

Do Democracies Select More Educated Leaders?*

Timothy Besley

LSE and CIFAR

Marta Reynal-Querol

Universitat Pompeu Fabra-ICREA, CEPR and CESifo

September 14, 2010

Abstract

This paper tests whether education levels differ between leaders selected in autocracies and democracies. We use a unique data set

*We are grateful to seminar participants at Oxford and CIFAR, UPF, LSE-UCL, Paris School of Economics, and Toulouse School of Economics for helpful feedback. We are also very grateful to Florencia Abiuso for excellent research assistance and to Anne Brockmeyer and Rohini Pande for helpful comments. Besley is grateful for funding from the ESRC and CIFAR. Reynal-Querol is grateful for funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n. 203576 and acknowledges the financial support of the grant SEJ2007-64340 from the Spanish Ministerio de Educación. Reynal-Querol also acknowledges the support of the Barcelona GSE Research Network and the Government of Catalonia.

on over 1300 world leaders between 1848 and 2004 and exploit within country variation from transitions to and from democracy to show that democracies pick more highly educated leaders. The results are robust to a wide range of specifications, controls and ways of measuring education and democracy.

1 Introduction

The question of how institutions shape policy making and economic performance is a central issue for political economy. A key issue within this concerns how democracies and autocracies compare. This remains important since, even though the past twenty years have seen the emergence of many new democracies, autocratic forms of government remain empirically important in the world today. Moreover, little consensus has emerged on how democratic and autocratic policy making differs either in theory or in practice.

Beginning from first principles, there are two broad dimensions along which political institutions can affect policy and economic performance.

First, institutions shape incentives. They affect how policy makers target taxes, transfers or local public goods towards particular groups of citizens. Autocracies will often be run by and for particular elite groups while democracies create a greater incentive to appeal to key swing voter groups.¹ Institutions also shape accountability – how policy makers are rewarded for good behavior or punished for misdemeanors. The absence of organized open contests for power may limit the extent of accountability in autocracies.² Related to this, it is also unusual for autocracies to have a free and independent media which can report on the behavior of office-holders.³

Second, institutions determine the process of political selection. Selection of leaders can matter because it affects their motivation, competence or honesty. Political systems differ in the way that they select their leaders. For example, monarchies put weight on hereditary factors in determining succession. Military dictatorships tend to select leaders whose credentials come from success in the armed forces. Democracies put weight on how

¹See, for example, Acemoglu and Robinson (2006).

²Besley and Kudamatsu (2008) develop a model along these lines.

³ Besley and Prat (2006) argue that this will affect both adverse selection and moral hazard in politics.

leaders appeal to the mass electorate.⁴

Differences in institutional arrangements have been used to motivate a variety of empirical tests for whether democracies and autocracies differ in practice. A large empirical literature has emerged on whether democracy enhances a country's economic performance, particularly growth. Among early contributions to the cross-country literature, Przeworski and Limongi (1993) and Barro (1996) conclude that the correlation is weak and not robust. However, a recent panel data analysis by Papaioannou and Siourounis (2008) based on within-country estimates of permanent transitions, i.e. those where democracy is consolidated, finds that on average democratizations are associated with an increase in growth of between 0.5% and 1% per annum. This finding is broadly consistent with Persson and Tabellini (2005)'s novel econometric approach which also finds support for the proposition that persistent democracy is associated with an improvement in economic performance.⁵ Consistent with this, Aghion et al (2008) argue that democracy is correlated with improved performance of advanced sectors, i.e. those that are closer to the technological frontier. There has also been a debate about whether democracy affects observable economic policy outcomes. Here the literature is somewhat equivocal. For example, Mulligan et al (2004) finds no evidence of any significant difference between autocracies and democracies on a variety of economic policies.

In interpreting these empirical findings, the main focus has been on incentives. In general, the political economy literature has been reticent in discussing the issue of political selection.⁶ However, in a world of incomplete contracts and limited commitment, we should expect the personal characteristics of leaders to matter.⁷ In effect, leadership is about the exercise of

⁴Political coalition formation could also differ between autocratic and democratic settings as recently emphasized by Acemoglu et al (2008).

