DISCUSSION PAPER SERIES

DP13659

DEMOCRACY DOES CAUSE GROWTH: COMMENT

Markus Eberhardt

MACROECONOMICS AND GROWTH

ISSN 0265-8003

DEMOCRACY DOES CAUSE GROWTH: COMMENT

Markus Eberhardt

Discussion Paper DP13659
Published 09 April 2019
Submitted 08 April 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

DEMOCRACY DOES CAUSE GROWTH: COMMENT

Abstract

I revisit the causal relationship between democracy and growth as recently studied in Acemoglu, Naidu, Restrepo, and Robinson (2019, ANRR). I demonstrate the sensitivity of their results to sample selection by dropping a small number of observations in a non-random fashion and use these findings to motivate a generalisation of their empirical approach. My own analysis relaxes the assumption of (i) a common democracy-growth relationship, and of (ii) the absence of strong cross-section correlation. Adopting novel methods for policy evaluation I find a robust positive long-run effect of democracy albeit with only around half the magnitude of that found in ANRR.

JEL Classification: O10, P16

Keywords: democracy, growth, Political development, Spillovers, Difference-in-Difference Estimator, Interactive Fixed Effects

Markus Eberhardt - markus.eberhardt@nottingham.ac.uk
School of Economics, University of Nottingham and CEPR

Acknowledgements
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 the 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

I Introduction

In a recent article Acemoglu, Naidu, Restrepo, and Robinson (2019, henceforth ANRR) study the causal link between democracy and growth[1] in a large sample of up to 175 countries. The authors adopt a variety of empirical strategies which account for country-specific fixed effects and the dynamics of per capita GDP. 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, the Arellano and Bond (1991, AB) and the Hahn, Hausman, and Kuersteiner (2001, HHK) estimators.[2] 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) the empirical 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).[3]

ANRR point to three modelling choices/assumptions which are crucial (my words) to successfully capture the causal effect of democracy on growth. First, their newly-developed consolidated dichotomous measure for democracy, which is suggested to address concerns over measurement error in existing democracy indices;[4] second, the importance of modelling the dynamics of GDP to account for its movements in the lead-up to, during, and after democratisation events which is argued to avoid the downward bias on the estimates of democracy in the previous empirical work (e.g. Barro, 1996); and third, the identification strategy of the 2SLS regressions exploiting regional waves of democratisation and reversal to address concerns of reverse causality. One can add to these *implicit* fourth and fifth assumptions, namely that the democracy-growth relationship and its dynamics are the same for all countries (parameter homogeneity), and that there is no bias due to strong cross-section dependence.

[1] I 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.

[2] The analysis of ANRR goes beyond what I describe in this paragraph: a number of semi-parametric estimators are employed to relax the assumption of linear GDP dynamics, yielding very similar results to those from the parametric models. ANRR further investigates the 'transmission channels' through which democracy affects growth.

[3] Chen, Chernozhukov and Fernandez-Val (2019, CCF-V) employ AB and FE estimators but with bias-correction for the many instruments and incidental parameter problems, respectively, confirming the AB/FE results of ANRR. CCF-V's sample (balanced panel, 1987-2009) yields an ANRR 2SLS long-run democracy estimate of 179% (t=1.57).

[4] Results presented in an Appendix suggest that using ANRR's preferred estimators typically yields statistically insignificant long-run estimates for democracy when using the democracy indicator devised by Cheibub, Gandhi, and Vreeland (2010; CGV). Results for the Chan and Kwok (2018) heterogeneous parameter models, however, yield long-run estimates between 8 and 12%, thus very close to those presented below using the ANRR definition.

This paper investigates the empirical consequences of relaxing these two assumptions. One can draw on existing arguments for a democratic legacy (Gerring, et al, 2005) or threshold levels in economic or human development (Aghion, et al, 2008; Madsen, et al, 2015; ANRR's own findings on the role of education) as necessary conditions for a positive democracy-growth nexus to motivate *parameter heterogeneity*. Alternatively, work by Cervellati and Sunde (2014) has indicated differential growth trajectories by 'democratization scenario,' namely peaceful versus violent transitions. Since *a priori* the magnitude and even the sign of the democracy effect is unknown, one can instead point to differential GDP dynamics across countries (speed of convergence to the long-run equilibrium) as an additional motivation for a heterogeneous parameter setup. The presence of *strong cross-section dependence* can arise from global spillovers or shocks with heterogeneous impact across countries. This would invalidate the identifying assumption inherent in ANRR's 2SLS estimation (Pesaran and Smith, 1995) and also in the Arellano and Bond (1991, AB) and Hahn, Hausmann, and Kuersteiner (2001, HHK) estimators the authors draw on to address concerns over 'Nickell bias' (Nickell, 1981) in the baseline fixed effects models. Taken together, my analysis presents a generalisation of ANRR's empirical approach: I relax the above two assumptions while adopting an identification strategy which (like theirs) allows me to interpret the findings as causal.

Sensitivity of empirical results to non-random changes in the sample can point to the presence of observed and unobserved heterogeneities and I devise two rule-based sample reduction strategies to demonstrate this in the ANRR sample: (i) I drop countries on the basis of their observation count, first omitting those with merely five observations, then those with six, etc.; 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 Global Financial Crisis and its aftermath — the most significant global macroeconomic shock since the 1930s — at the end of the ANRR sample period.[5]

My results for these exercises indicate that the magnitude and statistical significance of the effect of democracy on per capita GDP is sensitive to sample selection, in line with the arguments for heterogeneity laid out above: depending on the sample reduction strategy and the estimator used the 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; full sample estimates are between 16% and 32%) 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

[5]My presentation is limited to the parametric results. The semi-parametric results for 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.

exercises. Additional exercises provide some evidence that dynamic misspecification (heterogeneous GDP dynamics) may account for the sensitivity of results observed.

I then turn to an alternative empirical approach from the recent panel time series literature. Adopting heterogeneous panel estimators for policy evaluation with interactive fixed effects (Chan and Kwok, 2018) I demonstrate that the average long-run estimate for democracy is statistically significant positive and robust to substantial sample reductions. My results indicate that the economic magnitude of this effect is around 10%, thus more modest than in the results of ANRR, although differences in my sample make-up make such comparisons difficult.

The remainder of this note is organised as follows: Section II summarises results from sample reduction exercises for the main parametric estimators adopted in ANRR. Section III studies the democracy-growth nexus with heterogeneous treatment effects estimators.

II 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 (less than one quarter of the full sample result), 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 souce 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 disporportionate 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 FE, AB, HHK and 2SLS estimators — all results are for the specification with four lags of GDP, which is preferred by ANRR.[6] 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 statistcally insignificant, though at 15.6% still reasonably close to the full sample estimate of 21.2%.[7]

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.

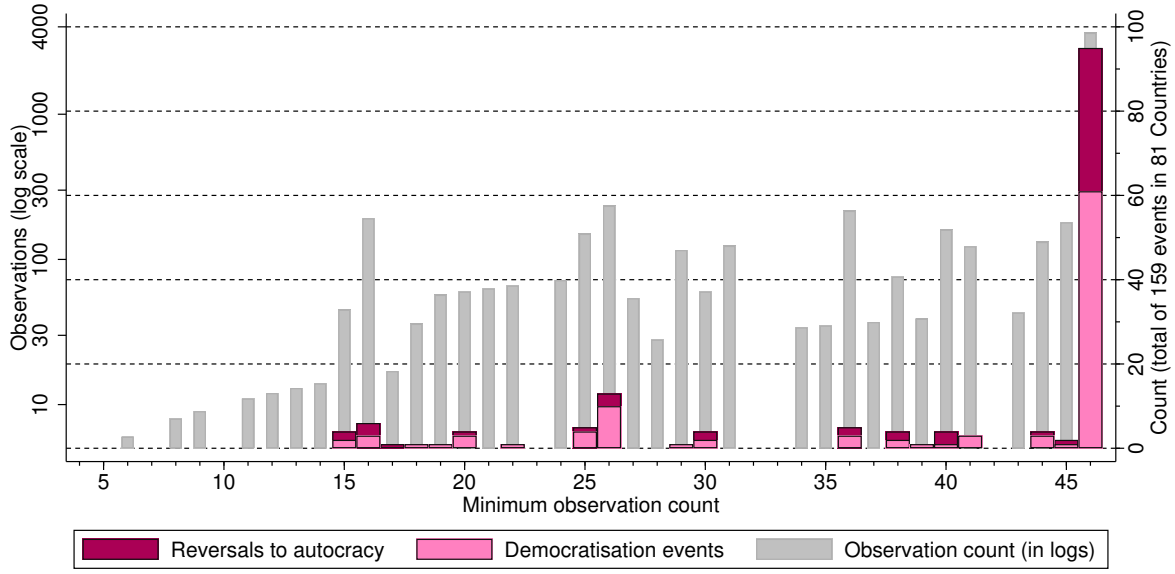
Democracy estimates based on the 2SLS estimator initially maintain a high and stable level in excess of 30%, but turn insignificant once countries with fewer than 21 observations are omitted (7% of the full sample of 6,300 observations).[8] The magnitude of 2SLS estimates drops quite rapidly, such that 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.

[6]Results 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.

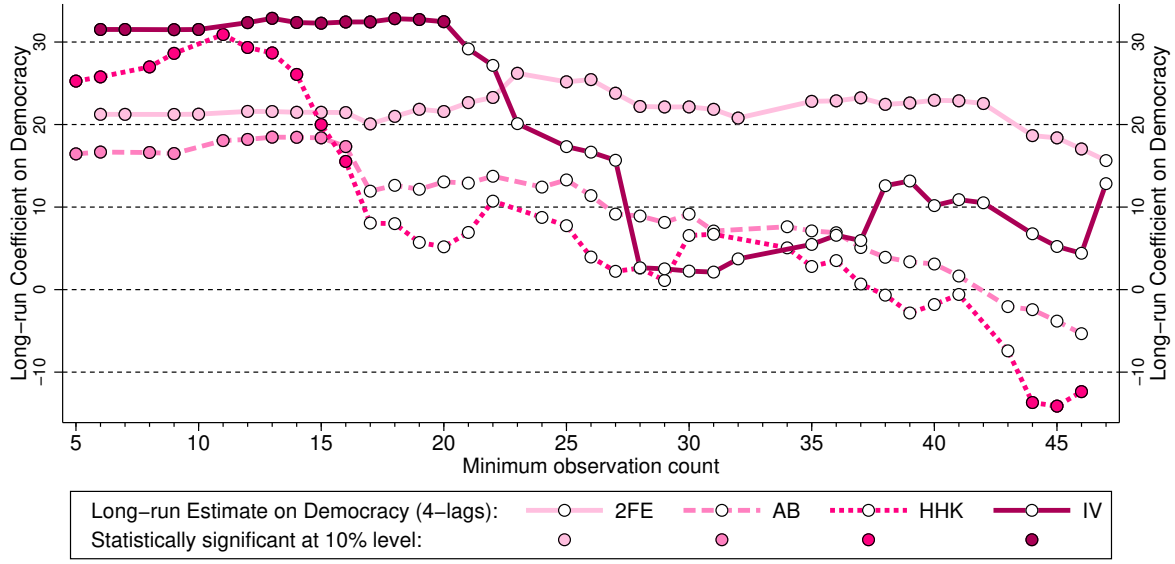
[7]Note that many researchers have serious reservations about the fixed effects estimator for causal inference in panel data (e.g. Gibbons, Suarez-Serratoz and Urbancic, 2018; Imai and Kim, 2018).

