V-Dem
INSTITUTE

Stairways to Denmark:
Does the Sequence of State-building and Democratization Matter for Economic Development?

Haakon Gjerlow, Carl Henrik Knutsen, Tore Wig, and Matthew Charles Wilson

August 2018

Working Paper

UNIVERSITY OF GOTHENBURG
DEPT OF POLITICAL SCIENCE

Stairways to Denmark:
Does the sequence of state-building and democratization matter for economic development?*

Haakon Gjerløw[1], Carl Henrik Knutsen[1], Tore Wig[1], and
Matthew Charles Wilson[2]

[1]Department of Political Science, University of Oslo
[2]Department of Political Science, West Virginia University

August 13, 2018

Abstract

Building effective state institutions before introducing democracy is widely presumed to improve different development outcomes. We discuss the assumptions that this prominent 'stateness-first' argument rests upon and how extant studies fail to correctly specify the counter-factual conditions required to test the argument. In extension, we subject the argument to three sets of tests focused on economic development as the outcome, leveraging new measures of democracy and state institutional features for almost 180 polities with time series extending back to 1789. First, we run standard panel regressions with interactions between state capacity and democracy. Second, we employ coarsened exact matching, specifying and testing different relevant counter-factuals embedded in the stateness-first argument. Finally, we employ sequencing methods to identify historically common sequences of institutional change, and use these sequences as growth predictors. We do not find any evidence supporting the stateness-first argument in either of these tests.

1 Introduction

Several studies point to how 'good institutions', such as democracy (Acemoglu et al., 2014) or a rule-following and capable state bureaucracy (Evans and Rauch, 1999), enhance economic development. While the presence or absence of particular institutions remains the core focus in the literature studying institutions and development, many prominent arguments focus on the *sequence* in which particular institutions are introduced. Perhaps the most widely regarded type of 'sequencing explanation' holds that building effective state institutions *before* introducing democracy has beneficial effects on a variety of outcomes (see, for instance, Huntington, 1968; Shefter, 1993; Zakaria, 2003; Mansfield and Snyder, 1995, 2007; Fukuyama, 2007, 2014a; D'Arcy and Nistotskaya, 2017). According to this view, the state-before-democracy pathway leads to 'Denmark'—the metaphor for an economically prosperous and politically stable country used by Fukuyama (2014a). Conversely, introducing democratic institutions, such as competitive elections and universal franchise, before effective and capable state institutions are in place, is often considered a path to political instability, violence, clientelism, and a stagnant economy. While this 'stateness-first' argument is plausible, it faces three major hurdles before it can be accepted as firm knowledge.

First, the argument is often presented in an insufficiently precise manner. In particular, studies often fail to outline the exact counter-factual institutional configurations and development patterns they have in mind when arguing for the benefits of building state capacity before democratization. This, in turn, presents difficulties for interpreting the evidence in favor of the conjecture. We propose that while extant case studies (e.g., Fukuyama, 2014a) and cross-national regressions (e.g., D'Arcy and Nistotskaya, 2017) may provide evidence that strong state institutions affect development outcomes, they do not provide evidence directly pertaining to the more complex sequencing explanation. Second, the stateness-first argument rests on several strong assumptions for which there are plausible counter-arguments.

1

Third, and perhaps most importantly, the stateness-first argument has not been systematically analyzed through the same types of stringent testing as many other propositions on the determinants of development. There are good reasons for why extant studies have not pursued such tests, notably the lack of extensive time-series data on relevant institutional features. But it remains unclear whether the stateness-first argument is valid, even if many scholars, policy makers, and others find it to be plausible.

This paper examines the relationship between stateness-first sequences of institutional development and economic growth, one of the key outcomes that stateness-first sequences are purported to explain. The paper makes two contributions. The first one is theoretical, as we specify and evaluate critical assumptions and counter-arguments, and elaborate on the specific observable implications following from the argument. While our empirical tests focus on economic growth—we report more preliminary results for alternative outcomes of interest in Appendix D—these theoretical insights draw on and speak to studies on other proposed consequences of stateness-first sequences, including regime stability, civil conflict, and broader human development.

We also make a second, empirical contribution. We systematically evaluate different assumptions and implications following from the stateness-first argument, focusing on economic growth as the dependent variable. To this end, we use data from the Varieties of Democracy (V-Dem) dataset (Coppedge et al., 2017), including the new Historical V-Dem data (Knutsen et al., 2017), which extends relevant V-Dem indicators back in time to 1789. These extensive time series allow us to track institutional developments throughout the course of 'modern history', covering important periods of state building and democratization in different regions of the world. In order to test different assumptions and implications of the argument, we employ a multi-approach strategy to testing. Specifically, we run panel regressions estimating how state capacity and democracy interact in affecting development. Next, we test several matching models that compare, for instance, high- and low-state capacity countries

2

that undergo democratization episodes. Finally, we employ sequencing analysis, classifying different historical patterns to investigate how common institutional sequences predict growth. Across all specifications, we find scant evidence to support the stateness-first argument. Some specifications even suggest that democracy-before-state sequences *enhance* economic development relative to stateness-first sequences.

We start out by introducing sequencing theories of development and the stateness-first argument in particular. Next, we critically discuss the argument's core assumptions and plausible counter-arguments. We then elaborate on the implications that follow from the argument and how these might be tested . After outlining the data, we introduce and present results from, respectively, panel regressions, matching models, and analyses on institutional sequences. In the conclusion, we summarize our findings and discuss how they may inform— or at least temper—normative debates and prescriptive policy advice on the (un)desirability of promoting democracy in countries with weak state institutions. We also highlight avenues for future research, emphasizing that similar, careful studies are needed in order to systematically assess whether stateness-first sequences have different effects on alternative outcomes of interest such as regime stability, civil war, and human development.

2 Critically evaluating the 'stateness-first' argument

2.1 Putting state capacity before democracy

Country rankings for important economic and political development outcomes, including democracy, bureaucratic quality, income level, and human development, often identify the same sets of countries as high- and low achievers. Why is it that some countries, such as Denmark, have followed favorable development trajectories, whereas others have not? Multiple deeper determinants of both political and economic development have been suggested, notably geography, climate, and natural resource endowment (Diamond, 1997; Sachs, 2005; Pomeranz, 2000), culture and ideology (Weber, 2002; Landes, 1998), as well as demographic

and genetic factors (Galor, 2012; Spolaore, 2013). Yet, one predominant view highlights the role of institutions in influencing development outcomes (North, 1990; Acemoglu, 2001; Acemoglu and Robinson, 2012; Rodrik et al., 2004). Despite the prominence of the institutionalist view, there remains no real consensus on exactly *which* institutions matter for spurring development, with suggestions ranging from institutions that guard against the concentration of political power and protect property rights (Acemoglu, 2001; De Long and Schleifer, 1993; North, 1990) to the establishment of a capable, impartial, and rule-following bureaucracy (Evans and Rauch, 1999; Fukuyama, 2014a; Rothstein, 2011).

An elaborate version of the institutionalist view on development highlights not only the importance of specific institutions, but also the order in which they are introduced. We term such explanations *sequencing explanations* of development. Sequencing explanations of successful democratization, for example, propose that certain institutional (and other) preconditions need to be in place prior to the establishment of democracy. Following decolonization processes in Africa and Asia after WWII, observations of newly democratized countries in which elected leaders abused their powers, or in which competing interests degenerated into conflicts, led prominent scholars to argue that the successful implementation of democracy depends on the relative timing of specific events and reforms (e.g., Huntington, 1968; Dahl, 1971). One notion is that the influence of interest groups under democracy can be detrimental to economic reform and modernization, such that popular pressures need to be restrained and channeled in the wake of a modernizing economy (Huntington, 1968; Wintrobe, 1998). Another variant of sequencing theory focuses on the need for establishing 'liberal' institutional features early on, citing the introduction of civil liberties prior to suffrage expansion as a condition that supports democratic deepening and prevents the emergence of illiberal democracies (Marshall and Bottomore, 1949; Møller and Skaaning, 2013; Zakaria, 2003).

Yet others have argued that a 'postponed transition' to democracy, after rule of law or a

rule-following, high-capacity bureaucratic apparatus have been achieved, will make countries less likely to experience conflict and violence, and better able to mitigate patronage or other bad outcomes associated with a rushed democratic transition (see, e.g., D'Arcy and Nistotskaya, 2017; Mansfield and Snyder, 2005, 2007; Shefter, 1993). This class of arguments, which we term 'stateness-first' arguments, generally holds that some aspects of the state must develop before the introduction of mass politics—including contested multi-party elections with extensive franchise—if democracy is to succeed and produce benevolent outcomes. While several scholars have rejected such notions of institutional sequencing, questioning the factual basis of the proposed risks of 'premature elections' and 'out-of-sequence' changes (Berman, 2007; Carothers, 2007; Hobson, 2012), this argument remains widely popular among scholars and in policy circles.

A recent and well-argued formulation of the stateness-first argument was proposed by Fukuyama (2012; 2014a), who maintains that a strong state—defined by *state capacity* and *rule of law*—is necessary to equip democracies for success.[1] Echoing the argument by Shefter (1993), Fukuyama (2014b) sums up the core logic of this benevolent sequence as follows: "when a modern, Weberian state has coalesced prior to the expansion of the democratic franchise, it tends to resist colonization by patronage-dispensing politicians because it develops around it a protective 'absolutist coalition'" (p. 1333). Fukuyama invokes the concepts of patronage—the reciprocal exchange of favors between two individuals of different status and power—and clientelism—patronage on a larger scale—treating clientelism as a consequence of unfettered democracy. This is based on the premise that only democratic politics

---

[1]State capacity, in turn, is a two-dimensional concept pertaining, respectively, to the *scope* of state functions and the *strength* of state institutions. Holding a monopoly of power, and having the strength to effectively enforce it, represent minimum requirements for any central authority (Fukuyama, 2014a, p. 54–59). For rule of law to be fully achieved, laws should be binding on even the most powerful political actors in society, without politicians being able to change them whenever it suits them (Fukuyama, 2014a, p. 11, 24).

requires the mobilization of large masses of voters (Fukuyama, 2014a, p. 86). If democracy is introduced *before* bars against clientelism—such as merit-based recruitment to, and impartial and rule-following behavior by, the bureaucracy—are in place, democratic leaders will offer government positions for political support. Thus, "[c]lientelism emerges in young democracies precisely because the state and its resources constitute useful piggy banks for democratic politicians seeking to mobilize supporters." (Fukuyama, 2014a, p. 532).

Where democracy meshes with clientelism, the result is a deterioration of governance outcomes and the pursuit of policies that may ultimately hamper economic development. The confluence of democracy and clientelism has the effect of directing government activities toward serving the private interests of a corrupt few, reducing the quality of governance by further eroding the capacity of the state, and eventually turning clientelism and elite entrenchment into self-reinforcing processes. The practice of providing rents through patron-client relationships in exchange for political support is widely regarded as a highly inefficient form of redistribution (e.g., Robinson and Verdier, 2013). Clientelism can negatively affect economic development by reducing productivity growth via inefficient allocation of resources and increased costs and uncertainty of entrepreneurial activities (e.g., North, 1990; Acemoglu, 2008). It can also diminish investment in physical capital, as investors become wary of increased expected costs and investment risks that follow from corrupt government and 'bad polices' (e.g., Knack and Keefer, 1995). Moreover, proponents of the stateness-first argument underscore that 'democratizing backwards' reduce the quality of public services such as education and health care (e.g., D'Arcy and Nistotskaya, 2017). This, in turn, hurts the accumulation of human capital, another key immediate determinant of growth (e.g., Mankiw et al., 1992). Hence, democracy-first sequences should negatively influence growth through different channels.

In sum, countries with weak state institutions at the time of democratization are anticipated to experience worse governance outcomes and poorer economic performance, and

they are anticipated to be locked into such situations over an extended period of time. In contrast, countries that followed a stateness-first sequence are more likely to be put on a development path that eventually make them resemble 'Denmark', Fukuyama's metaphor for a democratic, secure, well-governed, and prosperous country. In such a country "all three sets of political institutions [are] in perfect balance: a competent state, strong rule of law, and democratic accountability" (Fukuyama, 2014a, p. 25). Getting to Denmark, therefore, depends on a favorable historical pathway of institutional development whereby strong state institutions—including a rule-following bureaucracy with meritocratic recruitment practices—appear *before* democratization.

2.2 Assumptions and counterfactuals

Despite the plausibility of the stateness-first argument, the argument relies on a set of strong assumptions. We critically discuss three of these assumptions as well as issues of specifying the appropriate counterfactual conditions for evaluating the argument.

State building under democracy: One key assumption of the stateness-first argument is that state-building is relatively hard to do in democracies, especially when starting out in a low-capacity setting. Yet, Mazzuca and Munck (2014) note that state-building and processes of democratization have historically co-evolved in many instances, and that early democratization may even ease (nation- and) state-building. One proposed reason is that democratization provides the state with much needed legitimacy in the eyes of contending political elites and citizens. Furthermore, several scholars contend that democratically elected leaders face stronger incentives to provide public goods and services (e.g., Lake and Baum, 2001; Bueno de Mesquita et al., 2003). Providing public services to prospective voters in an efficient manner may increase re-election chances, thus incentivizing democratic leaders to build a competent bureaucratic apparatus for delivering such services. Public goods provision also requires taxation, which, in turn, requires well-functioning bureaucratic support functions. This creates another, albeit indirect, incentive for democratic politicians to build

7

state capacity. A handful of large-n studies have tested for a relationship between democracy and state capacity, mostly reporting a positive association (Adzera et al., 2003; Bäck and Hadenius, 2008; Carbone and Memoli, 2015; Wang and Xu, 2018). While these findings run counter to the assumption undergirding the stateness-first argument, we note that some recent studies have added qualifications, suggesting that democracy may only enhance capacity in rich-country contexts (Charron and Lapuente, 2010) or that competitive elections enhance capacity whereas suffrage expansions may have the opposite effect (Andersen and Cornell, 2018).

State building under autocracy: A second, and related, assumption is that autocratic leaders are both capable and willing to develop strong and capable state institutions. The above-mentioned empirical findings call this notion into question. On the theoretical side—and notwithstanding the question of whether autocratic regimes have the requisite knowledge and capacity to engage in such institution building—one important question is: how strong are the incentives of most autocrats to invest in state capacity? Indeed, many theoretical contributions highlight that autocratic regimes often have strong incentives to *under*-invest in building effective state institutions (e.g., Besley and Persson, 2009, 2010; Charron and Lapuente, 2010). What is more, autocrats sometimes have direct incentives to 'build down' the quality and capacity of state institutions to enhance personal control over access to public resources (Acemoglu et al., 2005; Knutsen, 2013), as indicated by the 'informalization' of politics under African strongman-rule in the post-colonial period (see, e.g., Chabal and Daloz, 1999). While there may be situations where autocrats (and democratic leaders) face stronger incentives to build state capacity—for instance in the presence of an external threat (e.g., Tilly, 1990; Fukuyama, 2014a)—we surmise that most autocratic regimes do not face strong incentives to do so.

Democratic transitions in consolidated autocracies: A third, and also often implicit, assumption is that autocratic governments are willing to yield power and oversee

transitions to democracy—or at least, that they are more easily pressured into doing so—after the initial building of state institutions. This is required for these stateness-first regimes to eventually end up like 'Denmark'. There is, however, little evidence that autocratic regimes more easily yield power after they have built capable and effective state institutions. In fact, there is evidence to the contrary. Andersen et al. (2014), for example, find that the expansion of certain types of state capacity, notably fiscal capacity and a firm monopoly on violence, significantly prolong the reign of autocratic regimes. Similarly, recent studies have found evidence that state capacity moderates the effect of elections on autocratic regime breakdown (Seeberg, 2015; van Ham and Seim, 2017). Autocratic regimes presiding over a state that is able to effectively extract resources (that can be used for co-optation) and repress threats are better able to bolster their own hold on power. This means that the final step in the prescribed stateness-first sequence (democratization) may be hard to achieve. Countries that build state capacity under autocracy may thus be stuck in a high-capacity–autocracy equilibrium for a long time, without reaping the anticipated development benefits following from a 'mature' democratic transition.

