

Democracy, Growth, and Heterogeneity*

Markus Eberhardt[†]

School of Economics, University of Nottingham, UK

Centre for Economic Policy Research, UK

This Version: January 8, 2021

Abstract: I revisit the causal relationship between democracy and growth as recently studied in Acemoglu, et al (2019, ANRR). These authors establish a 16-32% long-run effect of permanent democratic transition on income per capita. My analysis centres on two implicit assumptions in ANRR's implementations and in the existing literature on the democracy-growth nexus: (i) a common democracy-growth relationship across countries alongside common GDP dynamics; and (ii) the absence of strong cross-section dependence in the panel (e.g. spillovers or global shocks with heterogeneous impact). Adopting novel Difference-in-Difference estimators I relax these assumptions to show that democracy does have a positive average long-run effect on per capita income; in contrast to the estimates of ANRR this finding is robust to substantial sample reductions. My average long-run effect of democracy is less than half that of ANRR. Further analysis investigates the patterns of the heterogeneous 'democratic dividend' across countries.

JEL Classification: O10, P16

Keywords: Democracy, Growth, Political Development, Difference-in-Difference Estimator, Interactive Fixed Effects

*I am indebted to Matteo Cervellati, Marc Chan, Rodolphe Desbordes, Jonathan Dingel, Doug Gollin, Brian Knight, Simon Kwok, Gerard Padro i Miquel, Andrea Presbitero, Daniel Seidmann, Jon Temple, Daniel Treisman, Nicolas van de Sijpe, Dietz Vollrath, and Joakim Westerlund for comments and suggestions which helped improve the paper significantly. Seminar/session participants at Lund, Nottingham, Middlesex, and the Annual International Conference of the Research Group on Development Economics in Berlin provided useful feedback.

[†]Correspondence: Markus Eberhardt, School of Economics, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, UK; Tel: +44 115 8468416, Email: markus.eberhardt@nottingham.ac.uk

1 Introduction

Until recently the empirical literature on the long-term economic effects of democratisation provided mixed evidence on the significance and magnitude of any ‘democratic dividend’ (e.g. Barro, 1996; Gerring, et al, 2005; Papaioannou and Siourounis, 2008; Knutsen, 2013). A 2019 study by Acemoglu, et al (henceforth ANRR) now offers “significant and robust positive” (abstract) evidence for a link between democracy and growth,¹ accounting for the dynamics of GDP and devising an instrumentation strategy for causal identification. These 2SLS estimates support the findings from two-way fixed effects (2FE), the Arellano and Bond (1991, AB) and the Hahn, et al (2001, HHK) estimators: the common message is that the long-run effect of democratisation is sizeable, an increase in per capita GDP of 20% or more.

ANRR point to three choices/assumptions which are instrumental in capturing the causal effect of democracy on growth. First, their consolidated dichotomous measure for democracy;² second, the importance of modelling the dynamics of GDP to account for its movements before, during, and after democratisation; and third, the identification strategy of the 2SLS implementation to address concerns over reverse causality. One can add further *implicit* assumptions, namely that the democracy-growth relationship and GDP dynamics are the same for all countries (parameter homogeneity), and that there is no bias due to time-varying unobserved heterogeneity (cross-sectional independence).

This note investigates the empirical consequences of relaxing these additional assumptions, while adopting an identification strategy which allows me to interpret the findings as causal. To the best of my knowledge there are no existing studies on the economic effect of democratisation which consider these important aspects alongside dynamics, definitions of democracy, and causality. One can draw on arguments for a democratic legacy (Gerring, et al, 2005) or threshold levels in economic or human development (Madsen, et al, 2015; ANRR) as necessary conditions for a positive democracy-growth nexus to motivate *parameter heterogeneity*. The presence of *strong cross-section dependence* can arise from spillovers or global shocks with heterogeneous impact across countries, such as the country-response to the Global Financial Crisis (GFC), and the panel econometric literature has spent two decades emphasising the bias in empirical estimates if this property is ignored (e.g. Andrews, 2005; Pesaran, 2006; Bai, 2009).³ My analysis is in three parts.

First, sensitivity of empirical results to (minor) sample changes can point to the presence of observed and unobserved heterogeneities. I employ two rule-based sample reduction strategies to investigate this in the ANRR sample: (i) out of concern over influential outliers I drop countries on

¹I follow ANRR and use ‘growth’ as a short-hand for per capita GDP; see Eberhardt and Teal (2011) for a detailed discussion of dynamic growth empirics.

²See Appendix E for an analysis of Cheibub, et al’s (2010; CGV) democracy indicator.

³Parameter heterogeneity and cross-section dependence invalidate the identifying assumptions inherent in ANRR’s 2SLS estimation (Pesaran and Smith, 1995) and their AB/HHK implementations (Sarafidis and Robertson, 2008).

the basis of their observation count, first omitting those with five observations, then those with six, etc.; and (ii) motivated by the presence of the GFC near the end of ANRR's sample — the most significant global macroeconomic shock since the 1930s — I omit observations on the basis of the sample end year, dropping first 2010, then 2009, etc. My results⁴ indicate that the magnitude and statistical significance of ANRR's estimates are highly sensitive to sample selection: depending on the exercise and implementation the long-run estimates turn statistically insignificant when between 3% and 8% of the over 6,100 observations are dropped.

Second, adopting heterogeneous panel estimators for policy evaluation (Chan and Kwok, 2018) I demonstrate that the average long-run effect of democracy is statistically significant positive and robust to substantial sample reductions. My results indicate that the average effect of democratisation is 8-10%, thus around half that in ANRR.

Third, these results do not merely differ by degree: while a policy implication derived from ANRR's analysis would see a clear case for democratisation in *any* of their sample countries, my results suggest that such optimism is misplaced. Democracy has not universally brought higher income, and a positive outcome may be subject to necessary conditions. The final section of this letter studies existing suggestions for factors driving a heterogeneous democracy-growth nexus.

2 Sample Reduction Exercises

In this section I discuss results from two sample reduction exercises presented in Figure 1 and Table 1.⁵ ANRR adopt a variety of empirical implementations for a model given by:

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

where y is log GDP per capit (multiplied by 100), Dem is their democracy indicator, α_i and γ_t are country and time dummies, and ε is the error term. To allow for a causal interpretation they devise an instrumentation strategy based on regional waves of democratisation. The data used is taken from the World Development Indicators, the Polity IV project and a number of additional democracy indices, described in Appendix A.

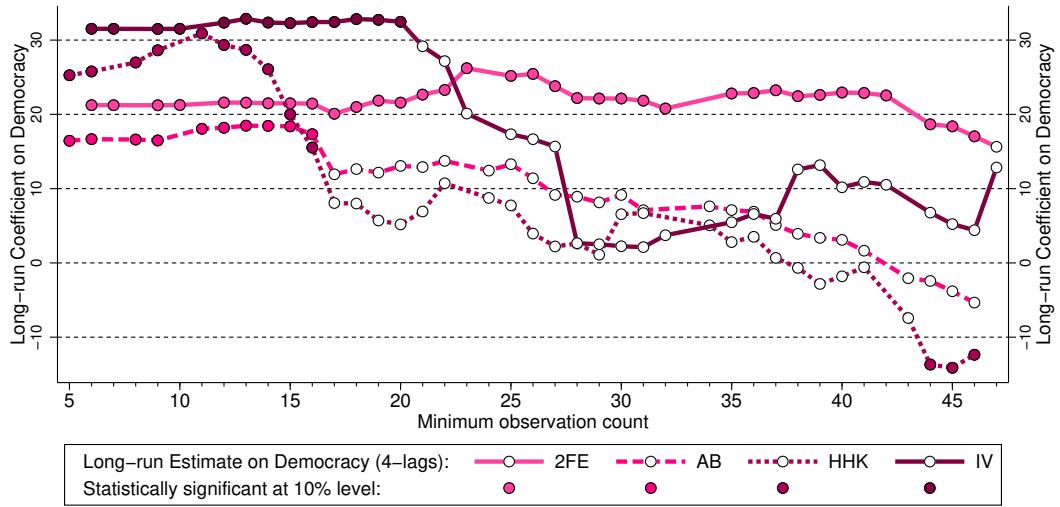
⁴Results for semi-parametric models are presented in Appendix D.

⁵In Appendix H I speculate about one potential source of the patterns observed.

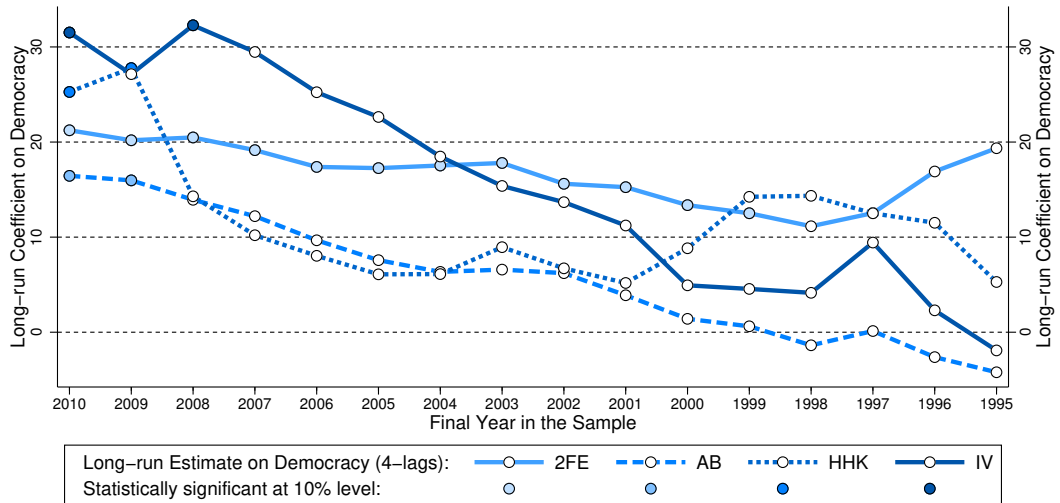
2.1 Sample reduction by minimum observation count

I begin with the strategy which drops countries by their observation count.⁶ Figure 1, Panel (a) presents the full and reduced sample results for the FE, AB, HHK and 2SLS estimators.⁷ The y -axis indicates the long-run effect of democracy on per capita GDP and the x -axis the minimum observation count for inclusion in the sample. The left-most estimates correspond to the full sample results, the right-most to the estimates for a balanced panel; a filled coloured (white) circle indicates statistical (in)significance at the 10% level.

