520+ (-1.73)	-0.00321 (-1.04)
Urbanization		0.0229* (2.23)	0.0250* (2.08)	0.0229+ (1.89)
Ln oil and gas income p.c.		0.000744+ (1.83)	0.000816+ (1.76)	0.000608 (1.19)
Country dummies?	no	no	yes	yes
Year dummies?	no	no	no	yes
N	10262	5399	5399	5399
R ²			0.00294	0.0768

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is the Polity Democracy index (Marshall, 2014). Education is measured using the average years of education indicator (for citizens between 15-64) from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.4: Democratic upturns and education.

	1 Electoral democracy (upturns)	2 Electoral democracy (upturns)	3 Electoral democracy (upturns)	4 Electoral democracy (upturns)
Electoral democracy index	-0.0235*** (-6.82)	-0.0252*** (-6.04)	-0.0445*** (-7.94)	-0.0472*** (-8.35)
enrollment	0.0216*** (5.92)	0.00810 (1.63)	-0.00484 (-0.43)	-0.0301+ (-1.86)
GDP p.c.(log)		-0.00195 (-1.12)	-0.00353 (-0.99)	-0.00150 (-0.42)
Urbanization		0.0236*** (3.52)	0.0712** (5.29)	0.0429*** (3.36)
Ln oil and gas income p.c.		-0.00411 (-1.30)	-0.00133 (-0.22)	-0.00164 (-0.25)
Country dummies?	no	no	yes	yes
Year dummies?	no	no	no	yes
N	12939	5522	5522	5522

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is positive changes in the Electoral democracy index from V-Dem. School enrollment is the average of primary, secondary and tertiary enrollment, from Barro and Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

9.3 Country-level education in authoritarian regimes

Table 9.5: Democratic upturns in authoritarian regimes.

	1 Electoral democracy (upturns)	2 Electoral democracy (upturns)	3 Electoral democracy (upturns)	4 Electoral democracy (upturns)
Electoral democracy index	0.000448 (0.07)	0.00447 (0.61)	-0.00358 (-0.42)	-0.0104 (-1.10)
School enrollment	0.0268*** (6.11)	0.0148+ (1.81)	-0.00142 (-0.10)	-0.0460* (-2.46)
GDP p.c.(log)		0.00166 (0.72)	0.00333 (1.06)	0.00741* (2.26)
Urbanization		0.0226* (2.48)	0.0618*** (3.62)	-0.00306 (-0.15)
Ln oil and gas income p.c.		-0.00825+ (-1.68)	-0.00537 (-0.73)	-0.00547 (-0.73)
Country dummies?	no	no	yes	yes
Year dummies?	no	no	no	yes
N	4331	3575	3575	3575

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is positive changes in the Electoral democracy index from V-Dem. School enrollment is the average of primary, secondary and tertiary enrollment, from Barro and Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.6: Democratization and regime breakdown in authoritarian regimes. Sample consisting of autocracies according to Geddes et al.

DV:	1 Electoral democracy	2 Electoral democracy	3 Electoral democracy	4 Electoral democracy	5 Autocratic breakdown	6 Autocratic breakdown	7 Autocratic breakdown	8 Autocratic breakdown
Electoral democracy_t-1	0.994*** (109.31)	0.996*** (103.04)	0.984*** (84.31)	0.972*** (76.75)	2.381*** (4.68)	2.289*** (4.32)	2.852*** (3.42)	2.997** (3.01)
School enrollment	0.0392*** (5.10)	0.0163 (1.29)	0.00737 (0.36)	-0.0547* (-2.08)	-1.331* (-2.51)	-1.692* (-2.28)	-6.888** (-2.96)	-5.629* (-2.03)
Ln GDP p.c.		-0.00271 (-0.83)	-0.00277 (-0.64)	0.00350 (0.78)		0.0428 (0.21)	-0.152 (-0.36)	-0.177 (-0.36)
Urbanization		0.0583*** (4.07)	0.111*** (4.46)	0.0327 (1.20)		0.896 (1.14)	4.493+ (1.70)	5.387 (1.57)
Ln oil and gas income		-0.0104 (-1.49)	-0.00585 (-0.58)	-0.00942 (-0.92)		-0.414 (-1.04)	-0.599 (-0.53)	-0.587 (-0.48)
Country dummies?			yes	yes			yes	yes
Year dummies?				yes			yes	yes
Time, Time ² , Time ³ ?					yes	yes	yes	yes
N	3034	2765	2765	2765	3000	2765	2190	2179

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8. Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.7: Democratization in authoritarian regimes. Democratic transitions according to Boix, Rosato and Miller (2013)

DV:	1 Electoral democracy	2 Electoral democracy	3 Electoral democracy	4 Electoral democracy	5 Electoral democracy
School enrollment	0.123*** (6.69)	0.0336 (0.71)	0.0167 (0.30)	0.0144 (0.53)	-0.189** (-2.69)
Ln GDP p.c.		0.00623 (0.57)	-0.00172 (-0.15)	0.0231** (3.15)	0.0113 (0.92)
Urbanization		0.123* (2.30)	0.156* (2.50)	-0.00495 (-0.19)	-0.0721 (-0.96)
Ln oil and gas income		-0.00882 (-0.35)	0.0180 (0.66)	-0.0502*** (-3.74)	0.0217 (0.77)
Country dummies?			yes		yes
Year dummies?				yes	yes
N	4367	3602	3602	3602	3602

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is democratic transitions from Boix, Miller and Rosato (2013). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.8: Democratization and regime breakdown in authoritarian regimes. Education measured using average years of education from Barro-Lee.

DV:	1 Electoral democracy	2 Electoral democracy	3 Electoral democracy	4 Electoral democracy	5 Autocratic breakdown	6 Autocratic breakdown	7 Autocratic breakdown	8 Autocratic breakdown
Electoral democracy_t-1	0.961*** (172.24)	0.957*** (145.48)	0.912*** (102.97)	0.881*** (89.75)	2.889*** (6.08)	2.596*** (5.17)	3.627*** (5.18)	4.524*** (5.35)
Years of education	0.00160*** (5.86)	0.00164*** (3.82)	0.00323** (2.99)	-0.00239+ (-1.93)	-0.0816** (-2.67)	-0.0721+ (-1.81)	-0.182 (-1.48)	-0.114 (-0.80)
Ln GDP p.c.		-0.00109 (-0.64)	0.000631 (0.22)	0.00798** (2.59)		0.126 (0.82)	-0.262 (-0.81)	-0.276 (-0.76)
Urbanization		0.0180** (2.77)	0.0344* (2.21)	-0.0129 (-0.70)		-0.0878 (-0.14)	2.059 (1.13)	3.061 (1.22)
Ln oil and gas income		-0.0113** (-0.013)	-0.00479 (-0.00479)	0.00267 (0.00267)		-0.472 (-0.472)	-0.699 (-0.699)	-0.754 (-0.754)
Country dummies?			yes	yes			yes	yes
Year dummies?				yes			yes	yes
Time, Time ² , Time ³ ?					yes	yes	yes	yes
N	5636	4634	4634	4634	3945	3682	2949	2905

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8.

Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). Education is measured the average years of education indicator (for age 15-64), from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.9: Democratization and regime breakdown in authoritarian regimes. Additional control variables

DV:	1 Electoral democracy	2 Electoral democracy	3 Electoral democracy	4 Electoral democracy	5 Autocratic breakdown	6 Autocratic breakdown	7 Autocratic breakdown	8 Autocratic breakdown
Electoral democracy_t-1	0.941*** (121.94)	0.895*** (50.26)	0.859*** (41.62)	0.850*** (38.37)	2.844*** (5.34)	2.416** (2.89)	3.486** (2.91)	4.050* (2.47)
School enrollment	0.0320*** (6.03)	0.0905*** (3.78)	0.0675* (2.13)	-0.0416 (-0.92)	-1.691** (-3.20)	-3.276** (-2.88)	-7.119* (-2.06)	-16.81** (-3.16)
Ln GDP p.c.		0.00208 (0.36)	-0.00205 (-0.29)	0.00868 (1.15)		1.077*** (3.63)	2.139** (2.70)	4.362*** (4.11)
Urbanization		0.000246 (0.01)	0.0867* (2.25)	-0.0126 (-0.27)		-2.114* (-2.04)	-0.730 (-0.17)	-6.306 (-1.11)
Ln oil and gas income		-0.00977 (-0.80)	-0.00391 (-0.25)	-0.00864 (-0.51)		-0.570 (-1.09)	-2.057 (-1.21)	-4.967* (-2.41)
Population total		-2.41e-11 (-0.99)	-2.93e-11 (-0.74)	-7.60e-11+ (-1.88)		-4.10e-09 (-1.19)	-1.09e-09 (-0.06)	-4.10e-08+ (-1.82)
Income inequality, Gini		0.000404 (1.60)	0.000727* (2.30)	0.000547+ (1.71)		-0.0203+ (-1.82)	0.0177 (0.58)	-0.0486 (-1.29)
GDP Growth (rescaled)		-0.000496+ (-1.65)	-0.000337 (-1.10)	-0.000264 (-0.85)		-0.0773*** (-4.47)	-0.101*** (-4.43)	-0.114*** (-4.34)
Armed conflict, internal		-0.00335 (-0.66)	-0.00523 (-0.99)	-0.0101+ (-1.87)		0.429 (1.62)	0.444 (1.08)	0.496 (1.06)
Country dummies?			yes	yes			yes	yes
Year dummies?				yes			yes	yes
Time, Time ² , Time ³ ?					yes	yes	yes	yes
N	4331	1542	1542	1542	2856	1524	1028	971

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8.

Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.10: Democratization and regime breakdown in authoritarian regimes. Controlling for authoritarian regime type, from Geddes, Wright and Frantz (2014).

DV:	1 Electoral democracy	2 Electoral democracy	3 Electoral democracy	4 Electoral democracy	5 Autocratic breakdown	6 Autocratic breakdown	7 Autocratic breakdown	8 Autocratic breakdown
Electoral democracy_t-1	0.976*** (95.77)	0.973*** (92.24)	0.931*** (63.88)	0.904*** (57.38)		2.324*** (3.77)	3.193*** (3.51)	3.875*** (3.45)
School enrollment	0.0290*** (4.01)	0.0242* (2.25)	0.0131 (0.58)	-0.0717* (-2.54)	-2.310*** (-4.02)	-3.172*** (-4.14)	-6.339** (-2.78)	-5.975* (-2.19)
Party regime	0.0236*** (4.76)	0.0244*** (4.78)	0.0279*** (4.37)	0.0275*** (4.35)	-2.207*** (-8.05)	-1.815*** (-6.14)	-1.260** (-3.13)	-1.304** (-2.96)
Military regime	0.0480*** (8.60)	0.0469*** (8.21)	0.0559*** (8.31)	0.0527*** (7.91)	-0.535* (-2.19)	-0.116 (-0.42)	0.170 (0.48)	0.234 (0.60)
Monarchy	0.0210** (3.17)	0.0200** (2.93)	0.0399*** (3.73)	0.0398*** (3.75)	-2.549*** (-5.39)	-1.985*** (-3.97)	-0.844 (-1.05)	-0.839 (-0.98)
Personalist regime	0.0272*** (5.24)	0.0268*** (4.98)	0.0305*** (4.58)	0.0239*** (3.60)	-1.163*** (-4.64)	-0.711* (-2.54)	-0.351 (-0.92)	-0.356 (-0.84)
Ln GDP p.c.		0.000154 (0.05)	-0.0000662 (-0.01)	0.00779 (1.59)		0.490* (2.47)	0.565 (1.33)	0.543 (1.10)
Urbanization		0.0167 (1.39)	0.104*** (3.69)	0.00436 (0.14)		-0.899 (-1.13)	1.132 (0.41)	1.355 (0.39)
Ln oil and gas income		-0.00891 (-1.48)	-0.00722 (-0.67)	-0.00654 (-0.59)		-0.133 (-0.33)	-0.989 (-0.97)	-1.255 (-1.15)
Country dummies?			yes	yes			yes	yes
Year dummies?				yes			yes	yes
Time, Time ² , Time ³ ?					yes	yes	yes	yes
N	2855	2743	2743	2743	2864	2743	2161	2133

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8.

Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

9.4 Educated protest campaigns in authoritarian regimes

Table 9.11: Democratic upturns in authoritarian regimes.

	1 Electoral democracy (upturns)	2 Electoral democracy (upturns)	3 Electoral democracy (upturns)	4 Electoral democracy (upturns)
Educated protesters (index)	0.0121*** (6.16)	0.0124*** (5.69)	0.0133*** (5.68)	0.0113*** (4.89)
Protest campaign	0.00960*** (3.45)	0.00963** (2.98)	0.00936** (2.64)	0.00704* (1.99)
School enrollment	0.0208*** (4.74)	0.00615 (0.74)	-0.0132 (-0.90)	-0.0443* (-2.40)
Ln GDP p.c.		0.00293 (1.27)	0.00519+ (1.68)	0.00846** (2.61)
Urbanization		0.0225* (2.46)	0.0582*** (3.47)	0.00417 (0.21)
Ln oil and gas income		-0.00813+ (-1.65)	-0.00562 (-0.78)	-0.00433 (-0.59)
Electoral democracy_t-1	-0.000104 (-0.02)	0.00415 (0.57)	-0.00441 (-0.53)	-0.0110 (-1.18)
Country dummies?	no	no	yes	yes
Year dummies?	no	no	no	yes
N	4331	3575	3575	3575

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is positive changes in Electoral democracy index from V-Dem. Education is measured using school enrollment from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.12: Democratization and regime breakdown in authoritarian regimes. Sample consisting of autocracies according to Geddes et al.

DV:	C1	2	3	4	5	6	7	8
	Electoral democracy	Electoral democracy	Electoral democracy	Electoral democracy	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown
Educated protesters (index)	0.0146*** (5.41)	0.0145*** (5.14)	0.0139*** (4.57)	0.0133*** (5.37)	0.442*** (-0.03)	0.461*** (3.13)	0.323 (3.10)	0.340 (1.44)
Protest campaign	0.00609 (1.51)	0.00626 (1.44)	0.00725 (1.51)	-0.000101 (-0.03)	0.463+ (1.99)	1.008* (1.84)	1.196** (2.48)	1.143 (2.76)
School enrollment	0.0322*** (4.22)	0.00431 (0.34)	-0.00931 (-0.45)	-0.01117 (-1.24)	-1.507** (-2.69)	-2.254** (-2.95)	-7.260** (-3.08)	-4.729+ (-1.69)
Ln GDP p.c.		-0.00155 (-0.48)	-0.00120 (-0.28)	0.00215 (0.88)		0.0960 (0.46)	-0.0495 (-0.12)	-0.178 (-0.37)
Urbanization		0.0563*** (3.98)	0.108*** (4.41)	0.0331*** (3.52)		0.944 (1.18)	2.984 (1.10)	6.530+ (1.86)
Ln oil and gas income		-0.00966 (-1.40)	-0.00476 (-0.48)	-0.0125** (-2.69)		-0.495 (-1.24)	-0.819 (-0.72)	-0.682 (-0.54)
Electoral democracy_t-1	0.992*** (110.41)	0.993*** (104.04)	0.980*** (84.88)	0.998*** (121.93)	2.320*** (4.29)	2.614** (3.16)	2.734** (2.74)	
Country dummies?		yes	yes	yes		yes	yes	yes
Year dummies?								
Time, Time ² , Time ³ ?		3034	2765	2765	3006	2765	2190	2179
N								

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8. Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.13: Democratization in authoritarian regimes. Democratic transitions according to Boix, Miller and Rosato (2013)

DV:	1 Electoral democracy	2 Electoral democracy	3 Electoral democracy	4 Electoral democracy	5 Electoral democracy
Educated protesters (index)	0.0412*** (5.46)	0.0423*** (4.91)	0.0414*** (4.74)	0.0409*** (4.66)	0.0364*** (4.73)
Protest campaign	0.0123 (1.13)	0.0108 (0.83)	0.0103 (0.78)	0.00347 (0.26)	0.00520 (0.46)
School enrollment	0.0974*** (5.27)	0.00456 (0.10)	-0.0135 (-0.25)	-0.186** (-2.66)	-0.000293 (-0.01)
Ln GDP p.c.		0.00993 (0.92)	0.00216 (0.19)	0.0133 (1.08)	0.0233** (3.19)
Urbanization		0.115* (2.18)	0.149* (2.40)	-0.0513 (-0.69)	0.00230 (0.09)
Ln oil and gas income		-0.00903 (-0.37)	0.0176 (0.65)	0.0245 (0.88)	-0.0513*** (-3.85)
Country dummies?			yes		yes
Year dummies?				yes	yes
N	4367	3602	3602	3602	3602

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses. Dependent variable is democratic transitions from Boix, Miller and Rosato (2013). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.14: Democratization and regime breakdown in authoritarian regimes. Controlling for years of education.

DV:	C1	2	3	4	5	6	7	8
	Electoral democracy	Electoral democracy	Electoral democracy	Electoral democracy	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown
Educated protesters (index)	0.0133*** (6.70)	0.0129*** (5.90)	0.0158*** (6.18)	0.0144*** (5.70)	0.509*** (4.02)	0.472*** (3.53)	0.484* (2.53)	0.434* (2.14)
Protest campaign	0.00114 (0.45)	0.000381 (0.13)	-0.000771 (-0.21)	-0.00327 (-0.91)	0.432* (2.07)	0.581** (2.62)	0.861** (2.78)	0.996** (3.05)
Years of education	0.00136*** (4.98)	0.00132** (3.07)	0.00222* (2.06)	-0.00307* (-2.48)	-0.0724* (-2.32)	-0.112** (-2.72)	-0.205+ (-1.69)	-0.104 (-0.73)
Ln GDP p.c.	-0.000982 (-0.57)	0.000694 (0.25)	0.00790* (2.57)	0.00409 (0.82)	0.209 (1.31)	-0.251 (-0.76)	-0.246 (-0.67)	-0.246 (-0.67)
Urbanization	0.0196** (3.03)	0.0386* (2.49)	-0.00409 (-0.22)	-0.0954 (-0.15)	0.821 (0.45)	3.270 (1.30)	3.270 (1.30)	3.270 (1.30)
Ln oil and gas income	-0.0117** (-3.22)	-0.00444 (-0.61)	0.00327 (0.44)	-0.570+ (-1.69)	-0.802 (-1.02)	-0.892 (-1.05)	-0.892 (-1.05)	-0.892 (-1.05)
Electoral democracy_t-1	0.962*** (173.53)	0.958*** (146.45)	0.911*** (103.67)	0.880*** (90.10)	2.897*** (5.72)	3.599*** (5.15)	4.411*** (5.13)	4.411*** (5.13)
Country dummies?					yes	yes	yes	yes
Year dummies?					yes	yes	yes	yes
Time, Time ² , Time ³ ?	4634	4634	4634	3953	3682	2949	2905	
5636								

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8. Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). Education is measured using the average years of education indicator from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.15: Democratization and regime breakdown in authoritarian regimes. Including additional control variables.

DV:	C1	2	3	4	5	6	7	8
Electoral democracy _{t-1}	0.940*** (122.03)	0.891*** (50.10)	0.853*** (41.52)	0.845*** (38.21)	2.574** (3.04)	3.394** (2.84)	4.168*	4.168*
Educated protesters (index)	0.0125*** (5.15)	0.00625 (1.49)	0.00694 (1.58)	0.00527 (1.21)	0.505*** (3.49)	0.474* (2.40)	0.838** (2.82)	0.981** (2.91)
Protest campaign	0.00421 (1.23)	0.0156* (2.19)	0.0157* (2.08)	0.0123 (1.62)	0.394 (1.55)	0.666+ (1.93)	0.130 (0.24)	-0.153 (-0.26)
School enrollment	0.0270*** (5.03)	0.0746*** (3.09)	0.0479 (1.51)	-0.0449 (-1.00)	-1.791** (-3.24)	-4.067*** (-3.56)	-6.904* (-2.04)	-15.95** (-2.91)
Ln GDP p.c.	0.00233 (0.41)	-0.00273 (-0.38)	0.00767 (1.02)	1.041*** (3.42)	2.102** (2.68)	4.005*** (3.75)		
Urbanization	0.00272 (0.11)	0.0884* (2.31)	-0.000538 (-0.01)	-2.073* (-1.98)	-2.964 (-0.68)	-9.321 (-1.55)		
Ln oil and gas income	-0.00916 (-0.75)	-0.00272 (-0.30)	-0.00513 (-0.86)	-0.461 (-1.30)	-2.191 (-2.32)	-4.753* (-2.32)		
Population total	-2.06e-11 (-0.84)	-1.29e-11 (-0.33)	-5.99e-11 (-1.48)	-6.59e-09 (-1.50)	-1.29e-08 (-0.68)	-4.70e-08* (-2.05)		
Income inequality, Gini	0.000396 (1.57)	0.000700* (2.23)	0.000533+ (1.68)	-0.0202+ (-1.77)	0.0186 (0.59)	-0.0412 (-1.06)		
GDP Growth (rescaled)	-0.000317 (-1.05)	-0.000162 (-0.53)	-0.000146 (-0.47)	-0.0705*** (-4.00)	-0.100*** (-4.35)	-0.111*** (-4.12)		
Armed conflict, internal	-0.0107* (-2.04)	-0.0125* (-2.29)	-0.0154** (-2.75)	-0.112 (-0.38)	0.0959 (0.22)	0.181 (0.36)		
Country dummies?	yes	yes	yes	yes	yes	yes	yes	yes
Year dummies?	yes	yes	yes	yes	yes	yes	yes	yes
Time, Time ² , Time ³ ?	4331	1542	1542	2864	1524	1028	971	
N								

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8. Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Figure 9.1: Predicted probabilities for uneducated campaigns vs highly educated campaigns

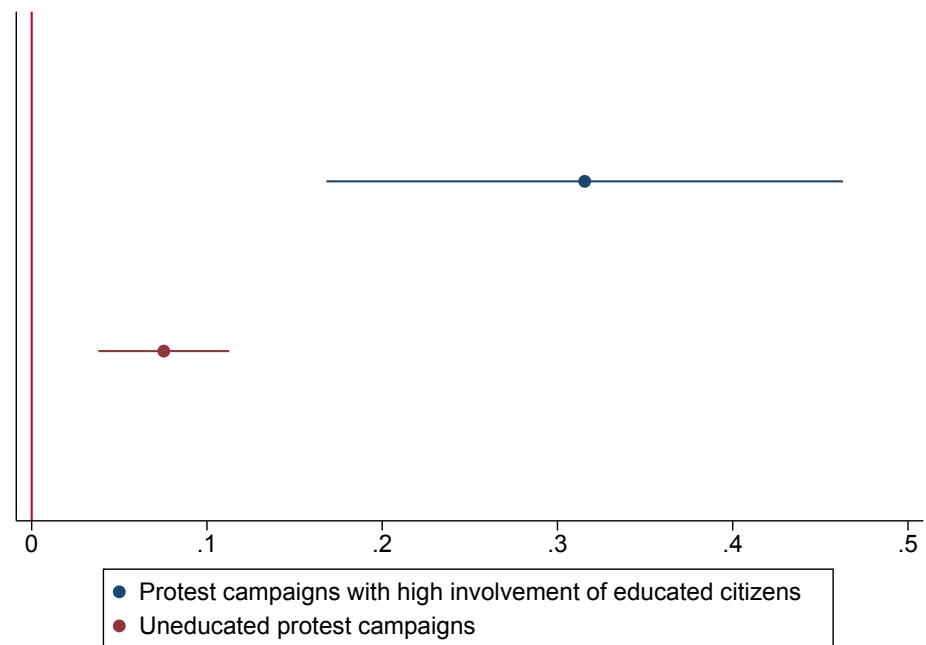


Table 9.16: Democratization and regime breakdown in authoritarian regimes. Controlling for authoritarian regime type

DV:	C1	2	3	4	5	6	7	8
Electoral democracy _{t-1}	0.975*** (95.39)	0.973*** (91.86)	0.930*** (64.46)	0.902*** (57.64)	2.483*** (3.98)	3.050*** (3.39)	3.630*** (3.20)	
Educated protesters (index)	0.0136*** (4.83)	0.0134*** (4.59)	0.0171*** (5.18)	0.0149*** (4.58)	0.565*** (3.80)	0.532*** (3.42)	0.688** (3.02)	0.713** (2.93)
Protest campaign	0.00276 (0.65)	0.00276 (0.61)	0.00122 (0.23)	-0.00150 (-0.29)	0.264 (1.02)	0.419 (1.54)	0.453 (1.12)	0.554 (1.30)
School enrollment	0.0237** (3.22)	0.0175 (1.60)	0.00494 (0.22)	-0.0709* (-2.52)	-2.593*** (-4.48)	-3.672*** (-4.74)	-6.116** (-2.67)	-5.065+ (-1.81)
Party regime	0.0246*** (4.95)	0.0253*** (4.95)	0.0294*** (4.64)	0.0284*** (4.51)	-2.160*** (-7.84)	-1.759*** (-5.94)	-1.103** (-2.73)	-1.125* (-2.56)
Military regime	0.0494*** (8.88)	0.0483*** (8.46)	0.0571*** (8.54)	0.0540*** (8.14)	-0.478+ (-1.92)	-0.0745 (-0.26)	0.231 (0.65)	0.336 (0.85)
Monarchy	0.0247*** (3.70)	0.0236*** (3.42)	0.0534*** (4.97)	0.0499*** (4.68)	-2.348*** (-5.03)	-1.771*** (-3.58)	-0.463 (-0.58)	-0.354 (-0.42)
Personalist regime	0.0274*** (5.30)	0.0271*** (5.03)	0.0295*** (4.48)	0.0231*** (3.50)	-1.148*** (-4.53)	-0.691* (-2.46)	-0.357 (-0.95)	-0.311 (-0.74)
Ln GDP p.c.	0.000101 (0.03)	-0.000310 (-0.07)	0.00757 (1.56)	0.561*** (2.69)	0.508 (1.15)	0.441 (0.87)		
Urbanization	0.0187 (1.53)	0.102*** (3.66)	0.0125 (0.40)	-0.903 (-1.12)	-0.753 (-0.27)	1.471 (0.41)		
Ln oil and gas income	-0.00860 (-1.41)	-0.00779 (-0.73)	-0.00557 (-0.51)	-0.223 (-0.55)	-0.926 (-0.92)	-1.166 (-1.06)		
Country dummies?	yes	yes	yes	yes	yes	yes	yes	
Year dummies?	yes	yes	yes	yes	yes	yes	yes	
Time, Time ² , Time ³ ?	yes	yes	yes	yes	yes	yes	yes	
N	2855	2743	2743	2743	2864	2743	2161	2133

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8. Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.17: Democratization and regime breakdown in authoritarian regimes. Independent variables lagged by five years.

DV:	C1	2	3	4	5	6	7	8
	Electoral democracy	Electoral democracy	Electoral democracy	Electoral democracy	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown
Electoral democracy_t-1	0.590*** (29.91)	0.582*** (26.08)	0.525*** (22.41)	0.454*** (18.21)	0.543 (-1.27)	-0.942 (-1.42)	-0.751 (-2.03)	-0.751 (-0.64)
Educated protesters (index)	0.0126* (2.17)	0.0152* (2.35)	0.0161* (2.47)	0.0103+ (1.68)	-0.263 (-1.27)	-0.297 (-1.42)	-0.569* (-2.03)	-0.541+ (-1.79)
Protest campaign	0.0483*** (5.78)	0.0494*** (5.10)	0.0474*** (4.82)	0.0334*** (3.56)	0.516* (2.04)	0.648* (2.47)	1.155*** (2.76)	1.357*** (2.84)
School enrollment	0.252*** (17.89)	0.141*** (4.42)	0.0375 (0.92)	-0.286*** (-5.83)	-1.823** (-3.19)	-2.096*** (-2.77)	-6.317** (-2.71)	-6.019* (-2.17)
Ln GDP p.c.		0.0102 (1.30)	0.0108 (1.26)	0.0461*** (5.35)		0.244 (1.19)	0.368 (0.86)	0.641 (1.29)
Urbanization		0.253*** (7.02)	0.418*** (8.93)	0.164** (3.12)		0.430 (0.54)	2.156 (0.80)	2.075 (0.59)
Ln oil and gas income		-0.0531** (-3.01)	-0.0240 (-1.20)	-0.0226 (-1.15)		-0.926* (-2.39)	-2.404* (-2.20)	-2.560* (-2.14)
Country dummies?		yes	yes	yes		yes	yes	yes
Year dummies?		yes	yes	yes		yes	yes	yes
Time, Time ² , Time ³ ?		4314	3562	3562		2822	2702	2133
N								1993

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8. Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 5 years. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.18: Democratization and regime breakdown in authoritarian regimes. Protest campaigns with educated participants described as having participated.