Specifying the counterfactual: The stateness-first argument is a causal argument—building state capacity before democratization is proposed to cause, among other outcomes, faster economic development. As all causal arguments, the validity of the stateness-first argument hinges on the relevant comparison, and thus assumptions about the proper counterfactual. An important question in this regard is: what is the relevant comparison to a state that democratized after high-capacity state institutions came in place (see also Knutsen, 2013)? This question is often neglected, or receives only a vague answer, in existing contributions, for understandable reasons. Explicitly specifying the appropriate counterfactual in the stateness-first argument is trickier than one might suppose and, we surmise, critically depends on how one interprets the broader theory.

First, if the theory is construed as the causal effect of democratization D (represented by

a binary variable in which 1 =democratization), conditional on the pre-existing level of state capacity, S ($0 = low; 1 = high$), on some outcome, Y, then the proper comparison would be: $Y = (Y_{(D=1|S=1)} - Y_{(D=0|S=1)}) - (Y_{(D=1|S=0)} - Y_{(D=0|S=0)})$. This compares the effects of democratization in high- vs. in low-capacity states. In this formulation, constructing the counterfactual outcome for a democratizing state under high capacity is non-trivial, since it involves comparing it to three counterfactual scenarios and not simply to democratization under low capacity. Below, we conduct such types of comparisons by using panel regressions (e.g., Figure 2 and Appendix A), but also by using matching techniques (e.g., Models 1 and 2, Table 2).

Alternatively, one might contend that some versions of the stateness-first argument focuses squarely on differences within the subset of observations that actually experience democratization; obtaining democracy under high state capacity should lead to stronger future development than democratizing in low-capacity contexts. This is equivalent to stating that $Y_{(D=1|S=1)} - (Y_{(D=1|S=0)} > 0$. While we highlight that this statement does not speak to the *causal effect of democratization*—no contrasts are made against counter-factual outcomes associated with remaining autocratic, under various realizations of S—this descriptive claim certainly exists in various formulations of the stateness-first argument. We thus also run tests aiming to evaluate this claim by only comparing observations that have undergone democratic transitions (e.g., Models 3-5, Table 2).

Yet, if we interpret the theory to say that the *sequence* of institutional changes matters (regardless of the effects of democracy on the development of a capable public administration, or vice versa) then the proper comparison is between a country that historically democratized *after* developing a capable public administration, and a country that democratized *before* developing a capable public administration. Importantly, such comparisons zooms in on the effect of the particular *historical sequence* of institutional adoptions and isolates it from the effects of *levels* of state-capacity and democracy. This type of comparison amounts

10

to saying that two *otherwise equal countries*, but with different institutional-sequencing histories, will differ in outcomes *due to these different historical sequences*. Tests that assess this specification of the theory are presented in Section 4.3.

A final complicating matter for evaluating the stateness-first argument relates to the complex inter-relationship between democracy and state capacity. Our discussions above suggested that state capacity is endogenous to regime type, but also that democratization may be a function of state capacity. These points are appreciated by proponents of the stateness-first argument, although the anticipated signs of the relationships between democracy and state building are often different from what our discussions above suggested. Nonetheless, if we anticipate that there are links between the two institutional factors, this also has implications for considerations on counterfactuals and empirical design. If we assume that subsequent regime developments and changes to state capacity are strongly linked to whether or not the first historical transition to democracy took place in a high- or low-capacity context—for example because 'premature democratization' leads to both political instability and difficulties in building capable administrations—it makes sense to only compare observations on the basis of their first democratic transition. Further, we should then measure Y with a substantial time lag, and *not* control for subsequently realized values on democracy and state capacity, since doing so will induce post-treatment bias (see, e.g., Model 4, Table 2). In contrast, if we believe that subsequent developments to regime type and state capacity (after the first historical transition) are driven mostly by other factors, it makes sense to control for these subsequent historical developments and even current realizations of state capacity and democracy (see, e.g., Model 5, Table 2; Models 3–5, Table 3).

2.3 Extant evidence

The body of evidence for the stateness-first argument largely consists of historical country narratives. Both early (e.g., Huntington, 1968) and recent contributions (e.g., Fukuyama, 2014a) draw heavily on case histories that comport with the prescribed sequence of building

state institutions before introducing mass politics, often from the historical experiences of Western countries. Møller (2015) questions the accuracy of the common narrative surrounding the sequencing of institutions in Western countries, however, arguing that "[t]he notion, so often taken for granted, that it went 'state-first, democracy later,' rests on an oversimplification of European history. What that history shows is that instances of either 'state-first' or 'democracy-first' sequencing were rare. What was much more common was for budding state institutions, the rule of law, and political accountability to grow alongside one another while interacting in messy ways. If there is any sequential pattern, it is for state-building to appear very late in the game." (p.111). Responding to this criticism, Fukuyama (2014b) argues that it employs a too inclusive definition of democracy, and that the historical legislative assemblies and rights highlighted by Møller (2015) represent rule-of-law institutions rather than democracy.

Still, this points to a broader problem with the extant (largely case-based) evidence used to support the stateness-first argument: When stringent operationalizations of the relevant institutional features are lacking, it is, in practice, difficult to reliably describe the actual sequence of institutional development postulated in the theory. This, in turn, makes it problematic to assess whether a case narrative—no matter how thoroughly laid out— actually corroborates the theory or not. The historical-narrative type of evidence also makes it hard to control for factors that contribute to the endogenous evolution of state-capacity and democracy. The problem of clearly identifying proper counterfactuals, discussed above, also raises issue for the interpretation of case narratives. To evaluate hypotheses on the detrimental consequences of the 'premature' introduction of democracy, the appropriate contrast class for the clientelistic, young democracy with low state capacity and weak rule of law not only includes countries that democratized under strong state capacity and well-functioning rule of law, but also the patronage-ridden autocracy with low state capacity and an equally weak rule of law. Such complex comparisons have typically not been made in the

case-based literature, at least not in an explicit and systematic manner.

What is more, given that explicit rules for selecting observations are missing, the case-based evidence is open to the charge of selective choice and interpretation of cases. These issues are compounded by the lack of clear criteria for how to select among different time periods, patterns, and events to exemplify a country's development. Concerning the selection of countries, how would, for example, the inclusion of narratives from Botswana or Mauritius—recent development miracles happening under democratic rule, in countries where democratization occurred under (initially) low levels of state capacity—alter the evaluation of the theory? (For numerous country cases, from different regions, that seem to contrast with the stateness-first argument, see Mazzuca and Munck, 2014). To exemplify other selection issues, the (relatively authoritarian) Prussian regime is (correctly) lauded by several scholars for its ability to modernize the military and state. These experiences have also been invoked as evidence for the stateness-first argument (see, e.g., Fukuyama, 2014a). However, a reading of somewhat more recent German history would highlight how autocratic forces contributed to the country entering into two world wars, with devastating effects on the infrastructure, economy and human development. This speaks to issues of unclear selection of outcome variables as well as potential selection biases related to the time period under study—whether or not Prussia/Germany is an unambiguous success story arguably depends on whether we end our investigations in 1885, 1920, 1945, or 1970.

Accompanying the wealth of case-narratives, a few large-n studies have aimed to evaluate stateness-first arguments. One recent example is D'Arcy and Nistotskaya (2017). These authors provide a novel justification based on rational choice theory for the hypothesis that sequencing state capacity before democratization enhances governance and provision of public goods. The authors draw on an impressive data collection effort on state-administered cadasters—"systematically arranged inventories of individual land parcels and land ownership" (p.2)—for 78 countries back to the year 1 A.D. They use these data to construct an

indicator of states' monitoring capacity. In a series of cross-section regressions, D'Arcy and Nistotskaya (2017) show that countries that scored high on this index at the time of democratization currently outperform countries that scored low on the index at democratization, on different public goods and development outcomes (quality of public services, education expenditures, infant mortality rates). However, D'Arcy and Nistotskaya (2017) do not include countries that remain autocratic, and consequently do not compare the performance of democratizers versus non-democratizers either in contexts of low or high state capacity. As our discussion on counter-factuals above suggests, this generates issues for evaluating any causal effect implied by the stateness-first argument. The analysis does show that countries such as Denmark or Sweden, which had high state capacity at democratization, are associated with better outcomes than countries with low capacity at democratization such as Benin or Mongolia.[2] But, this finding could stem from other factors, such as state capacity being persistent *and* affecting development (regardless of the timing of democratization).[3] Hence, these analyses do not provide direct evidence for the stateness-first argument.

Two studies speak somewhat more directly to the stateness-first argument, investigating whether the effects of democracy are conditional on level of state capacity, with economic growth (Knutsen, 2013) and health-care and education outcomes (Hanson, 2015) as dependent variables. Both studies suggest that democracy actually has a significantly *stronger* positive effect on the different outcomes in contexts of low state capacity, which contrasts with a core assumption of the stateness-first argument—that democracy has more benevolent effects in high-capacity contexts. Still, these studies rely on fairly limited time series or time-invariant measures of state capacity, and do not explicitly assess the historical sequences of democratization and state-building.

---

[2]The usual caveats related to drawing causal inferences from cross-sectional regressions apply, notably related to unobserved geographic, cultural or political-historical confounders.

[3]Indeed, D'Arcy and Nistotskaya (2017) show that their cadaster index measured at the time of democratization is a significant predictor of current levels of state capacity.

14

Below, we employ new data with longer time series that comprise most of what historians regard as 'modern history', including the late 18th and 19th centuries. This is the time period from which most case-based evidence supporting the 'stateness-first' argument is drawn. We examine the argument not only by re-assessing the above-mentioned core assumption tested in Knutsen (2013) and Hanson (2015) on longer time series, but also by testing various empirical implications that follow from the argument, notably including tests directly assessing the relevance of the temporal sequence in which different institutions were introduced historically.

3 Measuring institutions

Theories about institutional sequencing are difficult to test systematically. The hypotheses derived from such theories pertain to developments over long periods of time and involve different institutions that need to be clearly distinguished. Thus, in addition to covering many countries, data should a) have sufficiently long time series to capture pertinent historical changes and b) include detailed and distinct indicators on the relevant institutional features. Some datasets, such as Polity (Marshall and Jaggers, 2007), offer long time series but only include measures of democracy-relevant aspects and do not sufficiently distinguish between different democratic-institutional features (see, e.g., Coppedge et al., 2011). Datasets measuring state-capacity features have either been purely cross-sectional (e.g., Evans and Rauch, 1999) or based on short time series (e.g., Kaufmann et al., 2010).

The data situation has changed with the recent Varieties of Democracy (V-Dem) dataset (Coppedge et al., 2017). V-Dem includes more than 400 detailed measures, not only related to narrow conceptualizations of democracy (e.g., contested multi-party elections) but also features of rule of law and state capacity. Some measures are more objective and are coded by research assistants, whereas others are more evaluative and require expert judgments. Several strategies are pursued to limit measurement error and ensure cross-expert-, inter-

temporal-, and cross-country comparability.[4] V-Dem covers about 180 countries from 1900 to the present. Notably, the more recent 'Historical V-Dem' data extends the time series for many V-Dem indicators and indices back to the late 18th or early 19th century for 91 countries, thus covering a key period of state-building (Knutsen et al., 2017).

Our primary measure of democracy is V-Dem's Polyarchy index (Teorell et al., 2016). Stateness-first arguments focus on the introduction of mass politics through basic electoral features of democracy such as multi-party elections and extensive suffrage, which Polyarchy is well-suited to capture. It contains five sub-components on whether or not the chief executive is elected, how free and fair these elections are, and the extent to which there is freedom of association, freedom of speech, and universal suffrage. The inclusion of freedoms of association and speech reflect the importance of the free formation of opposition parties and open discussion for multi-party elections to be truly competitive (see Dahl, 1971). Polyarchy ranges from 0 to 1, but the empirical minimum and maximum values (for 22,406 country-year observations) are 0.01 and 0.95, respectively. The left panel of Figure 1 illustrates that most country-year observations cluster on the lower end of the scale, with differences in the middle- and higher ends of the scale reflecting variation between 'minimum-level electoral democracies' and 'high-quality democracies'. Given the focus of the 'stateness-first' argument, which considers the consequences of introducing core democratic features, we should use relatively low cut-off values in specifications that dichotomize Polyarchy. We therefore often divide Polyarchy by its median (0.18 in full sample), which also produces evenly split sub-samples and ensures higher-powered tests.

In robustness tests, we dichotomize Polyarchy by dividing at the mean (0.28). Switch-

---

[4]V-Dem uses a Bayesian Item-Response measurement model that leverages different kinds of information (e.g., cross country coding and anchoring vignettes) to generate comparable, interval-level scores from the ordinal indicators coded by multiple experts (Pemstein et al., 2017).

ing from a median- to mean-based dichotomization moves countries such as the United States in 1800 and United Kingdom in 1835 (after the Reform Act) from the high- to the low-democracy category. We also try out different cut-off points, and, where possible, run models with Polyarchy as a continuous variable. Further, we test the binary electoral democracy measure from Boix et al. (2012) (BMR), which registers the presence of 'free and fair' elections and requires that 1/4 of citizens are enfranchised. Finally, given the emphasis put on suffrage in stateness-first arguments, we also test V-Dem's *v2x_suffr* indicator, measuring the share of the adult population that is enfranchised.
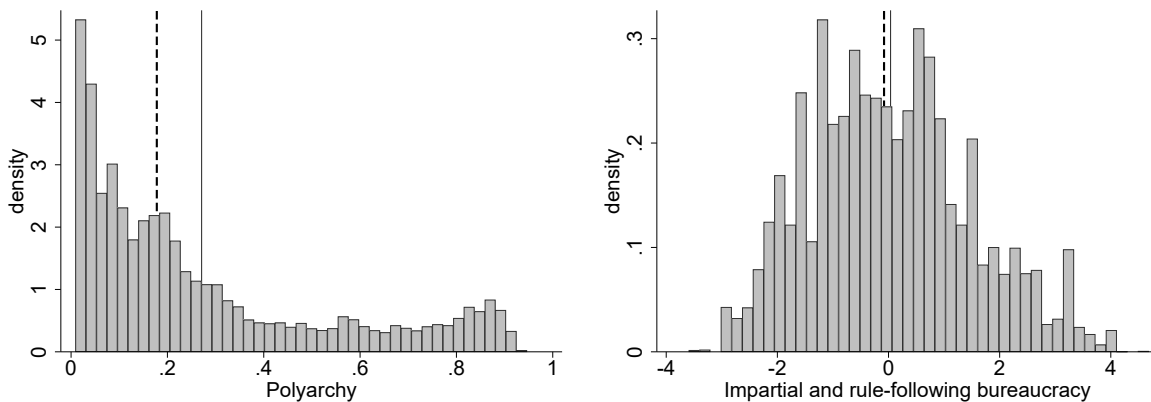
Figure 1 – Distribution of Polyarchy and impartial and rule-following administration (*v2clrspct*), 1789–2016. Dotted line represents the median; solid line indicates the mean.

State capacity is variously defined in the literature, but key to many conceptualizations is the ability of the public administration to effectively implement policies. This, in turn, relates to several 'Weberian' features of the bureaucracy, including recruitment processes for government officials (based on merit, as opposed to personal connections) and how officials are compensated (decent wages, which supposedly mitigate corruption) (e.g., Evans and Rauch, 1999). Other key features of Weberian bureaucracy is that decisions are based on impersonal rules rather than personal discretion and that they are implemented impartially. These features also tie into the concept of rule of law, which is explicitly considered a key

"pre-requisite" for democratization by many scholars proposing stateness-first arguments, including Fukuyama (2014a). While we test alternative proxies for state capacity from V-Dem such as meritocratic recruitment to the state administration (*v3stcritrecadm*) and extent of corruption (*v2x_corr*), we thus rely principally on *v2clrscpt*. This V-Dem indicator measures the extent to which public officials are impartial and rule-following in carrying out their duties. Expert coders originally score this item on a five-point ordinal scale, which is subsequently transformed to a continuous scale by the V-Dem measurement model (Pemstein et al., 2017). The impartial and rule-following bureaucracy measure covers 24,005 country-year observations across 1789–2016. It ranges from -3.55 to 4.67, and is close to normally distributed with a median of -0.09 and mean of 0.06 (Figure 1; right panel).

