

Democracy Does Cause Growth: Replicating ANRR (2019)*

Markus Eberhardt

*School of Economics, University of Nottingham, UK
Centre for Economic Policy Research, UK*

September 11, 2019

Abstract: I revisit the causal relationship between democracy and growth as recently studied in Acemoglu, Naidu, Restrepo, and Robinson (*Journal of Political Economy*, 2019; 127(1): 47-100), conducting a broad replication exercise. I demonstrate the sensitivity of their results to sample selection by dropping a small number of observations in a non-random fashion. Additional analysis suggests that one possible source of this lack of robustness may be their assumption of common GDP dynamics across all sample countries.

JEL Classification: O10, P16

Keywords: Democracy, Growth, Political Development

*I am indebted to Marc Chan, Matt Collin, Rodolphe Desbordes, Jonathan Dingel, Doug Gollin, Brian Knight, Simon Kwok, Gerard Padro i Miquel, Andrea Presbitero, Daniel Seidmann, Jon Temple, Nicolas van de Sijpe, Dietz Vollrath, and Joakim Westerlund for detailed comments and suggestions which helped improve the paper significantly. Seminar/session participants at Lund, Nottingham and 2019 CSAE conference provided useful feedback. The data and code of the original ANRR paper are available on Daron Acemoglu's personal website. The usual disclaimers apply. Correspondence: Markus Eberhardt, School of Economics, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, UK. Email: markus.eberhardt@nottingham.ac.uk

1 Introduction

In a recent article Acemoglu, Naidu, Restrepo, and Robinson (2019, henceforth ANRR) study the causal link between democracy and growth¹ in a large sample of up to 175 countries. The authors adopt a variety of empirical strategies for an empirical model which captures country-specific fixed effects and the dynamics of per capita GDP:²

$$y_{it} = \alpha_i + \gamma_t + \beta \text{Democracy}_{it} + \sum_{\ell=1}^p \rho_{i,t-\ell} y_{i,t-\ell} + \varepsilon_{it}, \quad (1)$$

where y is log per capita GDP (multiplied by 100), Democracy is a dummy variable, α_i and γ_t are country and time dummies, respectively, and ε is the error term. In order to allow for a causal interpretation of the results they devise an instrumentation strategy which builds on regional waves of democratisation and reversal. The findings from these 2SLS models are shown to be in line with results adopting country fixed effects (2FE), the Arellano and Bond (1991, AB) and the Hahn, Hausman, and Kuersteiner (2001, HHK) estimators. The collective message from these findings is that the *long-run* effects of democratisation are sizeable: an increase in per capita GDP of 20% or more.

ANRR's findings are remarkable on three counts: (i) previous research has often emphasised that 'democracy' as a political regime was perhaps "too blunt a concept" (Persson and Tabellini, 2006: 319) to provide robust empirical evidence; (ii) positive findings for a democracy-growth nexus were typically confined to a "short-run boost" (Rodrik and Wacziarg, 2005: 55); and (iii) ANRR's findings are robust across a host of empirical estimators with different assumptions about the Data Generating Process, including 2SLS adopting a new instrumental variable (regional waves of democratisation).

The aim of this broad replication is to study the robustness of the empirical results to changes in the sample. I devise two rule-based sample reduction strategies: (i) I drop countries on the basis of their observation count, first omitting those with merely five observations, then those with six, etc.; this sample reduction strategy is in line with the econometric aspiration to establish the *long-run effect* of democracy and also addresses concerns over the robustness of the assumed *common* GDP dynamics in highly unbalanced panels; and (ii) I omit observations on the basis of the sample year, moving the sample end year by one year at a time, i.e. dropping first 2010, then 2009, and so on. This strategy can be justified by the presence of the 2007/8

¹I follow the practice of ANRR (see their footnote 1) in using 'growth' as a short-hand for economic development (the level of per capita GDP). See Eberhardt and Teal (2011) for a more detailed discussion of growth empirics.

²My presentation is limited to the parametric results. The semi-parametric results for sample reduction strategy (i) yield confidence intervals which always include zero when around 20% of observations are omitted; for strategy (ii) results appear much less affected, if anything confidence intervals become *tighter* as respective end years are omitted. The source of this robustness relative to all other ANRR results is beyond the scope of this note, results are relegated to the Online Appendix.

Global Financial Crisis and its aftermath — the most significant global macroeconomic shock since the 1930s — at the end of the ANRR sample period.

My results demonstrate that ANRR’s long-run estimates become statistically insignificant when between 3% and 8% of the over 6,000 observations are dropped, and economically near-insignificant (long-run estimates below 5% in economic magnitude) when between 14% and 28% of observations are dropped. I demonstrate that there is no dramatic reduction in the count of democratisations/reversals in the reduced samples of my exercises. Additional analysis provides some evidence that dynamic misspecification (heterogeneous GDP dynamics) may account for the sensitivity of results observed.

2 Sample Reduction Exercises

In this section I discuss results from two sample reduction exercises presented in Figures 1 and 2. Table 1 summarizes the estimates and sample makeup of five *ad hoc* ‘thresholds’ in the long-run estimates for democracy: in Panel A for the full ANRR sample, in B the sample which yields an insignificant estimate, in C when the estimate falls below 5% in magnitude, in D when the reduced sample estimate is outside the confidence interval of the full sample one, and in E the balanced panel estimate. Columns [1]-[4] and [5]-[8] are for the respective sample reduction strategies. Using results in Figure 3 I speculate about one potential source of the patterns observed.

Sample reduction by minimum observation count I begin with the strategy which drops countries by their sample observation count. A major concern for this non-random sample reduction strategy is that even though the ‘small- T ’ countries may only account for a very small share of overall observations they may represent a disproportionate share of the democratisation and reversal events. If this were the case then the sample reduction strategy *by construction* makes it harder and harder for the estimators to identify a democracy effect. The histogram in Panel (a) of Figure 1 speaks to this concern — this plot is based on the AB/HHK sample (the 2SLS sample typically has one additional observation per country), detailed information about the countries dropped in these sample reduction exercises are contained in an Appendix. Along the x -axis we can see the minimum observation count for inclusion in the sample; the thin grey bars indicate the observation count (left scale, in logarithms). This highlights that over 60% of the full sample (around 4,000 observations) have data for all years, and for reference I will report the results for this ‘balanced’ panel below. The coloured bars indicate the distribution of democratisation and reversal events by minimum observation count: again roughly 60% of these events occur in the balanced panel sample, while the remainder are sprinkled thinly

across other minimum observation samples.

Panel (b) presents the full and reduced sample results for the 2FE, AB, HHK and 2SLS estimators — all results are for the specification with four lags of GDP, which is preferred by ANRR.³ In this and the equivalent plot in Panel (b) of Figure 2 a filled coloured (white) circle indicates statistical (in)significance at the 10% level. The left-most estimates correspond to the full sample results reported in the ANRR paper, the right-most to the estimates for a balanced panel. The x -axis is identical to the plot in panel (a), the y -axis indicates the long-run effect (in percent) of democracy on per capita GDP.

For the 2FE estimator, the sample reduction exercise has virtually no impact on the long-run democracy estimate: as we move to the right countries with fewer observations than the minimum number indicated on the x -axis are omitted from the regression sample, but the 2FE long-run estimate for democracy is virtually unchanged. The exception is the balanced panel result which is statistically insignificant, though at 15.6% still reasonably close to the full sample estimate of 21.2%.⁴

The patterns for the AB and HHK estimates are very different: both decline and turn statistically insignificant when the minimum observation count is 17 and thereafter fall (more or less monotonically) towards and beyond zero. Results in Columns [2] and [3] of Table 1 indicate that the AB and HHK estimates are statistically insignificant and reduced by a quarter and two-thirds, respectively, once 5% of the full sample observations are dropped. The balanced panel results for these two estimators (-5.3 and -12.4) are derived from a sample where just over 40% of observations are dropped.

Democracy estimates based on the 2SLS estimator initially maintain a high and stable level in excess of 30%, but turn insignificant once countries with fewer than 21 observations are omitted (7% of the full sample of 6,300 observations).⁵ The magnitude of 2SLS estimates drops quite rapidly, such that it falls below 5% in magnitude and also outside the full sample 90% confidence interval once 18% of observations are dropped. In contrast to the patterns for the AB and HHK estimators the 2SLS estimates increase again if further countries are dropped.

³Results for one and two lags are presented in an Appendix, where I also provide 2SLS estimates for the alternative construction of the long-run estimate with qualitatively identical results.

⁴Note that many researchers have serious reservations about the fixed effects estimator for causal inference in panel data (e.g. Gibbons, Suarez-Serratoz and Urbancic, 2018; Imai and Kim, 2018).