[8]The 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 signficance of the estimates.

Figure 1: Sample Reduction by T_i count



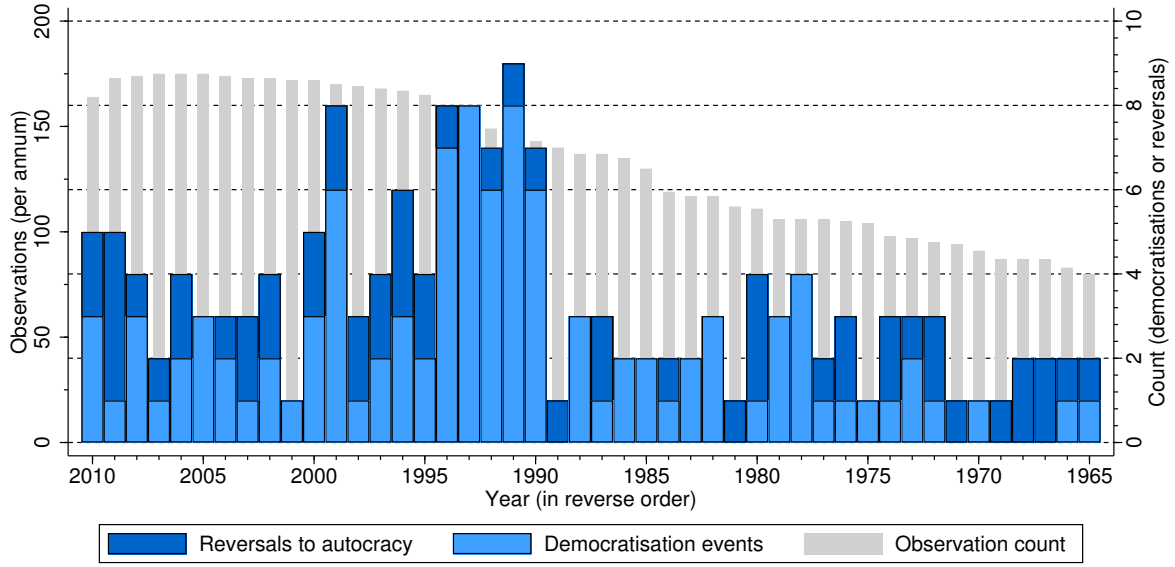
(a) Sample and Event Distribution



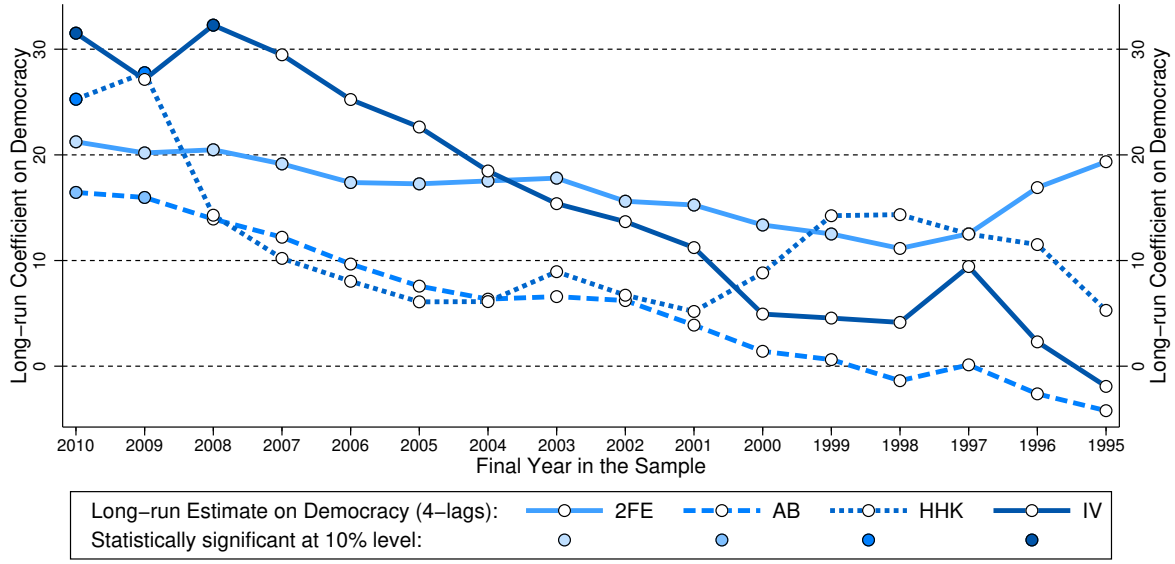
(b) Long-Run Estimates

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.

Table 1: Sample Reduction Estimates

	Sample reduction by T_i count				Sample reduction by end year			
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Estimator	2FE	AB	HHK	2SLS	2FE	AB	HHK	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], respectivly (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 prefered 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.

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.[9]

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.

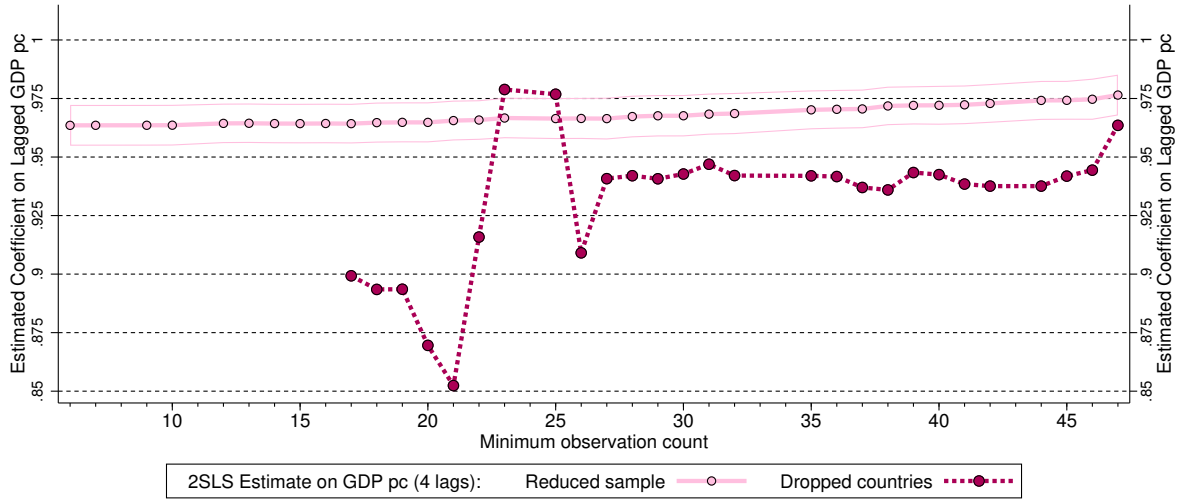
As before the 2FE estimates are found to be fairly robust to sample reduction, only turning insignificant when 30% of observations are dropped. The AB/HHK estimates, in contrast, turn insignificant if the post-GFC years 2009 and 2010 are omitted, thereafter declining and eventually diverging, with HHK 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. I

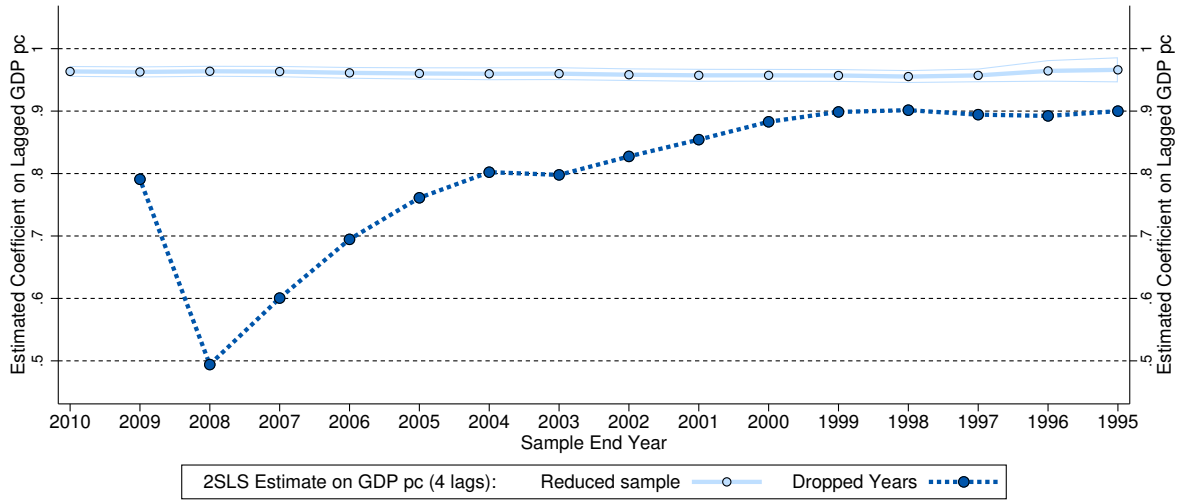
[9] This is once again the AB/HHK sample for the four-lag specification, hence the 1965 start year.

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.

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.

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: $\hat{\beta}^{LR} = \hat{\beta}/(1 - \sum_{\ell=1}^{4} \hat{\rho}_{i,t-\ell})$. Figures 1 and 2 plot (among others) the 2SLS $\hat{\beta}^{LR}$ for democracy, Figure 3 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 Figures 1 and 2.