Our dependent variables draw on the extensive Gross Domestic Product per capita (GDP p.c.) data from Fariss et al. (2017). To mitigate measurement error, which is present in all extant measures of GDP (see Jerven, 2013), Farris et al. employ a dynamic latent trait model to produce less error-prone estimates from several GDP (and population) sources. Specifically, we use the estimates benchmarked in the Maddison time series, which has the most extensive coverage among all GDP sources. In addition to mitigating different types of measurement error, the imputation of missing values by the Farris et al. routine helps reduce biases resulting from sample selection (Honaker and King, 2010). We employ (forward-lagged) Ln GDP p.c. (controlling for initial Ln GDP p.c.) and GDP p.c. growth, across different time intervals, as our dependent variables. For presentational reasons we introduce the control variables in the next section.

4 Empirical analysis

4.1 Panel regressions

Our first tests are carried out with standard panel regressions. These tests probe whether democracy is more beneficial for subsequent economic development when there is high versus

low state capacity. By gauging whether the impact of democracy differs in low- and high-capacity states, these models assess a key premise on which the 'stateness-first' argument rests. We also test a key implication of the argument by assessing whether the level of state capacity before or at the time of democratization relates to subsequent development.

We probed numerous panel specifications. The tests in Table 1 use Ln GDP p.c. as the dependent variable, whereas Appendix A reports similar specifications using annualized GDP p.c. growth. We start by running an OLS specification with country-year as unit of analysis and errors clustered by country to account for autocorrelation. We control only for lagged Ln GDP p.c. alongside country- and year-fixed effects. The year-fixed effects help account for time trends and shorter-term global shocks to economic development and institutional features that are common to all units. The country-fixed effects address country-specific and time-invariant factors related to geography, culture, etc., that may simultaneously affect institutional- and economic development The ability to control for such hard-to-observe confounders is a distinct advantage of the panel models, both relative to existing studies drawing on one or more historical country narratives, as well as large-n studies (including the models included in the next sections) that draw on cross-country comparisons.

We intentionally keep our benchmark sparse in order to mitigate post-treatment biases. Since democracy and state capacity may very well influence variables such as civil war or natural resource dependence, controlling for them eliminates (relevant) indirect effects. Yet, we do include these and other controls in subsequent tests that prioritize mitigating omitted variable bias over mitigating post-treatment bias. We begin by measuring the outcome 20 years after the covariates, in order to gauge the medium-term effects of democracy in contexts of high and low state capacity.

We first run our benchmark on sub-samples of low-capacity (Model 1) and high-capacity (Model 2) states, respectively. To produce balanced sub-samples, we split by the median sample value on *v2clrscpt* (-0.090). Polyarchy has a negative coefficient in the low-capacity

Table 1 – Panel regressions with Ln GDP per capita as dependent variable

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	LC	HC	Full	Full	Full	Full	Full	Full
Estimator	OLS	OLS	OLS	OLS	OLS	OLS	GMM	GMM
Panel length	1 yr	1 yr	1 yr	1 yr	1 yr	1 yr	10 yrs	10 yrs
Dep var measured in year	$t+20$	$t+20$	$t+20$	$t+20$	$t+20$	$t+40$	$t+20$	$t+10$
	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)
Polyarchy	-0.147	0.144	-0.101		-0.169	-0.101	0.029	0.181
	(0.175)	(0.119)	(0.119)		(0.115)	(0.198)	(0.331)	(0.167)
Polyarchy X Impartial adm			0.072		0.045	0.096	-0.103	-0.126**
			(0.046)		(0.037)	(0.090)	(0.100)	(0.050)
BMR democracy				0.009				
				(0.040)				
BMR X Impartial adm				0.011				
				(0.024)				
Impartial public admin			0.006	0.001	0.021	-0.007	0.078**	0.081***
			(0.018)	(0.021)	(0.024)	(0.035)	(0.036)	(0.027)
Civil war					0.011			
					(0.038)			
Resource dependence					-0.007***			
					(0.003)			
Ln Population					-0.010			
					(0.048)			
Ln GDP pc (lagged DV)	0.846***	0.611***	0.776***	0.809***	0.776***	0.663***	1.078***	0.999***
	(0.051)	(0.068)	(0.049)	(0.045)	(0.046)	(0.076)	(0.093)	(0.042)
Country dummies	Y	Y	Y	Y	Y	Y		
Year/period dummies	Y	Y	Y	Y	Y	Y	Y	Y
N	7318	7307	14643	11816	10073	11426	1530	1708
Countries	127	145	177	176	157	156	171	179
Maximum years/periods	202	206	206	195	179	186	21	22
R^2	0.727	0.882	0.829	0.853	0.844	0.761		
Number of instruments							169	178
Hansen J-test p-value							.359	.978
AR(2) test p-value							.011	.968

Notes: *p<0.1; **p<0.05; ***p<0.01. Constant, country- and year dummies omitted. Errors clustered
by country in OLS and robust in (System) GMM. LC: Low capacity. HC: High capacity, determined relative to
median-sample value on IPA (-0.090) for full sample for specification corresponding to Models 1 and 2.

sample and a positive in the high-capacity sample. Yet, none of the coefficients are close to conventional levels of statistical significance. Model 3 is our benchmark run on the full sample, but including Polyarchy, *v2clrscpt* and a multiplicative interaction term. While the estimated relationship between democracy and growth increases in level of state capacity, the pattern of interactions is not systematic. The same holds true in Model 4, which substitutes Polyarchy with the dichotomous democracy measure (BMR) from Boix et al. (2012), and in Model 5, which controls for population (Fariss et al., 2017), natural resources income/GDP (Miller, 2015), and civil war (Haber and Menaldo, 2011).

More generally, the (null) result holds across sets of plausible controls and for alternate measures of state capacity and democracy. The result is also robust when focusing more specifically on suffrage rather than broader democracy measures, using annualized GDP p.c. growth instead of Ln GDP p.c., and when changing the time frame across which we measure the dependent variable (both shorter and longer time frames; see Appendix A). Model 6 illustrates the latter, as it measures Ln GDP p.c. 40 years after the covariates. So far, there is no clear support for the core assumption of the stateness-first argument that democracy is more conducive to development in high-capacity than in low-capacity contexts.

Next, we employ the system Generalized Method of Moments (GMM) estimator, which is appropriate for dealing with 'sluggish' variables such as democracy and state capacity (Blundell and Bond, 1998). This estimator allows us to account for the potential endogeneity of our institutional variables (and interaction term), by using lagged levels (changes) as instruments for current changes (levels) in institutions (Roodman, 2009). Hence, when model assumptions are met, GMM estimates should not reflect any reverse effect from economic development on democracy and/or state capacity. GMM models are, however, originally constructed to handle relatively short time series, and long time series typically increase the number of instruments beyond advisable levels (the rule of thumb is fewer instruments than cross-section units, see Roodman, 2009). Thus, we follow the conventional practice of

growth economists and reduce our dataset to a ten-year panel structure.[5] In Model 7, which measures Ln GDP p.c. 20 years after the covariates, we find a negative, but insignificant, interaction term. However, the AR(2) test suggests that autocorrelation is an issue, and results are thus not credible. When measuring Ln GDP per capita only ten years after the covariates in Model 8, however, the negative interaction between democracy and state capacity turns statistically significant at 5%. The AR(2)- and Hansen J-tests suggest that this model gives consistent estimates. Hence, Model 8 indicates that democracy actually has a more benevolent (medium-term) effect on growth in low-capacity states. This result is in line with the findings in Knutsen (2013), which are based on data from more recent decades. Overall, however, our panel regressions do not reveal a robust interaction between state capacity and democracy on growth.

Another test of the stateness-first argument is to compare changes in growth before and after all recorded democratization episodes, and check whether post-transition increases in growth are more likely for democratization episodes in high-capacity states. Appendix A discusses such panel regressions, for instance using the measure by Boix et al. (2012) to identify democratic transition episodes. We do not find any evidence corroborating the stateness-first argument from these tests either. This is illustrated by Figure 2, which contains four scatter plots that map the difference in pre- and post-transition growth (both measured over twenty-year periods) along the y-axes and 10-year pre-transition averages on four proxies of state capacity along the x-axes. These measures include our main measure on impartial and rule-following administration, but also V-Dem measures on corruption, clientelism, and meritocratic recruitment to the bureaucracy. There is substantial variation in post-transition growth changes, both among low- and high-capacity states, and the best-fit lines

---

[5]Since we have three endogenous institutional variables and very long time series, we must also restrict the lags used for instrumentation (second and third lag) to keep the instrument-count down.
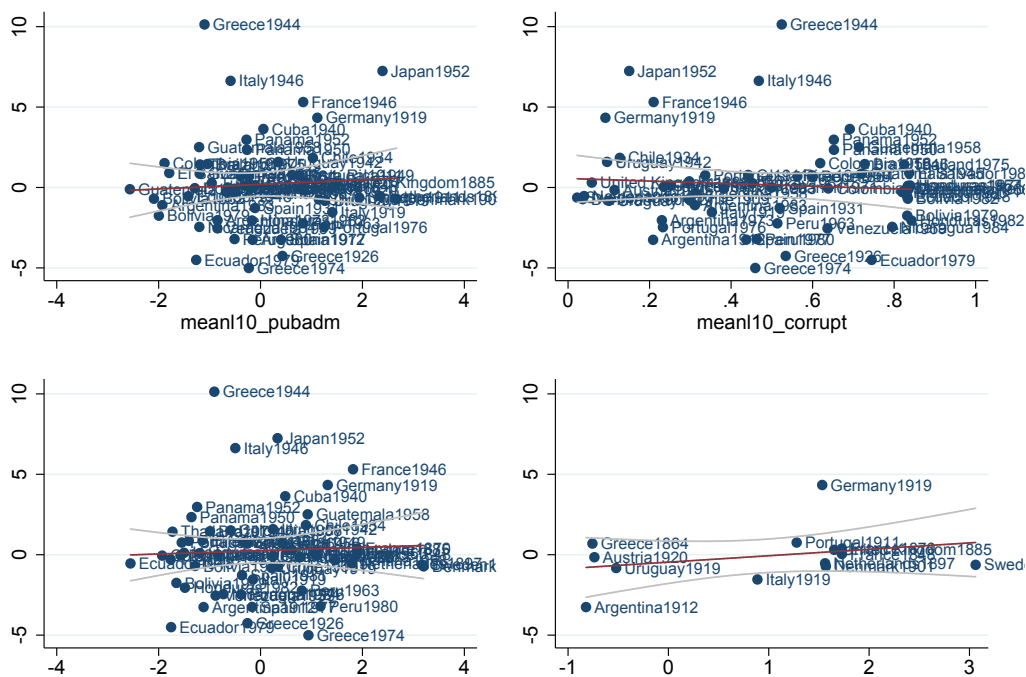
Figure 2 – Proxies of state capacity and change in growth rate from before to after transition. Notes: Average 20-years post-transition GDP p.c. growth minus average 20-years pre-transition GDP p.c. growth along the y-axis, and 10-year average pre-transition scores on rule-following bureaucracy (upper-left), corruption (upper-right), clientelism (lower-left), and meritocratic recruitment (lower-right; data from 1789–1920) along x-axes. The scatterplots are overlaid with best-fit lines and 95 percent confidence intervals.

and confidence intervals reveal no systematic patterns. This conflicts with the stateness-first argument, which predicts a greater growth-benefit from democratization in high-capacity contexts.

4.2 Matching analysis

Still, as discussed, it is not clear what is the most appropriate counterfactual comparison when probing the stateness-first argument. The choice of contrast class depends on specific assumptions related, e.g., to how democracy influences state building. To tackle this issue, we present tests making different relevant comparisons under alternative assumptions about appropriate counterfactuals, using Coarsened Exact Matching (CEM) models (see Iacus

et al., 2012). These models allow us to draw inferences from (only) comparing otherwise similar units (in their propensity to receive treatment) that differ on the treatment variable, and we assess both medium- and longer-term development effects of democratization in different state-institutional contexts.

All models include the following matching variables: year of democratization, GDP per capita at democratization, and—since we cannot account for country-fixed effects in this set-up—geographic region and score on the Ethnic Fractionalization Index from Alesina et al. (2003). CEM demands that all variables are categorized for the matching, and observations are only compared with observations placed in the exact same categories. Hence, there is a trade-off between only comparing very similar observations (fine-grained categories) and having many observations with available matches (broader categories). In some models we group year into three categories, with cut-offs chosen to reflect the various 'Waves of Democratization' (Huntington, 1991). In other specifications, we compare on the exact year of democratization. Ln GDP per capita is always recoded into three categories (<1st quartile; 1st–3rd quartiles; >3rd quartile). Ethnic fractionalization and the rule-following and impartial bureaucracy measure ($v2clsrcpt$) are recast into binary variables, with median values as thresholds. We always run (OLS) regressions after the matching to account for differences on the covariates within the categories, and then use the (logged) numeric version of these variables as covariates.[6] Our benchmark operationalizes 'democratization' as going from below- to above-median score on Polyarchy.

We start out with matching-designs that resemble the logic of the panel specifications. We compare countries that experienced democratization with autocratic countries—i.e., the treatment that we match on is democratization—but only allow for comparisons within

---

[6]CEM groups similar observations into different subclasses, and weights the observations based on the ratio of treatment-/control observations. We use these weights in all regressions, which also include subclass fixed effects.

groups of otherwise similar countries, following the rules outlined above and matching on the democratization year. Thus, we only compare countries with below-median scores on the rule-following and impartial public administration measure with each other, and similarly for above-median countries. In subsequent regressions, we estimate whether score on the impartial and rule-following administration measure *at the time of democratization* systematically predicts differences in logged income level 20 years later.[7] Initially, we allow countries to enter with more than one democratization episode, but to avoid mixing effects from multiple episodes we only include democratization episodes if the country had 20 years of consecutive autocratic rule prior to democratization. Results are reported in Model 1, Table 2. The coefficient on impartial and rule-following administration is very small (0.06) and far from statistically significant ($t = 0.7$). Hence, we do not find that a rule-following and impartial bureaucracy *at the time of democratization* enhances subsequent economic development.

Since our democratization measure is constructed by setting a cut-off on a continuous scale, one worry is that some countries could tip just above the median-value on Polyarchy one year, and then revert the next year. Such small changes could even result from measurement uncertainty (Pemstein et al., 2017). Therefore, we re-estimated Model 1 by restricting the definition of democratization to situations when a country passed the median-threshold and stayed above it for 10 consecutive years. The downside of using this '10-year rule' is that it introduces a form of post-treatment bias, since democracies with higher growth rates are more likely to avoid democratic regressions (Przeworski and Limongi, 1997; Kennedy, 2010), and because political instability and democratic regressions can be viewed as one proposed mechanism through which transitions under low state capacity could mitigate growth, per

---

[7]A downside to using such a long time lag is that estimated effects may be attenuated if the autocracies compared with democracies go through democratization experiences in the time interval between our independent and dependent variables are measured. We deal with this issue in alternative specifications below.

Table 2 – Coarsened exact matching analysis; results from post-matching OLS regressions.

	(1) Ln GDP 20 years after democratization	(2) Ln GDP in 2004	(3) Ln GDP 20 years after democratization	(4) Ln GDP in 2004	(5) Ln GDP in 2004
Bureaucracy-level at democratization	0.06	0.11	−0.03	−0.09	−0.01
	(0.07)	(0.08)	(0.05)	(0.06)	(0.07)
Ln GDP at democratization	0.94***	0.75***	0.52**	0.38	−0.26
	(0.15)	(0.16)	(0.18)	(0.20)	(0.32)
Ln year at democratization	−3.69	30.85	8.21	10.77	10.59
	(21.33)	(36.14)	(6.31)	(8.86)	(11.39)
Ethnic fractionalization at democratization	−0.75	0.81	1.14	−0.22	−0.33
	(0.70)	(0.70)	(0.70)	(0.63)	(1.00)
Constant	28.50	−228.51	−58.67	−75.58	−69.16
	(161.38)	(272.25)	(47.84)	(67.33)	(86.62)
Regional Dummies	Yes	Yes	Yes	Yes	Yes
Subclass Dummies	Yes	Yes	Yes	Yes	Yes
Comparison:	Demo. vs Aut.	Demo. vs Aut.	Democratizers	Democratizers	Democratizers
Democratization rule:	After 20 years of Aut.	First only	After 20 years of Aut.	First only	First only
Observations	91	188	69	57	28
R²	0.87	0.85	0.79	0.95	0.97
Adjusted R²	0.80	0.78	0.71	0.92	0.95

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors in parentheses.

the stateness-first argument. However, the insignificant impartial and rule-following administration coefficient barely changes when employing the 10-year rule (Model 1, Appendix Table B.1). Results are also fairly stable (with and without the 10-year rule) to using transitions from 0 to 1 on the BMR measure of electoral democracy or a dummy registering when countries pass a threshold of 50 percent of the adult population being enfranchised, although we note that one specification using majority suffrage as indicator of democratization lends support to the stateness-first argument (see Appendix B).