⁵The focus of the sample reduction exercises is primarily on the magnitudes of estimated coefficients, though statistical insignificance *can* indicate that underlying country estimates are heterogeneous and vary substantially across countries. If the focus of the exercise were more narrowly on the smallest sample reduction yielding a statistically insignificant long-run estimate for democracy, then the number of countries that would need to be dropped is very small: three for AB/HHK and four for 2SLS, amounting to fewer than 1% of observations in each case — see Appendix Table A-1. Recent work by Young (2018) has highlighted the fragility of IV estimates, demonstrating that many findings of statistical significance are driven by few observations. Here, it should be emphasised that the Appendix results derive from a purposeful exercise in sample selection (by trial and error), and further dropping a small number of countries may similarly *restore* the statistical significance of the estimates.

Sample reduction by sample end year Figure 2 presents the results when observations are omitted by sample end year. The primary focus here is on the impact of the Global Financial Crisis in 2007/8 and its aftermath. Panel (a) charts the distribution of sample observations and democratisation/reversal events by year — here and in panel (b) the x -axis is in reverse order. We can see that the annual sample observation count rises from the 1960s until peaking in the mid-2000s. The final three sample years 2008-10 account for around 8% of all observations (2010: 3%, 2009: 2%, 2008: 3%). The first 25 years of the sample indicate typically two to three democratisation/reversal events per annum, before a wave of events in the early 1990s following the collapse of the Soviet Union. The final three sample years 2008-10 indicate 14 events, around 9% of the total number of events over 1965-2010.⁶

Panel (b) presents the sample reduction results, where the x -axis indicates the final year included in the sample, and the y -axis indicates the long-run effect (in percent) of democracy on per capita GDP — again all estimates are for the 4-lag specification preferred by ANRR. I only chart end years down to 1995, since omitting 1996-2010 amounts to around 40% of observations, similarly to the 40% of observations omitted in the balanced panel of the ‘small T_i ’ exercise presented above.

As before the 2FE estimates are found to be fairly robust to sample reduction, only turning insignificant when 30% of observations are dropped. The AB/HHK estimates, in contrast, turn insignificant if the post-GFC years 2009 and 2010 are omitted, thereafter declining and eventually diverging, with HHK remaining positive (albeit insignificant throughout) while AB estimates turn negative (dto.). The 2SLS estimates are generally falling with earlier sample end years, but display curious patterns in the aftermath of the GFC: omitting only 2010 (3% of observations) yields a statistically insignificant long-run coefficient on democracy. Omitting both 2010 and 2009 (together 5% of observations) however restores the full sample coefficient in terms of magnitude and statistical significance, whereas the omission of further end years always yields statistically insignificant long-run democracy estimates. Table 1 provides all the details on estimates, standard errors and samples of the various ‘thresholds’ as defined above.

Some forensic analysis What are the reasons for this sensitivity of results to relatively small numbers of observations? In the following I indicate that the source of this puzzle is possibly related to parameter heterogeneity — my focus here is not on the estimates for democracy, but on the estimates for the GDP dynamics as the sample changes: *a priori* we do not know what the magnitude or even the sign of the democracy coefficient $\hat{\beta}$ in the dynamic 2SLS regression ‘should’ be (the literature has argued for positive or negative effects), but we know that the estimate for the GDP dynamics should be positive and somewhere below but fairly close to 1.

⁶This is once again the AB/HHK sample for the four-lag specification, hence the 1965 start year.

I limit my attention to the 4-lag 2SLS specification, where I plot the estimate and 90% confidence interval for the GDP dynamics (solid line), i.e. $\sum_{\ell=1}^4 \hat{\rho}_{i,t-\ell}$: in panel (a) of Figure 3 I drop countries by number of observations, and in panel (b) I drop observations by end year.

Since all of the parametric models studied above are pooled models, the democracy coefficient as well as the GDP dynamics are assumed to be *common* across countries. A high (low) coefficient on the GDP dynamics *ceteris paribus* implies a higher (lower) long-run coefficient on democracy in absolute terms: $\hat{\beta}^{LR} = \hat{\beta} / (1 - \sum_{\ell=1}^4 \hat{\rho}_{i,t-\ell})$. Figure 3 plots the estimated GDP dynamics used in computing these long-run democracy estimates and standard errors discussed above. For either sample reduction strategy the estimate on the GDP dynamics (solid line) is remarkably stable across samples, especially given the sensitivity of the long-run democracy coefficients in Figures 1 and 2.

But what is the estimate on the GDP dynamics in the countries or years I omit? The dashed line in panel (a) of Figure 3 represents the estimated GDP dynamics for all countries with a minimum observation count *lower* than that indicated on the x -axis:⁷ as we move to the right these countries are dropped from the sample estimating the solid line and included in the sample estimating the dashed line. It is noticeable that, with the exception of two, all of these estimates for GDP dynamics in the sample of ‘dropped’ countries are below those for the ‘included’ countries. For some samples toward the right end of the graph the confidence intervals of the two sets of estimates do not overlap.⁸ Similarly, in panel (b) the estimates on the GDP dynamics for the omitted end years are substantially below those of the included years, the patterns for 2008 and 2009 even speak to those of the results in Figure 2.

Thus, if GDP dynamics differ between countries in general, and between my samples of countries/years included and omitted in particular, then the inclusion of these ‘omitted’ countries or years may inflate the long-run democracy coefficients in the full sample results.

3 Concluding Remarks

My results for two sample reduction exercises indicate that the “economically and statistically significant positive effect” of democracy on per capita GDP uncovered in a recent paper by ANRR (2019: 27) is sensitive to sample selection: depending on the sample reduction strategy and the estimator used the long-run estimates become statistically insignificant when between 3% and 8% of observations are dropped, and economically insignificant when between 14% and 28% of observations are dropped. The sensitivity of these empirical results to non-random changes in the sample can point to the presence of observed and unobserved (time-varying)

⁷For $T_{min} = 6$ this estimate would be constructed from 6 observations in one country. I therefore only begin charting this estimate for countries with 17 or fewer observations (338 observations in 23 countries).

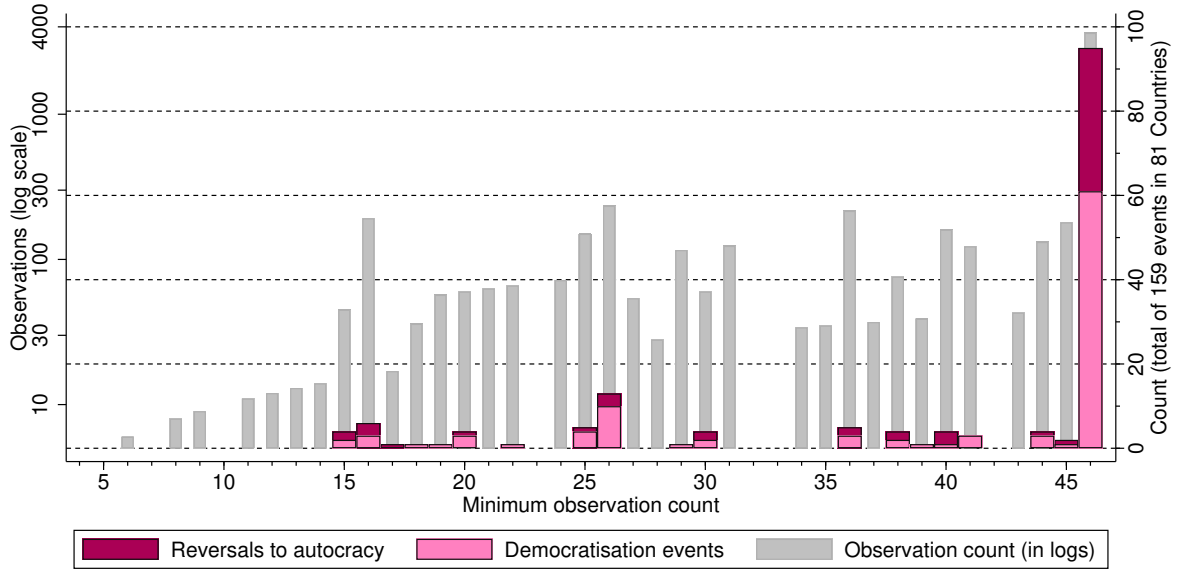
⁸I do not show the confidence intervals for the ‘drop-out’ estimates for ease of illustration.

heterogeneities: across different implementations ANRR always assume that the democracy-growth relationship and its dynamics are the same for all countries (parameter homogeneity), and that there is no bias due to time-variant unobserved heterogeneity (strong cross-section dependence). In ongoing work I plan to relax these assumptions adopting recent methods from the panel time series literature.