So 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:[10] 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.[11] 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.

III Heterogeneity

This section studies the impact of observable and unobservable heterogeneity on empirical estimates of the democracy-growth nexus. My empirical model builds on the panel time series econometric literature which has emphasised heterogeneous parameters across panel members (Pesaran and Smith, 1995)[12] and, more recently, the presence of strong cross-section dependence (e.g. Bai, 2009; Pesaran, 2006) — a form of unobserved, time-*varying* heterogeneity.[13] Strong correlation across panel members 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 (Phillips and Sul, 2003; Andrews, 2005). This literature has taken to specifying a multi-factor error structure, $\boldsymbol{\lambda}_i' \boldsymbol{f}_t$, where $\boldsymbol{f}$ is a set of common factors with associated

[10] For $T_{min} = 6$ 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).

[11] I do not show the confidence intervals for the 'drop-out' estimates for ease of illustration.

[12] The pitfalls of imposing common slope coefficients on heterogeneous equilibrium relationships have been highlighted for dynamic (Pesaran and Smith, 1995) and static specifications (Sul, 2016). It is also worth emphasising that any instrumentation strategy applied in a pooled panel (such as the IV strategy in ANRR) will be invalid *by construction* if the true underlying equilibrium relationship differs across countries. If the coefficient imposed on x is β yet the true relationship is $\beta_i x$ then $(\beta_i - \beta)x$ will be contained in the error term, thus violating the exclusion restriction that instrument z be uncorrelated with the error if $E[xz] \neq 0$.

[13] Eberhardt and Teal (2011) provide a detailed introduction to these models with discussion of empirical applications from the cross-country growth literature.

heterogeneous factor loadings $\boldsymbol{\lambda}$, to capture this strong dependence.

The most recent contributions to this literature have been able to build bridges to the empirical literature on policy evaluation using difference-in-difference specifications (Gobillon and Magnac, 2016; Chan and Kwok, 2018) and the synthetic control methodology (Xu, 2017). What distinguishes these latest approaches from their canonical predecessors is the adoption of a multi-factor error structure in order to address three challenges to identification in these popular methods: (i) the presence of uncommon trends prior to the policy change evaluated, (ii) endogeneous selection into 'treatment', and (iii) the possibility that, following the policy change, treated and control samples are affected by common shocks, albeit with heterogeneous impact (e.g. the differential effect of the Global Financial Crisis).

Previous work analysing the democracy-growth nexus using difference-in-difference specifications includes Giavazzi and Tabellini (2005), Papaioannou and Siourounis (2008) and Cervellati and Sunde (2014). My implementation follows the spirit of Chan and Kwok's (2018) estimator but adopts cross-section averages à la Pesaran (2006) instead of estimated factors à la Bai (2009) due to the strongly unbalanced nature of the panel data at hand — more details below. Crucially, this setup to investigate the long-run effect of democracy on growth allows for correlation between the unobserved determinants of growth (institutions, absorptive capacity, etc.) and selection into democratic transition or reversal. Since it may be suggested that the *static* Chan and Kwok (2018) estimator cannot be directly compared with the *dynamic* specifications investigated above, I present results for a dynamic 'CS-DL' version (cross-section-augmented distributed lag; Chudik, et al., 2016) of the Chan and Kwok estimator.

The implementation is straightforward: for the sample of countries which experienced variation in the ANRR democracy dummy I specify the following static regression model

$$y_{it} = \alpha_i + \beta_i \, \text{Dem}_{it} + \boldsymbol{\gamma}_i' \boldsymbol{X}_{it} + \delta_i^y \overline{y}_t + \boldsymbol{\delta}_i^{X'} \overline{\boldsymbol{X}}_t + \varepsilon_{it}, \tag{1}$$

where y is per capita GDP (in logs/100), Dem is the democracy dummy, and $\boldsymbol{X}$ is a set of additional controls (I adopt gross investment share of GDP and trade openness) — all of these are taken from the ANRR dataset.[14] $\overline{y}$ and $\overline{\boldsymbol{X}}$ are the cross-section averages of the observed variables *but for those countries which never experienced democracy during the sample period* (the control group). Following the insights from Pesaran (2006) and Chan and Kwok (2018) these cross-section averages capture the presence of uncommon and/or stochastic trends. Note that by construction there is no cross-section average for the democracy variable, since this is always zero in the control group from which these are computed. Observed covariates $\boldsymbol{X}$ are not included in what I refer to as the 'plain vanilla' Chan and Kwok (2018) implementation — the

[14] The focus of this analysis is on β_i and its cross-country average, not on $\boldsymbol{\gamma}_i$.

covariate cross-section averages from the control sample, $\overline{\boldsymbol{X}}$, are however always included.[15]

For comparison with ANRR's results I employ a dynamic variant of equation (1):

$$\begin{aligned} y_{it} &= \alpha_i + \theta_i \text{Dem}_{it} + \boldsymbol{\Gamma}_i' \boldsymbol{X}_{it} + \sum_{\ell=0}^{p-1} \omega_{i\ell}^{D} \Delta \text{Dem}_{i,t-\ell} + \sum_{\ell=0}^{p-1} \boldsymbol{\omega}_{i\ell}^{X'} \boldsymbol{\Delta X}_{i,t-\ell} \\ &\quad + \sum_{\ell=0}^{p_{\overline{y}}} \delta_{i\ell}^{y} \overline{y}_{t-\ell} + \sum_{\ell=0}^{p_{\overline{X}}} \boldsymbol{\delta}_{i\ell}^{X'} \overline{\boldsymbol{X}}_{t-\ell} + \varepsilon_{it}, \end{aligned} \tag{2}$$

where the two terms involving sums in the first line capture the short-run effects, while θ_i and $\boldsymbol{\Gamma}_i$ represent the long-run coefficients for the respective effects of democracy and additional controls on income per capita. The sums in the second line capture the multifactor error structure using cross-section averages, which like in the static model are constructed from those countries which never experienced democracy during the sample period. The use of this 'CS-DL' version of the Chan and Kwok (2018) approach is convenient since the long-run coefficient can be estimated in a single step rather than two like in an error-correction specification or the ANRR implementations.[16] Following suggestions in Chudik, et al. (2016) I adopt $p_{\overline{y}} = 0$ and $p = p_{\overline{X}} = int(T^{1/3}) = 3$, where T is the time dimension of the panel. My presentation below will focus on average estimates of $\hat{\theta}$ in the dynamic case (which can be interpreted as ATET estimates); in line with the literature I adopt robust regression (Hamilton, 1992) to compute outlier-robust means. Inference for this robust 'Mean Group' estimate is based on standard errors computed non-parametrically, following Pesaran and Smith (1995).

Panel (a) of Figure 4 presents the results from *dynamic* specifications of three heterogeneous parameter models for the first sample reduction exercise by country observation count.[17] The estimates from the empirical model ignoring any potential factor structure and thus selection, uncommon trends and/or common shocks with heterogenous impact (labelled 'Pesaran & Smith') are substantially higher than 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 very similar long-run effects of democracy on GDP to the static version (see Appendix), between 6 and 8%. The model including additional covariates (gross investment ratio and trade openness) lifts the estimate somewhat to around 10%.[18]

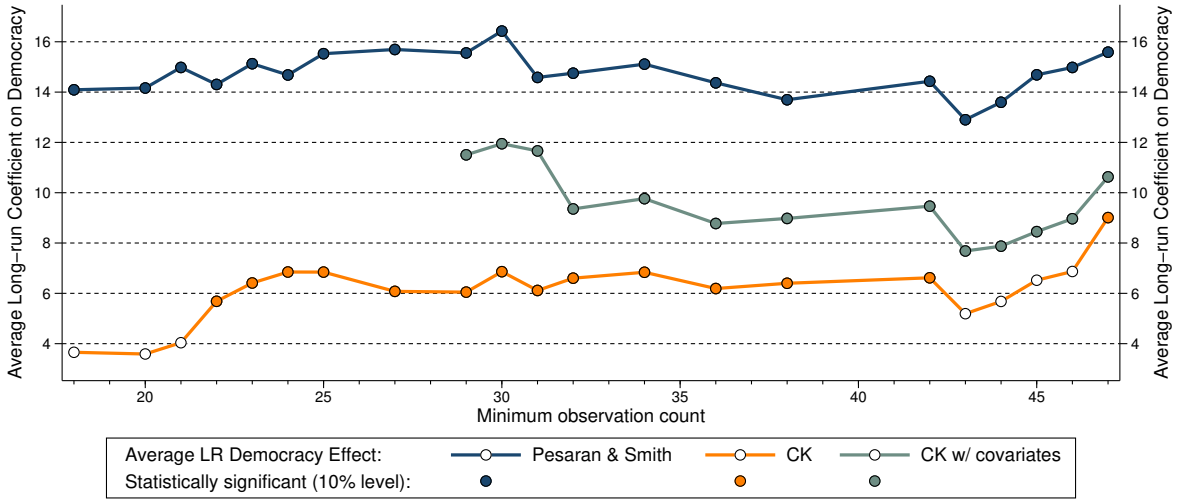
[15] Merely adding $\overline{y}$ allows for a single unobserved common factor f, whereas inclusion of $\overline{\boldsymbol{X}}$ allows for multiple common factors.

[16] In the ECM specification we obtain an estimate $\hat{\beta}_i$ for democracy and $\hat{\rho}_i$ for the lagged dependent variable (or $\sum_{\ell=1}^{p} \hat{\rho}_{i\ell}$ for p lags), from which the long-run coefficient $\hat{\theta}_i = \hat{\beta}_i / - \hat{\rho}_i$ has to be computed. It is apparent from this that any finite sample bias in $\hat{\rho}_i$ will carry over to $\hat{\theta}_i$ (Chudik and Pesaran, 2015). The CS-DL obtains these estimates in a single step by adopting an alternative specification and avoids potential bias from dynamic misspecification.