Yet, 20 years might be insufficient for capturing longer-term differences in development coming from experiencing democratization at different levels of state capacity. The outcome variable in Model 2 is Ln GDP p.c. measured in 2004, the year with the best data coverage after 2000. We now only use the *first* democratization episode recorded in a country's history. This is consistent with the notion that political instability and democratic deterioration are partly consequences of the first democratic transition happening under conditions of weak state capacity. Thus, even if subsequent transitions carry an effect on economic development, these effects may be viewed as 'indirect effects', and should thus not be controlled for. We match countries that experience their first democratization with an autocratic country observed in the exact same year, and otherwise follow the matching procedure as outlined above. Model 2 shows no evidence that impartial and rule-following bureaucracy at democratization is related to contemporary economic performance. The relationship is both weak (0.11) and statistically insignificant.

Proceeding with a different set of comparisons, the last three models in Table 2 only compare countries that undergo democratization experiences, but which do so under contexts of relatively high and low state capacity. In other words, the treatment that we match on is the dichotomized measure on impartial and rule-following bureaucracy. We follow the template from above and match these democratization observations on time period of democratization, income level at democratization, geographic region, and ethnic fractionalization, and

run OLS regressions post-matching. Model 3 measures the outcome 20 years after democratization. Model 4 measures the outcome in 2004, thus evaluating whether historically having transitioned under weak or strong state capacity relates to 'current' income. Once again, we find no support for the hypothesis that transition under state institutions with relatively high capacity corresponds with subsequent economic development, be it in the medium or long term.

Finally, in Model 5, we compare countries that are about equally democratic and have about equally capable bureaucracies today, but which have different institutional-sequencing histories. First, we subset the data to countries that were relatively democratic and had relatively capable states in 2004, and match on region, ethnic fractionalization, time period of democratization, and income at democratization. We then compare 2004 Ln GDP per capita scores across countries that had relatively high and low scores on the bureaucracy measure *at their first incidence of democratization.* Do these otherwise similar countries differ in income today, *based on their different democratization histories*? The results from Model 5 suggests that they do not. The results are robust to making various changes to the matching specification, including when we employ Entropy Balancing (Hainmueller, 2012) instead of Coarsened Exact Matching (see Appendix B).

4.3 Comparing sequences

Our final approach involves comparing observations on the different institutional sequences that they experienced historically and assessing potential consequences for economic growth. The long time span of our data enables us to differentiate country-histories based on whether high levels of state capacity preceded high levels of democracy or not. The particular mechanisms that explain changes in the relative levels of state strength and democracy may differ across countries and time—and indeed, they may be endogenously related to historical levels of economic development, which we control for. But, the fundamental implication that we want to assess is whether such stateness-first patterns, in general, are a boon for

later economic development. Compared to the panel-tests and the pre-parametric match-ing analysis, this approach has the benefit of leveraging the entire institutional history of a country and correlating it with growth, enabling what are arguably more direct—even if descriptive—assessments of the stateness-first sequencing argument.

We start by conceptualizing fairly simple patterns of development, dividing the roughly 22,000 country-years with coverage on both V-Dem measures into four states based on whether each observation fell above/below median scores for Polyarchy (0.178) and $v2clsrcpt$ (-0.080) in each year. The sub-samples constitute, respectively, 34 (low democracy–low state capacity), 17 (high democracy–low capacity), 16 (low democracy–high capacity), and 33 (high democracy–high capacity) percent of observations. When calculating transition prob-abilities between states (see Figure 3) we find that countries with low levels on both state capacity and democracy fairly often transition to high democracy–low capacity and that transitions from low to high state capacity among non-democracies are relatively infrequent. This pattern follows our earlier criticism of the assumption underlying the stateness-first argument, concerning prospects of state building in autocracies. Still, the observation that high democracy–low capacity situations often transition into low capacity–low democracy sit-uations is more in line with the stateness-first argument. Weak-capacity democracies fairly often go through 'de-democratization' processes. Yet, we note that this observation is also consistent with various other theories of political development, including theories proposing that low income (which correlates with low state capacity) destabilizes democracies (e.g., Przeworski and Limongi, 1997; Boix and Stokes, 2003; Inglehart and Welzel, 2005).

To test the relationship between institutional sequencing histories and growth, we specify OLS models with Ln GDP p.c. measured in $t + 20$ as dependent variable, and include current Ln GDP p.c as regressor. We add a dummy variable capturing the stateness-first pattern, identifying those countries that had previously transitioned from low democracy–high capacity to high levels of both. The indicator does not require that transitions between

29
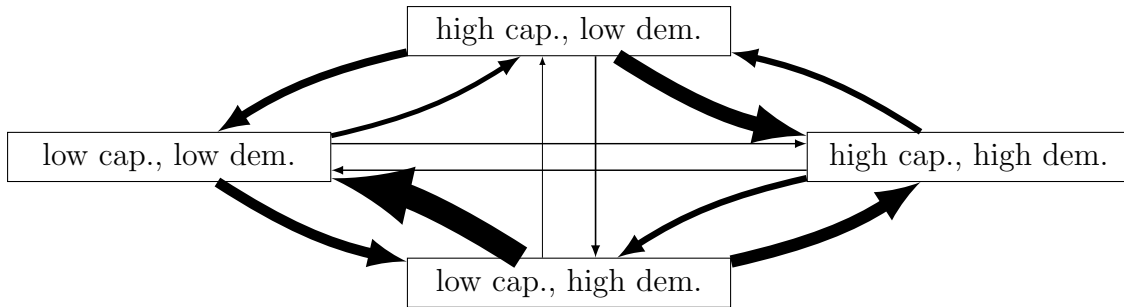
Figure 3 – Relative transition probabilities
(*line thickness indicates transition frequency*)
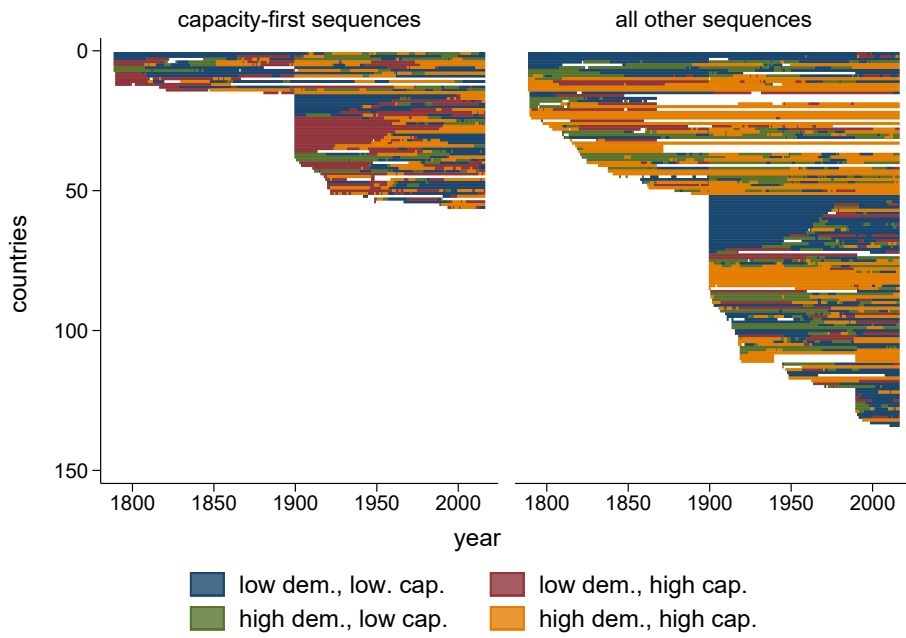
Figure 4 – Sequence index plot distinguishing 'stateness-first' sequences.

states are consecutive, but merely that high capacity and low democracy co-existed prior to attaining high capacity–high democracy status. To isolate the effect of the *sequencing* of institutional changes from the *levels* of state capacity and democracy at each point in time, as well as from the *duration* of time over which the institutional states have been in place, we use different control strategies.

Figure 4 illustrates the institutional sequencing histories—each country is represented by a horizontal line—for 'stateness-first' countries (left) and all other countries (right). There are 56 countries categorized as 'stateness-first' countries that include countries with (generally) high-growth track records, such as Japan and Taiwan, but also countries with more mixed development records, such as Turkey and Venezuela. OLS results comparing them to countries that did not follow a 'stateness-first' pattern are reported in Table 3

In Model 1, Table 3, we include three dummies representing the different 'institutional states' (low democracy–low capacity is reference category) alongside the dummy capturing whether a country experienced a stateness-first sequence. According to this model, having experienced a transition from high state capacity and low democracy to high levels of both is *negatively* associated with subsequent economic development, everything else equal. The negative stateness-first dummy is significant at 1%.

However, the time spent in each institutional state may also have lingering effects on growth patterns (see Gerring et al., 2005). In order to account not only for the current institutional make-up, but also for the cumulative institutional experience of a country, Model 2 includes a count of the number of years that a country had previously observed in each state. This model also controls for Ln GDP p.c. at the time of transition. The stateness-first coefficient remains negative and is considerably larger in size than in Model 1. Model 3 is a more constrained version of Model 2 that accounts for civil war, natural resource dependence, and population size. The sample is reduced to roughly 10000 observations due to missingness on these covariates, and among the added controls only natural resources

31

Table 3 – Stateness-first sequences and economic development

	(1)	(2)	(3)	(4)	(5)
	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)
Stateness-first	-0.185***	-0.625***	-0.708***		
	(0.066)	(0.104)	(0.134)		
Stateness-first,				-0.547***	
(20 years)				(0.144)	
Stateness-first,					-0.502***
(20 *consecutive* years)					(0.171)
Low dem., high cap.	0.051	0.064*	-0.030	-0.030	-0.030
	(0.039)	(0.038)	(0.056)	(0.056)	(0.056)
High dem., low cap.	-0.024	-0.018	-0.046	-0.046	-0.046
	(0.036)	(0.035)	(0.037)	(0.037)	(0.037)
High dem., high cap.	0.022	0.036	-0.005	-0.005	-0.005
	(0.048)	(0.046)	(0.054)	(0.054)	(0.054)
Σlow dem., low cap.		-0.000	0.000	0.000	0.000
		(0.002)	(0.001)	(0.001)	(0.001)
Σlow dem., high cap.		-0.001	0.003	0.003	0.003
		(0.002)	(0.002)	(0.002)	(0.002)
Σhigh dem., low cap.		-0.001	-0.000	-0.000	-0.000
		(0.001)	(0.002)	(0.002)	(0.002)
Σhigh dem., high cap.		0.001*	0.002*	0.002**	0.002*
		(0.001)	(0.001)	(0.001)	(0.001)
Ln GDP p.c. at transition		0.073***	0.072***	0.014	-0.001
		(0.012)	(0.016)	(0.011)	(0.012)
Polyarchy			-0.139	-0.137	-0.136
			(0.117)	(0.116)	(0.116)
Impartial administration			0.037	0.037	0.037
			(0.028)	(0.028)	(0.028)
Polyarchy x Impartial adm.			-0.018	-0.020	-0.019
			(0.049)	(0.049)	(0.049)
Civil War			0.012	0.013	0.013
			(0.039)	(0.039)	(0.039)
Resource dependence			-0.007***	-0.007***	-0.007***
			(0.002)	(0.002)	(0.002)
Ln population			0.007	0.006	0.005
			(0.044)	(0.043)	(0.044)
Ln GDP p.c.	0.803***	0.771***	0.767***	0.766***	0.768***
	(0.041)	(0.047)	(0.044)	(0.044)	(0.044)
Year fixed effects	yes	yes	yes	yes	yes
N	14643	14643	10073	10073	10073

Notes: *p<0.1; **p<0.05; ***p<0.01. OLS regressions with
Ln GDP p.c. in t+20 as dependent variable. Constant and year dummies omitted.
Country-year is unit. Errors are clustered by country.

exhibits a significant (negative) relationship with future income. The model also controls for current levels of Polyarchy and impartial public administration, as well as their interaction term. The negative stateness-first result is further strengthened in this specification.

Models 4 and 5 adjust on Model 3 by employing more conservative coding criteria for the stateness-first dummy. In Model 4, we require that a country experienced *at least* twenty years of high state capacity and low democracy before democratization to be registered as 'stateness-first', and in Model 5 we require that a country spent at least twenty *consecutive* years as a high-capacity autocracy before democratization. Consistent with prior tests, these more conservative stateness-first dummies are negative, thus providing evidence in discord with the stateness-first argument. We note that the specifications in Table 3 suggest that having spent numerous years as a high-capacity democracy is conducive to economic development. But, arriving at this institutional combination through first experiencing the building of state capacity and then democratization is *not* a positive factor.
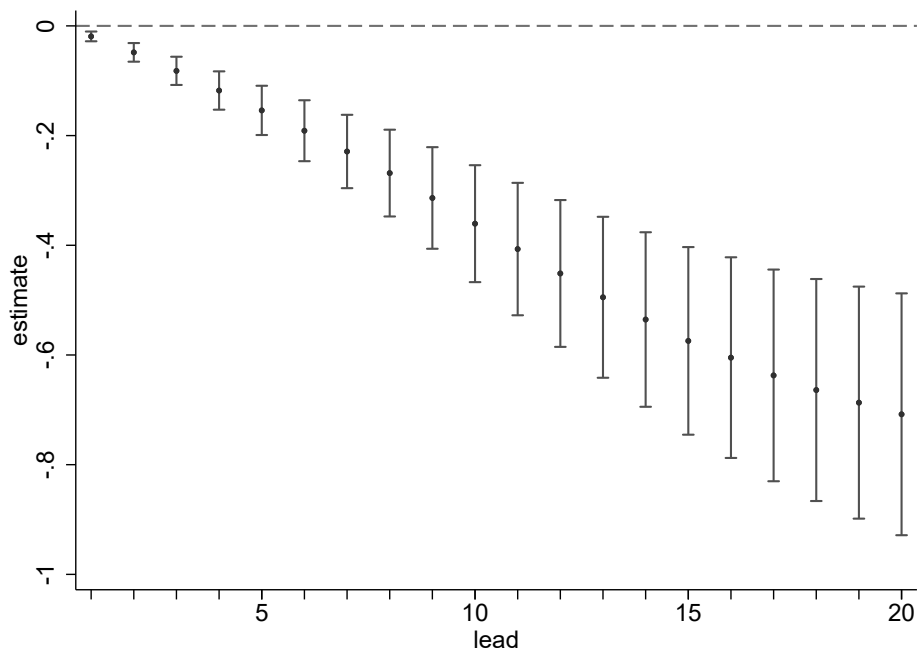
Figure 5 – Estimates for 'stateness-first' transitions, for varying leads.
(*based on Model 3, Table 3*)

We subjected the models in Table 3 to several other robustness tests. First, as Figure 5 suggests, the negative correlation persists for different forward leads. Following the specification in Model 3 of Table 3, the coefficients (surrounded by 90-percent confidence intervals) are estimated when varying the measurement of the dependent variable between $t + 2$ and $t + 20$. At higher leads—estimating GDP per capita further into the future—the uncertainty around the stateness-first estimate increases but the negative coefficient increases in size. Second, we restricted the comparison to include only the 90 countries that transitioned to high capacity–high democracy from *either* high capacity–low democracy *or* low capacity–high democracy. When doing so, the estimate is attenuated but remains statistically significant at the 1% level.