Figure 1: Sample Reduction Results – pooled specifications



(a) Sample Reduction by T_i count



(b) Sample Reduction by End Year

⁶One concern is that though ‘small- T ’ countries may only account for a small share of observations they may represent a disproportionate share of democratisation events. As Appendix Figure B-1 demonstrates, this is not the case.

⁷Results for one and two lags are presented in an Appendix. I also provide 2SLS estimates for an alternative construction of the long-run estimate.

Table 1: Sample Reduction Estimates – pooled specifications

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	16.448	25.268	31.521	21.240	16.448	25.268	31.521
Standard error	[7.215]***	[8.436]*	[10.869]**	[17.425]*	[7.215]***	[8.436]*	[10.869]**	[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
Panel B: Estimate insignificant (10% significance level)								
Long-Run Democracy Effect	15.637	11.932	8.066	29.168	12.516	3.891	14.293	27.145
Standard error	[9.867]	[8.071]	[7.047]	[17.733]	[7.386]	[8.131]	[11.504]	[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 full 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	3.949	2.651	1.160	3.891	-22.917	4.936
Standard error		[7.622]	[5.670]	[16.519]	[6.157]	[8.131]	[28.970]	[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 full 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	5.718	2.651	1.160	1.411	6.091	2.305
Standard error		[8.722]	[6.287]	[16.519]	[6.157]	[8.409]	[8.090]	[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 full 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	-5.337	-12.358	12.843	n/a	n/a	n/a	n/a
Standard error	[9.867]	[8.484]	[6.899]*	[23.009]				
min T_i	47	46	46	47				
Countries	79	79	79	78				
Observations	3,713	3,634	3,634	3,666				
Share of full ANRR sample	0.59	0.59	0.59	0.58				

For the 2FE estimator the sample reduction has virtually no impact on the democracy estimates, with the exception of the balanced panel result. The patterns for the AB and HHK estimates are very different: both decline and turn insignificant when the minimum observation count is 17 and thereafter fall towards and beyond zero. Table 1 indicates that once 5% of observations are dropped the AB and HHK estimates are statistically insignificant and reduced by a quarter and two-thirds, respectively. Democracy estimates based on the 2SLS estimator initially maintain a stable level above

30%, but turn insignificant once countries with fewer than 21 observations are omitted (7% of the full sample).⁸ The magnitude of 2SLS estimates drops rapidly, falling below 5% in magnitude if 18% of observations are dropped.

2.2 Sample reduction by sample end year

Figure 1, Panel (b) presents the long-run effects of democracy when observations are omitted by sample end year (x -axis, in reverse chronological order). As before the 2FE estimates are 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 2010 and 2009 (5% of observations) are omitted, thereafter declining and diverging. The 2SLS estimates display curious patterns in the aftermath of the GFC: omitting only 2010 (3% of observations) yields a statistically insignificant long-run coefficient for democracy. Omitting 2010 and 2009 restores the full sample coefficient in terms of magnitude and statistical significance, whereas the omission of further end years always yields statistically insignificant estimates.

2.3 Conclusions from these exercises

Two aspects are worth emphasising comparing these findings to the results in ANRR: first, the parity between results for the 2FE estimator on the one hand, and the AB, HHK and 2SLS estimators on the other is not given in my sample reduction exercises; second, all estimators intended to address endogeneity concerns show rapidly declining long-run growth implications of democracy as the sample is reduced.

3 Heterogeneity

3.1 Heterogeneous treatment effects

This section studies the impact of observable and unobservable heterogeneity on empirical estimates of the democracy-growth nexus. My model builds on the macro panel literature which has emphasised heterogeneous parameters across panel members (Pesaran and Smith, 1995) and, more recently, the presence of strong cross-section dependence (e.g. Bai, 2009; Pesaran, 2006). The latter is distinct from weaker forms of dependence, such as spatial correlation, and if ignored can lead to serious bias in the estimated coefficients on observable variables (Andrews, 2005). The literature has taken to specifying unobserved common factors with associated heterogeneous factor loadings to capture this strong dependence.⁹

⁸If we focus 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 (<1% of observations) — see Appendix Table C-1.

⁹For detailed discussions of how to motivate and address observed and unobserved heterogeneity in the context of cross-country growth empirics see Eberhardt and Teal (2011).

The most recent contributions to this literature have built bridges to the literature on policy evaluation using the synthetic control methodology (Xu, 2017) and difference-in-difference specifications (Gobillon and Magnac, 2016; Chan and Kwok, 2018). What distinguishes these latest approaches from their canonical predecessors is the adoption of a common factor setup to address three challenges to identification in these popular methods: (i) the presence of uncommon trends prior to the policy intervention, (ii) endogenous selection into ‘treatment’, and (iii) the possibility that, following the policy change, treated and control samples are affected by common shocks with heterogeneous impact (e.g. the differential country effect of the GFC).

My implementation follows the spirit of Chan and Kwok’s (2018) estimator but adopts cross-section averages (Pesaran, 2006) instead of estimated factors (Bai, 2009) due to the strongly unbalanced panel sample. Crucially, my setup to investigate the long-run effect of democracy allows for correlation between the unobserved determinants of growth (institutions, absorptive capacity, etc.) and selection into democracy. For the sample of countries which experienced variation in the ANRR democracy dummy I specify a dynamic ‘CS-DL’ (cross-section-augmented distributed lag; Chudik, et al, 2016) model, which is estimated separately for each country i using least squares:

$$y_{it} = \alpha_i + \theta_i \text{Dem}_{it} + \mathbf{\Gamma}'_i \mathbf{X}_{it} + \sum_{\ell=0}^{p-1} \omega_{i\ell}^D \Delta \text{Dem}_{i,t-\ell} + \sum_{\ell=0}^{p-1} \omega_{i\ell}^{X'} \Delta \mathbf{X}_{i,t-\ell} \quad (2)$$

$$+ \sum_{\ell=0}^{p_{\bar{y}}} \delta_{i\ell}^y \bar{y}_{t-\ell} + \sum_{\ell=0}^{p_{\bar{X}}} \delta_{i\ell}^{X'} \bar{\mathbf{X}}_{t-\ell} + \varepsilon_{it},$$

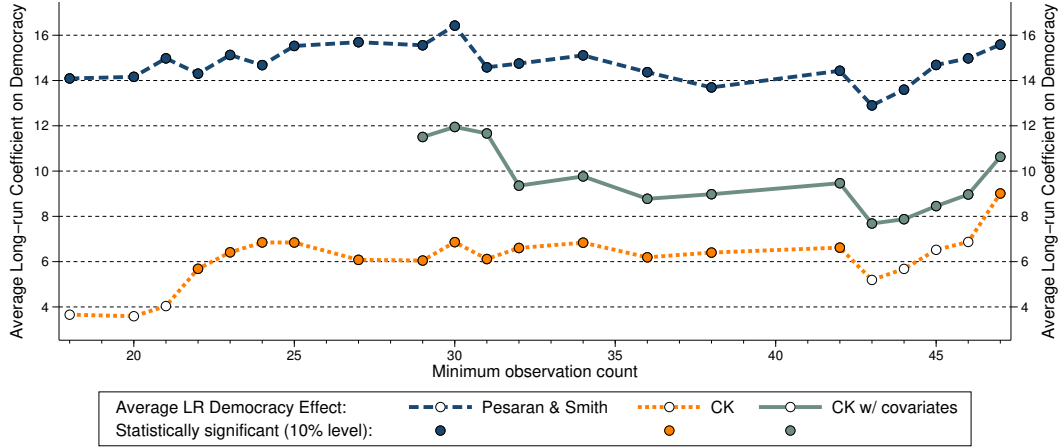
where y is per capita GDP (in logs and multiplied by 100), Dem is the democracy dummy, and \mathbf{X} is a set of additional controls (gross investment/GDP and trade openness). The two sums in the first line capture the short-run effects, θ_i and $\mathbf{\Gamma}_i$ represent the ATET long-run effects of democracy and additional controls on per capita GDP, respectively. \bar{y} and $\bar{\mathbf{X}}$ are cross-section averages of the observed variables *but for those countries which never experienced democracy during the sample period* (control group). These cross-section averages capture the presence of uncommon and stochastic trends (Pesaran, 2006; Chan and Kwok, 2018).¹⁰ This specification is convenient since the long-run coefficient can be estimated in a single step rather than two like in an error-correction or the ANRR specifications. I use $p_{\bar{y}} = 0$ and $p = p_{\bar{X}} = \text{int}(T^{1/3}) = 3$ (Chudik, et al, 2016), where T is the time dimension of the panel, and adopt robust regression (Hamilton, 1992) to compute outlier-robust means; inference is based on non-parametric standard errors (Pesaran, 2006).

Figure 2, Panel (a) presents the results from dynamic specifications of three heterogeneous parameter models for the sample reduction exercise by country observation count. The estimates from the model ignoring any potential factor structure and thus selection, uncommon trends and/or common shocks with heterogeneous impact (labelled ‘Pesaran & Smith’) are substantially higher than

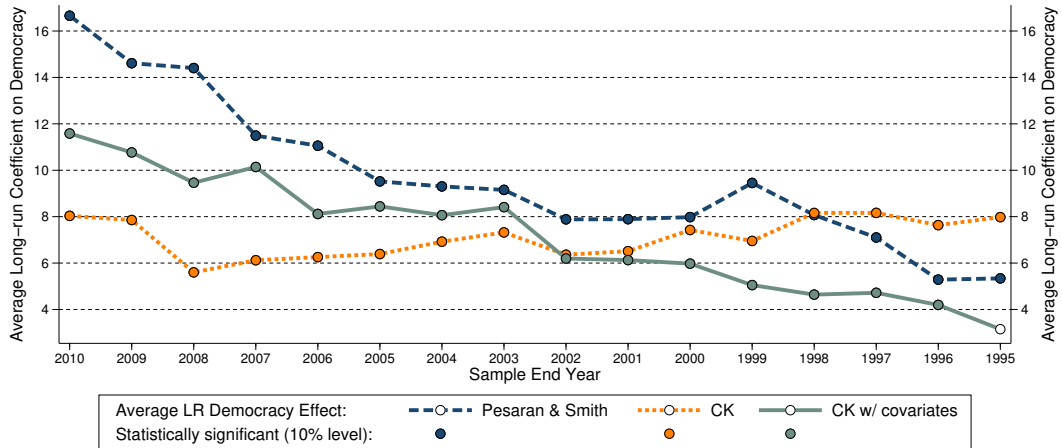
¹⁰Observed covariates \mathbf{X} are not included in the ‘plain vanilla’ implementation — the covariate cross-section averages from the control sample, $\bar{\mathbf{X}}$, are always included.

the two Chan & Kwok models, suggesting these aspects are ignored at one's peril. The plain vanilla Chan & Kwok model accounting for these distortions (in orange) yields long-run effects of democracy between 6 and 8%. The model with additional covariates lifts the estimates to 8-10%.¹¹ The sample reductions here are substantially *larger* than those in my robustness analysis of ANRR's models, yet still provide remarkably robust results.¹²

Figure 2: Sample Reduction Results — heterogeneous parameter specifications



(a) Sample reduction by T_i count



(b) Sample reduction by end year

The results in panel (b) for the sample reduction by end year indicate a fairly unchanged estimate for 'plain vanilla' CK and a declining one for the specification with additional covariates, for which all but the final estimate presented are statistically significant.

