Abstract: What are the economic consequences of changes to the system of governance? Though the existing empirical literature on the growth effect of democratic change covers a wide range of methods and specifications, we note the absence of three important aspects: First, an empirical specification which allows the democracy-growth relationship to differ across countries. Such heterogeneities can be motivated with arguments for democratic legacy or thresholds in economic or human development as necessary conditions for positive democracy effects. Second, the adoption of methods accounting for the distorting effect of time-varying unobserved heterogeneity, which can arise from common shocks and/or from spillover effects. And third, flexibility in the functional form of the democracy-growth nexus: existing studies either assume a linear effect relationship or adopt ad hoc single thresholds to distinguish ‘democratic’ from ‘autocratic’ regimes. Our specifications allow for (i) a linear, (ii) single threshold, or (iii) dual threshold relationship, for which we adopt a novel IV estimator (Norkuté et al. 2019), a difference-in-difference estimator (Chan & Kwok 2018), and a simple extension of the latter, respectively — all implementations are heterogeneous panel estimators and account for strong cross-sectional correlation. Afforded by the internal coherence of the Varieties of Democracy (V-Dem) data we are able to provide a closer mapping between an encompassing theory of democracy and its constituent components on the one hand, and our empirical analysis covering up to 159 countries (1960-2017) on the other. Employing high-, mid-, and low-level democracy indices we can trace the positive democratic dividend we establish to the most relevant constituent components which appear to drive these results.

JEL Classification: O10, P16

Keywords: Democracy, Growth, Political Development, Spillovers, Two-Stage Least Squares, Interactive Fixed Effects, Difference-in-Difference
1 Introduction

Until recently the empirical literature on long-term economic effects of democratisation has provided mixed evidence on the significance and magnitude of any ‘democratic dividend’ — prominent examples include Barro (1996, 1999), Tavares & Wacziarg (2001), Gerring et al. (2005), Giavazzi & Tabellini (2005), Rodrik & Wacziarg (2005), Papaioannou & Siourounis (2008), Persson & Tabellini (2009), and Knutsen (2013). A recent study by Acemoglu et al. (2019, henceforth ANRR) now offers positive and robust evidence of a link between democracy and growth in a large sample of countries. The authors adopt a variety of empirical strategies which account for country-specific fixed effects and the dynamics of GDP. In order to allow for a causal interpretation of the results they devise an instrumentation strategy which builds on regional waves of democratisation and reversal. The collective message from ANRR’s analysis is that the long-run effects of democratisation are sizeable: an increase in per capita GDP of 20% or more. These results were however shown to be highly sensitive to sample selection (e.g. exclusion of the data for 2010 rendered the 2SLS results statistically insignificant) and an extension to their analysis using policy evaluation methods with interactive fixed effects arrived at a robust long-run effect of a more modest 10% (Eberhardt 2019).

The ANRR paper results (and by implication those of the extension) are of particular significance since the definition of democracy, in form of a binary indicator based on various established democracy indices, is very simple and almost ‘blunt’: if such a crude categorization of democracy versus autocracy can successfully reveal significant quantitative effects, then the existence of a democratic growth dividend is surely difficult to reject. ANRR argue that their binary indicator successfully captures a “a bundle of institutions that characterize electoral democracies” (footnote 4), although they concede that it misses elements of a “broader set of inclusive institutions” (ibid) emphasized in other work by Acemoglu & Robinson (2012). Which elements of the ‘bundle’ matter most, if indeed they are not all of equal significance, is a question left unanswered.

In the comparative political science literature there is a near-consensus that a binary distinction between democracy and autocracy such as in ANRR is perhaps too blunt to capture the differences in level or quality of democratic regimes (see Bollen & Jackman, 1989; Bühlmann et al., 2008; Coppedge et al., 2011): a dichotomous distinction is deemed too insensitive, especially when examining the subtleties of hybrid regimes/anocracies and democratic or autocratic transitions (Mainwaring et al. 2007, Gates et al. 2007). Since changes in systems of governance frequently happen gradually and in small increments, the adequacy of a dichotomous measure is seriously questioned. Some scholars, however, have argued that the distributions of most polychotomous indices “are actually bimodal, with a high concentration of cases in their low and high ends” (Cheibub et al. 2010, 77) thus adding little information to a dichotomous classification. In contrast, Gates et al. (2007) demonstrate in the context of Huntington’s (1993) Third Wave hypothesis that this ‘little added information’ of a polychotomous index may be precisely what is needed for an examination of democratic and autocratic transitions.

In this paper we break up any notion of a blunt, dichotomous characterization of democ-

---

1We follow the practice of ANRR (see their footnote 1) in using ‘growth’ as a short-hand for economic development (the level of per capita GDP). See Eberhardt and Teal (2011) for a more detailed discussion of growth empirics. For brevity GDP is at times used instead of per capita GDP.
racy in two different ways: first, we study democracy as a multi-faceted concept and this multidimensional character is captured by adopting various indices and proxies from the recent Varieties of Democracy (V-Dem) dataset for up to 159 countries over the 1960-2017 period. We employ V-Dem data in order to trace the driving forces of the democracy-growth nexus from broad definitions of democracy to precise lower-level constituent components through a set of layers, which provide a closer mapping between the theory developed in Section 2 and the empirical analysis. The top layer is captured by a high-level index for democracy which encompasses the three characteristics of democracy highlighted by political scientists (see Figure 1): electoral participation, electoral competition, and executive constraints (incorporated in V-Dem’s high-level ‘liberal democracy’ index). The next layer splits these concepts into their constituent elements, namely an ‘electoral’ component of the former two (the electoral democracy or ‘polyarchy’ index following Dahl, 1971, widely recognised as closest in conceptual coverage to the polity2 variable from Polity IV), and a ‘liberal’ component covering rule of law and related aspects of judiciary independence. A third layer sees each of these mid-level indices split into low-level components, such as freedom of speech, freedom of association, suffrage, elected executives, and clean elections in case of the polyarchy index. Thus, from an encompassing institutional concept of democracy we can move all the way to much more tangible factors such as the extent to which elections are clean and fair. Our empirical analysis studies the proxies in these different layers to trace the positive democratic dividend we establish across all specifications to the most relevant constituent components.

Second, from the outset we are able to remain agnostic over the specification choice between polychotomous versus dichotomous proxies for democracy, as well as possible intermediates, by studying all of these options in a common empirical framework. Briefly, this common framework has two main features: (a) the potential for countries to have different long-run relationships between democracy and growth, motivated by existing arguments for a ‘democratic legacy’ (Gerring et al. 2005) or threshold levels in economic or human development (Aghion et al. 2007, Madsen et al. 2015, ANRR) as necessary conditions for a positive democracy-growth nexus; and (b) the potential presence of common shocks with heterogeneous impact across countries and/or of spillover effects which induce ‘cross-section correlation’ in the panel. The macro panel econometric literature has spent almost two decades emphasising the bias in empirical estimates if this data property is ignored (seemal contributions include Bai & Ng 2004, Andrews 2005, Pesaran 2006, Bai 2009).