Patterns of institutional change may vary based on differences stemming from geographic factors and colonial histories (Acemoglu et al., 2008), and such differences may also influence growth (although they might do so via affecting institutions; see, e.g., Acemoglu, 2001). Nevertheless, the clear negative relationship remains when controlling for region of the world and the year of independence. The relationship is somewhat sensitive to employing alternative operationalizations of democratization and the chosen threshold for achieving 'high' state capacity (See Appendix Table C.1). We replicated our design, but dividing by mean values for Polyarchy (0.271) and *v2clrscpt* (0.037) rather than the median when generating the four institutional states.[8] The 'stateness-first' dummy is significant at conventional levels and consistently negative in these specifications (see, e.g., Appendix Figure C.1). Our results are also similar when we categorize 'high' levels of democracy and impartial public administration based on membership in the top two-thirds or top one-third of each measure. When we substitute Polyarchy with the dichotomous BMR democracy measure from

---

[8]The categories respectively represent 42 (low democracy–low capacity), 20 (high democracy–low capacity), 11 (low democracy–high capacity), and 27 (high democracy–high capacity) percent of observations.

Boix et al. (2012) for constructing the institutional states, the coefficient on stateness-first sequences remains negative but is no longer statistically significant at conventional levels.

Finally, as a means of comparing the relevance of our pre-categorized 'stateness-first' path with alternative sequences of institutional change, we use optimal matching and hierarchical clustering methods to more inductively identify groups of countries displaying similar institutional-sequencing patterns (Blanchard, 2011). These methods consider countries' entire histories in terms of how they are placed in the four-fold state capacity–democracy categorization scheme. The purpose of doing so is to assess whether the explanatory power of the particular pattern that we have defined as 'stateness-first' is eclipsed by other, more refined institutional-sequencing patterns that are prevalent in the data. When sorting countries into four groups using Ward's clustering method, the groups tend to fall along the distinction of 'stateness-first' and 'democracy first' sequences. Adding these inductively identified sequence groups to our regressions does not have a major impact on the standard errors of the estimate; the negative association between stateness-first sequences and growth holds and remains statistically significant.

In sum, the analysis reported in this section—directly incorporating historical sequences of institutional change—provides some evidence that building state capacity before democratization, if anything, is associated with *slower* subsequent economic development.

5 Conclusion

We have scrutinized, criticized, and tested implications following from the so-called stateness-first argument. While different variants of the stateness-first argument have been forcefully and convincingly put forward by prominent scholars such as Huntington (1968) and Fukuyama (2014a), our analysis suggests that these arguments may not be correct, at least when it comes to explaining patterns of economic development. We have highlighted how these arguments rely on several quite strong assumptions about the causal linkages be-

tween democracy, state capacity, and growth. Further, we elaborated on how the argument is ambiguous when it comes to specifying the proper counterfactual conditions of a state experiencing the purportedly beneficial stateness-first sequence of institutional development. Using new data on countries from across the world, and with time series extending from 1789–2016, we find no systematic support for the stateness-first argument on economic development. The lack of empirical support for the argument appears despite our testing of different empirical implications, proxying for different plausible counterfactual conditions, and using several statistical methods and measures. Indeed, some specifications—notably those explicitly considering the sequencing of institutional changes—find indications that experiencing state-building prior to democratization is related to slower subsequent growth.

Our findings have several important implications for future research. First, one obvious and intriguing extension of this analysis is to test whether stateness-first sequences matter differently for other proposed outcomes of relevance, such as clientelism, civil conflict, or even the survival of democracies after the initial transition has been made. Absent systematic tests we cannot exclude the possibility that stateness-first transitions have important effects on other outcomes. Our preliminary analysis on several alternative outcomes in Appendix D do not provide much support for the notions that stateness-first transitions mitigate corruption or clientelism or enhance human development. However, we *do* find indications in this preliminary analysis that such transitions mitigate the risk of civil war. Second, our findings might suggest that other types of 'institutionalist explanations' that make strong claims about the role of deep institutional histories for economic development (prominent examples being Evans and Rauch, 1999; Acemoglu, 2001)—but rely on rough proxy indicators of historical institutions or measuring contemporary institutions and assuming that these have persisted across decades or centuries—should be more closely scrutinized by using refined and direct measures of different historical institutions. Third, our discussion of the different counterfactual comparisons required to draw inferences about the effects of stateness-first

36

sequences should inform research designs and theorizing in future studies of institutional sequencing.

Our findings also have potential policy implications, as they speak to long-standing debates over development policy and institution building with ramifications for millions of people across the world. For instance, our findings do not support the widespread notion that foreign development aid needs to be channeled towards building state institutions before, e.g., holding multi-party elections, if the ultimate goal is to ensure economic development. The notion that democracy is unsuitable in contexts where weaker state institutions prevail, such as Sub-Saharan Africa, Central Asia or the Middle East, is not uncommon in policy circles. If democratization is anticipated to come at the cost of a stagnant economy, then foreign ministries, international organizations, and aid agencies may find it preferable to continue to support authoritarian strongmen, and postpone the promotion of democracy until the distant future when state institutions are (hopefully) built and order is ensured. If, however, the trade-off between democratization and economic development is non-existent, policymakers may rather want to abandon the current strongman and work harder to promote democracy. Our findings provide some optimistic assessments for organizations, opposition movements, and other actors who want to promote democracy in countries where state administrations are currently far from the Weberian ideal.

References

Acemoglu, D. (2001). The colonial origins of comparative development: An empricial investigation. *The American Economic Review*, 91(5):1369–1401.

Acemoglu, D. (2008). *Introduction to Modern Economic Growth*. Princeton: Princeton University Press.

Acemoglu, D., Johnson, S., Robinson, J. A., and Yared, P. (2005). From education to democracy? *American Economic Association*, 95(2):44..49.

Acemoglu, D., Johnson, S., Robinson, J. A., and Yared, P. (2008). Income and democracy. *American Economic Review*, 98(3):808–842.

Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2014). Democracy does cause growth. Cambridge, MA.: NBER Working Paper No. 20004.

Acemoglu, D. and Robinson, J. A. (2012). *Why Nations Fail. The Origins of Power, Prosperity and Poverty*. London: Profile Books.

Adzera, A., Boix, C., and Payne, M. (2003). Are you being served? political accountability and quality of government. *Journal of Law, Economics and Organization*, 19(2):445–490.

Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(June):155–194.

Andersen, D. and Cornell, A. (2018). The ambiguous effects of democracy on bureaucratic quality. Working Paper.

Andersen, D., Møller, J., Rørbæk, L. L., and Skaaning, S.-E. (2014). State Capacity and Political Regime Stability. *Democratization*, 21(7):1305–1325.

Bäck, H. and Hadenius, A. (2008). Democracy and state capacity: Exploring a j-shaped relationship. *Governance - An International Journal of Policy, Administration, and Institutions*, 21(1):1–24.

Barro, R. J. and Sala-i Martin, X. (2004). *Economic Growth*. The MIT Press, Cambridge, MA.

Berman, S. (2007). How Democracies Emerge: Lessons from Europe. *Journal of Democracy*, 18(1):28–41.

Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation and politics. *American Economic Review*, 99(4):1218–1244.

Besley, T. and Persson, T. (2010). State capacity, conflict and development. *Econometrica*, 78(1):1–34.

Blanchard, P. (2011). Sequence analysis for political science. Working Papers of the Committee on Concepts and Methods, International Political Science Association.

Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1):115–143.

Boix, C., Miller, M., and Rosato, S. (2012). A Complete Data Set of Political Regimes, 1800-2007. *Comparative Political Studies*, 46(12):1523–1554.

Boix, C. and Stokes, S. (2003). Endogenous democratization. *World Politics*, 55(4):517–549.

Bueno de Mesquita, B., Smith, A., Siverson, R. M., and Morrow, J. D. (2003). *The Logic of Political Survival*. MIT Press, Cambridge, Ma.

Carbone, G. and Memoli, V. (2015). Does Democratization Foster State Consolidation? Democratic Rule, Political Order, and Administrative Capacity. *Governance*, 28(1):5–24.

Carothers, T. (2007). The "sequencing" fallacy. *Journal of Democracy*, 18(1):13–27.

Chabal, P. and Daloz, J.-P. (1999). *Africa Works: Disorder as Political Intstrument.* James Currey, Oxford.

Charron, N. and Lapuente, V. (2010). Does democracy produce quality of government? *European Journal of Political Research*, 49(4):443–470.

Coppedge, M., Gerring, J., Altman, D., Bernhard, M., Fish, S., Hicken, A., Kroenig, M., Lindberg, S. I., McMann, K., Paxton, P., Semetko, H., Skaaning, S.-E., Staton, J., and Teorell, J. (2011). Defining and measuring democracy: A new approach. *Perspectives on Politics*, 9(2):247–267.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., Bernhard, M., Fish, S., Glynn, A., Hicken, A., Knutsen, C. H., Krusell, J., Lührmann, A., Marquardt, K. L., McMann, K., Mechkova, V., Olin, M., Paxton, P., Pemstein, D., Pernes, J., Petrarca, C. S., von Romer, J., Saxer, L., Seim, B., Sigman, R., Staton, J., Stepanova, N., and Wilson, S. (2017). Varieties of Democracy, Dataset v7.

Dahl, R. A. (1971). *Polyarchy: Participation and Opposition.* Yale University Press.

D'Arcy, M. and Nistotskaya, M. (2017). State first, then democracy: Using cadastral records to explain governmental performance in public goods provision. *Governance*, 30(2):193–209.

De Long, J. B. and Schleifer, A. (1993). Princes and merchants: European city growth before the industrial revolution. *Journal of Law and Economics*, 36(2):671–702.

Diamond, J. (1997). *Guns, Germs and Steel: The Fates of Human Societies.* W. W. Norton & Company.

Evans, P. and Rauch, J. E. (1999). Bureaucracy and growth: A cross-national analysis of the effects of "weberian" state. *American Sociological Review*, 64(5):748–765.

Fariss, C. J., Crabtree, C. D., Anders, T., Jones, Z. M., Linder, F. J., and Markowitz, J. N. (2017). Latent Estimation of GDP, GDP per capita, and Population from Historic and Contemporary Sources. Working paper.

Fukuyama, F. (2007). Liberalism versus state-building. *Journal of democracy*, 18(3):10–13.

Fukuyama, F. (2012). *The Origins of Political Order*. Profile Books.

Fukuyama, F. (2014a). *Polticial Order and Political Decay: From the Industrial Revolution to the Globalisation of Democracy*. Profiles Books.

Fukuyama, F. (2014b). States and democracy. *Democratization*, 21(7):1326–1340.

Galor, O. (2012). The demographic transition: causes and consequences. *Cliometrica, Jounral of Historical Economics and Econometric Histroy*, 6(1):1–28.

Gerring, J., Bond, P., Barndt, W. T., and Moreno, C. (2005). Democracy and economic growth: A historical perspective. *World Politics*, 57(3):323–364.

Haber, S. and Menaldo, V. (2011). Do natural resources fuel authoritarianism? a reappraisal of the resource curse. *American Political Science Review*, 105(1):1–26.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1):25–46.

Hanson, J. (2015). Democracy and state capacity: Complements or substitutes? *Studies in Comparative International Development*, 50(3):304–330.

Hobson, C. (2012). Liberal democracy and beyond: Extending the sequencing debate. *International Political Science Review*, 33(4):441–454.

Honaker, J. and King, G. (2010). What to do about missing values in time-series cross-section data. *American Journal of Political Science*, 54(3):561–581.

Huntington, S. P. (1968). *Political Order in Changing Societies*. A Yale Paperbound, Political Science. Yale University Press.

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

Iacus, S. M., King, G., and Porro, G. (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis*, 20(1):1–24.

Inglehart, R. and Welzel, C. (2005). *Modernization, Cultural Change, and Democracy: The Human Development Sequence*. Cambridge University Press.

Jerven, M. (2013). *Poor Numbers: How We Are Misled by African Development Statistics and What to Do about It*. Cornell University Press.

Kaufmann, D., Kray, A., and Mastruzzi, M. (2010). The Worldwide Governance Indicators. Methodology and Analytical Issues. World Bank Policy Research Working Paper 5430.

Kennedy, R. (2010). The contradiction of modernization: A conditional model of endogenous democratization. *The Journal of Politics*, 72(3):785–798.

Knack, S. and Keefer, P. (1995). Institutions and economic performance: Cross-country tests using alternative measures. *Economics and Politics*, 7(3):207–227.

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

Knutsen, C. H., Teorell, J., Cornell, A., Gerring, J., Gjerløw, H., Skaaning, S.-E., Wig, T., Ziblatt, D., Marquardt, K., Pemstein, D., and Seim, B. (2017). Introducing the Historical Varieties of Democracy Dataset: Patterns and Determinants of Democratization in the Long 19th Century. Working Paper.

Lake, D. A. and Baum, M. A. (2001). The invisible hand of democracy: Political control and the provision of public services. *Comparative Political Studies*, 34(6):587–621.

Landes, D. (1998). *The Wealth and Poverty of Nations: Why Some are So Rich and Some So Poor*. W. W. Norton & Company.

Mankiw, G. N., Romer, D., and Weil, D. N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2):407–437.

Mansfield, E. D. and Snyder, J. (1995). Democratization and the danger of war. *International Security*, 20(1):5–38.

Mansfield, E. D. and Snyder, J. (2005). *Electing to Fight. Why Emerging Democracies Go to War*. MIT Press.

Mansfield, E. D. and Snyder, J. L. (2007). The sequencing "fallacy". *Journal of democracy*, 18(3):5–10.

Marshall, M. G. and Jaggers, K. (2007). Polity IV Project: Dataset Users' Manual.

Marshall, T. and Bottomore, T. (1949). *Citizenship and Social Class*. London: Pluto.

Mazzuca, S. and Munck, G. L. (2014). State or democracy first? Alternative perspectives on the state-democracy nexus. *Democratization*, 21(7):1221–1243.

Miller, M. K. (2015). Democratic Pieces: Autocratic Elections and Democratic Development since 1815. *British Journal of Political Science*, 45(3):501–530.

Møller, J. (2015). The medieval roots of democracy. *Journal of Democracy*, 26(3):110–123.

Møller, J. and Skaaning, S.-E. (2013). Regime types and democratic sequencing. *Journal of Democracy*, 24(1):142–155.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

Papaioannou, E. and Siourounis, G. (2008). Democratization and growth. *Economic Journal*, 118(532):1520–1551.

Pemstein, D., Marquardt, K. L., Tzelgov, E., Wang, Y.-t., Krussel, J., and Miri, F. (2017). The Varieties of Democracy Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data. Gothenburg: V-Dem Working Paper No. 21.

Pomeranz, K. (2000). *The Great Divergence: China, Europe and the Making og the Modern World Economy*. Princeton University Press.

Przeworski, A. and Limongi, F. (1997). Modernization: Theories and Facts. *World Politics*, 49(2):155–183.

Robinson, J. A. and Verdier, T. (2013). The Political Economy of Clientelism. *Scandinavian Journal of Economics*, 115(2):260–291.

Rodrik, D., Subramanian, A., and Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth*, 9(2):131–165.

Roodman, D. (2009). How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata. *Stata Journal*, 9(1):86–136.

Rothstein, B. (2011). *The Quality of Government: Corruption, Social Trust, and Inqeuality in International Perspective.* Chicago University Press.

Sachs, J. (2005). *The End of Poverty.* Penguin Books.

Seeberg, M. B. (2015). *The Contingent Effect of Authoritarian Elections.* PhD thesis, Aarhus University.

Shefter, M. (1993). *Political Parties and the State.* Princeton University Press.

Spolaore, E. (2013). What Is European Integration Really About? A Political Guide for Economists. *Journal of Economic Perspectives*, 27(3):125–144.

Teorell, J., Coppedge, M., Skaaning, S.-E., and Lindberg, S. I. (2016). Measuring electoral democracy with v-dem data: Introducing a new polyarchy index. Gothenburg: V-Dem Working Paper No. 25.

Tilly, C. (1990). *Coercion, Capital and European States, A.D. 990 - 1992.* Wiley-Blackwell.

van Ham, C. and Seim, B. (2017). Strong states, weak elections? how state capacity in authoritarian regimes conditions the democratizing power of elections. *International Political Science Review*, OnlineFirst.

Wang, E. H. and Xu, Y. (2018). Awakening leviathan: The effect of democracy on state capacity. *Research & Politics*, 5(2):2053168018772398.