The above setup is somewhat awkward, since it ignores that some countries switched repeatedly between democracy and autocracy. I estimate the average long-run democracy effect limiting

¹¹Since this model estimates an additional $2p + 2$ parameters the minimum feasible sample is $T = 29$.

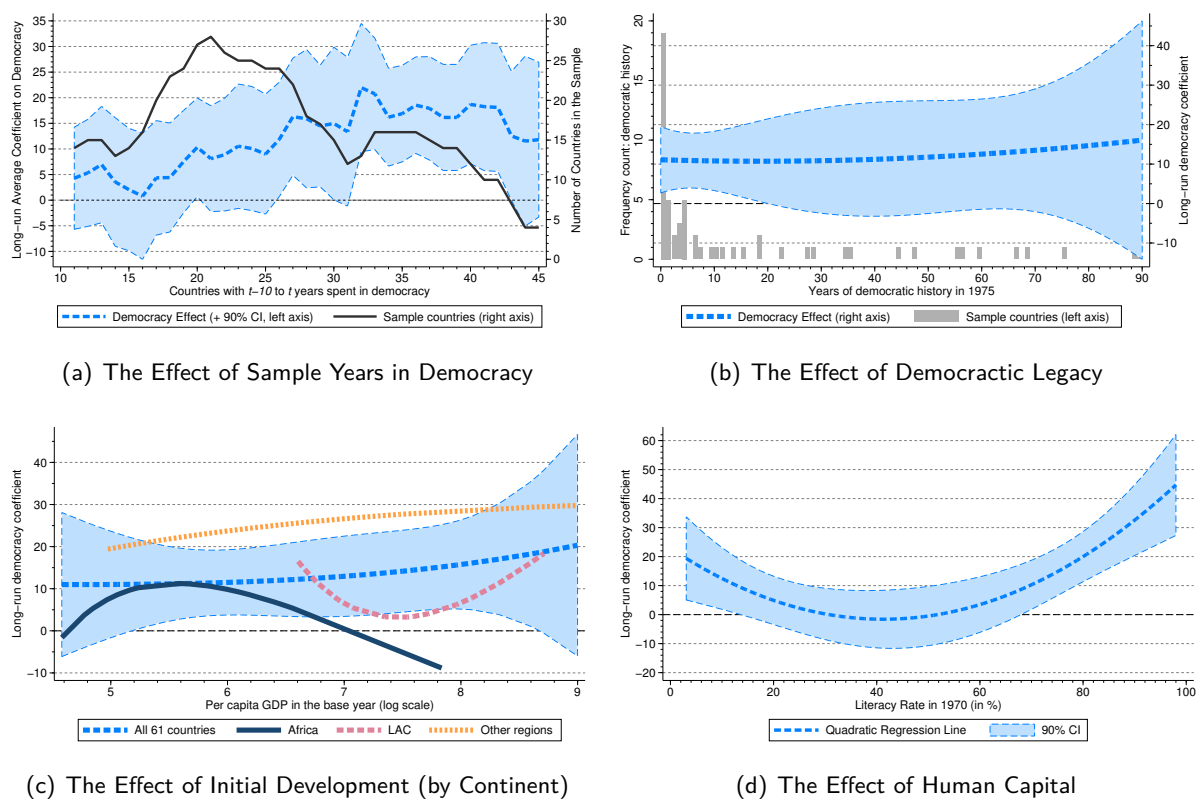
¹²A minimum observation count $T_{min} = 30$ excludes 13% of observations from countries which experienced democratisation, $T_{min} = 40$ 26% and $T_{min} = 47$ 43%.

the sample to countries which only transitioned into democracy *once* during the sample period. Appendix F suggests the exclusion of the repeated regime switchers moderately *raises* the average long-run effect of democracy.

3.2 Sources of Heterogeneity

Can we offer any insights beyond the above *average* effects of democracy? First, I assume that different length of ‘treatment’ (years in democracy) results in heterogeneous long-run estimates across countries. In Figure 3, Panel (a) I run separate regressions for country groups defined by the number of years a country spent in democracy. The band for each group is eleven years, increasing in single-year steps. The maximum year for each band is printed along the *x*-axis of the plot. The dashed line represents the robust mean estimate for the democracy effect, the shaded area its 90% CI. I indicate the sample size with a solid line. This analysis suggests that a positive and statistically significant effect of democracy could take over 25 years to materialise.

Figure 3: Analysing heterogeneity



Second, I gauge the significance of a democratic legacy, proxied by the years spent in democracy between 1800 and 1975 (to maximise data availability). Gerring, et al (2005) among others argue that political regimes are historical legacies, with cumulative effects of institutions (only) coming to bear over long time horizons. Panel (b) shows a fractional polynomial estimate for the

relationship between the long-run democracy coefficient and democratic legacy (dashed line, shaded 90% CI), together with a histogram for the latter. There is no clear advantage of democratic legacy, in fact, the 19 countries with no democratic legacy have a statistically significant long-run effect of democracy around 8%.

Third, I use a fractional polynomial regression of the long-run democracy coefficient on base year GDP (in logs). Panel (c) shows the resulting plot using a dashed line and shaded 90% CI: no matter whether countries were initially rich or poor, the average long-run democracy effect is 10-15% — I ignore the tails due to small numbers of observations. However, looking at separate geographical regions ('Africa' in black, 'Latin America and the Caribbean' (LAC) in red, and 'others' in orange) yields very different patterns for countries with similar base year GDP (log values 6.6 to 8): low and declining for Africa, moderately high and U-shaped for LAC, and high and moderately rising for 'other' regions. This finding needs to be interpreted with caution, since the way the data are sliced smacks of 'geographic determinism'.

Finally, in Panel (d) I study the relationship between human capital endowment and the long-run democracy effect: I adopt literacy data from Madsen, et al (2015) for 1970 to maximise sample coverage. I fit a quadratic regression line to reveal some evidence for a nonlinearity, suggesting that countries with high literacy rates were able to achieve a higher democratic dividend.

4 Concluding Remarks

In this letter I study the long-run causal relationship between democracy and growth. My robustness checks for the seminal study by ANRR suggest that pooled models — assuming a common democracy-growth nexus and common GDP dynamics across countries as well as cross-section independence — are fragile to even minor changes in the sample: their parametric models yield insignificant long-run democracy estimates when just 3-8% of observations are omitted in my rule-based exercises and fewer than 1% when I purposefully seek out influential countries (Appendix C).

I then adopt heterogeneous parameter regressions, which account for strong cross-section dependence, and provide evidence for a significant, positive and robust democracy-growth nexus. My long-run coefficients are roughly half those suggested by ANRR. Exploring the heterogeneity of the democracy-growth effect I contrast longer in-sample experience of democracy with democratic legacy, human capital endowment, and differences in the initial level of development.

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] **Andrews, Donald (2005)** 'Cross-section regression with common shocks.' *Econometrica* 73(5): 1551-1585.
- [3] **Arellano, Manuel, and Stephen 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.
- [4] **Bai, Jushan (2009)** 'Panel data models with interactive fixed effects.' *Econometrica* 77(4): 1229-1279.
- [5] **Barro, Robert (1996)** 'Democracy and growth.' *Journal of Economic Growth* 1(1): 1-27.
- [6] **Chan, Marc, and Simon Kwok (2018)** 'Difference-in-Difference when Trends are Uncommon and Stochastic.' Available at SSRN, <https://ssrn.com/abstract=3125890>.
- [7] **Cheibub, José, Jennifer Gandhi, and James Vreeland (2010)** 'Democracy and dictatorship revisited.' *Public Choice* 143(1-2): 67-101.
- [8] **Chudik, Alexander, Kamiar Mohaddes, Hashem Pesaran, and Mehdi Raissi (2016)** 'Long-run Effects in Large Heterogeneous Panel Data Models with Cross-Sectionally Correlated Errors.' In: Gloria Gonzalez-Rivera, Carter Hill, Tae-Hwy Lee (Eds.) *Advances in Econometrics: Essays in Honour of Aman Ullah*, Emerald Group Publishing, pp.85-135.
- [9] **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.
- [10] **Gerring, John, Philip Bond, William Barndt, and Carola Moreno (2005)** 'Democracy and economic growth: A historical perspective.' *World Politics* 57(3): 323-364.
- [11] **Gobillon, Laurent, and Thierry Magnac (2016)** 'Regional policy evaluation: Interactive fixed effects and synthetic controls.' *Review of Economics and Statistics* 98(3): 535-551.
- [12] **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.
- [13] **Hamilton, Lawrence (1992)** How robust is robust regression? *Stata Technical Bulletin* 1(2).

- [14] **Madsen, Jakob, Paul Raschky, and Ahmed Skali (2015)** 'Does democracy drive income in the World, 1500-2000?' *European Economic Review* 78: 175-195.
- [15] **Papaioannou, Elias, and Gregorios Siourounis (2008)** 'Democratisation and growth.' *Economic Journal* 118(532): 1520-1551.
- [16] **Pesaran, Hashem (2006)** 'Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure.' *Econometrica* 74(4): 967-1012.
- [17] **Pesaran, Hashem, and Ron Smith (1995)** 'Estimating long-run relationships from dynamic heterogeneous panels.' *Journal of Econometrics* 68(1): 79-113.
- [18] **Sarafidis, Vasilis, and Donald Robertson (2008)** 'On the impact of error cross-sectional dependence in short dynamic panel estimation.' *Econometrics Journal* 12(1): 62-81.
- [19] **Xu, Yiqing (2017)** 'Generalized synthetic control method: Causal inference with interactive fixed effects models.' *Political Analysis* 25(1): 57-76.

