# Democracy Doesn't Always Happen Over Night: Regime Change in Stages and Economic Growth*

Vanessa Boese-Schlosser[1,2] and Markus Eberhardt[3,4]

[1] *Geschwister Scholl Institute of Political Science, Ludwig-Maximilians-University, Munich, Germany*

[2] *V-Dem Institute, University of Gothenburg, Sweden*

[3] *School of Economics, University of Nottingham, UK*

[4] *Centre for Economic Policy Research, UK*

Word Count: 6,581 plus two exhibits

This version: December 16, 2022

**Abstract:** How substantial are the economic benefits from democratic regime change? We argue that democratisation is often not a discrete event but a two-stage process: autocracies enter into 'episodes' of political liberalisation which eventually culminate in regime change or not. To account for this chronology and the implicit counterfactual groups, we introduce a repeated-treatment difference-in-difference implementation capturing non-parallel trends and selection into treatment. We find that modelling regime change in two stages rather than a single event yields stronger long-run growth effects. Among democratizers, experiencing repeated episodes without regime change reduces growth in democracy whereas length of episode does not.

**JEL Classification:** O10, P16

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

*Correspondence: Markus Eberhardt, School of Economics, Sir Clive Granger Building, University Park, Nottingham NG7 2RD, UK. Email: markus.eberhardt@nottingham.ac.uk

**Acknowledgements:** We thank Jan Fidrmuc, Scott Gehlbach, Jonas Hjort, James Rockey, Valeria Rueda and audience members at the online workshop on 'The Political Economy of Democracy and Dictatorship' (2021), NICEP (University of Nottingham), the 2021 ISA annual convention, the 2021 RCEA 'Future of Growth' conference, Nottingham School of Economics brownbag, the University of Birmingham, the 2021 International Applied Econometrics Association annual conference, and the 2021 International Panel Data Conference for comments and suggestions. We are grateful for the recommendations and guidance from three anonymous referees and the handling editor, which have substantially improved the paper. The usual disclaimers apply.

# 1 Introduction

When Nelson Mandela became President of South Africa in 1994, the country had successfully overcome Apartheid following a decades-long struggle by the African National Congress (ANC) using guerrilla tactics and mass mobilisation in the form of demonstrations, strikes and boycotts. Lifting the ban on the ANC in 1990 then-President F.W. de Klerk embarked on negotiations with Mandela on behalf of the white minority to safeguard their dominant position in South African politics but ultimately the country adopted universal suffrage and became an electoral democracy in which De Klerk served as Deputy President alongside Thabo Mbeki.

Such a drawn-out liberalisation process eventually culminating in democratic regime change is far from uncommon: in the 62 countries in our sample (1950-2014) which experienced regime change the median length of time spent undergoing such a 'democratisation episode' — we elaborate below on definitions and data sources — is four years. An additional 43 countries spent a median of six years in episodes but never experienced regime change.

Existing studies on the growth effects of democracy do not account for this chronology and hence cannot consider (i) whether growth performance varies when we assume different counterfactual samples, (ii) the implications of repeated and/or lengthy episodes for subsequent growth under democracy, or (iii) a comparison of the growth experiences during ultimately failed versus successful episodes.

The first contribution of this paper is to accommodate the chronology of democratisation as a process rather than a discrete event (e.g. Geddes 1999, Epstein et al. 2006) in the empirical analysis of the democracy-growth nexus: countries select into democratisation episodes, and some select out of these episodes into democratic regime change. Our approach is situated between studies which favour binary democracy indicators (e.g. Giavazzi & Tabellini 2005, Rodrik & Wacziarg 2005, Papaioannou & Siourounis 2008, Acemoglu et al. 2019) and others which favour continuous measures (e.g. Knutsen 2013, Murtin & Wacziarg 2014, Madsen et al. 2015) in analysing the economic implications of democratic change.

Our second contribution is that we include countries with failed attempts at democratic regime change as a separate control group in our empirical analysis and study their

growth experience during episodes. Including these countries in our analysis has important implications in terms of control group choice for the study of democracy and growth: we compare and contrast the long-run growth performance of successful democratisers between alternative 'counterfactual cases'. It enables us to distinguish between those nations which attempted liberalisation and those that did not, whereas conventional operationalisations capturing 'democratic transitions-as-events' combine these two groups as a supposedly homogeneous counterfactual case for successful regime change (Wilson et al. 2022). Our setup provide for a deeper investigation of the heterogeneous economic effects of regime change by analysing the implications of repeated and/or lengthy democratisation episodes. We can compare growth performance during episodes in the two types of countries and are further able to highlight systemic determinants why some episodes do not culminate in regime change ('failed episode').

The third contribution of this paper is methodological: we extend previous causal inference in a heterogeneous Difference-in-Difference framework to our proposed two-stage setup. In addition, we introduce a new way to present results by tying them closer to individual countries, rather than the average across or common estimate for all countries in the sample (ATET) as is standard in much of the literature: length of time spent in democracy varies greatly across countries, so that a pooled or Mean Group (Pesaran & Smith 1995) estimate would implicitly or explicitly average across some countries which experienced decades and others which only experienced a few years of democracy. Using running line regressions we show the central tendencies in estimated country treatment effects *relative to the length of time spent in democracy*. We can further account for some of the difficulties in sample make-up which arise in cross-country data: differential sample start dates and the regime change histories of individual countries. By conditioning on the frequency of democratisation episodes, the years spent in episodes, and their estimated effect on development this approach furthermore enables us to account for the two-stage nature of democratic change we advocate.

The distinction between democratisation episode and regime change is quantified in the Varieties of Democracy (V-Dem) Episodes of Regime Transformation (ERT) dataset (Maerz

et al. 2021, Edgell et al. 2020),[1] which we analyse for 1950-2014, covering 227 episodes and 70 regime changes in 105 countries.[2]

Our empirical approach builds on previous studies adopting difference-in-difference specifications (e.g. Giavazzi & Tabellini 2005, Persson & Tabellini 2006, Papaioannou & Siourounis 2008). The novel empirical implementation by Chan & Kwok (2022) we use and extend estimates the *country-specific* treatment effects and allows for non-parallel pre-treatment trends as well as endogenous selection into treatment. This follows recent developments which saw the introduction of common factors popular in the panel time series literature (Pesaran 2006, Bai 2009, Chudik & Pesaran 2013) to the empirics of policy evaluation (e.g. Gobillon & Magnac 2016, Xu 2017).

Our implementations employ estimates of unobserved common factors to capture the endogeneity arising from selection into treatment and non-parallel trends. The single-treatment model (henceforth Single PCDID) includes one treatment dummy (regime change) and estimated factors from one control group (autocracies). Our extension to a repeated treatment (henceforth Double PCDID) uses two treatment dummies (episode, regime change) and estimated factors from two control groups: (i) autocracies which never experienced an episode, and (ii) autocracies which experienced an episode but not regime change.

The adoption of heterogeneous treatment effect models is a crucial part of our empirics enabling us to provide new insights into the differences in the democracy-growth nexus across countries: existing research has near-unanimously assumed a common democracy-growth effect, yet the same literature recognises the potential for cross-country differences as motivated by arguments for a 'democratic legacy' (Gerring et al. 2005) or threshold levels in economic or human development as necessary conditions for a positive democracy-growth nexus (Aghion et al. 2007, Madsen et al. 2015, Acemoglu et al. 2019).[3]

---

[1] ERT also captures democratisation episodes in democratic regimes, which are dropped in our analysis.

[2] Our treated (control) sample comprises 62 (43) countries experiencing 141 (86) episodes, the median rate of 2 episodes per country is identical across samples. Appendix A provides details.

[3] See Eberhardt (2022) for a detailed motivation of the heterogeneous democracy-growth nexus.

Our analysis offers a number of important new insights: first, modeling democratisation as a two-stage process yields higher long-run economic growth than assuming regime change 'over night'. Second, the magnitude of the democratic growth dividend decreases with the number of episodes a country experienced, but not with their duration in years. Third, countries that fail to successfully complete democratisation episodes appear to gain no growth benefits from these episodes. This suggests that growth dividends derive from the successful completion of an episode, not from experiencing an episode *per se*. Finally, failed episodes are strongly associated with oil booms, pointing to a variant of the 'natural resource curse' in this political economy analysis.

The remainder of this paper is organised as follows: in the next section we discuss the conceptual foundations for political regime change as a non-binary event, introduce the data and present descriptive analyses of democratisation episodes and regime change. Section 3 covers the common factor model setup and the empirical implementations in greater detail. Results and robustness checks are presented in Section 4, in Section 5 we conclude and speculate about the 'geographic origins' of our findings.

# 2 Regime Change as a Two-Stage Process

## 2.1 Conceptual Development

Our empirics capture two elements of democratisation: first, the notion that the initiation and completion of democratic liberalisation and regime change *takes time* (the rationale for 'episodes'); and second, a concern over those nations which initiated a process of liberalisation but were unable or unwilling to translate this into regime change (the rationale for considering an alternative counterfactual to regime change).

Empirical studies of democratisation are commonly focused on the analysis of electoral autocracies, so-called 'hybrid regimes' (Diamond 2002, Brownlee 2009, Geddes et al. 2014). These authors appear to tacitly agree that democratisation is an event, a single moment of "dramatic upheaval" (Gunitsky 2014, 561) as in Huntington's (1991) 'democratic waves'.

Democratic transitions, however, are the result of a potentially lengthy process of political struggle between several actors (Rustow 1970, Acemoglu & Robinson 2006, Brownlee 2007, Graham & Quiroga 2012). Many formal models of nondemocratic politics can speak to this notion of the passing of time (Gehlbach et al. 2016): Liberalisation represents a period of uncertainty over the political trajectory of a country due to mass mobilisation or coalition formation. 'Cascading' protests and revolutionary movements take time to foment regime-busting power in the face of repression. Existing research in the comparative case study literature provides a self-preserving rationale for autocracies to engage in liberalisation (Magaloni 2008, Levitsky & Way 2010, Frantz & Kendall-Taylor 2014), although they might end up as democracies 'by mistake' (Treisman 2020). We can further draw on existing work on the rational delay to stabilisation policy (Alesina & Drazen 1991), status-quo bias in the implementation of economic reforms (Fernandez & Rodrik 1991), and the advantage of gradual economic reform under uncertainty (Dewatripont & Roland 1995) to motivate the notion of political liberalisation episodes which 'take time.' Hence, while regime change as 'dramatic upheaval' undoubtedly does occur, these arguments suggest that establishing the political institutions of democracy frequently does not happen over night.

The conceptual distinction between episodes and regime change directly links to our second concern over the suitable control groups at each stage. Recent work by Geddes et al. (2014) highlights the relative ignorance in the empirical literature towards democratisation events which did *not* result in regime change. Levitsky & Way (2010, 52) point to the record of democratic transition during the 1990s which makes "the unidirectional implications of the word 'transitional' misleading". These thoughts create probing questions for the empirical literature on the democracy-growth nexus employing binary representations of democratic regime change: this practice assumes that within-category subjects are homogeneous (Wilson et al. 2022) and hence all 'negative' cases of transition are lumped together. A single regime change dummy picks out the 'winners' of the liberalisation process, the null category contains the 'losers' and those who never tried.[4] What if this heterogeneity is key for un-

[4]In the literature using *continuous* democracy measures (e.g. Knutsen 2013, Murtin & Wacziarg 2014, Madsen

derstanding when democratic institutions foster economic growth? There is ample evidence for heterogeneous growth effects (Cervellati et al. 2014) and particularly so for autocracies (the main group of interest when studying transitions to democracy): the variation in growth outcomes is substantially higher among autocratic regimes, i.e. some autocracies have very high and others very poor growth outcomes (Persson & Tabellini 2009, Knutsen 2012). For the poorly-performing autocracies, democracy can act as a 'safety net' against disastrous economic outcomes (Knutsen 2021) and hence they may attempt to undergo a process of liberalisation, while in the former an autocracy can perhaps 'grow itself out of' demands for political liberalisation, like China arguably has done for the past three decades.

## 2.2 Data Sources, Variable Transformations

We use measures from the Episodes of Regime Transformation (ERT) dataset (Edgell et al. 2020), real per capita GDP and population from Bolt & van Zanden (2020, the 'Maddison data'), and exports and imports from Fouquin & Hugot (2016, TRADHIST). For comparison we also employ the Regimes of the World (Lührmann et al. 2018, ROW) democracy measure.[5]

We log-transform real per capita GDP and multiply this by 100: results are estimates of the percentage change in income following regime change. We use population growth and export/trade, aggregated from bilateral export and import flows, as controls: population growth is justified by the use of *per capita* GDP as dependent variable, while controlling for trade was found to substantially affect the magnitude of the estimated democracy effect (Papaioannou & Siourounis 2008, Table 3 [5]; Acemoglu et al. 2019, Table 6 [6]).[6]

We adopt the democratic regime transformation dummy from ERT alongside the democratization episode dummy. The former builds on the ROW categorisation of democracy but

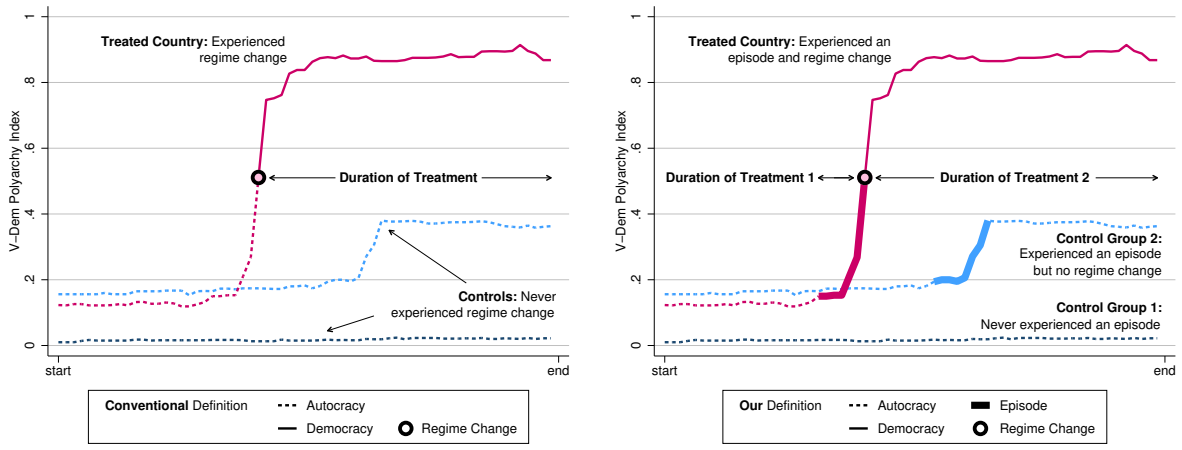
---

et al. 2015) failed liberalisations are likewise undistinguished.

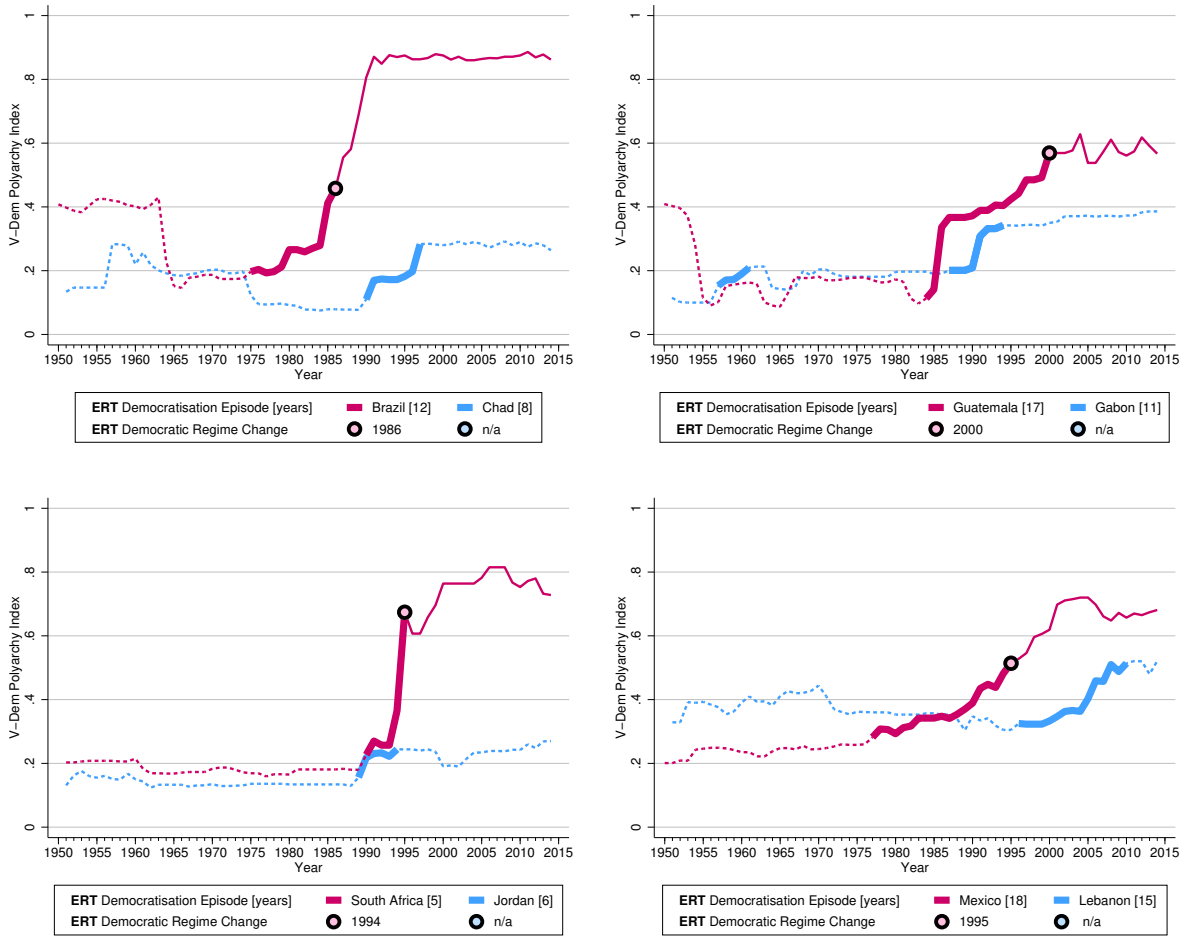
[5]Both, the ERT data and the ROW measure capture electoral democracy, i.e. free and fair elections, freedom of association and expression (Boese 2019).