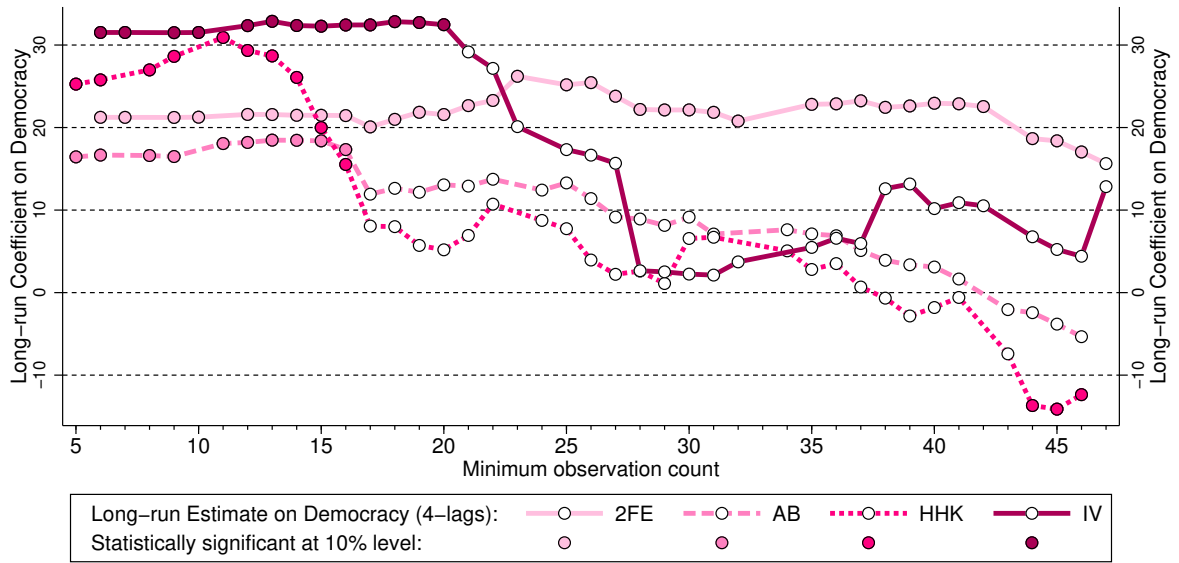
References

- [1] **Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson (2019)** ‘Democracy Does Cause Growth.’ *Journal of Political Economy* 127(1): 47-100.
- [2] **Arellano, Manuel, and Stephen R. Bond (1991)** ‘Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations.’ *Review of Economic Studies* 58(2): 277-297.
- [3] **Eberhardt, Markus, and Francis Teal (2011)** ‘Econometrics for grumblers: a new look at the literature on cross-country growth empirics.’ *Journal of Economic Surveys* 25(1): 109-155.
- [4] **Gibbons, Charles, Juan Carlos Suarez Serratoz, and Michael Urbancic (2018)** ‘Broken or Fixed Effects?’ *Journal of Econometric Methods*, forthcoming.
- [5] **Hahn, Jinyong, Jerry Hausman, and Guido Kuersteiner (2001)** *Bias corrected instrumental variables estimation for dynamic panel models with fixed effects*. MIT Department of Economics Working Paper 01-24, June 2001.
- [6] **Imai, Kosuke, and In Song Kim (2018)** ‘When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?’ *American Journal of Political Science*, forthcoming.
- [7] **Persson, Torsten, and Guido Tabellini (2006)** ‘Democracy and development: The devil in the details.’ *American Economic Review, Papers & Proceedings* 96(2): 319-324.
- [8] **Rodrik, Dani, and Romain Wacziarg (2005)** ‘Do democratic transitions produce bad economic outcomes?’ *American Economic Review, Papers & Proceedings* 95(2): 50-55.
- [9] **Young, Alwyn (2018)** ‘Consistency without Inference: Instrumental Variables in Practical Application.’ Unpublished mimeo, June.

Figure 1: Sample Reduction by T_i count



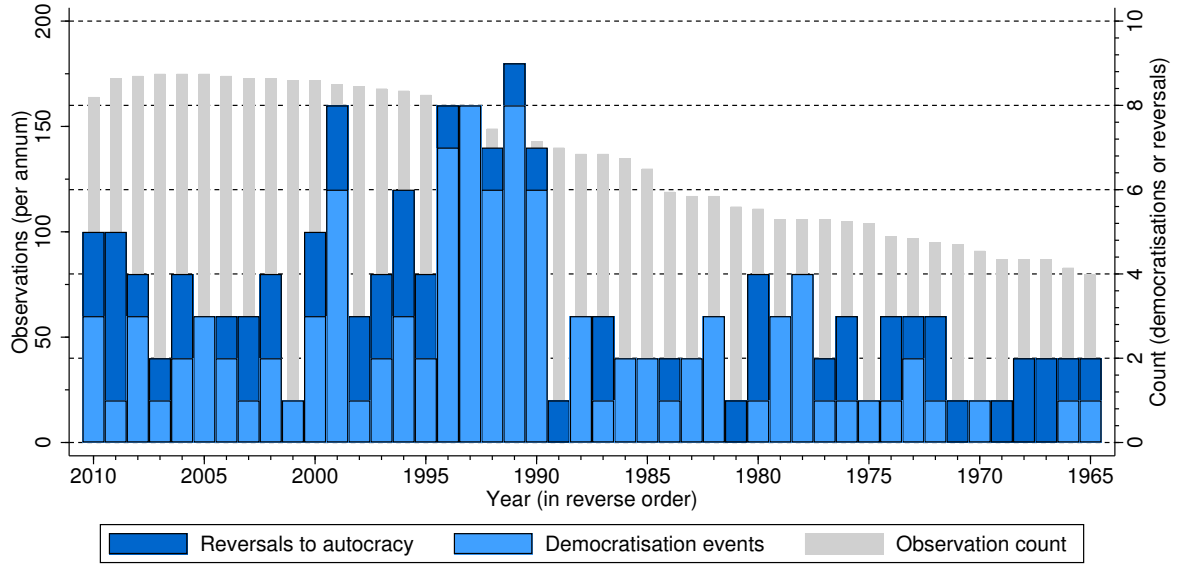
(a) Sample and Event Distribution



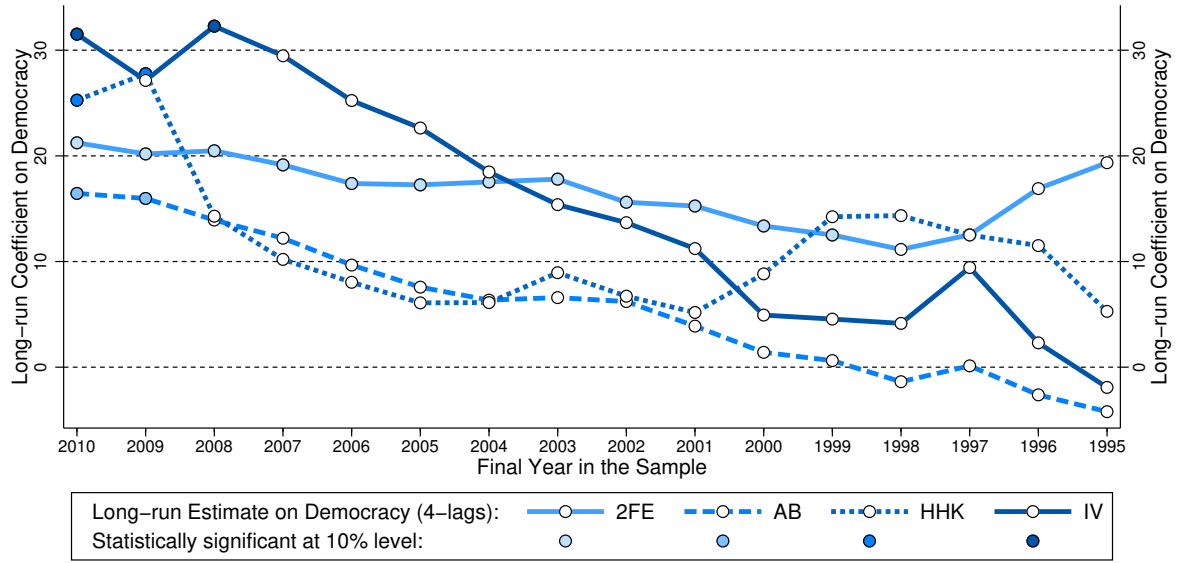
(b) Long-Run Estimates

Notes: The figure presents the sample distribution and long-run estimates for democracy from varying empirical samples for the 2FE, AB, HHK and 2SLS estimators, computed as $\hat{\beta}^{LR} = \hat{\beta} / (1 - \sum_{\ell=1}^4 \hat{\rho}_{i,t-\ell})$, where $\hat{\beta}$ is the estimate on the democracy dummy and the $\hat{\rho}$ are estimates for the lags of per capita GDP (standard errors are constructed via the Delta method). The x -axis in each plot indicates the minimum number of observations required to be included in the sample. In Panel (a) the thin grey bars indicate the distribution of observations (log scale, left axis) while the coloured bars indicate democratisation and reversal events (right scale). These distributions are for the AB/HHK samples. In Panel (b) a filled (white) circle marker indicates that the long-run coefficient is statistically (in)significant at the 10% level. All estimates are for the specification with four lags of GDP (and four lags of the instrument for 2SLS) preferred by ANRR — see Appendix for results of alternative specifications. The ‘left-most’ estimates replicate the results in ANRR’s Table 2, column (3) for 2FE, (7) for AB, and (11) for HHK, and Table 6, column (2) Panel A for 2SLS.