⁵ Persson and Tabellini (2008) further elaborate on the need to deal with heterogeneity in these debates.

⁶See Besley (2005) for a discussion.

⁷In addition to politics, there are other areas of public life where it is widely accepted

discretion. Confirming this, a number of studies show that selection in politics matters.⁸

Also consistent with the thesis that leaders matter are two papers by Jones and Olken (2005, 2007) who show that leadership change is correlated with economic and institutional performance. These studies are consistent with the idea that political leadership matters for policy, something which is commonplace in historical narratives. Almost every major historical episode of economic change has been associated with key personalities coming to power with a commitment to policy reform. In Germany of the 1870s and 1880s, for example, Otto von Bismarck was credited as the architect of early moves towards the creation of a welfare state. No account of the landmark economic reforms of Progressive Era in the United States fails to give a central role played by Theodore Roosevelt, and no account of the New Deal fails to give prominence to the role of Franklin Delano Roosevelt. The history of China's Great Leap Forward is written in terms of Mao's vision and how he imposed it. In the United Kingdom, Margaret Thatcher's economic legacy of privatization and de-regulation was in significant measure due to her single-minded pursuit of these goals. Barack Obama is similarly seen as a pivotal figure in meeting challenges faced during the current economic crisis.

Interest in these issues is not confined to political leadership. Recent accounts of corporate performance similarly place weight on charismatic chief executive officers (CEOs) and the way in which they shape corporate strategies. Indeed, Bertrand and Schoar (2003) argue that the data can be described in terms of CEO "fixed effects" indicative of management styles. A number of papers have observed that random shocks to CEOs affect firm performance consistent with the view that the identity of leaders matter – see, for example, Johnson et al (1985) and Bennedsen et al (2007).

This paper is motivated by the view that the quality of political leadership that selection matters including picking regulators, judges and central bankers.

⁸See, for example, Chattopadhyay and Duflo (2004) and Pande (2003) for evidence from India.

ship matters and looks for evidence of whether political institutions affect the characteristics of leaders, with a particular focus on educational attainment. We use a unique data set on world leaders for 198 countries for the period between 1848 and 2004 comprising 1329 leaders. Studying educational attainment is a good starting point for an exercise like this since it is possible to collect such data on a consistent basis for a fairly large group of leaders over a lengthy time period for a wide variety of countries.⁹ It is also arguable that educational attainment is an objectively measurable indicator of leader quality – we know from a whole range of economic studies that education has both economic and citizenship returns. In addition to education, we will also look at a wider set of leaders’ characteristics, but for a smaller sample of leaders.

Our paper is also related to a growing literature on the quality of leaders and its process of selection. Galasso and Nannici (2009) analyzes whether electoral competition is good for political selection. Their theoretical model predicts that electoral competition has beneficial effects since parties choose to send their high-quality politicians (measured by years of schooling, previous market income, and local government experience) to the most contestable districts. In similar vein, Ferraz and Finan (2009) examine whether higher wages attract better quality politicians and improve political performance using variation in the salaries of local legislators across municipal governments in Brazil. They find that higher wages increases political competition and improves the quality of legislators, as measured by education, type of previous profession, and political experience in office.

To preview our main finding, consider Figure 1 which gives the proportion of highly educated leaders in the world beginning in 1874 classified by whether a country is autocratic or democratic according to the Polity IV project.¹⁰ The graph makes clear that the proportion of highly educated selected leaders is persistently higher in democracies than in autocracies. However, it is clear

⁹Appendix Table 1 documents the sample of countries and the time period in detail.

¹⁰We discuss the definitions of these variables and sources in greater detail below.

that we need to control for many things to know whether this relationship is robust and hence of interest. In the remainder of the paper, we will explore this further and show that there is convincing evidence that the pattern in Figure 1 does indeed represent a selection difference between leaders picked in autocracies and democracies.

The remainder of the paper is organized as follows. In the next section, we discuss some relevant theoretical issues to motivate and interpret the empirical analysis. In section three, we discuss our data set and how it has been put together. Section four presents the core empirical results and a wide range of robustness checks. Section five offers some concluding comments.