[6]In robustness analysis we run the PCDID regressions without these controls.

**Figure 1:** Some (Stylised) Examples of Democratisation




(a) Conventional (left) vs Episode-Regime Change (right) Chronology of Democratisation




(b) Some Examples of Successful and Failed democratisation episodes

**Notes**: We present the V-Dem polyarchy index evolution for country pairs, where the country in dark pink experienced regime change and the country in light blue did not. The period highlighted by the thick line represents the democratisation episode, following ERT (the length of each episodes in years is indicated in the legend). The 'Eastern' end of the thick pink lines always coincides with the year of democratic regime change. A dashed (solid) thin line indicates the country regime is in autocracy (democracy) following the ERT definition. The circular marker indicates the year of democratic regime change (if applicable), which is required to include a 'founding election' (this explains the absence of regime change in Lebanon). We provide more examples in Appendix Figure A-2.

further requires a founding democratic election to occur. A democratisation episode[7] requires (i) a small increase (0.01) in the V-Dem polyarchy index[8] for a country classified as 'closed' or 'electoral autocracy' (following the ROW categorization: Lührmann et al. 2018); and (ii) a total increase of at least 0.1 in the same measure over the course of the episode. An episode ends after a final year with an increase of at least 0.01 if this is followed by a year-on-year drop of 0.03, a cumulative drop of 0.1 over several years, or a 5-year stasis. Appendix F provides results using a range of alternative parameter values to define episodes.

Figure 1 highlights the difference between thinking of democratisation as a binary event vs a two-stage process: Panel (a) contrasts the single treatment approach (left diagram), including the conflation of heterogeneous control groups, with the two-stage treatment approach suggested in this paper (on the right), highlighting democratisation episodes as first-stage treatments followed by democratic regime change as second-stage treatments along with respective control groups. Panel (b) charts the development of electoral democracy (V-Dem's polyarchy index) in four country pairs which experienced democratisation episodes (thick lines) but with differential outcomes (regime change, solid thin line, or not, dashed thin line) — Appendix Figure A-2 provides additional examples. These graphs demonstrate that the outcome of political episodes is uncertain: country pairs starting out with near-identical polyarchy scores in the 1950s at times end up at opposite ends of the scale in 2014.

All variables are available from 1901 to 2014, but here we limit our analysis to 1950-2014: our methodology, which relies on common factors extracted from two sets of control groups, would not yield reliable results for the longer panel since only a handful of countries

[7]Our analysis focuses on episodes of democratisation originating in autocracies. In order to obtain separate treatment effect estimates for episodes and regime changes we exclude episodes of democratic deepening from our analysis and adopt the ERT episode indicator for a 'liberalizing autocracy': our episode dummy always reverts to 0 in the first year of democracy.

[8]Polyarchy is also referred to as the Electoral Democracy Index. It is continuous, $\in [0, 1]$ and represents a minimal definition of democracy favored in political science (Teorell et al. 2019, Boese 2019). The 0.01 annual increment may seem small, 1% of the range of the index, yet between 1900 and 2018 over 70% of annual increments in the polyarchy index are between -0.01 and 0.01 (Wilson et al. 2022).

in the respective control groups have observations in the first half of the 20th century. This highlights that our approach forces us to consider the relative sample sizes of treated and various control groups — we regard this as a core strength of this methodology.

Our 1950-2014 sample covers 62 'treated' countries which experienced episodes and regime change (n=3,724), 43 autocratic countries which only experienced episodes (n=2,515; control group 2), and 15 autocratic countries which never experienced episodes (n=646; control group 1).[9] The median episode length in treated countries is four years (stdev. 3.3), and six years (stdev. 3.4) in countries where episodes did not lead to regime change; in either group there were a median of two episodes per country (stdev. 1.1). We provide descriptive statistics, graphs and further details on the three samples under analysis in Appendix A.

# 3 Empirical Strategies

In this section we introduce the novel empirical implementations we employ to study the economic effect of democratisation when regime change is modelled either as a single or a repeated 'treatment'. We discuss the Chan & Kwok (2022) Principal Component Difference-in-Difference estimator (Single PCDID) and our extension (Double PCDID) for these respective cases. The final part of the section introduces our novel presentation of heterogeneous treatment effects using predictions from running line regressions.

## 3.1 Single PCDID

In the Single PCDID approach democratisation is modelled as a binary event. The PCDID estimator allows for endogenous selection into regime change and potentially non-parallel pre-treatment trends between treated and non-treated (never-regime changing) countries. This is achieved by including estimated common factors — extracted via Principal Component

[9]We cannot use all 71 countries since nine of them have no pre-episodal observations which prevents separate identification of episode and regime change effects (see Appendix Table A-3).

Analysis (PCA) from a control sample regression[10] — in the treatment regression. The use of common factors has a long tradition in the macro panel literature to capture strong cross-section dependence (e.g. Pesaran 2006, Bai 2009), a form of unobserved, time-varying heterogeneity.[11] The most recent contributions extended the use of common factors to the empirics of policy evaluation (Gobillon & Magnac 2016, Xu 2017, Chan & Kwok 2022).

Assumptions of parallel trends and exogenous treatment in standard treatment effects models are violated if time-varying unobservables are correlated with the treatment variable.[12] Chan & Kwok's (2022) PCDID estimator by-passes these problems by adding proxies (estimated factors) for the time-varying unobservables as additional controls in heterogeneous treatment regressions.[13]

Since these factors are added to *country-specific* regressions, the proxied unobservables can have a different impact across countries. Most importantly: treated and control country outcomes can have different trends. Furthermore, because these factors can be correlated with the treatment variable, we can suggest that democratic regime change can be correlated with unobserved determinants of economic development (e.g. absorptive capacity, culture): regime change can be endogenous. We now discuss this more formally.

---

[10] The principal components are estimated from the residuals of a country-specific regression of income per capita on export/trade, population growth and an intercept. An alternative version omits these covariates.

[11] Strong cross-section correlation is distinct from weaker forms of dependence (e.g. spatial correlation) and can lead to omitted variable bias in the estimated coefficients (Phillips & Sul 2003, Andrews 2005).

[12] As an analogy, estimating production function regressions using OLS, the presence of unobserved TFP creates biased estimates for capital and labour due to the correlation between TFP and these inputs.

[13] The basic intuition for the PCDID follows that of the control function approach in microeconometric analysis of production functions (Olley & Pakes 1996) which use combinations of observed choice variables (like material inputs) to construct a proxy of how firms react to changes in unobserved TFP. Continuing the analogy, the common factor structure, $f$, in combination with a heterogeneous parameter regression, $\mu_i$, can proxy time-varying, heterogeneous TFP and hence eradicates the omitted variable bias problem (see Eberhardt & Teal 2011).

**Setup** Using the potential outcomes framework, the observed outcome of a single treatment $D_{it}$ for panel unit $i$ at time $T_0$ can be written as

$$y_{it} = D_{it}y_{it}(1) + (1 - D_{it})y_{it}(0) = \Delta_{it}\mathbf{1}_{\{i\in E\}}\mathbf{1}_{\{t>T_{0i}\}} + y_{it}(0) \tag{1}$$

$$\text{with} \qquad y_{it}(0) = \varsigma_i + \beta_i' x_{it} + \mu_i' f_t + \widetilde{\epsilon}_{it}, \tag{2}$$

where the indicator variables $\mathbf{1}_{\{\cdot\}}$ are for the panel unit and the time period treated, respectively, $\Delta_{it}$ is the time-varying heterogeneous treatment effect, $x$ the observed covariates with associated country-specific parameters $\beta_i$,[14] $\mu_i' f_t$ represents a set of unobserved common factors $f_t$ with country-specific factor loadings $\mu_i$, and $\widetilde{\epsilon}_{it}$ is the error term.

The treatment effect is assumed to decompose into $\Delta_{it} = \overline{\Delta}_i + \widetilde{\Delta}_{it}$, with $E(\widetilde{\Delta}_{it}|t > T_{0i}) = 0 \ \forall i \in E$ since $\widetilde{\Delta}_{it}$ is the demeaned, time-varying idiosyncratic component of $\Delta_{it}$; we refer to $\overline{\Delta}_i$ as ITET, the individual treatment effect averaged over the treatment period — our key parameter of interest. The reduced form model is

$$y_{it} = \overline{\Delta}_i\mathbf{1}_{\{i\in E\}}\mathbf{1}_{\{t>T_{0i}\}} + \varsigma_i + \beta_i' x_{it} + \mu_i' f_t + \epsilon_{it} \qquad \text{with} \qquad \epsilon_{it} = \widetilde{\epsilon}_{it} + \widetilde{\Delta}_{it}\mathbf{1}_{\{i\in E\}}\mathbf{1}_{\{t>T_{0i}\}}, \tag{3}$$

where given the treatment effect decomposition the composite error $\epsilon_{it}$ has zero mean but can be heteroskedastic and/or weakly dependent (spatially/serially correlated).

The combination of common factors and heterogeneous parameters allows for potentially non-parallel trends across panel units, most importantly between treated and control units. The above setup can accommodate endogeneity of treatment $D_{it}$ in the form of *inter alia* correlation between treated units and factor loadings, the timing of treatment and factor loadings, or between observed covariates and timing or units of treatment. Finally, the implementation allows for nonstationary factors $f_t$.

---

[14] As common in the literature (Pesaran 2006) we assume $\beta_i = \bar{\beta} + \tilde{\beta}_i$ where $E(\tilde{\beta}_i) = 0$ . Covariates $x$ and factors $f$ can be orthogonal or correlated.

**Implementation** The estimation of the country-specific treatment effect (ITET) $\overline{\Delta}_i$ proceeds in two steps: first, using PCA, we estimate proxies of the unobserved common factors from data in a control group equation; second, country-specific least squares regressions of treatment group countries are augmented with these factor proxies as additional regressors.

The estimation equation for treated country $i \in E$ is then:

$$y_{it} = b_{0i} + d_i \mathbf{1}_{\{t>T_{0i}\}} + a_i' \hat{f}_t + b_{1i}' x_{it} + u_{it}, \tag{4}$$

where $\hat{f}$ are the estimated factors obtained by PCA on the residuals $\hat{e}$ from the heterogeneous regression $y_{it} = b_{0i} + b_{1i}' x_{it} + e_{it}$ in the control group sample, and $d_i$ is the country-specific parameter of interest. We estimate (4) augmented with one to six common factors. See Section 3.3 for inference.

**Assumptions** The main assumptions required for the consistency of ITET estimates are that the unobservables can be represented by a low-dimensional multi-factor error structure,[15] $\mu_i' f_t$ (as in Pesaran 2006, Bai 2009, Athey et al. 2021), and that $\epsilon$ is orthogonal to all conditioning components in equation (4): this implies that all aspects of treatment endogeneity and non-parallel trends are assumed to be captured by the factors, the controls, and the deterministic term as well as their combinations/correlation with the treatment variable. We discuss threats to identification and how we test for these below.

## 3.2 Double PCDID

The 'double-treatment' case argues for democratic regime change as a repeated selection problem: (i) At time $T_0$ an autocracy starts democratic liberalisation, i.e. it endogenously selects into a democratisation episode as defined by ERT. The control group for this first

[15] Since factor proxies are measured with error, the idiosyncratic errors $\widetilde{\epsilon}_{it}$ of treated and non-treated units may be correlated — this asymptotic bias is removed with a condition that $\sqrt{T}/N_C \to 0$, where $T$ is the time series dimension of the treated sample and $N_C$ is the number of control sample units.

treatment are all autocracies which never experience an episode. (ii) Of those autocracies which experienced a democratisation episode we find two types: first, those which successfully transitioned into democracy, and second, those which failed. From the pool of autocracies experiencing an episode we thus have a country which at time $T_1$ endogenously selects into democratic regime change as defined by ERT. The control group for this second treatment constitutes all autocracies with at least one episode but which never transition into democracy. We posit that countries that tried and failed in their quest for democracy are a interesting and meaningful control group for countries which successfully transitioned.

Correcting for repeated treatment requires the use of estimated common factors from two control groups. The two sets of common factors account for non-parallel trends prior to the two treatments, and in analogy to the single treatment case above, these common factors can be correlated with treatments or observed covariates, amounting to treatment endogeneity.

**Setup** We extend the PCDID to a repeated-treatment 'Double PCDID' specification:

$$\begin{aligned} y_{it} &= \Delta_{it}^{\mathsf{A}} \mathbf{1}_{\{i \in E^*\}} \mathbf{1}_{\{t > T_{0i}\}} + \Delta_{it}^{\mathsf{B}} \mathbf{1}_{\{i \in E^*\}} \mathbf{1}_{\{t > T_{1i} > T_{0i}\}} \\ &\quad + \varsigma_i + \beta_i' x_{it} + \mu_i^{\mathsf{A}\prime} f_t^{\mathsf{A}} + \mu_i^{\mathsf{AB}\prime} f_t^{\mathsf{AB}} + \tilde{\epsilon}_{it}. \end{aligned} \tag{5}$$

We now distinguish two treatments: A for the treatment at $T_0$ and B for a second, later treatment at $T_1 > T_0$, yet conditional on having received treatment A. The treatment group is now made up of those panel units which experienced *both* treatments ($i \in E^*$). In analogy there are now two control groups: (1) all those units which never experienced treatment A, and (2) those units which experienced treatment A but not treatment B (see below for notation). We now assume two sets of multi-factor error terms: one for each counterfactual group. The reduced form is

$$\begin{aligned} y_{it} &= \overline{\Delta}_i^{\mathsf{A}} \mathbf{1}_{\{i \in E^*\}} \mathbf{1}_{\{t > T_{0i}\}} + \overline{\Delta}_i^{\mathsf{B}} \mathbf{1}_{\{i \in E^*\}} \mathbf{1}_{\{t > T_{1i} > T_{0i}\}} \\ &\quad + \varsigma_i + \beta_i' x_{it} + \mu_i^{\mathsf{A}\prime} f_t^{\mathsf{A}} + \mu_i^{\mathsf{AB}\prime} f_t^{\mathsf{AB}} + \epsilon_{it} \end{aligned} \tag{6}$$

using similar arguments as in the single intervention case. The assumptions from the Single PCDID case extend to this model.

**Implementation** The estimation of the regime change ITET $\overline{\Delta}_i^B$ again proceeds in two steps: first, using PCA we separately estimate proxies of the common factors in the two control groups; second, the estimation equation for treated country $i \in E^*$ is

$$y_{it} = b_{0i} + d_i^A \mathbf{1}^A_{\{t>T_0\}} + d_i^B \mathbf{1}^B_{\{t>T_1>T_0\}} + a_{1i}^{\cancel{A}\prime} \hat{f}_t^{\cancel{A}} + a_{2i}^{\cancel{AB}\prime} \hat{f}_t^{\cancel{AB}} + b_{1i}' x_{it} + e_{it}, \tag{7}$$

where $d_i^A$ and $d_i^B$ are the country-specific treatment parameters for episodes and regime change. The $\hat{f}$ with superscript $\cancel{A}$ are the estimated factors obtained by PCA from the residuals $\hat{e}$ of a heterogeneous regression $y_{it} = b_{0i} + b_{1i}' x_{it} + e_{it}$ in the first control group. The $\hat{f}$ with superscript ($\cancel{AB}$) are estimated from the residuals of the following regression in the second control group: $y_{it} = b_{0i} + d_i^A \mathbf{1}^A_{\{t>T_{0i}\}} + a_{1i}^{\cancel{A}\prime} \hat{f}_t^{\cancel{A}} + b_{1i}' x_{it} + e_{it}$, where the presence of the episode dummy and the $\hat{f}_t^{\cancel{A}}$ accounts for the endogeneous selection of countries into episodes.[16] We estimate (7) with one to six common factors extracted from each control group. See Section 3.3 for inference.

**Threats to Identification** One concern is the effect of idiosyncratic shocks which may induce countries to trigger regime change: a country experiencing a democratisation episode may transition to democracy because of a fortunate natural resource discovery, or it might have been hindered by a financial crisis or natural disaster. We know that oil exploration is guided by global prices, while financial crises have sizeable international dimensions (Cesa-Bianchi et al. 2019, Arellano et al. 2017) — all arguments in favour of our factor structure. In Appendix C we run separate event analyses for GDPpc growth and change in V-Dem's polyarchy index in treatment and control samples adopting event dummies constructed from data collated by Reinhart & Rogoff (2009), Cotet & Tsui (2013), Laeven & Valencia (2020)

[16] We are grateful to a referee for pointing out that our original auxiliary regression would lead to inconsistent estimates of $\hat{f}_t^{\cancel{AB}}$.

and EM-Dat. These suggest no systematic differences in the effects between the two groups.

We also study the parallel trend assumption in adjusting the parallel trend test under factor structure by Chan & Kwok (2022) for our Double PCDID setup. Appendix D introduces this test and results.

Finally, although we know that adding 'too many' estimated factor in principle does little harm to our treatment estimates (Moon & Weidner 2015), the Double PCDID requires substantially more degrees of freedom and we check its robustness using a range of factor augmentations in Appendix Figure B-2.