These features contrast with the practice in the literature to specify a democracy-growth nexus which is common to all countries and to ignore any of the features inducing strong cross-section correlation. Inherent choices over functional form (polychotomous versus dichotomous measures of democracy) are typically not given much thought in the literature either:

In spite of being available since 2014 only V-Dem data is already regarded as “the new standard in democracy measurement” (see Boese 2019, 26). The V-Dem indices have a number of distinct advantages over the popular Polity IV index which we discuss in detail in Section 4. In brief, the relative shortcomings of Polity IV are primarily in terms of (i) the absence of clear theoretical foundations, (ii) no justification for the weighting and aggregation rules applied to the constituent measures, (iii) ambiguities and vagueness in relation to periods of interregnum, interruption and transition (see Boese 2019, for a detailed comparison).

Since a priori the magnitude and even the sign of the democracy effect is unknown, one can instead point to differential GDP dynamics across countries (speed of convergence to the long-run equilibrium path) as an additional motivation for a heterogeneous parameter setup.
researchers focus either on single autocracy-democracy cut-offs/thresholds (e.g. Giavazzi & Tabellini 2005, Persson & Tabellini 2006, Papaioannou & Siourounis 2008, Cervellati et al. 2014, ANRR); or they employ index variables from Freedom House or Polity IV in levels, or transformations thereof (e.g. Leblang 1997, Knutsen 2013, Murtin & Wacziarg 2014, Madsen et al. 2015). Instead of opting for one or the other our specifications speak to both the levels and threshold approaches practiced in the literature, captured and extended as follows: we begin with a hypothesised linear relationship between democracy and growth, where a marginal increase in democracy has the same effect regardless of the level of democracy achieved; the specifications here simply adopt the V-Dem indices as independent variables. Next, we replicate the specifications using single democracy thresholds as a way to capture a nonlinear relationship using two strategies: first, we adopt the V-Dem Regimes in the World data, which represent high-quality indicators of two respective autocratic and democratic regimes, and hence qualitatively improve on crude numerical cut-offs ignorant of the specific country context. Second, we return to crude numerical cut-offs but ask whether growth effects vary systematically across different cut-off levels, adopting V-Dem indices (varying from 0 to 1) and all feasible cut-offs with treatment and control samples in excess of 20 countries, respectively. Finally, we attempt to capture more complex nonlinearities, e.g. an S-shaped or inverse S-shaped relationship between democracy and growth, by extending the single threshold specification to a dual threshold model. By not ruling out simple linear or more complex nonlinear specifications we avoid getting caught up in quibbles over appropriate cut-offs or accusations of overly simplistic modelling approaches.

Regarding our empirical implementation we can further highlight the adoption of a novel empirical method which combines the insights of the panel time series literature (see Chudik & Pesaran (2015) for a recent survey) with the established literature on instrumental variable (IV) regression. The Norkuté et al. (2019, henceforth NSYC) IVMG estimator is a heterogeneous parameter dynamic panel estimator with a multi-factor error structure. The identification strategy maintains that all observed variables — dependent variable, $y$; independent variables, $X$, including one or more endogenous variables; and potential external instruments, $Z$, for said endogenous variable(s) — are driven by some of the same unobserved common factors. Since these factors have heterogeneous factor loadings they cannot be captured using country fixed effects or common time dummies. The core insight of NSYC is that de-factored independent variables, $X^*$, are driven by (with reference to the omitted factors) exogenous variation, so the exclusion restriction is satisfied by construction. The IV estimator is valid and informative provided the correlation between $X$ and $X^*$ is not too weak — a characteristic which is easily measurable with conventional tests.

The commonality of unobserved common factors in treatment and control groups also drives the motivation for the recent policy evaluation literature (binary treatment variable) adopting a multi-factor error structure (Gobillon & Magnac 2016, Chan & Kwok 2018), which

---

4 A third approach is to accumulate index level data to construct discounted stocks of democracy (e.g. Gerring et al. 2005, Persson & Tabellini 2009). We leave this notion of democratic capital for later analysis with the newly-created V-Dem stock data forthcoming in 2019.

5 De-factoring can be achieved by adopting principal component analysis, like in the Bai (2009) estimator, or by the use of cross-section averages of $X$ in an auxiliary regression, as developed in the Pesaran (2006) estimator and related methods (Pesaran et al. 2013).
we implement for our single and dual threshold models (following Eberhardt 2019).

The remainder of this paper is organised as follows: in the next section we discuss the conceptual foundations of a link between democratic change and economic performance. Section 3 introduces the common factor model setup and the various empirical implementations in greater detail. Section 4 covers the data sources and presents a detailed descriptive analysis of the patterns of democratic change over time. Main results are presented in Section 5, Section 6 concludes.

2 Conceptual Development

We provide a brief discussion of the channels linking a country’s system of governance to its economic performance.

2.1 The rules of the game

(Rodrik et al. 2004, 132) argue “what matters are the rules of the game in a society and their conduciveness to desirable economic behaviour”. To analyze the interactions between democracy and economic outcomes it is central to understand these rules and how they shape said interactions.

Assume a society consists of two groups, the comparatively poorer people and the wealthier elite running the country. This is but an extension of the model developed in Acemoglu & Robinson (2001) in that the poor are not necessarily disenfranchised and the elite is not necessarily autocratic. In the spirit of Lake & Baum (2001) and Baum & Lake (2003) the elite then functions as a local monopolist of public services, education etc. who “seeks to exploit their market power to produce rents that can be redistributed to either the holders of the state power”, i.e. the elite, or the people. In such a framework the optimal quantity supplied by an unconstrained monopolist/elite is lower than the social welfare maximizing quantity.

In this framework the central argument shaping the rules of the game in a society is that “democracy is a system in which parties lose elections” (Przeworski 1991, p.10). The degree of contestation for office, judicial and legislative constraints placed on the elite, and participation of the people determines whether the monopolists produce in a contested or uncontested market, i.e. whether the elites need to fear being replaced at regular intervals. In a perfect democracy with competitive elections, full participation and constraints on the executive this will drive the monopolist elite to produce the socially optimal quantity (see Lake & Baum 2001).

In addition to providing business enhancing environments characterized by low uncertainty political institutions, when functioning in a democratic sense, hence provide a safety net limiting extractive powers of predatory rulers. Together political institutions provide state autonomy from particularistic pressures by special interest groups while at the same time ensuring the regime’s accountability to the people. Knutsen (2012), p.400. In democracies both influence from special interest groups, such as unions or pivotal voting blocs, as well as the regimes’ accountability to its citizens are higher than in autocracies

---

2.2 Elections

This substitutability argument has far-reaching economic consequences. Broadly speaking, it implies that the range of implementable policies in democratic countries is narrower than in autocratic ones. Competitive negotiation processes in democracies make resource mobilization and policy implementation more difficult. More autocratic leaders in contrast are less bound by judicial or legislative constraints and need to fear less for their re-election (Reuveny & Li (2003)). This introduces a heterogeneity in expected policy outcomes: a strong autocrat can implement policies inducing long-term development and growth not implementable in democracies (e.g. the substantial economic growth in non-democratic China) but at the same time they can cause severe harm to the economy and the country as a whole, for example, Bashir al-Assad in Syria.