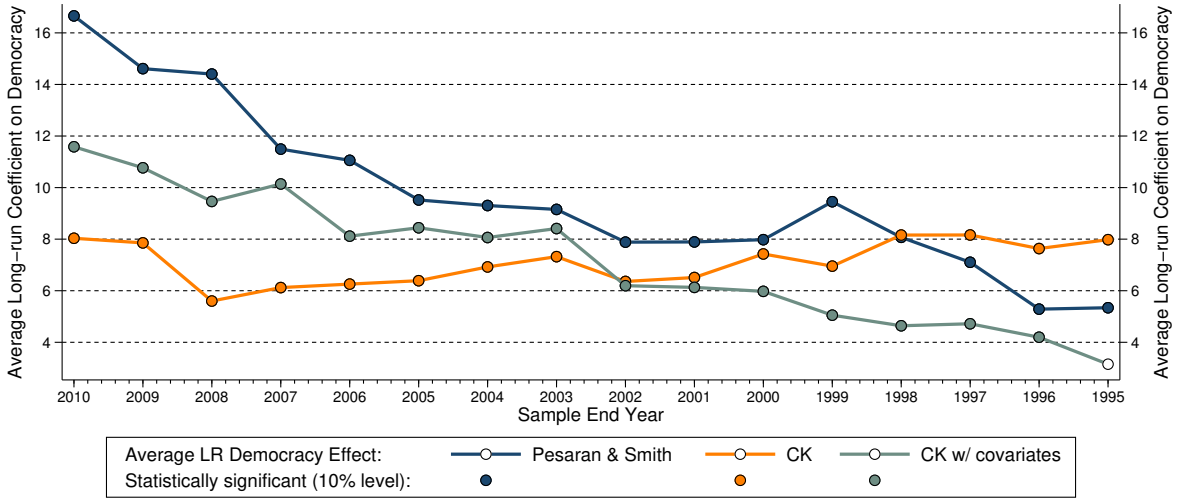
[17] As one would expect the static results provided in an Appendix are somewhat less stable than these dynamic estimates and yield a positive significant average effect of democracy in the range from 3% to 8%.

[18] Since this model requires estimation of an additional $2p + 2$ parameters the minimum feasible sample is $T = 29$ as opposed to $T = 18$ (for the plain vanilla version.)

Figure 4: Sample Reductions — heterogeneous parameter estimators



(a) Sample reduction by T_i count



(b) Sample reduction by end year

Notes: The figure presents the robust mean estimates for a variety of Difference-in-Difference estimators as the regression sample is constrained, using the minimum count of country observations as the selection mechanism in panel (a) and the sample end year in panel (b). The unconstrained sample is made up of a maximum of 86 economies which transitioned into or out of democracy at least once during the sample period. The estimates for the Chan and Kwok (CK) approaches further build on the information contained in a sample of 42 countries which *never* experienced democracy during the sample period. 45 countries which remained democracies throughout the sample period are omitted from the analysis — the remaining two countries compared with the ANRR sample of 175 are dropped since they do not have data on the gross investment and trade covariates used in parts of the analysis. 'Pesaran & Smith' presents results for models which ignore (strong) cross-section correlation and/or uncommon pre-democratisation trends; 'CK w/ covariates' presents results for a model including country observations for gross investment and trade as covariates to the ANRR democracy dummy and the various cross-section averages detailed in the main text; 'CK' only includes democracy as observed regressand alongside cross-section averages as detailed in the maintext. A filled (white) marker indicates that the coefficient on democracy is statistically (in)significant at the 10% level.

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

The above setup is a little awkward since it ignores that some countries repeatedly switched regimes between democracy and autocracy, e.g. during the sample period Thailand democratised four times, Pakistan and Ghana three times.[19] I try to address this concern by estimating the average long-run democracy coefficient limiting the sample to countries which only transitioned into democracy *once* during the sample period (2,223 observations for 63 countries). Appendix figures suggest the exclusion of the repeated regime switchers moderately *raises* the average estimate of the long-run effect of democracy.

Taken together, these empirical exercises suggest that allowing for parameter heterogeneity in the democracy dummy as well as the GDP dynamics arrives at robust results for a long-run 'democratic dividend' of around 10%.

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] **Aghion, Phillipe, Alberto Alesina, and Francisco Trebbi (2008)** 'Democracy, Technology, and Growth.' In: Elhanan Helpman (ed) *Institutions and Economic Performance*. Cambridge, MA: Harvard University Press.

[3] **Andrews, Donald W.K. (2005)** 'Cross-section regression with common shocks.' *Econometrica* 73(5): 1551-1585.

[4] **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.

[5] **Bai, Jushan (2009)** 'Panel data models with interactive fixed effects.' *Econometrica* 77(4): 1229-1279.

[6] **Barro, Robert J. (1996)** 'Democracy and growth.' *Journal of Economic Growth* 1(1): 1-27.

[7] **Cervellati, Matteo, and Uwe Sunde (2014)** 'Civil Conflict, Democratization, and Growth: Violent Democratization as Critical Juncture.' *Scandinavian Journal of Economics* 116(2): 482-505.

[19] It should be noted that the same applies for the semi-parametric models in ANRR.

[8] **Chan, Marc K., and Simon Kwok (2018)** 'Difference-in-Difference when Trends are Uncommon and Stochastic.' Unpublished mimeo.

[9] **Cheibub, José Antonio, Jennifer Gandhi, and James Raymond Vreeland (2010)** 'Democracy and dictatorship revisited.' *Public Choice* 143(1-2): 67-101.

[10] **Chen, Shuowen, Victor Chernozhukov and Ivan Fernandez-Val. 2019.** 'Causal Impact of Democracy on Growth: An Econometrician's Perspective.' Paper presented at the 2019 ASSA Meetings in Atlanta, GA.

[11] **Chudik, Alexander, Kamiar Mohaddes, M. Hashem Pesaran, and Mehdi Raissi (2016)** 'Long-run Effects in Large Heterogeneous Panel Data Models with Cross-Sectionally Correlated Errors.' In: Gloria Gonzalez-Rivera, R. Carter Hill, Tae-Hwy Lee (Eds.) *Advances in Econometrics: Essays in Honour of Aman Ullah*, Emerald Group Publishing, pp.85-135.

[12] **Chudik, Alexander, and M. Hashem Pesaran (2015)** 'Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors.' *Journal of Econometrics* 188(2): 393-420.

[13] **Eberhardt, Markus, and Andrea Presbitero (2015)** 'Public debt and growth: Heterogeneity and non-linearity.' *Journal of International Economics* 97: 45-58.

[14] **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.

[15] **Gerring, John, Philip Bond, William T. Barndt, and Carola Moreno (2005)** 'Democracy and economic growth: A historical perspective.' *World Politics* 57(3): 323-364.

[16] **Giavazzi, Francesco, and Guido Tabellini (2005)** 'Economic and political liberalizations.' *Journal of Monetary Economics* 52(7): 1297-1330.

[17] **Gibbons, Charles, Juan Carlos Suarez Serratoz, and Michael Urbancic (2018)** 'Broken or Fixed Effects?' *Journal of Econometric Methods*, forthcoming.

[18] **Gobillon, Laurent, and Thierry Magnac (2016)** 'Regional policy evaluation: Interactive fixed effects and synthetic controls.' *Review of Economics and Statistics* 98(3): 535-551.

[19] **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.

[20] **Hamilton, Lawrence (1992)** How robust is robust regression? *Stata Technical Bulletin* 1(2).

[21] **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.

[22] **Madsen, Jakob B., Paul A. Raschky, and Ahmed Skali (2015)** 'Does democracy drive income in the world, 1500- 2000?' *European Economic Review* 78: 175-195.

[23] **Papaioannou, Elias, and Gregorios Siourounis (2008)** 'Democratisation and growth.' *Economic Journal* 118(532): 1520-1551.

[24] **Persson, Torsten, and Guido Tabellini (2006)** 'Democracy and development: The devil in the details.' *American Economic Review, Papers & Proceedings* 96(2): 319-324.

[25] **Pesaran, M. Hashem (2006)** 'Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure.' *Econometrica* 74(4): 967-1012.

[26] **Pesaran, M. Hashem, and Ron P. Smith (1995)** 'Estimating long-run relationships from dynamic heterogeneous panels.' *Journal of Econometrics* 68(1): 79-113.

[27] **Phillips, Peter C.B., and Donggyu Sul (2003)** 'Dynamic panel estimation and homogeneity testing under cross section dependence.' *Econometrics Journal* 6(1): 217-259.

[28] **Rodrik, Dani, and Romain Wacziarg (2005)** 'Do democratic transitions produce bad economic outcomes?' *American Economic Review, Papers & Proceedings* 95(2): 50-55.

[29] **Sul , Donggyu (2016)** 'Pooling is harmful sometimes.' Unpublished mimeo, University of Texas at Dallas.

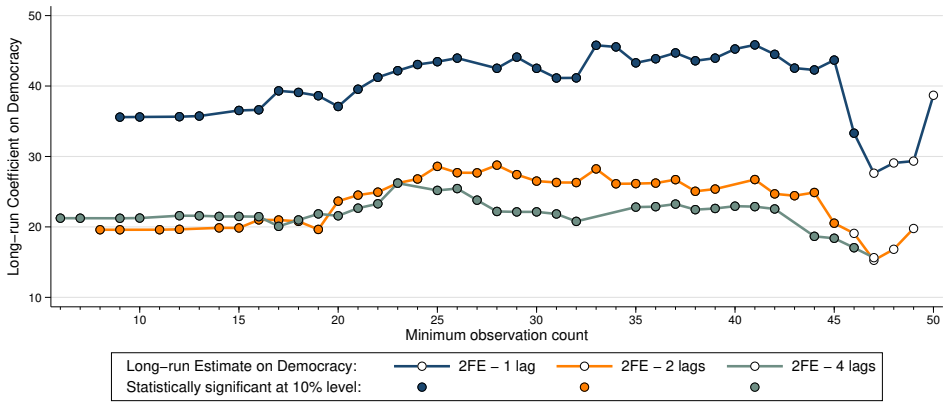
[30] **Xu, Yiqing (2017)** 'Generalized synthetic control method: Causal inference with interactive fixed effects models.' *Political Analysis* 25(1): 57-76.

[31] **Young, Alwyn (2018)** 'Consistency without Inference: Instrumental Variables in Practical Application.' Unpublished mimeo, June.