Appendix – Not intended for publication

A Data Sources, Descriptives and Sample Makeup

Income (ANRR) Data are taken from the World Bank World Development Indicators (WDI) database for real GDP per capita in year 2000 US\$. The GDPpc variable is transformed into logarithms and multiplied by 100, which eases the interpretation of the coefficients on the democracy dummy.

Democracy (ANRR) Data are combined from the Polity IV project, Freedom House and a number of alternative sources. The primary strategy for construction of the democracy dummy prescribes a positive Polity score in addition to a ‘free’ or ‘partially free’ label in Freedom House. Further strategies are described in detail in an appendix to ANRR.

Investment (ANRR) Data are taken from the World Bank World Development Indicators (WDI) database for the share of gross investment in GDP.

Trade Openness (ANRR) Data are taken from the World Bank World Development Indicators (WDI) database for the sum of imports and exports expressed as a share of GDP.

All of the above variables are compiled by ANRR and provided for download (along with the Stata do-files used in the analysis) from Daron Acemoglu’s personal website. Descriptive statistics for the ANRR variables are presented in Table A-1 below. In case of the income variable I provide statistics for the regression variable y (per capita GDP in logs multiplied by 100) and the raw per capita income data (in year 2000 US\$).

The sample makeup is illustrated in Table A-2, which arranges countries by sample observation count, and democratic status (transitioned, never/always a democracy).

Figure A-1 indicates the sample distribution across time (number of observations per annum) along with the distribution of democratisation events and reversals to autocracy. The different shadings of blue and pink reflect repeated democratisations/reversals, a characteristic which is largely ignored in the analysis of ANRR; in my own analysis I study the empirical results when the treatment sample is limited to countries which experience a single democratisation event.

Table A-1: Descriptive Statistics

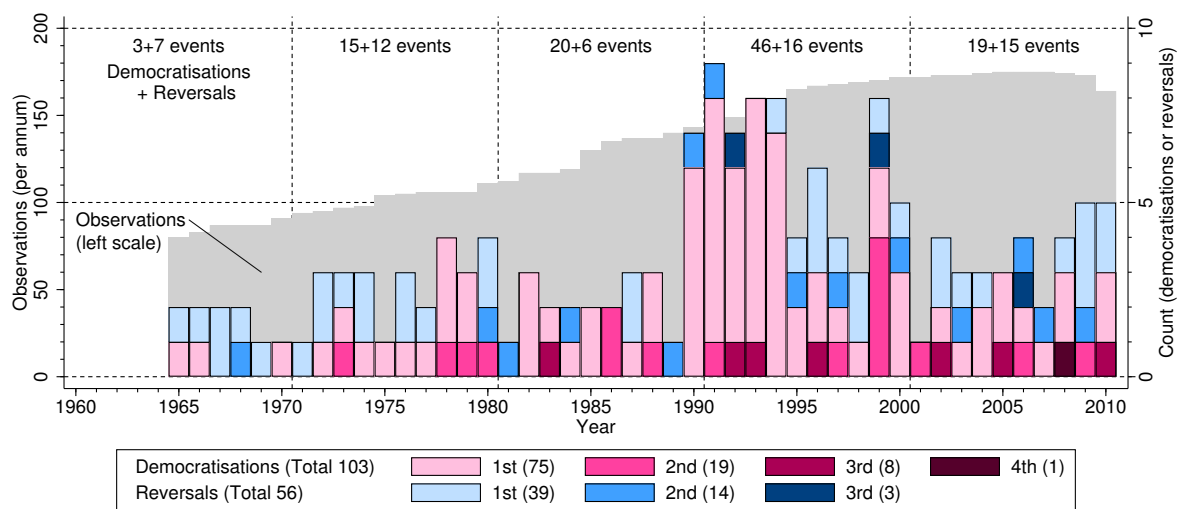
variable	Obs	mean	median	min	max	sd
dem	6,161	0.53	1.00	0.00	1.00	n/a
$\ln(\text{GDPpc}) \times 100$	6,161	753.1	746.8	405.7	1,094.0	155.4
Raw GDP pc	6,161	5,443	1,751	58	56,389	8,076

Table A-2: 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
								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								
27															
28															
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														
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		

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.

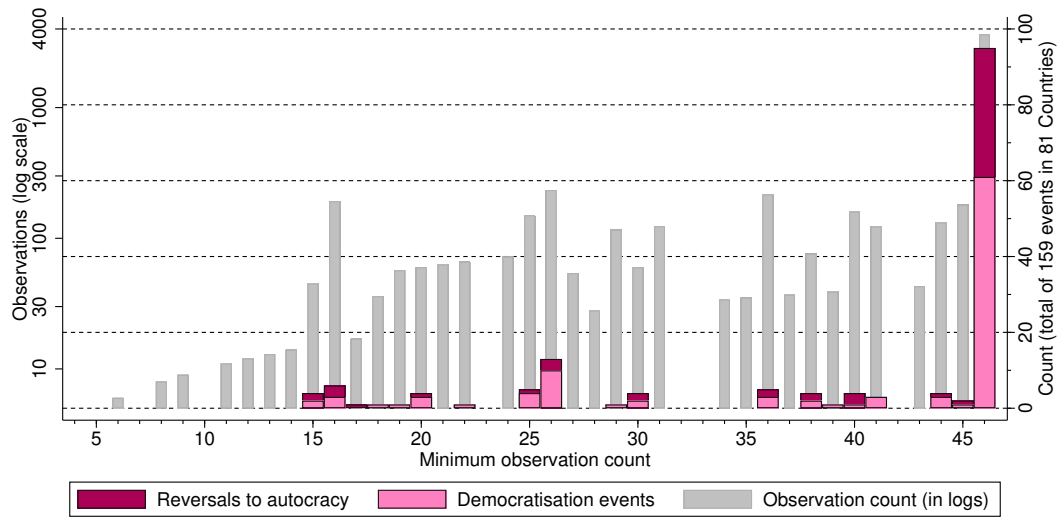
Figure A-1: Histogram – Full Sample – Democratisation and Reversal



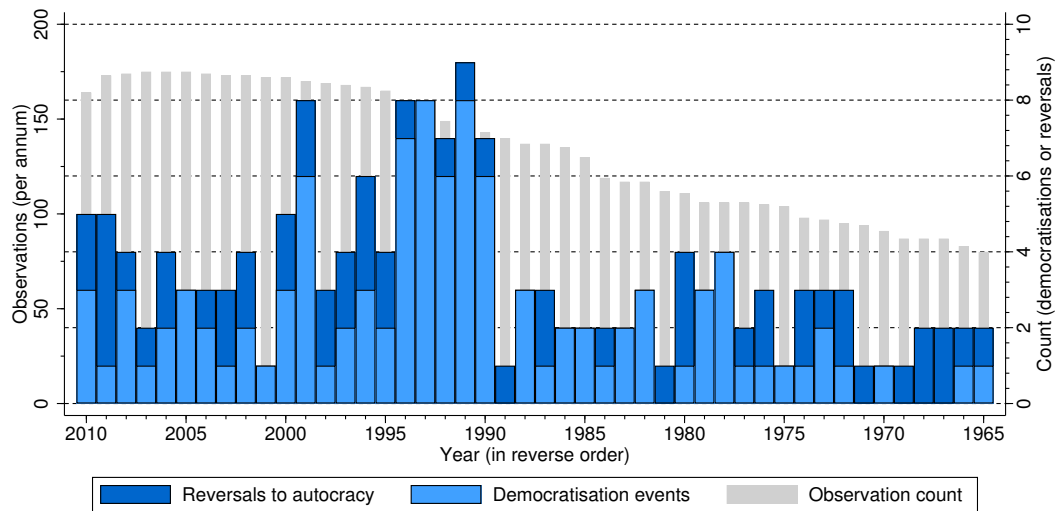
Notes: The figure presents the distribution of democratisation events (shades of pink) and reversals to autocracy (shades of blue) over the 1965 to 2010 period — the sample employed here is for the AB/HHK specification with four lags of GDP (hence the omission of 1960-64). The shades distinguish *first* 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).

B Sample Reduction Exercises – Event Distribution

Figure B-1: Sample and Event Distribution



(a) Sample Reduction by T_i count

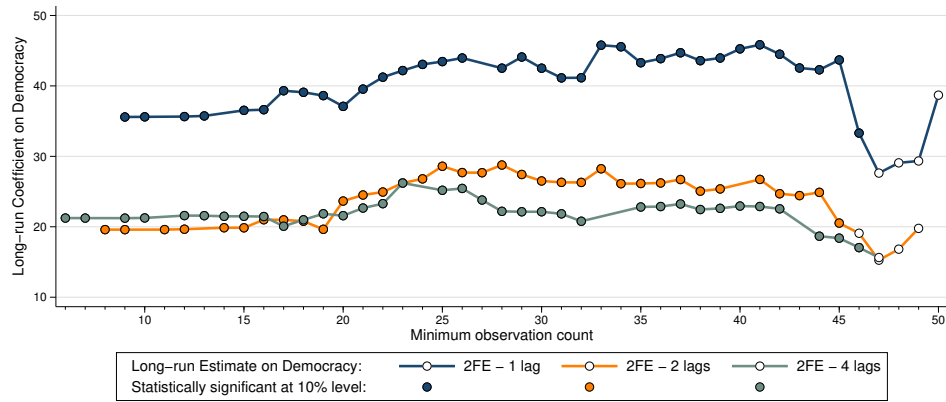


(b) Sample Reduction by end year

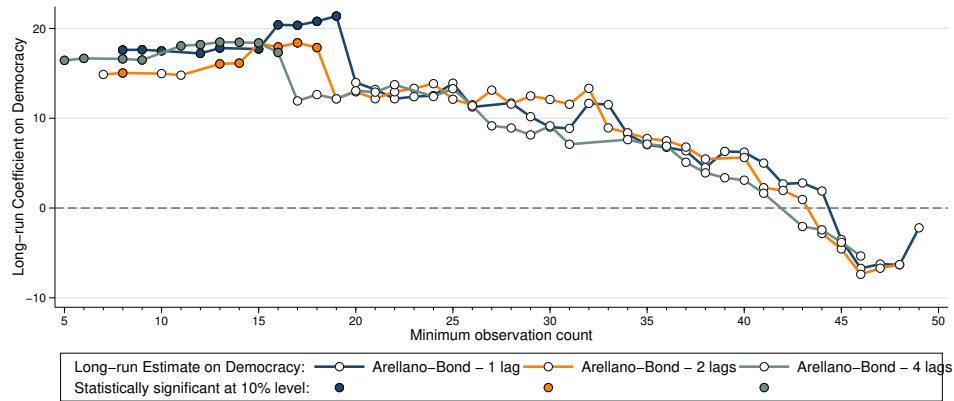
Notes: The figure presents the sample distribution for democratisation from varying empirical samples. The x -axis in panel (a) indicates the minimum number of observations required to be included in the sample, in panel (b) the sample end year (in reverse chronological order). The thin grey bars indicate the distribution of observations (log scale in panel (a), left axis) while the coloured bars indicate democratisation and reversal events (right scale). These distributions are for the AB/HHK samples.