With respect to development the substitutability argument implies that democracies are more responsive to the poor who in turn push for more redistributive policies, egalitarian distributions of income, and increased investments in human capital (public health and education). In addition, the substitutability argument and the consequential comparative stability of policy outcomes implies indirect positive effects of democratic institutions on investments (Tavares & Wacziarg 2001). This becomes apparent when separating the political from the economic aspects of democracy (such as protection of property rights, business and labour market regulations). These economic aspects of democratic systems are in themselves conducive to growth (Baum & Lake 2003). In addition, they raise the expectations about people’s returns to investments in the future exerting another positive effect on growth through investment (Persson & Tabellini 2009).

The higher variation of policy outcomes in autocracies makes inferring a “net-effect” of a country’s system of governance on its economy difficult. Even within the range of democratic countries this effect remains unclear. Tavares & Wacziarg (2001), for example, find that the increase in distributive policies in democracies comes at the expense of physical capital accumulation in turn exerting a negative effect on economic growth. While it is hard to identify such a “net-effect” the endogeneity between development and democracy is worth noting at this point. Several scholars in favour of the “modernization hypothesis” (attributed to Lipset 1959) argue that a certain level of development acts as a prerequisite for democracy. Increased levels of socio-economic development are both outcome and determinant of democracy (Acemoglu et al. 2019, Doucouliagos & Ulubaşoğlu 2008, Madsen et al. 2015, Narayan et al. 2011).

3 Empirical Strategies

This section introduces methods to capture the impact of observable and unobservable heterogeneity on empirical estimates of the democracy-growth nexus. Our models build on the panel time series econometric literature which since Pesaran & Smith (1995) has emphasised

---


8 The pitfalls of imposing common slope coefficients on heterogeneous equilibrium relationships have been highlighted for dynamic (Pesaran & Smith 1995) and static specifications (Sul 2016). It is also worth emphasising that any instrumentation strategy applied in a pooled panel (such as the IV strategy in ANRR) will be invalid by construction if the true underlying equilibrium relationship differs across countries. If the coefficient imposed on $x$ is
heterogeneous parameters across panel members, and, more recently, the presence of strong cross-section dependence (Pesaran 2006, Bai 2009, e.g.) — a form of unobserved, time-varying heterogeneity.\footnote{Eberhardt & Teal (2011) provide a detailed introduction to these models with discussion of empirical applications from the cross-country growth literature.}

Strong correlation across panel members is distinct from weaker forms of dependence, such as spatial correlation, and if ignored can lead to serious bias in the estimated coefficients on observable variables (Phillips & Sul 2003, Andrews 2005). This literature has taken to specifying a multi-factor error structure, \( X' f_t \), where \( f \) is a set of common factors with associated heterogeneous factor loadings \( \lambda \), to capture this strong dependence. Detailed discussions of how to motivate and implement the investigation of observed and unobserved heterogeneity in the context of the cross-country production function which underlies the empirical growth literature can be found in Eberhardt & Presbitero (2015) and Eberhardt & Teal (2019).

In the following we detail two different approaches we adopt to measure the impact of democracy on economic development: an approach assuming a linear relationship which identifies the causal effect via an instrumental variable setup, and a difference-in-difference approach which extracts unobservables from control countries to identify the causal effect in the face of endogenous selection into treatment, non-parallel pre-treatment trends, and common shocks with heterogeneous impact.

### 3.1 IV Estimation of Dynamic Linear Panel Data with Multifactor Error Structure

In the common factor framework we assume that all macro variables are driven by small set of common factors (see, for instance, Stock and Watson, 2002, in the context of forecasting). Any source of omitted variable bias is hence assumed to derive from the presence of common factors in dependent and independent variables, with instrumentation thwarted by the assumption that external instruments are also partly driven by the same common factors. A novel estimator by Norkute, et al. (2019) exploits this setup in an ingenious way to construct ‘internal’ instruments which are valid by (the above) assumption. We specify a dynamic ARDL model:

\[
\begin{align*}
  y_{it} & = \rho_i y_{it-1} + \beta_1 \text{Democracy}_{it} + \gamma_i X_{it} + u_{it} \\
  u_{it} & = X_i' f_t + \varepsilon_{it} \\
  X_{it} & = \phi_i f_t + \nu_{it} \\
  \text{Democracy}_{it} & = \phi_i f_t + \eta_{it},
\end{align*}
\]

where Democracy is a continuous variable and \( X \) is a set of additional controls. As is apparent, both Democracy and \( X \) are endogenous with respect to the error term \( u \), since they all contain the same set of common factors \( f \).

The implementation of this approach proceeds in three steps: first, all variables are 2FE-transformed (within-transformed and cross-sectionally de-meaned) by separately regressing them on a set of country and time fixed effects and extracting the residuals from this regression

\(\beta\) yet the true relationship is \(\beta x\) then \((\beta - \beta) x\) will be contained in the error term, thus violating the exclusion restriction that instrument \( z \) be uncorrelated with the error if \( E[zx] \neq 0\).
— this accounts for the presence of deterministic terms; second, the 2FE-transformed independent variables are de-factored — more on this below; third, the de-factored regressors, denoted as $X^*$, are used as instruments for the

Quite simple and intuitive idea for instrumentation: extract the common factors from Poly and $X$ (Poly*, $X^*$), which are valid (uncorrelated with $u_t$) by assumption (though assumed exogenous wrt $e$). Are they informative? Empirical question. Can use contemporaneous and lagged de-factored covariates as instruments for endogenous democracy index (here: Poly*, $X^*$; Poly*$_{-1}$, $X^*_{-1}$, Poly*$_{-2}$, $X^*_{-2}$).

This novel implementation is uniquely attractive for macro panel analysis for at least two reasons:

1. first, it conforms to the standard IV setup, where instruments are required to be informative, a property that is testable, and valid, which is less convincingly testable in the data — these conditions and related testing strategies are widely known and used in the applied empirical literature. Some of the most iconic papers in cross-country growth empirics — including the seminal work by Acemoglu et al. (2001) on the institutions-growth nexus — are based on (external) instrumentation but were crucially implemented in cross-section regressions. The IVMG, a panel estimator, brings to this a form of internal instrumentation as was practised using variable transformations and lags in the GMM estimators (Arellano & Bond (1991), AB; Blundell & Bond (1998) BB) which were wildly popular from the second half of the 1990s onwards (for seminal empirical applications see Caselli et al. 1996, Levine et al. 2000, Bond et al. 2001): while instruments for a cross-section regression can take advantage of ‘the long arm of history’ in selecting instruments (e.g. Spolaore & Wacziarg 2009), this form of instrumentation is difficult to operationalise in the panel setup, since exogenous variation would need to change exogenously over time. The IVMG can provide a credible panel instrumentation strategy where validity is achieved by construction.

Second, while the AB, BB and variations have now largely fallen out of favour, the NSYC IVMG addresses the most substantive concerns raised against these panel implementations of internal instrumentation: (i) the estimator can be implemented with homogeneous or heterogeneous slope coefficients; this deals with the well-known heterogeneity bias in dynamic panel estimation if common coefficients are imposed on a heterogeneous equilibrium relationship (Pesaran & Smith (1995)); (ii) the estimator is robust to variable nonstationarity including change points and integrated variable series, which raises serious challenges to the dynamic panel GMM estimators; (iii) the estimator overcomes the ubiquity of unobserved common factors, which thwarts attempts at a conventional panel IV strategy using lags or external time-varying instruments.