2 Theoretical Preliminaries

In this section, we discuss why political selection might differ between democracies and autocracies in theory. Selection of a leader has two components – determining who puts themselves forward for office, i.e. the set of candidates and a rule for determining who from among the candidates is successful. Both of these depend upon the formal and informal political institutions, such as networks, that are in place. We focus on a world in which the characteristics of leaders matter – in particular their loyalty, competence or honesty. To the extent that voters assess these to be related to observables, such as education, this will lead such observables to be correlated with leader choice.

In the case of democratic policy-making, the citizen-candidate approach of Osborne and Slivinski (2006) and Besley and Coate (1997) has developed a framework to model how political selection works.¹¹ However, there is no comparable framework for thinking through how autocracies select leaders. Given the heterogeneity of political systems that travel under the autocratic

¹¹For an extension of this approach with a particular focus on competence and honesty, see Caselli and Morelli (2004).

label (e.g. monarchies and military dictatorships), we will work with a reduced form approach as a means of motivating and interpreting the empirical results that follow.

A polity comprises a set of \mathcal{N} citizens indexed by j each of whom could, in principle, serve as its leader. We will consider an open contest to select a leader who will be endowed with some policy authority.

Each citizen in the polity has an underlying quality level q_j , such as education, that is valuable to other citizens if citizen j is chosen to be the leader. A system of leadership that selects on competence would pick the leader with the highest q . However, this requires that this citizen makes himself available for office and that he is picked from among the available candidates. We suppose that the characteristic that makes the citizen a good leader could also be valuable in the market place and hence that he or she can earn a wage $w(q_j)$ in his/her next best occupation. It is natural to think that $w(\cdot)$ is an increasing function.

Political selection in general requires modeling three stages of the political process: (i) Citizens first decide whether to be a candidate for leader; (ii) The leader is then picked from among the candidates; (iii) The winning leader is in office and can use his policy authority.

The set of available candidates is denoted by \mathcal{C} . In a democracy, where almost any voter can stand, there is typically a process of nomination (or a primary) to determine who is on the ballot. Hence the set of potential candidates for leader tends to be small and well-defined. In autocracies, the set of candidates is generally opaque, except perhaps in monarchies with well-defined succession rules. In military regimes, \mathcal{C} is presumably the set of senior generals. The composition of \mathcal{C} is determined by who is willing to stand.¹²

Candidates are motivated by the rewards (social and private) from holding office. This could include any altruistic benefit and direct financial

¹² It is possible that some countries are lead by reluctant leaders, but it is not possible to observe whether this is the case.

rewards including perquisites as well as “ego rents”. It may also reflect the risk of being assassinated while in office or being jailed afterwards. To reflect differences in motivation and actions while in office, let R_j be the rewards that j would gain from holding office. It is hard to know a priori how this depends on the leader’s quality. If higher quality leaders act with more probity, then this will reduce their personal rents but this may also reduce their risk of violent overthrow. But there is clearly a risk of adverse selection with lower quality leaders looking forward to a higher rent if they are in office. The absence of effective checks and balances in autocracies would lead us to expect this adverse election effect to be more pronounced in autocracies leading to more bad politicians being attracted to public office.

Potential candidates must also weigh up what would happen if they are not in office. Let $r_j(\mathcal{C})$ be j ’s expected payoff if he is not the leader and the candidate set is \mathcal{C} . This reflects the fact that the alternative leader will be drawn from \mathcal{C} excluding j . If the remaining potential leaders in \mathcal{C} are of low quality, then $r_j(\mathcal{C})$ is lower which could encourage j to stand. In effect, R_j and $r_j(\mathcal{C})$ summarize the policy making stage (iii) in reduced form. We now turn to stage (ii).