Weber, M. (2002). *The Protestant Ethic and "The Spirit of Capitalism".* Penguin, London.

Wintrobe, R. (1998). *The Political Economy of Dictatorship.* New York: Cambridge University Press.

Zakaria, F. (2003). *The Future of Freedom. Illiberal Democracy Home and Abroad.* W. W. Norton and Company.

Online Appendices for "Stairways to Denmark"

A Additional panel regressions

In this Appendix we present additional panel specifications than those presented in the paper. We start out by showing results for models using annualized GDP per capita growth as the dependent variable instead of Ln GDP per capita. The eight leftmost columns of Table 1 presents split-sample tests—separating on the median sample value on the impartial and rule-following bureaucracy (*v2clrscpt*) measure.[9] The five rightmost columns contain interaction tests, interacting different measures of democracy and state capacity. Since GDP per capita growth is continuous and close to normally distributed, all models in Table 1 are OLS with country- and year fixed effects with country-year as unit of analysis. As we remark in the paper, controlling for country-fixed effects is important since it accounts for different time-invariant features, such as geographic or cultural factors, that may influence both the institutional make-up of a country and growth (e.g., Acemoglu et al., 2008). The year-fixed effects account for (linear or more complex) global time trends and short-term shocks to growth and institutional features. We also control for ln GDP per capita to account for convergence effects (e.g., Barro and Sala-i Martin, 2004). Errors are clustered by country to mitigate concerns about serial correlation influencing the uncertainty estimates.

Model 1 ('low-capacity sample') and Model 2 ('high-capacity sample'), regresses five-year forward lagged growth outomes (i.e., GDP per capita growth from $t+4$ to $t+5$, on Polyarchy and covariates at t. Thus, we here focus on shorter-term effects than for most models presented in the paper. The Polyarchy coefficient is positive in the high-capacity sample and negative in the low-capacity sample, in line with predictions from the 'stateness-first' argument. However, as for the main specifications in the paper, Polyarchy fails to achieve

---

[9]This median (0.0101) concerns all observations entering Models 1 and 2 combined. The first and third quartile values—which demarcate the samples in Models 5 and 6—are -1.014 and 0.978, respectively.

conventional levels of significance in any of the two models. Models 1–2 only measures GDP per capita growth for one given year, and a five-year lag may be insufficient to capture possible longer-term developmental effects of democracy (although extant work suggests that even slags of about three years may be sufficient; see Papaioannou and Siourounis, 2008). Thus, Models 3–4 use as dependent variable the average annual growth rate over the entire twenty-year period after the independent variables are measured.[10] The Polyarchy coefficients is fairly comparable size in the low-capacity samples, but attenuated and even further from conventional levels of statistical significance in the high-capacity sample.

Models 5–6 replicate Models 1–2, but splitting on the lowest- (Model 5) and highest quartile observations (Model 6). While the point estimates suggest a positive effect in the very high capacity sample and a negative in the very low capacity sample, Polyarchy is not significant even at the 10% level in any specification. Models 7–8 again replicate Models 1–2, with the dichotomous (BMR) measure Boix et al. (2012). The relationship between democracy and growth is estimated to be positive, but is statistically insignificant in both samples. In sum, the split sample tests do not support the notion that democracy is more tightly linked to economic development in high-capacity states than in low capacity states (nor *vice versa*, cf. Knutsen, 2013).

Yet, split sample specifications are relatively inefficient for testing the hypothesis that the effect of democracy hinges on the level of state capacity, estimating the relationship using subsets of observations. Thus, Models 9–13 are interaction specifications run on the full set of observations. Model 9 interacts Polyarchy with *v2clrspct* and uses growth in $t + 5$ as dependent variable. Model 10 includes the same interaction term, but uses average annual growth over the subsequent twenty years as dependent variable. Model 11 adds additional covariates to Model 10 that might affect both growth and institutional features, namely

---

[10]The samples are truncated due to the twenty year lags, but we maintain the same cut-off on *v2clrscpt* (0.0101) to ensure comparable classifications.

Table A.1 – Effects of democracy on GDP per capita growth, for different levels of state capacity

Sample:	LC	HC	LC	HC	VLC	VHC	LC	HC	Full	Full	Full	Full	Full
	1	2	3	4	5	6	7	8	9	10	11	12	13
Polyarchy	-0.836	1.291	-0.729	0.654	-0.417	2.666			0.169	-0.566	-0.597	2.275	-0.311
	(0.981)	(0.842)	(0.845)	(0.607)	(2.724)	(1.829)			(0.665)	(0.600)	(0.555)	(2.413)	(0.545)
BMR regime							0.111	0.078					
							(0.377)	(0.270)					
Impartial pub. adm.									0.071	0.022	0.108		
									(0.103)	(0.092)	(0.118)		
Polyar. X Imp. adm.									0.225	0.386	0.215		
									(0.212)	(0.236)	(0.189)		
Corruption												1.028	
												(1.220)	
Polyarchy X Corr.												-4.071	
												(4.769)	
Clientelism													0.110
													(0.109)
Polyarchy X Client.													0.226
													(0.254)
Ln GDP pc	-0.897***	-2.286***	-0.827***	-2.004***	-0.863***	-3.516***	-1.098***	-2.012***	-1.270***	-1.147***	-1.232***	-1.194***	-1.113***
	(0.235)	(0.436)	(0.260)	(0.362)	(0.301)	(0.886)	(0.230)	(0.332)	(0.254)	(0.254)	(0.232)	(0.225)	(0.249)
Ln Population											-0.149		
											(0.238)		
Civil war											0.125		
											(0.182)		
Resource dep.											-0.045***		
											(0.011)		
Year dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Country dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	8719	8727	7607	7195	4367	4382	7247	8077	17402	14773	9117	14417	13901
R^2	0.042	0.033	0.297	0.420	0.058	0.032	0.151	0.073	0.027	0.324	0.411	0.332	0.326

Notes: *p<0.1; **p<0.05; ***p<0.01. Standard errors in parentheses. Constant, country- and year dummies omitted. All models are OLS with country- and year fixed effects with errors clustered by country. LC: Low capacity. HC: High capacity. Low vs High capacity is determined relative to median-sample value on IPA (*v2clrscpt* from V-Dem) for full sample for specification corresponding to Models 1 and 2. VLC: Very Large Capacity (lowest quartile on IPA). VHC: Very High Capacity (highest quartile on IPA). GDP per capita growth (DV) is measured in $t+5$ in Models 1–2 and 5–9 and as average growth rate from $t+1$ to $t+20$ in Models 3–4 and 10–13.

iii

(Ln) population size, civil war, and natural resource dependence. Models 12 and 13 also use twenty-year growth as dependent variable, but employ, respectively, measures of corruption (*v2x_corr*) and clientelism (*v2psprlnks*) as proxies for state capacity. None of these models yield evidence of any systematic interaction between democracy, on the one hand, and an impartial public administration, corruption, and clientelism, on the other.

Yet, the models in Table A.1 arguably do not offer precise tests of the effects of experiencing the specific *sequence* of state building before democratic transition. Thus, we shift to panel regressions considering the existing or past level of state capacity at the point of transition, using the BMR measure. Models 1–2, Table A.2 compare years of democratic transitions to all non-democratic country-years in, respectively, low and high-capacity states, using the median sample-value on *v2clrscpt* from Table A.1 as cut-off. The dependent variables is average annual growth rate over the following 20 years. Models 3–4 are similar split sample tests, but where average growth rate of the preceding 20 years is controlled for (thus resembling a difference–in–differences set-up).

One of these tests suggest that democratization has a systematically stronger link to growth in high-capacity states; the point estimates are larger for the low-capacity samples, but are highly uncertain in all specifications. Models 5–6, which control for growth over the preceding 20 years, are interaction specifications where the democratic transition dummy is multiplied with *v2clrscpt*. Models 7–8 replicate Models 5–6, but substitute *v2clrscpt* score at time of democratic transition with the historical average in the ten years preceding the transition. Models 9–10 replicate Model 6, but use the measures of corruption and clientelism, respectively, instead of *v2clrscpt*. None of these specifications suggest that democratic transitions are more detrimental to economic growth in contexts where the state has (or has had) a comparatively low level of capacity at (or before) transition. All interaction terms are insiginificant, even at 10%.

Finally, many contributions proposing that not building state capacity before opening up

Table A.2 – Robustness tests

Sample:	LC	HC	LC	HC	Full	Full	Full	Full	Full	Full
	1	2	3	4	5	6	7	8	9	10
BMR transition	0.130	-0.224	0.077	-0.118	-0.067	-0.100	-0.017	-0.076	0.101	-0.086
	(0.262)	(0.211)	(0.263)	(0.216)	(0.202)	(0.198)	(0.169)	(0.169)	(0.968)	(0.192)
Impartial public admin.					0.066	0.043				
					(0.122)	(0.114)				
Democratiz. X Imp. adm.					0.077	0.093				
					(0.147)	(0.146)				
Past 10 yrs imp adm							0.028	0.001		
							(0.144)	(0.132)		
Past imp. adm. X Democratiz.							0.038	0.033		
							(0.166)	(0.166)		
Corruption									0.791	
									(1.170)	
Democratization X Corr.									-0.450	
									(1.885)	
Clientelism										0.027
										(0.132)
Democratization X Client.										-0.081
										(0.153)
Ln GDP p.c.	-0.723***	-1.668***	-0.807***	-2.002***	-0.753***	-0.851***	-0.742**	-0.804***	-0.869**	-0.927***
	(0.242)	(0.457)	(0.271)	(0.504)	(0.280)	(0.305)	(0.298)	(0.307)	(0.335)	(0.327)
Growth past twenty years			0.024	0.044		0.011		0.011	0.002	0.006
			(0.096)	(0.084)		(0.078)		(0.073)	(0.081)	(0.081)
Year dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Country dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	5621	3582	5086	2940	8571	7531	8294	7426	7208	6976
R^2	0.317	0.493	0.323	0.552	0.331	0.350	0.324	0.344	0.351	0.357

Notes: *$p<0.1$; **$p<0.05$; ***$p<0.01$. Standard errors in parentheses. Constant, country- and year dummies omitted. All models are OLS with country- and year fixed effects with errors clustered by country. LC: Low capacity. HC: High capacity. Low vs High capacity is determined relative to median-sample value on IPA (*v2clrsrcpt* from V-Dem) for full sample for specification corresponding to Models 1 and 2, Table 1.

for mass politics has detrimental effects mainly focuses on expansions of the suffrage (from a narrow set of elites) to covering broader segments of the population (see, e.g., Fukuyama, 2014a). In order to test, more specifically, whether extensive suffrage is systematically less conducive to growth in contexts of low state capacity, we re-ran the OLS specifications from Table 1 in the paper—thus using Ln GDP per capita measured 20 years after the covariates as dependent variable—but substituting the Polyarchy index of democracy with a variable from V-Dem measuring the share of adults with voting rights (*v2x_suffr*). Since the varible is measured as a share, it extends from 0 to 1, and the distribution of the variable in our sample is shown in Figure A.1.
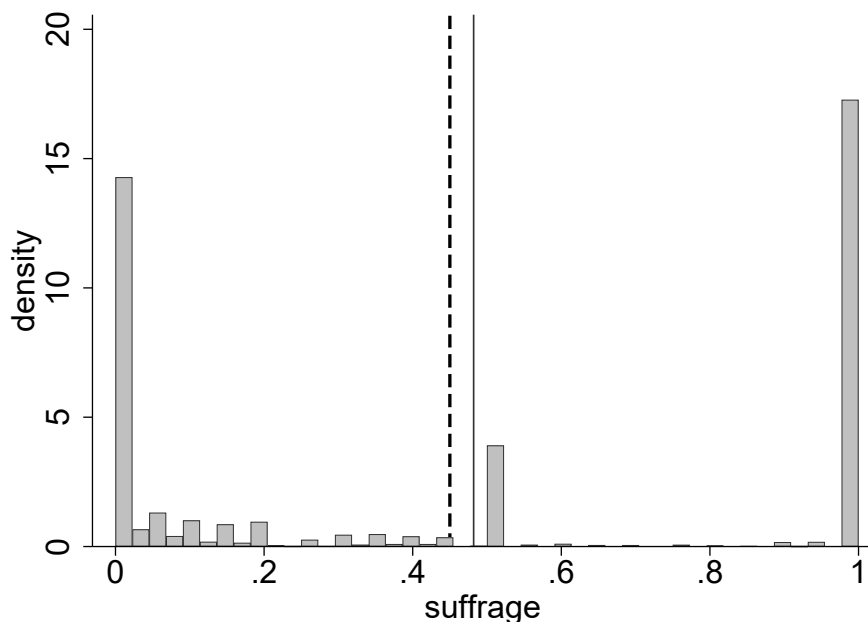
Figure A.1 – Distribution of suffrage, 1789–2017.
(*Dotted line represents the median; solid line indicates the mean.*)

The results for suffrage are reported in Table A.3. For the split-sample specifications— where the sample is split according to the median value on our measure of impartial and rule-following administration—the point estimates actually suggest a positive relationship between suffrage and income measured 20 years after in the low-capacity sample (Model 1)

vi

Table A.3 – Tests using share of adult population with suffrage as measure for democracy, with Ln GDP p.c. as dependent variable.

	1	2	3	4	5
Sample:	LC	HC	Full	Full	Full
Dep var measured in year	$t+20$	$t+20$	$t+20$	$t+20$	$t+40$
Suffrage	0.086	-0.016	0.041	-0.035	0.137*
	(0.109)	(0.068)	(0.064)	(0.094)	(0.082)
Impartial public administration			0.002	-0.017	-0.013
			(0.019)	(0.030)	(0.035)
Suffrage X IPA			0.029	0.042	0.040
			(0.020)	(0.028)	(0.035)
Ln GDP p.c.	0.848***	0.667***	0.792***	0.790***	0.679***
	(0.046)	(0.059)	(0.044)	(0.043)	(0.070)
Civil war				-0.005	
				(0.039)	
Resource dependence				-0.007***	
				(0.002)	
Ln Population				0.017	
				(0.048)	
Year dummies	Y	Y	Y	Y	Y
Country dummies	Y	Y	Y	Y	Y
N	7983	8254	15677	10385	12391
R^2	0.749	0.889	0.839	0.849	0.774

Notes: *p<0.1; **p<0.05; ***p<0.01. Standard errors in parentheses. Ln GDP p.c.
is dependent variable. OLS with country- and year fixed effects with errors clustered by country.
LC: Low capacity. HC: High capacity. Low vs High capacity is determined relative to median-sample
value on IPA for full sample for specification orresponding to Models 1 and 2, Table 1.

and a negative relationship in the high-capacity sample (Model 2). However, none of the coefficients are close to being statistically significant. Likewise, neither of the interaction specifications in Models 3 (benchmark), 4 (additional controls), or 5 (income measured 40 years after covariates) display a clear interaction between suffrage and state capacity. In conclusion, similar to the tests on broader measures of electoral democracy, tests focusing on suffrage do not provide any support for the stateness-first argument.

B   Alternative matching models

In this appendix, we display additional matching models mentioned, but not reported, in the paper. More specifically, we first report a table displaying variations of Model 1, Table 2 (with Ln GDP p.c. measured in $t + 20$ as outcome). In these models, which are estimated with Coarsened Exact Matching (Iacus et al., 2012), we alter the operationalization used for coding democratic transitions, testing median Polyarchy (applying the 10-year rule discussed in the paper, requiring that the country remained democratic at least 10 years after the initial transition), transition on the dichotomous BMR measure (with and without the ten-year rule), as well as a dummy registering whether or not a majority of the adult population has suffrage rights (with and without the ten-year rule).

The overall pattern is similar to the models reported in the paper, namely one of null results. But there is, in fact, a positive relationship (significant at 5%) between state capacity at the time of democratization when using the suffrage measure and applying the 10-year rule in Model 5. While this result is in line with the stateness-first argument, we note that the result is sensitive to minor tweaks to the specification, also when using suffrage expansion as the demarcator of democratization episodes (such as not employing the 10-year rule, in Model 4). More generally, this result contrasts with results from the numerous other tests that we have conducted. As such, this single result does not alter our main conclusion regarding the (lack of) validity of the stateness-first argument when it comes to explaining economic development.