C Sample Reduction Exercises – Full Parametric Results

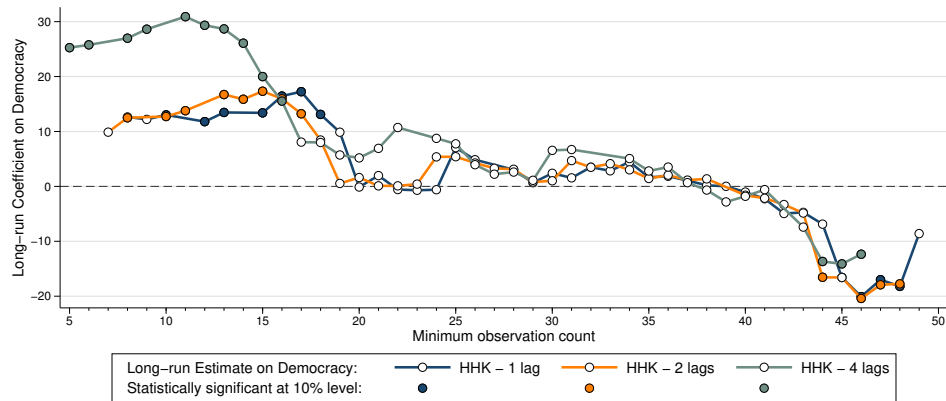
Figure C-1: Sample Reductions by T_i count (FE, AB, HHK)



(a) Within Estimates



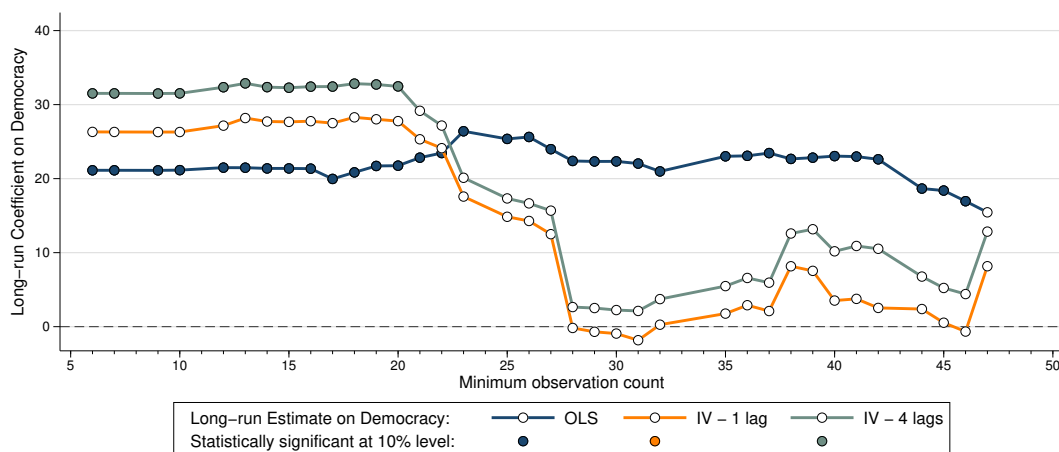
(b) Arellano-Bond Estimates



(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 B-1 in the maintext for all other details.

Figure C-2: Sample Reductions by T_i count (2SLS)



(a) OLS and 2SLS Estimates

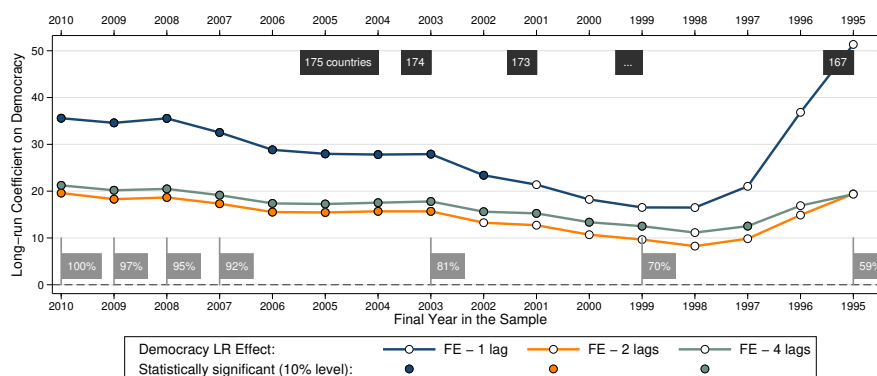
Notes: The figure presents long-run estimates for democracy from varying empirical samples for the 2SLS estimator. There are results for 2SLS using one and four lags of regional democratisation waves, respectively; in addition I report simple OLS ('within'/2FE) estimates for reference. The 'left-most' estimates for each specification replicate the results presented in ANRR's Table 6 (Panel A), columns (1) and (2) for single and four lags of the instrument. See Figure C-3 for all other details. Note that ANRR further adopt a HHK-2SLS implementation, which is omitted here since I was unable to replicate their full sample results.

Table C-1: Minimal Sample Reduction

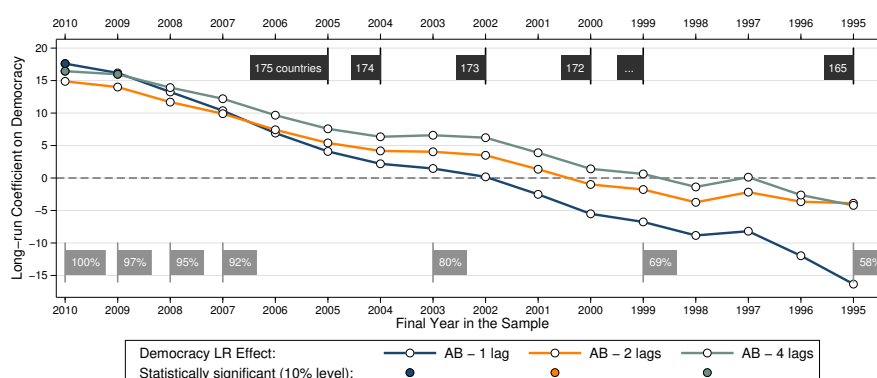
ANRR Reference	Sample reduction by T_i count							
	[1] 2FE		[2] AB		[3] HHK		[4] 2SLS	
	Table 2(3)		Table 2(7)		Table 2(11)		Table 6(2)A	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
Long-run effect of democracy	21.240 [7.215]***	15.637 [9.867]	16.448 [8.436]*	12.846 [8.023]	25.032 [10.581]***	9.221 [5.830]	31.521 [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.

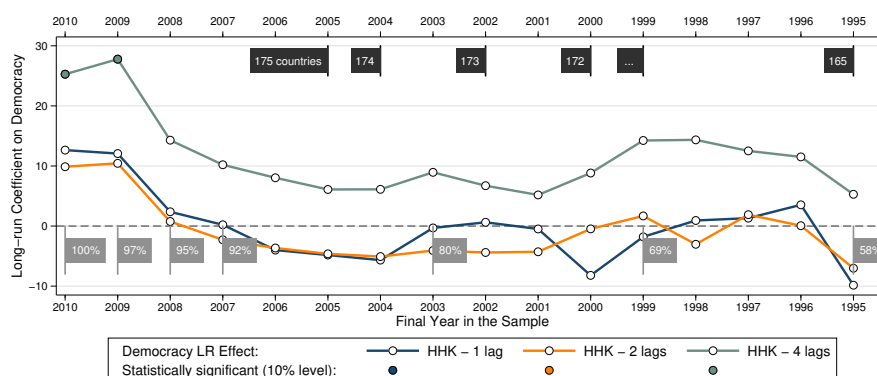
Figure C-3: Sample Reductions by end year (FE, AB, HHK)