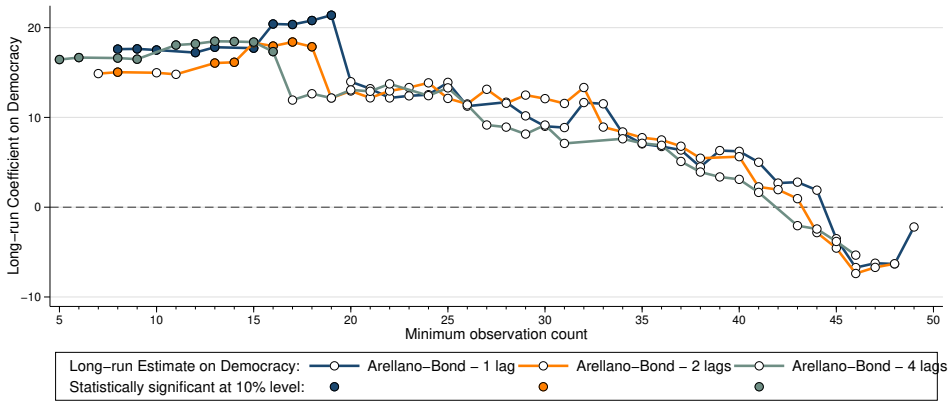
Appendix

A Sample Reduction Exercises – Full Parametric Results

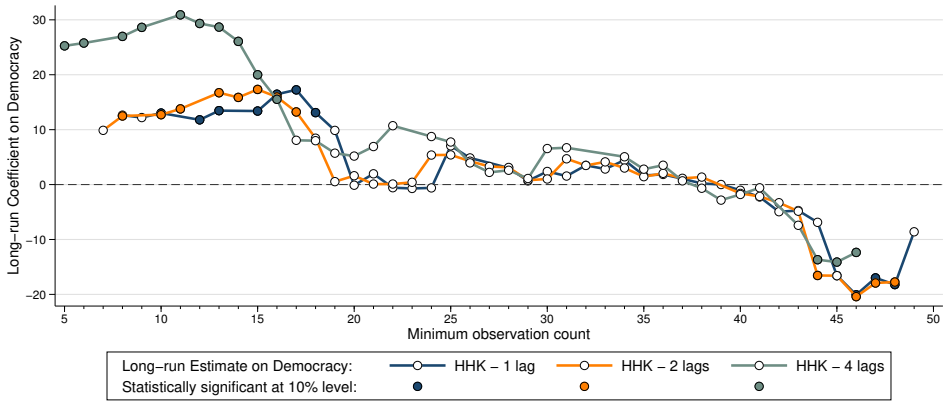
Figure A-1: Sample Reductions by T_i count (FE, AB, HHK)



(a) Within Estimates



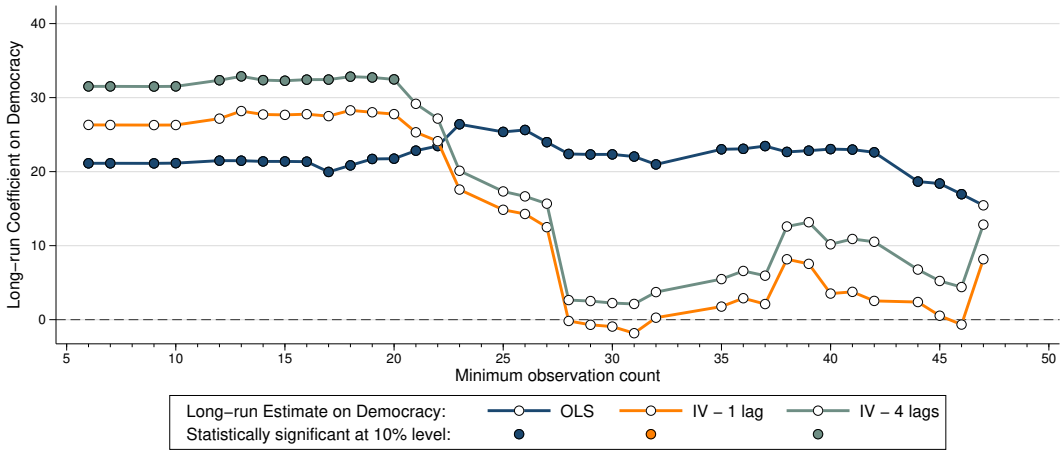
(b) Arellano-Bond Estimates



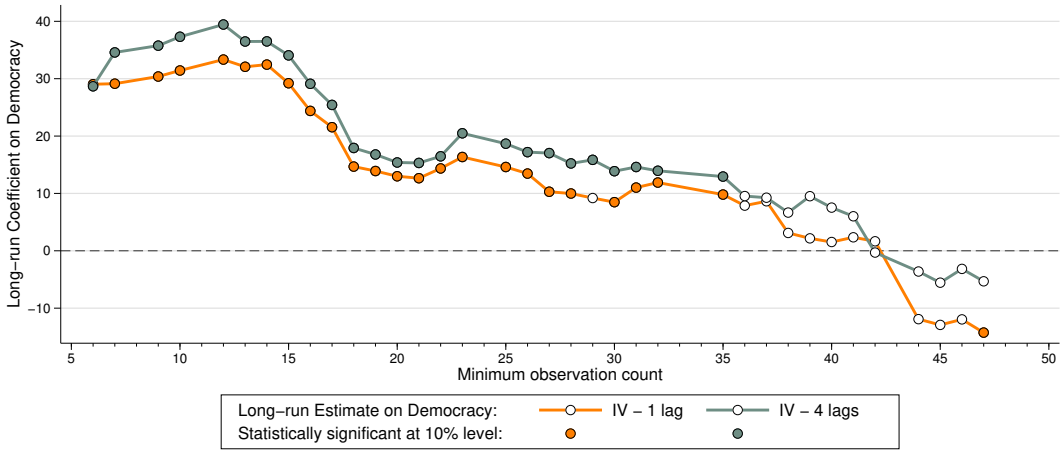
(c) Hahn, Hausman & Kuersteiner (HHK) Estimates

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

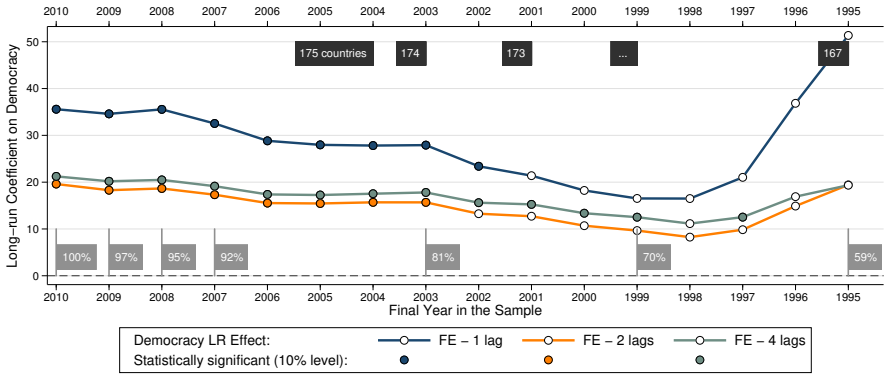
Notes: The figure presents long-run estimates for democracy from varying empirical samples for the (a) 2SLS, and (b) HHK-2SLS estimators. In each plot there are results for 2SLS using one and four lags of regional democratisation waves, respectively; plot (a) in addition reports simple OLS ('within') estimates for reference. The 'left-most' estimates for eachspecification 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

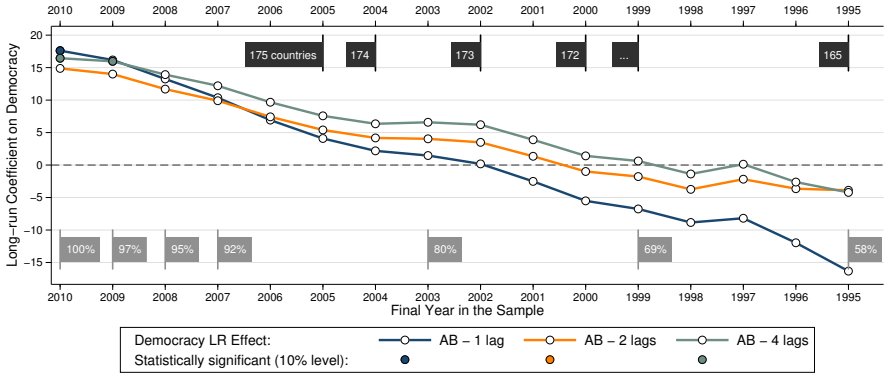
	Sample reduction by dropping countries (unsystematically)							
	[1]		[2]		[3]		[4]	
	2FE		AB		HHK		2SLS	
ANRR Reference	Table 2(3)		Table 2(7)		Table 2(11)		Table 6(2)A	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
Long-run effect of democracy	21.240 [7.215]***	15.637 [9.867]	16.448 [8.436]*	12.846 [8.023]	25.032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

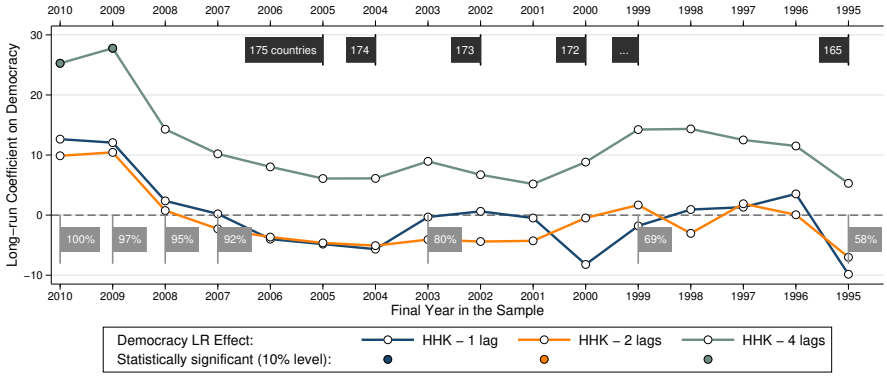
Figure A-3: Sample Reductions by end year (FE, AB, HHK)