3.2 Difference-in-Difference Estimation with Single or Multiple Interventions

We now turn to methods which assume that there is a single or multiple thresholds in the relationship between democracy and growth. This adopts a certain value $\tau$ from an underlying continuous democracy index as a threshold and implicitly postulates that any effects below this cut-off are more or less negligible, and that any effects after this cut-off are more or less of identical magnitude.

10Note that this empirical strategy is distinct from the weakly exogenous instrumental variable CCE of Harding & Lamarche (2011) which does not focus on internal instruments.
The most recent contributions to the macro panel econometric literature have been able to build bridges to the literature on policy evaluation using difference-in-difference specifications (Gobillon and Magnac, 2016; Chan and Kwok, 2018) and the synthetic control methodology (Xu, 2017) most suited to the present empirical setup. What distinguishes these latest approaches from their canonical predecessors is the adoption of a multi-factor error structure in order to address three challenges to identification in these popular methods: (i) the presence of uncommon trends prior to the policy change evaluated, (ii) endogeneous selection into ‘treatment’, and (iii) the possibility that, following the policy change, treated and control samples are affected by common shocks, albeit with heterogeneous impact (e.g. the differential effect of the Global Financial Crisis across countries).

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

Underlying the approach is a treatment effect model with interactive fixed effects

\[ y_{it} = \Theta_i 1_{\{i \in I\}} 1_{\{t > T_{bi}\}} + y_{it}^0, \]  

where \( \Theta_i \) refers to the treatment effect on the treated unit \( i \), \( 1_{\{i \in I\}} \) is a dummy for the treatment group, and \( 1_{\{t > T_{bi}\}} \) is a dummy for the (heterogeneous) intervention date. The DGP for \( y \) is given by

\[ y_{it}^0 = \beta_i x_{it} + u_{it} \quad u_{it} = \lambda_i f_t + \epsilon_{it} \]
\[ x_{it} = \Lambda_i f_t + \nu_{it} \]
\[ \Rightarrow y_{it} =: \Theta_i 1_{\{i \in I\}} 1_{\{t > T_{bi}\}} + \mu_i f_t + \epsilon_{it}, \]

where \( \mu \) is some combination of the \( \lambda \) and \( \Lambda \) parameters, and \( f \) is a set of unobserved common factors (note that country and year FE are accommodated as special cases of this multifactor error structure). The expected value of \( \Theta_i \), \( E[\Theta_i | i \in I]\), is then the average treatment effect on the treated (ATET).

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

\[ y_{it} = \alpha_i + \beta_i \text{Dem}_{it} + \gamma_i X_{it} + \delta y_t + \delta_i X_{it} + \epsilon_{it}, \]

where \( y \) is per capita GDP (in logs and multiplied by 100), \( \text{Dem} \) is the democracy dummy, and
$X$ is a set of additional controls (we adopt gross investment share of GDP and trade openness).\textsuperscript{11} $\overline{y}$ and $\overline{X}$ are the cross-section averages of the observed variables but for those countries which never experienced democracy during the sample period (the control group). Following the insights from Pesaran (2006) and Chan and Kwok (2018) these cross-section averages capture the presence of uncommon and stochastic trends. Note that by construction there is no cross-section average for the democracy variable, since this is always zero in the control group from which these are computed.

Since we are concerned with establishing long-run relationships we also employ a dynamic variant of equation (8):

$$
y_{it} = \alpha_i + \theta_i Dem_{it} + \Gamma_i' X_{it} + \sum_{\ell=0}^{p-1} \omega_{i\ell}^D \Delta Dem_{i,t-\ell} + \sum_{\ell=0}^{p-1} \omega_{i\ell}^X \Delta X_{i,t-\ell} + \sum_{\ell=0}^{p-1} \delta_{i\ell}^y \overline{y}_{i,t-\ell} + \sum_{\ell=0}^{p-1} \delta_{i\ell}^X \overline{X}_{i,t-\ell} + \varepsilon_{it}, \tag{9}
$$

where the two terms involving sums in the first line capture the short-run effects, while $\theta_i$ and $\Gamma_i$ represent the long-run coefficients for the respective effects of democracy and additional controls on income per capita. The sums in the second line capture the multifactor error structure using cross-section averages, which like in the static model are constructed from those countries which never experienced democracy during the sample period. The use of this ‘CS-DL’ version of the Chan and Kwok (2018) approach is convenient since the long-run coefficient can be estimated in a single step rather than two like in an error-correction specification.\textsuperscript{12} Following suggestions in Chudik, et al. (2016) we adopt $p\overline{y} = 0$ and $p = p\overline{X} = int(T^{1/3})$, where $T$ is the time dimension of the panel. Our presentation below will focus on average estimates of $\hat{\theta}$ in the dynamic case; in line with the literature we adopt robust regression (Hamilton, 1992) to compute outlier-robust means. Inference for this robust ‘Mean Group’ estimate is based on standard errors computed non-parametrically, following Pesaran and Smith (1995).

Above we described the empirical implementation for a single threshold, but in principle (data permitting) this methodology can be extended to cover the analysis of $K$ thresholds.\textsuperscript{13} A static implementation of this multiple threshold difference-in-difference model is then:

$$
y_{it} = \alpha_i + \sum_{k=1}^{K} \beta_{ki} Dem_{kit} + \gamma_i' X_{it} + \sum_{k=1}^{K} \delta_{ki}^y \overline{y}_{kt} + \sum_{k=1}^{K} \delta_{ki}^X \overline{X}_{kt} + \varepsilon_{it}, \tag{10}
$$

where $Dem_{kit}$ is a set of treatment dummy, one each for the $k$th threshold, $X$ are the additional controls, $\overline{y}_{kt}$ is the cross-section average of $y$ for all those countries which never crossed the $k$th threshold — note that for $k > 1$ this subset of control countries does not include those which never crossed the threshold $k - 1$: take for instance $k = 2$, then $\overline{y}_{1t}$ is constructed from countries which never crossed threshold 1, while $\overline{y}_{2t}$ is the average for countries which did cross

\textsuperscript{11} The focus of this analysis is on $\beta_i$ and its cross-country average, not on $\gamma_i$.

\textsuperscript{12} In the ECM specification we obtain an estimate $\hat{\beta}_i$ for democracy and $\hat{\rho}_i$ for the lagged dependent variable (or $\sum_{\ell=1}^{p} \hat{\rho}_i \Delta Dem_{i,t-\ell}$ for $p$ lags), from which the long-run coefficient $\hat{\theta}_i = \hat{\beta}_i / - \hat{\rho}_i$ has to be computed. It is apparent from this that any finite sample bias in $\hat{\rho}_i$ will carry over to $\hat{\theta}_i$ (Chudik and Pesaran, 2015). The CS-DL obtains these estimates in a single step by adopting an alternative specification and avoids potential bias from dynamic misspecification.

\textsuperscript{13} We thank Marc Chan for comments on this idea.
threshold 1 but never crossed threshold 2. The construction of the $X_{kt}$ terms follows the same principle. Crucially, all treatment countries for which equation (10) is estimated have successfully overcome all $k$ thresholds, with the counterfactual constructed from countries which did not cross threshold $k,$ but for $k > 1$ did successfully cross a lower threshold of $k − 1.$

Despite the large sample size, even considering just two thresholds as we do below, already pushes the methodology to its limits (treatment or control samples below 15 countries): investigating additional thresholds, although methodologically feasible, is difficult in this cross-country context.