## 3.3 Heterogeneous Treatment Effects and Inference

We estimate both of our models country by country. Consequently, the Single and Double PCDID models yield $N$ country-specific treatment estimates (INET) for regime change. A typically useful estimate to present is the ATET, which in our setup would be $\overline{\overline{\Delta}} = E(\overline{\Delta}_i)$, the average of the ITET across treated units $i \in E$ or $E^*$.[17] Focusing on the ATET would make sense when studying a treatment effect that manifests itself *in its entirety* after a small number of years, as would be the case for many medical interventions.[18] In the context of the democracy-growth nexus we propose an alternative means of presentation, namely predictions from running line plots of the estimated ITET for democratic regime change, $\hat{d}_i$ or $\hat{d}_i^B$ for Single and Double PCDID, relative to the years of treatment.

We chose this form of presentation since the effect of democracy on growth potentially differs within countries over time: New democracies may suffer from 'democratic overload', drawn to short-termism, and with too many processes not yet formalised they frequently represent "boisterous, obstreperous affairs" (Gerring et al. 2005, 335). But *over time*, politi-

[17] Results for the ATET from Single and Double PCDID models are presented in Appendix Table B-1.

[18] We also point to the recent insights regarding the decomposition of a 'pooled' DID ATET estimate in the context of variation in treatment timing (Goodman-Bacon 2021). Heterogeneous estimators do not face similar ambiguities of interpretation (weighting) and our running line regressions put 'treatment length' (early vs late treatment) at the heart of the results.

cians, bureaucrats and citizens learn how democracy works, while decisions and bureaucratic processes become formalised and hence predictable (ibid).

A running line regression smooths the dependent variable against an independent variable by using subsets of nearest neighbours in local linear regressions. Using predictions from *multivariate* running line regression allows us to simultaneously smooth on *multiple* independent variables. This form of presentation has a number of advantages: (i) we do not average across different countries with dozens or just a few years in democracy; (ii) we can account for differential sample observations and for multiple regime changes in each country;[19] and, for the Double PCDID, (iii) we can condition on the novel two-stage setup advocated here, by controlling for the number of episodes, the years spent in these episodes, and the magnitude of the episode effect $\hat{d}_i^A$.

In analogy to a standard Mean Group estimator, the ATET in the Chan & Kwok (2022) PCDID is simply the average across all treated units, $\hat{d}^{MG} = N^{-1} \sum_i \hat{d}_i$, with a nonparametric variance estimator following Pesaran (2006): $\widehat{\text{var}}(\hat{d}^{MG}) = [N(N-1)]^{-1} \sum_{i=1}^{N} (\hat{d}_i - \hat{d}^{MG})^2$. We view running line regressions as 'local ATET', where 'local' refers to a similar number of years spent in democracy, and simply adopt the standard errors from this methodology.[20]

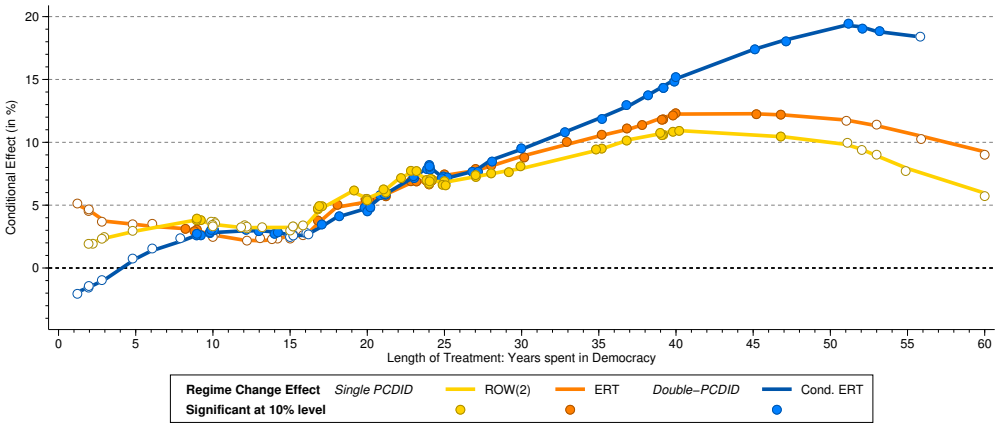
# 4 Empirical Results

**Visual Presentation of Results** We estimate the both PCDID models country by country. Thus, we obtain individual coefficients for each country (ITET) rather than a single treatment

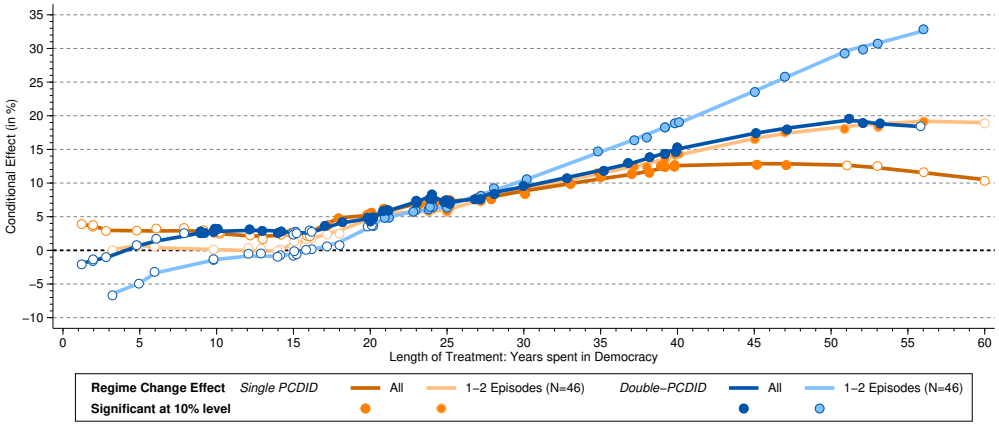
[19] Most of the existing literature on democracy and growth models democratisation as a one-off event, ignoring the empirical reality that some countries flip back and forth between regimes. Exceptions include Przeworski et al. (2000), Papaioannou & Siourounis (2008) and Eberhardt (2022).

[20] Since there may be concerns that these standard errors do not fully account for the correlation amongst the regressors we employ bootstrap methods to show that using bias-corrected confidence intervals (Appendix Figure B-2) the patterns of statistical significance are similar to those in the uncorrected results.

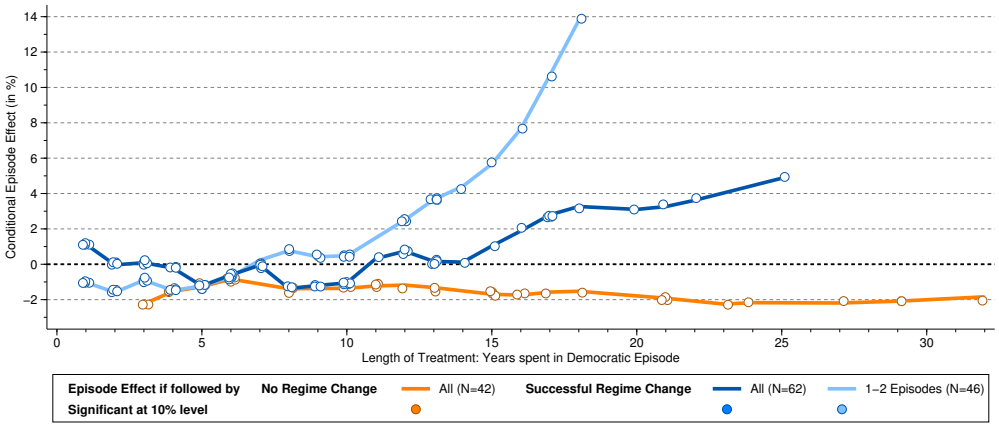
**Figure 2:** Regime Change, Episodes and Economic Growth — Single and Double PCDID




(a) Comparison of Single and Double PCDID Results




(b) Dto., with Focus on Countries with at most two Episodes




(c) Comparing Episode Effects in Countries with and without Regime Change

**Notes**: These plots present the causal effect of time spent in democracy (or, in Panel (c), in an episode) on income per capita. These are predictions from multivariate running line regressions of country-specific democracy or episodes effects ($y$-axis) on years spent in democracy or episodes ($x$-axis) and additional controls (see maintext). The sample matches that of the Double PCDID estimates for ERT (62 treated countries, unless indicated), all results are for PCDID models augmented with 4 common factors for each control group. Panel (a) presents Single PCDID alongside Double PCDID (in blue) results. Panel (b) contrasts Single and Double PCDID results for all 62 countries with those for 46 which experienced at most 2 liberalisation episodes. In Panel (c) we report the estimates for episodes in 62 countries with regime change and those in 43 without. In the former we further distinguish countries with at most 2 episodes like in Panel (b).

effect for all countries.[21] Regardless of whether we think of democratisation as a one-step or two-step process, individual countries enter our sample at different times, spend different periods of time in democracy, and may or may not experience (temporary) reversion to autocracy. These aspects matters for assessing the effect of democracy on growth and simply displaying the individual coefficients without accounting for them would be misleading. Instead, we employ predictions from running line regressions to condition on these country-specific differences and to display our findings.

Panel (a) of Figure 2 presents the results from both approaches: (i) the findings from the Single PCDID model are displayed for two democracy measures: a dummy for the ROW measure (yellow line), and a dummy for the ERT regime type dummy (orange line) and (ii) the results from the Double PCDID model (dark blue line). In all cases the democracy effect (in percent, $y$-axis) is smoothed over the years the country spent in democracy ($x$-axis) using multivariate running line regression. In the Single PCDID we control for (i) the start year of the country series, and (ii) the number of times a country moved into or out of democracy; in the Double PCDID we additionally control for (iii) the number of democratisation episodes, (iv) the years spent in episodes, and (v) the coefficient estimate on the episodes dummy, $\hat{d}_i^A$.

The interpretation of these graphs is that the years spent in democracy indicated on the $x$-axis cause the percentage increase in income per capita indicated on the $y$-axis. Filled (white) markers indicate statistical (in)significance at the 10% level.[22]

**Democratic Regime Change** The treatment effects and their relationship with time spent in democracy are very similar for the two democracy indicators using Single PCDID (orange and yellow lines): effects are moderately positive and statistically insignificant for the first 15 years, whereupon additional years spent in democracy lead to a rise in income up until a peak around 40 years of 'treatment', which is associated with 11-13% higher per capita GDP. Thereafter the effect plateaus.

[21] We present *average* effects (ATET) in Appendix Table B-1.

[22] The sample size is limited to the same 62 'treated' countries in the Double PCDID analysis.

When accounting for the episodic nature of democratisation in the Double PCDID (the dark blue line) regime change implies a more substantial long-run effect on development: in the early years these estimates are very similar to those when episodes are ignored, but from around thirty years onwards the effect continues to increase to reach around 20% higher income after 50 years in democracy.[23] Standard ATET estimates (Appendix Table B-1) fail to provide this insight offered by our running line predictions.

Our Double PCDID approach yields identical results if we exclude countries with very short episodes ($\leq 2$ years) — see Panel (a) of Appendix Figure B-1. Indeed, additional analysis in Panel (b) of the same Appendix Figure indicates that *length of time* spent in episodes is not linked to subsequent growth performance in democracy. However, the *number* of episodes experienced plays an important role: Panel (b) of Figure 2 suggests that countries with at most two episodes (light blue line) have considerably higher long-run income effects of around 30% after 50 years in democracy compared with the full sample (dark blue line).

**Successful and Failed Episodes** Panel (c) of Figure 2 compares income effects of episodes in countries which did (dark and light blue lines for full sample and countries with one or two episodes, respectively) and did not (orange line) experience regime change.[24] While none of the three line plots are statistically significant, the contrast is still indicative: not only deprived from a boost to income *in democracy*, countries which did not experience regime change also failed to benefit economically from their time in episodes. It is not an episode per se, but its successful completion that matters for growth.

Given the significance we put on successful versus failed episodes, we provide additional insights into the dominant determinants of 'episodal failure'. In Appendix E we develop an empirical Early Warning System inspired by the literature on financial crises (Eberhardt & Presbitero 2021). Across a range of specifications we find that oil booms (rather than coup

[23] Panel (a) of Appendix Figure B-2 presents 90% confidence intervals for the ERT estimates using Single and Double PCDID, which overlap.

[24] The latter is derived from regressions in the control sample of 43 countries.

attempts or natural disasters, among others) are associated with large and significant increases in the propensity for an episode to end without democratic regime change.

**Robustness** All of the above estimates are constructed from PCDID models where we include four common factors estimated from each control group. We could be concerned that this choice fails to capture all the unobserved heterogeneity. In Panel (a) of Appendix Figure B-2 we show the regime change estimate for the augmentation with four common factors (from each control group) in dark blue alongside alternative specifications with 1 to 6 common factors (dto). Augmented with only one or two factors the estimate for the democracy-growth nexus is attenuated but still reaches 20% higher per capita GDP. Including three or more common factors leads to qualitatively very similar results, as predicted by theory (Moon & Weidner 2015).

The running line predictions based on local linear regression presented above do not account for all the correlation between the underlying variables (here: estimates) and we therefore use the bootstrap to address this concern. Panel (b) of Appendix Figure B-2 suggests that patterns of statistical significance are similar to those in our main results.

All of our results above include export/trade and population growth as additional covariates, raising concerns that these may represent *outcomes* of democratic regime change. A version of the Single and Double PCDID excluding these produces identical *relative* patterns of results (see Appendix Figure B-3).

Finally, we rely in our definition of episodes on the parameters spelled out in Section 2.2. In Appendix F we present results for a wide range of alternative parameterisations (see Table note for details), which yield qualitatively very similar results.

# 5 Conclusion

Recent efforts in the analysis of the democracy-growth nexus have emphasised that there is substantial heterogeneity in the growth performance across democratizers (Cervellati & Sunde 2014, Eberhardt 2022) and that great care needs to be taken in defining democratic regime

change (Papaioannou & Siourounis 2008, Acemoglu et al. 2019). Building on this literature, our paper motivates and empirically implements democratisation as a two-stage process, made up of a liberalisation episode and regime change. This chronology enables us to provide a more nuanced analysis of the long-term growth implications of democratisation.

Our results suggest that modeling democracy as a two-stage process yields even higher economic growth in the long-run. Repeated failed episodes prior to a successful democratic transition diminish subsequent growth in democracy, but the length of episodes does not. Countries that fail to successfully complete an episode appear to derive no growth benefits, which suggests that growth dividends hinge on the successful completion of an episode, not on experiencing an episode *per se*. Avoiding episode failure is clearly important. We identify a version of the natural resource curse as the most significant culprit for episode failure.

We report standard ATET for episodes and regime change (see Appendix Table B-1), but caution that these obscure the differences between results for democracy 'over night' versus a two-stage process. Our main insights derive from running line regressions which predict *the trajectory of economic growth* over the years spent in democracy and additionally account for the idiosyncracies of individual countries' data availability, their episode and regime change dynamics, as well as the implications of the episode and regime change chronology.

Our analysis highlights the importance of episode completion and, more generally, the heterogeneity of the democratic growth dividend. Why is it then that some countries experience repeated failed episodes whereas others just need one 'attempt'? What drives the differential patterns of growth under democracy? It stands to reason that factors related to the 'deep determinants of comparative development' may play an important role in answering these questions. We can think of three ways in which the 'unequal favours' of geography (Landes 1999) may influence the magnitude of the democracy-growth effect: first, democracy fosters structural change (Acemoglu et al. 2015), yet geography (climate, crop type) can lead to differential *speeds* of structural transformation and hence development (Vollrath 2011, Eberhardt & Vollrath 2018); second, political institutions foster financial development (Rajan & Zingales 2003, Degryse et al. 2018), but 'bad' geography limits investment opportunities in countries lacking market access (Malik & Temple 2009) and/or with a narrow

range of (primary) exports; third, regime change interacts with and possibly fosters economic reforms (Giavazzi & Tabellini 2005, Prati et al. 2013). 'Bad' geography (e.g. unequal land rights, incomplete land reform; Bhattacharya et al. 2019, Albertus 2021) may correlate with elite capture in new democracies (Acemoglu & Robinson 2008, Albertus & Menaldo 2018), hampering economic reforms crucial for the realisation of the democratic dividend. We seek to investigate these factors in future research.

# References


Acemoglu, D., Naidu, S., Restrepo, P. & Robinson, J. A. (2015), Democracy, redistribution, and inequality, *in* F. B. Anthony B. Atkinson, ed., 'Handbook of income distribution', Vol. 2, Elsevier, pp. 1885–1966.

Acemoglu, D., Naidu, S., Restrepo, P. & Robinson, J. A. (2019), 'Democracy does cause growth', *Journal of Political Economy* **127**(1), 47–100.

Acemoglu, D. & Robinson, J. A. (2006), *Economic origins of dictatorship and democracy*, Cambridge University Press.

Acemoglu, D. & Robinson, J. A. (2008), 'Persistence of power, elites, and institutions', *American Economic Review* **98**(1), 267–93.

Aghion, P., Alesina, A. & Trebbi, F. (2007), Democracy, technology, and growth, Working Paper 13180, National Bureau of Economic Research.

Albertus, M. (2021), *Property without rights: Origins and consequences of the property rights gap*, Cambridge University Press.

Albertus, M. & Menaldo, V. (2018), *Authoritarianism and the elite origins of democracy*, Cambridge University Press.

Alesina, A. & Drazen, A. (1991), 'Why are stabilizations delayed?', *American Economic Review* **81**, 1170–1188.

Andrews, D. W. (2005), 'Cross-section regression with common shocks', *Econometrica* **73**(5), 1551–1585.

Arellano, C., Bai, Y. & Lizarazo, S. V. (2017), Sovereign risk contagion, NBER Working Paper Series 24031, National Bureau of Economic Research.