Figure 2: Sample Reduction by end year



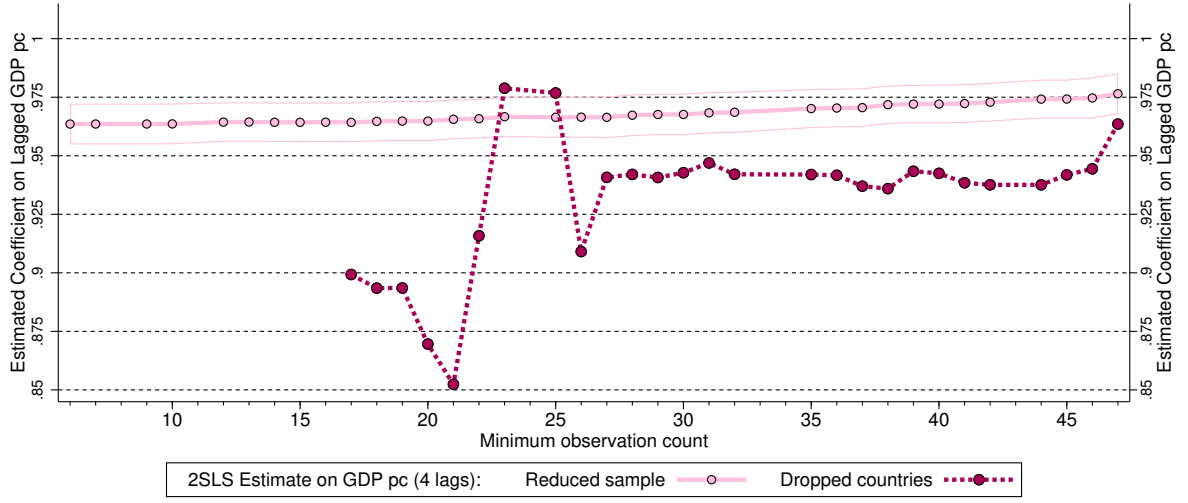
(a) Sample and Event Distribution



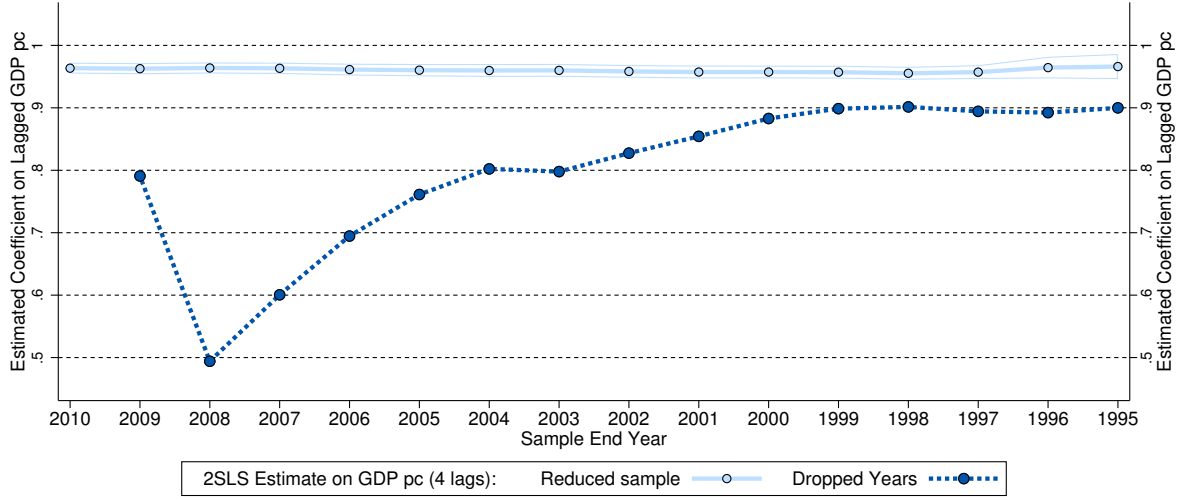
(b) Long-Run Estimates

Notes: The figure presents the sample distribution and long-run estimates for democracy from varying empirical samples for the 2FE, AB, HHK and 2SLS estimators, computed as $\hat{\beta}^{LR} = \hat{\beta} / (1 - \sum_{\ell=1}^4 \hat{\rho}_{i,t-\ell})$, where $\hat{\beta}$ is the estimate on the democracy dummy and the $\hat{\rho}$ are estimates for the lags of per capita GDP (standard errors are constructed via the Delta method). The x -axis in each plot indicates the end year included in the sample. In Panel (a) the thin grey bars indicate the distribution of observations (left axis) while the coloured bars indicate democratisation and reversal events (right scale). These distributions are for the AB/HHK samples. In Panel (b) a filled (white) circle marker indicates that the long-run coefficient is statistically (in)significant at the 10% level. All estimates are for the specification with four lags of GDP preferred by ANRR — see Appendix for results of alternative specifications. The ‘left-most’ estimates replicate the results in ANRR’s Table 2, column (3) for 2FE, (7) for AB, and (11) for HHK, and Table 6, column (2) Panel A for 2SLS.

Figure 3: The GDP Dynamics of ‘Omitted’ Countries



(a) Sample reduction by T_i count



(b) Sample reduction by end year

Notes: The plots present estimates on the sum of lagged GDP terms in the 4-lag 2SLS regressions for the sample reduction by T_i count in panel (a) and for the reduction by end year in panel (b): $\sum_{\ell=1}^4 \hat{\rho}_{i,t-\ell}$, where the $\hat{\rho}$ are the coefficients on the lags of per capita GDP (standard errors are constructed via the Delta method). Each panel plots two series, the estimates (i) for the reduced sample (solid line with 90% CI), and (ii) for those countries or end years which are dropped (dashed line), e.g. the 2009 estimate in panel (b) is for the years 2010 and 2009.

Table 1: Sample Reduction Estimates

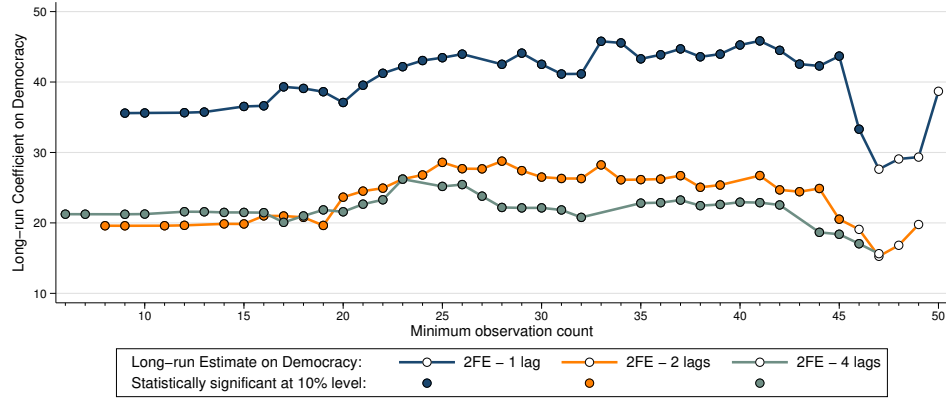
Estimator	Sample reduction by T_i count				Sample reduction by end year			
	[1] 2FE	[2] AB	[3] HHK	[4] 2SLS	[5] 2FE	[6] AB	[7] HHK	[8] 2SLS
Panel A: Full ANRR sample estimates								
Long-Run Democracy Effect	21.240 [7.215]***	16.448 [8.436]*	25.268 [10.869]**	31.521 [17.425]*	21.240 [7.215]***	16.448 [8.436]*	25.268 [10.869]**	31.521 [17.425]*
min T_i /End year	6	5	5	6	2010	2010	2010	2010
Countries	175	175	175	174	175	175	175	174
Observations	6,336	6,161	6,161	6,309	6,336	6,161	6,161	6,309
Share of ANRR sample	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Panel B: Estimate insignificant (10% significance level)								
Long-Run Democracy Effect	15.637 [9.867]	11.932 [8.071]	8.066 [7.047]	29.168 [17.733]	12.516 [7.386]	3.891 [8.131]	14.293 [11.504]	27.145 [17.309]
min T_i /End year	47	17	17	21	1999	2001	2008	2009
Countries	79	152	152	146	172	172	175	174
Observations	3,713	5,846	5,846	5,873	4,433	4,605	5,824	6,146
Share of ANRR sample	0.59	0.95	0.95	0.93	0.70	0.75	0.95	0.97
Panel C: Estimate below 5% in magnitude								
Long-Run Democracy Effect	n/a	3.918 [7.622]	3.949 [5.670]	2.651 [16.519]	1.160 [6.157]	3.891 [8.131]	-22.917 [28.970]	4.936 [17.275]
min T_i /End year		38	26	28	1991	2001	1994	2000
Countries		97	128	119	149	172	152	171
Observations		4,387	5,325	5,202	3,119	4,605	3,422	4,588
Share of ANRR sample		0.71	0.86	0.82	0.49	0.75	0.56	0.73
Panel D: Estimate outside 90% CI of full sample estimate								
Long-Run Democracy Effect	n/a	1.650 [8.722]	5.718 [6.287]	2.651 [16.519]	1.160 [6.157]	1.411 [8.409]	6.091 [8.090]	2.305 [23.466]
min T_i /End year		41	19	28	1991	2000	2005	1996
Countries		90	149	119	149	172	175	166
Observations		4,112	5,793	5,202	3,119	4,433	5,300	3,908
Share of ANRR sample		0.67	0.94	0.82	0.49	0.72	0.86	0.62
Panel E: Estimate for balanced panel								
Long-Run Democracy Effect	15.637 [9.867]	-5.337 [8.484]	-12.358 [6.899]*	12.843 [23.009]	n/a	n/a	n/a	n/a
min T_i	47	46	46	47				
Countries	79	79	79	78				
Observations	3,713	3,634	3,634	3,666				
Share of ANRR sample	0.59	0.59	0.59	0.58				