(a) Within Estimates



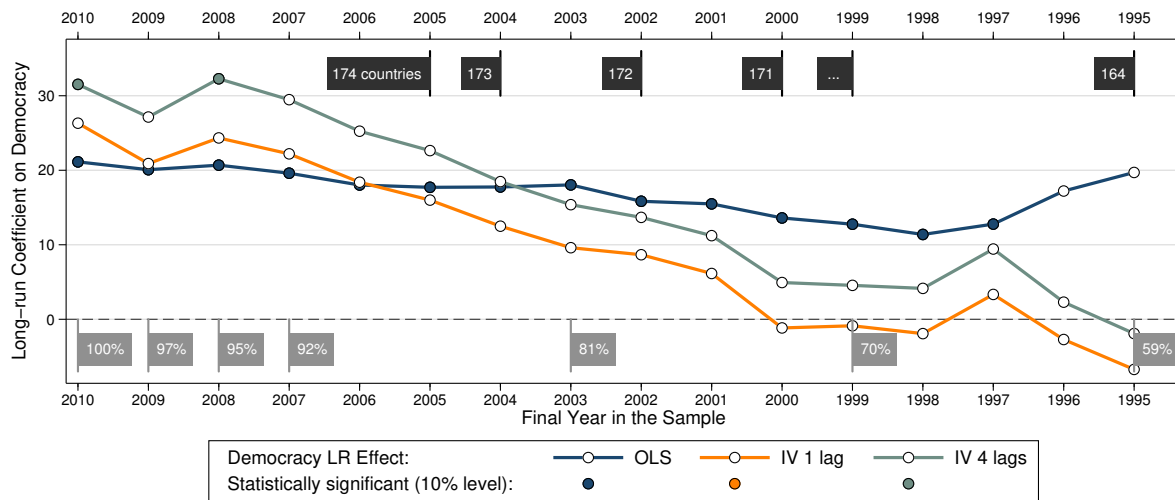
(b) Arellano-Bond (AB) Estimates



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

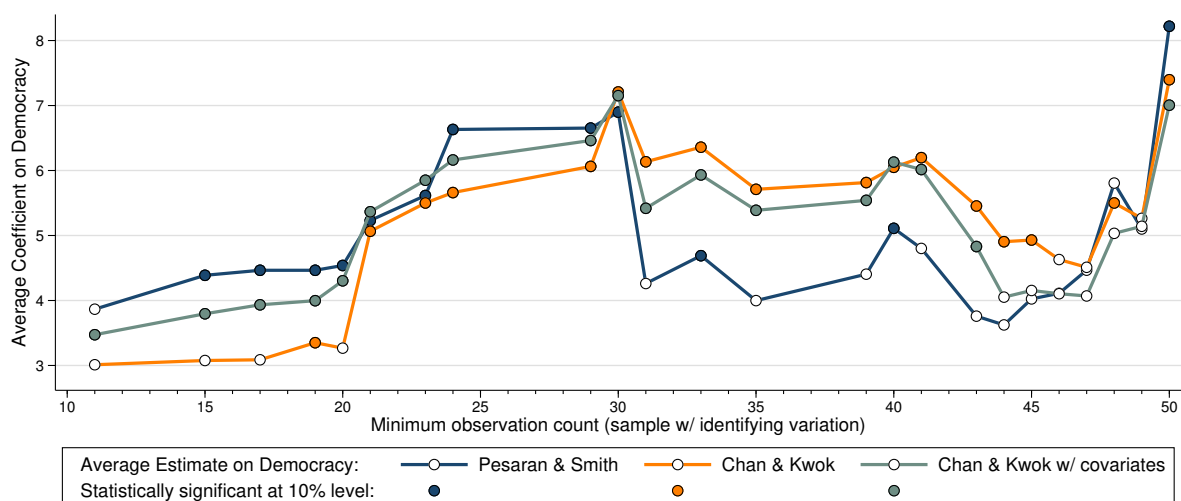
Notes: The figure presents long-run estimates for democracy from varying empirical samples for the (a) within, (b) AB, and (c) HHK estimators. The x -axis in each plot indicates the sample end year. A filled (white) circle marker indicates that the long-run coefficient is statistically (in)significant 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 C-4: Sample Reductions by end year (2SLS)

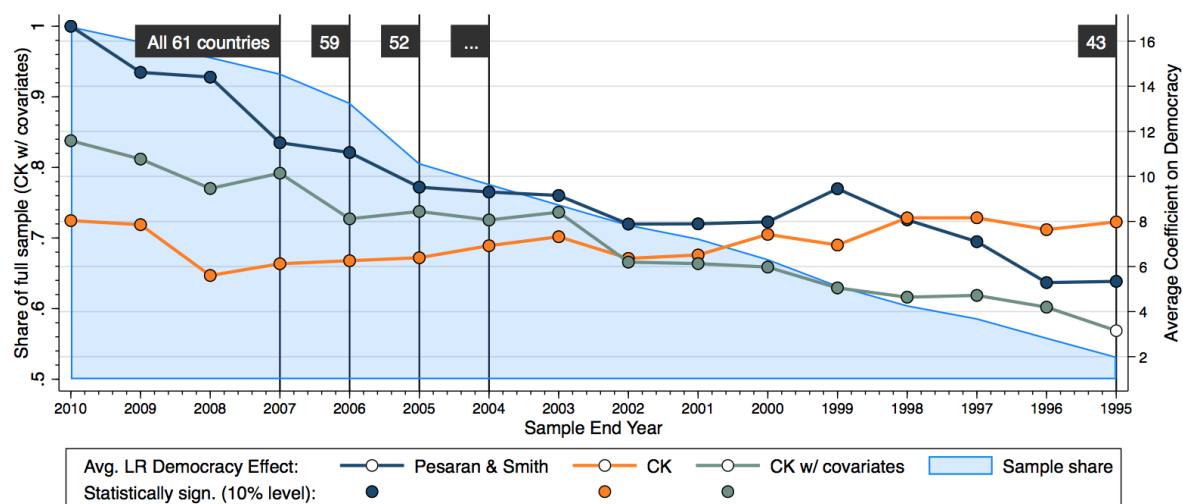


Notes: The figure presents long-run estimates for democracy from varying empirical samples for the 2SLS estimator. In addition to results for the the ANRR 2SLS specification with one and four lags of GDP I include the FE-4 lags specification for illustration (labelled 'OLS'). The former correspond to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

Figure C-5: Heterogeneous parameter models — static models



(a) Sample reductions by T_i count

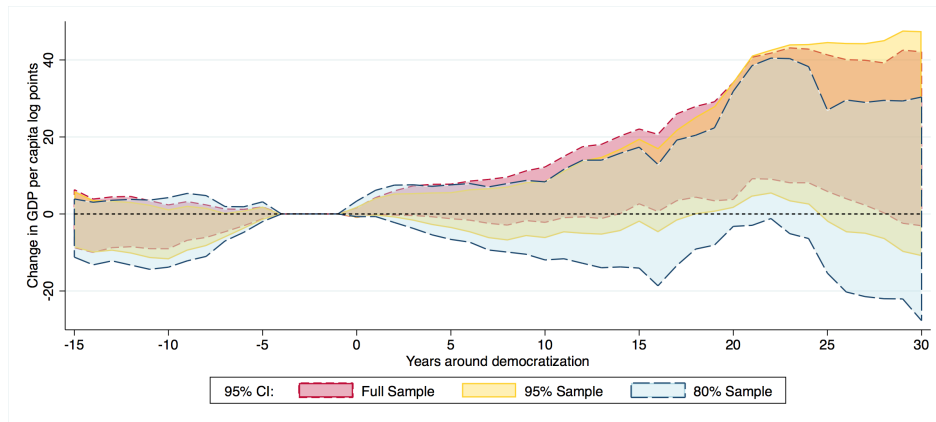


(b) Sample reductions by end year

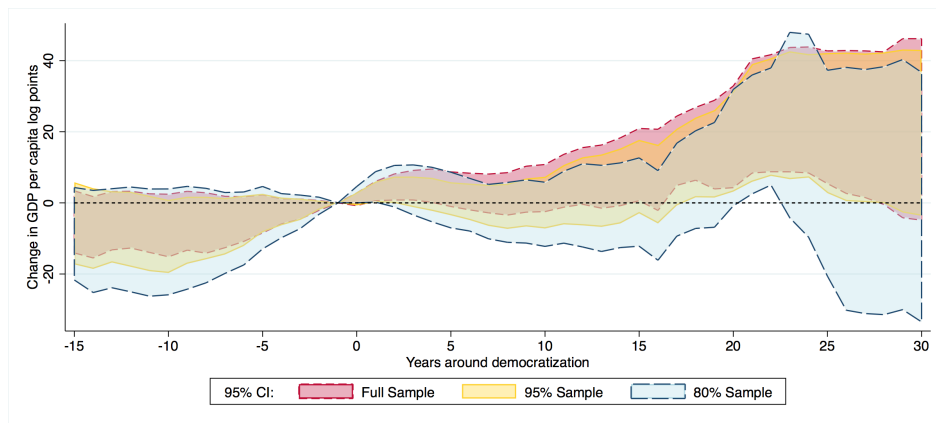
Notes: The figure presents robust mean estimates for democracy from varying empirical samples for the MG and Chan and Kwok type estimators in a static model. Panel (a) presents results the sample reduction by observation count, panel (b) from the sample reduction by end year.

D Sample Reduction Exercises – Semi-Parametric Results

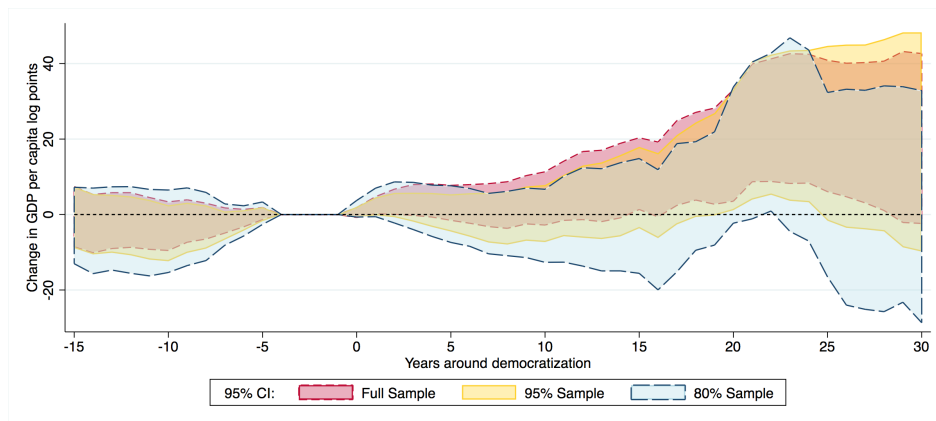
Figure D-1: Sample Reductions by T_i count — Semi-parametric models



(a) Adopting Linear Regression



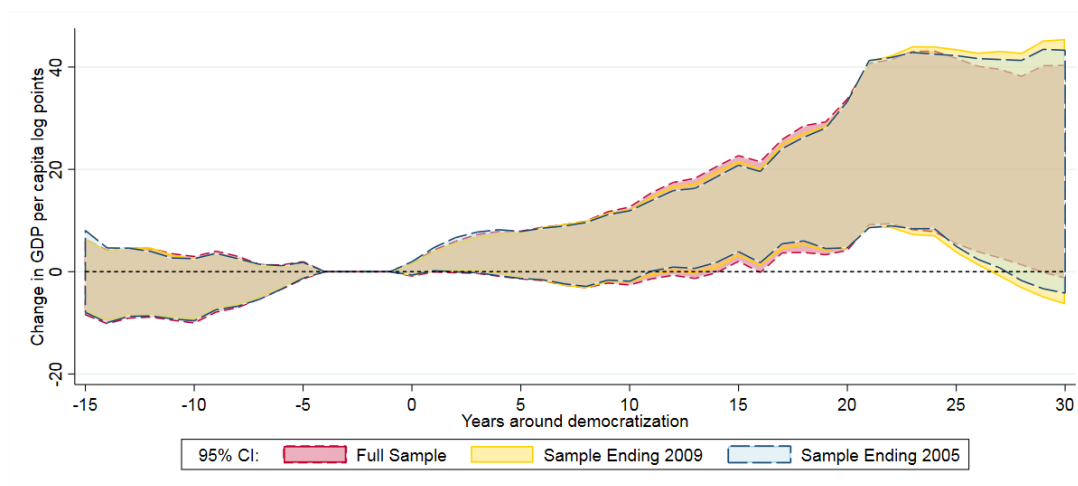
(b) Adopting Inverse Propensity Score Reweighting