Athey, S., Bayati, M., Doudchenko, N., Imbens, G. & Khosravi, K. (2021), 'Matrix completion methods for causal panel data models', *Journal of the American Statistical Association* **116**(536), 1716–1730.

Bai, J. (2009), 'Panel data models with interactive fixed effects', *Econometrica* **77**(4), 1229–1279.

Bhattacharya, P. S., Mitra, D. & Ulubaşoğlu, M. A. (2019), 'The political economy of land reform enactments: New cross-national evidence (1900–2010)', *Journal of Development Economics* **139**, 50–68.

Boese, V. A. (2019), 'How (not) to measure democracy', *International Area Studies Review* **22**(2), 95–127.

Bolt, J. & van Zanden, J. L. (2020), Maddison Style Estimates of the Evolution of the World Economy: A New 2020 Update, Maddison Project Working Paper 15, University of Groningen, Groningen Growth and Development Centre.

Brownlee, J. (2007), *Authoritarianism in an Age of Democratization*, Cambridge University Press Cambridge.

Brownlee, J. (2009), 'Portents of pluralism: How hybrid regimes affect democratic transitions', *American Journal of Political Science* **53**(3), 515–532.

Cervellati, M., Jung, F., Sunde, U. & Vischer, T. (2014), 'Income and democracy: Comment', *American Economic Review* **104**(2), 707–19.

Cervellati, M. & Sunde, U. (2014), 'Civil conflict, democratization, and growth: Violent democratization as critical juncture', *Scandinavian Journal of Economics* **116**(2), 482–505.

Cesa-Bianchi, A., Martin, F. E. & Thwaites, G. (2019), 'Foreign booms, domestic busts: The global dimension of banking crises', *Journal of Financial Intermediation* **37**, 58–74.

Chamberlain, G. (1982), 'Multivariate regression models for panel data', *Journal of Econometrics* **18**(1), 5–46.

Chan, M. K. & Kwok, S. (2022), 'The PCDID Approach: Difference-in-Differences when Trends are Potentially Unparallel and Stochastic', *Journal of Business and Economic Statistics* **40**(3), 1216–1233.

Chudik, A. & Pesaran, M. H. (2013), Large Panel Data Models with Cross-Sectional Dependence: A Survey, CESifo Working Paper Series 4371.

Cotet, A. M. & Tsui, K. K. (2013), 'Oil and conflict: What does the cross country evidence really show?', *American Economic Journal: Macroeconomics* **5**(1), 49–80.

Degryse, H., Lambert, T. & Schwienbacher, A. (2018), 'The political economy of financial systems: Evidence from suffrage reforms in the last two centuries', *The Economic Journal* **128**(611), 1433–1475.

Dewatripont, M. & Roland, G. (1995), 'The design of reform packages under uncertainty', *American Economic Review* **85**(5), 1207–1223.

Diamond, L. (2002), 'Elections without democracy: Thinking about hybrid regimes', *Journal of Democracy* **13**(2), 21–35.

Eberhardt, M. (2022), 'Democracy, growth, heterogeneity, and robustness', *European Economic Review* **147**(104173).

Eberhardt, M. & Presbitero, A. F. (2021), 'Commodity prices and banking crises', *Journal of International Economics* **131**, 103474.

Eberhardt, M. & Teal, F. (2011), 'Econometrics for grumblers: a new look at the literature on cross-country growth empirics', *Journal of Economic Surveys* **25**(1), 109–155.

Eberhardt, M. & Vollrath, D. (2018), 'The effect of agricultural technology on the speed of development', *World Development* **109**, 483–496.

Edgell, A., Maerz, S., Maxwell, L., Morgan, R., Medzihorsky, J., Wilson, M., Boese, V. A., Hellmeier, S., Lachapelle, J., Lindenfors, P., Lührmann, A. & Lindberg, S. I. (2020), 'Episodes of regime transformation dataset, version 1.0', *Available at https://github.com/vdeminstitute/ERT* .

Epstein, D. L., Bates, R., Goldstone, J., Kristensen, I. & O'Halloran, S. (2006), 'Democratic transitions', *American Journal of Political Science* **50**(3), 551–569.

Fernandez, R. & Rodrik, D. (1991), 'Resistance to reform: Status quo bias in the presence of individual-specific uncertainty', *American Economic Review* **81**(5), 1146–1155.

Fouquin, M. & Hugot, J. (2016), Two centuries of bilateral trade and gravity data: 1827-2014, CEPII Working Paper 2016-14.

Frantz, E. & Kendall-Taylor, A. (2014), 'A dictator's toolkit: Understanding how co-optation affects repression in autocracies', *Journal of Peace Research* **51**(3), 332–346.

Geddes, B. (1999), 'What do we know about democratization after twenty years?', *Annual Review of Political Science* **2**(1), 115–144.

Geddes, B., Wright, J. & Frantz, E. (2014), 'Autocratic breakdown and regime transitions: A new data set', *Perspectives on Politics* **12**(2), 313–331.

Gehlbach, S., Sonin, K. & Svolik, M. W. (2016), 'Formal models of nondemocratic politics', *Annual Review of Political Science* **19**, 565–584.

Gerring, J., Bond, P., Barndt, W. T. & Moreno, C. (2005), 'Democracy and economic growth: A historical perspective', *World Politics* **57**(3), 323–364.

Giavazzi, F. & Tabellini, G. (2005), 'Economic and political liberalizations', *Journal of Monetary Economics* **52**(7), 1297–1330.

Gobillon, L. & Magnac, T. (2016), 'Regional policy evaluation: Interactive fixed effects and synthetic controls', *Review of Economics and Statistics* **98**(3), 535–551.

Goodman-Bacon, A. (2021), 'Difference-in-differences with variation in treatment timing', *Journal of Econometrics* **225**(2), 254–277.

Graham, H. & Quiroga, A. (2012), After the fear was over? what came after dictatorships in spain, greece, and portugal, *in* 'The Oxford Handbook of Postwar European History'.

Gunitsky, S. (2014), 'From shocks to waves: Hegemonic transitions and democratization in the twentieth century', *International Organization* **68**(3), 561–597.

Huntington, S. P. (1991), 'Democracy's third wave', *Journal of Democracy* **2**(2), 12–34.

Huntington, S. P. (1993), *The third wave: Democratization in the late twentieth century*, Vol. 4, University of Oklahoma Press, Norman, USA.

Knutsen, C. H. (2012), 'Democracy and economic growth: A survey of arguments and results', *International Area Studies Review* **15**(4), 393–415.

Knutsen, C. H. (2013), 'Democracy, state capacity, and economic growth', *World Development* **43**, 1–18.

Knutsen, C. H. (2021), 'A business case for democracy: regime type, growth, and growth volatility', *Democratization* **28**(8), 1505–1524.

Laeven, L. & Valencia, F. (2020), 'Systemic banking crises database II', *IMF Economic Review* **68**(2), 307–361.

Landes, D. S. (1999), *The Wealth and Poverty of Nations*, London: Abacus.

Levitsky, S. & Way, L. A. (2010), *Competitive authoritarianism: Hybrid regimes after the Cold War*, Cambridge University Press.

Lührmann, A., Tannenberg, M. & Lindberg, S. I. (2018), 'Regimes of the World (RoW): Opening New Avenues for the Comparative Study of Political Regimes', *Politics & Governance* **6**(1), 60–77.

Madsen, J. B., Raschky, P. A. & Skali, A. (2015), 'Does democracy drive income in the world, 1500–2000?', *European Economic Review* **78**, 175–195.

Maerz, S. F., Edgell, A., Wilson, M. C., Hellmeier, S. & Lindberg, S. I. (2021), 'A framework for understanding regime transformation: introducing the ert dataset', *V-Dem Working Paper* **113**.

Magaloni, B. (2008), 'Credible power-sharing and the longevity of authoritarian rule', *Comparative Political Studies* **41**(4-5), 715–741.

Malik, A. & Temple, J. R. (2009), 'The geography of output volatility', *Journal of Development Economics* **90**(2), 163–178.

Moon, H. R. & Weidner, M. (2015), 'Linear regression for panel with unknown number of factors as interactive fixed effects', *Econometrica* **83**(4), 1543–1579.

Mundlak, Y. (1978), 'On the pooling of time series and cross section data', *Econometrica* **46**(1), 69–85.

Murtin, F. & Wacziarg, R. (2014), 'The democratic transition', *Journal of Economic Growth* **19**(2), 141–181.

Olley, G. S. & Pakes, A. (1996), 'The dynamics of productivity in the telecommunications equipment industry', *Econometrica* **64**(6), 1263–1297.

Papaioannou, E. & Siourounis, G. (2008), 'Democratisation and growth', *Economic Journal* **118**(532), 1520–1551.

Persson, T. & Tabellini, G. (2006), 'Democracy and development: The devil in the details', *American Economic Review, Papers & Proceedings* **96**(2), 319–324.

Persson, T. & Tabellini, G. (2009), 'Democratic capital: The nexus of political and economic change', *American Economic Journal: Macroeconomics* **1**(2), 88–126.

Pesaran, M. H. (2006), 'Estimation and inference in large heterogeneous panels with a multifactor error structure', *Econometrica* **74**(4), 967–1012.

Pesaran, M. H. & Smith, R. (1995), 'Estimating long-run relationships from dynamic heterogeneous panels', *Journal of Econometrics* **68**(1), 79–113.

Phillips, P. C. & Sul, D. (2003), 'Dynamic panel estimation and homogeneity testing under cross section dependence', *Econometrics Journal* **6**(1), 217–259.

Prati, A., Onorato, M. G. & Papageorgiou, C. (2013), 'Which reforms work and under what institutional environment? evidence from a new data set on structural reforms', *Review of Economics and Statistics* **95**(3), 946–968.

Przeworski, A., Alvarez, M., Cheibub, J. A. & Limongi, F. (2000), *Democracy and development: political institutions and well-being in the world, 1950-1990*, Vol. 3, Cambridge University Press, Cambridge, UK.

Rajan, R. G. & Zingales, L. (2003), 'The great reversals: the politics of financial development in the twentieth century', *Journal of Financial Economics* **69**(1), 5–50.

Reinhart, C. M. & Rogoff, K. S. (2009), *This time is different*, Princeton University Press.

Rodrik, D. & Wacziarg, R. (2005), 'Do democratic transitions produce bad economic outcomes?', *American Economic Review, Papers & Proceedings* **95**(2), 50–55.

Rustow, D. A. (1970), 'Transitions to democracy: Toward a dynamic model', *Comparative Politics* **2**(3), 337–363.

Teorell, J., Coppedge, M., Lindberg, S. & Skaaning, S.-E. (2019), 'Measuring polyarchy across the globe, 1900–2017', *Studies in Comparative International Development* **54**(1), 71–95.

Treisman, D. (2020), 'Democracy by mistake: How the errors of autocrats trigger transitions to freer government', *American Political Science Review* **114**(3), 792–810.

Vollrath, D. (2011), 'The agricultural basis of comparative development', *Journal of Economic Growth* **16**(4), 343–370.

Wilson, M. C., Medzihorsky, J., Maerz, S. F., Lührmann, A., Lindenfors, P., Edgell, A. B., Boese, V. A. & Lindberg, S. I. (2022), 'Episodes of liberalization in autocracies: A new approach to quantitatively studying democratization', *Political Science Research and Methods* (forthcoming).

Xu, Y. (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 Appendix

### A.1 Sources and Sample Makeup

Our choice of data sources (Maddison, TRADHIST) enables analysis from 1950-2014, but excludes a number of countries which are available in ERT from inclusion in the treatment or control groups: ten small treated economies (Bhutan, Fiji, Guyana, Kosovo, Maldives, PNG, Solomon Islands, Suriname, Timor-Leste, Vanuatu); five small (historical) economies with failed episodes: Zanzibar, Somaliland, Somalia, Republic of (South) Vietnam, GDR; and three autocratic economies with no episodes: South Yemen, Gaza/Palestine, Eritrea.

Our 1950-2014 sample covers 62 'treated' countries which experienced episodes and regime change (n=3,724 observations — see for Table A-3 sample makeup), 43 autocratic countries which only experienced democratisation episodes (n=2,515; control group 2 — Table A-2), and 15 autocratic countries which never experienced episodes (n=646; control group 1 — Table A-1). Four democracies reverted to autocracy and subsequently had unsuccessful democratisation episodes (n=75 observations); 9 countries had episodes and regime change but no pre-episode data (n=399) — both sets of observations are excluded from the analysis. The balance to arrive at 161 countries in the full available sample (n=8,770) is made up by 28 countries which were democracies throughout the sample period, which are also excluded. In practice the minimum number of time series observations for inclusion in our analysis is n=21. This is in line with the practice in Giavazzi & Tabellini (2005), Persson & Tabellini (2006) and Papaioannou & Siourounis (2008).

Figure A-1 provides an overview of the distribution of episodes and regime changes in our sample. In the top panel the histogram in light blue highlights two peaks of democratisation episodes in the late 1950s/early 1960s, and in the 1990s, coinciding with the second and third waves of democratisation (Huntington 1993). The lowest rate of ongoing democratisation episodes is in the mid-1960s and 1970s. The regime change events, in dark pink, clearly

match these patterns for the second peak in the 1990s, but less so for the earlier period. The middle panel supports this notion of differential rates of episodes and their outcomes over time: the share of failed episodes (in teal) is particularly strong in the 1950s and early 1960s, and again in the 1990s. Episodes culminating in regime change (in dark pink) are only substantial in the late 1970s to early 1990s and are otherwise dominated by the former group.

The bottom panel in Figure A-1 charts the mean episode length over time and the evolution of each episode in our sample. It shows substantial variation in episode length over time as well as temporal clusters of episodes with and without regime change. The graphs for successful episodes are frequently very steep (implying short episodes), yet it would be misleading to claim that these trajectories *dominate* the treatment sample.

**Table A-1:** Sample Makeup: Control Group 1 (never experienced a democratisation episode)

| Country | ISO | Total obs | Country | ISO | Total obs |
|---|---|---|---|---|---|
| United Arab Emirates | ARE | 21 | North Korea | PRK | 35 |
| Azerbaijan | AZE | 21 | Qatar | QAT | 40 |
| China | CHN | 64 | Saudi Arabia | SAU | 64 |
| Cuba | CUB | 65 | Tajikistan | TJK | 21 |
| Djibouti | DJI | 64 | Turkmenistan | TKM | 21 |
| Iran | IRN | 64 | Uzbekistan | UZB | 21 |
| Kazakhstan | KAZ | 21 | Viet Nam | VNM | 60 |
| Mozambique | MOZ | 64 | | | |

*Notes*: This table provides details on the sample-makeup of the first control group sample, made up of the 15 countries which never experienced a democratisation episode (and of course also no regime change).

**Table A-2:** Sample Makeup: Control Group 2 (never democratised)

| | | | Episodes (all failed) | | | | | | | | | Autocracy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | ISO | Total obs | Years in ep | Share | Avg length | Count | 1st | 2nd | 3rd | 4th | 5th | Years in auto | Share |
| Afghanistan | AFG | 59 | 5 | 8% | 5.0 | 1 | 2002 | | | | | 54 | 92% |
| Angola | AGO | 39 | 4 | 10% | 4.0 | 1 | 2008 | | | | | 35 | 90% |
| Burundi | BDI | 55 | 17 | 31% | 5.7 | 3 | 1982 | 1992 | 1999 | | | 38 | 69% |
| Bahrain | BHR | 44 | 6 | 14% | 3.0 | 2 | 1972 | 2000 | | | | 38 | 86% |
| Central African Republic | CAF | 64 | 21 | 33% | 5.3 | 4 | 1956 | 1987 | 2005 | 2014 | | 43 | 67% |
| Cameroon | CMR | 52 | 4 | 8% | 4.0 | 1 | 1990 | | | | | 48 | 92% |
| DR of Congo | COD | 64 | 18 | 28% | 9.0 | 2 | 1955 | 1998 | | | | 46 | 72% |
| Congo | COG | 64 | 11 | 17% | 3.7 | 3 | 1957 | 1990 | 2002 | | | 53 | 83% |
| Algeria | DZA | 44 | 6 | 14% | 2.0 | 3 | 1977 | 1990 | 1995 | | | 38 | 86% |
| Egypt | EGY | 64 | 10 | 16% | 10.0 | 1 | 1956 | | | | | 54 | 84% |
| Ethiopia | ETH | 64 | 6 | 9% | 6.0 | 1 | 1987 | | | | | 58 | 91% |
| Gabon | GAB | 64 | 13 | 20% | 6.5 | 2 | 1957 | 1987 | | | | 51 | 80% |
| Guinea | GIN | 64 | 24 | 38% | 8.0 | 3 | 1957 | 1985 | 2010 | | | 40 | 63% |
| Gambia | GMB | 64 | 13 | 20% | 3.3 | 4 | 1960 | 1966 | 1996 | 2014 | | 51 | 80% |
| Guinea-Bissau | GNB | 64 | 21 | 33% | 5.3 | 4 | 1973 | 1990 | 2005 | 2014 | | 43 | 67% |
| Equatorial Guinea | GNQ | 55 | 15 | 27% | 7.5 | 2 | 1968 | 1982 | | | | 40 | 73% |
| China, Hong Kong | HKG | 64 | 8 | 13% | 8.0 | 1 | 1985 | | | | | 56 | 88% |
| Haiti | HTI | 65 | 12 | 18% | 2.4 | 5 | 1951 | 1987 | 1991 | 1993 | 2006 | 53 | 82% |
| Iraq | IRQ | 64 | 8 | 13% | 8.0 | 1 | 2004 | | | | | 56 | 88% |
| Jordan | JOR | 64 | 6 | 9% | 6.0 | 1 | 1989 | | | | | 58 | 91% |
| Kenya | KEN | 64 | 29 | 45% | 9.7 | 3 | 1956 | 1990 | 2010 | | | 35 | 55% |
| Kyrgyzstan | KGZ | 23 | 11 | 48% | 11.0 | 1 | 2003 | | | | | 12 | 52% |
| Cambodia | KHM | 60 | 11 | 18% | 11.0 | 1 | 1990 | | | | | 49 | 82% |
| Kuwait | KWT | 40 | 16 | 40% | 8.0 | 2 | 1981 | 1991 | | | | 24 | 60% |
| Lao PDR | LAO | 60 | 4 | 7% | 4.0 | 1 | 1955 | | | | | 56 | 93% |
| Lebanon | LBN | 64 | 15 | 23% | 15.0 | 1 | 1996 | | | | | 49 | 77% |
| Libya | LBY | 62 | 3 | 5% | 3.0 | 1 | 2011 | | | | | 59 | 95% |
| Morocco | MAR | 64 | 15 | 23% | 7.5 | 2 | 1963 | 1993 | | | | 49 | 77% |
| Myanmar | MMR | 64 | 8 | 13% | 8.0 | 1 | 2010 | | | | | 56 | 88% |
| Mauritania | MRT | 55 | 10 | 18% | 3.3 | 3 | 1987 | 2007 | 2010 | | | 45 | 82% |
| Malaysia | MYS | 65 | 27 | 42% | 13.5 | 2 | 1972 | 1999 | | | | 38 | 58% |
| Oman | OMN | 57 | 4 | 7% | 4.0 | 1 | 2000 | | | | | 53 | 93% |
| Pakistan | PAK | 64 | 32 | 50% | 10.7 | 3 | 1962 | 1985 | 2002 | | | 32 | 50% |
| Rwanda | RWA | 55 | 21 | 38% | 7.0 | 3 | 1979 | 1991 | 2003 | | | 34 | 62% |
| Sudan | SDN | 64 | 23 | 36% | 7.7 | 3 | 1965 | 1986 | 1996 | | | 41 | 64% |
| Singapore | SGP | 55 | 1 | 2% | 1.0 | 1 | 1960 | | | | | 54 | 98% |
| Swaziland | SWZ | 55 | 6 | 11% | 6.0 | 1 | 1964 | | | | | 49 | 89% |
| Seychelles | SYC | 55 | 29 | 53% | 9.7 | 3 | 1963 | 1979 | 1991 | | | 26 | 47% |
| Syrian Arab Rep. | SYR | 64 | 5 | 8% | 2.5 | 2 | 1953 | 1961 | | | | 59 | 92% |
| Chad | TCD | 64 | 8 | 13% | 8.0 | 1 | 1990 | | | | | 56 | 88% |
| Uganda | UGA | 64 | 16 | 25% | 5.3 | 3 | 1953 | 1981 | 1989 | | | 48 | 75% |
| Yemen | YEM | 52 | 6 | 12% | 6.0 | 1 | 1988 | | | | | 46 | 88% |
| Zimbabwe | ZWE | 64 | 3 | 5% | 3.0 | 1 | 1979 | | | | | 61 | 95% |