DV:	C1	2	3	4	5	6	7	8
	Electoral democracy	Electoral democracy	Electoral democracy	Electoral democracy	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown
Electoral democracy_t-1	0.939*** (121.86)	0.942*** (107.23)	0.915*** (88.36)	0.886*** (77.16)	2.808*** (5.03)	3.335*** (4.19)	3.934*** (3.94)	
Educated protesters participated	0.018*** (3.76)	0.0188*** (3.40)	0.0204*** (3.34)	0.0152** (2.50)	1.041** (3.06)	0.925** (2.60)	1.501** (2.83)	1.476** (2.63)
Protest campaign	0.00542 (1.43)	0.00525 (1.19)	0.00504 (1.01)	0.00260 (0.52)	0.296 (1.00)	0.472 (1.51)	0.407 (0.88)	0.537 (1.11)
School enrollment	0.0287*** (5.36)	0.0246* (2.51)	0.0101 (0.56)	-0.0554* (-2.43)	-1.725** (-3.13)	-2.934*** (-4.05)	-7.030** (-3.13)	-6.572* (-2.43)
Ln GDP p.c.	0.00120 (0.43)	0.00193 (0.51)	0.0123** (3.09)	0.470* (2.25)	0.494 (1.17)	0.405 (0.84)		
Urbanization	0.0208+ (1.91)	0.0556** (2.68)	0.00254 (0.10)	-0.108 (-0.14)	1.798 (0.68)	3.274 (0.96)		
Ln oil and gas income	-0.0115+ (-1.96)	-0.00209 (-0.23)	0.000202 (0.02)	-0.534 (-1.37)	-1.102 (-1.11)	-1.203 (-1.20)		
Country dummies?	yes	yes	yes	yes	yes	yes	yes	
Year dummies?	yes	yes	yes	yes	yes	yes	yes	
Time, Time ² , Time ³ ?								
N	4331	3575	3575	3575	2864	2743	2161	2133

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8. Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.19: Democratization and regime breakdown in authoritarian regimes. Protest campaigns dominated by educated participants.

DV:	C1	2	3	4	5	6	7	8
	Electoral democracy	Electoral democracy	Electoral democracy	Electoral democracy	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown
Electoral democracy_t-1	0.942*** (122.21)	0.945*** (107.58)	0.918*** (88.77)	0.889*** (77.53)	2.893*** (5.19)	3.374*** (4.23)	3.288*** (4.99)	
Educated protesters dominated	0.0368*** (5.13)	0.0382*** (4.99)	0.0413*** (5.16)	0.0370*** (4.68)	0.714+ (1.81)	0.839* (2.08)	0.359 (0.67)	0.628 (1.49)
Protest campaign	0.0118*** (4.46)	0.0116*** (3.84)	0.0125*** (3.85)	0.00709* (2.17)	0.843*** (4.48)	0.944*** (4.74)	1.384*** (4.86)	0.998*** (4.65)
School enrollment	0.0278*** (5.21)	0.0220* (2.24)	0.03668 (0.20)	-0.0578* (-2.54)	-1.620** (-2.97)	-3.006*** (-4.16)	-7.181** (-3.24)	-2.656** (-3.18)
Ln GDP p.c.	0.0154 (0.56)	0.00213 (0.56)	0.0124** (3.12)	0.0124** (2.48)	0.513 (1.20)	0.513 (2.14)	0.499* (2.14)	
Urbanization	0.0222* (2.04)	0.0628** (3.04)	0.0102 (0.42)	-0.150 (-0.19)	-0.150 (0.83)	2.180 (-0.18)	-0.146 (-0.18)	
Ln oil and gas income	-0.0116* (-1.98)	-0.00395 (-0.44)	-0.00157 (-0.17)	-0.521 (-1.35)	-1.093 (-1.10)	-0.568 (-1.43)		
Country dummies?	yes	yes	yes	yes	yes	yes	yes	
Year dummies?	yes	yes	yes	yes	yes	yes	yes	
Time, Time ² , Time ³ ?								
N	4331	3575	3575	3575	2864	2743	2161	2701

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8. Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.20: Democratization and regime breakdown in authoritarian regimes. Alternative education index.

DV:	C1	2	3	4	5	6	7	8
	Electoral democracy	Electoral democracy	Electoral democracy	Electoral democracy	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown	Autocratic breakdown
Electoral democracy t-1	0.259*** (3.55)	0.161* (2.07)	0.00535 (0.06)	-0.200* (-2.06)	-1.631 (-0.81)	-6.481 (-1.54)	-30.95+ (-1.70)	
Education index 2	-0.00112 (-0.16)	0.00235 (0.32)	-0.00630 (-0.77)	0.00243 (0.24)	-0.0809 (0.57)	-0.0620 (-0.42)	-0.733 (-1.14)	
School enrollment	0.669*** (11.13)	0.409*** (3.94)	0.187 (1.34)	-0.411* (-2.09)	-2.587+ (-1.66)	-3.186 (-1.47)	-12.83 (-1.40)	
Ln GDP p.c.		0.0183 (0.59)	0.0280 (0.72)	0.0559 (1.29)	-0.405 (-0.72)	0.114 (0.04)	19.54* (2.14)	
Urbanization		0.428*** (3.52)	0.770*** (4.36)	0.671** (2.64)	3.035 (1.45)	30.10 (1.47)	149.7* (2.06)	
Ln oil and gas income		-0.0685 (-1.04)	0.0714 (0.82)	0.143 (1.44)	-1.083 (-1.12)	1.277 (0.26)	17.09 (1.18)	
Country dummies?	yes	yes	yes	yes	yes	yes	yes	
Year dummies?		yes	yes	yes	yes	yes	yes	
Time, Time ² , Time ³ ?	478	435	435	387	363	184	94	
N								

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. T-values in parentheses in models 1-4 and z-values in parentheses in models 5-8. Dependent variable in models 1-4 is Electoral democracy index from V-Dem, and dependent variable in models 5-8 is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

9.5 Empirical extensions

Table 9.21: Democratization and regime breakdown in authoritarian regimes. Controlling for governance index.

DV:	C1	2	3	4	5	6	7	8
	Autocratic breakdown							
School enrollment	-1.498* (-1.98)	-2.309* (-2.11)	-2.459 (-0.94)	-3.656 (-1.14)	-1.340+ (-1.71)	-2.061+ (-1.89)	-1.940 (-0.73)	-3.401 (-1.06)
Bureaucratic quality					-0.103 (-0.60)	-0.192 (-1.07)	-0.0523 (-0.24)	-0.0190 (-0.08)
Economic policies					-0.0423 (-0.31)	-0.205 (-1.45)	-0.212 (-1.36)	-0.181 (-1.06)
Electoral democracy_t-1	3.232*** (4.33)	2.898** (3.17)	4.270*** (3.66)		3.554*** (4.70)	3.107*** (3.33)	4.310*** (3.65)	
Ln GDP p.c.	0.230 (0.87)	0.286 (0.59)	0.263 (0.48)		0.371 (1.37)	0.436 (0.87)	0.436 (0.72)	0.404 (0.72)
Urbanization	0.000968 (0.00)	-0.932 (-0.28)	-4.234 (-0.89)		-0.148 (-0.14)	-0.860 (-0.25)	-4.315 (-0.91)	
Ln oil and gas income	-0.449 (-0.81)	-0.00624 (-0.00)	-0.502 (-0.37)		-0.562 (-1.03)	-0.141 (-0.11)	-0.587 (-0.43)	
Country dummies?	yes	yes	yes		yes	yes	yes	
Year dummies?								
Time, Time ² , Time ³ ?	yes	yes	yes		yes	yes	yes	
II	-530.8 2412	-505.2 2339	-314.5 1712	-286.1 1712	-530.5 2412	-502.7 2339	-313.3 1712	-285.4 1712
N								

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Dependent variable is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

Table 9.22: Democratization and regime breakdown in authoritarian regimes. Controlling for governance index.

DV:	C1	2	3	4	5	6	7	8
	Autocratic breakdown							
School enrollment	-1.357+ (-1.83)	-1.956+ (-1.82)	-2.319 (-0.84)	-3.665 (-1.06)	-1.177 (-1.58)	-1.836+ (-1.76)	-1.871 (-0.68)	-3.124 (-0.89)
Governance index					-5.017*** (-3.57)	-5.017** (-3.57)	-4.138* (-2.35)	-3.554+ (-1.88)
GDP p.c.(log)	0.114 (0.42)	0.345 (0.67)	0.515 (0.86)	0.477+ (1.68)	0.621 (1.18)	0.745 (1.23)		
Urbanization	0.0550 (0.05)	-1.738 (-0.49)	-5.892 (-1.16)	-0.576 (-0.54)	-3.682 (-1.01)	-6.786 (-1.33)		
Ln oil and gas income p.c.	-0.157 (-0.28)	0.246 (0.18)	-0.104 (-0.07)	-0.459 (-0.83)	0.0567 (0.04)	-0.247 (-0.17)		
Electoral democracy index	3.175*** (4.16)	2.210* (2.28)	3.295* (2.56)	4.358*** (5.25)	3.487*** (3.15)	4.308*** (3.07)		
Country dummies?	yes							
Year dummies?								
Time, Time ² , Time ³ ?	yes							
N	2311	2240	1548	1548	2311	2240	1548	1548
ll	-484.8	-460.8	-279.2	-247.8	-454.4	-454.4	-276.5	-246.1

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Dependent variable is authoritarian regime breakdown from Geddes, Wright and Frantz (2014). School enrollment is the average of primary, secondary and tertiary enrollment, from Barro-Lee (2013). All independent variables are lagged by 1 year. Constant, country dummies, year dummies and time coefficients are omitted from table.

10 Appendix to Chapter 5

10.1 Overview

This appendix contains the following. Section 10.2 evaluates the (relative) substantive effect of our education variable. Section 10.3 presents descriptive statistics for the main variables in the baseline models. Additional information about the data is presented in 10.4. Subsequently, section 10.5 estimates the baseline models with a number of additional controls included; we control for alternative measures for local wealth (all taken from the latest version of the PRIO-GRID database), literacy and media consumption (based on DHS variables), additional measures of urbanity, and a range of types of land-cover (taken from PRIO-GRID) to capture geographic features. In section 10.6 we estimate models where we drop cells with few respondents. Finally, section 10.7 evaluates whether our results are specific to francophone countries, where unilingual education policies are increasingly practiced.

10.2 Evaluating the relative importance of education

This section evaluates the substantive importance of education as a driver of protest activity compared to other relevant factors. This is explicitly comparative, since there is no absolute metric with which to gauge the “importance” of a predictor in this context.

Table 10.1 displays the change in predicted probability of protest for a set of models. In the first row, for each model, we simulate the probability of protest as the value of variable in question (denoted by the heading) moves from its (empirical) minimum to its (empirical) maximum value, holding the other variables (in the baseline model) at their means. This gives us an overview of the *maximum change in protest that each*

variable could potentially yield given our model. In the second row, we simulate the same change in probability as each variable increases by one standard deviation, all other variables kept at their means. This exercise is performed for four variables corresponding to different aspects of economic and social development, that should be expected to be important for affecting the likelihood of protest in a locality. We include urbanization, since that has been shown to induce social instability (Raleigh, 2015). Furthermore, we include unemployment, and the mean value of the “wealth index” used in the DHS surveys (registering the economic resources of respondents), since poverty and economic grievances are often linked to conflict and instability (Collier and Hoeffler, 2004; Miguel, Satyanath and Serengeti, 2004). In other words, we compare the substantive impact of education to some of the most well-established predictors in the literature on political instability and mobilization.

Table 10.1: Substantive effects of different predictors

	Education	Urbanization	Unemployment	Literacy	Wealth index
Min-max increase	0.098	0.011	0.037	0.016	-0.063
Standard deviation increase	0.014	0.005	0.015	0.007	-0.042

Table 10.1 clearly shows the relative import of education when it comes to the risk of protest activity. When education increases from its minimum to its maximum value, the risk of mass protest in a grid-cell increases by 10 percentage points. This is more than the increased risk yielded when we move any of the other development-related variables from their minimum to their maximum values. A min-max increase in urbanization produces a 1 percent increase, while unemployment, urbanization, literacy and the wealth index yield increases of .1, .037, and .016 percent respectively. When we move the given variables by one standard deviation (from the mean) we see a similar pattern: Education is one of the substantively most important variables, yielding a .014 increase in protest probability, larger than both urbanization and literacy. Using the standard deviation-increase as a metric, unemployment and the wealth index perform better, yielding changes of .015 and -.042 respectively. Nevertheless, this exercise clearly shows that education is a substantively important variable in terms of the change in predicted probability of protest, also when compared to other (expectedly) important variables tapping key aspects of socioeconomic development at the local level.

10.3 Descriptives

Table 10.2: Descriptives for variables used in baseline models

Variable	Obs	Mean	Std. Dev.
Year	6816	2003.531	5.785
Education	6816	1.102	.872
Urban	6816	.211	.329
Unemployment	6363	.332	.252
Age	6816	28.623	1.978
Christian	5864	.557	.406
Protest	6816	.022	.173
L(Christian missions)	6816	-5.672	2.738
Travel time to city	6816	318.605	210.365
L(capital distance)	6816	5.652	.792
Polity score	6458	1.787	4.381

Table 10.3: Countries included and democracy scores (Polity)

Country	Polity score
Guinea-Bissau	5.50
Gambia	-2.83
Mali	6.66
Senegal	4.76
Benin	5.99
Mauritania	-4.21
Niger	0.56
Cote DâIvoire	-4.53
Guinea	-1.03
Burkina Faso (Upper Volta)	-2.10
Liberia	5.79
Sierra Leone	6.36
Ghana	3.56
Togo	-2.19
Cameroon	-4.76
Nigeria	2.33
Gabon	-0.50
Central African Republic	4.18
Chad	-2.63
Congo	-4.00
Congo, Democratic Republic of (Zaire)	4.69
Uganda	-1.82
Kenya	7.20
Tanzania/Tanganyika	-1.00
Burundi	5.65
Rwanda	-3.40
Somalia	0.00
Eritrea	-6.75
Angola	-2.05
Mozambique	5.00
Zambia	4.95
Zimbabwe (Rhodesia)	-1.96
Malawi	5.99
South Africa	9.00
Namibia	6.00
Lesotho	8.00
Botswana	8.00
Swaziland	-9.00
Madagascar	3.89
Madagascar (Malagasy)	3.89
Algeria	2.00
Sudan	-4.60

10.4 The data

The DHS surveys is a global survey operation intended to gauge important development metrics such as women's health, education, poverty and malnutrition in developing countries across the globe. The surveys are conducted at the household-level, are designed to be nationally representative and typically involve high numbers of respondents (5000-30000). The first surveys were conducted in 1986 and the first countries in Africa that were surveyed (in this year) were Liberia, Nigeria and Senegal. Since then, it has been expanded to cover 40+ countries in Africa. Since the SCAD data do not start until 1991, our empirical DHS time-period starts in 1990. Table 10.4 below shows the number of respondents and years that are included for each country in this period.

Table 10.4: Countries and time-periods included in the DHS

Country	Earliest year included	Latest year included
Algeria	2006	2006
Angola	2000	2011
Benin	1990	2010
Botswana	1999	2010
Burkina Faso (Upper Volta)	1992	2010
Burundi	1999	2011
Cameroon	1991	2011
Central African Republic	1991	2011
Chad	1991	2011
Congo	2007	2011
Congo, Democratic Republic of (Zaire)	1994	2011
Cote DIvoire	1993	2010
Eritrea	1997	2003
Ethiopia	1992	2009
Gabon	2004	2011
Gambia	1992	2011
Ghana	1993	2010
Guinea	1992	2011
Guinea-Bissau	1997	2005
Kenya	1997	2011
Lesotho	2004	2010
Liberia	1999	2009
Madagascar	1997	2009
Madagascar (Malagasy)	1997	2009
Malawi	1999	2010
Mali	1993	2011
Mauritania	1992	2011
Mozambique	1999	2011
Namibia	2000	2007
Niger	1990	2010
Nigeria	1990	2011
Rwanda	2000	2011
Senegal	1992	2011
Sierra Leone	1999	2009
Somalia	1992	2008
South Africa	1999	2011
Sudan	1992	2011
Swaziland	2006	2007
Tanzania/Tanganyika	1999	2011
Togo	1993	2008
Uganda	2000	2011
Zambia	1999	2011
Zimbabwe (Rhodesia)	1999	2011

The Social Conflict in Africa Database (now expanded to more countries, and labeled the Social Conflict Analysis Database), includes information on protests, riots, strikes and other “social disturbances” in Africa since 1991. It uses the search engine LexisNexis to identify events, and the coding procedure is described in the codebook (Salehyan and Hendrix, 2016). It covers all countries included in our dataset, but not all countries experience a protest event in a given year. To operationalize our “protest” variable, we combine the events: demonstrations and riots, both “organized” and “unorganized”. In the codebook (Salehyan and Hendrix, 2016, 4), these are described thus:

- **Organized Demonstration.** Distinct, continuous, and largely peaceful action directed toward members of a distinct âotherâ group or government authorities. In this event, clear leadership or organization(s) can be identified.
- **Spontaneous Demonstration.** Distinct, continuous, and largely peaceful action directed toward members of a distinct “other” group or government authorities. In this event, clear leadership or organization cannot be identified.
- **Organized Violent Riot.** Distinct, continuous and violent action directed toward members of a distinct âotherâ group or government authorities. The participants intend to cause physical injury and/or property damage. In this event, clear leadership or organization(s) can be identified.
- **Spontaneous Violent Riot.** Distinct, continuous and violent action directed toward members of a distinct âotherâ group or government authorities. The participants intend to cause physical injury and/or property damage. In this event, clear leadership or organization(s) cannot be identified.

Table 10.5: SCAD protest events in our sample (column 2) and all countries (in our sample) over 1990-2011 period (column 3)

Country	protests in grid-cells with DHS data	protests in total
Algeria	0	81
Angola	1	19
Benin	3	27
Botswana	0	16
Burkina Faso (Upper Volta)	7	19
Burundi	3	26
Cameroon	9	65
Central African Republic	0	21
Chad	2	22
Congo	0	10
Congo, Democratic Republic of (Zaire)	1	54
Cote DIvoire	3	35
Eritrea	0	8
Ethiopia	2	48
Gabon	0	33
Gambia	0	4
Ghana	12	29
Guinea	2	27
Guinea-Bissau	0	19
Kenya	11	106
Lesotho	0	5
Liberia	2	22
Madagascar	2	14
Madagascar (Malagasy)	2	14
Malawi	12	40
Mali	4	27
Mauritania	0	18
Mozambique	2	45
Namibia	3	16
Niger	4	49
Nigeria	27	327
Rwanda	3	15
Senegal	11	35
Sierra Leone	2	32
Somalia	0	26
South Africa	3	168
Sudan	0	66
Swaziland	3	22
Tanzania/Tanganyika	5	38
Togo	2	15
Uganda	2	29
Zambia	0	48
Zimbabwe (Rhodesia)	10	83

Column 2 captures all the protest events occurring in sampled (DHS-surveyed) grid-cells, and thus represents a fraction of the actual protest events in a given country. The reason why some protest events do not overlap with grid-cells are twofold: Either the DHS survey does not cover the given locality, or it does cover the given locality but not in the time-window when the protest occurs.

10.5 Additional controls

10.5.1 Alternative controls for local wealth

Table 10.6: Probit and OLS models of protest events

	Probit (A1)	Probit (A2)	Probit (A3)	Probit (A4)	OLS (A5)	OLS (A6)	OLS (A7)	OLS (A8)
Education	0.225*** (3.73)	0.233*** (3.63)	0.150* (2.37)	0.145* (2.42)	0.141*** (4.10)	0.142*** (4.03)	0.071* (2.04)	0.101** (3.04)
Infant mortality rate	0.000 (0.01)				0.000 (0.02)			
Child malnutrition		0.002 (0.29)				0.001 (0.20)		
Night light density			4.125*** (3.81)				6.966*** (4.67)	
Gross Cell Product				0.227*** (4.58)				0.128*** (4.82)
Protest _{t-1}	0.643*** (6.05)	0.637*** (5.98)	0.580*** (5.12)	0.625*** (5.89)	0.965*** (6.24)	0.957*** (6.17)	0.824*** (5.06)	0.943*** (6.15)
Age	0.175 (0.45)	0.172 (0.44)	0.178 (0.49)	0.241 (0.73)	-0.020 (-0.22)	-0.022 (-0.24)	-0.023 (-0.27)	-0.037 (-0.43)
Age ²	-0.004 (-0.57)	-0.004 (-0.56)	-0.004 (-0.61)	-0.005 (-0.88)	0.000 (0.09)	0.000 (0.11)	0.000 (0.12)	0.000 (0.27)
Urban	0.561*** (5.42)	0.576*** (5.54)	0.520*** (4.92)	0.619*** (5.78)	0.309*** (4.71)	0.317*** (4.83)	0.270*** (4.35)	0.326*** (4.99)
Unemployment	0.150 (0.80)	0.154 (0.81)	0.081 (0.42)	0.206 (1.06)	0.081 (1.21)	0.082 (1.22)	0.056 (0.78)	0.107 (1.59)
L(Distance to capital)	-0.175** (-3.14)	-0.176*** (-3.51)	-0.156** (-3.12)	-0.110* (-2.18)	-0.112** (-3.14)	-0.112*** (-3.37)	-0.090** (-2.87)	-0.067* (-2.03)
Travel time to city	-0.001*** (-3.48)	-0.001*** (-3.37)	-0.001*** (-3.64)	-0.001* (-2.11)	-0.000** (-2.82)	-0.000** (-2.78)	-0.000** (-3.08)	-0.000 (-0.47)
N	5716.000	5667.000	5474.000	5718.000	5902.000	5853.000	5654.000	5904.000
Log likelihood	-896.877	-889.978	-865.085	-884.738	-8343.321	-8280.330	-7985.175	-8324.273
R ²					0.134	0.134	0.151	0.140

Probit and OLS models with country- and year-fixed effects. Intercept excluded from table. Z-scores (Probit) and T-values (OLS) in parentheses.

Table 10.7: IV Probit and 2SLS models of protest events

	IVProbit (A9)	IVProbit (A10)	IVProbit (A11)	IVProbit (A12)	2SLS (A13)	2SLS (A14)	2SLS (A15)	2SLS (A16)
Education	0.918** (2.74)	1.004** (2.69)	0.807 ⁺ (1.88)	0.783 ⁺ (1.92)	0.747** (2.88)	0.809** (2.75)	0.582 ⁺ (1.92)	0.691* (2.34)
Infant mortality rate	0.000 (1.49)				0.000 ⁺ (1.87)			
Child malnutrition		0.025* (2.31)				0.018* (2.50)		
Night light density			0.893 (0.44)				3.888* (2.32)	
Gross Cell Product				0.082 (1.05)				0.037 (0.82)
Age	-0.003 (-0.01)	-0.031 (-0.11)	0.008 (0.03)	0.040 (0.15)	-0.034 (-0.29)	-0.049 (-0.39)	-0.032 (-0.29)	-0.032 (-0.28)
Age ²	-0.000 (-0.01)	0.000 (0.09)	-0.000 (-0.06)	-0.001 (-0.20)	0.001 (0.34)	0.001 (0.44)	0.001 (0.31)	0.001 (0.32)
Urban	-0.035 (-0.11)	-0.054 (-0.16)	0.008 (0.02)	0.088 (0.23)	-0.168 (-0.80)	-0.182 (-0.79)	-0.090 (-0.39)	-0.127 (-0.54)
L(Distance to capital)	-0.111 (-1.59)	-0.110 ⁺ (-1.70)	-0.086 (-1.15)	-0.072 (-1.13)	-0.064 (-1.51)	-0.064 (-1.59)	-0.047 (-1.08)	-0.030 (-0.75)
Travel time to city	-0.001 ⁺ (-1.76)	-0.001 (-1.44)	-0.001 ⁺ (-1.80)	-0.000 (-1.35)	-0.000 (-0.35)	0.000 (0.09)	-0.000 (-0.71)	-0.000 (-0.13)
christian	-0.799* (-2.01)	-0.776 ⁺ (-1.91)	-0.792 (-1.57)	-0.708 (-1.57)	-0.707* (-2.48)	-0.707* (-2.37)	-0.635 ⁺ (-1.86)	-0.684* (-2.22)
Protest _{t-1}	0.475*** (3.68)	0.465*** (3.53)	0.478*** (3.91)	0.498*** (3.81)	0.845*** (5.64)	0.839*** (5.58)	0.775*** (5.15)	0.847*** (5.69)
N	5154.000	5079.000	4931.000	5156.000	5317.000	5242.000	5092.000	5319.000
Log likelihood	-4143.176	-3973.255	-3958.299	-4104.860				
R ²					0.063	0.055	0.096	0.073

IVProbit and 2SLS models with country- and year-fixed effects. Intercept excluded from table. Z-scores (Probit) and T-values (OLS) in parentheses.

10.5.2 Literacy and media consumption

Table 10.8: Probit, OLS, IVProbit and 2SLS models of protest events

	Probit (A17)	Probit (A18)	OLS (A19)	OLS (A20)	IVProbit (A21)	IVProbit (A22)	2SLS (A23)	2SLS (A24)
Education	0.186** (2.68)	0.277*** (3.75)	0.108** (2.83)	0.166*** (4.08)	1.114* (2.37)	1.629** (3.08)	0.973* (2.47)	1.420** (2.67)
communication	0.170 (1.26)		0.140* (2.13)		-0.619 (-1.41)		-0.599+ (-1.81)	
Literacy		-0.225 (-1.01)		-0.108 (-1.03)		-2.747** (-2.84)		-2.368* (-2.50)
Protest _{t-1}	0.637*** (5.95)	0.640*** (6.02)	0.958*** (6.19)	0.964*** (6.23)	0.489*** (3.88)	0.410** (2.65)	0.870*** (5.86)	0.826*** (5.29)
Age	0.165 (0.43)	0.185 (0.47)	-0.021 (-0.22)	-0.021 (-0.23)	-0.014 (-0.05)	0.011 (0.04)	-0.036 (-0.33)	-0.029 (-0.22)
Age ²	-0.004 (-0.54)	-0.004 (-0.58)	0.000 (0.10)	0.000 (0.10)	0.000 (0.03)	-0.000 (-0.07)	0.001 (0.38)	0.001 (0.25)
Urban	0.491*** (4.47)	0.552*** (5.39)	0.258*** (4.20)	0.306*** (4.72)	0.117 (0.51)	-0.235 (-0.64)	-0.018 (-0.13)	-0.338 (-1.20)
Unemployment	0.144 (0.76)	0.154 (0.82)	0.083 (1.25)	0.081 (1.21)				
L(Distance to capital)	-0.167*** (-3.40)	-0.174*** (-3.58)	-0.105** (-3.23)	-0.112*** (-3.45)	-0.103+ (-1.75)	-0.040 (-0.52)	-0.065+ (-1.78)	-0.015 (-0.30)
Travel time to city	-0.001*** (-3.37)	-0.001*** (-3.56)	-0.000* (-2.53)	-0.000** (-2.95)	-0.001* (-2.24)	-0.001* (-2.04)	-0.000* (-1.98)	-0.000 (-1.51)
christian					-0.960+ (-1.91)	-0.941* (-2.42)	-0.906* (-2.28)	-0.876* (-2.46)
N	5719.000	5719.000	5905.000	5905.000	5157.000	5157.000	5320.000	5320.000
Log likelihood	-896.065	-896.365	-8342.935	-8345.359	-3477.816	-2861.090		
R ²			0.135	0.134			0.025	.

Probit, OLS, IVProbit and 2SLS models with country- and year-fixed effects. Intercept excluded from table. Z-scores (Probit) and T-values (OLS) in parentheses.

10.5.3 Additional proxies for urban settings

Table 10.9: Probit, OLS, IVProbit and 2SLS models of protest events

	Probit (A25)	Probit (A26)	OLS (A27)	OLS (A28)	IVProbit (A29)	IVProbit (A30)	2SLS (A31)	2SLS (A32)
Education	0.304*** (5.06)	0.266*** (4.55)	0.157*** (4.67)	0.180*** (5.75)	0.551+ (1.74)	0.944+ (1.75)	0.381* (1.99)	0.539** (3.07)
Urban area (Globcover)	0.309*** (3.41)		0.416*** (3.82)		0.258* (2.23)		0.344** (3.22)	
L(Population)		0.303*** (6.06)		0.113*** (5.72)		0.235*** (4.16)		0.082*** (4.05)
Protest _{t-1}	0.590*** (5.38)	0.601*** (5.49)	0.872*** (5.55)	0.947*** (6.12)	0.529*** (4.35)	0.511*** (3.56)	0.803*** (5.43)	0.848*** (5.71)
Age	0.784* (2.13)	0.066 (0.22)	0.020 (0.29)	-0.070 (-0.81)	0.327 (0.97)	-0.061 (-0.25)	0.024 (0.32)	-0.063 (-0.59)
Age ²	-0.014* (-2.25)	-0.002 (-0.40)	-0.001 (-0.50)	0.001 (0.65)	-0.006 (-1.05)	0.001 (0.13)	-0.000 (-0.38)	0.001 (0.61)
Unemployment	0.349+ (1.88)	0.382* (2.02)	0.132* (1.97)	0.180* (2.57)				
L(Distance to capital)	-0.127* (-2.40)	-0.109* (-2.28)	-0.076* (-2.31)	-0.079* (-2.50)	-0.105 (-1.54)	-0.078 (-1.10)	-0.053 (-1.40)	-0.039 (-1.04)
Travel time to city	-0.001*** (-3.84)	-0.001+ (-1.84)	-0.000* (-2.28)	-0.000 (-0.43)	-0.001** (-2.63)	-0.000 (-0.75)	-0.000 (-0.82)	0.000 (1.04)
Christian					-0.563 (-1.45)	-0.645+ (-1.85)	-0.378+ (-1.72)	-0.549** (-2.75)
N	5589.000	5719.000	5772.000	5905.000	5038.000	5157.000	5198.000	5320.000
Log likelihood	-855.384	-882.540	-8063.112	-8335.493	-4533.871	-3289.065		
R ²			0.153	0.137			0.136	0.103

Probit, OLS, IVProbit and 2SLS models with country- and year-fixed effects. Intercept excluded from table. Z-scores (Probit) and T-values (OLS) in parentheses.