We further probed the sensitivity of the results by employing a very different matching approach. More specifically, we re-estimate models 3 - 5 from table 2 using entropy balancing (EB) (Hainmueller, 2012) instead of CEM. In brief, EB uses a re-weighting scheme that ensures adjustments in inequalities between treatment and control groups with respect to the first, second, and potentially higher moments of the covariate distribution. Recent research indicates that EB, on a number of dimensions, outperforms traditional approaches to matching based estimating the propensity score by maximum likelihood (see, e.g., Zhao and Percival, 2016; Harvey et al., 2017). Furthermore, while EB requires a dichotomous treatment variable, it does not demand the categorization of the covariates, which often involves an element of arbitrariness, including in our setting. Thus, there are clear benefits to using EB also when it comes to analyzing the stateness-first argument.

Nevertheless, EB cannot replicate that tests performed by Models 1 and 2 in Table 2 of the paper. With CEM, observations that does not find a match within the bins of categories across treatment and control groups are discarded. In effect, countries experiencing democratization are matched with countries that are equal *within that same year/time period*, and all other comparisons are discarded. With EB, however, observations are not discarded. Without further restrictions in our panel set-up, country `i` experiencing democratization in year `t` would, effectively, be compared with itself in `t-k` and `t+k`. Furthermore, country `i` in year `t` would also be compared to country `j`, which is not experiencing democratization, perhaps because it is already democratic, in year `t`, year `t-k`, and year `t+k`. Some of these issues could be alleviated, for example by removing all observations from a country after democratization or all observations X years before and after democratization. Preliminary tests (not reported) suggest that results are very sensitive to such specification choices, but overall providing little evidence in support of the stateness-first argument. Yet, since there are multiple issues and arbitrariness involved with selecting a specific subset of years for countries that do not experience democratization as well as a sensible subset of autocratic

ix

Table B.1 – CEM Balancing – Alternatives to Model 1 in Table 2

	(1)	(2)	(3)	(4)	(5)
			Ln GDP 20 years after democratization		
Bureaucracy-level at democratization	−0.04	0.09	0.07	−0.22	0.17*
	(0.15)	(0.05)	(0.05)	(0.28)	(0.07)
Ln GDP at democratization	1.09	1.03***	0.80***	0.45	0.97***
	(0.55)	(0.20)	(0.16)	(0.47)	(0.16)
Ln year at democratization	−12.29	−1.53	9.87	73.80	57.60
	(26.92)	(13.03)	(12.42)	(132.47)	(29.18)
Ethnic fractionalization at democratization	0.83	−0.09	−0.14	−1.19	−1.22*
	(1.51)	(0.31)	(0.30)	(3.46)	(0.53)
Constant	91.63	11.50	−73.08	−551.00	−434.28
	(201.86)	(99.25)	(94.26)	(993.51)	(220.54)
Regional Dummies	Yes	Yes	Yes	Yes	Yes
Subclass Dummies	Yes	Yes	Yes	Yes	Yes
Democracy var:	10 year Polyarchy	BMR	10 year BMR	Suffrage	10 year suffrage
Comparison:	Demo. vs Aut.	Demo. vs Aut.	Demo. vs Aut.	Demo. vs Aut.	Demo. vs Aut.
Democratization rule:	After 20 years of Aut.	After 20 years of Aut.	After 20 years of Aut.	After 20 years of Aut.	After 20 years of Aut.
Observations	26	31	68	19	78
R²	0.81	0.86	0.92	0.85	0.90
Adjusted R²	0.63	0.79	0.88	0.61	0.83

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors in parentheses.

x

years for a country that later experiences democratization, we do not focus on these results. (These issues are avoided with CEM, because the sample is effectively restricted as a part of the matching procedure.)

Table B.2 – Entropy Balancing

	(1)	(2)	(3)
	Ln GDP 20 years after democratization	Ln GDP in 2004	
Bureaucracy-level	−0.04	0.04	−0.04
	(0.04)	(0.06)	(0.07)
Ln GDP pc at democratization	0.87***	0.48***	0.63***
	(0.06)	(0.09)	(0.11)
Year	0.005**	−0.002	−0.003
	(0.001)	(0.002)	(0.002)
Ethnic fractionalization	0.05	−0.16	−0.50
	(0.18)	(0.35)	(0.45)
Constant	−7.33**	8.76**	9.26*
	(2.78)	(3.00)	(3.57)
Regional Dummies	Yes	Yes	Yes
Comparison:	Democratizers	Democratizers	Democratizers
Democratization rule:	After 20 years of Aut.	First only	First only
Observations	99	80	46
R^2	0.85	0.85	0.88
Adjusted R^2	0.83	0.83	0.86

Notes: *$p<0.05$; **$p<0.01$; ***$p<0.001$. Standard error in parentheses.

But, this is not a problem for replicating Models 3-5 from Table 2 by using EB. These models already restrict the sample to country-years when democratization happens. The EB-versions of these models are thus presented in Table B.2. In all three models, bureaucracy-level is matched on GDP per capita, year, ethnic fractionalization, and regional dummies. In Models 1 and 3, the coefficient for impartial and rule-following bureaucracy is negative, while it iss positive in Model 2. In all models the association is highly uncertain, and statistically insignificant at conventional levels.

C   Alternative sequencing models

In this appendix, we include a table with some of the different robustness tests that we ran on our sequencing specifications. Here, we present results altering the operationalizations of democratic transition as well as the transition from low to high capacity. In the benchmark specifications presented in the paper, we opted for splitting both Polyarchy and the impartial and rule-following administration measures at their medians, ensuring fairly balanced sub-samples of observations in different 'institutional states' (low democracy–low capacity, high democracy–low capacity, etc.) and also aiming to set thresholds that appropriately captures extant arguments about stateness-first sequences. For instance, we highlighted the importance of not setting the bar for 'democratization' too high, as existing studies mainly focus on the introduction of elementary aspects of mass politics (e.g., opening up for multi-party competition or extending the franchise beyond elites).

Overall, the results in Table C.1 show that the statistical significance of the relationship between stateness-first sequences and subsequent economic development is sensitive to such opertationalization choices. While results are quite similar to our benchmark—which finds a negative and significant stateness-first coefficient—when using a mean cut-off rule for Polyarchy and the impartial public administration measure (see also Figure C.1), the results change for alternative measurement strategies. For instance, the relationship is indistinguishable from zero once we use the BMR measure to identify democratization episodes rather than V-Dem's Polyarchy measure. However, the sign of the relationship is consistently negative. Hence, what is also consistent across these (and other) specifications is that there is no evidence that stateness-first sequences correspond with higher, subsequent economic growth.

Finally, we tested similar sequencing specifications, but focusing on suffrage instead of overall democraitzation. Once again, we leverage the extensive coding of suffrage from V-

Dem, and test different ways of operationalizing high/low suffrage (mean and median in the 1789–2016 sample). While the coefficient for the resulting stateness-first dummy—registering sequences where countries achieve high state capacity before extensive suffrage—is positive in the twospecifications using the median values as cut-offs. Yet, the stateness-first dummy never achieves statistical significance even at the 10% level. In contrast, the models that employ the mean values as cut-offs report a negative and (weakly) significant stateness first dummy. Thus, although results are less clear when focusing on suffrage rather than composite democracy measures, there is no evidence supporting the stateness-first argument from any of the analyses.
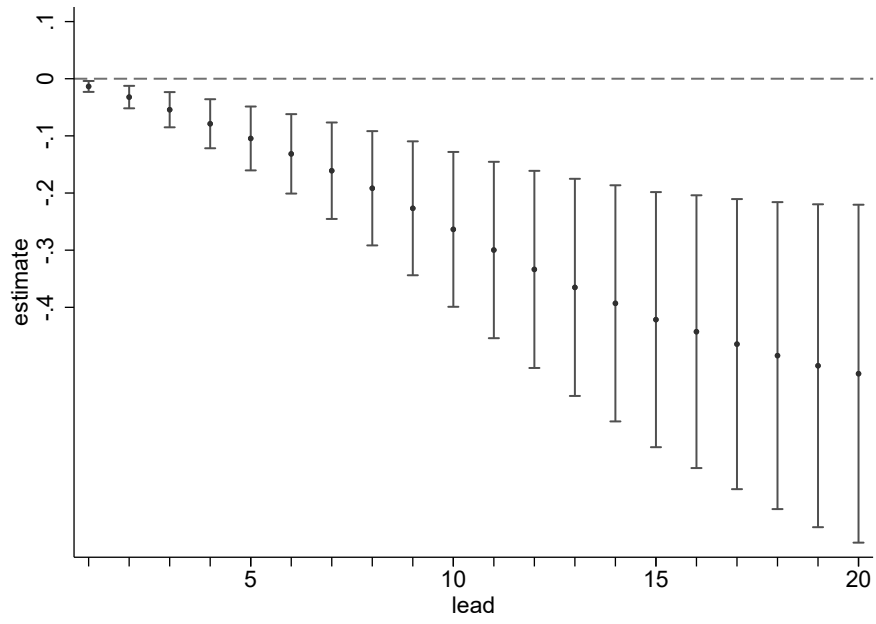
Figure C.1 – Estimates for 'stateness-first' transitions, for varying leads.
(*Cut-offs for institutional variables based on mean values in sample.*)

Table C.1 – Alternative measurements of democracy and state capacity when operationalizing stateness-first transitions

	(1)	(2)	(3)	(4)	(5)
	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)
Democracy measure	Polyarchy	BMR	BMR	Polyarchy	Polyarchy
Cut-off rule for continuous measures of state cap./democ.	mean	median	mean	top two-thirds	top third
Stateness-first	-0.516***	-0.49	-0.085	-0.514***	-0.169
	-0.18	-0.671	-0.455	-0.154	-0.25
Low dem., high cap.	-0.043	0.021	-0.01	0.003	-0.065
	-0.06	-0.044	-0.049	-0.069	-0.065
High dem., low cap.	-0.036	0.039	0.03	-0.077*	-0.028
	-0.042	-0.045	-0.041	-0.046	-0.031
High dem., high cap.	-0.003	0.035	0.002	-0.024	-0.004
	-0.053	-0.049	-0.058	-0.065	-0.052
Σlow dem., low cap.	0	0	0	0.001	0
	-0.001	-0.001	-0.001	-0.002	-0.001
Σlow dem., high cap.	0.002	0.004**	0.004***	0.003	0.003*
	-0.002	-0.002	-0.001	-0.003	-0.002
Σhigh dem., low cap.	0	0.008***	0.007***	0	0.002
	-0.002	-0.002	-0.002	-0.002	-0.002
Σhigh dem., high cap.	0.003**	0.002	0.002	0.001	0.002**
	-0.001	-0.001	-0.001	-0.001	-0.001
Ln GDP p.c.	0.770***	0.761***	0.761***	0.779***	0.777***
	(0.043)	(0.044)	(0.042)	(0.048)	(0.044)
Ln GDP p.c. at transition	0.063***	0.067	0.022	0.058***	0.023
	-0.023	-0.086	-0.059	-0.022	-0.035
Polyarchy	-0.150	-0.221**	-0.217*	-0.099	-0.166
	(0.133)	(0.111)	(0.113)	(0.109)	(0.126)
Impartial administration	0.047	0.026	0.036	0.019	0.044
	(0.028)	(0.027)	(0.029)	(0.028)	(0.027)
Polyarchy x Impartial adm.	-0.040	0.019	0.005	0.019	-0.018
	(0.052)	(0.049)	(0.050)	(0.047)	(0.048)
Civil War	0.011	0.005	0.002	0.011	0.011
	(0.038)	(0.038)	(0.038)	(0.039)	(0.039)
Resource dependence	-0.007***	-0.006***	-0.006***	-0.007***	-0.007***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
Ln population	0.006	-0.002	0.001	0.008	0.008
	(0.045)	(0.040)	(0.038)	(0.040)	(0.045)
Intercept	1.546***	1.677***	1.660***	1.387**	1.431**
	(0.569)	(0.578)	(0.527)	(0.546)	(0.581)
Year fixed effects	yes	yes	yes	yes	yes
N	10073	10071	10071	10073	10073

Notes: *p<0.1; **p<0.05; ***p<0.01. OLS regressions with
Ln GDP p.c. in t+20 as dependent variable. Year dummies omitted.
Country-year is unit. Errors are clustered by country.

Table C.2 – Suffrage and state capacity, operationalizing stateness-first transitions

	(1) b/(se)	(2) b/(se)	(3) b/(se)	(4) b/(se)
Cut-off rule for continuous measures of state cap./suffr.	median	median	mean	mean
Stateness-first	0.168	0.354	-0.356*	-0.605*
	(0.251)	(0.348)	(0.201)	(0.320)
Low dem., high cap.	0.059*	-0.032	0.057	-0.079
	(0.034)	(0.044)	(0.048)	(0.085)
High dem., low cap.	0.095**	0.143**	0.015	-0.057
	(0.043)	(0.069)	(0.046)	(0.078)
High dem., high cap.	0.119**	0.196**	0.061	-0.056
	(0.049)	(0.081)	(0.053)	(0.094)
Σlow dem., low cap.	-0.001	-0.001	-0.001	0
	(0.001)	(0.001)	(0.001)	(0.001)
Σlow dem., high cap.	-0.001	-0.002	0.001	0.003
	(0.001)	(0.002)	(0.001)	(0.002)
Σhigh dem., low cap.	0	0.002	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Σhigh dem., high cap.	0.002**	0.004***	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Ln GDP p.c.	0.788***	0.782***	0.796***	0.797***
	(0.044)	(0.040)	(0.043)	(0.039)
Ln GDP p.c. at transition	-0.009	-0.027	0.039	0.064
	(0.034)	(0.047)	(0.025)	(0.041)
Suffrage		-0.180		-0.022
		(0.114)		(0.121)
Impartial administration		-0.006		0.025
		(0.030)		(0.036)
Suffrage x Impartial adm.		0.029		-0.011
		(0.031)		(0.036)
Civil War		-0.007		-0.002
		(0.037)		(0.039)
Resource dependence		-0.007***		-0.007***
		(0.002)		(0.002)
Ln population		-0.001		0.012
		(0.041)		(0.036)
Intercept	1.194***	1.354***	1.250***	1.399***
	(0.296)	(0.525)	(0.297)	(0.476)
Year fixed effects	yes	yes	yes	yes
N	15627.000	10385.000	15627.000	10385.000

Notes: *p<0.1; **p<0.05; ***p<0.01. OLS regressions with
Ln GDP p.c. in t+20 as dependent variable. Year dummies omitted.
Country-year is unit; Errors are clustered by country.

D Preliminary analysis on alternative outcomes

As we discuss in the paper, various formulations of the stateness-first argument have been proposed to explain a range of outcomes. As we also note in the paper, our theoretical discussion on the strong assumptions of the stateness-first argument, and our discussions on how to appropriately consider counter-factual scenarios when designing tests, pertain to a number of these outcomes. However, the more specific design of the most appropriate tests (e.g., in terms of choosing the right estimator and the lag-specification on the dependent variable) could certainly differ for other reasons across outcomes. Empirically, the focus of our paper was put squarely on economic development, testing specifications that assess both the medium- and long-term consequences for economic growth. We thus reiterate that the fact that our findings do not corroborate the stateness-first argument for economic development, should not lead to strong conclusions on the irrelevance of stateness-first sequences for *any* outcome. Indeed, this word of caution is borne out by the preliminary analysis contained in this appendix, where some tests—for instance when civil war onset is the dependent variable—reveal patterns that are more in line with the stateness-first argument than the results we find for economic development. That being said, also the majority of models that we have tested for alternative, relevant outcomes discussed in the extant literature fail to provide clear evidence in line with the stateness-first argument.