*Notes*: This table provides details on the sample-makeup of the second control group sample, made up of the 43 countries which experienced at least one democratisation episode but never realised democratic regime change.

**Table A-3:** Sample Makeup: Treated Countries

| | | | Episodes (successful or failed) | | | | | | | | | | | Regime change to democracy | | | | | | | Autocracy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | ISO | Total obs | Years in ep | Share | Avg length | Count | 1st | 2nd | 3rd | 4th | 5th | Count Failed | Avg length Failed | Years in dem | Share | Count | 1st | Ep Length | 2nd | Ep Length | Years in auto | Share |
| **Albania** | ALB | 60 | 10 | 17% | 3.3 | 3 | 1991 | 1998 | 2005 | | | 2 | 5.0 | 10 | 17% | 1 | 2005 | 0 | | | 40 | 67% |
| **Argentina** | ARG | 65 | 10 | 15% | 2.5 | 4 | 1957 | 1963 | 1972 | 1983 | | 2 | 4.0 | 33 | 51% | 2 | 1964 | 1 | 1984 | 1 | 22 | 34% |
| Armenia | ARM† | 21 | 8 | 38% | 4.0 | 2 | 1998 | 2010 | | | | 2 | 4.0 | 2 | 10% | 0 | | | | | 11 | 52% |
| **Benin** | BEN | 64 | 15 | 23% | 7.5 | 2 | 1952 | 1990 | | | | 1 | 13.0 | 23 | 36% | 1 | 1992 | 2 | | | 26 | 41% |
| **Burkina Faso** | BFA | 55 | 14 | 25% | 4.7 | 3 | 1960 | 1978 | 1990 | | | 2 | 2.0 | 15 | 27% | 1 | 2000 | 10 | | | 26 | 47% |
| **Bangladesh** | BGD | 42 | 17 | 40% | 4.3 | 4 | 1973 | 1977 | 1984 | 2009 | | 3 | 3.0 | 10 | 24% | 1 | 1992 | 8 | | | 15 | 36% |
| **Bulgaria** | BGR | 64 | 1 | 2% | 1.0 | 1 | 1990 | | | | | 0 | | 24 | 38% | 1 | 1991 | 1 | | | 39 | 61% |
| Bosnia & Herzeg. | BIH[a] | 19 | 1 | 5% | 1.0 | 1 | 1996 | | | | | 0 | | 18 | 95% | 1 | 1997 | 1 | | | 0 | 0% |
| Belarus | BLR† | 21 | 0 | 0% | | 0 | | | | | | 0 | | 3 | 14% | 0 | | | | | 18 | 86% |
| **Bolivia** | BOL | 65 | 13 | 20% | 6.5 | 2 | 1952 | 1983 | | | | 1 | 11.0 | 30 | 46% | 1 | 1985 | 2 | | | 22 | 34% |
| **Brazil** | BRA | 65 | 12 | 18% | 12.0 | 1 | 1975 | | | | | 0 | | 28 | 43% | 1 | 1987 | 12 | | | 25 | 38% |
| Barbados | BRB[b] | 64 | 9 | 14% | 9.0 | 1 | 1951 | | | | | 0 | | 55 | 86% | 1 | 1960 | 9 | | | 0 | 0% |
| Botswana | BWA[c] | 55 | 7 | 13% | 7.0 | 1 | 1960 | | | | | 0 | | 48 | 87% | 1 | 1967 | 7 | | | 0 | 0% |
| **Chile** | CHL | 65 | 3 | 5% | 1.5 | 2 | 1958 | 1988 | | | | 0 | | 39 | 60% | 2 | 1959 | 1 | 1990 | 2 | 23 | 35% |
| **Côte d'Ivoire** | CIV | 64 | 17 | 27% | 4.3 | 4 | 1990 | 1995 | 2001 | | | 3 | 3.3 | 2 | 3% | 1 | 2008 | 7 | | | 45 | 70% |
| **Colombia** | COL | 65 | 25 | 38% | 8.3 | 3 | 1958 | 1972 | 1982 | | | 2 | 8.0 | 24 | 37% | 1 | 1991 | 9 | | | 16 | 25% |
| **Comoros** | COM | 55 | 7 | 13% | 2.3 | 3 | 1990 | 1997 | 2002 | | | 2 | 1.5 | 9 | 16% | 1 | 2006 | 4 | | | 39 | 71% |
| **Cabo Verde** | CPV | 57 | 7 | 12% | 2.3 | 3 | 1972 | 1980 | 1990 | | | 2 | 3.0 | 24 | 42% | 1 | 1991 | 1 | | | 26 | 46% |
| **Cyprus** | CYP | 64 | 2 | 3% | 2.0 | 1 | 1960 | | | | | 0 | | 45 | 70% | 1 | 1960 | 0 | | | 17 | 27% |
| **Dominican Rep.** | DOM | 64 | 13 | 20% | 3.3 | 4 | 1961 | 1966 | 1978 | 1995 | | 2 | 4.0 | 27 | 42% | 2 | 1982 | 4 | 1996 | 1 | 24 | 38% |
| **Ecuador** | ECU | 65 | 9 | 14% | 3.0 | 3 | 1950 | 1967 | 1978 | | | 2 | 3.5 | 35 | 54% | 1 | 1980 | 2 | | | 21 | 32% |
| **Spain** | ESP | 65 | 2 | 3% | 2.0 | 1 | 1976 | | | | | 0 | | 37 | 57% | 1 | 1978 | 2 | | | 26 | 40% |
| Georgia | GEO[b] | 21 | 10 | 48% | 10.0 | 1 | 1994 | | | | | 0 | | 11 | 52% | 1 | 2004 | 10 | | | 0 | 0% |
| **Ghana** | GHA | 64 | 7 | 11% | 1.8 | 4 | 1951 | 1969 | 1979 | 1993 | | 3 | 2.0 | 21 | 33% | 1 | 1994 | 1 | | | 36 | 56% |
| **Greece** | GRC | 65 | 5 | 8% | 2.5 | 2 | 1950 | 1974 | | | | 1 | 4.0 | 40 | 62% | 1 | 1975 | 1 | | | 20 | 31% |
| **Guatemala** | GTM | 65 | 16 | 25% | 16.0 | 1 | 1984 | | | | | 0 | | 15 | 23% | 1 | 2000 | 16 | | | 34 | 52% |
| **Honduras** | HND | 65 | 13 | 20% | 4.3 | 3 | 1950 | 1971 | 1980 | | | 2 | 1.5 | 18 | 28% | 1 | 1990 | 10 | | | 34 | 52% |
| Croatia | HRV[d] | 22 | 7 | 32% | 7.0 | 1 | 1993 | | | | | 0 | | 15 | 68% | 1 | 2000 | 7 | | | 0 | 0% |
| **Hungary** | HUN | 65 | 2 | 3% | 2.0 | 1 | 1988 | | | | | 0 | | 25 | 38% | 1 | 1990 | 2 | | | 38 | 58% |
| **Indonesia** | IDN | 65 | 10 | 15% | 5.0 | 2 | 1950 | 1997 | | | | 1 | 7.0 | 15 | 23% | 1 | 2000 | 3 | | | 40 | 62% |

Continued overleaf

**Table A-3:** Sample Makeup: Treated Countries (continued)

| | | | Episodes (successful or failed) | | | | | | | | | | | Regime change to democracy | | | | | | | Autocracy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | ISO | Total obs | Years in ep | Share | Avg length | Count | 1st | 2nd | 3rd | 4th | 5th | Count Failed | Avg length Failed | Years in dem | Share | Count | 1st | Ep Length | 2nd | Ep Length | Years in auto | Share |
| India | IND[e] | 65 | 2 | 3% | 2.0 | 1 | 1950 | | | | | 0 | | 63 | 97% | 1 | 1952 | 2 | | | 0 | 0% |
| **Jamaica** | JAM | 64 | 2 | 3% | 2.0 | 1 | 1953 | | | | | 0 | | 56 | 88% | 1 | 1955 | 2 | | | 6 | 9% |
| Japan | JPN[e] | 65 | 2 | 3% | 2.0 | 1 | 1950 | | | | | 0 | | 63 | 97% | 1 | 1952 | 2 | | | 0 | 0% |
| **South Korea** | KOR | 64 | 13 | 20% | 6.5 | 2 | 1964 | 1976 | | | | 1 | 1.0 | 27 | 42% | 1 | 1988 | 12 | | | 24 | 38% |
| **Liberia** | LBR | 64 | 7 | 11% | 2.3 | 3 | 1985 | 1997 | 2005 | | | 2 | 3.0 | 9 | 14% | 1 | 2006 | 1 | | | 48 | 75% |
| **Sri Lanka** | LKA | 65 | 8 | 12% | 4.0 | 2 | 1983 | 2011 | | | | 1 | 4.0 | 51 | 78% | 1 | 1987 | 4 | | | 6 | 9% |
| **Lesotho** | LSO | 55 | 12 | 22% | 4.0 | 3 | 1960 | 1992 | 2002 | | | 2 | 5.5 | 12 | 22% | 1 | 2003 | 1 | | | 31 | 56% |
| **Moldova** | MDA | 23 | 4 | 17% | 4.0 | 1 | 2006 | | | | | 0 | | 16 | 70% | 1 | 2010 | 4 | | | 3 | 13% |
| **Madagascar** | MDG | 64 | 20 | 31% | 5.0 | 4 | 1956 | 1985 | 2003 | 2013 | | 2 | 4.0 | 10 | 16% | 2 | 1994 | 9 | 2006 | 3 | 34 | 53% |
| **Mexico** | MEX | 65 | 18 | 28% | 18.0 | 1 | 1977 | | | | | 0 | | 20 | 31% | 1 | 1995 | 18 | | | 27 | 42% |
| **Macedonia** | MKD | 22 | 6 | 27% | 6.0 | 1 | 1993 | | | | | 0 | | 14 | 64% | 1 | 1999 | 6 | | | 2 | 9% |
| **Mali** | MLI | 55 | 3 | 5% | 1.5 | 2 | 1960 | 1992 | | | | 1 | 2.0 | 20 | 36% | 1 | 1993 | 1 | | | 32 | 58% |
| **Malta** | MLT | 57 | 1 | 2% | 1.0 | 1 | 1962 | | | | | 0 | | 52 | 91% | 1 | 1963 | 1 | | | 4 | 7% |
| Montenegro | MNE† | 10 | 0 | 0% | | 0 | | | | | | 0 | | 3 | 30% | 0 | | | | | 7 | 70% |
| Mongolia | MNG[a] | 24 | 1 | 4% | 1.0 | 1 | 1990 | | | | | 0 | | 23 | 96% | 1 | 1992 | 2 | | | 0 | 0% |
| **Mauritius** | MUS | 64 | 3 | 5% | 1.5 | 2 | 1959 | 1968 | | | | 1 | 3.0 | 47 | 73% | 1 | 1968 | 0 | | | 14 | 22% |
| **Malawi** | MWI | 58 | 13 | 22% | 6.5 | 2 | 1992 | 2005 | | | | 1 | 9.0 | 6 | 10% | 1 | 2009 | 4 | | | 39 | 67% |
| **Namibia** | NAM | 35 | 3 | 9% | 1.5 | 2 | 1989 | 1995 | | | | 1 | 3.0 | 20 | 57% | 1 | 1995 | 0 | | | 12 | 34% |
| **Niger** | NER | 64 | 12 | 19% | 3.0 | 4 | 1957 | 1988 | 1993 | 1997 | | 3 | 3.7 | 15 | 23% | 1 | 1994 | 1 | | | 37 | 58% |
| **Nigeria** | NGA | 64 | 11 | 17% | 3.7 | 3 | 1976 | 1998 | 2010 | | | 2 | 4.0 | 2 | 3% | 1 | 2013 | 3 | | | 51 | 80% |
| **Nicaragua** | NIC | 65 | 10 | 15% | 10.0 | 1 | 1980 | | | | | 0 | | 17 | 26% | 1 | 1990 | 10 | | | 38 | 58% |
| **Nepal** | NPL | 52 | 5 | 10% | 2.5 | 2 | 1990 | 2006 | | | | 1 | 3.0 | 5 | 10% | 1 | 2008 | 2 | | | 42 | 81% |
| **Panama** | PAN | 65 | 6 | 9% | 2.0 | 3 | 1950 | 1953 | 1990 | | | 2 | 2.5 | 24 | 37% | 1 | 1991 | 1 | | | 35 | 54% |
| **Peru** | PER | 65 | 17 | 26% | 4.3 | 4 | 1950 | 1964 | 1976 | 1994 | | 3 | 4.0 | 25 | 38% | 1 | 1981 | 5 | | | 23 | 35% |
| **Philippines** | PHL | 65 | 9 | 14% | 4.5 | 2 | 1982 | 2007 | | | | 0 | | 21 | 32% | 2 | 1988 | 6 | 2010 | 3 | 35 | 54% |
| **Poland** | POL | 64 | 10 | 16% | 10.0 | 1 | 1980 | | | | | 0 | | 25 | 39% | 1 | 1990 | 10 | | | 29 | 45% |
| **Portugal** | PRT | 65 | 6 | 9% | 6.0 | 1 | 1970 | | | | | 0 | | 39 | 60% | 1 | 1976 | 6 | | | 20 | 31% |
| **Paraguay** | PRY | 65 | 4 | 6% | 4.0 | 1 | 1990 | | | | | 0 | | 21 | 32% | 1 | 1994 | 4 | | | 40 | 62% |
| **Romania** | ROU | 60 | 1 | 2% | 1.0 | 1 | 1990 | | | | | 0 | | 24 | 40% | 1 | 1991 | 1 | | | 35 | 58% |
| Russia | RUS† | 23 | 0 | 0% | | 0 | | | | | | 0 | | 2 | 9% | 0 | | | | | 21 | 91% |