A key determinant of leadership quality concerns how the leader is chosen from among the available candidates. We use $P^j(q; \mathcal{C})$ to denote j ’s probability of winning in a candidate set \mathcal{C} if he is of quality q with $\sum_{j \in \mathcal{C}} P^j(q; \mathcal{C}) = 1$. Besley and Coate (1997) discuss how this function can be derived from an underlying voting equilibrium in a simple majoritarian model of democracy. The form that this might take in an autocracy is not clear-cut. Typically, there is a decisive group who makes the decision. Bueno de Mesquita (2003) use the term *selectorate* to describe that group. Occasionally, the rules according to which they operate are clear. This might be true for a monarchy which works on the basis of inheritance. However, in many autocratic systems, it is far from clear what the rules are for picking a future leader from among the candidates. This is particularly

true in one-party states and military dictatorships.¹³ Our function $P^j(q; \mathcal{C})$ summarizes this process in reduced form including any randomness in it.

One key issue is the extent to which selection from among the candidates is based on candidate quality. All kinds of political systems may face a potential trade-off between selection based on loyalty and selection based on competence. In a democracy, voters may prefer to elect a leader from their own ethnic group or party rather than picking the most competent candidate. Military elites may prefer to select a leader based on who will be most docile to military interests while party elites in a communist dictatorship may prefer to select an ideologue. To the extent that loyalty is more prized in autocracies given the informality of the rules, we might expect less selection on competence under autocracy.

Finally, we turn to candidate entry – stage (i) above. We suppose that every candidate faces a cost (or benefit) of becoming a candidate which we denote by δ_j . Included in this cost is any intrinsic like or dislike from standing for office as well as any pressure brought to bear to encourage or discourage a particular candidate. In a reduced form way, this could capture the importance of a citizen’s membership of political networks. Repression of particular groups could be represented by a high δ_j while favoritism would be the opposite. In autocracies, costs of becoming a candidate for leader tend to be lower for those who are close to members of the selectorate, e.g. generals. Political networks in democracies are often based around university and party activists who are repressed in autocratic settings. In politically violent societies, δ_j could include the risk of being harmed, even killed, during a campaign. In general, entry costs will vary across political systems with greater formal entry barriers expected in autocratic systems. It is notable that Polity IV uses openness of executive recruitment as one of its criteria for assessing the extent of democracy.

Putting these factors together we can now consider the factors that shape

¹³Party systems in democracies vary widely in the clarity of rules that they have used to pick their leaders who, in Parliamentary systems, are then the main candidates for leader.

the equilibrium pool of candidates for office, essentially as an extension of the approach taken in the citizen-candidate model of democracy. Since payoffs depend on who else is a candidate, this is best modeled as a game played between all citizens whose strategies are whether or not to put themselves forward to become leader. We will require that these decisions form a Nash equilibrium. In this context, this means that the equilibrium candidate set must satisfy two conditions.

First, for each candidate $j \in \mathcal{C}$, that candidate must be willing to stand for office, i.e.:

$$P^j(q_j; \mathcal{C}) R_j - [1 - P^j(q_j; \mathcal{C})] r_j(\mathcal{C}) - \delta_j \geq w(q_j) + r_j(\mathcal{C} \setminus \{j\}) \quad (1)$$

The left hand side of (1) is the expected payoff from being a candidate reflecting the probability of becoming the leader and the payoff from holding office. Clearly if $P^j(q_j; \mathcal{C}) = 0$, a citizen is not likely to be a candidate so the process of selection at stage (ii) is central to this decision. The right hand side of (1) is the opportunity cost in terms of foregone earnings or other rewards from not being a candidate. Candidates of higher quality may be deterred from putting themselves forward to the extent that there are larger gains to other careers. For instance, countries with market liberalization may find it harder to attract higher quality candidates.

Second, we require that the proposed candidate set be *entry-proof*, i.e. that no other candidate who is not in \mathcal{C} wishes to enter. Formally, for all $k \notin \mathcal{C}$, we require that:

$$P^k(q_k; \mathcal{C} \cup \{k\}) R_k + [1 - P^k(q_k; \mathcal{C} \cup \{k\})] r_k(\mathcal{C} \cup \{k\}) - \delta_k < w(q_k) + r_k(\mathcal{C}). \quad (2)$$

The left hand side is the expected payoff to being in the race while the right hand side is the payoff from staying out. To make this decision, citizen k must conjecture the likelihood of being picked given who else is standing as represented by $P^k(q_k; \mathcal{C} \cup \{k\})$.