10.5.4 Different types of land cover

Table 10.10: Probit and OLS models of protest events

	Probit (A33)	Probit (A34)	Probit (A35)	OLS (A36)	OLS (A37)	OLS (A38)
Education	0.211*** (3.45)	0.233*** (3.67)	0.207*** (3.42)	0.132*** (3.82)	0.147*** (4.15)	0.134*** (3.95)
Agricultural area (Globcover)	0.002			0.002		
Forested area (Globcover)		-0.002 (-0.79)			-0.001 (-1.12)	
Mountainous terrain			0.369* (1.98) (0.72)			0.117 (1.58)
Protest _{t-1}	0.625*** (5.86)	0.625*** (5.83)	0.654*** (6.06)	0.951*** (6.14)	0.953*** (6.14)	0.973*** (6.27)
Age	1.010** (2.61)	1.019** (2.63)	0.144 (0.36)	0.069 (1.03)	0.076 (1.11)	-0.020 (-0.22)
Age ²	-0.018** (-2.68)	-0.018** (-2.70)	-0.003 (-0.47)	-0.001 (-1.19)	-0.001 (-1.27)	0.000 (0.09)
Urban	0.641*** (6.07)	0.623*** (5.90)	0.609*** (5.78)	0.343*** (5.11)	0.336*** (4.99)	0.332*** (5.03)
Unemployment	0.228 (1.17)	0.213 (1.10)	0.194 (1.01)	0.098 (1.45)	0.091 (1.36)	0.096 (1.42)
L(Distance to capital)	-0.161** (-3.21)	-0.166*** (-3.36)	-0.173*** (-3.57)	-0.103** (-3.11)	-0.110*** (-3.35)	-0.112*** (-3.45)
Travel time to city	-0.001*** (-3.58)	-0.002*** (-3.87)	-0.002*** (-3.99)	-0.000** (-2.78)	-0.000** (-3.08)	-0.000*** (-3.35)
N	5589.000	5589.000	5682.000	5772.000	5772.000	5865.000
Log likelihood	-856.942	-856.936	-883.825	-8119.392	-8120.827	-8285.287
R ²				0.136	0.136	0.136

Probit and OLS models with country- and year-fixed effects. Intercept excluded from table. Z-scores (Probit) and T-values (OLS) in parentheses.

Table 10.11: IV Probit and 2SLS models of protest events

	IV-Probit (A39)	IV-Probit (A40)	IV-Probit (A41)	2SLS (A42)	2SLS (A43)	2SLS (A44)
Education	0.777* (1.98)	0.769* (2.04)	0.693* (1.97)	0.673* (2.40)	0.686* (2.57)	0.641** (2.62)
Agricultural area (Globcover)	-0.002 (-0.53)			-0.000 (-0.15)		
Forested area (Globcover)		-0.004 (-1.36)			-0.003* (-2.15)	
Mountainous terrain			0.515** (2.88)			0.229* (2.55)
Age	0.482 (1.35)	0.480 (1.35)	-0.046 (-0.16)	0.074 (0.87)	0.080 (0.95)	-0.035 (-0.30)
Age ²	-0.009 (-1.37)	-0.008 (-1.36)	0.000 (0.10)	-0.001 (-0.82)	-0.001 (-0.89)	0.001 (0.33)
Urban	0.166 (0.46)	0.174 (0.51)	0.222 (0.69)	-0.066 (-0.30)	-0.075 (-0.35)	-0.055 (-0.28)
L(Distance to capital)	-0.092 (-1.25)	-0.097 (-1.38)	-0.106 (-1.51)	-0.044 (-1.02)	-0.049 (-1.17)	-0.051 (-1.22)
Travel time to city	-0.001** (-2.70)	-0.001* (-2.49)	-0.001** (-2.78)	-0.000 (-1.64)	-0.000 (-0.80)	-0.000 (-1.50)
christian	-0.725 (-1.59)	-0.678 (-1.57)	-0.672 (-1.56)	-0.669* (-2.18)	-0.644* (-2.25)	-0.658* (-2.33)
Protest _{t-1}	0.496*** (3.68)	0.502*** (3.85)	0.536*** (4.29)	0.839*** (5.55)	0.838*** (5.57)	0.867*** (5.81)
Country-FE	YES	YES	YES	YES	YES	YES
Year-FE	YES	YES	YES	YES	YES	YES
N	5038.000	5038.000	5122.000	5198.000	5198.000	5282.000
Log likelihood	-4009.562	-3965.547	-4104.097			
R ²				0.081	0.081	0.086

IV-Probit and 2SLS models with country- and year-fixed effects. Intercept excluded from table.
Z-scores (Probit) and T-values (OLS) in parentheses.

10.5.5 Models with spatial lags

Table 10.12: Probit and OLS models of protest events

	Probit (A45)	Probit (A46)	Probit (A47)	Probit (A48)	OLS (A49)	OLS (A50)	OLS (A51)	OLS (A52)
Education	0.245*** (9.75)	0.468*** (8.22)	0.205*** (3.48)	0.225*** (3.74)	0.138*** (6.12)	0.245*** (7.62)	0.123*** (3.83)	0.076** (2.95)
Spatial lag	0.177** (2.82)	0.008 (0.09)	0.022 (0.25)	0.009 (0.10)	0.115 (1.41)	0.019 (0.24)	0.020 (0.25)	0.090 (1.04)
Protest $t - 1$	0.610*** (7.15)	0.730*** (7.18)	0.645*** (6.27)	0.641*** (6.01)	1.143*** (7.87)	0.999*** (6.66)	0.950*** (6.41)	1.051*** (7.08)
Age			0.032 (0.11)	0.175 (0.45)			-0.045 (-0.53)	-0.034 (-0.36)
Age ²			-0.001 (-0.22)	-0.004 (-0.57)			0.001 (0.41)	0.000 (0.23)
Urban			0.592*** (5.88)	0.562*** (5.39)			0.328*** (5.13)	0.383*** (6.03)
L(Distance to capital)			-0.168*** (-3.60)	-0.174*** (-3.57)			-0.109*** (-3.50)	-0.108*** (-3.76)
Travel time to city			-0.001** (-3.02)	-0.001*** (-3.54)			-0.000* (-2.42)	-0.000*** (-4.50)
Unemployment				0.150 (0.79)				0.098* (2.08)
β_0	-1.557*** (-41.02)	-2.405*** (-7.08)	-1.113 (-0.26)	-2.996 (-0.53)	-4.546*** (-208.79)	-4.370*** (-21.01)	-3.016* (-2.28)	-2.803 (-1.94)
Country-FE	NO	YES	YES	YES	NO	YES	YES	YES
Year-FE	NO	YES	YES	YES	NO	YES	YES	YES
N	6699.000	6097.000	6097.000	5719.000	6272.000	6272.000	6272.000	5905.000
Log likelihood	-2243.494	-1031.155	-980.568	-896.909	-9116.770	-8964.894	-8913.687	-8421.449
R ²					0.068	0.112	0.126	0.112

Probit and OLS models with country- and year-fixed effects. Intercept excluded from table. Z-scores (Probit) and T-values (OLS) in parentheses.

Table 10.13: IV Probit and 2SLS models of protest events

	IV-probit (A53)	IV-probit (A54)	IV-probit (A55)	IV-probit (A56)	2SLS (A57)	2SLS (A58)	2SLS (A59)	2SLS (A60)
Education	0.419** (2.68)	1.104*** (10.08)	0.243 (1.69)	0.899*** (3.64)	0.692*** (5.04)	0.704*** (5.94)	0.653*** (4.26)	0.739*** (3.62)
Christian	-0.130 (-0.72)	-1.102*** (-4.75)	0.113 (0.43)	-0.804 (-1.72)	-0.903*** (-3.96)	-0.663** (-2.83)	-0.862** (-3.20)	-0.714* (-1.98)
Spatial lag	0.109* (1.99)	-0.036 (-0.52)	0.112* (2.16)	-0.026 (-0.40)	0.048 (0.76)	-0.006 (-0.09)	0.033 (0.53)	-0.014 (-0.23)
Protestt - 1	0.478*** (4.69)	0.449*** (4.41)	0.479*** (4.93)	0.480*** (4.16)	0.942*** (5.80)	0.843*** (5.26)	0.940*** (5.73)	0.848*** (5.25)
Age			-0.343* (-2.00)	0.005 (0.04)			-0.013 (-0.18)	-0.025 (-0.60)
Age ²			0.005* (2.18)	-0.000 (-0.08)			0.000 (0.17)	0.001 (0.74)
Urban			0.101 (0.23)	-0.020 (-0.05)			-0.092 (-0.54)	-0.162 (-0.65)
L(Distance to capital)			-0.083*** (-4.72)	-0.076 (-1.92)			-0.096** (-2.75)	-0.038 (-1.30)
Travel time to city			-0.000 (-1.11)	-0.001*** (-5.40)			-0.000 (-0.57)	-0.000 (-1.23)
β_0	-1.594*** (-7.23)	-2.335*** (-3.74)	4.560 (1.41)	-1.768 (-0.93)	-4.639*** (-140.52)	-5.188*** (-44.78)	-3.849*** (-3.35)	-4.622*** (-7.19)
Country-FE	NO	YES	YES	YES	NO	YES	YES	YES
Year-FE	NO	YES	YES	YES	NO	YES	YES	YES
N	5747.000	5157.000	5747.000	5157.000	5320.000	5320.000	5320.000	5320.000
Log likelihood	-7954.007	-4966.732	-7345.774	-4160.171				
R ²					.	0.068	0.004	0.063

IV Probit and 2SLS models with country- and year-fixed effects. Intercept excluded from table.
Z-scores (Probit) and T-values (OLS) in parentheses.

10.6 Dropping cells with few respondents respondents (< 30 respondents)

Table 10.14: Probit and OLS models of protest events

	Probit (A67)	Probit (A68)	Probit (A69)	Probit (A70)	OLS (A71)	OLS (A72)	OLS (A73)	OLS (A74)
Education	0.244*** (5.45)	0.503*** (5.27)	0.248* (2.48)	0.304* (2.58)	0.194*** (4.24)	0.292*** (4.58)	0.166* (2.44)	0.149* (2.53)
Protestt - 1	0.717*** (3.75)	1.004*** (4.16)	0.972*** (3.69)	0.961*** (3.32)	1.322*** (4.05)	1.220*** (3.76)	1.157*** (3.70)	1.325*** (4.16)
Age			-0.994 (-1.13)	0.421 (0.39)			-0.236 (-0.65)	0.049 (0.16)
Age ²			0.017 (1.11)	-0.008 (-0.43)			0.004 (0.64)	-0.001 (-0.21)
Urban			0.984*** (4.74)	0.915*** (3.97)			0.411** (3.06)	0.402** (2.93)
L(Distance to capital)			-0.165 (-1.89)	-0.196* (-2.06)			-0.109 (-1.96)	-0.105* (-2.10)
Travel time to city			0.000 (0.85)	-0.000 (-0.16)			0.000 (0.77)	-0.000 (-0.72)
Unemployment			0.051 (0.14)				0.042 (0.48)	
β_0	-1.438*** (-24.58)	-0.435 (-0.63)	14.670 (1.16)	-5.579 (-0.36)	-4.562*** (-135.29)	-4.511*** (-9.06)	-0.550 (-0.10)	-4.183 (-0.92)
Country-FE	YES	YES	YES	YES	YES	YES		
Year-FE	YES	YES	YES	YES	YES	YES		
N	1766.000	1307.000	1307.000	1152.000	1637.000	1637.000	1637.000	1475.000
Log likelihood	-617.433	-246.166	-233.609	-189.948	-2310.239	-2270.922	-2256.821	-1996.788
R ²					0.078	0.122	0.137	0.139

Probit and OLS models with country- and year-fixed effects. Intercept excluded from table. Z-scores (Probit) and T-values (OLS) in parentheses.

10.7 Education in Anglophone versus Francophone countries

As pointed out in the main text, education systems in Africa differ. One crucial difference is their varying emphasis on promoting the national language in national education. As carefully detailed in Albaugh (2014), there has been a trend over the past decades towards a focus on multi-lingual education in Francophone countries – reversing the previous French focus on learning French in education – contrasted with a focus on unilingual education in Anglophone countries. As Albaugh points out, there has been a push towards “mother tongue” education policies in Africa, mainly from foreign NGOs, supported by the French government and increasingly widespread in Francophone Africa. This could be relevant if, for example, learning *one* language makes a difference to collective action. While we have no strong theoretical reason for believing this, plausible narratives exist: For example, it could be that facilitating collective action on a widespread scale is easier when everyone speaks the same language (note, however that a school practicing “mother tongue” policies does not mean that the pupils in that school or citizens in the local vicinity do not speak the same (one) language).

To probe the possibility that effects of education vary across this divide in education policies, we split the sample into two groups; Francophone and non-Francophone countries (most of these are former British colonies, while a handful have other colonial histories). Table A11 investigates whether we find different effects of education across these two groups. Further supporting the indication of a strong general effect of education on protest, we find similar results in both Francophone and non-Francophone countries. This suggests that the impact of education is not conditional on this crucial difference in education policy.

Table 10.15: Probit models of protest events

	Probit (A75)	Probit (A76)	Probit (A77)	Probit (A78)	Probit (A79)	Probit (A80)	Probit (A81)	Probit (A82)
Francophone Africa	Yes	Yes	Yes	Yes	No	No	No	No
Education	0.544*** (11.39)	0.672*** (5.30)	0.199 (1.52)	0.218 (1.52)	0.202*** (6.30)	0.429*** (6.70)	0.228*** (3.44)	0.258*** (3.86)
protspatlag	0.349** (2.71)	0.134 (0.74)	0.100 (0.57)	0.089 (0.52)	0.118 (1.78)	-0.036 (-0.38)	-0.015 (-0.17)	-0.038 (-0.41)
Protest _{t-1}	0.833*** (3.49)	0.940*** (3.46)	0.813** (3.19)	0.809** (3.02)	0.583*** (6.56)	0.676*** (6.27)	0.604*** (5.34)	0.601*** (5.15)
Age			0.310 (0.42)	0.592 (0.78)			0.003 (0.01)	0.150 (0.41)
Age ²			-0.005 (-0.41)	-0.010 (-0.76)			-0.001 (-0.15)	-0.004 (-0.55)
Urban			0.963*** (5.60)	1.070*** (5.50)			0.425*** (3.34)	0.388** (3.00)
L(Distance to capital)			-0.037 (-0.51)	-0.050 (-0.65)			-0.224*** (-3.70)	-0.222*** (-3.56)
Travel time to city			-0.002** (-3.04)	-0.003*** (-3.44)			-0.001* (-2.19)	-0.001* (-2.49)
Unemployment				-0.394 (-1.00)				0.233 (0.98)
Country-FE	YES	YES	YES	YES	YES	YES	YES	YES
Year-FE	YES	YES	YES	YES	YES	YES	YES	YES
N	2374	2099	2099	2003	4325	3993	3993	3711
Log likelihood	-767.783	-328.732	-305.759	-267.930	-1435.111	-685.082	-655.484	-606.001

Probit models with country- and year-fixed effects. Intercept excluded from table. Z-scores in parentheses.

11 Appendix to Chapter 6

11.1 Introduction

This appendix offers additional details about the data collection and coding procedure, as well as several robustness tests. In brief, the latter indicate that the main findings from the paper are generally robust to alternative model specifications. In section 11.2, I present the complete codebook used to collect indicators on the education profile of protest movements, and in section 11.3 I present summary statistics for these variables. Section 11.4 discusses robustness tests to assess whether the main results from the paper are driven by reliability issues related to the coding procedure. Sections 11.5-11.7 contain various robustness tests related to models with non-violence as dependent variable. Section 5 presents models with alternative sets of control variables, including models where all of the controls in the main analysis are left out at a time and the inclusion of additional controls. In section 11.6, I presents numerous models based on different specifications of the variable measuring involvement of students and graduates. Alternative specifications of the general sensitivity analysis discussed in the paper are presented in section 11.7. Section 11.8 discusses robustness tests of models with successful outcome as dependent variable, and section 11.9 contains a discussion and presentation of robustness tests of the mediation analysis. Finally, section 11.10 contains test where the education profile of the campaign leader is also accounted for and in section 11.11 I present results from estimations on a sample where missing values have been imputed.

11.2 Codebook for indicators of education

Below is the codebook for the variables recording the educational background of participants in protest campaigns. For each campaign, I always consulted several different sources, including both databases and encyclopedias such as “The international encyclopedia of revolution and protest” (Ness, 2015) and the “Global Non-Violent Action Database” (Swarthmore, 2015), partly on various reports by international or non-governmental organizations such as UNHCR or Freedom House, where the relevant events were described, and partly on books and journal articles that cover the campaigns, such as e.g. Ackerman and DuVall (2000) and Nepstad (2011). I coded each campaign based on the participants that were reported to have been a part of the campaign from the start year listed in NAVCO 1.0, to the end year.

As noted in the paper, the sources are usually relatively clear as to whether educated individuals or students participated at all, as there is usually some kind of description of the protest participants. Likewise, there is usually some description of where and amongst who the campaign originated. When it comes to the variable recording whether students or graduates dominated the campaign, I relied on a coding rule requiring that the movement should at least either have students or educated as the majority of its members, or, if not, they should have had a critical impact on the strategies and outcomes of the campaign. The details are described below:

1. Participated in the campaign

These variables record whether students or educated groups or individuals are mentioned in the sources as having participated in the campaign at some point in time. As noted, I always consult several sources to determine the profile of each campaign. To get assigned the value “1”, at least 2 sources should mention that students and educated groups participated in the campaign. To get assigned the value “0”, at 4-5 sources should *not* mention these groups as having participated in the campaign, to ensure that the no-mentioning is not simply a missing data or poor data quality problem. Campaigns were also coded as 0 if at least 2 sources described the campaign as consisting exclusively of other groups, such as peasants or workers.

1A. *At least students*: Variable measuring whether the campaign consisted at least partly of students, at any point in time.

- 0: Students are *not* reported to have participated in the campaign
- 1: Students, at a minimum, are reported to have participated in the campaign at some point in time

1B. At least graduates: Variable measuring whether the movement consisted at least partly of university graduates

- 0: Educated individuals/groups are *not* reported to have participated in the campaign
- 1: Educated individuals/groups, at a minimum, are reported to have participated in the campaign at some point in time

2. Dominated the campaigns

These variables record whether the campaign was dominated by either students or educated. The dominance criteria should be considered fulfilled if the campaign is in line with at *least one* of the following criterias: 1) Students or educated made up a majority of campaign members or 2) Students or educated had a critical impact on the outcome and/or the strategies of the campaign. To be considered to be in majority, the sources should mention this explicitly, for instance through formulations such as “most of the protesters were students”. Conversely, if sources mention explicitly that other social groups were in majority, the majority criteria is considered as not being fulfilled. The “critical impact” criteria is more reliant on subjective judgment. It should be considered fulfilled if the sources describe students or educated as having been the most important force behind the strategies and/or the outcome of the movement.

2A. Students dominate: Variable measuring whether the campaign was dominated by students.

- 0: campaign was not dominated by students
- 1: The campaign was dominated by students

2B. Graduates dominate: Variable measuring whether the campaign was dominated by students.

- 0: campaign was not dominated by educated
- 1: The campaign was dominated by educated

3. Initiated the campaign

These variables record whether the campaign originated among or was initiated by students or educated individuals/groups. To code this variable, the year that is listed as “start year” in NAVCO is considered as the campaign’s starting point. Hence, events taking place before this starting point are excluded from the consideration. If the movement originated among several groups, group can be considered

as having initiated the campaign although other groups also helped to initiate it. Hence, if the movement was described as originating among students/educated as well as other groups, the campaign will still be assigned a “1” on the initiation variable.

3A. Originated among students: Considers whether the campaign was established primarily by or among students. For instance, did it originate at a university campus? Were the initial protests led by students?

- 0: The campaign did not originate among students
- 1: The campaign originated among students

3B. Originated among graduates: Considers whether the campaign was established by or among graduates. For instance, were initial protests led by a professional association? Did criticism from graduates and/or intellectuals spur the initial acts of resistance?

- 0: The campaign did not originate among educated
- 1: The campaign originated among educated

4. Leader educated

Considers whether the leader of the campaign was a student or a university graduate. If there are leadership alterations during the campaign, the leader that was in charge for the longest period of time should be considered. If there are two or several leaders, the coding should be based on the person with the lowest education level. If no prominent leader is identified in the sources, the campaign should be coded as having no clear leadership. In order to count as a graduate, the leader must have *completed* a degree from a tertiary education institution.

- 0 : no clear leadership
- 1 : clear leadership, but not educated
- 2 : clear leadership, student (not from religious or military institutions)
- 3 : clear leadership, student, from religious institution
- 4 : clear leadership, student, from military academy
- 5 : clear leadership, graduate (not from religious or military institutions)
- 6 : clear leadership, graduate, from religious institution
- 7 : clear leadership, graduate, from military academy

11.3 Summary statistics for main independent and dependent variables

Table 11.1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
At least students	195	.5794872	.494912	0	1
At least graduates	195	.5897436	.4931463	0	1
Students dominate	196	.122449	.3286431	0	1
Graduates dominate	197	.0913706	.2888695	0	1
Originate among students	197	.2284264	.4208883	0	1
Originate among graduates	195	.1846154	.3889839	0	1
Nonviolent	323	.3281734	.4702769	0	1
Aim: Regime change	323	.5944272	.4917644	0	1
Success	323	.3467492	.4766735	0	1

11.4 Sensitivity tests leaving out observations coded as uncertain

This section contains robustness tests in which I exclude from the analysis all the observations (68 in total) that were coded as fairly uncertain. In brief, the results are robust to this. Table 11.2 contains results from models with non-violence as dependent variable, based on table 6.1 in the paper. Although the sample size drops quite a lot when excluding the 68 “uncertain” observations, the coefficient estimate for the education index is always positive and significant at the 0.00 level.¹ Table 11.3 contains results from models with success as outcome variable, based on table 6.2 from the paper. It yields similar results to the findings in the paper: There is some, although weak, evidence that the education index is positively associated with sucess (in models A1-A4), but this effect is almost completely soaked up when controlling for non-violence (in models A5-A8).

Table 11.2: Non-violent outcome on campaign involvement by educated protesters (logit models). Leaving out cases for which coding was considered uncertain.

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Education index	1.049*** (5.64)	1.208*** (5.09)	1.262*** (3.63)	1.487*** (4.06)	1.756*** (4.46)	2.833*** (3.68)	3.947** (2.59)
Membership size (log)					0.531* (2.05)	0.558* (2.37)	0.628 (1.17)
Target capacity (log)					-0.449 (-1.61)	-1.052+ (-1.72)	0.522 (0.52)
Target support					1.461+ (1.67)	1.321 (1.16)	3.397* (2.37)
Campaign support					-3.225** (-2.79)	-5.675* (-2.10)	-12.96** (-2.70)
Regime violence					-0.280 (-0.18)	0.878 (0.31)	4.671 (1.44)
Polity index					-0.0688 (-0.94)	-0.144 (-0.77)	-0.149 (-1.28)
GDP p.c.(log)					1.128* (2.50)	3.197*** (4.10)	3.419+ (1.68)
University enrollment					-0.0109 (-1.23)	-0.0308* (-2.33)	-0.0939* (-2.38)
Industrial activity					-0.000819 (-1.12)	-0.00168 (-0.73)	-0.00200+ (-1.68)
Constant	-2.505*** (-7.19)	-4.997*** (-4.78)	-1.612* (-2.06)	-4.653*** (-3.44)	-18.00*** (-3.33)	-44.90*** (-3.85)	-30.29 (-1.38)
R2 (pseudo)	0.339	0.409	0.483	0.591	0.623	0.745	0.734
ll	-55.72	-49.86	-39.67	-31.36	-18.73	-12.66	-10.30
N	124	124	111	111	72	72	56
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.3: Successful outcome on campaign involvement by educated protesters (logit models). Leaving out cases for which coding was considered uncertain.

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit	A8 logit
Education index	0.246* (2.24)	0.441* (2.49)	0.287+ (1.83)	0.333 (1.06)	-0.0318 (-0.21)	0.149 (0.70)	0.00572 (0.03)	0.0484 (0.13)
Non-violent					1.567** (3.15)	1.749* (2.15)	2.078* (2.31)	2.813* (2.10)
Polity index		0.0521 (1.20)	0.0546 (1.27)	0.0288 (0.51)		0.0730 (1.62)	0.0945+ (1.91)	0.0967 (1.36)
GDP p.c. (log)		0.458+ (1.75)	0.336 (1.43)	0.791+ (1.83)		0.221 (0.79)	0.165 (0.70)	0.546 (0.97)
Target capacity				-40.12+ (-1.83)				-35.85 (-1.47)
Duration					-0.311 (-1.52)			-0.146 (-0.63)
Target support					0.388 (0.45)			0.126 (0.16)
Campaign support					-0.756 (-0.67)			0.282 (0.23)
Military defection					0.791 (0.91)			0.992 (1.13)
Sanctions against regime					0.537 (0.40)			-0.895 (-0.60)
Regime violence					-3.744*** (-3.41)			-3.477** (-2.71)
Constant	-0.854*** (-3.33)	-5.873** (-2.81)	-1.722 (-0.95)	-2.649 (-0.76)	-0.994*** (-3.74)	-4.169+ (-1.93)	-1.531 (-0.71)	-2.302 (-0.59)
R2 (pseudo)	0.0319	0.248	0.165	0.417	0.0931	0.294	0.221	0.460
ll	-81.31	-49.04	-53.51	-32.27	-76.17	-46.02	-49.92	-29.89
N	124	96	94	83	124	96	94	83
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

¹Due to fewer degrees of freedom, the models did not converge when adding region and decade dummies to the same models as the full set of controls.

11.5 Alternating control variables

As the main models and finding in the paper deal with non-violent outcome as dependent variable, the next sections present robustness tests related to this finding, and models with success as dependent variable are returned to in section 11.8. As noted, it is necessary with some agnosticism when choosing models specifications and especially control variables, due to the fact that several of the potential controls could potentially both induce omitted variable bias if excluded and post-treatment bias if included. This section tries to account for this. First, table 11.4 replicates model A5 from table 6.1 in the paper, but leaves out one control variable at a time. This does not produce any major changes to the findings. Second, table 11.5 is also a replication of table 6.1 in the paper, but with the campaign size control completely left out, and the findings are robust to this. Third, table 11.6 presents results from replications of model 5 from table 6.1 in the paper, but adding several additional control variables. Including additional controls reduces the sample size due to empty cells and reduces the degrees of freedom, but the coefficient estimate for the education index is still always positive and significant at least at the 0.05 level. Hence, the main finding from the paper should generally considered robust to these adjustments when it comes to control variables.

Table 11.4: Non-violent outcome on campaign involvement by educated protesters (logit models). Leaving out one control at a time.

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit	A8 logit
Education index	1.053*** (4.75)	1.091** (3.07)	1.003*** (3.76)	1.340*** (4.39)	1.046*** (3.83)	1.265*** (4.52)	1.246*** (4.25)	1.263*** (4.49)
Membership size (log)	0.543** (2.61)	0.626** (3.19)	0.672** (3.19)	0.701** (3.26)	0.768*** (3.87)	0.668** (3.03)	0.690** (3.22)	0.668** (3.02)
Target support	1.530* (2.39)		0.758 (1.13)	1.966** (2.66)	1.532* (2.21)	1.910** (3.01)	1.891* (2.57)	1.907** (3.02)
Campaign support	-3.104*** (-3.38)	-2.581** (-2.73)		-3.690*** (-4.11)	-2.730** (-3.27)	-3.728*** (-3.64)	-3.421*** (-3.51)	-3.722*** (-3.63)
Regime violence	-1.177 (-1.03)	-1.042 (-0.76)	-1.502 (-1.15)		-1.301 (-0.83)	-1.086 (-0.74)	-1.041 (-0.72)	-1.089 (-0.74)
Polity index	-0.0892 (-1.63)	-0.0960* (-2.13)	-0.0986+ (-1.86)	-0.108* (-2.11)		-0.117* (-2.00)	-0.120* (-2.05)	-0.118* (-1.98)
GDP p.c.(log)	0.432 (1.08)	0.655 (1.27)	0.714+ (1.91)	0.691 (1.37)	0.637 (1.51)		0.597 (1.23)	
University enrollment	-0.00137 (-0.36)	-0.00130 (-0.32)	0.00136 (0.40)	-0.00367 (-0.85)	-0.000477 (-0.12)	-0.000116 (-0.03)		
Industrial activity	-0.000512 (-0.64)	-0.000413 (-0.40)	-0.000552 (-0.69)	-0.000741 (-0.85)	-0.000514 (-0.61)	0.000121 (0.16)	-0.000670 (-0.78)	0.000116 (0.15)
Target capacity (log)		-0.452+ (-1.67)	-0.455+ (-1.75)	-0.550* (-2.12)	-0.550+ (-1.92)	-0.455* (-2.50)	-0.540* (-2.11)	-0.455* (-2.50)
R2 (pseudo)	0.555	0.547	0.494	0.583	0.536	0.569	0.586	0.569
ll	-36.62	-34.64	-38.64	-31.88	-40.92	-32.95	-31.66	-32.95
N	119	111	111	111	130	111	111	111

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Replicating model A5 from Table 6.1 in the paper.

Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.5: Non-violent outcome on campaign involvement by educated protesters (logit models). Without controlling for membership size.

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Education index	0.947*** (5.85)	1.029*** (5.79)	1.061*** (4.94)	1.146*** (4.68)	1.201*** (3.61)	1.102*** (4.20)	1.170** (3.04)
Target capacity (log)					-0.287 (-1.61)	-0.339 (-1.57)	-0.248 (-1.32)
Target support					1.181+ (1.81)	1.748** (2.64)	2.221** (2.98)
Campaign support					-3.365*** (-3.98)	-3.505*** (-4.84)	-5.519*** (-4.20)
Regime violence					-1.595* (-2.23)	-1.456+ (-1.86)	-1.004 (-0.99)
Polity index					-0.109* (-2.27)	-0.130* (-2.48)	-0.117* (-2.29)
GDP p.c.(log)					0.615* (2.04)	1.070** (3.01)	0.750+ (1.73)
University enrollment					0.00248 (0.58)	0.00127 (0.28)	-0.0104 (-1.12)
Industrial activity					-0.000170 (-0.25)	0.000109 (0.13)	0.00128 (0.78)
Constant	-2.128*** (-5.99)	-4.486*** (-5.72)	-1.040+ (-1.67)	-3.768*** (-3.78)	-6.945* (-2.39)	-14.22*** (-5.21)	-7.030* (-2.01)
R2 (pseudo)	0.281	0.340	0.397	0.486	0.510	0.579	0.597
ll	-93.76	-86.09	-71.62	-61.10	-48.95	-42.04	-34.06
N	192	192	172	172	145	145	122
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.6: Non-violent outcome on campaign involvement by educated protesters (logit models). Additional controls.

	A1 logit	A2 logit	A3 logit
Education index	6.236* (2.06)	1.998* (2.47)	1.085*** (3.45)
Number of protesters (log)	4.024** (2.65)	1.215* (2.02)	0.685** (3.27)
Target capacity (log)	-5.056** (-3.04)	-3.439 (-0.98)	-0.482+ (-1.73)
Target support	14.59** (2.58)	2.810 (1.29)	1.471+ (1.86)
Campaign support	-46.96** (-3.28)	-6.708*** (-3.52)	-3.649*** (-3.36)
Regime violence	0.974 (0.33)	-2.838+ (-1.74)	-1.122 (-0.74)
Polity	-0.832* (-2.25)	-0.220 (-1.22)	-0.0573 (-0.95)
GDP p.c.(log)	7.939** (3.06)	2.553 (1.53)	0.733 (1.54)
University enrollment	-0.121* (-2.06)	-0.0360** (-3.11)	-0.00809+ (-1.70)
Duration	2.354 (1.19)		
Military defection	-8.523* (-2.51)		
Sanctions against target	14.08*** (3.42)		
Interstate conflict	-1.382 (-0.22)		
Simultaneous violent campaign	-7.160 (-1.54)		
Civil wars	15.07** (2.74)		
Industrial activity		0.000447 (0.34)	
Urbanization		0.00500 (1.47)	
GDP growth p.c. (log)		0.211+ (1.83)	
Population (log)		1.718 (0.59)	
Cold war			2.237* (2.38)
Post cold war			2.456* (2.00)
R2 (pseudo)	0.821	0.733	0.626
ll	-10.30	-12.94	-28.77
N	83	70	112
Region dummies?	no	no	no
Decade year dummies?	no	no	no

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Replicating model A5 in table 6.2 in the paper. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

11.6 Alternating the operationalization of the education index

This section presents numerous models based on different specifications of the variables measuring involvement of students and graduates. It starts by presenting results from models where the education index is replaced by dichotomous variables capturing various degrees of involvement: Whether the campaign originated among (table 11.7), consisted of (table 11.8) and was dominated by educated participants (table 11.9). Next, models are presented with separate indexes for graduate involvement and student involvement (and these are both added separately and to the same model) (see tables 11.10, 11.11 and 11.12). I continue by replacing the initial education index by a dichotomous education simply recording whether at least either students or graduates participated in the movement (table 11.13). Finally, models are presented with a multiplicative education index (instead of the original additive index) (table 11.14). In brief, the main findings from the paper are robust to alternative ways of constructing the education variable, but there are some nuances. For instance, there is more evidence that campaigns that either originated among or consist of educated participants (tables 11.7 and 11.8) are associated with non-violence, than there is evidence that campaigns dominated by educated participants have a higher likelihood of non-violence (table 11.9). Moreover, when adding the student index and the graduate index to the same model, it seems as if the effect of education on non-violence is mainly driven by university graduates rather than students.

Table 11.7: Non-violent outcome and whether campaign **originated** among educated protesters (logit models)

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Originated among educated	1.385*** (3.65)	1.445*** (3.57)	1.658*** (3.77)	1.669*** (3.32)	1.835** (3.17)	1.661** (3.08)	1.806* (2.40)
Membership size (log)					0.702*** (3.57)	0.868*** (3.51)	0.675* (2.39)
Target capacity (log)					-0.280 (-1.46)	-0.516+ (-1.93)	-0.218 (-0.97)
Target support					1.220* (2.13)	1.571* (2.56)	2.998* (2.36)
Campaign support					-2.825** (-3.13)	-3.836*** (-3.71)	-6.032*** (-3.36)
Regime violence					-1.264 (-0.94)	-0.977 (-0.63)	-0.615 (-0.31)
Polity index					-0.101* (-2.25)	-0.144* (-2.37)	-0.155* (-2.37)
GDP p.c.(log)					0.509 (1.34)	1.500** (2.72)	1.462 (1.48)
University enrollment					0.00277 (0.58)	0.000572 (0.12)	-0.0187* (-2.24)
Industrial activity					-0.000560 (-0.62)	0.000194 (0.16)	0.000366 (0.16)
Constant	-0.825*** (-3.53)	-2.597*** (-4.95)	0.651 (1.20)	-1.174+ (-1.72)	-11.76** (-3.17)	-28.75*** (-4.00)	-15.64+ (-1.85)
R2 (pseudo)	0.0748	0.145	0.195	0.309	0.462	0.601	0.608
ll	-122.0	-112.8	-96.26	-82.57	-41.46	-30.68	-23.31
N	194	194	173	173	112	112	86
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region and decade dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.8: Non-violent outcome and whether campaigns had **at least some** educated protesters (logit models)

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Educated protesters participated	2.736*** (5.98)	2.798*** (6.06)	3.135*** (6.50)	3.165*** (6.21)	2.768** (3.23)	3.432* (2.57)	7.605** (2.68)
Membership size (log)					0.603*** (3.43)	0.861*** (3.46)	0.944+ (1.80)
Target capacity (log)					-0.345+ (-1.75)	-0.892* (-2.40)	0.0994 (0.41)
Target support					1.297 (1.41)	1.233 (1.17)	9.708** (2.74)
Campaign support					-2.999** (-2.73)	-4.149** (-2.68)	-14.31** (-3.06)
Regime violence					-1.410 (-1.19)	-1.394 (-1.09)	1.063 (0.40)
Polity index					-0.143** (-2.73)	-0.186** (-2.72)	-0.248+ (-1.96)
GDP p.c.(log)					0.371 (1.06)	1.135+ (1.90)	0.749 (0.93)
University enrollment					0.00456 (1.06)	0.00749 (1.14)	-0.0346* (-2.29)
Industrial activity					-0.000707 (-0.85)	-0.00236 (-1.57)	0.000494 (0.29)
Constant	-2.183*** (-5.12)	-3.890*** (-5.91)	-1.476* (-2.15)	-3.282*** (-3.44)	-11.61** (-3.09)	-26.34*** (-3.79)	-18.81* (-2.16)
R2 (pseudo)	0.231	0.287	0.373	0.460	0.514	0.654	0.794
ll	-101.4	-93.98	-75.02	-64.61	-37.38	-26.65	-12.27
N	194	194	173	173	112	112	86
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region and decade dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.9: Non-violent outcome and whether campaign was **dominated** by educated protesters (logit models)

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Dominated by educated	2.886*** (5.74)	2.921*** (5.81)	3.251*** (4.52)	3.289*** (4.72)	2.996*** (3.50)	2.436* (2.48)	2.545* (2.53)
Membership size (log)					0.488** (3.15)	0.563** (3.28)	0.580** (2.67)
Target capacity (log)					-0.156 (-1.11)	-0.298* (-2.07)	-0.106 (-0.56)
Target support					0.811+ (1.68)	0.912+ (1.79)	1.138+ (1.87)
Campaign support					-2.107** (-2.99)	-2.358** (-3.06)	-2.360** (-2.94)
Regime violence					-1.909* (-2.28)	-1.851+ (-1.88)	-1.860 (-1.57)
Polity index					-0.0857* (-2.20)	-0.134** (-2.61)	-0.0741 (-1.60)
GDP p.c.(log)					0.637* (2.51)	1.159*** (3.48)	0.937** (2.98)
University enrollment					0.00208 (0.63)	0.00130 (0.37)	-0.00827* (-2.13)
Industrial activity					-0.000664 (-0.90)	-0.00111+ (-1.67)	0.000613 (0.87)
Constant	-0.976*** (-6.36)	-2.400*** (-4.85)	0.818 (1.53)	-0.700 (-0.92)	-9.008*** (-4.10)	-17.43*** (-5.19)	-11.10** (-3.10)
R2 (pseudo)	0.108	0.160	0.272	0.346	0.425	0.528	0.528
ll	-179.2	-168.8	-142.3	-127.8	-60.39	-49.62	-41.29
N	316	316	303	303	162	162	130
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region and decade dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.10: Non-violent outcome on campaign involvement by graduates (logit models)

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Graduates index	1.485*** (5.63)	1.621*** (5.96)	2.141*** (4.66)	2.367*** (4.48)	2.011*** (3.73)	2.538*** (3.33)	3.756** (2.70)
Membership size (log)					0.692*** (3.34)	0.931*** (3.31)	0.724* (2.00)
Target capacity (log)					-0.468+ (-1.91)	-1.046* (-2.56)	-0.0637 (-0.28)
Target support					1.456+ (1.66)	2.075* (2.16)	6.054* (2.09)
Campaign support					-2.998* (-2.46)	-3.998*** (-3.84)	-10.45** (-2.80)
Regime violence					-1.638 (-1.36)	-1.720 (-1.38)	0.0373 (0.02)
Polity index					-0.105* (-2.06)	-0.155* (-2.29)	-0.205+ (-1.91)
GDP p.c.(log)					0.547 (1.24)	1.475* (2.33)	1.169 (1.19)
University enrollment					-0.00153 (-0.31)	-0.00518 (-0.70)	-0.0374* (-2.39)
Industrial activity					-0.000602 (-0.62)	0.000325 (0.16)	-0.00144 (-0.67)
Constant	-1.646*** (-5.96)	-4.103*** (-7.90)	-1.076 (-1.57)	-4.466*** (-3.72)	-13.54** (-2.66)	-33.96*** (-3.58)	-15.74* (-2.04)
R2 (pseudo)	0.211	0.277	0.384	0.484	0.536	0.684	0.725
ll	-103.6	-94.92	-73.64	-61.69	-35.46	-24.13	-16.37
N	193	193	173	173	111	111	86
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.11: Non-violent outcome on campaign involvement by student protesters (logit models)

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Student index	1.346*** (5.35)	1.417*** (5.40)	1.370*** (4.65)	1.360*** (4.56)	1.648*** (3.69)	1.388** (3.24)	1.717** (3.14)
Membership size (log)					0.644** (3.13)	0.729** (3.23)	0.627+ (1.96)
Target capacity (log)					-0.335+ (-1.78)	-0.562+ (-1.93)	-0.343 (-1.24)
Target support					1.598* (2.39)	1.306+ (1.95)	3.628* (2.37)
Campaign support					-3.063** (-3.21)	-3.334*** (-3.42)	-6.345** (-3.09)
Regime violence					-1.063 (-0.75)	-0.928 (-0.57)	0.320 (0.15)
Polity index					-0.132* (-2.13)	-0.150* (-2.31)	-0.177* (-2.49)
GDP p.c.(log)					0.513 (1.24)	1.180* (2.46)	1.377 (1.56)
University enrollment					0.00215 (0.60)	0.00340 (0.79)	-0.0198* (-2.14)
Industrial activity					-0.000685 (-0.84)	-0.00148 (-1.43)	0.00240 (1.26)
Constant	-1.637*** (-5.80)	-3.375*** (-4.22)	-0.294 (-0.55)	-1.910* (-2.05)	-12.95*** (-3.41)	-23.35*** (-4.44)	-17.82+ (-1.92)
R2 (pseudo)	0.226	0.287	0.312	0.404	0.537	0.625	0.654
ll	-101.3	-93.35	-81.71	-70.82	-35.62	-28.83	-20.61
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.12: Non-violent outcome on campaign involvement by educated protesters (logit models). Separate indicators for students and graduates, included in the same model

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Graduate index index	0.951*** (3.72)	1.056*** (3.90)	1.590*** (3.79)	1.846*** (3.66)	1.465+ (1.95)	2.168* (2.49)	3.311* (2.24)
Student index	0.943*** (3.86)	1.007*** (3.61)	0.696* (2.44)	0.677* (2.17)	1.215* (2.45)	0.870* (2.26)	0.720 (1.06)
Membership size (log)					0.701*** (3.39)	0.894*** (3.44)	0.690* (1.97)
Target capacity (log)					-0.569+ (-1.94)	-1.105** (-2.60)	-0.181 (-0.73)
Target support					1.950** (2.63)	2.277** (2.69)	6.059* (1.98)
Campaign support					-3.572*** (-3.88)	-4.359*** (-3.93)	-10.54** (-2.59)
Regime violence					-0.991 (-0.77)	-1.229 (-0.99)	0.752 (0.42)
Polity index					-0.112+ (-1.79)	-0.151* (-2.24)	-0.223+ (-1.87)
GDP p.c.(log)					0.662 (1.34)	1.518* (2.37)	1.163 (1.26)
University enrollment					-0.00343 (-0.73)	-0.00633 (-0.88)	-0.0390* (-2.43)
Industrial activity					-0.000627 (-0.71)	-0.000713 (-0.46)	0.0000199 (0.01)
Constant	-2.128*** (-5.99)	-4.504*** (-5.94)	-1.212+ (-1.81)	-4.492*** (-3.88)	-16.29** (-2.83)	-33.75*** (-3.65)	-17.38* (-2.29)
R2 (pseudo)	0.281	0.340	0.408	0.501	0.589	0.699	0.730
ll	-93.76	-86.09	-70.29	-59.33	-31.44	-22.96	-16.09
N	192	192	172	172	111	111	86
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.13: Non-violent outcome on campaign involvement by educated protesters (logit models). Dichotomous education index.

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Education index (dichotomous)	3.460*** (4.50)	3.703*** (4.59)	3.796*** (4.57)	3.820*** (5.67)	3.735*** (3.62)	3.724** (3.10)	6.899* (2.35)
Membership size (log)					0.573** (2.92)	0.696** (2.60)	0.420 (1.61)
Target capacity (log)					-0.325 (-1.50)	-0.609* (-2.21)	0.152 (0.56)
Target support					1.519+ (1.78)	1.889* (2.25)	4.589** (3.07)
Campaign support					-3.022** (-3.19)	-3.794*** (-3.63)	-8.946* (-2.35)
Regime violence					-1.379 (-0.86)	-1.135 (-0.60)	2.530 (1.00)
Polity index					-0.157** (-2.95)	-0.194** (-2.67)	-0.274* (-2.18)
GDP p.c.(log)					0.463 (1.16)	1.430* (2.19)	0.823 (1.23)
University enrollment					0.00399 (1.01)	0.00191 (0.31)	-0.0272 (-1.55)
Industrial activity					-0.000176 (-0.18)	0.000443 (0.31)	0.00255 (0.92)
Constant	-3.029*** (-3.93)	-4.697*** (-5.08)	-2.035* (-2.15)	-3.683*** (-3.55)	-13.05** (-2.74)	-28.48*** (-3.74)	-14.85* (-2.11)
R2 (pseudo)	0.254	0.307	0.387	0.454	0.558	0.670	0.748
ll	-97.29	-90.35	-72.85	-64.84	-33.80	-25.20	-15.01
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

Table 11.14: Non-violent outcome on campaign involvement by students and graduates (logit models). Multiplicative education index.

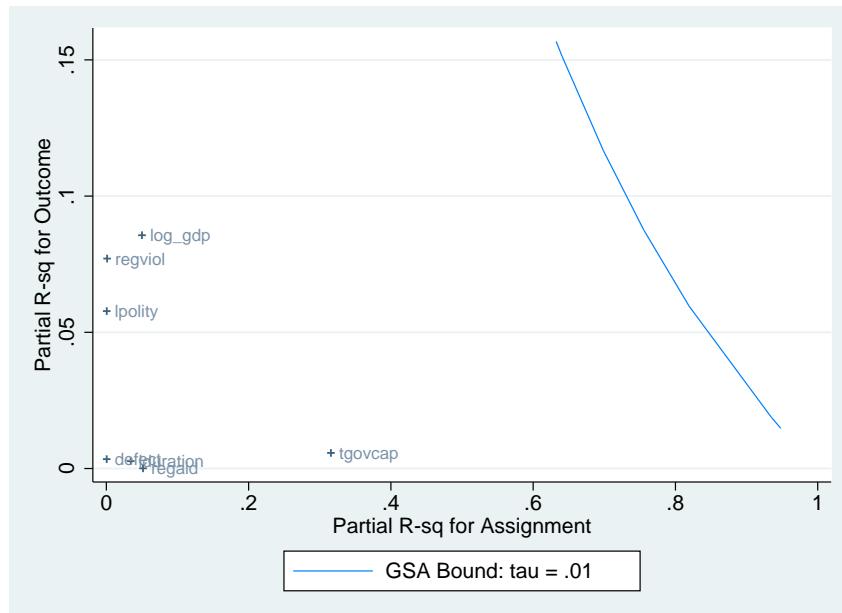
	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Education index 2	1.893*** (5.85)	2.058*** (5.79)	2.122*** (4.94)	2.292*** (4.68)	2.627*** (4.37)	2.708*** (4.51)	3.458** (3.13)
Membership size (log)					0.697*** (3.32)	0.846** (3.27)	0.663+ (1.83)
Target capacity (log)					-0.557* (-2.07)	-0.999* (-2.52)	-0.410 (-1.43)
Target support					1.965** (2.74)	2.225** (2.70)	5.238+ (1.80)
Campaign support					-3.578*** (-3.98)	-4.311*** (-4.13)	-9.301* (-2.13)
Regime violence					-0.961 (-0.69)	-0.939 (-0.66)	0.964 (0.66)
Polity index					-0.114* (-2.00)	-0.154* (-2.30)	-0.215+ (-1.78)
GDP p.c.(log)					0.677 (1.32)	1.575* (2.43)	1.378 (1.39)
University enrollment					-0.00319 (-0.74)	-0.00496 (-0.87)	-0.0337+ (-1.96)
Industrial activity					-0.000646 (-0.73)	-0.000819 (-0.59)	0.00237 (0.94)
Constant	-2.128*** (-5.99)	-4.486*** (-5.72)	-1.040+ (-1.67)	-3.768*** (-3.78)	-16.30** (-2.85)	-32.79*** (-3.50)	-20.34* (-2.09)
R2 (pseudo)	0.281	0.340	0.397	0.486	0.588	0.690	0.711
ll	-93.76	-86.09	-71.62	-61.10	-31.48	-23.71	-17.18
N	192	192	172	172	111	111	86
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily non-violent throughout its activities.

11.7 Alternative specifications of general sensitivity analysis

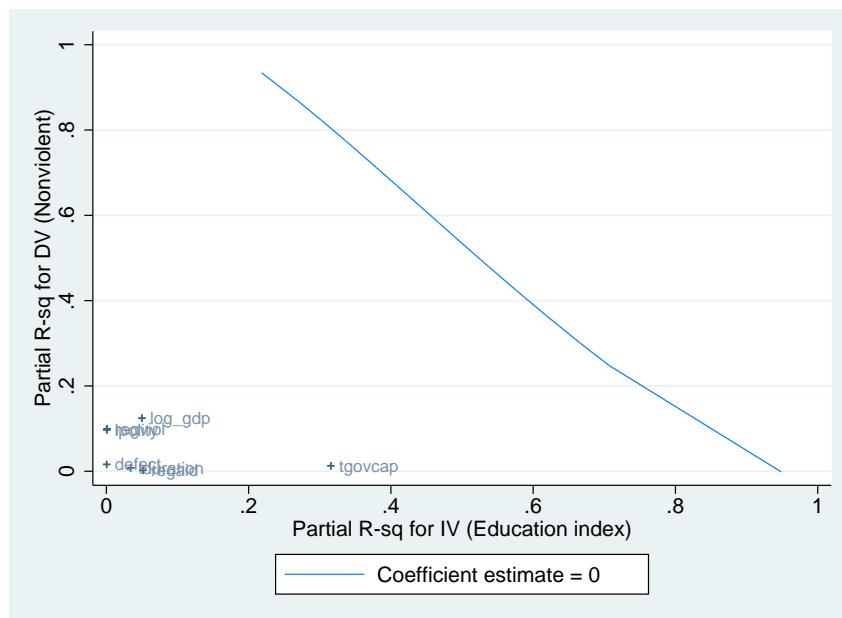
This section includes alternative specifications of the general sensitivity test that is presented and discussed in the paper. The analysis is performed using the generalized sensitivity analysis (GSA) package in Stata (Harada, 2015). The graphs are also created using this package. Figure 11.1 is based on the same model specifications as the GSA figure in the paper, but investigates how large the omitted variable bias would be to push the coefficient estimate for the education index down to zero. As advised by Harada (2015), the main model in the paper and most of the models below are estimated using a linear probability model (ylpm), as this performs well for models with binary outcomes and a low number of observations. However, figures 11.2 and 11.3 present results from GSA analysis estimated using logit models, with the graph in figure 11.2 reducing the coefficient estimate to zero and figure 11.3 reducing it to half of its size. Finally, figure 11.4 differs from the figure in the paper by leaving out all controls but log gdp and polity level, and figure 11.5 is similar to figure 11.4 but also includes region dummies. In brief, all these specifications yield similar conclusions: The line is always far from the included covariates, suggesting that the omitted variable would have to explain a lot more than all the included controls in order to reduce the coefficient estimate by the chosen magnitudes.

Figure 11.1: Assessing the threat of omitted variable bias. Coefficient estimate = 0



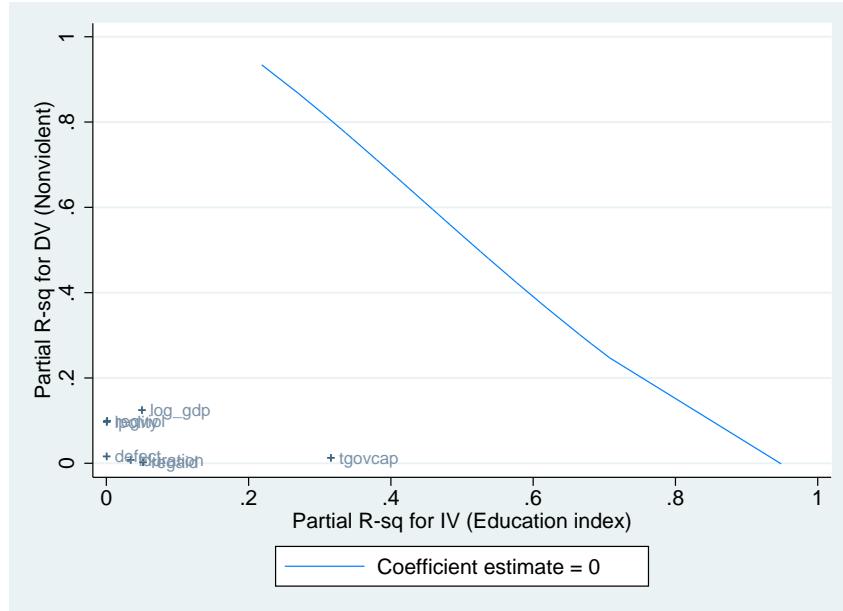
Notes: The plotted line shows where the coefficient estimate for the graduates index (in models explaining nonviolent outcome) would be zero. Included control variables are logged gdp (log gdp), the use of violence of by the regime (regviol), the polity level (lpolity), log government capacity (log govcap), foreign aid to the regime (regaid) and the duration of the campaign (lduration). The model is estimated using a linear probability model.

Figure 11.2: Assessing the threat of omitted variable bias. Coefficient estimate = 0.
Estimation using logit model.



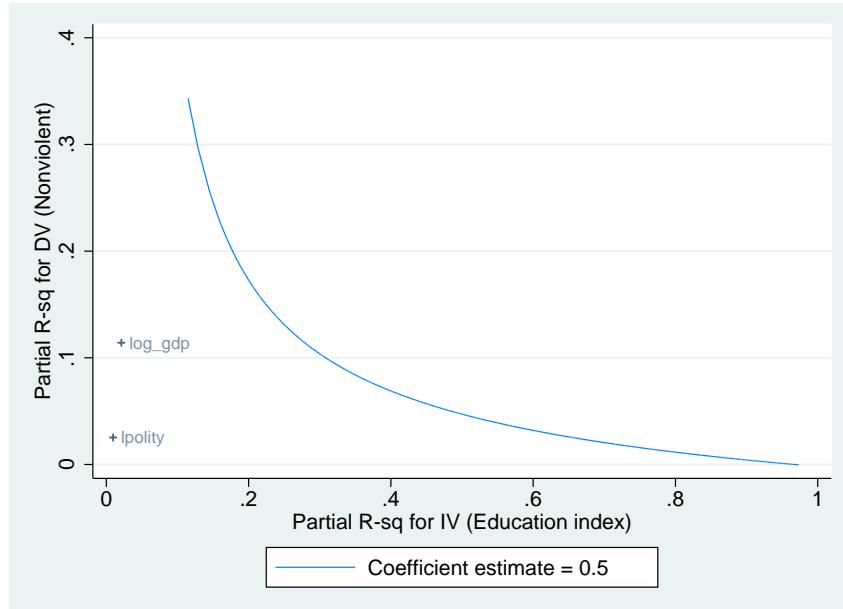
Notes: The plotted line shows where the coefficient estimate for the graduates index (in models explaining nonviolent outcome) would be 0 (reduced from 0.93 with these model specifications). Included control variables are logged gdp (log gdp), the use of violence of by the regime (regviol), the polity level (lpolity), log government capacity (log govcap), foreign aid to the regime (regaid) and the duration of the campaign (lduration). The model is estimated using a linear probability model.

Figure 11.3: Assessing the threat of omitted variable bias. Coefficient estimate is halved. Estimation using logit model.



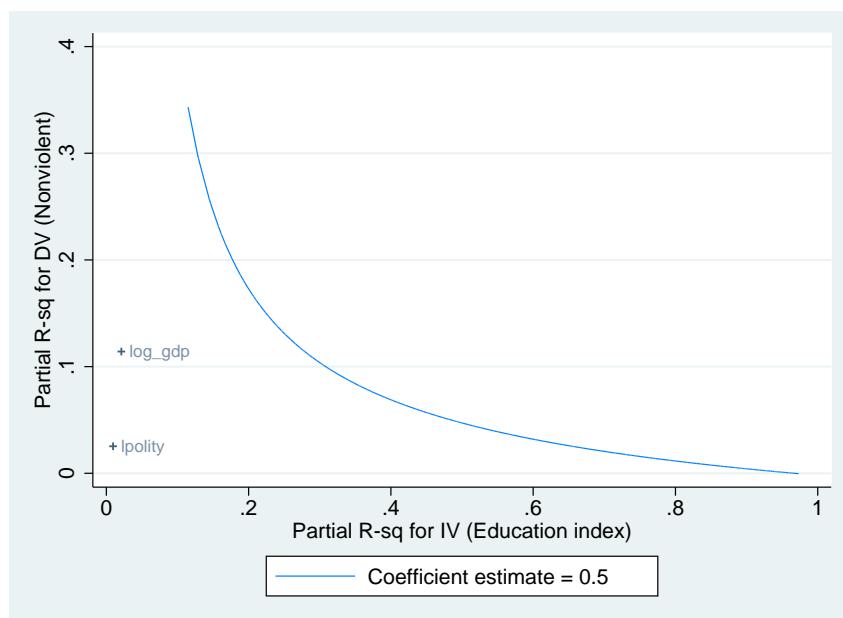
Notes: The plotted line shows where the coefficient estimate for the graduates index (in models explaining nonviolent outcome) would be 0, i.e. approximately halved from its original value (0.93 with these model specifications). Included control variables are logged gdp (log gdp), the use of violence by the regime (regviol), the polity level (lpolity), log government capacity (log govcap), foreign aid to the regime (regaid) and the duration of the campaign (lduration). The model is estimated using a linear probability model.

Figure 11.4: Assessing the threat of omitted variable bias. Coefficient estimate is halved. Simple model.



Notes: The plotted line shows where the coefficient estimate for the graduates index (in models explaining nonviolent outcome) would be 0, i.e. approximately halved from its original value (0.93 with these model specifications). Included control variables are logged gdp (log gdp) and the polity level (lpolity). The model is estimated using a linear probability model.

Figure 11.5: Assessing the threat of omitted variable bias. Coefficient estimate is halved. With region dummies.