(a) Within Estimates



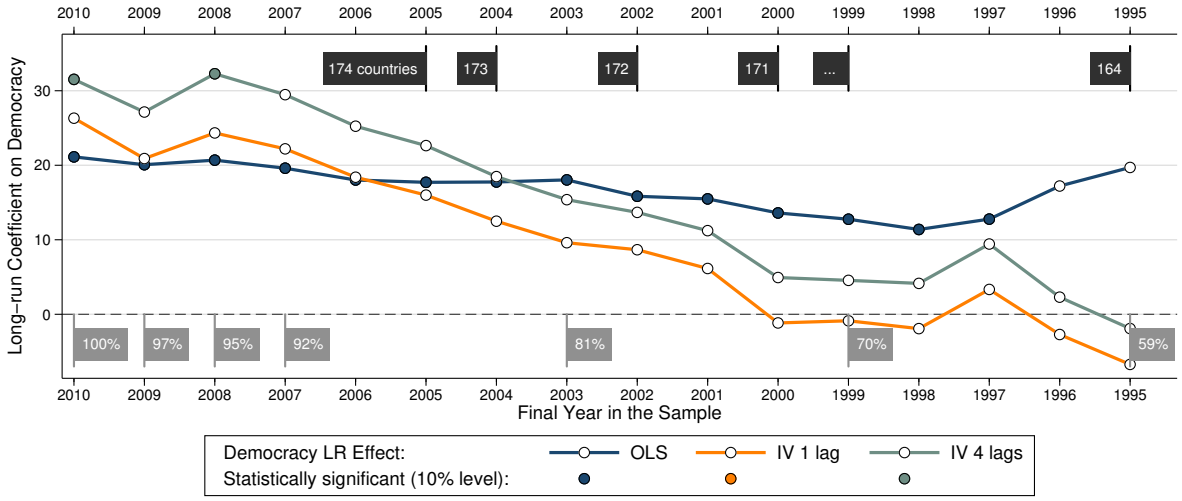
(b) Arellano-Bond (AB) Estimates



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

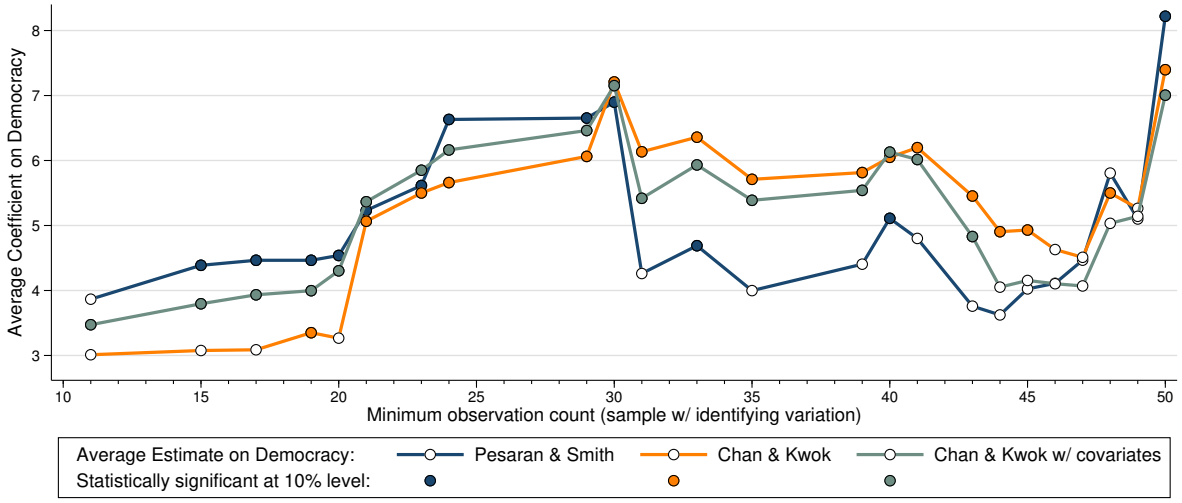
Notes: The figure presents long-run estimates for democracy from varying empirical samples for the (a) within, (b) AB, and (c) HHK estimators. The x-axis in each plot indicates the sample end year. A filled (white) circle marker indicates that the long-run coefficient is statistically (in)significant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

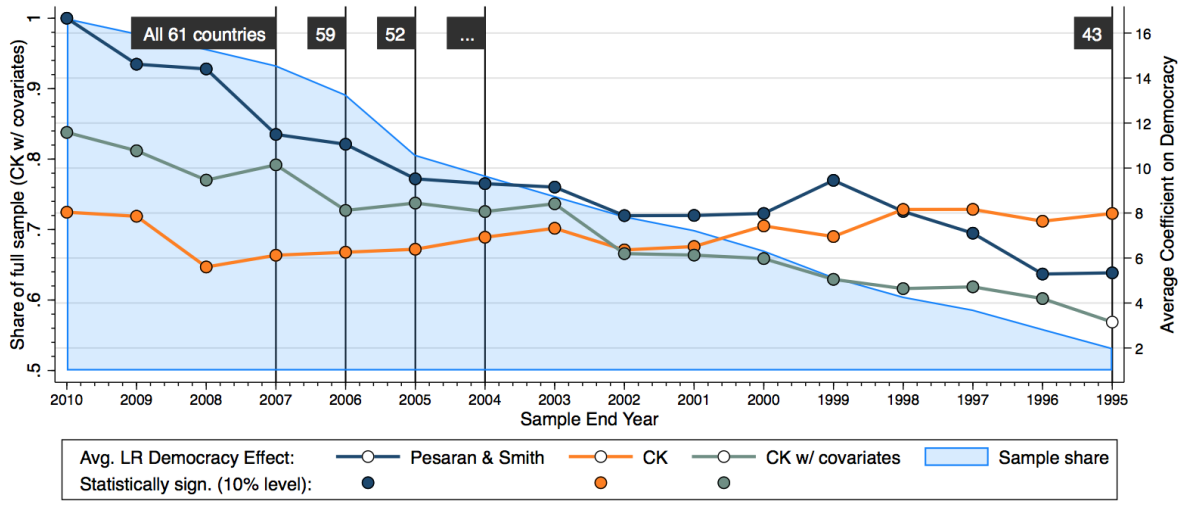


Notes: The figure presents long-run estimates for democracy from varying empirical samples for the 2SLS estimator. In addition to results for the the ANRR 2SLS specification with one and four lags of GDP I include the FE-4 lags specification for illustration (labelled 'OLS'). The former correspond to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

Figure A-5: Heterogeneous parameter models — static models



(a) Sample reductions by T_i count



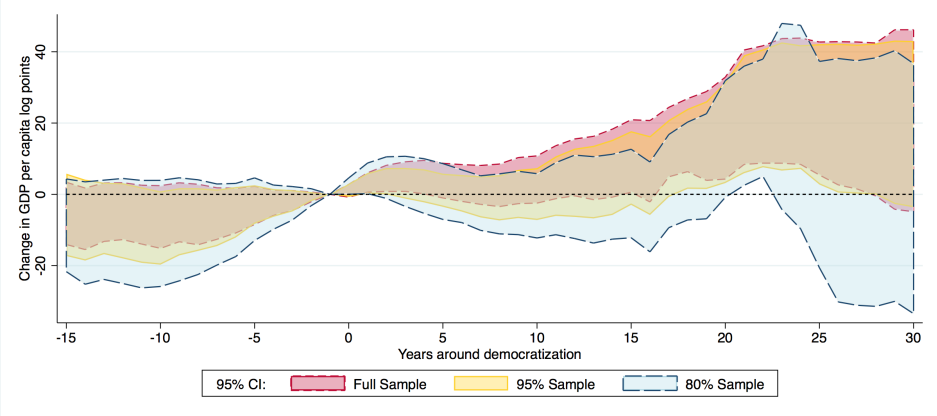
(b) Sample reductions by end year

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.

B Sample Reduction Exercises – Semi-Parametric Results

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

(a) Adopting Linear Regression

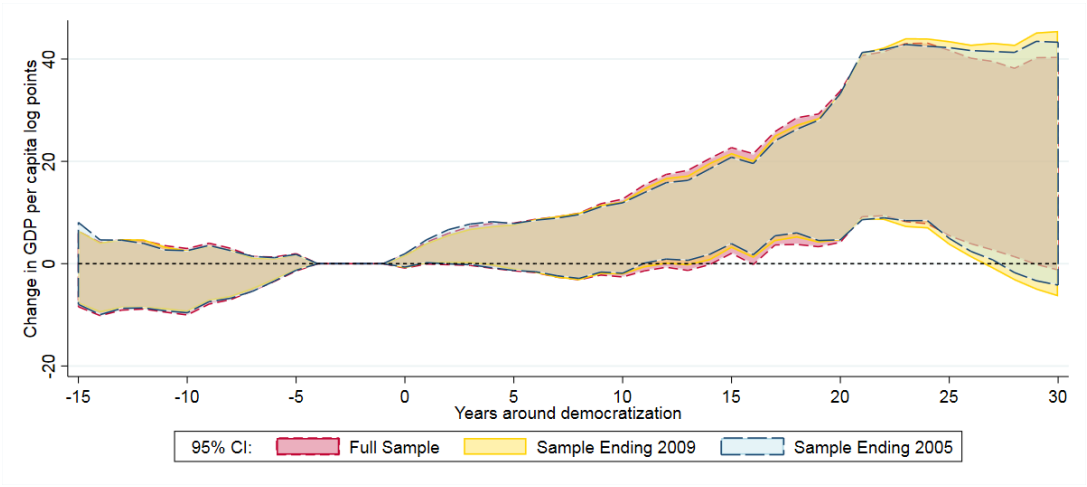


(b) Adopting Inverse Propensity Score Reweighting

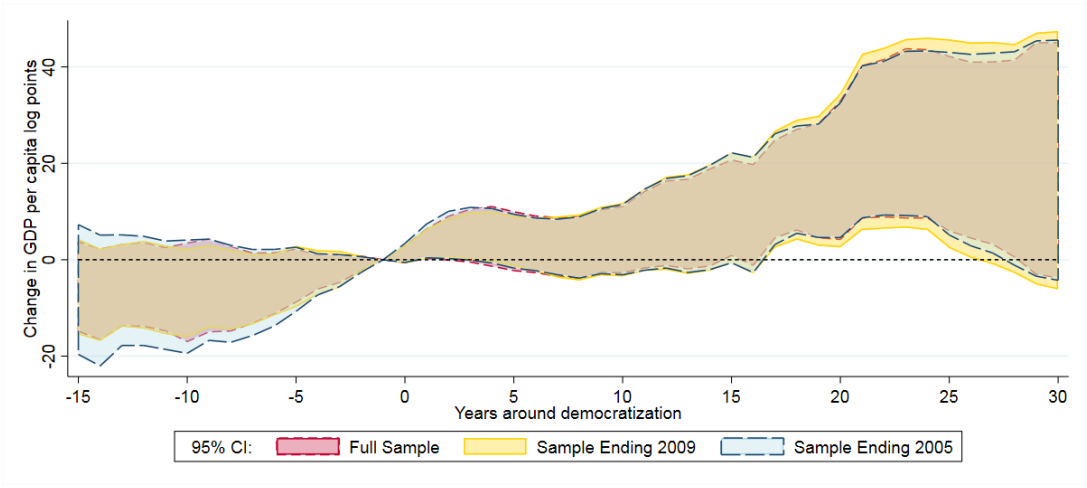
(c) Adopting Both Strategies

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).