Notes: The table presents estimates for the two sample reduction exercises in columns [1]-[4] and [5]-[8], respectively (estimator as indicated). All estimates are based on specifications with four lags of per capita GDP and in case of the 2SLS using four lags of the instrument — these are the preferred specifications by ANRR. Long-run estimates are computed as $\hat{\beta}^{LR} = \hat{\beta} / (1 - \sum_{\ell=1}^4 \hat{\rho}_{i,t-\ell})$, where $\hat{\beta}$ is the estimate on the democracy dummy and the $\hat{\rho}$ are estimates for the lags of per capita GDP (standard errors are computed via the Delta method). Results in Panel A are identical to those in ANRR Tables 2 (2FE, AB, HHK) and 6 (2SLS). The 2FE estimate in column [1] never drops below 5% in magnitude or outside the 90% confidence interval of the full sample estimate. The sample end year reduction strategy in columns [5]-[8] does not lead to a balanced panel like the sample reduction by minimum observation count in columns [1]-[4]. Statistical significance at the 10%, 5% and 1% level are indicated as *, **, and ***, respectively.

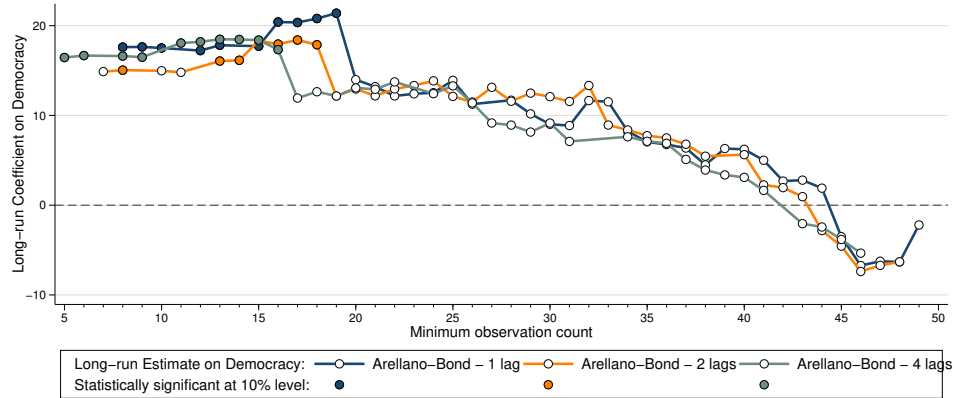
Online Appendix — Not Intended for Publication

A Sample Reduction Exercises – Full Parametric Results

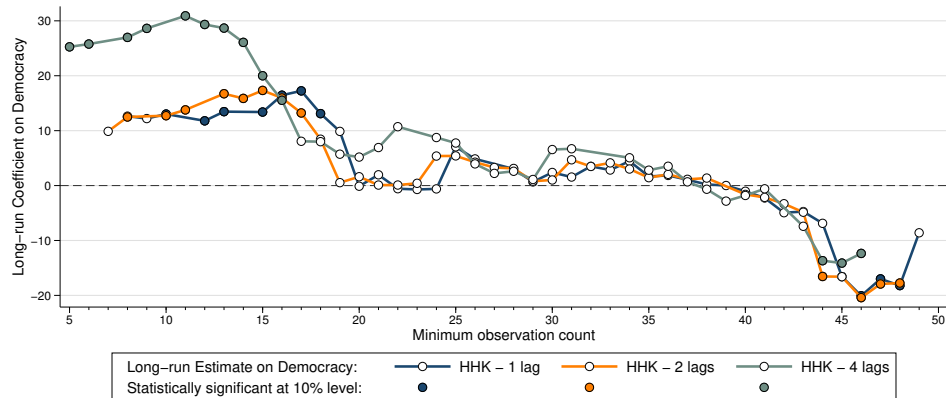
Figure A-1: Sample Reductions by T_i count (FE, AB, HHK)



(a) Within Estimates



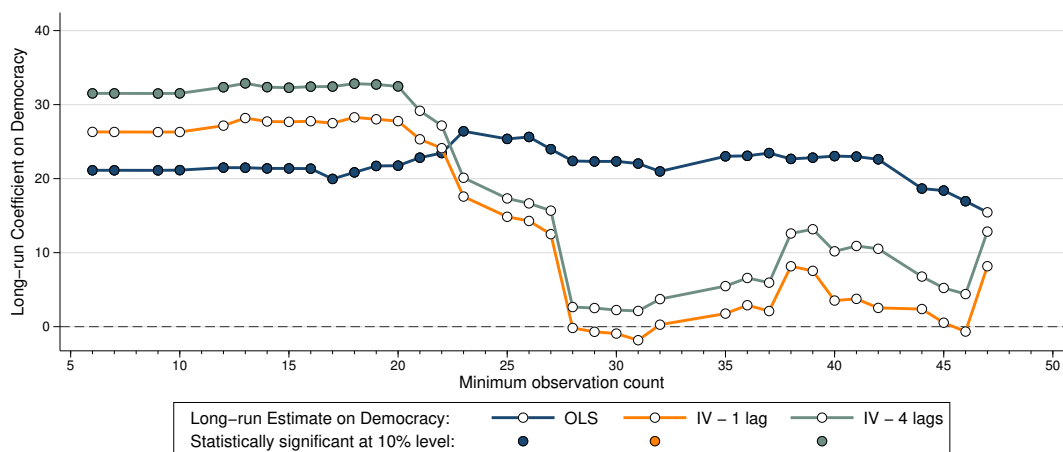
(b) Arellano-Bond Estimates



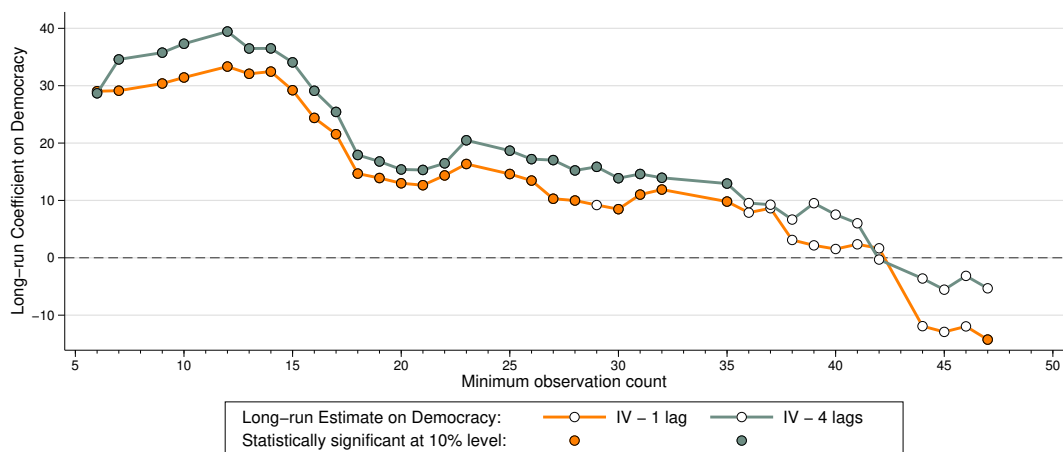
(c) Hahn, Hausman & Kuersteiner (HHK) Estimates

The figure presents long-run estimates for democracy from varying empirical samples for the (a) within, (b) Arellano-Bond, and (c) HHK estimators. See Figure 1 in the maintext for all other details.