Together, equations (1) and (2) guarantee that a proposed candidate set \mathcal{C}

can be supported as a Nash equilibrium in entry strategies.¹⁴ Who actually holds office will be governed by the selection function $P^j(q_j; \mathcal{C})$ for $j \in \mathcal{C}$ which determines who wins from among this set of candidates.

It is clear that there are two ways that institutions can matter for leadership quality – institutions can change the mix of available candidates or they can affect who is picked from among the candidates. The two are linked since selection from among the candidates affects the incentive to stand as (1) and (2) make clear. Our discussion above emphasizes that there are good theoretical reasons why democracies and autocracies differ. However, whether this is true in practice requires an empirical investigation.

Using this observation, our empirical approach will be to look at the reduced form relationship between measures of realized leader quality and a country’s political institutions. Suppose that leader ℓ from country c is in office at date t . Then the probability that he is of quality Q_{ct} when institutions are I_{ct} is:

$$\text{Prob}(Q_{ct} = q_\ell; I_{ct}) = \text{Prob}(\ell \in \mathcal{C}(I_{ct})) P^\ell(q_\ell; \mathcal{C}(I_{ct}), I_{ct}) \quad (3)$$

where we have made the candidate set \mathcal{C} and the selection function an explicit function of institutions I_{ct} . Institutions can affect either the set of candidates who are available or the probability that a given candidate is selected from among the pool as governed by (1) and (2).

It is the reduced form relationship, (3), that we study empirically since we are not able to get data on \mathcal{C} directly. We therefore look at leader characteristics as a function of political institutions. Interpreting the findings, we will discuss whether there is any evidence of the stage at which institutions are having an impact on quality.

¹⁴These are essentially as in Besley and Coate (1997) which characterizes candidate entry in a representative democracy where the function $P(\cdot)$ is induced by voting. Whether a Nash equilibrium exists is not clear. However, as Besley and Coate (1997) shows, a model of mixed strategies can straightforwardly be introduced to deal with this issue.

3 Data

Our core data set which identifies the political leader in each country and year comes from the Archigos data. Archigos collects data from 1875 to 2004.¹⁵ Archigos provides information on the exact dates for which leaders have been in power. For each state, Archigos identifies the primary ruler, the way in which rulers enter and leave political power, the post-tenure fate of the ruler, and personal variables such as birth date or year, date of death and gender. As Archigos explains, many countries have more than one head of state. In some cases, the formal head of state may be a ceremonial position, as in many present day European monarchies. Archigos attempts to identify the actual effective ruler based on their knowledge of the particularities of each state. In most cases, identifying these rulers is clear and uncontroversial. But in a few, it is not. Two simple rules are generally followed: (i) in Parliamentary regimes, the prime minister is coded as the ruler while in Presidential systems, it is the president; (ii) in communist states the Chairman of the Party is coded as the effective ruler. However, there is a small number of exceptions. Countries are based on the Gleditsch and Ward sample. If a country is conquered or occupied, but is governed by an autonomous leader, as was the case in Denmark from 1940-43, and Estonia in 1940, those leaders are included in this data.¹⁶ For each leader the Archigos data provides information on the start and end date creating a leader-spell. Since some

¹⁵Archigos has two datasets: the long one, which gives information on leader-year-country, and the short one, which gives information on leader-country. In the short dataset there are 95 leader-country points that do not appear in the long dataset. We include these 95 points in the long dataset, and in the long format. (leader-year-country). (These 95 country-leaders points correspond to the following countries: Barbados, Bahamas, Belize, Brunei, Cape Verde, Iceland, Luxemburg, Maldives, Malta, Montenegro, Solomon Islands, Suriname, Tibet, Transvaal, Zanzibar)

¹⁶This creates a few incompatibilities with the Polity IV data which coded these cases as regime transitions.

leaders have more than one spell in office, the same leader may have more than one start date and end date.