Notes: The plotted line shows where the coefficient estimate for the graduates index (in models explaining nonviolent outcome) would be 0, i.e. approximately halved from its original value (0.93 with these model specifications). Included control variables are logged gdp (log_gdp), the use of violence by the regime (regviol), the polity level (lpolity), log government capacity (log_govcap), foreign aid to the regime (regaid), the duration of the campaign (lduration) and region dummies. The model is estimated using a linear probability model.

11.8 Robustness tests of the relationship between education and successful outcome

This section contains numerous robustness tests of models with successful outcome as dependent variable. Most of the models are replications of table 6.2 in the paper, but with some adjustments to the models specifications. Tables 11.15-11.21 contain results from models with alternative operationalizations of the education index. Table 11.22 contains results from multinomial logit models, where the dependent variable is a variable from NAVCO recording whether the campaign failed (0), acccomplished limited success (1), or full success (2). Finally, in table 11.23 one control is left out at a time, and in models 11.23 additional controls are included.

In brief, most of these robustness tests yield similar results to the findings in the paper, confirming that there is some, but not very robust, evidence that participation by educated protesters is related to success. Moreover, these tests also add some nuances to the main finding. For instance, the results in table 11.15-11.17 suggest that the effect of involvement by educated protesters on success is driven by variations in to what extent educated protesters at least participated in or dominated the campaign, and not by variations in to what extent campaigns originated among educated protesters. Moreover, results from table 11.22 suggests that participation by educated protesters only matters for the likelihood of full success, and not for limited success.

Table 11.15: Successful outcome and whether campaigns had some educated protesters (logit models)

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
At least some educated	0.999*** (3.42)	0.951* (2.09)	0.977* (2.44)	0.716 (1.41)	0.369 (1.02)	0.382 (0.75)	0.344 (0.70)	0.00340 (0.01)
Non-violent					1.205** (3.21)	1.469** (2.81)	1.326* (2.14)	2.235** (2.78)
Polity index		0.0230 (0.78)	0.0338 (1.14)	0.0104 (0.26)		0.0475 (1.61)	0.0591+ (1.83)	0.0560 (1.28)
GDP p.c. (log)		0.192 (0.93)	0.296+ (1.67)	0.405 (1.34)		0.0464 (0.21)	0.227 (1.26)	0.326 (1.04)
Target capacity				10.74+ (1.93)				9.327 (1.55)
Duration				-0.235 (-1.50)				-0.186 (-1.13)
Target support				0.693 (1.24)				0.240 (0.40)
Campaign support				0.642 (0.90)				1.612+ (1.68)
Military defection				1.621** (2.74)				1.810** (2.87)
Sanctions against regime				0.259 (0.34)				0.0835 (0.12)
Regime violence				-1.450* (-2.15)				-1.311+ (-1.89)
R2 (pseudo)	0.0403	0.151	0.121	0.321	0.0846	0.201	0.152	0.365
ll	-125.5	-89.32	-91.07	-65.89	-119.7	-84.09	-87.90	-61.67
N	194	156	153	144	194	156	153	144
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies, year dummies and constant are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

Table 11.16: Successful outcome and whether campaigns were dominated by educated protesters (logit models)

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
Dominated by educated	1.066* (2.28)	1.262* (2.53)	1.007+ (1.71)	0.920 (1.27)	0.409 (0.85)	0.642 (1.25)	0.448 (0.77)	0.538 (0.74)
Non-violent					1.169*** (4.09)	1.317** (3.13)	1.202** (2.65)	1.370* (2.57)
Polity index		0.00341 (0.13)	0.00258 (0.10)	-0.0218 (-0.69)		0.0209 (0.80)	0.0162 (0.61)	-0.00191 (-0.06)
GDP p.c. (log)		0.263 (1.49)	0.146 (1.11)	0.147 (0.56)		0.0718 (0.36)	0.0439 (0.30)	0.0170 (0.06)
Target capacity				3.221 (0.85)				3.249 (0.83)
Duration					-0.233* (-2.01)			-0.174 (-1.52)
Target support					0.0962 (0.21)			-0.0727 (-0.15)
Campaign support					0.0768 (0.14)			0.453 (0.73)
Military defection					1.715*** (3.35)			1.754*** (3.42)
Sanctions against regime					0.795 (1.50)			0.757 (1.34)
Regime violence					-0.952 (-1.63)			-0.761 (-1.31)
Constant	-0.741*** (-5.61)	-4.252** (-2.69)	-0.287 (-0.27)	0.0394 (0.01)	-1.101*** (-6.54)	-3.105+ (-1.87)	-0.554 (-0.49)	-0.529 (-0.20)
R2 (pseudo)	0.0190	0.108	0.0782	0.250	0.0655	0.153	0.110	0.274
ll	-200.3	-126.8	-129.0	-98.22	-190.8	-120.4	-124.5	-95.05
N	316	218	213	202	316	218	213	202
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies, year dummies and constant are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

Table 11.17: Successful outcome and whether campaigns originated among educated protesters (logit models)

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
Originated among educated	0.329 (1.00)	0.630 (1.43)	0.473 (1.21)	0.545 (0.98)	-0.149 (-0.43)	0.290 (0.66)	0.110 (0.27)	0.366 (0.68)
Non-violent					1.439*** (4.39)	1.532** (3.06)	1.449** (2.66)	2.111** (2.71)
Polity index		0.0146 (0.47)	0.0228 (0.77)	-0.00385 (-0.09)		0.0425 (1.41)	0.0508 (1.60)	0.0413 (0.95)
GDP p.c. (log)		0.316 (1.50)	0.343* (2.04)	0.448 (1.50)		0.0959 (0.39)	0.221 (1.24)	0.317 (0.99)
Target capacity					10.19+ (1.86)			9.223 (1.61)
Duration					-0.199 (-1.40)			-0.139 (-0.89)
Target support					0.652 (1.13)			0.210 (0.35)
Campaign support					0.490 (0.71)			1.468 (1.61)
Military defection					1.714** (2.90)			1.872** (3.01)
Sanctions against regime					0.376 (0.53)			0.153 (0.23)
Regime violence					-1.506* (-2.39)			-1.302+ (-1.90)
R2 (pseudo)	0.00436	0.141	0.0954	0.314	0.0817	0.201	0.142	0.361
ll	-130.2	-90.36	-94.24	-66.90	-120.0	-84.04	-89.38	-62.34
N	194	156	154	145	194	156	154	145
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies, year dummies and constant are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

Table 11.18: Successful outcome on campaign involvement by graduates (logit models)

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
Graduate index	0.397* (2.09)	0.397 (1.64)	0.365 (1.54)	0.229 (0.80)	0.0127 (0.06)	0.0838 (0.33)	0.0183 (0.07)	-0.106 (-0.33)
Non-violent					1.369*** (3.82)	1.496** (2.93)	1.505* (2.57)	2.320** (2.81)
Polity index		0.0248 (0.81)	0.0354 (1.18)	0.00819 (0.19)		0.0465 (1.55)	0.0608+ (1.88)	0.0561 (1.26)
GDP p.c. (log)		0.292 (1.36)	0.319+ (1.82)	0.452 (1.51)		0.109 (0.45)	0.233 (1.29)	0.328 (1.07)
Target capacity				9.163+ (1.69)				10.16+ (1.78)
Duration					-0.246 (-1.63)			-0.189 (-1.17)
Target support					0.670 (1.22)			0.255 (0.43)
Campaign support					0.615 (0.87)			1.597+ (1.66)
Military defection					1.646** (2.79)			1.821** (2.89)
Sanctions against regime					0.350 (0.47)			0.0467 (0.07)
Regime violence					-1.529* (-2.35)			-1.295+ (-1.86)
R2 (pseudo)	0.0214	0.150	0.108	0.315	0.0797	0.200	0.150	0.365
ll	-127.4	-89.04	-92.47	-66.49	-119.8	-83.72	-88.13	-61.61
N	193	155	153	144	193	155	153	144
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

Table 11.19: Successful outcome on campaign involvement by students (logit models)

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
Student index	0.347*	0.537*	0.423*	0.352	-0.0103	0.237	0.151	0.135
	(2.13)	(2.13)	(2.03)	(1.18)	(-0.05)	(0.87)	(0.64)	(0.44)
Non-violent					1.382***	1.376*	1.348*	2.137**
					(3.76)	(2.55)	(2.25)	(2.70)
Polity index	0.0260	0.0350	0.00710			0.0466	0.0592+	0.0519
	(0.84)	(1.16)	(0.17)			(1.54)	(1.84)	(1.20)
GDP p.c. (log)	0.250	0.349*	0.450			0.0712	0.238	0.328
	(1.18)	(2.00)	(1.49)			(0.30)	(1.31)	(1.05)
Target capacity				9.178			8.900	
				(1.56)			(1.44)	
Duration				-0.219			-0.177	
				(-1.47)			(-1.09)	
Target support				0.604			0.194	
				(1.03)			(0.32)	
Campaign support				0.632			1.602+	
				(0.90)			(1.68)	
Military defection				1.667**			1.808**	
				(2.79)			(2.95)	
Sanctions against regime				0.411			0.174	
				(0.57)			(0.26)	
Regime violence				-1.408*			-1.302+	
				(-2.15)			(-1.89)	
R2 (pseudo)	0.0210	0.159	0.114	0.320	0.0785	0.200	0.148	0.365
ll	-127.1	-87.69	-91.05	-65.42	-119.6	-83.48	-87.52	-61.05
N	193	155	152	143	193	155	152	143
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

Table 11.20: Successful outcome on campaign involvement by educated protesters (logit models). Multiplicative education index.

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
Non-violent					1.369*** (3.51)	1.361* (2.51)	1.386* (2.25)	2.209** (2.71)
Education index 2	0.455* (2.17)	0.602* (1.99)	0.496+ (1.87)	0.381 (1.10)	-0.00994 (-0.04)	0.219 (0.66)	0.117 (0.38)	0.0252 (0.07)
Polity index		0.0269 (0.87)	0.0375 (1.24)	0.00882 (0.21)		0.0456 (1.50)	0.0596+ (1.85)	0.0527 (1.20)
GDP p.c. (log)		0.269 (1.22)	0.319+ (1.81)	0.445 (1.47)		0.103 (0.42)	0.228 (1.27)	0.330 (1.07)
Target capacity				8.558 (1.47)			9.227 (1.52)	
Duration					-0.227 (-1.50)			-0.185 (-1.14)
Target support					0.588 (1.03)			0.199 (0.33)
Campaign support					0.661 (0.94)			1.602+ (1.67)
Military defection					1.650** (2.80)			1.816** (2.96)
Sanctions against regime					0.418 (0.58)			0.142 (0.21)
Regime violence					-1.476* (-2.28)			-1.331+ (-1.93)
R2 (pseudo)	0.0252	0.162	0.114	0.318	0.0773	0.200	0.147	0.364
ll	-126.0	-87.00	-91.08	-65.56	-119.3	-83.08	-87.65	-61.16
N	192	154	152	143	192	154	152	143
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

Table 11.21: Successful outcome on campaign involvement by educated protesters (logit models). Dichotomous education index.

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
Education index (dichotomous)	0.928** (3.01)	0.983+ (1.79)	1.027* (2.18)	0.929 (1.63)	0.222 (0.57)	0.215 (0.32)	0.258 (0.41)	0.127 (0.19)
Non-violent					1.247*** (3.36)	1.454** (2.61)	1.351* (2.06)	2.163** (2.80)
Polity index		0.0145 (0.49)	0.0239 (0.84)	0.000495 (0.01)		0.0426 (1.41)	0.0557+ (1.67)	0.0509 (1.19)
GDP p.c. (log)		0.255 (1.20)	0.307+ (1.76)	0.433 (1.46)		0.100 (0.43)	0.228 (1.26)	0.328 (1.06)
Target capacity				9.706+ (1.68)				9.284 (1.53)
Duration				-0.216 (-1.40)				-0.184 (-1.11)
Target support				0.657 (1.14)				0.210 (0.35)
Campaign support				0.713 (0.96)				1.599+ (1.67)
Military defection				1.636*** (2.78)				1.808** (2.93)
Sanctions against regime				0.425 (0.59)				0.150 (0.22)
Regime violence				-1.440* (-2.09)				-1.326+ (-1.91)
R2 (pseudo)	0.0318	0.153	0.118	0.323	0.0785	0.197	0.147	0.364
ll	-125.2	-87.94	-90.61	-65.09	-119.2	-83.31	-87.64	-61.14
N	192	154	152	143	192	154	152	143
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

Table 11.22: Successful outcome on campaign involvement by educated protesters. Multi-nomial logit models (base category = failure)

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
<i>Limited success</i>								
Education index	0.0949 (0.63)	0.126 (0.69)	0.0361 (0.19)	0.0313 (0.13)	-0.391* (-1.99)	-0.479+ (-1.93)	-0.401+ (-1.73)	-0.437 (-1.49)
Non-violent					2.758*** (4.62)	3.680*** (4.39)	2.680*** (3.44)	3.986*** (3.86)
Polity index		-0.0446 (-1.04)	-0.0742 (-1.51)	-0.0886+ (-1.71)		-0.0140 (-0.25)	-0.0416 (-0.78)	-0.0169 (-0.25)
GDP p.c. (log)		0.208 (0.92)	-0.0556 (-0.20)	0.409 (0.95)		-0.0459 (-0.21)	-0.263 (-0.98)	0.234 (0.57)
Target capacity (log)				-0.378+ (-1.74)				-0.361+ (-1.72)
Duration					0.215 (0.70)			0.379 (1.19)
Target support					1.011 (1.23)			0.381 (0.40)
Campaign support					-0.770 (-0.76)			0.502 (0.41)
Military defection					0.237 (0.30)			0.866 (0.99)
Sanctions against regime					-1.583 (-1.12)			-2.116 (-1.46)
Regime violence					-0.688 (-0.71)			-1.174 (-1.15)
<i>Full success</i>								
Education index	0.254* (2.02)	0.339* (2.00)	0.254 (1.58)	0.253 (1.37)	-0.135 (-0.89)	-0.0167 (-0.09)	-0.0906 (-0.48)	-0.104 (-0.46)
Non-violent					2.281*** (4.69)	2.453*** (3.58)	2.308** (3.19)	3.576** (3.15)
Polity index		0.0147 (0.45)	0.0164 (0.48)	-0.0226 (-0.54)		0.0476 (1.36)	0.0504 (1.34)	0.0487 (0.95)
GDP p.c. (log)		0.323 (1.42)	0.300 (1.54)	0.539 (1.51)		0.0502 (0.19)	0.118 (0.57)	0.243 (0.59)
Target capacity (log)				-0.144 (-0.80)				-0.138 (-0.64)
Duration					-0.176 (-1.05)			-0.0833 (-0.41)
Target support					0.857 (1.40)			0.227 (0.32)
Campaign support					0.671 (0.90)			1.919+ (1.85)
Military defection					1.777* (2.51)			2.268* (2.56)
Sanctions against regime					0.0923 (0.11)			-0.516 (-0.57)
Regime violence					-1.706* (-2.31)			-2.004* (-2.30)
r2 (pseudo)	0.0180	0.141	0.171	0.334	0.111	0.238	0.230	0.412
ll	-191.8	-137.1	-132.3	-99.65	-173.7	-121.5	-122.8	-88.07
N	192	154	154	145	192	154	154	145
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a categorical indicator from NAVCO 1.0 indicating whether the campaign failed in terms of accomplishing their main goals (0), achieved a limited success (1) or a full success (2).

Base category is 0.

Table 11.23: Successful outcome on campaign involvement by educated protesters (logit models). Omitting one control at a time.

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
Education index	0.265+ (1.86)	0.291* (2.03)	0.273+ (1.95)	0.248+ (1.73)	0.248+ (1.78)	0.262+ (1.88)	0.288* (2.02)	0.156 (1.32)
Duration	-0.290** (-2.61)		-0.260* (-2.12)	-0.239* (-1.98)	-0.223+ (-1.68)	-0.252* (-2.13)	-0.273* (-2.37)	-0.207* (-1.98)
Target support	0.385 (0.89)	0.214 (0.47)		0.290 (0.62)	0.223 (0.51)	0.277 (0.59)	0.214 (0.47)	0.272 (0.73)
Campaign support	0.718 (1.45)	0.292 (0.56)	0.548 (1.01)		0.400 (0.80)	0.618 (1.20)	0.339 (0.65)	0.392 (0.85)
Military defection	1.768*** (3.92)	1.728*** (3.71)	1.758*** (3.82)	1.748*** (3.78)		1.824*** (3.96)	1.694*** (3.78)	1.167*** (3.54)
Sanctions against regime	0.597 (0.96)	0.470 (0.86)	0.625 (0.99)	0.705 (1.15)	0.988 (1.57)		0.614 (0.93)	1.232* (2.36)
Regime violence	-1.350* (-2.27)	-1.564** (-2.79)	-1.545** (-2.64)	-1.473* (-2.44)	-1.339** (-2.63)	-1.579* (-2.55)		-1.470* (-2.55)
Polity index	0.0478 (1.32)	0.0302 (0.76)	0.0348 (0.89)	0.0334 (0.87)	0.0248 (0.73)	0.0344 (0.89)	0.0511 (1.42)	
GDP p.c. (log)	0.405* (2.16)	0.487* (2.35)	0.464* (2.20)	0.430* (1.99)	0.341+ (1.77)	0.474* (2.27)	0.382+ (1.95)	
Target capacity		2.209 (0.49)	2.246 (0.48)	2.599 (0.60)	1.728 (0.41)	2.737 (0.57)	2.513 (0.59)	0.555 (0.14)
R2 (pseudo)	0.256	0.231	0.249	0.246	0.148	0.245	0.217	0.157
ll	-76.50	-74.73	-72.99	-73.24	-83.57	-73.33	-76.04	-96.89
N	153	145	145	145	146	145	145	169

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies, year dummies and constant are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

Table 11.24: Successful outcome on campaign involvement by educated protesters (logit models). Additional controls

	B1 logit	B2 logit
Education index	0.277 (1.45)	0.243 (0.92)
Target capacity (log)	-0.163 (-0.91)	-0.680 (-1.35)
Duration	-0.310+ (-1.95)	-0.562** (-3.19)
Target support	0.0193 (0.04)	-0.0604 (-0.07)
Campaign support	0.323 (0.39)	-0.952 (-1.41)
Military defection	1.756*** (3.31)	2.013* (2.47)
Sanctions against regime	0.819 (1.12)	0.178 (0.20)
Regime violence	-0.272 (-0.43)	-1.737* (-2.53)
Polity index	0.0118 (0.27)	0.000273 (0.01)
GDP p.c. (log)	0.211 (0.87)	-0.106 (-0.28)
Target in war	-1.133 (-0.63)	
Simultaneous violent campaign	-0.721 (-1.31)	
Number of wars	0.346 (0.68)	
Industrial activity		0.00410** (2.75)
University enrollment		-0.00504 (-0.89)
Urbanization		0.00128 (0.50)
GDP growth p.c.		-0.0286 (-0.61)
Population (log)		0.228 (0.37)
R2 (pseudo)	0.236	0.343
ll	-53.29	-37.64
N	106	86
Region dummies?	no	no
Decade year dummies?	no	no

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies, year dummies and constant are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

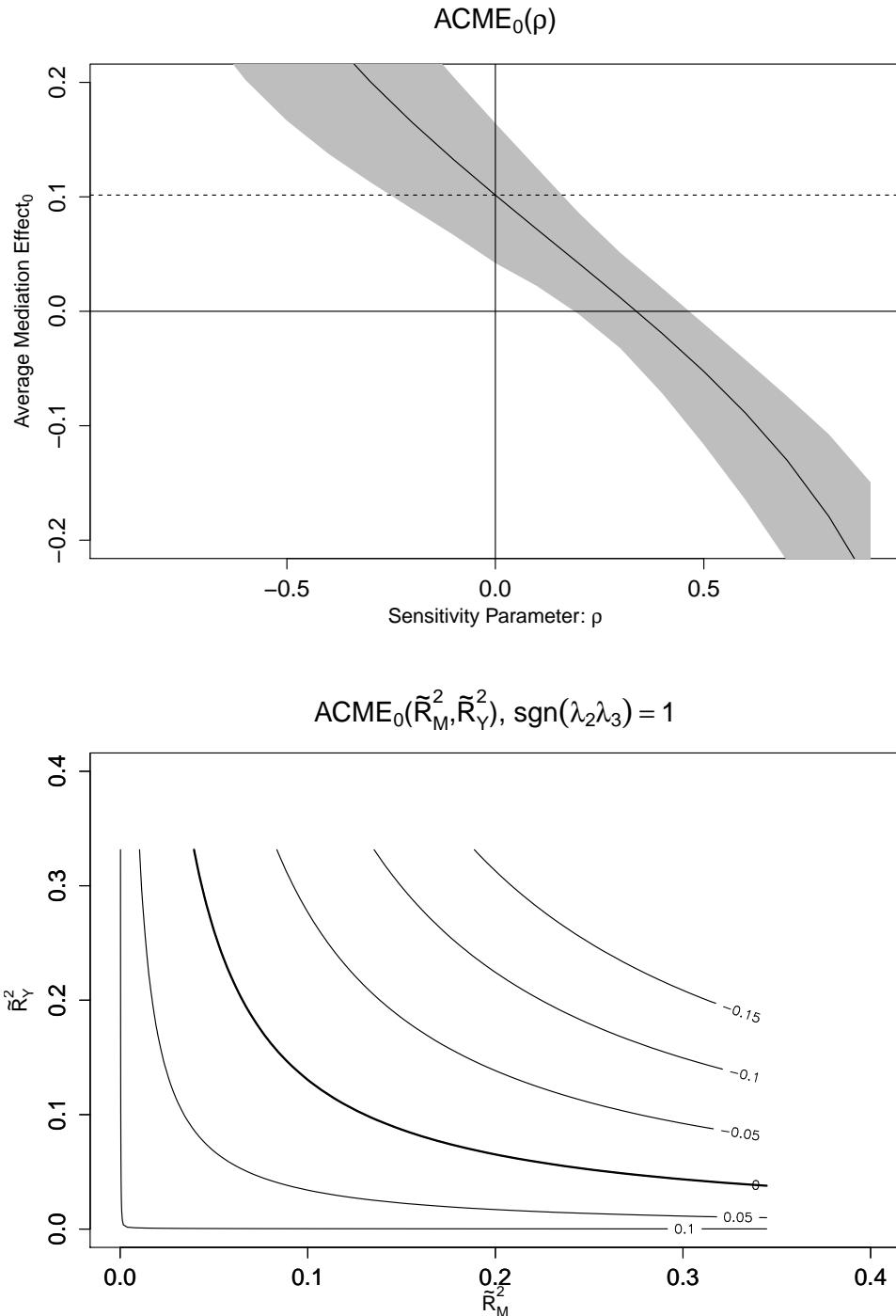
11.9 Robustness tests of mediation analysis

Section 6.5 in the paper estimated that the mediation effect from education on the likelihood of success through nonviolence is positive and statistically significant, relying on the causal mediation analysis procedure by (Imai et al., 2011). The graph in the left panel shows the size of the real ACME plotted against values of a sensitivity parameter, that equals the correlation between the error terms in the mediator and outcome models. When the sensitivity parameter is zero, the real ACME will coincide with the value reported in figure 6.6 (0.13). As can be seen from the graph, for the real ACME to be zero, this sensitivity parameter must exceed 0.3. In other words, to conclude that the true ACME is not significantly different from zero, we must assume an omitted variable that makes the correlation between the error terms larger than 0.3. In order to interpret this in more substantial terms, we can inspect the contours in the right panel which represent the true ACME as a function of the variance in the mediator (on the x-axis) and the outcome (on the y-axis) that is explained by the omitted variable. The true ACME becomes zero when, for instance, the omitted variable explains 20 percent of the variance in the likelihood of success (the outcome) and 8 percent of the variance in non-violence (the mediator). Given that the model controls for many plausible confounding factors such as socio-economic indicators and characteristics of the campaign and the target it is unlikely that an unobserved factor of that magnitude exists, but it can not be ruled out, and the results from the mediation analysis should therefore be treated with caution.

11.10 Education profile of campaign leadership

As described in the codebook, I also identified the leader or leadership of each campaign and coded whether the leader was a student or educated. To account for the fact that educated leaders may be crucial to the campaign even in instances when the rest of the campaign was not dominated by educated participants, I also conduct tests where I add information on whether the leader was educated or student to the education index. First, I create a dichotomous variable with the value 1 if leader was educated or a students. Second, I add this variable to the education index, which in practice means adding one point to the index for every campaign with educated leader. Table 6.1 in the paper was replicated with this alternative education index, and the results are presented below. In brief, there is still strong and robust evidence that educated campaigns are more likely

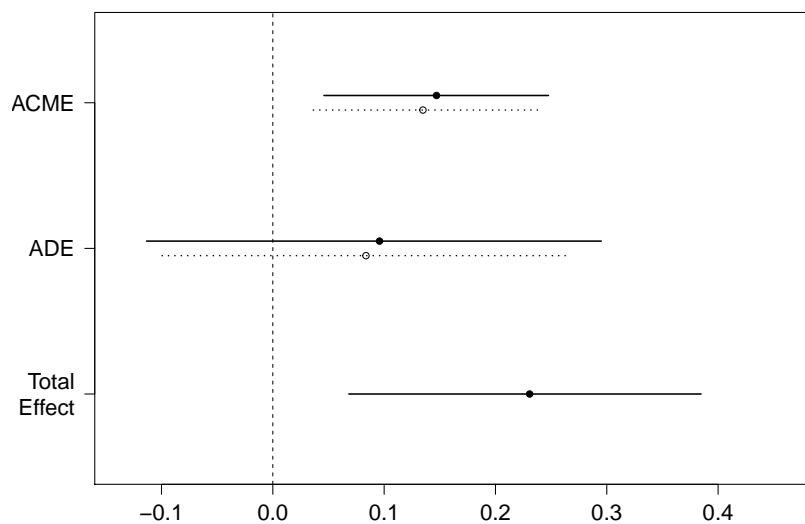
Figure 11.6: Sensitivity analysis of mediation effect



In the left panel, the true ACME is plotted against sensitivity parameter, which is the correlation between the error terms in the mediator and outcome regression models. In the right panel, the contours represent the true ACME plotted as a function of the proportion of the total mediator variance (the y-axis) and the total outcome variance (the x-axis), that are explained by the unobserved confounder included in the regression models.

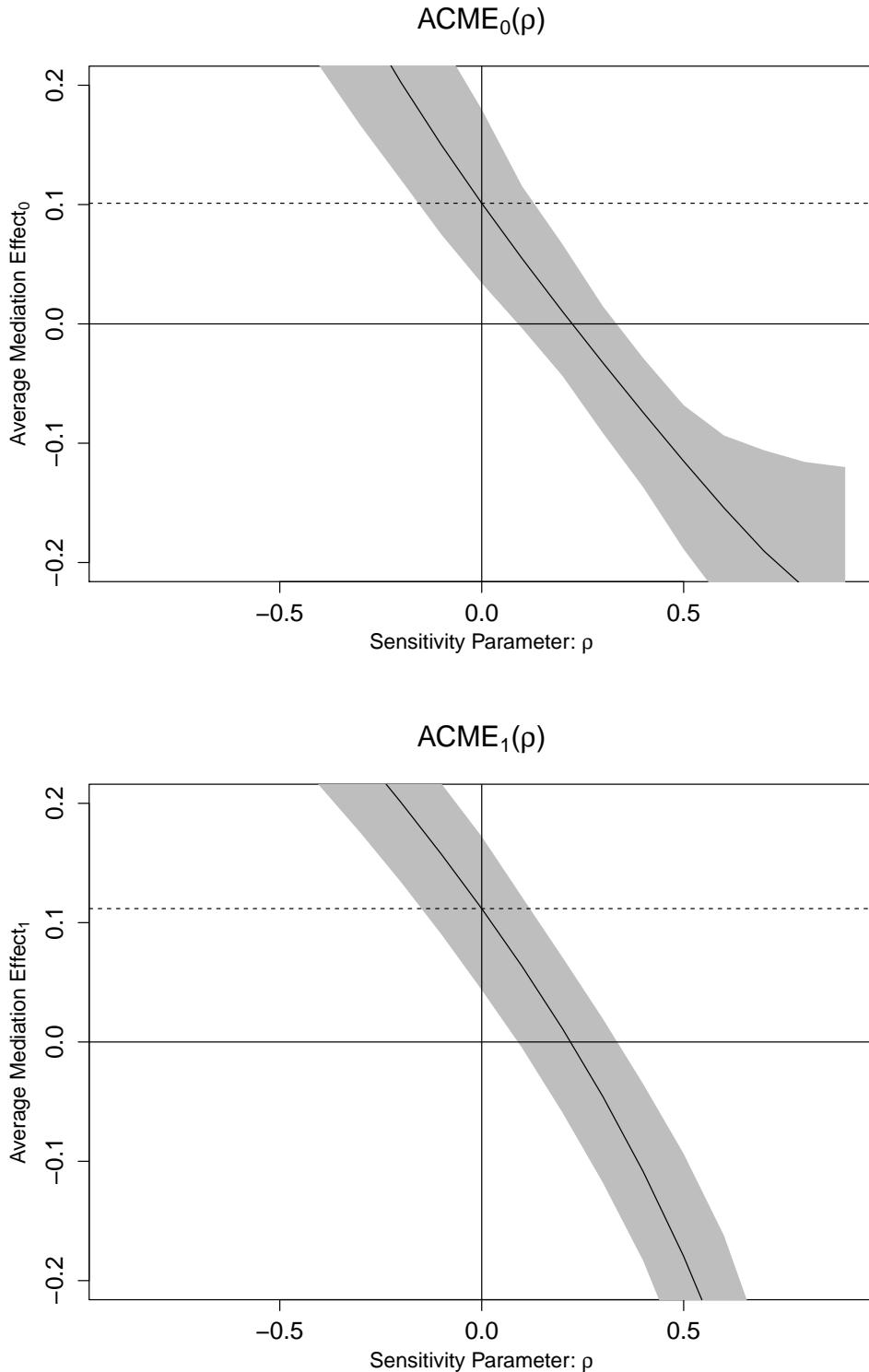
to become nonviolent.