We take advantage of the rich V-Dem data in the application of this empirical methodology: we first present results pertaining to the three regime thresholds created by the V-Dem researchers, which take uncertainty in the V-Dem democracy index into account and hence represent non-arbitrary cut-offs. Then we study a large number of potential cut-offs (from around 0.1 to 0.9 in the V-Dem indices defined between 0 and 1), arbitrary and exogenously-imposed cut-offs, which will paint a richer picture of the patterns of growth effects by democracy threshold.

4 Data

4.1 Sources, Variable Transformations

Our empirical analysis relies on two main data sources, the Varieties of Democracy (V-Dem) data (Coppedge et al. 2017, version 8) for high-, medium- and low-level indicators for democracy, and World Bank World Development Indicator (WDI) data for income per capita (real GDP per capita in 2010 US$), along with additional controls, namely gross fixed capital formation and total trade (both in nominal values in local currency units and expressed in terms of nominal GDP in local currency units). For ease of interpretation in the threshold regressions we log-transform the income variable and multiply it with 100, so that threshold effects (shifts from ‘autocracy’ to ‘democracy’, using various definitions/cut-offs) can be interpreted as the long-term change in per capita income (in percent).

Much more detail on the individual V-Dem indices, their rationale, the coherent construction of higher-level from lower-level indices will be provided here in due course, for now refer to Boese (2019) for details. These aspects will be developed with reference to the Polity IV polity2 index, until recently the standard indicator for democracy.

4.2 Descriptive Statistics

This is work in progress. We want to highlight a number of aspects which have significant bearings on the interpretation of all our empirical results and will therefore be accounted for in our discussion of the results. First and foremost among these is the unbalanced nature of the panel used in this analysis. This is nothing new in the study of the democracy-growth nexus, but as the replications by Eberhardt (2019) for the ANRR paper highlight pooled empirical results may be driven by a (very) small number of influential country series: in that paper it is shown that restricting the sample by number of observations per country, i.e. excluding first all countries with just 5 observations, then those with 6, etc., leads to a steady decline in the
magnitude of the democracy coefficient and a rapid loss of statistical significance. Reducing the sample by excluding the final year 2010 yields the same outcome.\textsuperscript{14}

The top panel of Figure 2 provides information on when individual countries enter the sample,\textsuperscript{15} presenting the annual, quinquennial and decadal statistics (these start dates are identical for polyarchy and the liberal democracy indices). While for all three levels of aggregation the 1960-1969 period accounts for the highest share of country start dates, this amounts to at most 35% of the total sample of 159 countries (decadal aggregates). The second most significant wave of countries entering the sample takes place in the 1990s, which matches our priors about the collapse of the Soviet Union and the subsequent independence of many Warsaw Pact states; just under 25% of the sample enters during this decade. The two decades of the 1970s and 1980s together account for around 35% of the total sample, yet the hypothesised reasons for this delayed start vis-à-vis 1960 are very mixed – some advanced countries such as France or the Netherlands fall into this category. The remaining 5% of observations are for countries which entered the sample in the 2000s.

The plot in the lower panel of Figure 2 asks whether countries entering the sample at different points in time had very different secular patterns of polyarchy (using medians by entry group) – this analysis is by construction very crude, but nevertheless highlights that countries for which data started in the 1980s remained significantly below the full sample median (light-blue shading) throughout the remainder of the sample, whereas the 1960s and 1970s entrants stayed (more or less) above the full sample median.

The top panel of Figure 3 reproduces the sample median evolution of polyarchy (the solid pink line from panel (b) of Figure 2 (the evolution looks less dramatic because Figure 3 uses the full scale from 0 to 1) and further highlights the respective annual shares of the sample in which polyarchy is stable or improving, compared with the previous year (rendered in light-blue), and in which it is declining (dark-blue). This suggests that there were broadly two ‘eras’ over the past 50 years, an earlier period until 1990 when the share of countries in which polyarchy deteriorated steadily declined; and a later period from 1990 onwards in which this pattern was reversed and the share of countries in which polyarchy deteriorated steadily increased. The final five years from 2013 onwards represent the first consecutive years in the entire sample period during which polyarchy deteriorated in more countries than not.

The middle panel of Figure 3 plots every single one of the 159 country series for polyarchy, along with the superimposed median evolution (in blue). Although we know that the number of countries in the sample prior to 1990 amounted to at most 120, while after this point in time we reach the sample maximum of 159, it is nevertheless striking to see how entrenched country polyarchy series were in the former period: with a few countries at the top of the scale, the bulk of country series was below the median and only very few countries saw democratic change in the range between the median and the level of 0.8. This changed dramatically after 1990, with country series now covering the entire range of values between 0 and 0.9. Democratic change (for better or worse) to a very large extent happened after 1990, but any empirical

\textsuperscript{14}Excluding just three countries (less than 1% of the total number observations) renders the IV estimates of ANRR insignificant.

\textsuperscript{15}This analysis here is based on a sample for which we ran a panel regression including per capita GDP, polyarchy, gross fixed capital formation and trade openness. Given the same data restrictions in our regressions below this would seem more appropriate than using the V-Dem data in isolation.
analysis of the democratic dividend would do well to include the earlier period in the sample since this constitutes the largely autocratic counterfactual we need to assess the effect of democratic change.

The lower panel of the same figure provides details on the sample makeup (blue bars in the background, left scale), showing that the sample steadily increased from around 40 countries in 1960 to in excess of 125 by 1990. The pink shaded areas present the shares of countries by V-Dem regime over time. The earlier ‘era’ until 1990 appears more volatile, the decline in median polyarchy is reflected here by the increased and subsequently stable sample share for the lowest, ‘closed autocracy’, category. At the same time the late entry into the sample of a small number of advanced countries (e.g. France and the Netherlands) in 1970 compressed the share of the electoral democracy regime and expanded the share in the highest, ‘liberal democracy’, regime (which over the subsequent 47 years witnessed only a marginally increase). Once again everything changes around 1990, when the lowest regime saw its share decline from almost 40% to just over 10% in a matter of half a dozen years. The next higher two categories of electoral autocracy and electoral democracy also saw some relative changes during this period, increasing their respective shares, while this was only very marginally the case for the highest, ‘liberal democracy’, regime.

This descriptive analysis has highlighted a number of interesting aspects: first, it is important to account for the unbalanced nature of the panel, given that later entry may imply less time to establish the ‘treatment’ effect of democratisation or indeed the control effect of economic performance prior to significant democratic change; we achieve this by estimating the democracy-growth nexus at the country-level while controlling for the sample entry date and time period in subsequent analysis to aid interpretation of the results. Second, the failure of earlier studies on the democracy-growth nexus to identify a significant positive effect can be assigned to the relatively limited secular democratic change prior to the 1990s. Third, while median rates of democracy have experienced a step-change in 1990 and subsequently increased steadily yet marginally, the share of countries in which democracy deteriorated has steadily increased over time, and for the first time represented the majority of countries for consecutive years in 2013-2017.