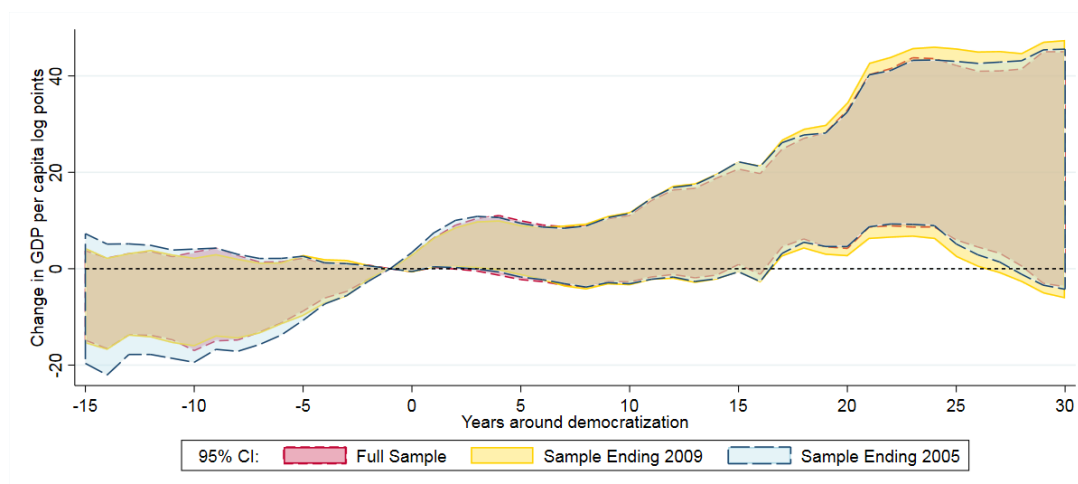
(c) Adopting Both Strategies

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).

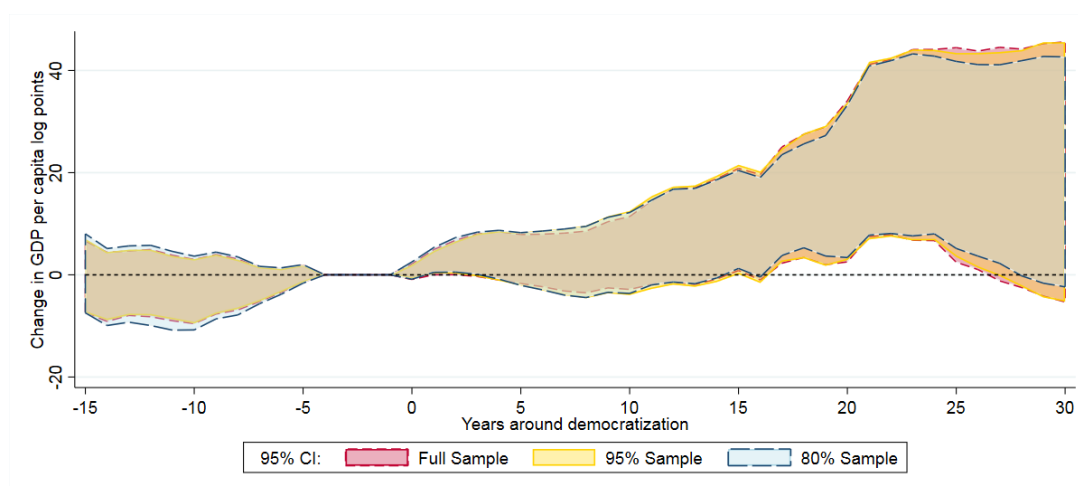
Figure D-2: Sample Reductions by end year — Semi-parametric models



(a) Adopting Linear Regression



(b) Adopting Inverse Propensity Score Reweighting

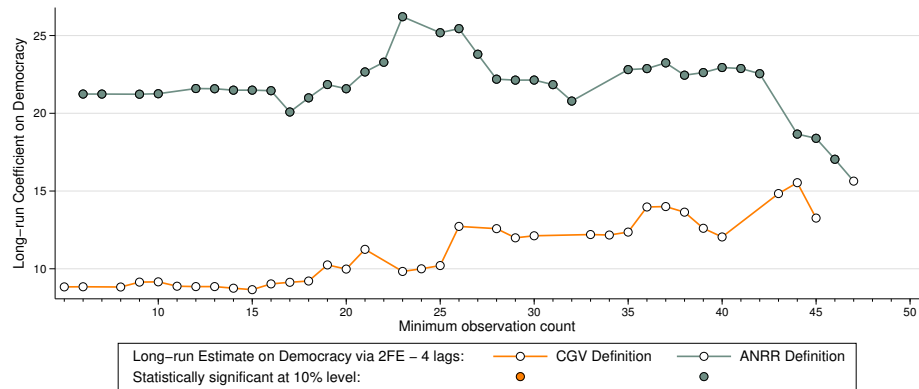


(c) Adopting Both Strategies

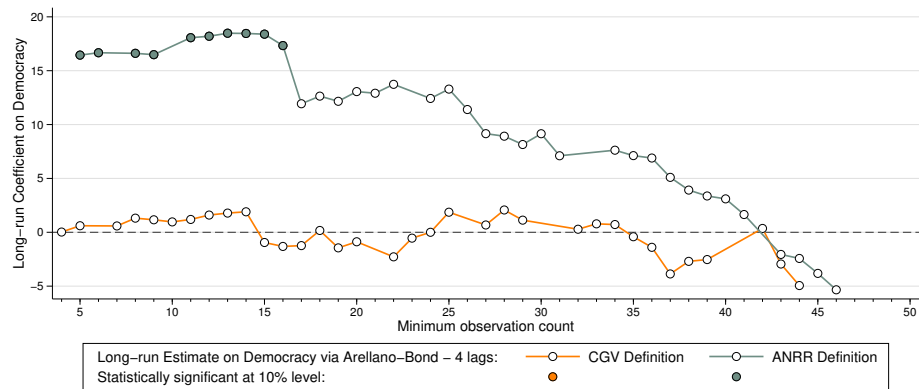
Notes: The figure presents the estimates for the ‘over time’ effect of democracy on per capita GDP using three different methods to estimate the relationship semi-parametrically. In each case the red shading indicates the full ($n = 3,029, N = 126$) sample 95% confidence interval (CI), and yellow and 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).

E Alternative Definitions for Democracy

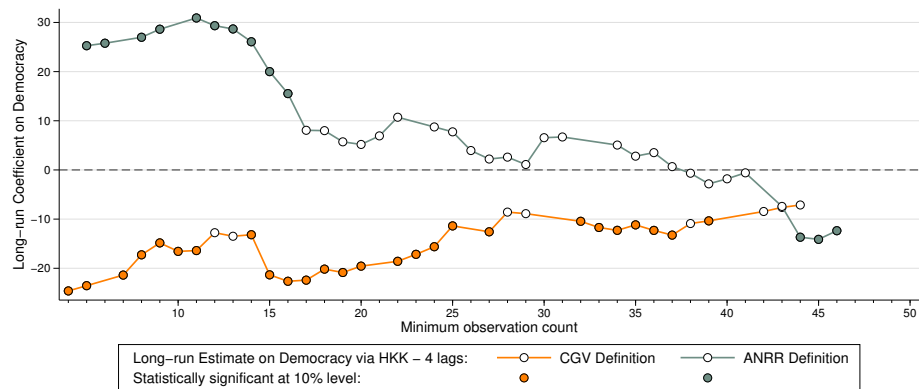
Figure E-1: Sample Reductions — Parametric Models (FE, AB, HHK)



(a) Within Estimates



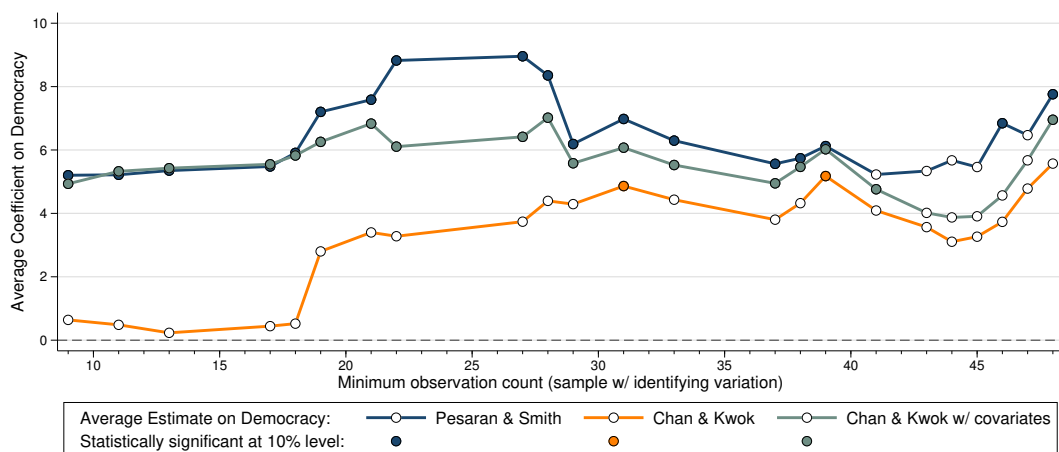
(b) Arellano-Bond Estimates



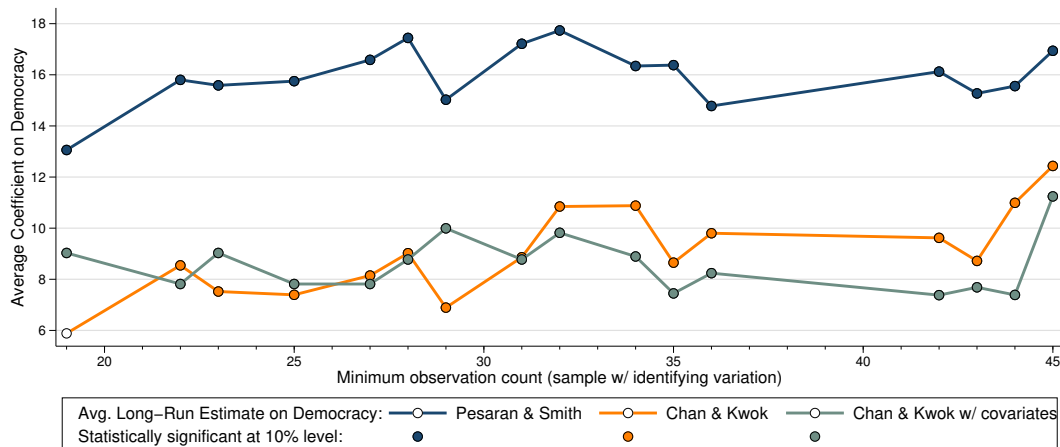
(c) Hahn, Hausman & Kuersteiner (HHK) Estimates

Notes: This figure provides sample reduction results for the 2FE, Arellano and Bond, and HHK estimators (in each case specification with 4 lags of the dependent variable), contrasting long-run estimates for the ANRR definition (results in teal) of democracy with that of CGV (results in yellow).