While it is infeasible for us to carefully treat each and every interesting outcome potentially affected by stateness-first sequences within the scope of one article, we present some preliminary results for a handful of outcomes in the tables below. For the sake of easing comparability, we present these results following the same model specification templates as we do for economic development in the paper. This choice—in addition to the absence of numerous robustness tests—is also why we consider these results to be preliminary; we are likely not identifying the optimal specification for each outcome. For instance, when study-

ing civil war onset, we should perhaps have used a logit estimator rather than a fixed effects OLS specification and the 20-year lags is perhaps too long when assessing the relationship with civil war. Thus, we underscore that more careful empirical studies of the stateness-first argument for alternative outcomes should be considered fruitful topics for future research.
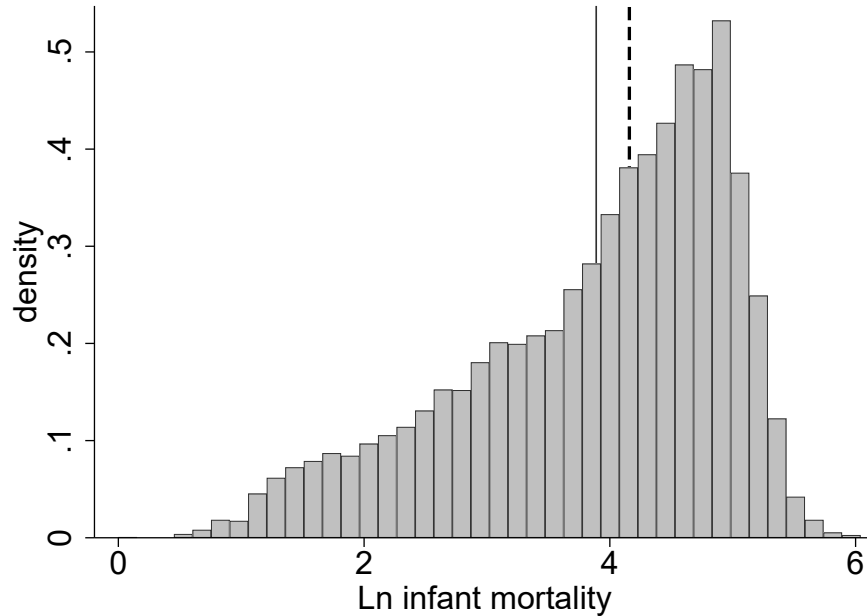

Figure D.1 – Distribution Ln infant mortality rate, 1900–2015.
(*Dotted line represents the median; solid line indicates the mean.*)

All this being said, Table D.1 presents fixed effects OLS specifications akin to those in Models 1–3, Table 1 for, respectively, Ln infant mortality rate, Programmatic vs. clientelist linkages between candidates/parties and citizens/voters, and, civil war onset. Data for all these dependent variables are taken from the V-Dem dataset, version 7.1. To reiterate, these specifications test the core assumption that democracy functions better in producing 'good outcomes' in contexts of high state capacity than in contexts of low, controlling for initial income level as well as country- and year fixed effects.

The findings vary according to the outcome. For Ln infant mortality rate—the perhaps most widely used proxy for human development in the literature—we actually replicate the
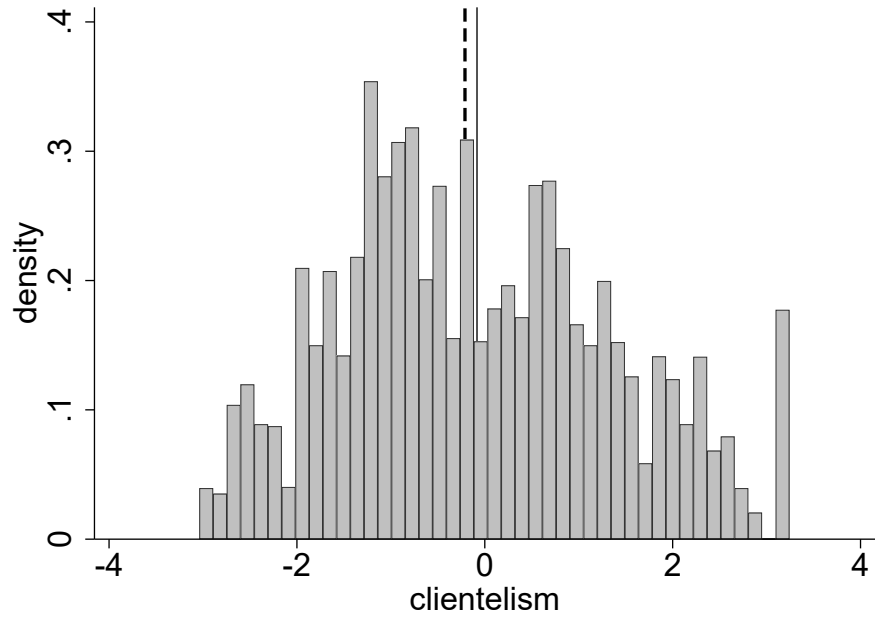
Figure D.2 – Distribution programmatic vs clientelist linkages, 1789–2016.
(*Dotted line represents the median; solid line indicates the mean.*)
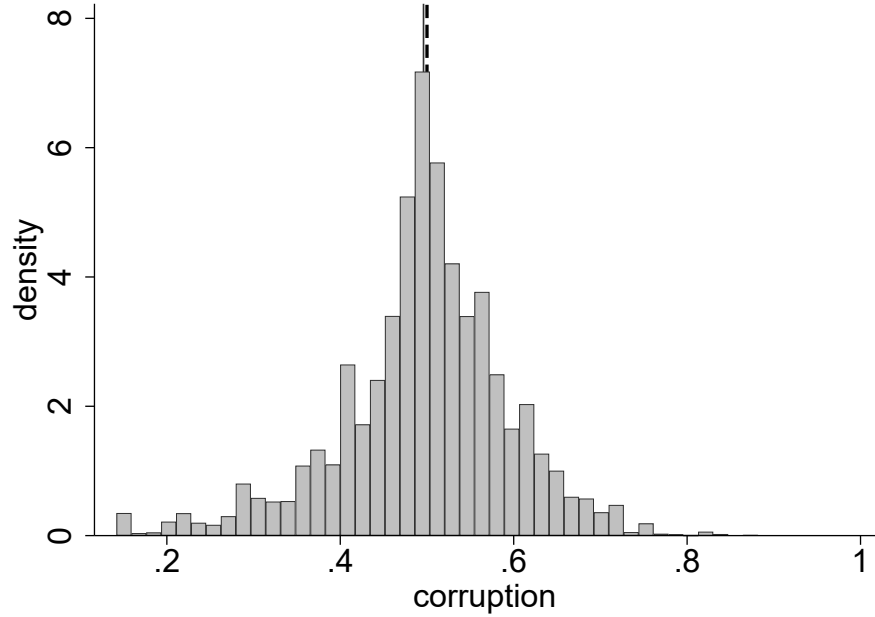
Figure D.3 – Distribution Corruption Index, 1789–2017.
(*Dotted line represents the median; solid line indicates the mean.*)

interaction results produced by Hanson (2015); according to Model 3, democracies have a systematically more benevolent effect in reducing infant mortality the lower state capacity is. This runs counter to the expectation related to the stateness-first argument. Adding to this, a sequencing specification—resembling the specification in Model 2, Table 3 of the paper, but with Ln infant mortality as dependent variable—finds that stateness-first sequences are systematically associated with higher subsequent mortality rates. This latter result is reported in Table .

Yet, neither the split-sample specifications (Models 1 and 2, which do not account for level of state capacity other than when dividing the sample in two) nor the CEM matching models, presented further below, on features of the bureaucracy at time of democratic transition and long-term reductions in infant mortality, find any systematic pattern. Nonetheless, there at least is no evidence *supporting* the stateness-first argument for this proxy of human development.

Next, we discussed in the paper how benevolent effects on reducing clientelism and corruption are proposed as key mechanisms linking stateness-first sequences to long-term development, for instance in Fukuyama (2014a). Yet, we find no evidence supporting the notion that democratization/democracy bears a stronger (or weaker) relationship with future clientelism or corruption in low-capacity contexts than in high-capacity contexts, neither from panel-, CEM, nor sequencing analysis.

In contrast, we do find indications from the panel specifications comporting with the notion that democracy is associated with increased risk of civil war onset in low-capacity contexts (Model 7,Table D.1). This relationship is non-existent in high-capacity contexts (Model 8). Further, the interaction specification (Model 9) finds that the relationship between democracy and civil war risk is systematically reduced as state capacity increases. While we once again remind that we regard these as preliminary results, and caution against reading too much into them, they do suggest that stateness-first type of arguments *might* be

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Ln infant mortality rate			Programmatic vs Clientelist links			Civil war onset		
Sample	LC	HC	Full	LC	HC	Full	LC	HC	Full
	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)	b/(se)
Polyarchy	-0.037	-0.051	0.185	0.018	-0.040	-0.097	0.058*	-0.003	0.023
	(0.147)	(0.163)	(0.139)	(0.411)	(0.379)	(0.363)	(0.034)	(0.011)	(0.017)
Impartial Pub. Admin.			0.058**			0.080			0.003
			(0.027)			(0.053)			(0.003)
Polyarchy X IPA			-0.280***			0.051			-0.010**
			(0.047)			(0.145)			(0.005)
Ln GDP per capita	-0.268***	-0.359***	-0.339***	0.172	-0.067	0.129	0.001	-0.009	-0.002
	(0.074)	(0.095)	(0.064)	(0.143)	(0.187)	(0.114)	(0.006)	(0.006)	(0.004)
Country dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	4473	5300	9799	7035	7310	14343	4003	4428	8457
R^2	0.823	0.916	0.881	0.111	0.258	0.181	0.025	0.027	0.013

Notes: *p<0.1; **p<0.05; ***p<0.01. Constant, country- and year dummies omitted. Errors clustered by country.
LC: Low capacity. HC: High capacity, relative to median value on IPA (-0.090) for full sample Models 1 and 2, Table 1.
Data for all dependent variables (see top row) are taken from the V-Dem dataset.

of more relevance to mitigating internal conflict than to other outcomes.

Table D.2 – CEM Balancing

	(1) Ln Inf. Mort. 20 years after democratization	(2) Ln Inf. Mort. in 2004	(3) Ln Inf. Mort. 20 years after democratization	(4) Ln Inf. Mort. in 2004	(5) Ln Inf. Mort. in 2004
Bureaucracy-level at democratization	−0.03 (0.06)	−0.001 (0.03)	−0.06 (0.04)	−0.06 (0.05)	0.18 (0.09)
Ln GDP at democratization	−0.08 (0.18)	−0.17* (0.08)	0.17 (0.12)	0.05 (0.17)	−0.13 (0.22)
Ln year at democratization	796.35 (618.30)	133.00 (352.50)	0.96 (10.82)	52.72** (16.65)	89.75* (27.09)
Ethnic fractionalization at democratization	0.79 (0.59)	1.14*** (0.29)	−0.49 (0.43)	0.16 (0.57)	3.13* (1.11)
Infant mortality at democratization	0.25 (0.22)	0.29* (0.12)	0.36* (0.15)	0.58* (0.23)	1.20* (0.41)
Constant	−6,042.45 (4,693.80)	−1,005.06 (2,669.22)	−6.53 (82.39)	−400.00** (127.11)	−683.05* (206.02)
Regional Dummies	Yes	Yes	Yes	Yes	Yes
Subclass Dummies	Yes	Yes	Yes	Yes	Yes
Comparison:	Demo. vs Aut.	Demo. vs Aut.	Democratizers	Democratizers	Democratizers
Democratization rule:	After 20 years of Aut.	First only	After 20 years of Aut.	First only	First only
Observations	54	156	39	41	17
R²	0.87	0.93	0.86	0.93	0.99
Adjusted R²	0.80	0.91	0.79	0.91	0.97

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors in parentheses.

Table D.3 – CEM Balancing

	(1) Corruption 20 years after democratization	(2) Corruption in 2004	(3) Corruption 20 years after democratization	(4) Corruption in 2004	(5) Corruption in 2004
Bureaucracy-level at democratization	−0.02 (0.02)	−0.02 (0.01)	0.002 (0.01)	−0.01 (0.02)	−0.03 (0.02)
Ln GDP at democratization	0.005 (0.03)	−0.02 (0.02)	0.08 (0.04)	0.02 (0.07)	0.08 (0.09)
Ln year at democratization	29.70 (23.83)	0.33 (19.20)	−0.50 (1.21)	2.97 (2.55)	3.40 (3.98)
Ethnic fractionalization at democratization	−0.04 (0.14)	0.12 (0.09)	−0.37* (0.14)	0.22 (0.14)	0.39 (0.27)
Corruption at democratization	0.60* (0.24)	0.44** (0.13)	0.40 (0.26)	0.14 (0.23)	−0.14 (0.35)
Constant	−223.90 (179.78)	−2.07 (144.82)	3.59 (9.35)	−22.24 (19.39)	−25.84 (30.06)
Regional Dummies	Yes	Yes	Yes	Yes	Yes
Subclass Dummies	Yes	Yes	Yes	Yes	Yes
Comparison:	Demo. vs Aut.	Demo. vs Aut.	Democratizers	Democratizers	Democratizers
Democratization rule:	After 20 years of Aut.	First only	After 20 years of Aut.	First only	First only
Observations	52	117	40	28	16
R^2	0.67	0.59	0.75	0.77	0.82
Adjusted R^2	0.38	0.36	0.53	0.56	0.47

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors in parentheses.

Table D.4 – CEM Balancing

	(1) Clientelism 20 years after democratization	(2) Clientelism in 2004	(3) Clientelism 20 years after democratization	(4) Clientelism in 2004	(5) Clientelism in 2004
Bureaucracy-level at democratization	−0.03	0.10	0.04	0.10	0.13
	(0.14)	(0.10)	(0.09)	(0.13)	(0.22)
Ln GDP at democratization	−0.04	−0.06	−0.49	−0.17	−0.05
	(0.28)	(0.20)	(0.30)	(0.42)	(0.88)
Ln year at democratization	−20.95	−35.09	5.82	2.63	8.62
	(252.42)	(217.07)	(13.96)	(20.27)	(33.74)
Ethnic fractionalization at democratization	−1.60	−1.23	−1.67	−1.05	1.72
	(1.32)	(0.90)	(1.27)	(1.60)	(3.42)
Corruption at democratization	0.28	0.35**	0.18	0.32	0.28
	(0.17)	(0.12)	(0.18)	(0.18)	(0.25)
Constant	157.88	267.71	−40.83	−17.40	−63.43
	(1,903.68)	(1,644.63)	(106.14)	(153.89)	(256.98)
Regional Dummies	Yes	Yes	Yes	Yes	Yes
Subclass Dummies	Yes	Yes	Yes	Yes	Yes
Comparison:	Demo. vs Aut.	Demo. vs Aut.	Democratizers	Democratizers	Democratizers
Democratization rule:	After 20 years of Aut.	First only	After 20 years of Aut.	First only	First only
Observations	70	146	56	48	23
R²	0.53	0.74	0.75	0.76	0.79
Adjusted R²	0.19	0.62	0.60	0.63	0.54

Notes: *p<0.05; **p<0.01; ***p<0.001. Standard errors in parentheses.

Table D.5 – Comparing various outcomes measured in $t + 20$, operationalizing stateness-first transitions

	(1)	(2)	(3)
	b/(se)	b/(se)	b/(se)
Dep. variable	Ln Infant mortality	Clientelism	Corruption
Stateness-first	0.251**	-0.197*	-0.006
	(0.115)	(0.114)	(0.008)
Low dem., high cap.	-0.015	0.131	-0.001
	(0.051)	(0.089)	(0.007)
High dem., low cap.	-0.014	0.016	-0.003
	(0.022)	(0.071)	(0.006)
High dem., high cap.	0.010	0.019	-0.005
	(0.034)	(0.084)	(0.007)
Σlow dem., low cap.	0.000	0.001	0.000
	(0.001)	(0.001)	(0.000)
Σlow dem., high cap.	-0.003	0.001	-0.000*
	(0.003)	(0.003)	(0.000)
Σhigh dem., low cap.	0.003	0.001	-0.000
	(0.002)	(0.002)	(0.000)
Σhigh dem., high cap.	-0.004***	0.006***	-0.000**
	(0.001)	(0.001)	(0.000)
Ln Infant mortality	0.543***		
	(0.041)		
Clientelism		0.483***	
		(0.042)	
Corruption			0.404***
			(0.036)
Year fixed effects	yes	yes	yes
N	8442	17567	18379

Notes: *p<0.1; **p<0.05; ***p<0.01. OLS regressions.
Year dummies omitted; Country-year is unit; Errors are clustered by country.