Continued overleaf

**Table A-3:** Sample Makeup: Treated Countries (continued)

| | | | Episodes (successful or failed) | | | | | | | | | | | Regime change to democracy | | | | | | | Autocracy | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | ISO | Total obs | Years in ep | Share | Avg length | Count | 1st | 2nd | 3rd | 4th | 5th | Count Failed | Avg length Failed | Years in dem | Share | Count | 1st | Ep Length | 2nd | Ep Length | Years in auto | Share |
| **Senegal** | SEN | 64 | 7 | 11% | 2.3 | 3 | 1960 | 1978 | 1990 | | | 1 | 7.0 | 25 | 39% | 2 | 1960 | 0 | 1990 | 0 | 32 | 50% |
| **Sierra Leone** | SLE | 64 | 13 | 20% | 3.3 | 4 | 1958 | 1994 | 2002 | 2013 | | 3 | 4.0 | 9 | 14% | 1 | 2003 | 1 | | | 42 | 66% |
| **El Salvador** | SLV | 65 | 12 | 18% | 6.0 | 2 | 1982 | 1991 | | | | 1 | 4.0 | 16 | 25% | 1 | 1999 | 8 | | | 37 | 57% |
| **São Tomé & Principe** | STP | 55 | 9 | 16% | 4.5 | 2 | 1972 | 1987 | | | | 1 | 4.0 | 23 | 42% | 1 | 1992 | 5 | | | 23 | 42% |
| **Togo** | TGO | 64 | 21 | 33% | 5.3 | 4 | 1956 | 1991 | 2005 | 2012 | | 3 | 6.3 | 1 | 2% | 1 | 2014 | 2 | | | 42 | 66% |
| **Thailand** | THA | 64 | 22 | 34% | 4.4 | 5 | 1974 | 1978 | 1992 | 2008 | 2010 | 4 | 4.0 | 8 | 13% | 1 | 1998 | 6 | | | 34 | 53% |
| Trinidad & Tobago | TTO[f] | 64 | 9 | 14% | 9.0 | 1 | 1951 | | | | | 0 | | 55 | 86% | 1 | 1960 | 9 | | | 0 | 0% |
| **Tunisia** | TUN | 64 | 6 | 9% | 3.0 | 2 | 1956 | 2011 | | | | 1 | 5.0 | 3 | 5% | 1 | 2012 | 1 | | | 55 | 86% |
| **Turkey** | TUR | 65 | 13 | 20% | 4.3 | 3 | 1950 | 1962 | 1983 | | | 1 | 2.0 | 38 | 58% | 2 | 1966 | 4 | 1990 | 7 | 14 | 22% |
| **Tanzania** | TZA | 64 | 17 | 27% | 8.5 | 2 | 1958 | 1986 | | | | 1 | 7.0 | 13 | 20% | 1 | 1996 | 10 | | | 34 | 53% |
| **Ukraine** | UKR | 21 | 1 | 5% | 1.0 | 1 | 2005 | | | | | 0 | | 10 | 48% | 1 | 2006 | 1 | | | 10 | 48% |
| **Uruguay** | URY | 65 | 4 | 6% | 4.0 | 1 | 1981 | | | | | 0 | | 53 | 82% | 1 | 1985 | 4 | | | 8 | 12% |
| **Venezuela** | VEN | 65 | 5 | 8% | 5.0 | 1 | 1958 | | | | | 0 | | 40 | 62% | 1 | 1963 | 5 | | | 20 | 31% |
| **South Africa** | ZAF | 64 | 5 | 8% | 5.0 | 1 | 1990 | | | | | 0 | | 20 | 31% | 1 | 1995 | 5 | | | 39 | 61% |
| **Zambia** | ZMB | 58 | 8 | 14% | 2.7 | 3 | 1961 | 1990 | 2000 | | | 1 | 8.0 | 14 | 24% | 2 | 1961 | 0 | 2000 | 0 | 36 | 62% |

*Notes*: This table provides details on the sample-makeup of the 'treated' sample, i.e. the set of countries which experienced at least one democratisation episode followed by a regime change. There are 75 countries in this table, but only 62 of them have estimates for both the episode dummy and the democratic regime change dummy — these countries have their country names highlighted in bold. The remaining 13 countries have the following characteristics: four countries (marked in the 'ISO' column with †) experienced a reversal to autocracy, in one case followed by an unsuccessful democratisation episode. Nine countries (marked in the 'ISO' column with superscripts a-f) do not have any pre-episode data (and in some cases additionally experienced episodes lasting only one or two years), hence the regime change or the episode dummy is unidentified. In some more detail:

a) No pre-episode data, one-year episode before democratic regime change (BIH,MNG)

b) No pre-episode data, ten-year episode before democratic regime change, no regime change estimate (BRB, GEO)

c) No pre-episode data, seven-year episode before democratic regime change, no episode estimate (BWA)

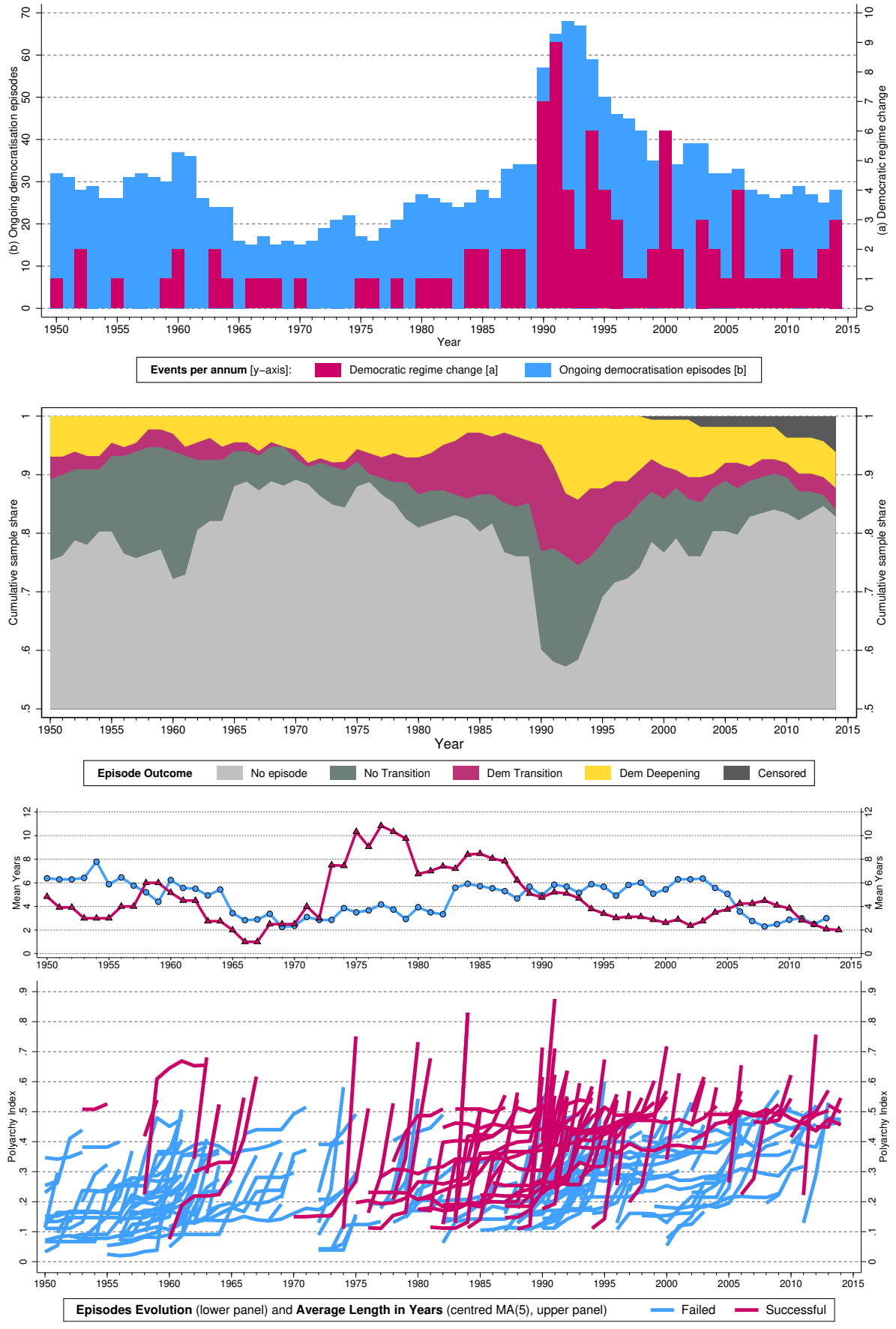
d) No pre-episode data, seven-year episode before democratic regime change, no regime change estimate (HRV)

e) No pre-episode data, two-year episode before democratic regime change, no regime change estimate (IND, JPN)

f) No pre-episode data, nine-year episode before democratic regime change, no regime change estimate (TTO)

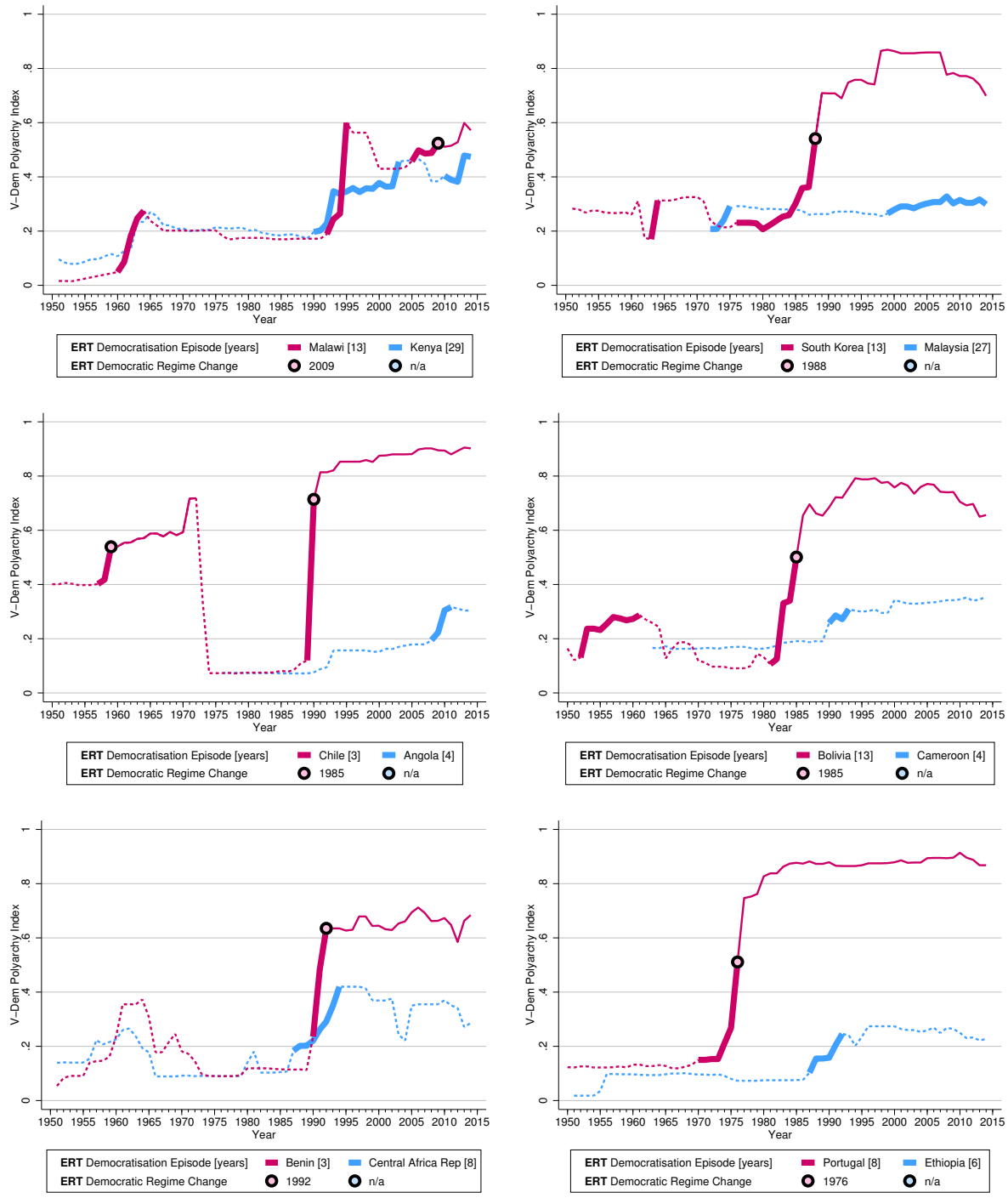
The first set of columns after the country name, ISO code and total observation count refers to information on the total number of episodes, their average length, and timing as well as the count and average length for failed episodes. The next set of columns refers to successful regime changes, how long countries spent in democracy (the 'years of treatment') and the length of the associated democratisation episodes (in years). The final two columns report the information on the pre-episode data. For years in episodes, democracy, and autocracy we report the share of total years, which adds up to 100% (even though of course episodes are *nominally* within the autocratic regime but we separate them out here).

**Figure A-1:** Episodes and Regime Change (1950-2014)




**Notes**: We present the distribution of democratisation episodes and regime changes in the top panel, the share of episode type in the middle panel, and the individual evolution of each episode in the lower plot along with the smoothed annual mean episode length (computed for episode start years) in the bottom panel.

**Figure A-2:** More Examples of Episodes and Democratic Regime Change




**Notes**: We present the V-Dem polyarchy index evolution for country pairs, where the country in dark pink experienced regime change and the country in light blue did not. The period highlighted by the thick line represents the democratisation episode, following ERT (the length of each episodes in years is indicated in the legend). The 'Eastern' end of the thick pink lines always coincides with the year of democratic regime change. A dashed (solid) thin line indicates the country regime is in autocracy (democracy) following the ERT definition. The circular marker indicates the year of democratic regime change (if applicable), which is required to include a 'founding election'.

# B Additional Figures and Tables

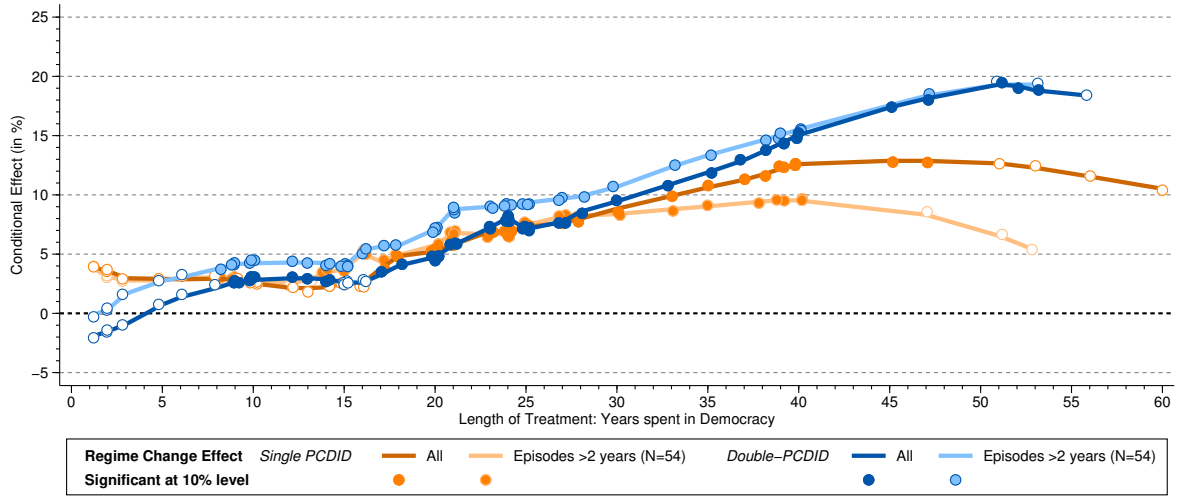
**Table B-1:** ATET Estimates: Single and Double PCDID

| **Panel (a) Double PCDID Results** | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Factors included | $1\times2$ | $2\times2$ | $3\times2$ | $4\times2$ | $5\times2$ | $6\times2$ |
| Democratic Episode | -2.832 | -1.170 | 0.959 | 0.351 | 0.548 | -0.310 |
| | [2.582] | [2.003] | [2.077] | [2.136] | [1.977] | [1.992] |
| Democratic Regime Change | 3.157 | 5.497* | 10.165*** | 6.845** | 6.645** | 6.785** |
| | [3.952] | [3.341] | [3.927] | [3.321] | [3.359] | [3.311] |
| Export/Trade Ratio | -0.212 | -0.213 | -0.172 | -0.224** | -0.166** | -0.180** |
| (in percent) | [0.147] | [0.134] | [0.116] | [0.091] | [0.084] | [0.088] |
| Population Growth Rate | -5.000** | -7.775*** | -7.540*** | -6.206*** | -7.054*** | -7.844*** |
| (in percent) | [1.991] | [2.100] | [1.883] | [1.593] | [1.694] | [1.994] |
| Treated Countries | 62 | 62 | 62 | 62 | 62 | 62 |
| Observations | 3660 | 3660 | 3660 | 3660 | 3660 | 3660 |
| Control Countries 1 | 15 | 15 | 15 | 15 | 15 | 15 |
| Observations | 631 | 631 | 631 | 631 | 631 | 631 |
| Control Countries 2 | 43 | 43 | 43 | 43 | 43 | 43 |
| Observations | 2472 | 2472 | 2472 | 2472 | 2472 | 2472 |