5 Results

5.1 Linear Democracy-Growth Nexus

Table 1 presents empirical results when assuming a linear relationship between democracy and growth. All estimates presented are for models adopting the Pesaran (2006) cross-section average augmentation approach — adopting the PCA approach for unbalanced panels following Bai (2009) yields qualitatively similar results. The two high-level democracy indices employed are ‘polyarchy’ and ‘liberal democracy’, where the former captures the elements of participation and electoral competition (in analogy to the commonly used polity IV index), while the latter in addition captures constraints on executive. IVMG estimates yield positive significant long-run effects for either high-level concept of democracy, the magnitude of coefficients are roughly similar, though the diagnostic checks in the lower panel of the table suggest that a
much higher share of countries ‘passes’ the weak instrument tests in the specifications using more lags of the instruments in column [7] and [8]. For comparison, the results from a conventional Pesaran (2006) CCE Mean Group estimation approach in columns [1] and [2] are only marginally significant in case of the polyarchy and insignificant in case of the liberal democracy measure.16

These results capture robust means across all countries in the sample. But we also present the patterns arising from the IVMG with 2 lags (in [7] and [8]) with reference to the change in the respective index over the country-specific sample period in Figure 4: this representation plots a running line regression of the country-estimates on change in the respective index (i.e. these are not scatter plots), further controlling for the unbalanced nature of the panel (year in which country series starts and its length of time in the sample) and the standard deviation of the respective democracy index. In this and all subsequent results plots the grey bars indicate the distribution of country-estimates, while country-estimates which are statistically (in)significant are presented using filled (white) markers. In cases of both polyarchy and liberal democracy the long-run effects of democratic change on growth are small and insignificant when change is negative or negligible, and the effects rise and turn statistically significant with the magnitude of democratic change experienced. This evolution is arguably fairly linear in case of the polyarchy index and displays some diminishing returns in case of the liberal democracy index.

The message from these IV results is that not only is the average effect of either measure for democracy positive and significant, but they also suggest that more substantial democratic change is associated with a higher boost to long-run growth.

5.2 Non-linearities and Thresholds

We begin our analysis of single growth thresholds of democracy by adopting the pre-defined V-Dem ‘regime’ variable — as was detailed above V-Dem researchers spend considerable time analysing the evolution of each country in the sample rather than adopting crude regime cut-offs (as we will apply below). Table 2 shows robust mean results for static and dynamic single threshold effects for the thresholds indicated in panels (A) and (B), respectively. In all cases the treatment and control samples feature at the very least 27 countries, hence we need not worry about small sample results.

The magnitude of the coefficients here translate into a meaningful economic magnitude, namely the long-run change in per capita GDP as a result of a permanent transition from below to above the democracy threshold studied. These estimates suggest that a shift from closed autocracy to electoral autocracy or a higher regime is associated with a permanent increase of between 2.2 and 4.9% of GDP pc – note that for the preferred specification in the columns marked [b], which include covariates for trade openness and investment in the model, both estimates are statistically insignificant. Overcoming the threshold between the two autocracy and two democracy regimes (in column [2]) yields long-run estimates between 3.7 and 6%, while moving across the electoral-liberal democracy regime divide (from any regime below) leads to between 10 and 16% higher per capita income in the long run. These results indicate a relative ranking (higher regimes yield higher returns), but also clearly mark out the difference

16Two-way fixed effects regressions (results not shown) yield insignificant results for both.
between the top regime (liberal democracy) and the rest.

The above estimates are for V-Dem regimes, in the following we discuss the findings for continuous polyarchy and liberal democracy indices, where we (exogenously) vary the autocracy/democracy threshold: in Figure 5 each of the markers represents a Diff-in-Diff ATET estimate for a sample of treated countries which overcame the democracy threshold (as highlighted on the $x$-axis) against a control group of countries which did not, where a filled marker indicates statistical significance. The gray bars indicate the number of countries in the ‘treatment’ group. Starting with the patterns arising from the liberal democracy results, although there are some minor idiosyncracies for very low threshold values, the boost from a higher value of the index declines near-monotonically to turn insignificant around a value of .7 — note that beyond this value the treatment samples contain typically fewer than 20 countries, hence the negative significant results are not based on a conventionally acceptable sample size. For polyarchy the patterns are somewhat more complicated, following two consecutive inverted U-shapes. As the horizontal dashed lines repeat the dynamic CCE Diff-in-Diff estimate from Panel (B) of Table 2, column [2][b] (moving from autocracy to democracy), and it is interesting to see that these represent rough averages across the shifting threshold estimates presented.

Analysis along the lines of the IVMG estimates using running line regressions will be added in due course.

While single threshold analysis with exogenously imposed thresholds can reveal some unusual patterns of the growth effect of democracy, it nevertheless does not allow for individual countries to undergo a more complex nonlinear relationship as they improve their level of democracy over time. Table 3 addresses this concern by studying the difference-in-difference effects of countries overcoming two V-Dem regime thresholds during the sample period. As the figures in the lower panels indicate, even though our sample amounts to over 150 countries, the sizes of treatment and two control samples are not always acceptable, and as low as 12 or 13: precious few countries either stayed at a very low level of democracy or successfully moved across consecutive regime thresholds – an insight which is easily missed when studying pooled regressions with multiple thresholds or higher polynomials of the variable of interest. Since as suggested the sample sizes are typically not very large the following findings have to be taken with a pinch of salt.

Although the mean estimates for consecutive threshold effects are never statistically significantly different from each other, there nevertheless is a clear pattern whereby higher thresholds lead to higher growth effects than previously-overcome lower thresholds. This provides evidence against the presence of substantial diminishing returns to democracy, at least within the confines of the regime definition adopted here: from a growth standpoint, more democracy is better.

Analysis along the lines of the IVMG estimates using running line regressions will be added in due course.
5.3 Channels: Lower Level Democracy Indices

This part of the paper is as yet incomplete. The plan is to follow down the constituent elements of each high-level index to identify the patterns of results for low(er)-level indices when adopting identical methods and specifications as above. So far we have extended this analysis to the constituents of the polyarchy variable, for which the running line regression results for IVMG estimates are presented in Figure 6: polyarchy (via additive and multiplicative sub-indices) is based on five lower-level indices, capturing the freedom of expression (index), freedom of association (index), clean elections (index), elected officials (index) and the share of population with suffrage (not an index). For elected officials 70% of sample observations are at 1, with 19% at 0 during the 1960-2017 period, leading to very limited time series variation for each country. Similarly, if more extreme, universal suffrage is the case in 93% of all sample observations, with 2% at zero and the remainder inbetween – neither of these facets of democracy can play an important part in the post-WWII growth process since the vast majority of countries are already at the top value. For the remaining three indices we carry out the IVMG estimation and subject the resulting country specific coefficients to a running line regression against the change in the respective index and the additional controls related to the unbalancedness of the panel. However, in contrast to the equivalent plots in Figure 4 in the graph for low-level index \( k \) (e.g. freedom of expression) we further control for (i) the change in the index for the other two indices \(-k\) (freedom of association, and clean elections), (ii) the country-specific IV long-run estimates for the other two indices \(-k\), and (iii) the standard deviation of all three low-level indices. The results in Figure 6 can thus be interpreted as ceteris paribus with regard to the rival low-level democracy indices.