Figure 11.7: The mediation effect of education on successful outcome, mediated through nonviolence. Simple model.



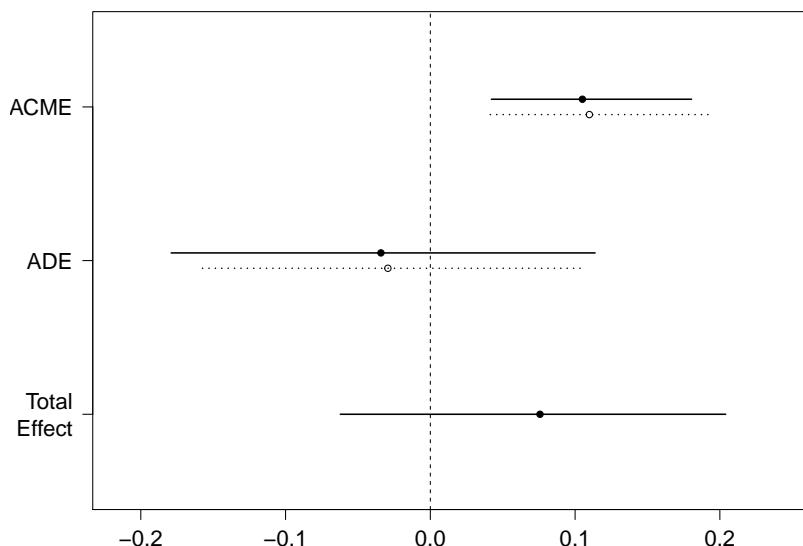
Notes: The total effect of education on the likelihood of success is decomposed into the average direct effect (ADE) and the average causal mediation effect (ACME). The effects are plotted with their 95 percent confidence intervals. Education is measured using a binary version of the combined education index. The pre-treatment covariates are log gdp and the polity index.

Figure 11.8: Sensitivity analysis of mediation effect. Simple model



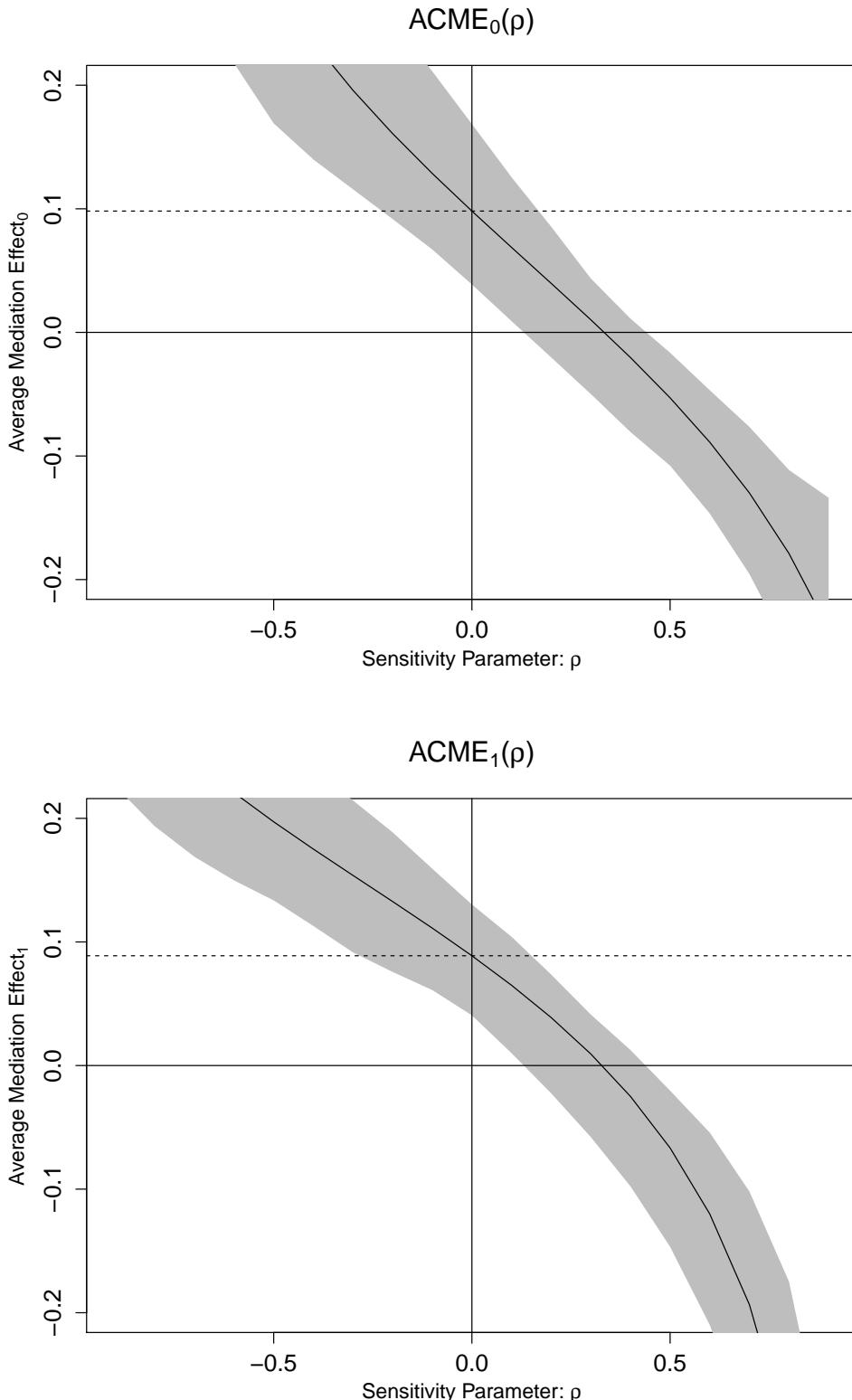
In the left panel, the true ACME is plotted against sensitivity parameter, which is the correlation between the error terms in the mediator and outcome regression models. In the right panel, the contours represent the true ACME plotted as a function of the proportion of the total mediator variance (the y-axis) and the total outcome variance (the x-axis), that are explained by the unobserved confounder included in the regression models.

Figure 11.9: The mediation effect of education on successful outcome, mediated through nonviolence. Categorical education index.



Notes: The total effect of education on the likelihood of success is decomposed into the average direct effect (ADE) and the average causal mediation effect (ACME). The effects are plotted with their 95 percent confidence intervals. Education is measured using the education index (ranging from 0-3). The pre-treatment covariates are log gdp and the polity index.

Figure 11.10: Sensitivity analysis of mediation effect. Categorical education index.



In the left panel, the true ACME is plotted against sensitivity parameter, which is the correlation between the error terms in the mediator and outcome regression models. In the right panel, the contours represent the true ACME plotted as a function of the proportion of the total mediator variance (the y-axis) and the total outcome variance (the x-axis), that are explained by the unobserved confounder included in the regression models.

Table 11.25: Non-violent outcome on campaign involvement by students and graduates (logit models)

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Education index 2	0.737*** (5.55)	0.813*** (5.49)	0.827*** (4.55)	0.878*** (4.51)	0.971*** (4.02)	1.153*** (4.23)	1.104*** (4.07)
Membership(log)					0.692*** (3.33)	0.901*** (3.34)	0.662* (2.27)
Target capacity (log)					-0.458+ (-1.90)	-0.992* (-2.30)	-0.373 (-1.48)
Target support					1.606* (2.21)	2.051* (2.50)	4.071* (2.13)
Campaign support					-3.288** (-3.19)	-4.123*** (-4.22)	-7.528** (-2.82)
Regime violence					-1.186 (-0.87)	-1.080 (-0.76)	0.252 (0.17)
Polity					-0.116* (-2.05)	-0.162* (-2.15)	-0.194* (-2.40)
GDP per capita(log)					0.695 (1.33)	1.753* (2.36)	1.621+ (1.83)
University enrollment					-0.00201 (-0.45)	-0.00440 (-0.90)	-0.0258* (-2.31)
Industrial activity					-0.000516 (-0.54)	-0.00124 (-0.97)	0.000609 (0.30)
Constant	-2.010*** (-5.68)	-4.165*** (-6.48)	-0.809 (-1.33)	-3.146*** (-3.70)	-15.35** (-3.00)	-33.84*** (-3.34)	-20.46* (-2.49)
R2 (pseudo)	0.237	0.301	0.344	0.436	0.554	0.680	0.680
ll	-99.55	-91.19	-77.92	-66.95	-34.05	-24.42	-19.07
ll0							
N	192	192	172	172	111	111	86
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-scores in parentheses. Standard errors are clustered on country. Constant term, region and decade dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily nonviolent.

11.11 Multiple imputation of missing values

Table 11.26: Non-violent outcome on campaign involvement by students and graduates (logit models)

	A1 logit	A2 logit	A3 logit	A4 logit	A5 logit	A6 logit	A7 logit
Education index	0.950*** (5.79)	1.010*** (5.77)	1.070*** (4.48)	1.178*** (4.38)	1.151*** (5.01)	1.415*** (5.91)	1.179*** (4.25)
Membership size (log)					0.813*** (4.31)	0.836*** (4.03)	0.857*** (4.01)
Target capacity (log)					-0.526* (-2.31)	-0.661* (-2.33)	-0.542+ (-1.92)
Target support					1.083+ (1.70)	1.437* (2.35)	1.234 (1.34)
Campaign support					-3.014*** (-3.46)	-3.670*** (-3.51)	-3.389*** (-3.84)
Regime violence					-2.002* (-2.29)	-1.909* (-2.02)	-1.457 (-1.40)
Polity index					-0.0901+ (-1.81)	-0.0929+ (-1.83)	-0.0845 (-1.34)
GDP p.c.(log)					0.625+ (1.67)	0.926* (2.46)	0.665 (1.57)
University enrollment					0.00287 (0.82)	0.00241 (0.67)	-0.00268 (-0.57)
Industrial activity					-0.000746 (-0.97)	-0.00169* (-2.04)	0.0000417 (0.03)
N	203	203	181	181	203	203	181
Region dummies?	no	yes	no	yes	no	yes	no
Decade year dummies?	no	no	yes	yes	no	no	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-scores in parentheses. Standard errors are clustered on country. Constant term, region and decade dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign was primarily nonviolent.

Table 11.27: Successful outcome on campaign involvement by graduates (logit models)

	B1 logit	B2 logit	B3 logit	B4 logit	B5 logit	B6 logit	B7 logit	B8 logit
Education index	0.270* (2.57)	0.276* (2.56)	0.236* (2.19)	0.257* (2.14)	0.0321 (0.25)	0.0613 (0.47)	0.0212 (0.17)	0.0363 (0.25)
Non-violent					1.420*** (3.73)	1.466*** (3.34)	1.535** (3.13)	1.994*** (3.54)
Polity index		0.0271 (0.90)	0.0323 (1.12)	0.0316 (0.85)		0.0453 (1.52)	0.0504+ (1.70)	0.0597 (1.47)
GDP p.c. (log)		0.101 (0.68)	0.110 (0.78)	0.00623 (0.03)		-0.0324 (-0.20)	0.0175 (0.12)	-0.0507 (-0.25)
Target capacity					-0.0396 (-0.34)			-0.0408 (-0.29)
Duration					-0.263* (-2.21)			-0.190 (-1.56)
Target support					0.569 (1.33)			0.432 (1.02)
Campaign support					0.696 (1.47)			1.305* (2.41)
Military defection					1.553*** (3.91)			1.710*** (4.13)
Sanctions against regime					0.799 (1.56)			0.622 (1.17)
Regime violence					-1.323** (-2.65)			-0.942+ (-1.73)
Constant	-0.912*** (-4.25)	-2.592* (-2.07)	-0.995 (-0.96)	0.141 (0.07)	-1.108*** (-4.99)	-1.554 (-1.19)	-0.992 (-0.88)	-0.764 (-0.35)
N	203	203	203	203	203	203	203	203
Region dummies?	no	yes	no	yes	no	yes	no	yes
Decade year dummies?	no	no	yes	yes	no	no	yes	yes

Notes: + $p < .10$; * $p < .05$; ** $p < .01$. Z-values in parentheses. Standard errors are clustered on country. Region dummies are omitted from the table. Dependent variable is a dichotomous indicator from NAVCO 1.0 indicating whether the campaign succeeded in terms of accomplishing their main goals.

11.12 List of campaigns included in the analysis

Campaigns aiming for regime change

eyear	byear	location	campaign	target	nonviol	success
1959	1956	Cuba	Cuban Revolution	Batista regime	0	1
1985	1985	Haiti		Jean Claude Duvalier	1	1
1965	1965	Dominican Republic	leftists	Loyalist regime	0	0
1920	1910	Mexico	Liberals and Radicals rebellion	Diaz regime	0	1
1924	1923	Mexico	Huerta led rebels	Obregon regime	0	0
1930	1926	Mexico	Cristeros rebellion	Mexican regime	0	0
1929	1929	Mexico	Escobar-led rebellion	Calles government	0	0
2000	1987	Mexico		corrupt govt	1	1
2006	2006	Mexico		Calderon regime	1	0
1944	1944	Guatemala	October Revolutionaries	Ubico dictatorship	1	1
1954	1954	Guatemala	Conservative movement	Arbenz leftist regime	0	1
1996	1961	Guatemala	Marxist rebels (URNG)	government of Guatemala	0	0
1924	1924	Honduras	Conservative movement	Carias regime	0	0
1932	1932	El Salvador	leftist rebellion	authoritarian Martinez regime	0	0
1944	1944	El Salvador	Strike of Fallen Arms	Martinez dictatorship	1	1
1981	1979	El Salvador		Mil/civ junta	1	0
1991	1979	El Salvador	Farabundo Marti (FMLN)	El Salvador government	0	0
1932	1925	Nicaragua	Nicaraguan Guerrillas	US intervening troops	0	1
1979	1978	Nicaragua	FSLN	Nicaraguan regime	0	0
1990	1980	Nicaragua	Contras	Sandinista regime	0	0
1948	1948	Costa Rica	National Union Party	Calderon regime	0	1
1989	1987	Panama		Noriega regime	1	0
1949	1948	Colombia	Conservative movement	Liberal government	0	1
1949	1949	Colombia	Liberals of 1949	Conservative govt	0	0
2006	1964	Colombia	Revolutionary Armed Forces of Colombia	Colombia govt and US influence	0	0
1903	1901	Venezuela	La Revolucion Libertador	Venezuela government	0	0
1958	1958	Venezuela		Jimenez dictatorship	1	1

1963	1958	Venezuela	Armed Forces for National Liberation (FALN)	Betancourt regime	0	0
2002	2002	Venezuela	anti-coup	anti-Chavez coup	1	1
1992	1990	Guyana		Burnham/Hoyte autocratic regime	1	1
1932	1932	Peru	Aprista rebels	Cerro regime	0	0
1995	1980	Peru	Senderista Insurgency (Sendero Luminoso)	Peruvian government	0	0
1997	1996	Peru	Senderista Insurgency (Tupac Amaru)	Peruvian government	0	0
2000	2000	Peru		Fujimori govt	1	1
1932	1932	Brazil	Padistas	Brazilian regime	0	0
1985	1984	Brazil	diretas já	Military rule	1	1
1952	1952	Bolivia	leftists	military junta	0	1
1982	1977	Bolivia		Mil juntas	1	1
1947	1947	Paraguay	leftist rebellion	Morinigo regime	0	0
1931	1931	Chile		Ibanez regime	1	1
1973	1973	Chile	Pinochet-led rebels	Allende regime	0	1
1989	1983	Chile		Augusto Pinochet	1	1
1977	1973	Argentina	ERP/Monteneros	Argentina regime	0	0
1981	1977	Argentina	pro-democracy movement	military junta	1	1
1904	1904	Uruguay	Blancos Rebellion	Ordonez regime	0	0
1972	1963	Uruguay	Tupamaros	Uruguay government	0	0
1985	1984	Uruguay		Military rule	1	1
1934	1934	Spain	Asturian miners	right-wing government	0	0
1939	1936	Spain	Fascists	Republican government	0	1
1974	1974	Portugal	Carnation Revolution	Military rule	1	1
1953	1953	East Germany		Communist regime	1	0
1956	1956	East Germany		Communist regime	1	0
1989	1989	East Germany	pro-dem movement	Communist regime	1	1
1956	1956	Poland		Communist regime	1	0
1970	1968	Poland		Communist regime	1	0

1989	1981	Poland	Solidarity	Communist regime	1	1
1934	1934	Austria	socialists	Dolfuss government	0	0
1920	1919	Hungary	anti-communist movement (Whites)	communist regime	0	1
1989	1989	Hungary	pro-dem movement	Communist regime	1	1
1989	1989	Czechoslovakia	Velvet Revolution	Communist regime	1	1
1992	1989	Slovakia	People Against Violence	Czech communist government	1	1
1989	1989	Albania		Communist regime	1	0
2000	1999	Croatia		semi-presidential system	1	1
1968	1968	Yugoslavia	student protests	Communist regime	1	0
2000	2000	Yugoslavia		Milosevic regime	1	1
1990	1989	Slovenia		Communist regime	1	1
1963	1963	Greece		Karamanlis regime	1	1
1974	1974	Greece		Military rule	1	1
1923	1923	Bulgaria	Agrarian League movement	Military regime	0	0
1989	1989	Bulgaria		Communist regime	1	1
1989	1987	Romania		Ceasescu regime	1	0
1989	1989	Romania	anti-Ceasescu rebels	Ceacescu regime	0	1
1953	1946	USSR	Ukrainian rebellion	Communist regime	0	0
1906	1905	Russia	Peasant/Worker Rebellion	Romanov dynasty	0	0
1917	1916	Russia	Kirghiz and Kazables rebels	Romanov regime	0	1
1921	1917	Russia	anti-Bolsheviks	Bolshevik regime	0	0
1989	1989	Estonia	Singing Revolution	Communist regime	1	1
1989	1989	Latvia	pro-dem movement	Lithuanian regime	1	1
1991	1989	Lithuania	pro-democracy movement/Sajudis	Kuchma regime	1	1
2004	2001	Ukraine	Orange Revolution	Communist regime	1	0
1989	1989	Belarus		Belarus government	1	0
2006	2006	Belarus		Shevardnadze regime	1	1
2003	2003	Georgia	Rose Revolution			

1918	1918	Finland	communist rebels	Finnish government	0	0
1994	1989	Mali	Tauregs	Mali regime	0	0
1992	1989	Mali		Military rule	1	1
2000	2000	Senegal		Diouf govt	1	0
1990	1989	Benin		Communist regime	1	0
1992	1991	Niger		Military rule	1	0
2005	2002	Ivory Coast	PMIC	incumbent regime	0	0
1990	1989	Liberia	anti-Doe rebels	Doe regime	0	1
1995	1992	Liberia	NPFL & ULIMO	Johnson regime	0	0
1996	1996	Liberia	national patriotic forces	Liberian govt	0	0
2003	2003	Liberia	LURD	Taylor regime	0	1
1996	1991	Sierra Leone	RUF	Republcan government	0	0
2000	2000	Ghana		Rawlings govt	1	1
1984	1980	Nigeria	Muslim fundamentalists	Nigerian govt	0	0
1999	1993	Nigeria		Military rule	1	1
1997	1994	CAR	multiple factions	CAR regime	0	1
1990	1966	Chad	Frolinat	Chadian government	0	1
1998	1994	Chad	rebels	Chadian regime	0	0
1999	1997	Congo-Brazzaville (ROC)	Denis Sassou Nguesso	Lissouba regime	0	1
1978	1977	Zaire/DRC	FLNC	DRC/Zaire regime	0	0
1993	1993	Zaire/DRC	rebels (People's Revolutionary Party)	Mobutu regime	0	0
1997	1996	Zaire/DRC	Kabila-ADFL	Mobutu regime	0	1
1988	1980	Uganda	National Resistance Army	Okello regime	0	1
2006	1996	Uganda	LRA	Museveni government	0	0
1989	1989	Kenya		Daniel Arap Moi	1	0
1995	1992	Tanzania	pro-democracy movement	Mwinyi regime	1	0
2002	1972	Burundi	Hutu rebellion	Tutsi hegemony in government	0	0
1992	1991	Burundi	Tutsi supremacists	Hutu regime	0	0

1964	1963	Rwanda	Watusi	Hutu regime	0	0
1993	1990	Rwanda	Tutsi rebels	Hutu regime	0	0
1994	1994	Rwanda	Patriotic Front	Hutu regime and genocide	0	1
1997	1982	Somalia	clan factions; SNM	Siad Barre regime	0	1
1994	1991	Djibouti	Afar insurgency	Djibouti regime	0	0
2001	1975	Angola	UNITA	Angolan government	0	0
1992	1979	Mozambique	Renamo	Mozambique government	0	0
1991	1990	Zambia		One-party rule	1	0
2001	2001	Zambia		Chiluba regime	1	1
1980	1974	Zimbabwe	Zimbabwe African People's Union	Smith/Muzorewa regime	0	1
1987	1983	Zimbabwe	PF-ZAPU guerrillas	Mugabe regime	0	0
1994	1992	Malawi		Banda regime	1	1
1993	1991	Madagascar	Active Voices	Didier Ratsiraka	1	1
2003	2002	Madagascar	pro-democracy movement	Ratsiraka regime	1	1
1963	1962	Algeria	former rebel leaders	Ben Bella regime	0	0
2006	1992	Algeria	Islamic Salvation Front	Algerian government	0	0
2005	1983	Sudan	SPLA-Gatang faction	Sudanese government	0	0
1985	1985	Sudan		Jaafar Nimiery	1	1
1909	1908	Iran	Constitutionalists	shah regime	0	1
1979	1977	Iran	Iranian Revolution	Shah Reza Pahlavi	1	1
1996	1979	Iran	KDPI	Iranian regime	0	0
1982	1981	Iran	Mujahideen	Khomenei regime	0	0
1927	1924	Turkey	Sheikh Said insurgency	Kemal regime	0	0
1959	1959	Iraq	Shammar Tribe and pro-Western officers	Qassim regime	0	0
1991	1991	Iraq	Shiite rebellion	Hussein regime	0	0
2006	2003	Iraq	Iraqi insurgency	Iraqi government	0	0
2005	2000	Egypt	Kifaya	Mubarak regime	1	0
1982	1980	Syria	Muslim Brotherhood	Syrian regime	0	0

1958	1958	Lebanon	leftists	Shamun regime	0	0
1975	1975	Lebanon	leftists	Lebanese government	0	0
1930	1929	Saudi Arabia	Ikhwan rebellion	Ibn Saud regime	0	0
1948	1948	Yemen Arab Republic	Yahya Family revolt	counter coup	0	1
1957	1955	Yemen	Yemeni insurgency	British and Aden administration	0	0
1969	1962	Yemen Arab Republic	Royalists	al-Sallal regime	0	0
1986	1986	Yemen People's Republic	leftists	Ali Nasir regime	0	0
1976	1964	Oman	Popular Front for the Liberation of Oman	Oman government	0	0
1929	1924	Afghanistan	Anti-Reformist movement	Amanullah Khan regime	0	0
1979	1978	Afghanistan	Afghans	Afghan government	0	1
1996	1992	Afghanistan	Taliban	Afghan regime	0	1
2006	2001	Afghanistan	Taliban	Afghan government	0	0
1997	1992	Tajikistan	Popular Democratic Army (UTO)	Rakhmanov regime	0	0
1989	1989	Kyrgyzstan	Kyrgyzstan Democratic Movement	Communist regime	1	1
2005	2005	Kyrgyzstan	Tulip Revolution	Akayev regime	1	1
1913	1911	China	Republicans	Imperial regime; military dictatorship	0	0
1949	1922	China	Chinese communist movement	Kuomintang regime	0	1
1957	1956	China	Hundred Flowers Movement	Communist regime	1	0
1979	1976	China	Democracy Movement	Communist regime	1	0
1989	1989	China		Communist regime	1	0
1990	1989	Mongolia		Communist regime	1	0
1985	1979	Taiwan		autocratic regime	1	0
1960	1960	South Korea	Student Revolution	Rhee regime	1	1
1980	1979	South Korea		military junta	1	0
1987	1987	South Korea		Mil govt	1	0
1922	1921	India	Moplah Rebellion	local Hindu leaders	0	0
1971	1967	India	Naxalite rebellion	Indian regime	0	0
1969	1968	Pakistan		Khan regime	1	0

1983	1983	Pakistan	pro-dem movement	Zia al-Huq	1	0
1995	1994	Pakistan	Mohajir	Pakistani government	0	0
1990	1989	Bangladesh		Military rule	1	0
1988	1988	Burma	pro-dem movement	military junta	1	0
1971	1971	Sri Lanka	JVP	Sri Lankan government	0	0
1990	1989	Nepal	The Stir	Monarchy/Panchayat regime	1	0
2006	1996	Nepal	CPN-M/UPF	Nepalese government	0	0
2006	2006	Nepal		Nepalese govt; martial law	1	0
1973	1970	Thailand	communist rebels	Thai government	0	0
1973	1973	Thailand	student protests	military dictatorship	1	1
1992	1992	Thailand	pro-dem movement	Suchinda regime	1	0
2006	2005	Thailand		Thaksin regime	1	1
1975	1970	Cambodia	Khmer Rouge	Cambodian government	0	1
1997	1978	Cambodia	Khmer Rouge	Cambodian government	0	0
1975	1960	Laos	Pathet Lao	Laotian government	0	1
1975	1958	Vietnam	North Vietnam (National Liberation Front)	government of South Vietnam	0	1
1954	1946	Philippines	Hukbalahap Rebellion	Filipino government	0	0
1980	1970	Philippines	Moro Islamic Liberation Front	Filipino government	0	0
2006	1972	Philippines	New People's Army	Filipino government	0	0
1983	1986	Philippines	People Power	Ferdinand Marcos	1	1
2001	2001	Philippines	Second People Power Movement	Estrada regime	1	1
1953	1953	Indonesia	Darul Islam	Indonesian government	0	0
1960	1956	Indonesia	leftists	Sukarno regime	0	0
1998	1997	Indonesia		Suharto rule	1	1
1998	1988	Papua New Guinea	Bougainville Revolt	Papuan regime	0	0

Campaigns aiming for "other goals"

eyear	byear	location	campaign	target	nonviol	success
1972	1966	Guatemala	Indian resistance	government of Guatemala	0	0
1986	1986	Argentina		attempted coup	1	1
1962	1958	France	Secret Army Organization	French withdrawal from Algeria	0	0
1907	1907	Romania	Peasant Rebellion	land dist system	0	0
1935	1930	USSR	Central Asian rebels	Soviet policies of conscription	0	0
1991	1990	Russia	pro-dem movement	Anti-coup	1	1
1961	1952	South Africa	Defiance Campaign	Apartheid	1	0
1994	1984	South Africa	Anti-Apartheid	Apartheid	1	1
2006	2003	Sudan	JEM/SLA	Janjaweed militia	0	0
1970	1970	Jordan	Palestinian activists	Jordanian rule	0	0
1968	1967	China	Red Guard	anti-Maoists	0	0

12 Bibliography

- Abrams, Dominic and Michael A Hogg. 2006. *Social identifications: A social psychology of intergroup relations and group processes*. Routledge.
- Acemoglu, Daron. 2005. “Constitutions, Politics and Economics: A Review Essay on Persson and Tabellini’s “The Economic Effects of Constitutions”.” *Journal of Economic Literature* 43(4):1025–1048.
- Acemoglu, Daron, Francisco Gallego and James A. Robinson. 2014. “Institutions, Human Capital and Development.” *Annual Reviews of Economics* 6:875–912.
- Acemoglu, Daron and James A. Robinson. 2001. “A Theory of Political Transitions.” *American Economic Review* 91(4):938–963.
- Acemoglu, Daron and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.
- Acemoglu, Daron and James Robinson. 2013. “Middle class rising?”. Post at the “Why nations fail?” blog, 2nd of July.
URL: <http://whynationsfail.com/blog/?currentPage=18>
- Acemoglu, Daron, Simon Johnson and James A. Robinson. 2001. “The Colonial Origins of Comparative Development: An Empirical Investigation.” *American Economic Review* 91(5):1369–1401.
- Acemoglu, Daron, Simon Johnson, James A. Robinson and Pierre Yared. 2005. “From Education to Democracy.” *The American Economic Review* 95(2):44–49.