Figure A-2: Sample Reductions by T_i count (2SLS, HHK-2SLS)



(a) OLS and 2SLS Estimates



(b) HHK-2SLS Estimates

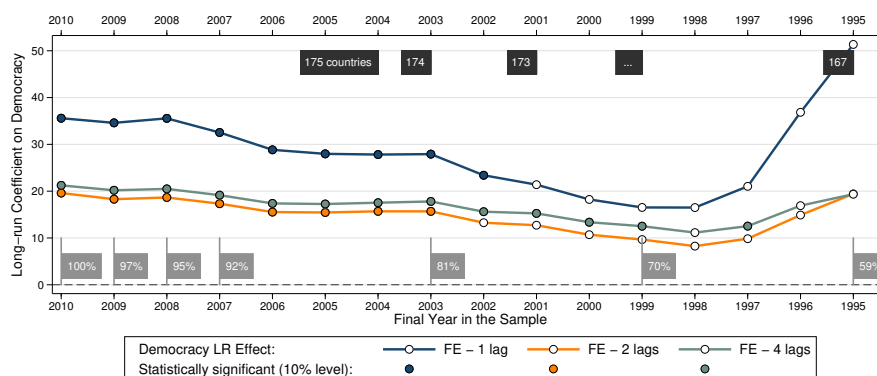
Notes: The figure presents long-run estimates for democracy from varying empirical samples for the (a) 2SLS, and (b) HHK-2SLS estimators. In each plot there are results for 2SLS using one and four lags of regional democratisation waves, respectively; plot (a) in addition reports simple OLS ('within') estimates for reference. The 'left-most' estimates for each specification replicate the results presented in ANRR's Table 6, columns (1) and (2) for single and four lags of the instrument, 2SLS (Panel A) and HHK-2SLS (Panel C). See Figure A-3 for all other details.

Table A-1: Minimal Sample Reduction

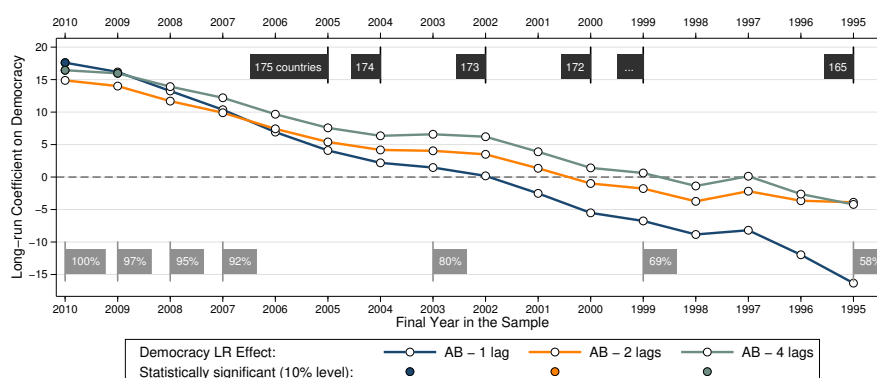
ANRR Reference	Sample reduction by dropping countries (unsystematically)							
	[1] 2FE		[2] AB		[3] HHK		[4] 2SLS	
	Table 2(3)		Table 2(7)		Table 2(11)		Table 6(2)A	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
Long-run effect of democracy	21.240 [7.215]***	15.637 [9.867]	16.448 [8.436]*	12.846 [8.023]	25.032 [10.581]***	9.221 [5.830]	31.521 [17.425]*	28.574 [17.394]
Observations	6,336	3,713	6,161	6,113	6,161	6,100	6,309	6,249
Obs dropped	none	2,623	none	48	none	61	none	60
dto. (in %)	0%	41.4%	0%	0.78%	0%	0.99%	0%	0.95%
Countries	175	79	175	172	175	171	174	171
Countries dropped	none	96	none	ARM, AZE, SLB	none	AZE, BLR, ERI, HTI	none	TKM, UKR, UZB
dto. (in %)	0%	54.9%	0%	1.7%	0%	2.3%	0%	1.7%

Notes: The table presents full sample estimates in columns marked (a) and reduced sample estimates in columns marked (b) for the 2FE, AB, HHK and 2SLS estimators. In a purposeful exercise I determine (via trial and error) the minimum set of countries that need to be dropped from the sample for the long-run democracy estimate to turn statistically insignificant (AB, HHK and 2SLS only). The countries dropped are indicated in the bottom of the table — for instance, the 2SLS estimate turns insignificant if Turkmenistan (TKM; 20 sample years in autocracy, none in democracy), the Ukraine (UKR; 3, 17), and Uzbekistan (UZB; 20, 0) are dropped from the sample. Statistical significance at the 10%, 5% and 1% level are indicated as *, **, and ***, respectively.

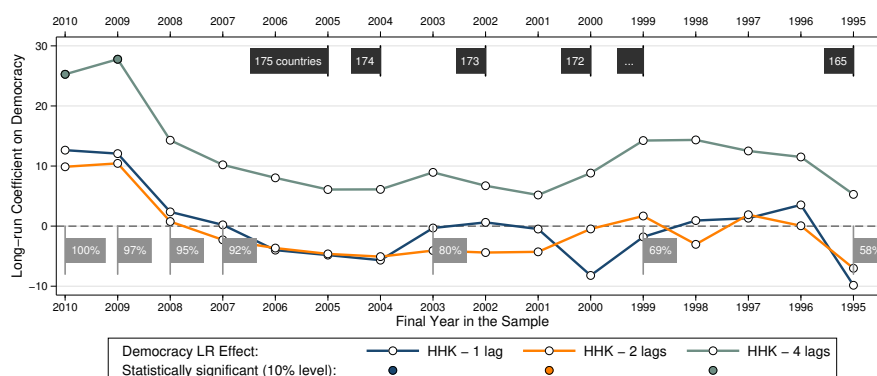
Figure A-3: Sample Reductions by end year (2FE, AB, HHK)



(a) Within Estimates



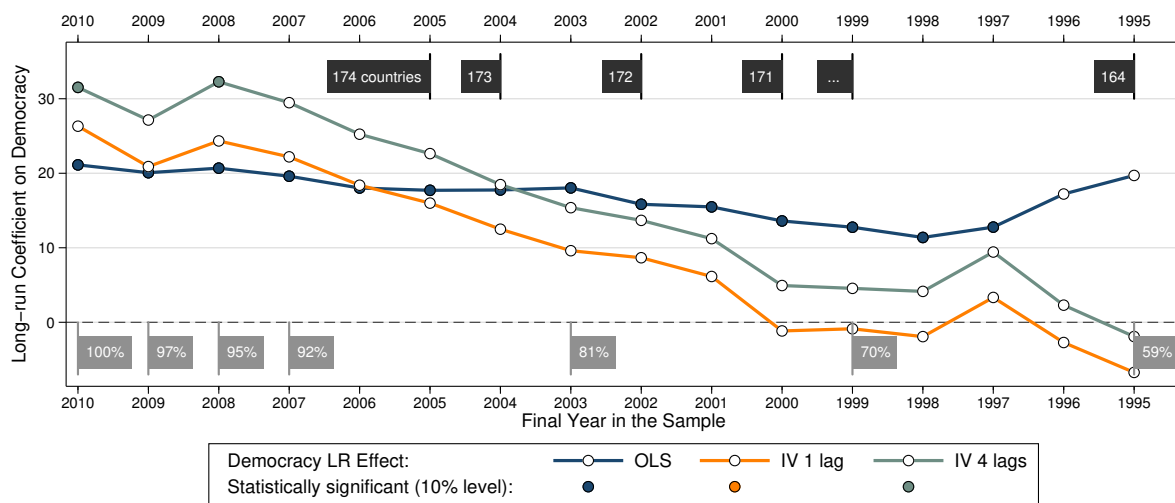
(b) Arellano-Bond (AB) Estimates



(c) Hahn, Hausman & Kuersteiner (HHK) Estimates

Notes: The figure presents long-run estimates for democracy from varying empirical samples for the (a) within, (b) AB, and (c) HHK estimators. The x -axis in each plot indicates the sample end year. A filled (white) circle marker indicates that the long-run coefficient is statistically (in)significant at the 10% level. Grey labels indicate the share of full sample observations, black labels the number of countries — for each estimator these numbers are based on the most restrictive 4-lag model. In each plot there are results for three specifications: with one lag of per capita GDP (in logs), two lags and four lags, corresponding to the specifications in ANRR's Table 2, columns (1)-(3) for FE, (5)-(7) for AB, and (9)-(11) for HHK.