This strategy yields some very interesting patterns across the three lower-level indicators: while the relationships between increases in freedom of expression and freedom of association on the one hand and the conditional effect of these respective indices on long-run growth on the other is generally flat or negative, positive changes in the clean election index are associated with a rising growth effect. If the results for polyarchy (a positive association between change in polyarchy and the long-run growth effect) were driven by either of the former aspects of democratic participation and competition, then we would expect the lower-level indices to display positive associations as well. This is only the case for the clean election index, where the relationship is near-linear monotonic and highlights a negative effect for some countries in which the clean election index deteriorated over time.

Results for other lower-level indices as well as alternative specifications (threshold models) will be added in due course.

6 Concluding Remarks

<to be added>
References


17


## Tables and Figures

### Table 1: Dynamic Linear Models of Democracy and Growth

<table>
<thead>
<tr>
<th>Estimator</th>
<th>CCEMG</th>
<th>IVMG</th>
<th>IVMG</th>
<th>IVMG</th>
<th>IVMG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excl. instruments</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Instrument set</td>
<td>$X^*$</td>
<td>$X^<em>, X^</em>_{-1}$</td>
<td>$X^<em>, X^</em><em>{-1}, X^*</em>{-2}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ln(GDPpc) $(t - 1)$)</td>
<td>0.791</td>
<td>0.781</td>
<td>0.898</td>
<td>0.896</td>
<td>0.894</td>
</tr>
<tr>
<td>Polyarchy</td>
<td>1.470</td>
<td>1.770</td>
<td>2.629</td>
<td>1.963</td>
<td></td>
</tr>
<tr>
<td>Liberal Democracy</td>
<td>0.266</td>
<td>2.275</td>
<td>2.766</td>
<td>2.366</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Observations</td>
<td>6,470</td>
<td>6,449</td>
<td>6,394</td>
<td>6,373</td>
<td>6,391</td>
</tr>
<tr>
<td>Countries</td>
<td>159</td>
<td>159</td>
<td>159</td>
<td>159</td>
<td>159</td>
</tr>
<tr>
<td>LR Effect Polyarchy</td>
<td>7.030</td>
<td>15.279</td>
<td>21.791</td>
<td>16.809</td>
<td></td>
</tr>
<tr>
<td>LR Effect Liberal Democracy</td>
<td>1.210</td>
<td>20.346</td>
<td>23.427</td>
<td>20.458</td>
<td></td>
</tr>
<tr>
<td>Countries</td>
<td>159</td>
<td>159</td>
<td>154</td>
<td>151</td>
<td>156</td>
</tr>
<tr>
<td>Share 1st-Stage F &gt; 10</td>
<td>0.82</td>
<td>0.83</td>
<td>0.71</td>
<td>0.73</td>
<td>0.70</td>
</tr>
<tr>
<td>Share Hansen J p&gt;0.1</td>
<td>0.69</td>
<td>0.68</td>
<td>0.76</td>
<td>0.76</td>
<td>0.83</td>
</tr>
<tr>
<td>Share K-P LM p&lt;0.1</td>
<td>0.90</td>
<td>0.84</td>
<td>0.71</td>
<td>0.74</td>
<td>0.60</td>
</tr>
<tr>
<td>Share S-W LM p&gt;0.1</td>
<td>0.44</td>
<td>0.45</td>
<td>0.47</td>
<td>0.51</td>
<td>0.60</td>
</tr>
<tr>
<td>Observations</td>
<td>6,470</td>
<td>6,449</td>
<td>6,394</td>
<td>6,373</td>
<td>6,391</td>
</tr>
<tr>
<td>Countries</td>
<td>159</td>
<td>159</td>
<td>154</td>
<td>151</td>
<td>156</td>
</tr>
<tr>
<td>LR Effect Polyarchy</td>
<td>12.10</td>
<td>20.346</td>
<td>23.427</td>
<td>20.458</td>
<td></td>
</tr>
<tr>
<td>LR Effect Liberal Democracy</td>
<td>1.210</td>
<td>20.346</td>
<td>23.427</td>
<td>20.458</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** The table reports robust mean estimates from dynamic CCEMG (columns [1] and [2]) and the NSYC IVMG (all other columns). In the IV models all variables have been within- and between-transformed, and defactored variables $X^*$ (in [5]-[9] their first lags $X^*_{-1}$, and first and second lags $X^*_{-1}, X^*_{-2}$) are used as instruments for the endogenous democracy variable (the polyarchy or liberal democracy index). De-factoring here is carried out using cross-section averages for all regressors (excluding the lagged dependent variable); results are qualitatively identical if we use de-factoring via principal components instead (results available on request). The upper part of the table reports the estimates from the ARDL(1,0) specification, the implied ‘long-run estimates’ are reported in the lower part, computed via the Delta method and using the weights employed in the computation of the robust means across countries – since these weights attach zero importance to some outlier estimates the sample size is reduced for these long-run estimates in all IVMG cases. IV-related diagnostic tests are available for each individual country regression, and the last four rows of the table report the share of sample countries which have desirable properties on the basis of these tests. Statistical significance at the 10%, 5% and 1% level is indicated as *, **, and *** respectively.
Table 2: Static and Dynamic Single Regime Threshold Models of Democracy and Growth

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[a]</td>
<td>[b]</td>
<td>[a]</td>
</tr>
<tr>
<td>Covariates</td>
<td>none</td>
<td>open, gfcf</td>
<td>none</td>
</tr>
</tbody>
</table>

(A) Static Threshold Effects:

- closed autocracy → electoral autocracy/higher: 3.337 [1.645]**
- closed/electoral autocracy → electoral/liberal democracy: 6.034 [1.616]**
- electoral democracy/lower → liberal democracy: 10.582 [3.148]**

Treatment Sample:

| Countries | 85 | 85 | 81 | 81 | 27 | 27 |
| Total observations | 3,621 | 3,613 | 3,399 | 3,398 | 1,183 | 1,183 |
| dto. per country | 42.6 | 42.5 | 42.0 | 42.0 | 43.8 | 43.8 |

(B) Dynamic Threshold Effects:

- closed autocracy → electoral autocracy/higher: 4.919 [2.253]**
- closed/electoral autocracy → electoral/liberal democracy: 4.937 [2.466]**
- electoral democracy/lower → liberal democracy: 16.188 [3.305]**

Treatment Sample:

| Countries | 75 | 73 | 71 | 70 | 25 | 24 |
| Total observations | 3,174 | 3,131 | 2,974 | 2,952 | 1,063 | 1,041 |
| dto. per country | 42.3 | 42.9 | 41.9 | 42.2 | 42.5 | 43.4 |

Control Sample:

| Countries | 42 | 42 | 59 | 59 | 132 | 132 |
| Total observations | 970 | 970 | 1,953 | 1,953 | 5,389 | 5,389 |

Notes: The table reports robust mean estimates for single threshold models based on a static and dynamic version of the Chan and Kwok (2018) diff-in-diff specification in Panels A and B, respectively. The estimator can be implemented with or without control variables (here: trade openness and gross fixed capital formation, both in logs). We adopt V-Dem regimes to create thresholds: 0 – closed autocracy, 1 – electoral autocracy, 2 – electoral democracy, 3 – liberal democracy. The thresholds are thus between 0 and 1, 1 and 2, and 2 and 3. Countries which always were above the threshold are excluded from the analysis. The estimates are (long-run) ATET effects (based on a CS-DL model in the dynamic case) and represent the percentage increase in per capita GDP over the control group at each respective threshold. The lower panels report the number of countries and observations which make up treatment and control samples — the treatment samples in the static models are marginally larger. Statistical significance at the 10%, 5% and 1% level is indicated as *, **, and *** respectively.
### Table 3: Static Double Regime Threshold Models of Democracy and Growth