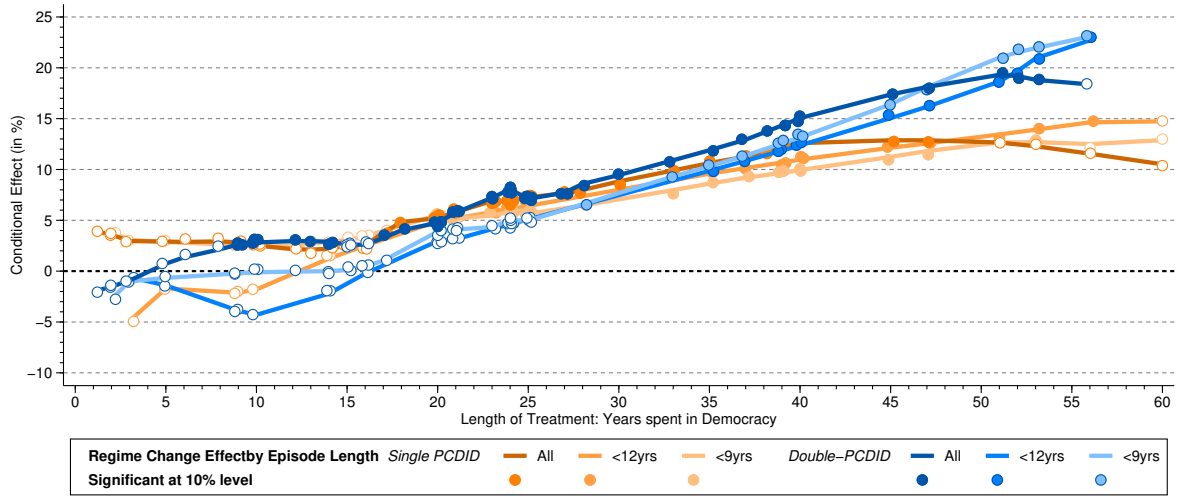
| **Panel (b) Single PCDID Results** | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Factors included | 1 | 2 | 3 | 4 | 5 | 6 |
| Democratic Regime Change | 5.914* | 3.877 | 8.601*** | 6.247*** | 6.710** | 7.738*** |
| (ERT definition) | [3.595] | [3.455] | [2.816] | [2.286] | [2.744] | [2.769] |
| Export/Trade Ratio | -0.304* | -0.369** | -0.363** | -0.179 | -0.076 | -0.092 |
| (in percent) | [0.175] | [0.150] | [0.157] | [0.121] | [0.107] | [0.101] |
| Population Growth Rate | -6.721** | -6.709*** | -7.059*** | -5.564*** | -6.445*** | -6.200*** |
| (in percent) | [2.891] | [2.584] | [2.661] | [2.025] | [2.083] | [2.090] |
| Treated Countries | 62 | 62 | 62 | 62 | 62 | 62 |
| Observations | 3724 | 3724 | 3724 | 3724 | 3724 | 3724 |
| Control Countries | 58 | 58 | 58 | 58 | 58 | 58 |
| Observations | 3161 | 3161 | 3161 | 3161 | 3161 | 3161 |

Notes: The table presents the Mean Group estimates from the Double and Single PCDID treatment regressions in Panels (a) and (b), respectively. The regime change effects can be interpreted as ATET. There are six different models for augmentation with 1 to 6 common factors — for the Double PCDID in Panel (a) there are separate factors from each of the two control samples, hence the number of factors is double that included in the Single PCDID models in Panel (b). Statistical significance is indicated using * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Figure B-1:** Episodes and their Implications for the Regime Change Effect



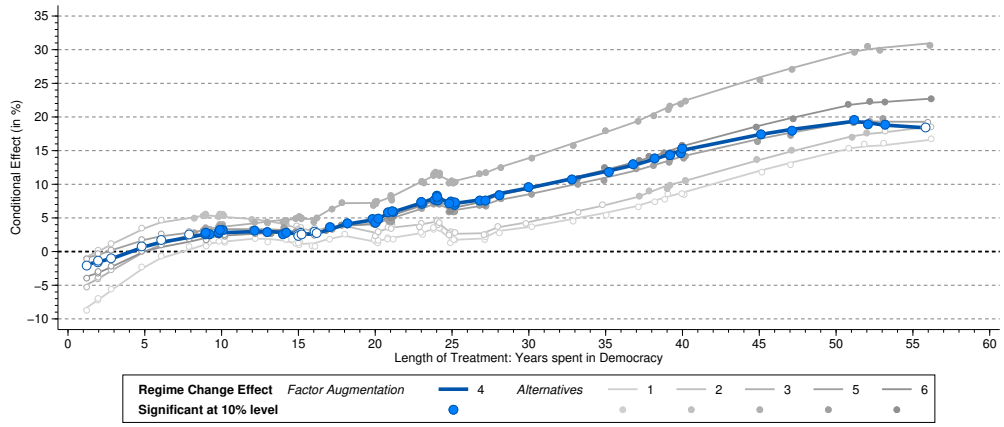

(a) Robustness: Exclude 'short' episodes ($\leq$2yrs)



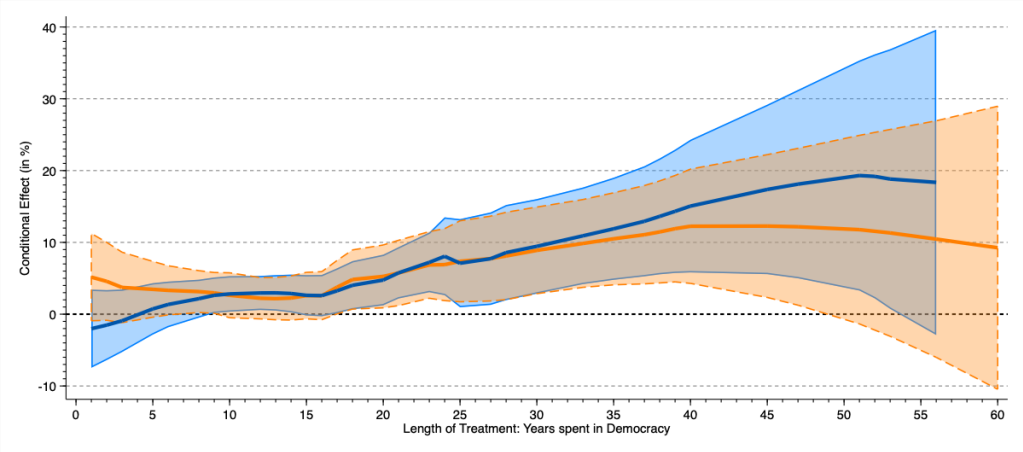

(b) Countries with <9 and <12yrs in Episodes: Regime Change Effect

**Notes**: These plots presents the results from running line regressions of country-specific coefficients on the democracy (ERT) dummy, derived from Single and Double PCDID estimates. In Panel (a) we compare Single (orange lines) and Double PCDID results (blue lines) for ERT in the full sample with those where countries with just one or two years spent in episodes are dropped. In Panel (b) we distinguish countries which had episodes lasting up to 9 years ($N = 33$) or up to 12 years ($N = 43$), respectively the median and 70th percentile, and find qualitatively no difference to the full sample ($N = 62$) results.

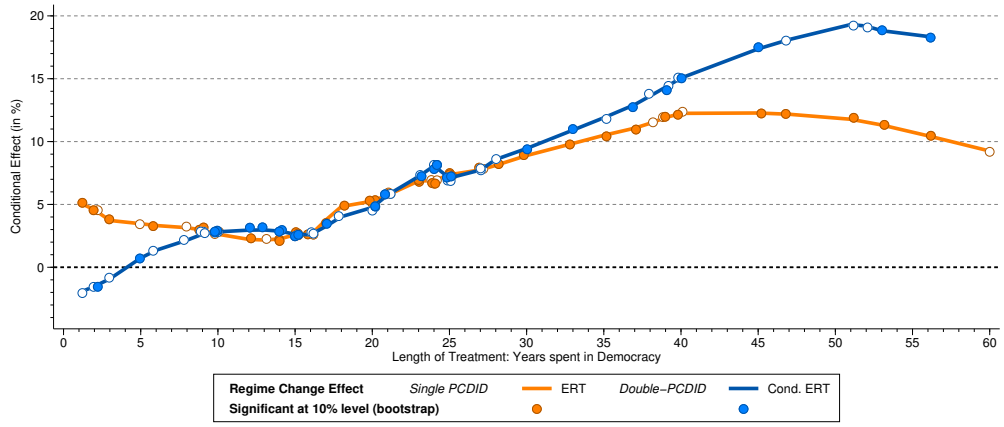
**Figure B-2:** Alternative Factor Augmentations, Confidence Intervals and Bootstrapped CI




(a) Alternative Factor Augmentations



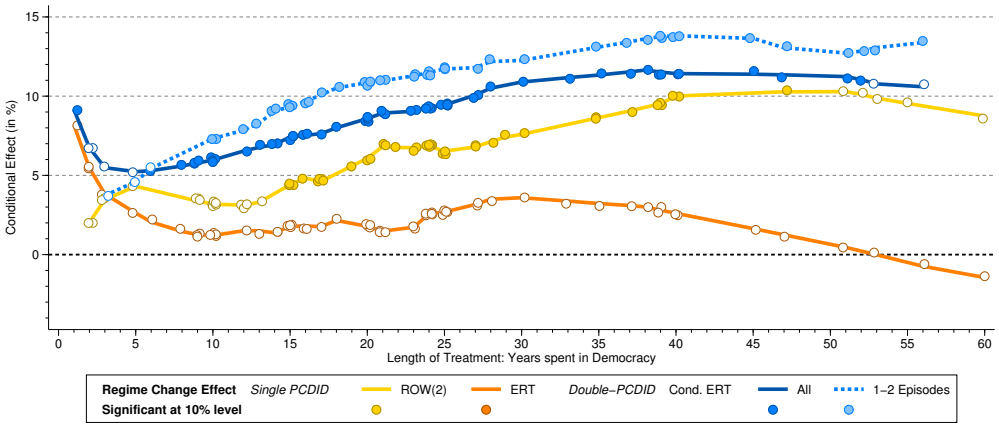

(b) Comparison Single and Double PCDID (90% CI)



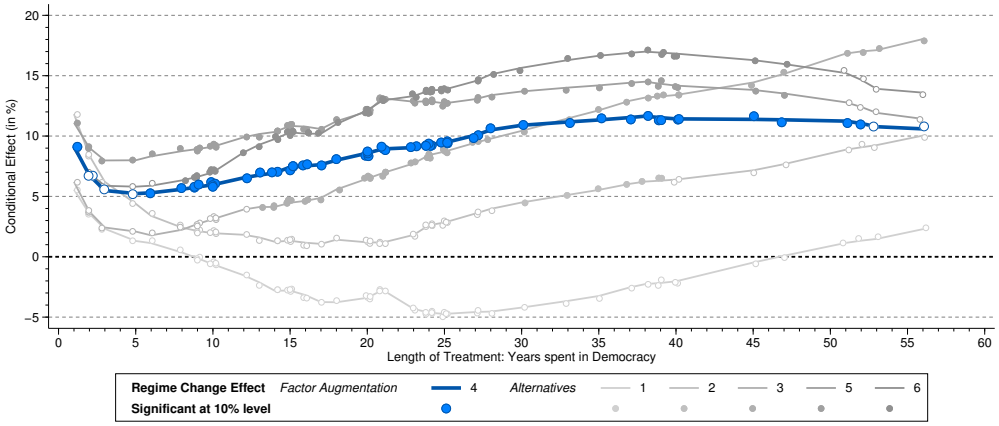

(c) Statistical Significance Based on Bootstrap

**Notes**: These plots presents the results from running line regressions of country-specific coefficients on the democracy (ERT) dummy, derived from Single and Double PCDID estimates. In Panel (a) we present the conditional ERT results from Double PCDID models augmented with 1 to 6 factors from each of the respective control groups: the blue line is for the model augmented with 4 estimated factors (from each respective control sample), grey lines present alternative augmentations using one to six factors (dto.). In Panel (b) we report the full sample results for ERT (Single and Double PCDID) but plot the 90% confidence intervals for each running line regression. In Panel (c) we signal statistically significant difference from zero in the running line regression adopting the bootstrap 90% confidence interval (250 replications). A hollow (filled) marker indicates that the bootstrapped 90% confidence interval does (not) include zero.

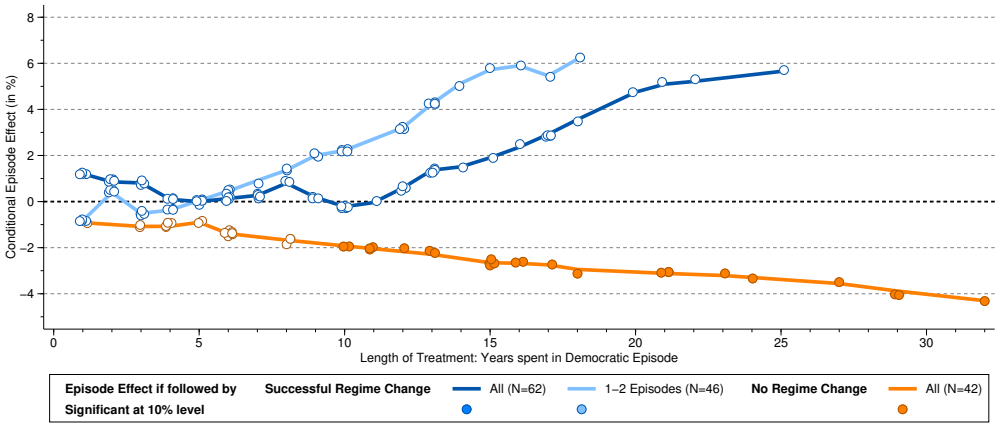
**Figure B-3:** Single and Double PCDID — Excluding Covariates




(a) Comparison of Single and Double PCDID Results




(b) Alternative Factor Augmentations: Conditional ERT




(c) Episodes effect in the Double PCDID

**Notes**: The results presented in this figure do not include any controls in the PCDID regressions. We present results from running line regressions of country-specific democracy coefficients on years spent in democracy or years spent in episodes, respectively. Additional controls in these running line regressions are the same as those in the analysis in the main text. The full sample matches that of the Double PCDID estimates for ERT (62 treated countries unless indicated), and with the exception of Panel (b) all results are for PCDID models augmented with 4 common factors for each control group. Panel (a) presents the results for Single PCDID alongside those for Double PCDID estimates — for the latter we contrast results for all countries with those which experienced only 1 or 2 liberalisation episodes (dark and light blue, respectively). Panel (b) presents results for the Double PCDID for 1-6 factors per control sample. Panel (c) focuses on the Episode effect, distinguishing countries which eventually experienced regime change (in blue) from those which did not (in orange).

# C Event Analysis

In this section we study the potential for idiosyncratic events, such as natural resource discoveries, natural disasters, or financial crises, exerting undue influence/bias on our PCDID estimates. Adopting dummies for each of the aforementioned events we employ event analysis to investigate the evolution of GDP per capita growth and the change in the V-Dem Polyarchy measure (the index underlying our episode and regime change data) up to five years before and after the event/crisis: we estimate country fixed effects models separately for each variable $k$ (growth, polyarchy) and event type:

$$y_{it}^{k} = \alpha_{i}^{k} + \sum_{s=-5}^{5} \beta_{\tau+s}^{k} \delta_{i,\tau+s} + \varepsilon_{it}^{k}, \tag{8}$$

where $\delta_{i,\tau+s}$ is a dummy equal to one if country $i$ is $s$ years away from the event at time $\tau$, $t$ indexes the years between 1950 and 2014, $\alpha_i$ is the country fixed effect and $\varepsilon$ is the error term. $s$ varies from $-5$ to $+5$, such that we evaluate each variable in the lead-up and aftermath of the event *relative to* the observations outside this 11-year window, with the latter interpreted as 'normal' times. Importantly, we compare the sample of countries which experienced regime change with the sample which experienced liberalisation episodes but no regime change, presenting results separately. Finally, we do not study crises/events at just any point in time, but focus on those which occur during democratisation episodes: if individual liberalising countries get bumped into or are prevented from realising democracy by a natural resource find, a financial crisis or a natural disaster, then this amounts to the type of idiosyncratic shock which threatens our identification strategy. The number of events in treated and control groups during episodes are tabulated in Table C-1 below. Since the event analysis includes a country fixed effects only countries which experienced a crisis/shock during a democratisation episode are included in the sample.

Although there are ample reasons for spillovers across countries for each event type, our primary reason for selecting these economic events/crises is that they are typically regarded as *country-specific* events, with the respective literatures (at least for the economic crises) seeking to explain their prevalence largely with country-specific determinants.

We adopt data on new oil discoveries from Cotet & Tsui (2013): we define a boom as the point in time when either (i) the 3-year moving average of the growth rate of new oil discoveries (in billion barrels) is at least 100% and the magnitude of the discovery is at least half a billion barrels; or (ii) when the 3-year moving average of the growth rate of new oil discoveries (in billion barrels *per capita*) is at least 100% and the magnitude of the discovery is at least half a million barrels *per 1,000 population*.

**Table C-1:** Sample Makeup: Event Analysis samples

| | Oil boom | Banking Crisis | Currency Crisis | Natural Disaster | Full Sample |
|---|---|---|---|---|---|
| Treated $N$ | 16 | 18 | 27 | 10 | 62 |
| observations | 947 | 1104 | 1674 | 551 | 3660 |
| Control $N$ | 19 | 19 | 18 | 10 | 43 |
| observations | 1080 | 602 | 1085 | 606 | 2472 |

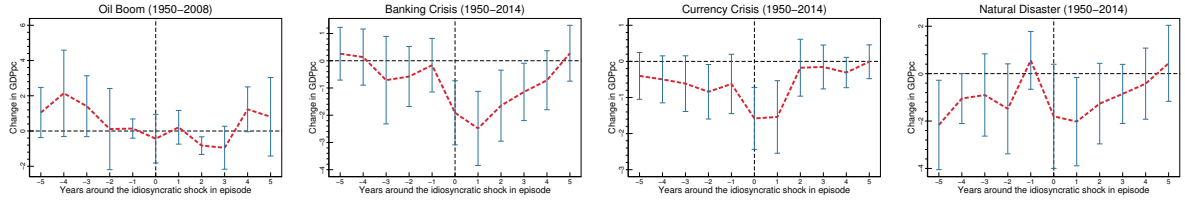
*Notes*: This table provides details on the crisis/event count in the treated and control groups for the episode-regime change event analysis. The full treated (control) sample (analysed in Figure 2 of the maintext) contains 62 (43) countries and 3,660 (2,472) observations.

For financial crises we augment the data collated by Carmen Reinhart — the expanded Reinhart & Rogoff (2009, RR) database — with information from Laeven & Valencia (2020, LL) — additional search established no further crises in the 1950s and 1960s (LL only starts in 1970). In all cases we mark the crisis start year; for banking crises we do not exclude 'ongoing crisis years' from the event analysis sample, in line with existing practice in the literature.