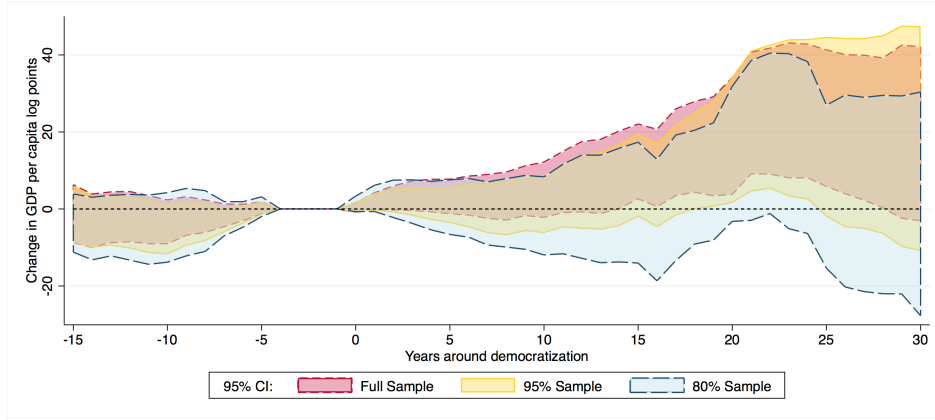
Figure A-4: Sample Reductions by end year (2SLS)



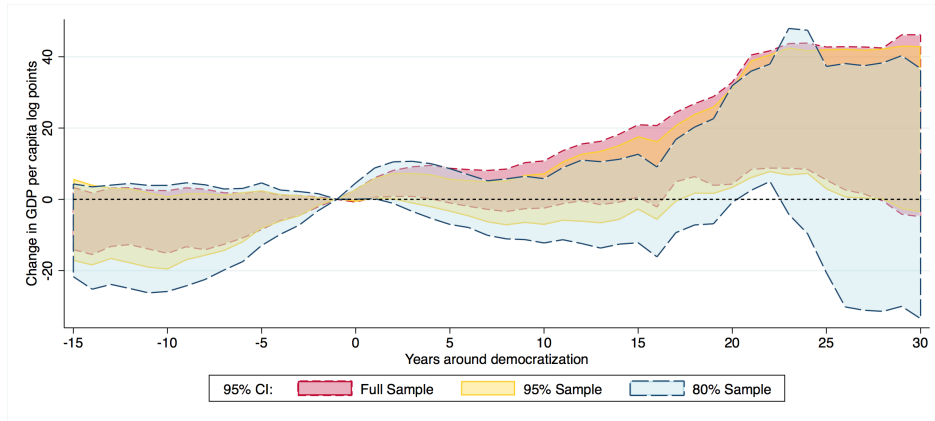
Notes: The figure presents long-run estimates for democracy from varying empirical samples for the 2SLS estimator. In addition to results for the the ANRR 2SLS specification with one and four lags of GDP I include the FE-4 lags specification for illustration (labelled 'OLS'). The former correspond to the results in ANRR's Table 6, Panel A, Columns (1) and (2). See Figure A-3 for all other details.

B Sample Reduction Exercises – Semi-Parametric Results

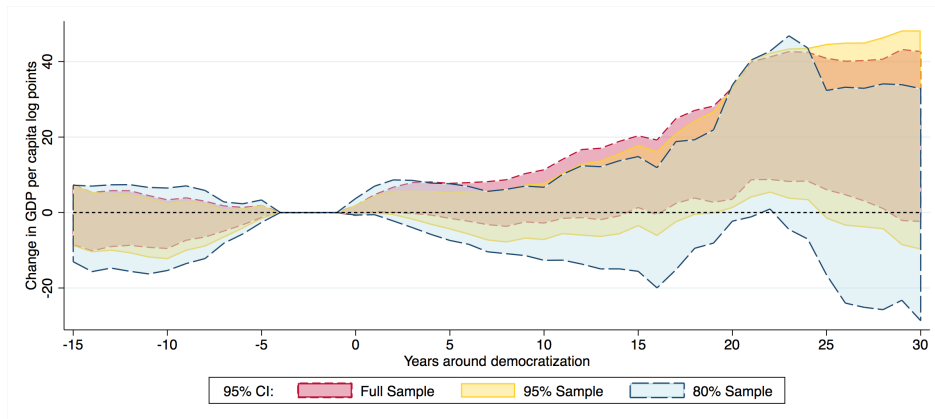
Figure B-1: Sample Reductions by T_i count — Semi-parametric models



(a) Adopting Linear Regression



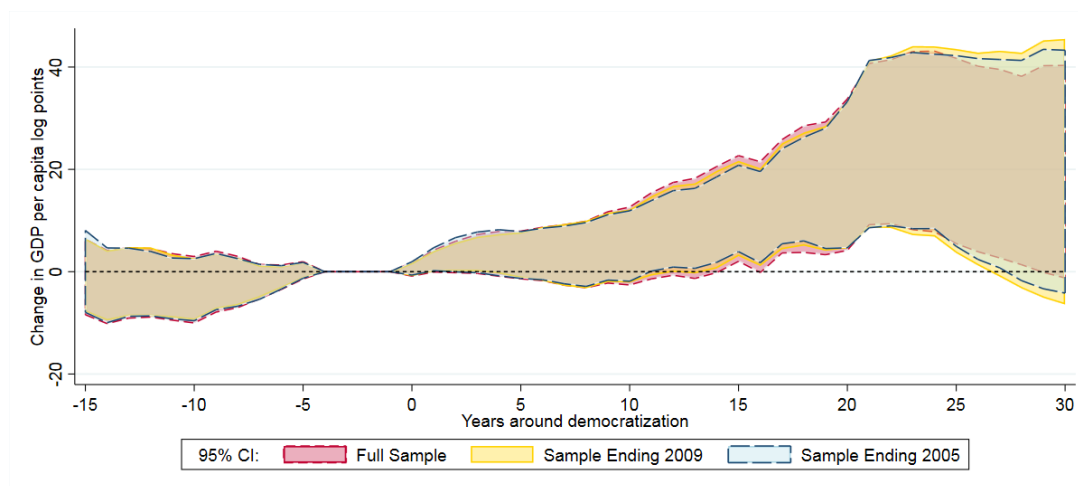
(b) Adopting Inverse Propensity Score Reweighting



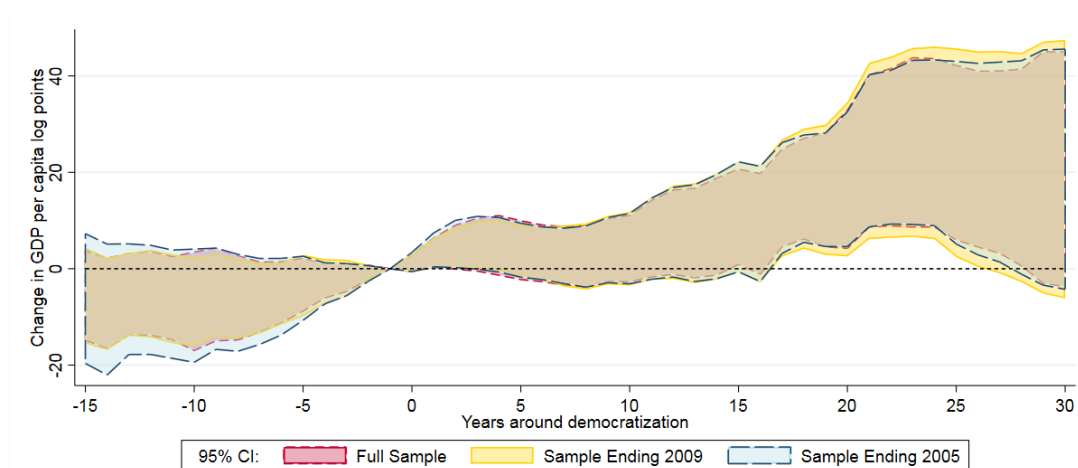
(c) Adopting Both Strategies

Notes: The figure presents the estimates for the ‘over time’ effect of democracy on per capita GDP using three different methods to estimate the relationship semi-parametrically. In each case the red shading indicates the full ($n = 3,029, N = 126$) sample 95% confidence interval (CI), and yellow and blue shading the *reduced-sample* CIs: in the former I drop 5% of observations ($n = 2,871, N = 102$), in the latter 20% ($n = 2,415, N = 72$) — results for a 10%-reduced sample ($n = 2,735, N = 91$) are qualitatively identical to the 5%-reduced sample. As above the criterion for exclusion is the number of a country’s time series observations (starting with the smallest samples).