- Acemoglu, Daron, Simon Johnson, James A. Robinson and Pierre Yared. 2008. "Income and Democracy." *American Economic Review* 98(3):808–842.
- Acemoglu, Daron, Suresh Naidu, Pascual Restrepo and James A. Robinson. 2015. Democracy, Redistribution and Inequality. Vol. 2 of *Handbook of Income Distribution* Elsevier pp. 1885 – 1966.
- Achen, Christopher H. 2002. "Toward a new political methodology: Microfoundations and ART." *Annual Review of Political Science* 5(1):423–450.
- Achen, Christopher H. 2005. "Let's Put Garbage-Can Regressions and Garbage-Can Probits Where They Belong." *Conflict Management and Peace Science* 22(4):327–339.
- Ackerman, Peter and Jack DuVall. 2000. *A force more powerful: A century of non-violent conflict*. New York: Palgrave Macmillan.
- Afrobarometer. 2008. "Afrobarometer Data Codebook for Round 3 Afrobarometer Surveys in 18 African Countries." Ann Arbor: Michigan State University.
- A'Hearn, Brian, Jörg Baten and Dorothee Crayen. 2009. "Quantifying quantitative literacy: Age heaping and the history of human capital." *The Journal of Economic History* 69(3):783–808.
- Albaugh, Ericka A. 2014. *State-building and multilingual education in Africa*. Cambridge: Cambridge University Press.
- Albertus, Michael and Victor Menaldo. 2012. "If you're against them you're with us: The effect of expropriation on autocratic survival." *Comparative Political Studies* 45(8):973–1003.
- Alesina, Alberto, Arnaud Devlaeminck, William Easterly, Sergio Kurlat and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8(2):155–194.
- Alesina, Alberto and Bryony Reich. 2013. "Nation Building." National Bureau of Economic Research (NBER) Working Paper (NBER) Working Paper.
- Alesina, Alberto and Paola Giuliano. 2011. "Family ties and political participation." *Journal of the European Economic Association* 9(5):817–839.
- Almond, Gabriel and Sidney Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton: Princeton University Press.
- Altinok, Nadir and Hatidje Murseli. 2007. "International database on human capital quality." *Economics Letters* 96(2):237–244.

- Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press.
- Angrist, Joshua D and Jörn-Steffen Pischke. 2010. "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics." *The Journal of Economic Perspectives* 24(2):3–30.
- Angrist, Noam; Patrinos, Harry Anthony; Schlotter Martin. 2013. "An expansion of a global data set on educational quality: a focus on achievement in developing countries." *Policy Research working paper* WPS 6536.
- Ansell, Ben. 2008. "Traders, teachers, and tyrants: democracy, globalization, and public investment in education." *International Organization* 62(2):289–322.
- Ansell, Ben and David Samuels. 2010. "Inequality and Democratization: A Contractarian Approach." *Comparative Political Studies* 45(12):1543–1574.
- Ansell, Ben and David Samuels. 2014. *Inequality and democratization: an elite-competition approach*. Cambridge: Cambridge University Press.
- Aristotle. 2000. *Politics*. New York: Dover Publications.
- Arjona, Ana and Stathis Kalyvas. 2006. "Preliminary Results of a Survey of Demobilized Combatants in Colombia." Unpublished working paper. New Haven: Yale University.
- Bahgat, Gawdat. 1999. "Education in the Gulf monarchies: Retrospect and prospect." *International Review of Education* 45(2):127–136.
- Balzer, Harley. 1993. Science, Technology and Education in the former U.S.S.R. In *The Former Soviet Union in Transition*, ed. Richard Kaufman and John P. Hardt. Armonk: M.E. Sharpe pp. 889–908.
- Barakat, Bilal and Henrik Urdal. 2009. "Breaking the Waves? Does Education Mediate the Relationship Between Youth Bulges and Political Violence?" World Bank Policy Research Working Paper 5114. Washington, DC: The World Bank.
- Barro, Robert J. 1999. "Determinants of Democracy." *Jounal of Political Economy* 107(S6):158–183.
- Barro, Robert J and Jong Wha Lee. 2013. "A new data set of educational attainment in the world, 1950–2010." *Journal of development economics* 104:184–198.

- Baten, Joerg and Dácil Juif. 2014. “A story of large landowners and math skills: Inequality and human capital formation in long-run development, 1820–2000.” *Journal of comparative economics* 42(2):375–401.
- Bates, Robert H. 1981. *States and Markets in Tropical Africa: The Political Basis of Agricultural Policy*. Series on Social Choice and Political Economy Berkeley: University of California Press.
- Baum, Matthew A. and David A. Lake. 2003. “The Political Economy of Growth: Democracy and Human Capital.” *American Journal of Political Science* 47(2):333–347.
- Beck, Nathaniel, David Epstein, Simon Jackman and Sharyn O’Halloran. 2001. Alternative Models of Dynamics in Binary Time-Series-Cross-Section Models: The Example of State Failure. In *Annual Meeting of the Society for Political Methodology*. Atlanta: Emory University: .
- Beck, Nathaniel and Jonathan N. Katz. 1995. “What to Do (and Not to Do) with Time-Series Cross-Section Data.” *American Political Science Review* 89(3):634–647.
- Beetham, David. 1992. “Liberal democracy and the limits of democratization.” *Political Studies* 40(s1):40–53.
- Benavot, Aaron. 1996. “Education and political democratization: Cross-national and longitudinal findings.” *Comparative education review* 40(4):377–403.
- Benos, Nikos and Stefania Zoumou. 2014. “Education and economic growth: A meta-regression analysis.” *World Development* 64:669–689.
- Berman, Eli, Michael Callen, Joseph H Felter and Jacob N Shapiro. 2011. “Do working men rebel? Insurgency and unemployment in Afghanistan, Iraq, and the Philippines.” *Journal of Conflict Resolution* 55(4):496–528.
- Bermeo, Nancy. 1997. “Myths of moderation: confrontation and conflict during democratic transitions.” *Comparative Politics* 29(3):305–322.
- Bils, Mark and Peter J. Klenow. 2000. “Does Schooling Cause Growth?” *The American Economic Review* 90(5):1160–1183.
- Bleck, Jaimie. 2015. *Education and Empowered Citizenship in Mali*. Baltimore: John Hopkins University Press.
- Blundell, Richard and Stephen Bond. 1998. “Initial conditions and moment restrictions in dynamic panel data models.” *Journal of econometrics* 87(1):115–143.

- Blundell, Richard and Stephen Bond. 2000. “GMM estimation with persistent panel data: an application to production functions.” *Econometric reviews* 19(3):321–340.
- Bobba, Matteo and Decio Coviello. 2007. “Weak instruments and weak identification, in estimating the effects of education, on democracy.” *Economics Letters* 96(3):301–306.
- Boesman, William. 1993. Science and Technology in the Former Soviet Union: Capabilities and Needs. In *The Former Soviet Union in Transition*, ed. Richard Kaufman and John P. Hardt. Armonk: M.E. Sharpe pp. 610–628.
- Boix, Carles. 2003. *Democracy and Redistribution*. Cambridge: Cambridge University Press.
- Boix, Carles. 2011. “Democracy, Development, and the International System.” *American Political Science Review* 105(4):809–828.
- Boix, Carles, Michael Miller and Sebastian Rosato. 2013. “A complete data set of political regimes, 1800–2007.” *Comparative Political Studies* 46(12):1523–1554.
- Boudreau, Vincent. 2009. *Resisting dictatorship: Repression and protest in Southeast Asia*. Cambridge: Cambridge University Press.
- Brady, Henry E. 2008a. *Causation and explanation in social science*. Vol. The Oxford Handbook of Political Methodology Oxford: Oxford University Press.
- Brady, Henry E. 2008b. “Causation and explanation in social science.” *Oxford handbook of political methodology* pp. 530–569.
- Bratton, Michael and Nicholas Van de Walle. 1997. *Democratic experiments in Africa: Regime transitions in comparative perspective*. Cambridge: Cambridge University Press.
- Bratton, Michael, Robert B Mattes and Emmanuel Gyimah-Boadi. 2005. *Public opinion, democracy, and market reform in Africa*. Cambridge: Cambridge University Press.
- Brown, David S. 1999. “Reading, writing and regime type: Democracy’s impact on primary school enrollment.” *Political Research Quarterly* 52(4):681–707.
- Brown, David S. and Wendy A. Hunter. 2004. “Democracy and Human Capital Formation: Educational Expenditures in Latin America.” *Comparative Political Studies* 37(7):842–864.
- Bueno de Mesquita, Bruce and Alastair Smith. 2009. “Political Survival and Endogenous Institutional Change.” *Comparative Political Studies* 42(2):167–197.

- Bueno de Mesquita, Bruce and Alastair Smith. 2010. “Leader Survival, Revolutions, and the Nature of Government Finance.” *American Journal of Political Science* 54(4):936–950.
- Bueno de Mesquita, Bruce and Alistair Smith. 2003. *The Logic of Political Survival*. New York: MIT Press.
- Bueno de Mesquita, Bruce and Alistair Smith. 2012. *The Dictator’s Handbook. Why Bad Behavior is Almost Always Good Politics*. New York: PublicAffairs.
- Bueno de Mesquita, Bruce, James D. Morrow, Randolph M. Siverson and Alastair Smith. 2002. “Political Institutions, Policy Choice and the Survival of Leaders.” *British Journal of Political Science* 32(4):559 – 590.
- Buhaug, Halvard, Kristian Skrede Gleditsch, Helge Holtermann, Gudrun Østby and Andreas Forø Tollefsen. 2011. “Itâs the local economy, stupid! Geographic wealth dispersion and conflict outbreak location.” *Journal of Conflict Resolution* 55(5):814–840.
- Butcher, Charles and Isak Svensson. 2014. “Manufacturing Dissent Modernization and the Onset of Major Nonviolent Resistance Campaigns.” *Journal of Conflict Resolution*. 0022002714541843.
- Cagé, Julia and Valeria Rueda. 2016. “The long-term effects of the printing press in Sub-Saharan Africa.” *American Economic Journal: Applied Economics* 8(3):69–99.
- Campante, Filipe and Edward L Glaeser. 2009. “Yet another tale of two cities: Buenos Aires and Chicago.”
- Campante, Filipe R and Davin Chor. 2012a. “Schooling, political participation, and the economy.” *Review of Economics and Statistics* 94(4):841–859.
- Campante, Filipe R and Davin Chor. 2012b. “Why was the Arab world poised for revolution? Schooling, economic opportunities, and the Arab Spring.” *The Journal of Economic Perspectives* 26(2):167–187.
- Card, David. 1999. “The causal effect of education on earnings.” *Handbook of labor economics* 3:1801–1863.
- Carter, David B. and Curtis S. Signorino. 2010. “Back to the Future: Modeling Time Dependence in Binary Data.” *Political Analysis* 18(3):271–292.
- Castelló-Climent, Amparo. 2008. “On the distribution of education and democracy.” *Journal of Development Economics* 87(2):179–190.

- Cederman, Lars-Erik, Andreas Wimmer and Brian Min. 2010. "Why do ethnic groups rebel? New data and analysis." *World Politics* 62(1):87–119.
- Cederman, Lars-Erik and Kristian Skrede Gleditsch. 2009. "Introduction to Special Issue on "Disaggregating Civil War"." *Journal of Conflict Resolution* 53(4):487–495.
- Cederman, Lars-Erik, Kristian Skrede Gleditsch and Halvard Buhaug. 2013. *Inequality, grievances, and civil war*. Cambridge: Cambridge University Press.
- Cederman, Lars-Erik, Kristian Skrede Gleditsch and Simon Hug. 2013. "Elections and Ethnic Civil War'." *Comparative Political Studies* 46(3):387–417.
- Cederman, Lars-Erik, Nils B Weidmann and Nils-Christian Bormann. 2015. "Triangulating horizontal inequality Toward improved conflict analysis." *Journal of Peace Research* 52(6):806–821.
- Celestino, Mauricio Rivera and Kristian Skrede Gleditsch. 2013. "Fresh carnations or all thorn, no rose? Nonviolent campaigns and transitions in autocracies." *Journal of Peace Research* 50(3):385–400.
- Charron, Nicholas and Victor Lapuente. 2010. "Does democracy produce quality of government?" *European Journal of Political Research* 49(4):443–470.
- Che, Yi, Yi Lu, Zhigang Tao and Peng Wang. 2013. "The impact of income on democracy revisited." *Journal of Comparative Economics* 41(1):159–169.
- Cheibub, Jose, Jennifer Gandhi and James Vreeland. 2010. "Democracy and dictatorship revisited." *Public Choice* 143(1–2):67–101.
- Chengze, Simon Fan, Jody Overland and Michael Spagat. 1999. "Human Capital, Growth, and Inequality in Russia." *Journal of Comparative Economics* 27(4):618–643.
- Chenoweth, Erica and Jay Ulfelder. 2017. "Can Structural Conditions Explain the Onset of Nonviolent Uprisings?" *Journal of Conflict Resolution* 61(2):298–324.
- Chenoweth, Erica and Kathleen Gallagher Cunningham. 2013. "Understanding nonviolent resistance. An introduction." *Journal of Peace Research* 50(3):271–276.
- Chenoweth, Erica and Maria J. Stephan. 2011. *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*. New York: Columbia University Press.
- Chenoweth, Erica and Orion A. Lewis. 2013. "Unpacking nonviolent campaigns. Introducing the NAVCO 2.0 dataset." *Journal of Peace Research* 50(3):415–423.

- Cho, Wendy K Tam, Charles F Manski, Henry Brady et al. 2008. "Cross level/ecological inference." *Oxford handbook of political methodology* pp. 530–569.
- Cincotta, Richard P. 2008. "Half a Chance: Youth Bulges and Transitions to Liberal Democracy." *New Directions in Demographic Security. ECSP Report* 13:10–18.
- Clark, Damon and Heather Roayer. 2013. "The effect of education on adult mortality and health: Evidence from Britain." *The American Economic Review* 103(6):2087–2120.
- Clark, Tom S and Drew A Linzer. 2015. "Should I use fixed or random effects?" *Political Science Research and Methods* 3(2):399–408.
- Clark, William Roberts, Matt Golder and Sona Nadenichek Golder. 2012. *Principles of comparative politics*. Thousand Oaks: CQ Press.
- Clarke, Christopher. 2006. *Iron Kingdom: The Rise and Downfall of Prussia, 1600-1947*. London: Penguin Books.
- Cohen, Daniel and Marcelo Soto. 2007. "Growth and human capital: good data, good results." *Journal of economic growth* 12(1):51–76.
- Colclough, Christopher and Samer Al-Samarrai. 2000. "Achieving schooling for all: Budgetary expenditures on education in sub-Saharan Africa and South Asia." *World Development* 28(11):1927–1944.
- Coleman, James S. 1986. "Social theory, social research, and a theory of action." *American journal of Sociology* 91(6):1309–1335.
- Collier, Paul and Anke Hoeffler. 2004. "Greed and Grievance in Civil War." *Oxford Economic Papers* 56(4):563–595.
- Collier, Paul and Anke Hoeffler. 2005. "Resource Rents, Governance and Conflict." *Journal of Conflict Resolution* 49(4).
- Collier, Paul, Anke Hoeffler and Dominic Rohner. 2009. "Beyond greed and grievance: feasibility and civil war." *oxford Economic papers* 61(1):1–27.
- Collier, Paul, Anke Hoeffler and Måns Söderbom. 2004. "On the Duration of Civil War." *Journal of Peace Research* 41(3):253–273.
- Collier, Paul, Lani Elliot, Håvard Hegre, Anke Hoeffler, Marta Reynal-Querol and Nicholas Sambanis. 2003. *Breaking the Conflict Trap. Civil War and Development Policy*. Oxford: Oxford University Press.

- Collier, Ruth Berins. 1999. *Paths Towards Democracy: The Working Class and Elites in Western Europe and South America*. Cambridge: Cambridge University Press.
- Cooper, Harris, Jorgianne Civey Robinson and Erika A. Patall. 2006. "Does Homework Improve Academic Achievement? A Synthesis of Research, 1987-2003." *Review of Educational Research* 76(1):1–62.
- Coppedge, Michael. 2012. *Democratization and research methods*. Cambridge: Cambridge University Press.
- Coppedge, Michael, John Gerring, Staffan Lindberg, Jan Teorell, David Altman, Michael Bernhard, Steven Fish, Adam Glynn, Allen Hicken and Matthew Kroenig. 2015. "Varieties of Democracy (V-Dem) Codebook." Gothenburg: Varieties of Democracy Institute.
- Crayen, Dorothee and Joerg Baten. 2010. "Global trends in numeracy 1820–1949 and its implications for long-term growth." *Explorations in Economic History* 47(1):82–99.
- Croke, Kevin, Guy Grossman, Horacio A Larreguy and John Marshall. 2016. "Deliberate disengagement: how education can decrease political participation in electoral authoritarian regimes." *The American Political Science Review* 110(3):579.
- Cunningham, Kathleen Gallagher. 2013a. "Actor Fragmentation and Civil War Bargaining: How Internal Divisions Generate Civil Conflict." *American Journal of Political Science* 57(3):659–672.
- Cunningham, Kathleen Gallagher. 2013b. "Understanding strategic choice : The determinants of civil war and nonviolent campaign in self-determination disputes." *Journal of Peace Research* 50(3):291–304.
- Cunningham, Philip J. 2014. *Tiananmen moon: Inside the Chinese student uprising of 1989*. Lanham: Rowman & Littlefield.
- Dahl, Marianne, Scott Gates, Kristian Skrede Gleditsch and Belén González. 2013. "Apples and Oranges? Actor Profiles and The Choice of Violent and Non-Violent Tactics.".
- URL:** <http://forms2.la.psu.edu/pss/pss-presentations/posters/dahl-and-gonzales-apples-and-oranges-actor-profiles-and-violent-and-non-violent-tactics/view>
- Dahl, Robert A. 1971. *Polyarchy: Political Participation and Opposition*. New Haven: Yale University Press.

Dahlum, Sirianne and Carl Henrik Knutsen. 2016. "Democracy by demand? Reinvestigating the effect of self-expression values on political regime type." *British Journal of Political Science* Online first.

URL: DOI: <https://doi.org/10.1017/S0007123415000447>

Davies, James C. 1962. "Towards a Theory of Revolution." *American Sociological Review* 27(1):5–19.

Dee, Thomas S. 2004. "Are there civic returns to education?" *Journal of Public Economics* 88(9):1697–1720.

Denny, Elaine K and Barbara F Walter. 2014. "Ethnicity and civil war." *Journal of Peace Research* 51(2):199–212.

Desimone, Laura. 1999. "Linking Parent Involvement With Student Achievement: Do Race and Income Matter?" *The Journal of Educational Research* 93(1):11–30.

Dewey, John. 1938. *Education and experience*. New York: Macmillan.

Di Palma, Giuseppe. 1990. *To craft democracies: An essay on democratic transitions*. Berkeley and Los Angeles: Univ of California Press.

Diamond, Larry. 2012. "Why East Asia – Including China – Will Turn Democratic Within a Generation." *The Atlantic*, January 24th.

URL: <https://www.theatlantic.com/international/archive/2012/01/why-east-asia-including-china-will-turn-democratic-within-a-generation/251824>

Diamond, Larry. 2015. *In Search of Democracy*. New York: Routledge.

Doucouliagos, Hristos and Mehmet Ali Ulubasoglu. 2008. "Democracy and Economic Growth: A Meta-Analysis." *American Journal of Political Science* 52(1):61–83.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.

Easterlin, Richard A. 1981. "Why isn't the whole world developed?" *The Journal of Economic History* 41(01):1–17.

Economist, The. 2015. "Argentina's inflation problem. The price of cooking the books." February 25th.

URL: <http://www.economist.com/node/21548229>

El-Said, Hamed and James E Rauch. 2015. "Education, Political Participation, and Islamist Parties: The Case of Jordan's Islamic Action Front." *The Middle East Journal* 69(1):51–73.

- Elster, Jon. 2009. "Excessive Ambitions." *Capitalism and Society* 4(2):1–30.
- Engerman, Stanley L., Elisa V. Mariscal and Kenneth L. Sokoloff. 2009. The Evolution of Schooling in the Americas, 1800-1925. In *Human Capital and Institutions: The Long Run View*, ed. David Eltis, Frank D. Lewis and Kenneth L. Sokoloff. Cambridge: Cambridge University Press pp. 93–142.
- Engerman, Stanley L. and Kenneth L. Sokoloff. 2005. "The Evolution of Suffrage Institutions in the New World." *Journal of Economic History* 65(4):891–921.
- Epstein, David L., Robert Bates, Jack Goldstone, Ida Kristensen and Sharyn O'Halloran. 2006. "Democratic Transitions." *American Journal of Political Science* 50(3):551–569.
- Faria, Hugo J, Hugo M Montesinos-Yufa and Daniel R Morales. 2014. "Should the Modernization Hypothesis Survive Acemoglu, Johnson, Robinson, and Yared? Some More Evidence." *Scholarly Comments on Academic Economics* 11(1):17–36.
- Fearon, James D. and David D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.
- Fjelde, Hanne. 2010. "Generals, Dictators, and Kings: Authoritarian Regimes and Civil Conflict, 1973–2004." *Conflict Management and Peace Science* 27(3):195–218.
- Fjelde, Hanne and Gudrun Østby. 2014. "Socioeconomic inequality and communal conflict: A disaggregated analysis of sub-Saharan Africa, 1990–2008." *International Interactions* 40(5):737–762.
- Freeman, Richard. 2004. "What, Me Vote?" *Social Inequality, Vol 1*.
- Friedman, Willa, Michael Kremer, Edward Miguel and Rebecca Thornton. 2011. "Education as Liberation?"
- Fuchs, Thomas and Ludger Woessmann. 2008. What accounts for international differences in student performance? A re-examination using PISA data. In *The Economics of Education and Training. Studies in Empirical Economics*, ed. Christian Dustmann, Bernd Fitzenberger and Stephen Machin. Heidelberg: Springer pp. 209–240.
- Fullan, Michael. 2000. "The three stories of education reform." *Phi Delta Kappan* 81(8):581–584.
- Gallego, Francisco A. 2010. "Historical Origins of Schooling: The Role of Democracy and Political Decentralization." *Review of Economics and Statistics* 92(2):228–243.

- Galston, William A. 2001. "Political knowledge, political engagement, and civic education." *Annual review of political science* 4(1):217–234.
- Gandhi, Jennifer. 2008. *Political Institutions under Dictatorship*. Cambridge: Cambridge University Press.
- Gandhi, Jennifer and Ellen Lust-Okar. 2009. "Elections under Authoritarianism." *Annual Review of Political Science* 12:403–422.
- Garcia, Blake E and Cameron Wimpy. 2016. "Does Information Lead to Emulation? Spatial Dependence in Anti-Government Violence." *Political Science Research and Methods* 4(1):27–46.
- Gat, Azar. 2008. *War in human civilization*. Oxford University Press.
- Gates, Scott, Håvard Hegre, Mark P. Jones and Håvard Strand. 2006. "Institutional Inconsistency and Political Instability: Polity Duration, 1800–2000." *American Journal of Political Science* 50(4):893–908.
- Geddes, Barbara. 1999. "What Do We Know About Democratization After Twenty Years?" *Annual Review of Political Science* 2:115–144.
- Geddes, Barbara, Joseph Wright and Erica Frantz. 2014. "Autocratic breakdown and regime transitions: A new data set." *Perspectives on Politics* 12(2):313–331.
- Geddes, Barbara, Joseph Wright and Erica Franz. 2012. "Autocratic Regimes Code Book, Version 1.0."
- Gelman, Andrew. 2009. *Red state, blue state, rich state, poor state: Why Americans vote the way they do*. Princeton: Princeton University Press.
- George, Alexander L and Andrew Bennett. 2005. *Case studies and theory development in the social sciences*. Cambridge: Mit Press.
- Gerring, John. 2005. "Causation A Unified Framework for the Social Sciences." *Journal of Theoretical Politics* 17(2):163–198.
- Gerring, John. 2011. *Social science methodology: A criterial framework*. Cambridge: Cambridge University Press.
- Glaeser, Edward L, Giacomo AM Ponzetto and Andrei Shleifer. 2007. "Why does democracy need education?" *Journal of economic growth* 12(2):77–99.

- Glaeser, Edward L, Rafael La Porta, Florencio Lopez-de Silanes and Andrei Schleifer. 2004. "Do institutions cause growth?" *Journal of Economic Growth* 9(3):271–303.
- Gleditsch, Kristian S. 2002. "Expanded Trade and GDP data." *Journal of Conflict Resolution* 46(5):712–724.
- Gleditsch, Kristian Skrede and Andrea Ruggeri. 2010. "Political opportunity structures, democracy, and civil war." *Journal of Peace Research* 47(3):299–310.
- Glewwe, Paul, Eugenie Maiga and Haochi Zheng. 2014. "The contribution of education to economic growth: A review of the evidence, with special attention and an application to Sub-Saharan Africa." *World Development* 59:379–393.
- Goemans, Hein E. 2000. *War and Punishment*. Princeton: Princeton University Press.
- Goldin, Claudia Dale and Lawrence F Katz. 2009. *The race between education and technology*. Cambridge and London: Harvard University Press.
- Goldstone, Jack A. 1994. "Is revolution individually rational? Groups and individuals in revolutionary collective action." *Rationality and Society* 6(1):139–166.
- Golebiowska, Ewa A. 1995. "Individual value priorities, education, and political tolerance." *Political Behavior* 17(1):23–48.
- Granovetter, Mark. 1983. "The strength of weak ties: A network theory revisited." *Sociological Theory* 1(1):201–233.
- Granovetter, Mark S. 1973. "The strength of weak ties." *American Journal of Sociology* pp. 1360–1380.
- Green, Andy. 2013. "Education and state formation: Europe, East Asia and the USA.".
- Green, Donald P., Soo Yeon Kim and David Yoon. 2001. "Dirty Pool." *International Organization* 55(2):441–468.
- Gundlach, Erich and Martin Paldam. 2009. "A farewell to critical junctures: Sorting out long-run causality of income and democracy." *European Journal of Political Economy* 25(3):340–354.
- Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton: Princeton University Press.
- Gurr, Ted Robert. 2000. *People Versus States, Minorities at Risk in the New Century*. Washington D.C.: United States Institute of Peace Press.

- Gurr, Ted Robert and Raymond Duvall. 1973. "Civil Conflict in the 1960s: A Reciprocal Theoretical System with Parameter Estimates." *Comparative Political Studies* 6(2):135–170.
- Gylfason, Thorvaldur. 2001. "Natural resources, education, and economic development." *European economic review* 45(4):847–859.
- Haber, Stephen. 2006. *Authoritarian Government*. Vol. The Oxford Handbook of Political Economy Oxford: Oxford University Press.
- Hadenius, Axel and Jan Teorell. 2005. "Cultural and Economic Prerequisites of Democracy: Reassessing Recent Evidence." *Studies in Comparative International Development* 39(4):87–106.
- Hadenius, Axel and Jan Teorell. 2006. "Authoritarian Regimes: Stability, Change, and Pathways to Democracy, 1972-2003." . Lund University: Working Paper no. 331.
- Hadenius, Axel and Jan Teorell. 2007. "Pathways from Authoritarianism." *Journal of Democracy* 18(1):143 –156.
- Haggard, Stephan and Robert R Kaufman. 2008. *Development, Democracy, and Welfare States: Latin America, East Asia, and Eastern Europe*. Princeton: Princeton University Press.
- Hanushek, Eric A. 2002. "Publicly provided education." *Handbook of public economics* 4:2045–2141.
- Hanushek, Eric A. 2016. "School human capital and teacher salary policies." *Journal of Professional Capital and Community* 1(1):23–40.
- Hanushek, Eric A and Dennis D Kimko. 2000. "Schooling, labor-force quality, and the growth of nations." *American Economic Review* 90(5):1184–1208.
- Hanushek, Eric A and Ludger Woessmann. 2008. "The role of cognitive skills in economic development." *Journal of economic literature* 46(3):607–668.
- Hanushek, Eric A and Ludger Woessmann. 2011. "How much do educational outcomes matter in OECD countries?" *Economic Policy* 26(67):427–491.
- Hanushek, Eric A and Ludger Woessmann. 2012a. "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation." *Journal of Economic Growth* 17(4):267–321.

- Hanushek, Eric A and Ludger Woessmann. 2012b. "Schooling, educational achievement, and the Latin American growth puzzle." *Journal of Development Economics* 99(2):497–512.
- Hanushek, Eric A. and Ludger Woessmann. 2015. "Universal Basic Skills. What Countries Stand to Gain."
- Hanushek, Eric A, Ludger Woessmann, Eliot A Jamison and Dean T Jamison. 2008. "Education and economic growth." *Education Next* 8(2).
- Harada, Masataka. 2015. Generalized sensitivity analysis and application to quasi-experiments. Technical report Working Paper New York University.
- Harding, Robin. 2015. "Attribution and Accountability: Voting for Roads in Ghana." *World Politics* 67(4):656–689.
- Harding, Robin and David Stasavage. 2014. "What Democracy Does (and Doesn't Do) for Basic Services: School Fees, School Inputs, and African Elections." *Journal of Politics* 76(1):229–245.
- Harris, Douglass N. and Tim R. Sass. 2011. "Teacher training, teacher quality and student achievement." *Journal of Public Economics* 95(7–8):798–812.
- Hayhoe, Ruth. 1995. "An Asian multiversity? Comparative reflections on the transition to mass higher education in East Asia." *Comparative Education Review* 39(3):299–321.
- Hecock, R Douglas. 2014. "Democratization, Education Reform, and the Mexican Teachers' Union." *Latin American Research Review* 49(1):62–82.
- Hedström, Peter. 2008. "Studying mechanisms to strengthen causal inferences in quantitative research." *The Oxford Handbook of Political Methodology*. Oxford University Press, Oxford pp. 319–335.
- Hegre, Håvard, Carl Henrik Knutsen and E Geelmuyden Rød. 2012. "The Determinants of Democracy: A Sensitivity Analysis." Oslo: Department of Political Science, University of Oslo: Working Paper .
- Hegre, Håvard, Gudrun Østby and Clionadh Raleigh. 2009. "Poverty and Civil War Events: A Disaggregated Study of Liberia." *Journal of Conflict Resolution* 53(4):598–623.
- Hegre, Håvard and Håvard Mokleiv Nygård. 2014. "Governance and Conflict Relapse." *Journal of Conflict Resolution* 0:0.