Figure E-2: Sample Reductions — heterogeneous parameter estimators (CGV definition)



(a) Static Models

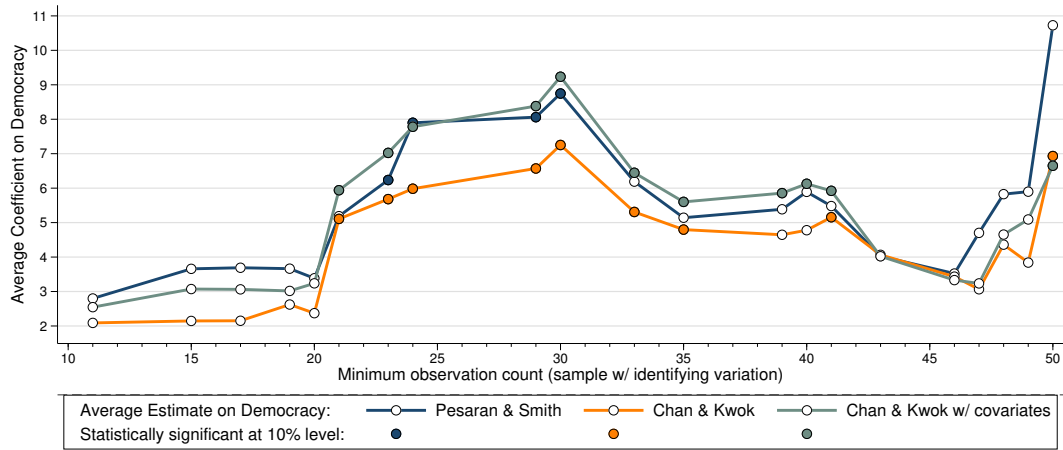


(b) Dynamic Models

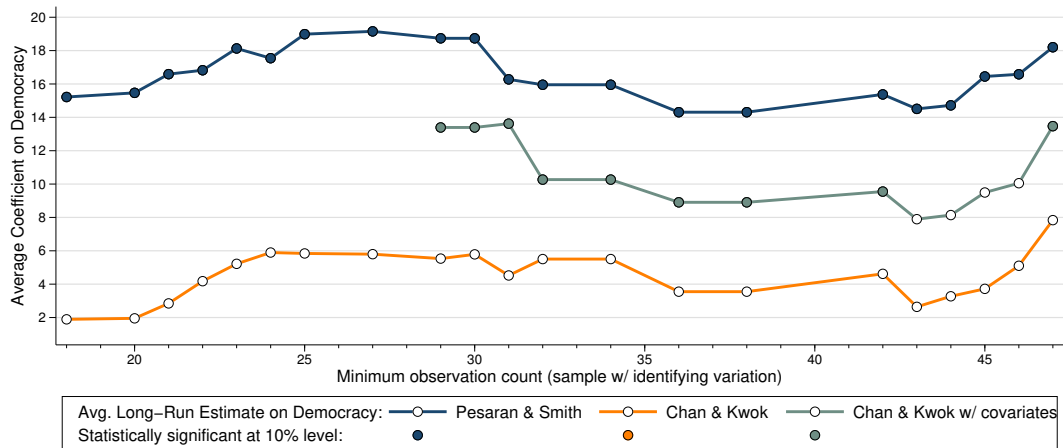
Notes: This figure provides sample reduction results for the static and dynamic Diff-in-Diff estimators using the alternative definition for democracy by Cheibub, et al (2010, CGV). This figure needs to be contrasted with Figure 2 in the maintext for a comparison with the results for the ANRR definition of democracy.

F Alternative Specifications

Figure F-1: Sample Reductions — heterogeneous parameter estimators;
excluding countries with more than one transition into democracy



(a) Static Models

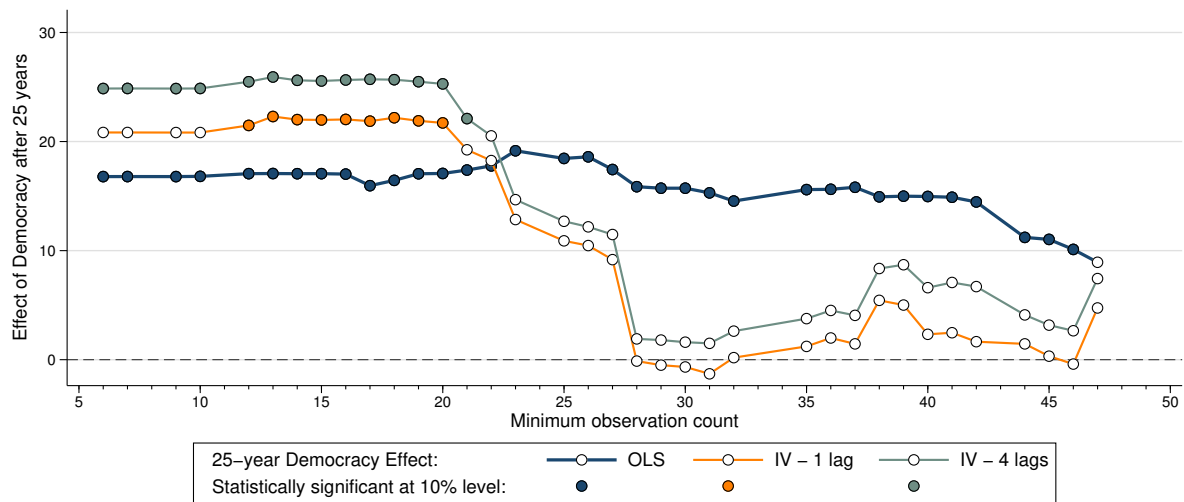


(b) Dynamic Models

Notes: This figure provides sample reduction results for the static and dynamic Diff-in-Diff estimators using a sample which excludes countries which transitioned into democracy more than once during the sample period. The resulting sample analysed in the above plots is thus made up of 63 countries.

G Alternative Definition of the ‘Long-run’

Figure G-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 (2FE) 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).

H Some forensic analysis of the sources of bias in ANRR

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. 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 H-1 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 1 plots (among others) the 2SLS $\hat{\beta}^{LR}$ for democracy, Figure H-1 plots the estimated GDP dynamics used in computing these long-run democracy estimates and standard errors. 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 Figure 1.

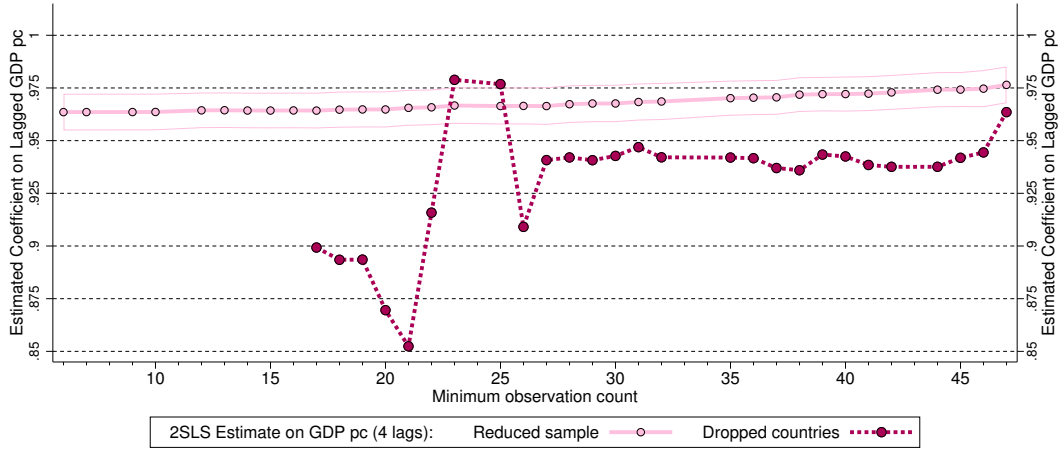
So what is the estimate on the GDP dynamics in the countries or years I omit? The dashed line in panel (a) of Figure H-1 represents the estimated GDP dynamics for all countries with a minimum observation count *lower* than that indicated on the *x*-axis:¹³ as we move to the right these countries are 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.¹⁴ 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 panel (b) of Figure 1.

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.

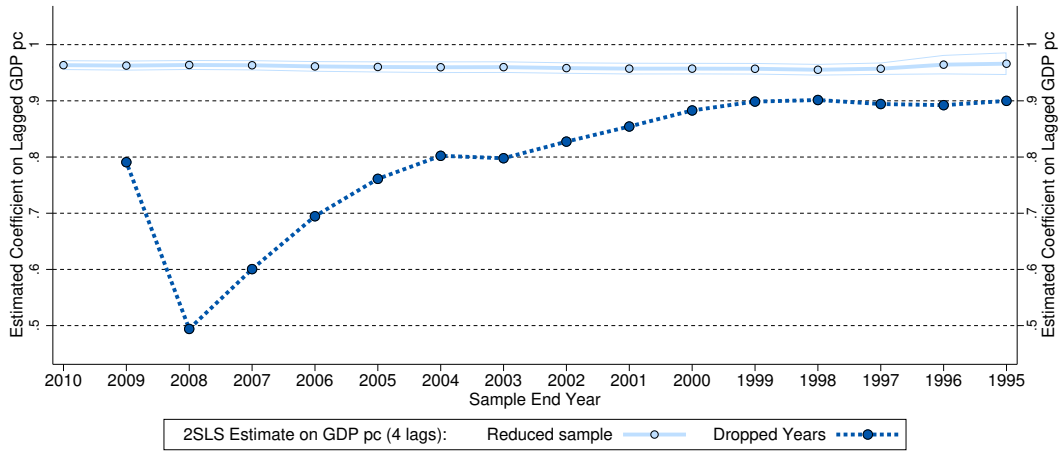
¹³For $T_{min} = 6$ this estimate would be constructed from 6 observations in one country. I therefore only begin charting this estimate for countries with 17 or fewer observations (338 observations in 23 countries).

¹⁴I do not show the confidence intervals for the ‘drop-out’ estimates for ease of illustration.

Figure H-1: 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.