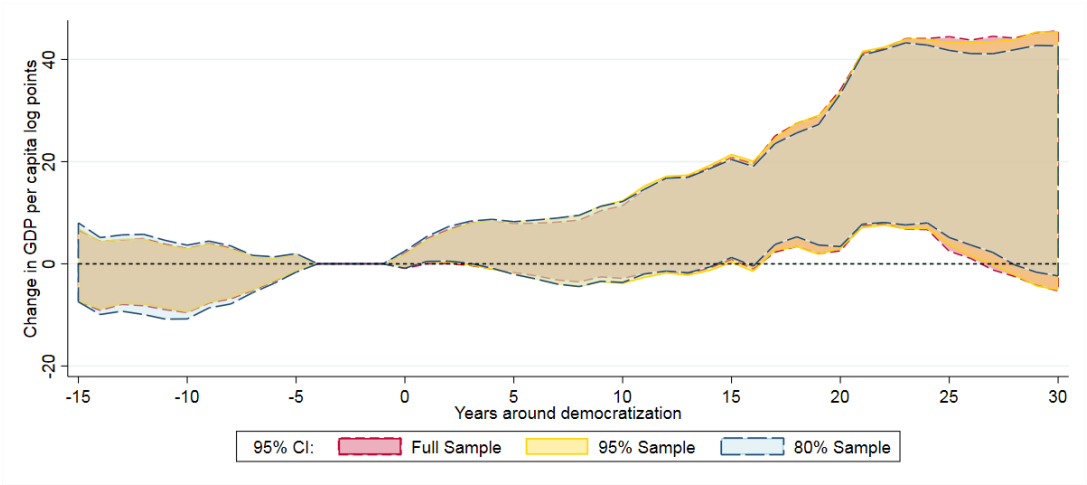
Figure B-2: Sample Reductions by end year — Semi-parametric models



(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).

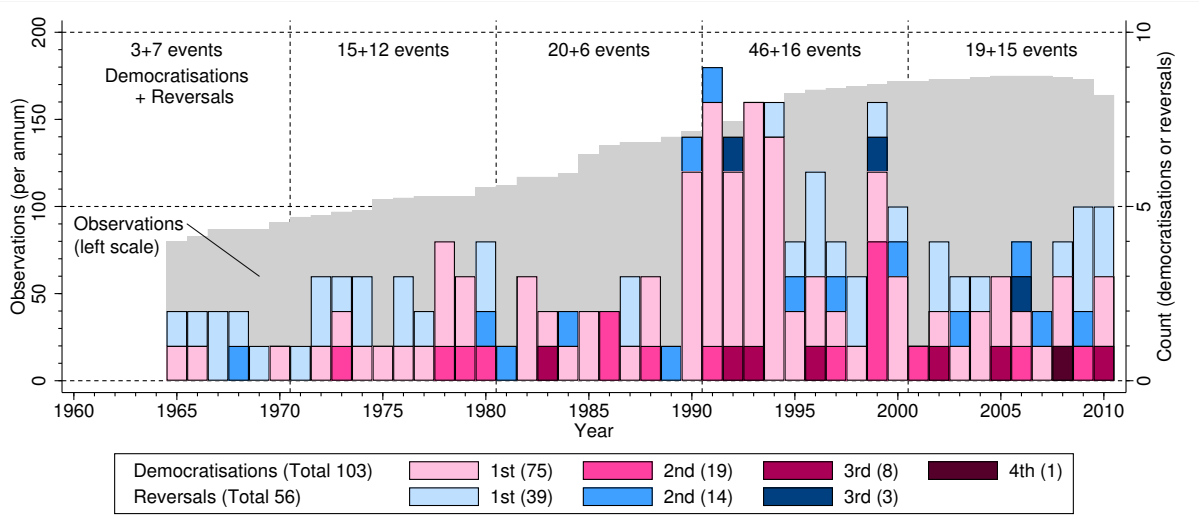
C Sample Reduction Exercises – Sample Makeup

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.

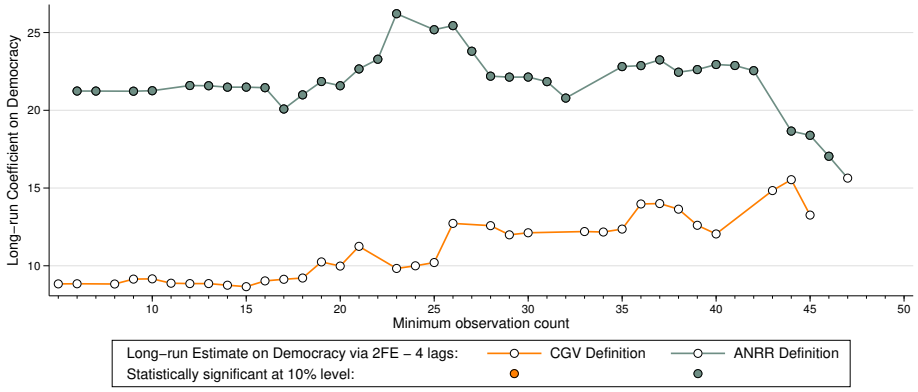
Figure C-1: Histogram – Full Sample – Democratisation and Reversal



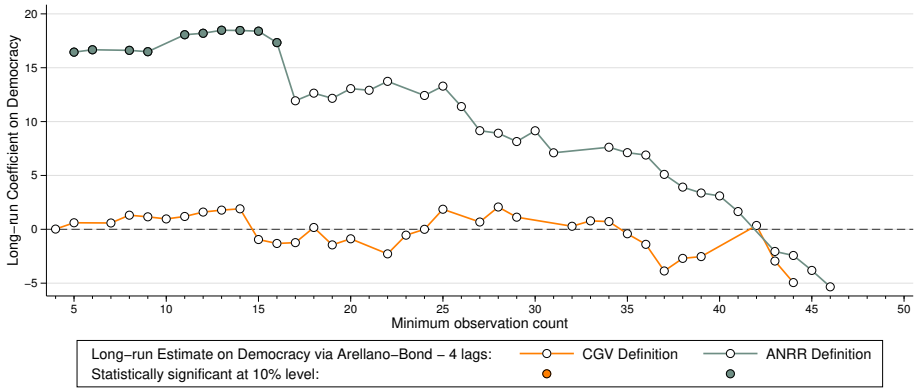
Notes: The figure presents the distribution of democratisation events (shades of pink) and reversals to autocracy (shades of blue) over the 1965 to 2010 period — the sample employed here is for the AB/HHK specification with four lags of GDP (hence the omission of 1960-64). The shades distinguish *first*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

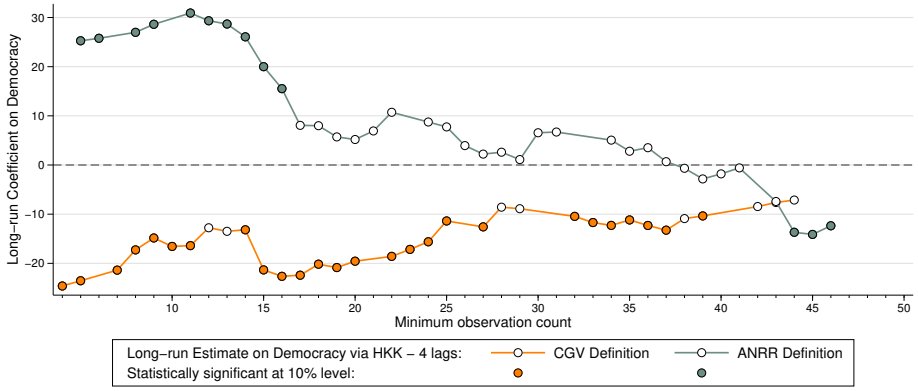
Figure D-1: Sample Reductions — Parametric Models (FE, AB, HHK)



(a) Within Estimates



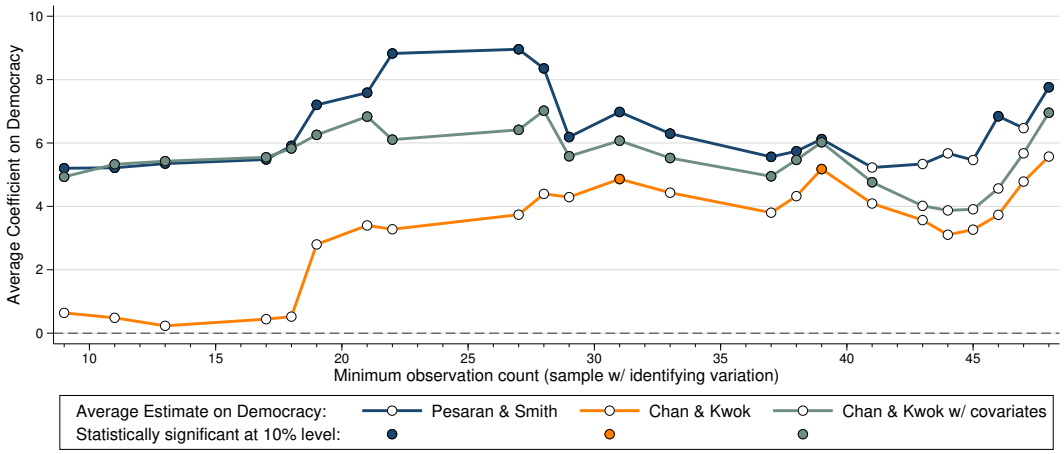
(b) Arellano-Bond Estimates



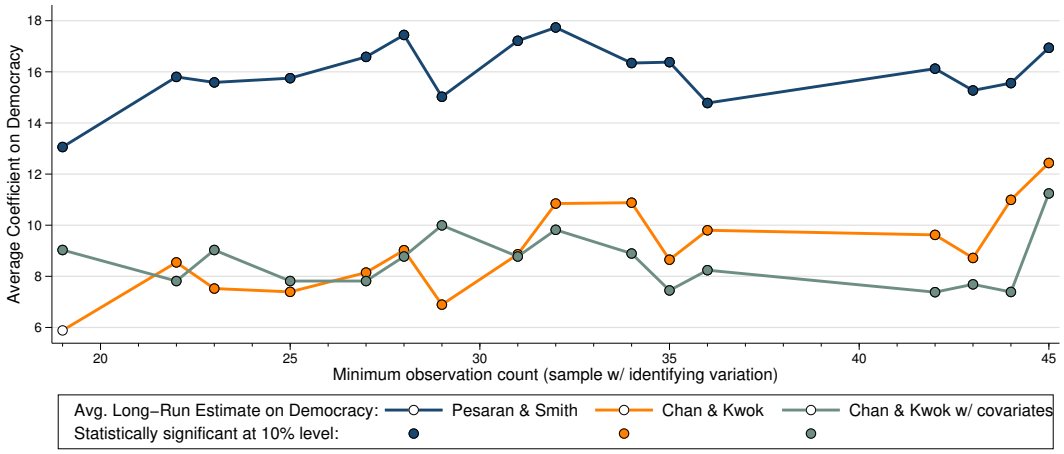
(c) Hahn, Hausman & Kuersteiner (HHK) Estimates