- Heid, Benedikt, Julian Langer and Mario Larch. 2012. "Income and democracy: Evidence from system GMM estimates." *Economics Letters* 116(2):166–169.
- Held, David. 2006. *Models of democracy*. Cambridge: Polity Press.
- Helliwell, John F. 1994. "Empirical Linkages between Democracy and Economic Growth." *British Journal of Political Science* 24(2):225–248.
- Hendrix, Cullen S. 2010. "Measuring State Capacity: Theoretical and Empirical Implication for the Study of Civil Conflict." *Journal of Peace Research* 47(3):271–285.
- Hendrix, Cullen S and Idean Salehyan. 2012. "Climate change, rainfall, and social conflict in Africa." *Journal of Peace Research* 49(1):35–50.
- Hillygus, D Sunshine. 2005. "The missing link: Exploring the relationship between higher education and political engagement." *Political Behavior* 27(1):25–47.
- Holland, Paul W. 1986. "Statistics and causal inference." *Journal of the American Statistical Association* 81(396):945–960.
- Holtermann, Helge. 2012. "Explaining the Development-Civil War Relationship." *Conflict Management and Peace Science* 29(1):56–78.
- Honaker, James and Gary King. 2010. "What to Do about Missing Values in Time-Series Cross-Section Data." *American Journal of Political Science* 54(3):561–581.
- Honaker, James, Gary King, Matthew Blackwell et al. 2011. "Amelia II: A program for missing data." *Journal of statistical software* 45(7):1–47.
- Houle, Christian. 2009. "Inequality and Democracy: Why Inequality Harms Consolidation but Does Not Affect Democratization." *World Politics* 61(4):589–622.
- House, Freedom. 2003. *Freedom in the World 2002. Country report: Malawi*. Washington, DC: Freedom House.
- House, Freedom. 2013. *Freedom in the World. Country report: Qatar*. Washington, DC: Freedom House.
- Hoxby, Caroline M. 2000. "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *Quarterly Journal of Economics* 115(4):1239–1285.
- Huber, Evelyn and John D. Stephens. 2012. *Democracy and the Left: Social Policy and Inequality in Latin America*. Chicago: University of Chicago Press.

- Humphreys, Macarthur and Jeremy M. Weinstein. 2008. "Who fights? The determinants of participation in civil war." *American Journal of Political Science* 52(2):436–455.
- Huntington, Samuel P. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. Norman: University of Oklahoma Press.
- Imai, Kosuke, Luke Keele, Dustin Tingley and Teppei Yamamoto. 2010. Causal mediation analysis using R. In *Advances in social science research using R*. Springer pp. 129–154.
- Imai, Kosuke, Luke Keele, Dustin Tingley and Teppei Yamamoto. 2011. "Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies." *American Political Science Review* 105(04):765–789.
- Imbens, Guido W. 2003. "Sensitivity to exogeneity assumptions in program evaluation." *The American Economic Review* 93(2):126–132.
- Immigration and Refugee Board of Canada. 1991. "The Fall of President Ershad.". URL: <http://www.refworld.org/cgi-bin/texis/vtx/rwmain?docid=3ae6a8050>
- Inglehart, Ronald and Christian Welzel. 2005. *Modernization, cultural change, and democracy: The human development sequence*. Cambridge: Cambridge University Press.
- Inglehart, Ronald F, Bi Puranen and Christian Welzel. 2015. "Declining willingness to fight for one's country. The individual-level basis of the long peace." *Journal of Peace Research* 52(4):418–434.
- Jackson, Linda A., Alexander von Eye, Frank A. Biocca, Gretchen Barbatsis, Yong Zhao and Hiram E. Fitzgerald. 2006. "Does home internet use influence the academic performance of low-income children?" *Developmental Psychology* 42(3):429–435.
- Jamison, Eliot A, Dean T Jamison and Eric A Hanushek. 2007. "The effects of education quality on income growth and mortality decline." *Economics of Education Review* 26(6):771–788.
- Jennings, Colin. 2015. "Collective choice and individual action: Education policy and social mobility in England." *European Journal of Political Economy* 40:288–297.
- Kahneman, Daniel. 2011. *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kalyvas, Stathis N. 2006. *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.

- Kam, Cindy D and Carl L Palmer. 2008. “Reconsidering the effects of education on political participation.” *The Journal of Politics* 70(03):612–631.
- Kennedy, Peter. 2008. *A Guide to Modern Econometrics*. Oxford: Blackwell Publishing.
- Kennedy, Ryan. 2010. “The Contradiction of Modernization: A Conditional Model of Endogenous Democratization.” *The Journal of Politics* 72(3):785–798.
- King, Gary. 2010. A hard unsolved problem? Post-treatment bias in big social science questions. Paper presented at the “Hard Problems in Social Science” Symposium, Harvard University.
- King, Gary. 2013. *A solution to the ecological inference problem: Reconstructing individual behavior from aggregate data*. Princeton: Princeton University Press.
- King, Gary, Robert O. Keohane and Sidney Verba. 1994. *Designing Social Inquiry. Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- King, Gary, Robert O. Keohane and Sidney Verba. 1995. “The Importance of Research Design in Political Science.” *American Political Science Review* 89(2):475–481.
- Knutsen, Carl Henrik. 2011a. “Democracy, Dictatorship and Protection of Property Rights.” *Journal of Development Studies* 47(1):164–182.
- Knutsen, Carl Henrik. 2011b. The Economic Effects of Democracy and Dictatorship PhD thesis Department of Political Science, University of Oslo.
- Knutsen, Carl Henrik. 2011c. “Security Threats, Enemy-Contingent Policies and Economic Development in Dictatorships.” *International Interactions* 37(4):414–440.
- Knutsen, Carl Henrik. 2015. “Reinvestigating the Reciprocal Relationship between Democracy and Income Inequality.” *Review of Economics & Institutions/Economia, Società e Istituzioni* 6(2).
- Knutsen, Carl Henrik and Hanne Fjelde. 2013. “Property rights in dictatorships: kings protect property better than generals or party bosses.” *Contemporary Politics* 19(1):94–114.
- Knutsen, Carl Henrik, Håvard Mokleiv Nygård and Tore Wig. 2017. “Autocratic Elections: Stabilizing Tool or Force for Change?” *World Politics* 69(1):98–143.
- Kohlberg, Lawrence. 1976. “Moral stages and moralization: The cognitive-developmental approach.” *Moral development and behavior: Theory, research, and social issues* pp. 31–53.

- Kosack, Stephen. 2014. "The logic of pro-poor policymaking: Political entrepreneurship and mass education." *British Journal of Political Science* 44(2):409–444.
- Krueger, Alan B and Jitka Maleckova. 2003. "Education, poverty and terrorism: Is there a causal connection?" *The Journal of Economic Perspectives* 17(4):119–144.
- Kumah, Stephanie and Samuel Brazys. 2016. "Democracy or Accountability? Governance and Social Spending in Africa." *Journal of Development Studies* 52(2):286–299.
- Kuran, Timur. 1991. "Now out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44(1):7–48.
- Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge: Harvard University Press.
- Kurtines, William M, Jacob Gewirtz and Jacob L Lamb. 2014. *Handbook of Moral Behavior and Development: Volume 2: Research*. Psychology Press.
- Lake, David A. and Matthew A. Baum. 2001. "The Invisible Hand of Democracy: Political Control and the Provision of Public Services." *Comparative Political Studies* 34(6):587–621.
- Leamer, Edwar. 1985. "Sensitivity analyses would help." *American Economic Review* 75(3):308–313.
- Lerner, Daniel. 1958. *The passing of traditional society: Modernizing the Middle East*. New York: Free Press.
- Lichbach, Mark Irving. 1995. *The Rebel's Dilemma*. Ann Arbor: University of Michigan Press.
- Lindert, Peter H. 2005. *Growing Public: Social Spending and Economic Growth since the Eighteenth Century. Volume 1*. Cambridge: Cambridge University Press.
- Linz, Juan and Alfred Stepan. 1996. *Problems of Democratic Transition and Consolidation*. Baltimore: Johns Hopkins University Press.
- Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53(1):69–105.
- Little, Daniel. 2005. "Causal Mechanisms in Comparative Historical Sociology.". Working paper.
- URL:** <http://www-personal.umd.umich.edu/~delittle/SSHA.htm>

- Little, Daniel. 2012. "Lecture 4. Frameworks for the social sciences.".
URL: <http://www-personal.umd.umich.edu/~delittle/Lectures204.pdf>
- Londregan, John B. and Keith T. Poole. 1996. "Does High Income Promote Democracy?" *World Politics* 49(3):1–30.
- Lott, Jr, John R. 1999. "Public schooling, indoctrination, and totalitarianism." *Journal of Political Economy* 107(S6):127–157.
- Loveless, Tom. 2011. *Conflicting missions? Teachers unions and educational reform*. Washington D.C.: Brookings Institution Press.
- Lucas, Robert E. 1988. "On the Mechanisms of Economic Development." *Journal of Monetary Economics* 22:3–42.
- MacCulloch, Robert. 2004. "The impact of income on the taste for revolt." *American Journal of Political Science* 48(4):830–848.
- Magaloni, Beatriz. 2008. *Voting for Autocracy: Hegemonic Party Survival and its Demise in Mexico*. Cambridge: Cambridge University Press.
- Magaloni, Beatriz and Ruth Kricheli. 2010. "Political Order and One-Party Rule." *Annual Review of Political Science* 13(1):123–143.
- Mankiw, Gregory N., David Romer and David N. Weil. 1992. "A Contribution to the Empirics of Economic Growth." *The Quarterly Journal of Economics* 107(2):407–437.
- Mares, Isabela and Matthew E Carnes. 2009. "Social policy in developing countries." *Annual Review of Political Science* 12:93.
- Marshall, Monty G. 2014. "Polity IV Project: Political Regime Characteristics and Transitions, 1800–2013."
URL: <http://www.systemicpeace.org/polity/polity4.htm>. Accessed January 16, 2014.
- Marshall, Monty G., Ted Robert Gurr and Keith Jagers. 2013. "POLITY IV PROJECT: Political Regime Characteristics and Transitions, 1800–2012." Center for Systemic Peace.
- Martin, Bradley K. 2004. *Under the Loving Care of the Fatherly Leader: North Korea and the Kim Dynasty*. New York: Thomas Dunne.
- Maslow, Abraham Harold. 1943. "A theory of human motivation." *Psychological review* 50(4):370.

- Mattes, Robert and Dangalira Mughogho. 2009. "The limited impacts of formal education on democratic citizenship in Africa."
- McAdam, Doug. 1986. "Recruitment to high-risk activism: The case of freedom summer." *American Journal of Sociology* 92(1):64–90.
- McAdam, Doug, Sidney Tarrow and Charles Tilly. 2001. *Dynamics of Contention*. Cambridge: Cambridge University Press.
- McCarthy, John D and Mayer N Zald. 1977. "Resource mobilization and social movements: A partial theory." *American Journal of Sociology* 82(6):1212–1241.
- McClintock, Cynthia. 1984. "Why Peasants Rebel: The Case of Peru's Sendero Luminoso." *World Politics* 37(1):48–84.
- McClurg, Scott D. 2003. "Social networks and political participation: The role of social interaction in explaining political participation." *Political Research Quarterly* 56(4):449–464.
- Meltzer, Allan and Scott Richards. 1981. "A rational theory of the size of government." *Journal of Political Economy* 89(5):914–927.
- Meredith, Martin. 2006. *The State of Africa – A History of Fifty Years of Independence*. London: The Free Press.
- Michaelowa, Katharina. 2001. "Primary education quality in Francophone Sub-Saharan Africa: Determinants of learning achievement and efficiency considerations." *World Development* 29(10):1699–1716.
- Miguel, Edward, S. Satyanath and E Serengeti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112(4):725–753.
- Miller, Michael K. 2012. "Economic Development, Violent Leader Removal, and Democratization." *American Journal of Political Science* 56(4):1002–1020.
- Miller, SM. 1977. "Economic crisis and oppositional movements in the USA." *International Journal of Urban and Regional Research* 1(1-4):126–131.
- Milligan, Kevin, Enrico Moretti and Philip Oreopoulos. 2004. "Does education improve citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics* 88(9):1667–1695.

- Moe, Terry M. 2011. *Special interest: Teachers unions and America's public schools*. Washington D.C.: Brookings Institution Press.
- Moore, Barrington. 1966. *Social Origins of Democracy and Dictatorship: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.
- Moral-Benito, Enrique and Cristian Bartolucci. 2012. "Income and democracy: revisiting the evidence." *Economics Letters* 117(3):844–847.
- Morrisson, Christian and Fabrice Murtin. 2009. "The century of education." *Journal of Human capital* 3(1):1–42.
- Munck, Gerardo L. 2016. Democracy and Democratization in the Modern Age: The Evolution of Knowledge Since World War II. Working paper University of Southern California.
- Munck, Gerardo L. and Jay Verkuilen. 2002. "Conceptualizing and Measuring Democracy: Evaluating Alternative Indices." *Comparative Political Studies* 35(1):5–34.
- Murray, Michael P. 2006. "Avoiding invalid instruments and coping with weak instruments." *Journal of Economic Perspectives* 20(4):111–132.
- Murtin, Fabrice and Romain Wacziarg. 2014. "The democratic transition." *Journal of Economic Growth* 19(2):141–181.
- Nepstad, Sharon Erickson. 2011. *Nonviolent revolutions: Civil resistance in the late 20th century*. Oxford: Oxford University Press.
- Ness, Immanuel. 2015. "The International Encyclopedia of Revolution and Protest.". **URL:** <http://www.revolutionprotestencyclopedia.com>
- Newton, Kenneth. 1997. "Social capital and democracy." *American behavioral scientist* 40(5):575–586.
- Nickell, Stephen. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49(6):1417–1426.
- Nie, Norman H. 1987. *Participation in America: Political democracy and social equality*. Chicago: University of Chicago Press.
- Nkonyangi, John A. 1991. "Student protests in sub-Saharan Africa." *Higher Education* 22(2):157–173.

- North, Douglass C., John Joseph Wallis and Barry R. Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge: Cambridge University Press.
- Nunn, Nathan. 2010. "Religious conversion in colonial Africa." *The American Economic Review* 100(2):147–152.
- O'Donnell, Guillermo and Phillip C. Schmitter. 1986. *Transitions from Authoritarian Rule: Tentative Conclusions About Uncertain Democracy*. Baltimore: Johns Hopkins University Press.
- OECD, UNESCO and Eurostat. 2006. "Definitions, Explanations and Instructions." 2001 Data Collection on Education Systems.
- Okasha, Samir. 2016. *Philosophy of Science: Very Short Introduction*. Oxford: Oxford University Press.
- Olson, Mancur. 1965. *The Logic of Collective Action*. Cambridge: Harvard University Press.
- Olson, Mancur. 1982. *The Rise and Decline of Nations. Economic Growth, Stagflation and Social Rigidities*. New Haven: Yale University Press.
- Oreopoulos, Philip and Kjell G Salvanes. 2011. "Priceless: The nonpecuniary benefits of schooling." *The journal of economic perspectives* 25(1):159–184.
- Østby, Gudrun and Henrik Urdal. 2011. "Education and civil conflict: A review of the quantitative, empirical literature." *Background paper prepared for UNESCO's Education for All Global Monitoring Report* .
- Oyefusi, Aderoju. 2008. "Oil and the probability of rebel participation among youths in the Niger Delta of Nigeria." *Journal of Peace Research* 45(4):539–555.
- Papaioannou, Elias and Gregorios Siourounis. 2008. "Democratisation and growth." *The Economic Journal* 118(532):1520–1551.
- Paterson, Lindsay. 2009. "Civic values and the subject matter of educational courses." *Oxford Review of Education* 35(1):81–98.
- Pérez-Armendáriz, Clarisa and David Crow. 2010. "Do migrants remit democracy? International migration, political beliefs, and behavior in Mexico." *Comparative political studies* 43(1):119–148.

- Persson, Mikael. 2012. "Does Type of education affect Political Participation? results From a Panel survey of swedish adolescents." *Scandinavian Political Studies* 35(3):198–221.
- Persson, Torsten and Guido Tabellini. 2003. *The Economic Effect of Constitutions*. Cambridge, Massachusetts: MIT Press.
- Pierskalla, Jan H. and Florian M. Hollenbach. 2013. "Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa." *American Political Science Review* 107:207–224.
- Pinker, Steven. 2011. *The Better Angels of Our Nature. Why Violence has Declined*. New York: Viking.
- Plümper, Thomas and Vera E Troeger. 2007. "Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects." *Political Analysis* 15(2):124–139.
- Plümper, Thomas, Vera E Troeger and Philip Manow. 2005. "Panel data analysis in comparative politics: Linking method to theory." *European Journal of Political Research* 44(2):327–354.
- Political Risk Group. 2010. "International Country Risk Guide Methodology.".
URL: <http://www.prsgroup.com/PDFS/icrgmethodology.pdf>
- Pritchett, Lant. 2002. "When Will They Ever Learn? Why All Governments Produce Schooling." Working paper. Cambridge: John F. Kennedy School of Government, Harvard University.
- Przeworski, Adam and Fernando Limongi. 1997. "Modernization: Theory and Facts." *World Politics* 49(2):155–183.
- Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub and Fernando Limongi. 2000. *Democracy and Development. Political Institutions and Well-Being in the World, 1950–1990*. Cambridge: Cambridge University Press.
- Puryear, Jeffrey. 1994. *Thinking politics: intellectuals and democracy in Chile, 1973-1988*. Baltimore: John Hopkins University Press.
- Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.

- Raleigh, Clionadh. 2015. "Urban violence patterns across African states." *International Studies Review* 17(1):90–106.
- Rodrik, Dani. 2008. *One Economics, Many Recipes: Globalization, Institutions, and Economic Growth*. Princeton: Princeton University Press.
- Roome, William. 1925. *Ethnographic Survey of Africa. Showing the Tribes and Languages; Also the Stations of Missionary Societies*. E. Stanford Limited.
- Rosenberg, Alexander. 2012. *Philosophy of Science: A Contemporary Introduction*. New York: Routledge.
- Ross, Michael. 2001. "Does Oil Hinder Democracy." *World Politics* 53(April):325 –61.
- Ross, Michael. 2012. *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton: Princeton University Press.
- Rothstein, Bo. 2011. *The quality of government: Corruption, social trust, and inequality in international perspective*. Chicago: University of Chicago Press.
- Rothstein, Bo and Jan Teorell. 2008. "What Is Quality of Government? A Theory of Impartial Government Institutions." *Governance* 21(2):165–190.
- Rubin, Donald B. 2005. "Causal inference using potential outcomes: Design, modeling, decisions." *Journal of the American Statistical Association* 100(469):322–331.
- Rueschemeyer, Dietrich, Evelyne Huber Stephens and John D. Stephens. 1992. *Capitalist Development and Democracy*. Chicago: University of Chicago Press.
- Rustow, Dankwart A. 1970. "Transitions to democracy: Toward a dynamic model." *Comparative Politics* 2(3):337–363.
- Saint-Paul, G. and T. Verdier. 1993. "Education, democracy and growth." *Journal of Development Economies* 42(2):399–407.
- Sala-i-Martin, Xavier X. 1997. "I just ran two million regressions." *American Economic Review* 87(2):178–183.
- Salehyan, Idean and Cullen Hendrix. 2016. "Codebook for Social Conflict Analysis Database."
- Salehyan, Idean, Cullen S. Hendrix, Jesse Hamner, Christina Case, Christopher Linebarger, Emily Stull and Jennifer Williams. 2012. "Social Conflict in Africa: A New Database." *International Interactions* 38(4):503–511.

- Sanborn, Howard and Clayton L Thyne. 2014. “Learning democracy: Education and the fall of authoritarian regimes.” *British Journal of Political Science* 44(4):773–797.
- Schock, Kurt. 2005. *Unarmed Insurrection. People Power Movements in Nondemocracies*. Minneapolis: University of Minnesota Press.
- Schumpeter, Joseph A. 1976. *Capitalism, Socialism and Democracy*. New York: Harper-Perennial.
- Sen, Amartya. 1999. *Development as Freedom*. New York: Anchor Books.
- Sharp, Gene. 2005. *Waging nonviolent struggle: 20th century practice and 21st century potential*. New York: Extending Horizons Books.
- Shayo, Moses. 2007. *Education, militarism and civil wars*. Maurice Falk Institute for Economic Research in Israel.
- Shepsle, Kenneth A. and Mark S. Boncheck. 1997. *Analyzing Politics: Rationality, Behavior and Institutions*. New York: Norton.
- Siegel, David A. 2009. “Social networks and collective action.” *American Journal of Political Science* 53(1):122–138.
- Solis, Alex. 2013. Does higher education cause political participation? Evidence from a regression discontinuity design. Technical report Working Paper, Department of Economics, Uppsala University.
- Solt, Frederick. 2009. “Standardizing the World Income Inequality Database.” *Social Science Quarterly* 90(2):231–242.
- Sondheimer, Rachel Milstein and Donald P Green. 2010. “Using experiments to estimate the effects of education on voter turnout.” *American Journal of Political Science* 54(1):174–189.
- Sovey, Allison J and Donald P Green. 2011. “Instrumental variables estimation in political science: A readers’guide.” *American Journal of Political Science* 55(1):188–200.
- Spilimbergo, Antonio. 2009. “Democracy and foreign education.” *The American Economic Review* 99(1):528–543.
- Stasavage, David. 2005. “Democracy and Educational Spending in Africa.” *American Journal of Political Science* 49(2):343–358.
- Stiglitz, Joseph E. 1997. *Economics. Second Edition*. New York: W.W. Norton.

- Stock, James H. 2010. “The other transformation in econometric practice: Robust tools for inference.” *The Journal of Economic Perspectives* 24(2):83–94.
- Stock, James H and Motohiro Yogo. 2005a. “Testing for weak instruments in linear IV regression.” *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* .
- Stock, James H and Motohiro Yogo. 2005b. “Testing for weak instruments in linear IV regression.” *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* .
- Svolik, Milan. 2012. *The Politics of Authoritarian Rule*. Cambridge: Cambridge University Press.
- Svolik, Milan W. 2009. “Power Sharing and Leadership Dynamics in Authoritarian Regimes.” *American Journal of Political Science* 53(2):477–494.
- Swarthmore. 2015. “Global Nonviolent Action Database.”
- URL:** <http://nvdatabase.swarthmore.edu>
- Tajfel, Henri and John C Turner. 1979. “An integrative theory of intergroup conflict.” *The social psychology of intergroup relations* 33(47):74.
- Tarrow, Sidney. 1994. *Power in Movement: Social Movements, Collective Action and Politics*. Cambridge: Cambridge University Press.
- Tavares, Jose and Romain Wacziarg. 2001. “How democracy affects growth.” *European Economic Review* 45(8):1341–1378.
- Taylor, Michael. 1988. *Rationality and Revolution*. Reading: Cambridge University Press.
- Tenn, Steven. 2007. “The effect of education on voter turnout.” *Political Analysis* 15(4):446–464.
- Teorell, Jan. 2010. *Determinants of Democracy: Explaining Regime Change in the World, 1972–2006*. Cambridge: Cambridge University Press.
- Teorell, Jan, Marcus Samanni, Søren Holmberg and Bo Rothstein. 2011. “The Quality of Government Dataset, version 6Apr11.”
- URL:** <http://www.qog.pol.gu.se/data/qogstandarddataset/>
- Teorell, Jan, Michael Coppedge, Svend-Erik Skaaning and Staffan I Lindberg. 2016. “Measuring Electoral Democracy with V-Dem Data: Introducing a New Polyarchy Index.” V-Dem Working Paper 2016:25. Gothenburg: Varieties of Democracy Institute.

- Thyne, Clayton. 2006. “ABC’s, 123’s, and the Golden Rule: The Pacifying Effect of Education on Civil War, 1980–1999.” *International Studies Quarterly* 50(4):733–754.
- Tilly, Charles. 2003. *The Politics of Collective Violence*. Cambridge: Cambridge University Press.
- Tilly, Charles and Sidney Tarrow. 2006. *Contentious politics*. Oxford: Oxford University Press.
- Tingley, Dustin, Teppei Yamamoto, Kentaro Hirose, Luke Keele and Kosuke Imai. 2014. “mediation: R Package for Causal Mediation Analysis.” *Journal of Statistical Software* 59.
- Tollefsen, Andreas Forø, Håvard Strand and Halvard Buhaug. 2012. “PRIO-GRID: A unified spatial data structure.” *Journal of Peace Research* 49(2):363–374.
- Treisman, Daniel. 2011. “Income, democracy, and the cunning of reason.” National Bureau of Economic Research (NBER) Working Paper.
- Tsoi, Grace. 2014. “Did ‘Liberal Studies’ enable Hong Kong’s youth awakening?” *Foreign Policy* November 18th.
- URL:** <http://foreignpolicy.com/2014/11/18/did-liberal-studies-enable-hong-kongs-youth-awakening/>
- Ulfelder, Jay. 2005. “Contentious collective action and the breakdown of authoritarian regimes.” *International Political Science Review* 26(3):311–334.
- Unesco. 2014. “Teaching and learning: Achieving quality for all.” Education for All global Monitoring Report 2013/2014.
- Unesco. 2015. “Unesco Institute for Statistics Database.”
- URL:** http://data UIS.unesco.org/Index.aspx?DataSetCode=EDULIT_DSpopupcustomise=true&lang=en
- Urdal, Henrik. 2006. “A Clash of Generations? Youth Bulges and Political Violence.” *International Studies Quarterly* 50(3):607–630.
- Van Fleet, Joerg. 2012. “Africa’s education crisis: In school but not learning.”
- URL:** <http://www.brookings.edu/blogs/up-front/posts/2012/09/17-africa-education-crisis-van-fleet>
- Verba, Sidney and Norman H Nie. 1987. *Participation in America: Political democracy and social equality*. Chicago: University of Chicago Press.

- Wallace, Jeremy. 2013. "Cities, redistribution, and authoritarian regime survival." *The Journal of Politics* 75(03):632–645.
- Wantchekon, Leonard. 2004. "The paradox of "warlord" democracy: A theoretical investigation." *American Political Science Review* 98(1):17–33.
- Weakliem, David L. 2002. "The effects of education on political opinions: An international study." *International Journal of Public Opinion Research* 14(2):141–157.
- Weber, Max. 2002. *The Protestant Ethic and "The Spirit of Capitalism"*. London: Penguin Books.
- Weidmann, Nils B. 2009. "Geography as Motivation and Opportunity: Group Concentration and Ethnic Concentration." *Journal of Conflict Resolution* 53(4):526–543.
- Weidmann, Nils B and Sebastian Schutte. 2016. "Using night light emissions for the prediction of local wealth." *Journal of Peace Research* Online first.
- URL:** DOI: <https://doi.org/10.1177/0022343316630359>
- Wells, H.G. 1937. "Joseph Stalin. Marxism versus liberalism. An interview with H.G. Wells." *Works* 14.
- Welzel, Christian. 2013. *Freedom rising*. Cambridge: Cambridge University Press.
- Weyland, Kurt. 2005. "Theories of Policy Diffusion Lessons from Latin American Pension Reform." *World Politics* 57(2):262–295.
- Wig, Tore and Andreas Forø Tollefson. 2016. "Local institutional quality and conflict violence in Africa." *Political Geography* 53:30–42.
- Wilson, Andrew. 2005. *Ukraine's Orange revolution*. New Haven: Yale University Press.
- Wintrobe, Ronald. 1998. *The Political Economy of Dictatorship*. Cambridge: Cambridge University Press.
- Wischnath, Gerdis and Halvard Buhaug. 2014. "On climate variability and civil war in Asia." *Climatic change* 122:709–721.
- Wolfinger, Raymond E. 1980. *Who votes?* Vol. 22 New Haven: Yale University Press.
- Wood, Elisabeth Jean. 2003. *Insurgent collective action and civil war in El Salvador*. Cambridge studies in comparative politics Cambridge and New York: Cambridge University Press.

- Woodberry, Robert D. 2012. “The missionary roots of liberal democracy.” *American Political Science Review* 106(2):244–274.
- Woodward, Jim. 2002. “What is a mechanism? A counterfactual account.” *Philosophy of Science* 69(S3):366–377.
- Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.
- World Bank. 2011. “World Development Indicators.” Washington, DC: World Bank.
- Wright, Joseph. 2008. “Do Authoritarian Institutions Constrain? How Legislatures Affect Economic Growth and Investment.” *American Journal of Political Science* 52(2):322–343.
- Wright, Joseph and Abel Escriba-Folch. 2012. “Authoritarian Institutions and Regime Survival: Transitions to Democracy and Subsequent Autocracies.” *British Journal of Political Science* 42(2):283–309.
- Wright, Joseph, Erica Frantz and Barbara Geddes. 2015. “Oil and autocratic regime survival.” *British Journal of Political Science* 45(2):287–306.
- Zelig, Leo. 2008. “Student Politics and Activism in Zimbabwe: The Frustrated Transition.” *Journal of Asian and African Studies* 43(2):215–237.
- Zhao, Yong. 2014. *Who’s Afraid of the Big Bad Dragon?: Why China Has the Best (and Worst) Education System in the World*. San Francisco: John Wiley & Sons.