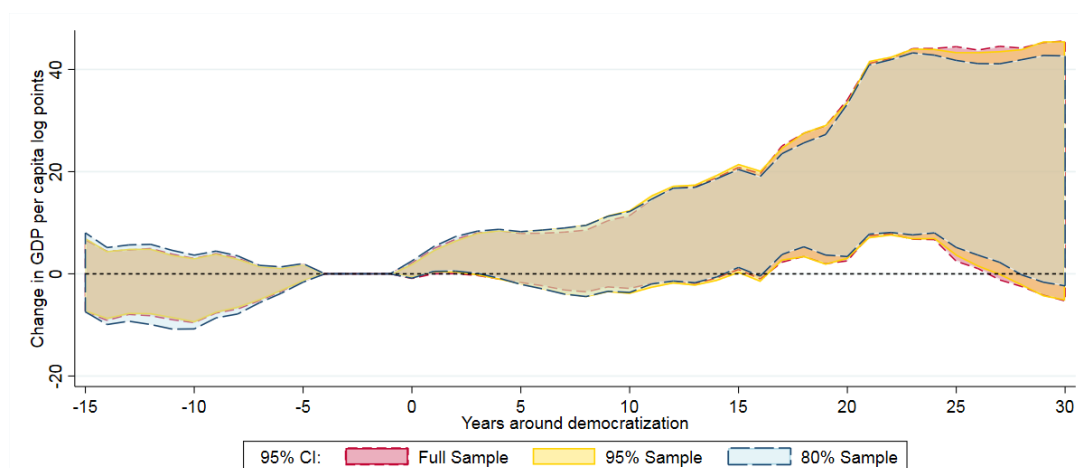
Figure B-2: Sample Reductions by end year — Semi-parametric models



(a) Adopting Linear Regression



(b) Adopting Inverse Propensity Score Reweighting



(c) Adopting Both Strategies

Notes: The figure presents the estimates for the ‘over time’ effect of democracy on per capita GDP using three different methods to estimate the relationship semi-parametrically. In each case the red shading indicates the full ($n = 3,029$, $N = 126$) sample 95% confidence interval (CI), and yellow and blue shading the *reduced-sample* CIs: in the former I drop 5% of observations ($n = 2,871$, $N = 102$), in the latter 20% ($n = 2,415$, $N = 72$) — results for a 10%-reduced sample ($n = 2,735$, $N = 91$) are qualitatively identical to the 5%-reduced sample. As above the criterion for exclusion is the number of a country’s time series observations (starting with the smallest samples).

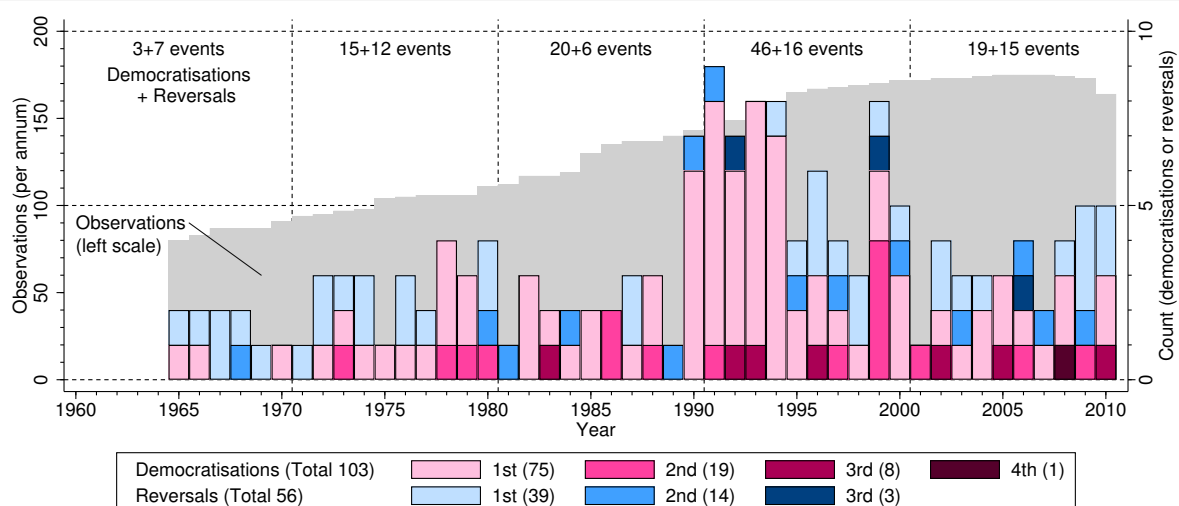
C Sample Reduction Exercises – Sample Makeup

Table C-1: Regression Sample (AB/HHK 4-lag specification)

obs	Transitioned into/out of democracy							Never a democracy			Always a democracy			
5								QAT						
6								LBY						
8								KWT						
9								IRQ						
11								MDV						
12								BIH						
13								KHM						
14								ERI						
15	DJI	HTI									PLW			
16	ARM	HRV	SLB					AZE	BLR	KAZ	CZE	LTU	MKD	
								YEM			POL	SVN		
17	RUS													
18	LBN							TZA						
19	UKR							TKM	UZB					
20	GIN	KGZ									NAM			
21								AGO	GNQ	TJK				
22	SVK							LAO	VNM					
24								BHR	UGA	WSM				
25	BTN	CPV	ETH	MNG				BRN	TON					
26	ALB	BGR	COM	EST	MDA	MOZ	ROM				CHE	LCA		
27											KNA	VUT		
28											NZL			
29	GRD										ATG	BLZ	DMA	
30	SUR										MUS			
31								JOR			CYP	KIR	VCT	
34								SYC						
35											PNG			
36	GNB							CUB	SWZ			DEU	IRL	MLT
37											BHS			
38	BGD							SAU						
39	MLI													
40	FJI	GMB						IRN			JAM			
41	GEO	HUN	LVA											
43											BRB			
44	GUY	LSO									BWA			
45	ZWE							OMN	SIN	TUN				
46	ARG	BDI	BEN	BFA	BOL	BRA	CAF	CHN	CMR	DZA	AUS	AUT	BEL	
	CHL	CIV	COG	DOM	ECU	ESP	GHA	EGY	GAB	MAR	CAN	COL	CRI	
	GRC	GTM	HND	IDN	KEN	KOR	LBR	MYS	RWA	SYR	DNK	FIN	FRA	
	MDG	MEX	MRT	MWI	NER	NGA	NIC	TCD	TGO	ZAR	GBR	IND	ISL	
	NPL	PAK	PAN	PER	PHL	PRT	PRY				ISR	ITA	JPN	
	SDN	SEN	SLE	SLV	THA	TUR	URY				LKA	LUX	NLD	
	VEN	ZAF	ZMB								NOR	SWE	TTO	
											USA			

Notes: The three samples contain 80, 46, and 49 countries, respectively. The analysis is based on the AB/HHK samples; for the 2SLS estimates the minimum observation count is typically increased by one observation.

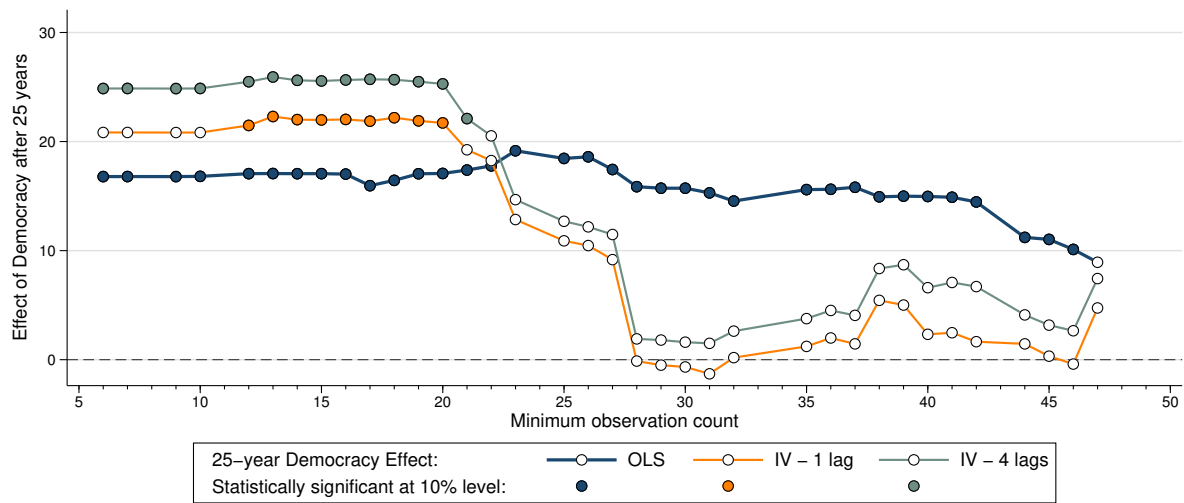
Figure C-1: Histogram – Full Sample – Democratisation and Reversal



Notes: The figure presents the distribution of democratisation events (shades of pink) and reversals to autocracy (shades of blue) over the 1965 to 2010 period — the sample employed here is for the AB/HHK specification with four lags of GDP (hence the omission of 1960-64). The shades distinguish *first* democratisations or reversals (during the sample period) from repeated events; around 27% of all democratisations are for countries which previously had already experienced an episode of democratisation during the sample period. One country – Thailand – democratised four times. For reversals this figure is 30%. Events during the 1990s account for 39% of all events, for comparison the figures for the 1980s and 2000s are 17% and 15%, respectively. The figure also indicates the number of observations per sample year (in grey, left axis).

D Alternative Definitions of the ‘Long-run’

Figure D-1: Sample Reductions — 2SLS estimator



(a) Static Models

Notes: This figure provides sample reduction results for the ANRR 2SLS specifications with one and four lags of (waves of) democracy alongside results for an OLS model. In contrast to the results presented in the maintext, where I focus on the ‘long-run’ computed from the dynamic specification, these results report the outcomes for income pc after 25 years of democracy (computation as defined by ANRR).