We supplement data in Archigos from other sources. The main additional data source is Ludwig (2002) which includes all political leaders from independent states who held power as the “chief executive” for any length of time during the period mainly between 1900 and 2000.¹⁷ De facto leaders, i.e. those judged to have the greatest political authority, with or without formal titles or positions, took precedence over any constitutionally appointed or elected officials who had nominal authority or ceremonial positions. For inclusion in the database, the chief executive heads may or may not share power with other branches of government such as the legislature or council. However, they could not share power equally with other individuals, as in the case of a junta. The leader has to be “in charge”. In order to establish the list of actual leaders Ludwig (2002) uses Lentz’s *Encyclopedia of Heads of States and Governments* as well as the “Rulers” database from the geocities webpage.¹⁸ In order to decide whether then real executive power was vested in a monarch, president or prime minister, the data made use of *Britannica Online*, the *Library of the Congress Country Studies*, and a number of country level studies. The *Europa World Year Book* for 1997, 1998, 1999 and *Lexis-Nexis Academic Universe* were consulted for information on many late-century leaders that were neglected by *Lentz* and *Britannica*.

We have a potential sample of 227 countries. Among them, the Archigos data has listed the leaders for 183 of the 227 countries. Ludwig also listed the leaders of 12 more countries which are not listed in Archigos. This leaves a total of 197 countries for which a leader is listed.

Our core leader sample is for 198 countries from 1848 to 2004. We pick one leader per year to give us a total of 2075 leaders, and a total of 2486 leader-spells in office.¹⁹ Among the final list of 2075 leaders, 2018 are in the Archigos list, and 1522 are in the Ludwig list. Upon termination of his

¹⁷The data does list some leaders from in office between 1848 and 1900.

¹⁸<http://www.geocities.com/Athens/1058/rulers.html>

¹⁹In cases where more than one leader is in office in a given year, we focus on the leader

project, Ludwig (2002) had 1953 individuals from 199 countries. A special sub-sample contained 377 rulers with more detailed information.

The Ludwig (2002) data provides information on the education of only 333 leaders of our sample of 2075.²⁰ We have constructed a new dataset on education of leaders following Ludwig (2002)'s criteria and we now have data for 1654 leaders of the 2075 leaders we have in the sample. For 1451 leaders of the 1654 we use the *Encyclopedia of Heads of States and Governments*, *Oxford Political Biography: Who is Who in the Twentieth Century World Politics*, *Encyclopedia Britannica*, other online sources, as well as individual biographies from *Lexis-Nexis*.²¹ This completes our high data quality sample of leaders. For 203 of the 1654 leaders, we also collected information on education from less reliable sources. Our core results are based on the high quality sample, although we will assess their robustness to using the larger sample. We also collected information on whether the leader was educated abroad from the same sources.

To measure education, we follow Ludwig (2002) to construct the following eight-value discrete variable denoting the educational attainment of the leader:²² illiterate (no formal education) – 2 (3) leaders; literate (no formal education) – 39 (57) leaders; grade /elementary /primary school or tutors – 196 (224) leaders; high /finishing /secondary /trade school – 128 (142) leaders; special training (beyond high school), such as mechanical, nursing, art, music or military school – 44 (48) leaders; college – 643 (700) leaders; graduate or professional school (e.g. master's degree) – 270 (332) leaders; doctorate (e.g. PhD) – 129 (148) leaders. Our core variable measuring whether or not a leader is highly educated is a dummy variable equal to one

who has been in office for the longest time period during the year.

²⁰We are grateful to Arnold Ludwig and Gregory Gunthner for generously agreeing to make their data available to us.

²¹ In completing this exercise, we were careful to exclude honorary degrees obtained during or after a leader's spell in office.

²²The numbers in parentheses refer to the larger sample of leaders where we deem the data to be less reliable.

if the leader is in either category 7 or 8, i.e. has a post-graduate qualification which we will refer to as “Graduate Education”.