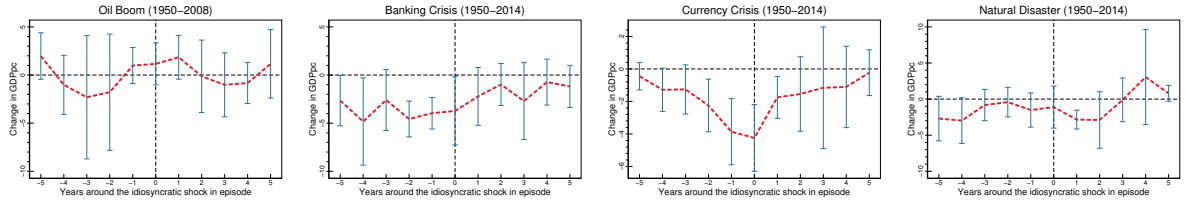
For natural disasters we use the EM-DAT database: EM-DAT, CRED/UCLouvain, Brussels, Belgium – www.emdat.be which covers primarily 'natural' disasters like earth quakes, floods or epidemics, but also large-scale industrial accidents and air/rail/road disasters. We construct a dummy for large-scale disasters by combining the EM-DAT information on associated deaths with Maddison (Bolt & van Zanden 2020) population data and select events with a death rate of 1 in 10,000 population.

The event analysis plots for per capita GDP growth and the annual change in polyarchy are presented in Figure C-1. Timings differ at times minimally, but the patterns of sign and statistical significance of the effects on growth and change in polyarchy between the treated and control samples are in general closely matched.

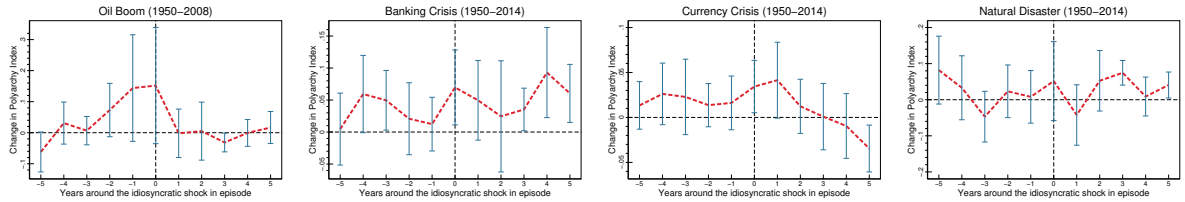
**Figure C-1:** Idiosyncratic Shocks in their Effect on Growth (a,b) and Polyarchy (c,d)



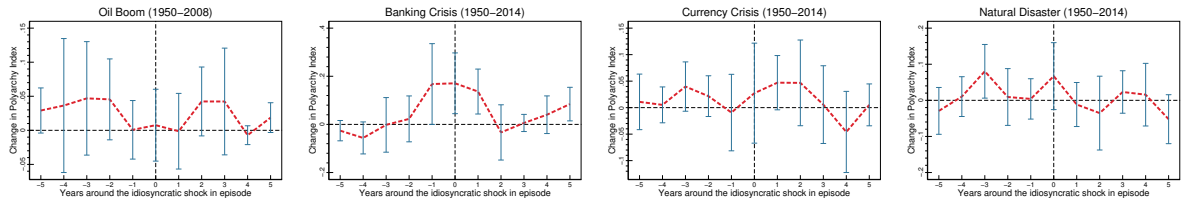

(a) Effect of Shock on *per capita Growth*: Treated Sample




(b) Effect of Shock on *per capita Growth*: Control Sample




(c) Effect of Shock on *Polyarchy*: Treated Sample




(d) Effect of Shock on *Polyarchy*: Control Sample

**Notes**: We present the results from event analyses for the GDP pc growth or polyarchy variables and the event as indicated. Event plots for growth are presented in panels (a) and (b), those for polyarchy in panels (c) and (d). In each case the first panel is for the treated sample, the second for the control sample. All of these are within-country estimates with standard errors clustered at the country-level. The vertical bars are the 90% confidence intervals.

# D Alpha Test of the Parallel Trend Assumption

We carry out the Alpha test for the parallel trend assumption in the Double PCDID models. The test is introduced in Chan and Kwok (2022), section 4.4, and works with the residuals from the auxiliary regression in the control sample. In the standard PCDID we estimate the treatment sample regression with factors estimated from $\hat{e}_{it}$ via PCA. In the Alpha Test, we would compute the cross-section average of the $\hat{e}_{it}$, say $\bar{e}_t$. and would enter this term in the PCDID regression *instead of* the estimated factors: $y_{it} = b_{0i} + d_i \mathbf{1}_{\{t>T_{0i}\}} + a_i' \bar{e}_t + b_{1i}' x_{it} + u_{it}$. We adjust this test to our new empirical setup with two control samples and estimate

$$y_{it} = b_{0i} + d_i^A \mathbf{1}^A_{\{t>T_0\}} + d_i^B \mathbf{1}^B_{\{t>T_1>T_0\}} + a_{1i}^{\cancel{A}} \bar{e}_t^{\cancel{A}} + a_{2i}^{\cancel{AB}} \bar{e}_t^{\cancel{AB}} + b_{1i}' x_{it} + e_{it},$$

where $\bar{e}_t^{\cancel{A}}$ and $\bar{e}_t^{\cancel{AB}}$ are the cross-section averages of the residuals from the auxiliary regressions in the control samples (a) for countries which never experienced an episode and (b) for countries which experienced episodes but not regime change, respectively. Since the second auxiliary regression includes between one and six estimated factors our Double PCDID Alpha Test has different results for different numbers of factors included. The null hypothesis is that the respective Mean Group estimates of $a_{1i}^{\cancel{A}}$ (for episodes) and $a_{2i}^{\cancel{AB}}$ (for regime change) are equal to 1, which would constitute 'weak parallel trends'. Results suggest by and large that this assumption is satsfied.

**Table D-1:** Alpha test for the weak parallel trends assumption

| Factors included | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **Panel (A)** Full Sample (N=62) | | | | | | |
| Episodes Test Statistic | -0.134* | -0.025 | -0.042 | -0.186*** | -0.191*** | -0.232*** |
| [p-value] | [0.07] | [0.77] | [0.61] | [0.01] | [0.01] | [0.00] |
| Regime Change Test Statistic | -0.105 | 0.096 | 0.115 | -0.166 | -0.139 | -0.295*** |
| [p-value] | [0.41] | [0.53] | [0.45] | [0.10] | [0.21] | [0.00] |
| **Panel (B)** Countries with only 1 or 2 Episodes (N=46) | | | | | | |
| Episodes Test Statistic | -0.042 | 0.080 | 0.060 | -0.112 | -0.112 | -0.154* |
| [p-value] | [0.65] | [0.44] | [0.54] | [0.20] | [0.19] | [0.07] |
| Regime Change Test Statistic | 0.088 | 0.323* | 0.342* | -0.035 | 0.022 | -0.130 |
| [p-value] | [0.58] | [0.09] | [0.07] | [0.78] | [0.87] | [0.25] |

Notes: We report the Alpha test statistics for weak parallel trends, with $p$ values in brackets. Panel (A) uses the full treated sample, Panel (B) the reduced sample for countries which experienced only one or two episodes. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# E An Early Warning System for Episode Failure

We study the determinants of episode failure with an 'early warning system' approach borrowed from the literature on banking crises (e.g. Eberhardt & Presbitero 2021). Our empirical results derive from a random effects logit model augmented with country-specific means of all covariates (Mundlak 1978, Chamberlain 1982) — this allows us to interpret coefficients as estimate 'within-country' estimates. We estimate a latent 'crisis' model, where the observed variable is a realized systemic event when the latent variable exceeds some threshold: we code the dependent variable as equal to one in the year a country exits an episode without transition to a democratic regime, and zero otherwise. With reference to our PCDID analysis, our sample includes all 'treated' countries and those in 'control sample 2'.

**Determinants** Our analysis focuses on economic and social conditions alongside 'upheaval' in an early warning system for episode failure.[25] We consider determinants from six different groups (for data sources see Appendix C): First, we control for *changes in economic conditions* using per capita GDP growth and the change in export/trade. Second, given the relevance of the *military* in the context of regime fragility, we capture the buildup of military personnel using data from COW. Third, *natural disasters* may devastate economic and social conditions in a country and foster the return to a 'strong' (autocratic) leader or accelerate a transition to democracy if incumbent systems are shown to be failing. Fourth, *financial crises* (banking or currency crises) can similarly undermine episodes if the uncertainty created by the political process is blamed for financial fragility. Fifth, the relationship between *natural resources* and political institutions is a fraught one, and we include an indicator for oil booms to capture this variant of the natural resource curse. Finally, since *coups attempts* relate to democratic failure they make a natural candidate to study episode failure.[26]

[25] Since all regressions capture 'within-country' effects, levels variables (e.g. per capita income) are discounted in favour of changes or 'trigger' crisis variables (see Eberhardt & Presbitero 2021).

[26] Our regressions do not include a measure for episodes since this would per fectly predict failure.

**Variable transformation** An important aspect of our empirical investigation is how we model the pre-failure period: simply lagging all regressors by one time period may ignore slow-moving processes contributing to episode failure, e.g. lacklustre economic performance *over several years*. We therefore provide additional results following the practice in Eberhardt & Presbitero (2021), among others, using moving averages of regressors to capture pre-failure dynamics: we adopt MA(2) to MA(5) transformations, with the MA(2) defined as the mean of values at $t-1$ and $t-2$, and the higher-order MAs adding further lags accordingly. Higher-order MA-transformations come with the potential caveats that they 'wash out' spikes, while for shorter episodes this transformation may conflate events *during* and *before* the episode.[27]

**Results** Our results in Table E-1 are presented as average marginal effects (in %): for continuous variables we multiply the margins with the standard deviation while for binary indicators we report magnitudes for a shift from 0 to 1. Standard errors are computed using the Delta method, those in the underlying RE-Logit regressions are clustered at the country-level. Wald tests uniformly confirm the inclusion of within-country averages.

Across specifications the effect of military buildup has a consistent and statistically significant attenuating effect on episode failure: a one standard deviation increase in military personnel is associated with a 0.2-0.3 percentage point decrease in the propensity of a failure. A second consistent and statistically significant trigger of episode failure is the presence of an oil boom, underlining the detrimental effects of natural resources in the political process. Finally, in models allowing for a longer time horizon in effects to accumulate — the MA(3)-MA(5) models — banking crises seem to reduce the propensity of episode failure. In economic terms the effects of banking crises and oil booms are an order of magnitude larger than that of a change in military personnel. Economic effects of all other variables studied are small and/or imprecisely estimated. Most notably, although with the correct positive sign, the presence of coup attempts has a comparatively small effect and is estimated imprecisely.

[27] We estimate these models for the most elaborate specification. Using a comparison of ROC areas suggests that this model has borderline ($p = .10$) higher predictive power than the model excluding coup attempts.

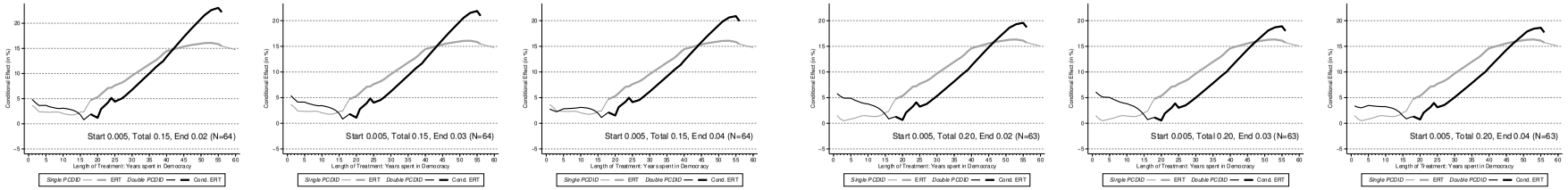
**Table E-1:** Determinants of Episode Failure (RE Mundlak Logit)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Transformation | Lagged by one period | | | | | MA(2) | MA(3) | MA(4) | MA(5) |
| Dep. Variable | Failure of Episode at time $t$ | | | | | | | | |
| **Continuous Covariates (magnitudes in percent for 1sd increase)** | | | | | | | | | |
| $\Delta$ Military personnel | -0.180 (1.89)* | -0.190 (1.95)* | -0.172 (1.86)* | -0.183 (2.05)** | -0.171 (1.95)* | -0.138 (1.17) | -0.228 (1.70)* | -0.285 (1.83)* | -0.361 (2.34)** |
| $\Delta$ Export/Trade | -0.161 (0.86) | -0.161 (0.87) | -0.160 (0.88) | -0.153 (0.84) | -0.148 (0.86) | -0.111 (0.59) | -0.110 (0.51) | 0.049 (0.24) | -0.013 (0.06) |
| GDPpc growth rate | 0.213 (1.27) | 0.227 (1.31) | 0.198 (1.21) | 0.200 (1.20) | 0.221 (1.32) | 0.104 (0.41) | 0.083 (0.37) | 0.026 (0.13) | 0.007 (0.03) |
| **Binary Indicators (magnitudes in percent)** | | | | | | | | | |
| Natural disaster | | 0.269 (0.19) | 0.251 (0.18) | 0.278 (0.20) | 0.255 (0.18) | -0.104 (0.12) | -0.624 (0.76) | -0.047 (0.08) | 0.429 (1.07) |
| Banking crisis | | | -2.596 (1.03) | -2.556 (1.02) | -2.577 (1.03) | -1.579 (1.07) | -2.738 (1.85)* | -2.823 (2.27)** | -2.149 (2.43)** |
| Currency crisis | | | -0.277 (0.23) | -0.382 (0.31) | -0.354 (0.29) | -0.342 (0.36) | 0.395 (0.55) | 0.21 (0.34) | 0.662 (1.38) |
| Oil boom | | | | 3.676 (1.89)* | 3.655 (1.90)* | 1.768 (0.94) | 2.770 (1.84)* | 2.329 (1.71)* | 2.858 (2.14)** |
| Coup Attempt | | | | | 0.646 (0.86) | 0.329 (0.67) | 0.315 (0.88) | 0.250 (0.83) | 0.292 (1.09) |
| Observations | 4,905 | 4,905 | 4,905 | 4,905 | 4,905 | 5,008 | 5,111 | 5,214 | 5,316 |
| Countries | 102 | 102 | 102 | 102 | 102 | 102 | 102 | 102 | 102 |
| Events | 125 | 125 | 125 | 125 | 125 | 127 | 130 | 132 | 135 |
| LogL | -577.11 | -575.04 | -571.53 | -569.86 | -566.82 | -577.34 | -586.11 | -595.23 | -605.27 |
| AUROC | 0.584 | 0.591 | 0.615 | 0.628 | 0.647 | 0.641 | 0.664 | 0.663 | 0.668 |
| se(AUROC) | 0.025 | 0.026 | 0.024 | 0.025 | 0.024 | 0.024 | 0.023 | 0.022 | 0.023 |
| Wald $chi^2$ (FE) | 7.45 | 12.22 | 19.10 | 22.55 | 26.57 | 25.82 | 27.09 | 26.20 | 22.19 |
| Wald $p$-value | 0.059 | 0.016 | 0.004 | 0.002 | 0.001 | 0.001 | 0.001 | 0.001 | 0.005 |
| Uncond. Prob. | 2.55% | 2.55% | 2.55% | 2.55% | 2.55% | 2.55% | 2.55% | 2.55% | 2.55% |

Notes: The table presents results (marginal effects) of a Random Effects 'Mundlak-Chamberlain' Logit regression, see text for details. We run 'early warning system' regressions akin to the analysis in the literature on banking crises. Our dependent variable is a dummy for the failure of a democratic episode in year $t$. All regressors in columns (1) to (5) are lagged one period ($t-1$), those in columns (6)-(9) are MA-tranformed as indicated at the top of the table. All results for continuous variables are marginal effects (in percent) for a 1 standard deviation increase in the covariate, e.g. -.180 in column (1) indicates that a 1sd increase in military personnel is associated with a 0.18 percentage point reduction in the propensity of an episode failure. The results for binary indicators are also in percent and for a shift from 0 to 1. The unconditional propensity is 2.55%. 'Events' refers to the number of episode failures in the sample, which comprises the treatment sample and control sample 2 from our main analysis. LogL reports the maximised log likelihood. AUROC is the Area under the receiver operating curve (a higher value indicates higher predictive power; 0.5 is the lower benchmark, 1 the maximum), with the associated standard error also reported. The Wald test analyses the within-country average terms and implicitly tests whether country FE 'matter' ($H_0$: FE jointly insignificant). Absolute t-ratios in parentheses. Statistical significance is indicated using * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# F Alternative Definitions for Episodes

**Figure F-1:** Alternative Episode Definition




(a) Trigger ≥0.005, Total Change ≥0.15 (C 1/2: N=20/38)

(b) Trigger ≥0.005, Total Change ≥0.20 (C 1/2: N=31/27)

(c) Trigger ≥0.010, Total Change ≥0.15 (C 1/2: N=22/36)

(d) Trigger ≥0.010, Total Change ≥0.20 (C 1/2: N=33/25)

(e) Trigger ≥0.015, Total Change ≥0.15 (C 1/2: N=22/36)

(f) Trigger ≥0.015, Total Change ≥0.20 (C 1/2: N=35/23)

**Notes**: We present predictions from running line regression results for Single and Double PCDID estimates using different definitions for episodes. Each graph indicates the number of treated countries, each panel header (a) to (f) the number of countries in the two control samples ('C' 1/2). 'Trigger' is for minimal annual change to start an episode (0.02-0.04), 'Total Change' is 0.15 or 0.20. Each panel has three plots for annual change to 'terminate' an episode: 0.02, 0.03, and 0.04 (from left to right). The ERT default parameters were developed by the data authors to best capture actual episodes of political change. Alternative episode definitions consequently deviate from this first best with implications for the validity of the democracy-growth nexus represented in these figures.