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Regime &gt;= 1</td>
<td>Regime &gt;= 2</td>
<td>none</td>
<td>2.660</td>
<td>3.320</td>
<td>5.761</td>
<td>4.553</td>
</tr>
<tr>
<td></td>
<td></td>
<td>open, gfcf</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threshold Effects:</td>
<td></td>
<td>closed autocracy →</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>electoral autocracy</td>
<td>[1.680]</td>
<td>[1.588]**</td>
<td>[5.350]</td>
<td>[3.402]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>electoral autocracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>electoral/liberal democracy</td>
<td>[1.975]**</td>
<td>[1.843]**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>closed/electoral autocracy → electoral democracy</td>
<td></td>
<td>5.634</td>
<td>4.158</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>electoral autocracy</td>
<td>[2.692]*</td>
<td>[1.399]*****</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>liberal democracy</td>
<td>[2.017]****</td>
<td>[1.993]****</td>
<td></td>
<td></td>
</tr>
<tr>
<td>electoral autocracy/democracy → liberal democracy</td>
<td></td>
<td>8.021</td>
<td>8.664</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>democratic</td>
<td>[2.865]**</td>
<td>[3.200]****</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Treatment Sample:**

- **Countries:** 54, 15, 12
- **Total observations:** 2,456, 681, 577
- **Observations below 1st Threshold:** 609, 213, 129
- **dto. per country:** 11.3, 14.2, 10.8
- **Observations btw 1st & 2nd Threshold:** 825, 183, 205
- **dto. per country:** 15.3, 12.2, 17.1
- **Observations above 2nd Threshold:** 1,022, 285, 243
- **dto. per country:** 18.9, 19.0, 20.3

**Control Sample:**

- **Countries 1st Threshold:** 13, 46, 13
- **dto. observations:** 506, 1,739, 506
- **Countries 2nd Threshold:** 33, 67, 100
- **dto. observations:** 1,233, 2,752, 3,985

**Notes:** The table reports robust mean estimates for double threshold models based on a static Chan and Kwok (2018) diff-in-diff specification with two thresholds. This estimator can be implemented with or without control variables (here: trade openness and gross fixed capital formation, both in logs). We adopt V-Dem regimes to create thresholds: 0 – closed autocracy, 1 – electoral autocracy, 2 – electoral democracy, 3 – liberal democracy. The thresholds are thus between 0 and 1, 1 and 2, and 2 and 3, and we present results for all double-threshold combinations of these. Crucially, all estimates are for the ‘treatment’ sample of countries which crossed both thresholds 1 and 2, and the control samples of countries which never cross threshold 1 and never crossed threshold 2, respectively. Countries which always were above the higher threshold are excluded from the analysis. The estimates are ATET effects and represent the percentage increase in per capita GDP over the control group at each respective threshold, e.g. in column [1], specification with covariates included, countries which moved from closed to electoral autocracy and eventually to electoral democracy (or beyond) had 3.3% and 4.5% higher per capita GDP than control group countries which always stayed below the respective thresholds. The lower panels report the number of countries and observations which make up treatment and control samples. Statistical significance at the 10%, 5% and 1% level is indicated as *, **, and ***, respectively.
Figure 1: How political institutions affect economic growth

Notes: The figure needs revision as it presently does not reflect the theoretical ‘split’ we advocate (building on the work by Douglass North and others). Further we will add the direct mapping to specific high-, mid- and low-level indices provided by V-Dem in the same diagram.
Figure 2: Polyarchy — timings and evolution

(a) Date (annual, quinquennial, decadal) each sample country first appeared in the dataset

(b) Evolution by decade of appearing in the sample

Notes: The figure presents the timings when countries appear in the sample (upper panel), as well as the median of polyarchy, separating out samples with data starting in the 1960s, 1970s, 1980s, and 1990s.
Figure 3: Evolution of Democracy – 159 countries (1960-2017)

Notes: The figure presents (a) the median evolution of the V-Dem polyarchy index in 159 countries over the 1960-2017 time horizon (solid line) along with information about the share of the sample for which the respective polyarchy index has improved/remained stable, or declined, (b) country series for polyarchy, and (c) the numbers of countries in the sample (left scale) along with the evolution of the sample share of the four regimes from the V-Dem Regimes in the World dataset.
Figure 4: Change in Democracy (high-level indices) and IV Estimates

(a) Analysing the ‘polyarchy’ index

(b) Analysing the ‘liberal democracy’ index

Notes: We present the country-specific IV estimates for the specification with $X^*, X_{-1}^*, X_{-2}^*$ (average IVMG results in Table 1) in an alternative way, taking into account the change experienced in the democracy index over time (x-axis). The estimates presented are from running line regressions, which linearly condition on (a) the year in which the country series started, (b) the standard deviation of the democracy index during the sample period, and (c) the length of time (in years) the country spent in-sample. The grey bars indicate the distribution of the 159 country coefficients, the filled (white) markers indicate statistical (in)significance at the 10% level. Note that the markers are not a scatter of the individual estimates, they are included here to indicate statistical significance. The running line is constructed adopting KNN local regressions.
Notes: The figure presents the running line estimates for dynamic Difference-in-Difference estimates Chan & Kwok (2018) as the definition for a democracy ‘threshold’ shifts (along the x-axis) from 0.1 to 0.9: these graphs are constructed from the country-specific long-run ATET estimates for democratisation, conditioning on country fixed effects — this accounts for the fact that the sample makeup is entirely different as we move from left to right. The estimates presented show the average effect of democratisation on per capita GDP (in percent) accounting for the difference in the treatment sample across different thresholds applied. For comparison we also present an average long-run democracy effect when adopting the V-Dem ‘regime’ indicators (we group scores 0 and 1 as ‘autocracies’, scores 2 and 3 as ‘democracies’): this is the dashed horizontal line in blue for polyarchy and in dark pink for liberal democracy ($\hat{\beta} = 4.39, t = 2.13$). The country-specific long-run ATET estimates are in turn based on a dynamic CS-DL version of the Chan & Kwok (2018) difference-in-difference estimator – treated countries are those which crossed the threshold, control countries are those which always stayed below the threshold. The grey bars (left scale) indicate the sample size of the treatment group (number of countries).
Notes: The figure presents the running line fits for IV estimates of three lower-level democracy indices — Freedom of Expression (top panel), Freedom of Association (middle panel), and Free and Fair Elections (bottom panel) — which are plotted against the change in the respective democracy index (end year minus start year). The IV estimates are from separate regressions with up to two lags of instruments (average results presented in Table X), but the running line results further condition on: (a) the year in which the country series started, (b) the length of time (in years) the country spent in-sample, (c) the standard deviations of all three democracy indices during the sample period, (d) the change in the other two low-level democracy indices, and (e) the coefficients for the other two low-level democracy indices. In the bottom panel we exclude the observation for West Bank and Gaza as an outlier ($< -0.8$).