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 D-2: Sample Reductions — heterogeneous parameter estimators



(a) Static Models



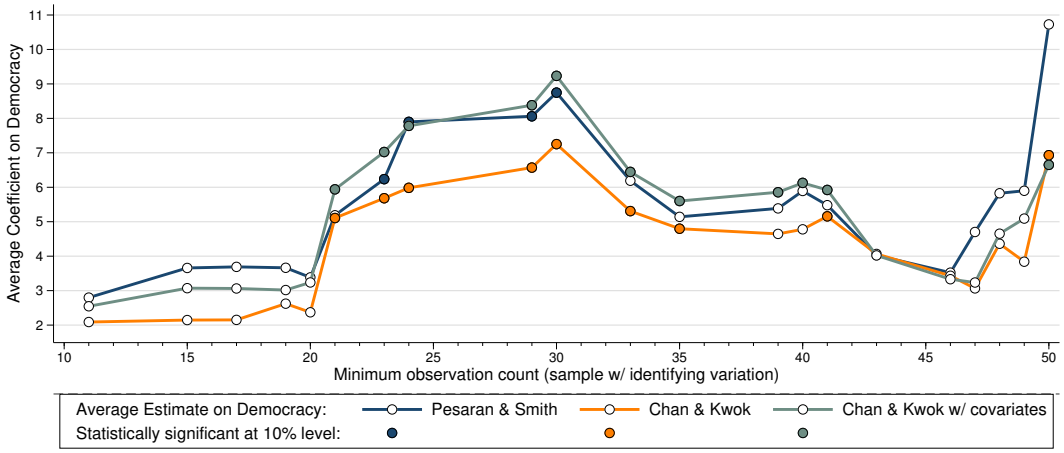
(b) Dynamic Models

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). This figure needs to be contrasted with Figure 4 in the maintext for a comparison with the results for the ANRR definition of democracy.

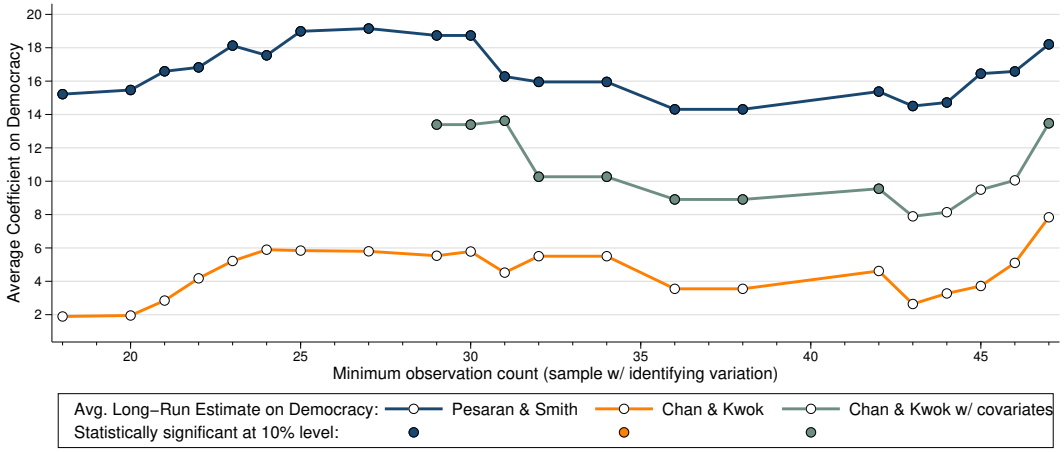
E Alternative Specifications

Figure E-1: Sample Reductions — heterogeneous parameter estimators;

excluding countries with more than one transition into democracy



(a) Static Models

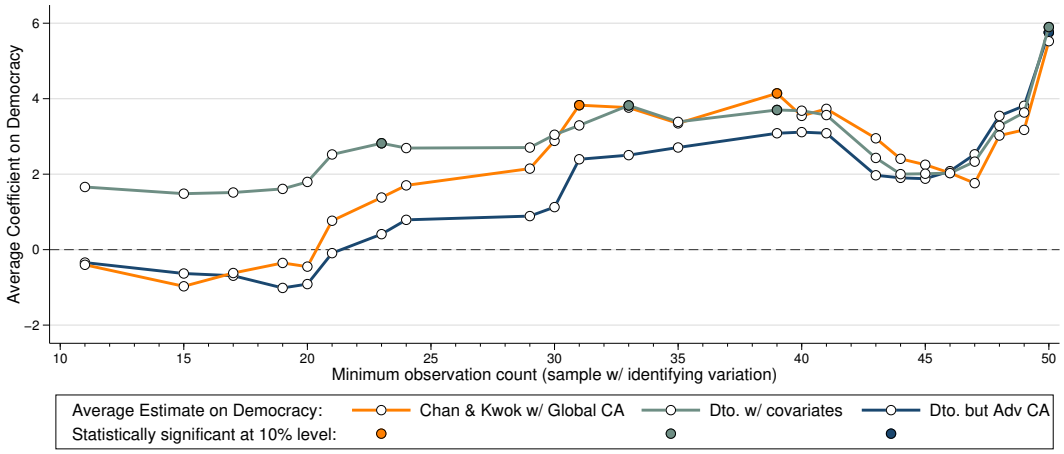


(b) Dynamic Models

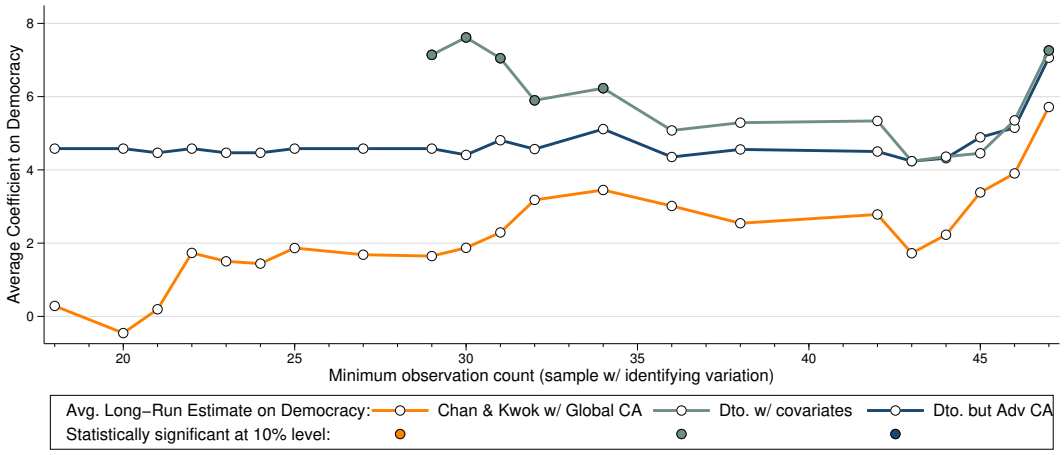
Notes: This figure provides sample reduction results for the static and dynamic Diff-in-Diff estimators using a sample which excludes countries which transitioned into democracy more than once during the sample period. The resulting sample analysed in the above plots is thus made up of 63 countries.

Figure E-2: Sample Reductions — heterogeneous parameter estimators;

adding 'global' cross-section averages (CA) or CA for countries which are always democracies



(a) Static Models

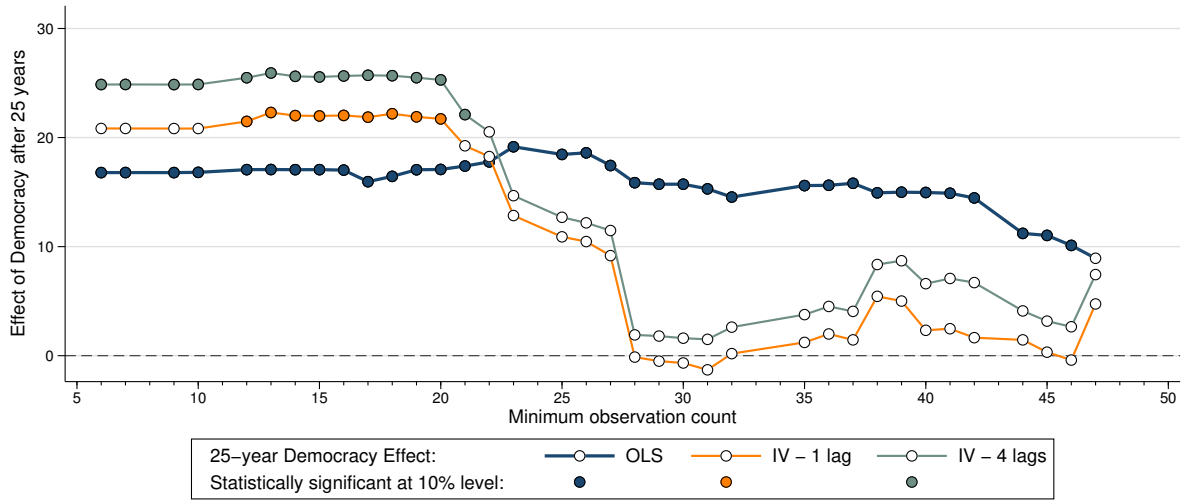


(b) Dynamic Models

Notes: This figure provides sample reduction results for the static and dynamic Diff-in-Diff estimators using a specificatio which includes a cross-section average of per capita GDP in all countries ('Global CA') or in those countries which were democracies throughout the sample period ('Adv CA').

F Alternative Definitions of the 'Long-run'

Figure F-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).