Our core measure of democracy is from the widely-used Polity IV data base. Polity data provides a definition of democracy that captures different dimensions: how competitive and open the recruitment of chief executives is; the extent to which the chief executive is constrained institutionally; and how competitive and regulated political participation is. The main summary variable that we use takes on values between -10 and $+10$. Following a long line of research by economists, e.g. Persson and Tabellini (2005), our core definition classifies a country as democratic if the variable POLITY2 is positive.

4 Results

In this section, we present our empirical results. We do so in a number of sections, beginning with our empirical specification and core results.

4.1 Empirical Specification

Our core empirical specification is for 1329 leaders, i.e. only the high data quality sample. We estimate a linear probability model as a version of (3) so that we can focus on within country variation for the education of leader ℓ first selected to serve in country c at date t

$$e_{\ell ct} = \alpha_c + \alpha_t + \beta d_{ct} + \gamma x_{ct} + \varepsilon_{\ell ct} \quad (4)$$

where α_c is a country fixed effect, α_t is a year dummy and x_{ct} are other controls. We cluster the standard errors by country to allow for arbitrary within country correlations in the errors.

In this specification, we enter the democracy variable in the year in which the leader is first selected to hold office. This is important since our hypothesis that political selection is at work pinpoints the institutions at the time

of selection to be of key relevance rather than institutional changes while a leader is in office. We will test the robustness of our results to different timing assumptions.

Given the length of the time series and wide set of countries, the main time-varying regressor which we control for consistently is income per capita which comes from Maddison's data. Potentially this is important as it could represent economic opportunities for potential leaders outside of government. For a more limited sample of countries/time periods that we discuss below, we can also include general measures of educational attainment in the country concerned. Given the paucity of time varying regressors over our long time period, we check whether our results are robustness to including country-specific time trends that are likely to do a reasonable job in proxying for a wide variety of economic changes within countries.

4.2 Core Results

The core results are in Table 1. In column (1), we look at the relationship between high education and democracy without including either country or year dummies. There is a positive correlation between the leader being highly educated and democracy. There is also a positive correlation with income per capita. In column (2), we add country dummies which allow us to control for time-invariant country characteristics. The correlation between the leader's education and democracy remains positive and significant. The democracy coefficient remains significant in column (3) where we include year dummy variables to capture global macro-economic shocks and trends. However, income per capita at the country level is now no longer significantly correlated with the leader being highly educated. Finally, column (4) adds country specific time trends. This would, among other things, pick up the rate at which general rates of educational attainment are growing in each country. Again, we find that the correlation between having a highly educated leader and being democratic is strongly significant. Across the first

four columns in Table 1, the size of the coefficient on democracy is largely of the same magnitude with democratic elections delivering leaders who are around 25% more likely to be highly educated.

Columns (5) and (6) in Table 1 repeat the specifications of columns (3) and (4) excluding the income per capita variable. This recognizes the fact that we lose almost one sixth of our observations by including it. The results are essentially unchanged from those in the first four columns and hence are not sensitive to whether or not we control for income per capita.

Given that the mean of the left hand side variable is around 0.35, we are likely to be safe using a linear probability model. However, columns (7) and (8) assess this by using a fixed-effects logit model. The link between democracy and highly-educated leaders survives and the effect is even a little larger in magnitude.

Finally, columns (9) and (10) repeat the specifications of column (5) and (6), but including information on the 203 leaders from the lower quality data sources. The results are robust to using this larger sample and the size and significance of the democracy effect is identical.

To summarize, these core results suggest that more educated leaders are found in democracies.

4.3 Robustness

We now assess the robustness of these results to a variety of alternative specifications.

Timing: If what we are capturing is political selection, then we should not be able to predict leader's education by lagged democracy (conditional on contemporaneous democracy). This is because it should be the institutions in place at the time of selection that count rather than some general trend towards democracy within a country. Table 2 assesses this by including a number of lags of democracy in the regression. Columns (1) through (6) include successively longer lags – from one to six years. All specifications

include both country and year dummy variables. Looking at the top row, it is clear that adding these lags does not in any way disturb the correlation between democracy and selecting a highly educated leader, the size and significance of which remains the same. In no case does the lag of democracy predict whether the leader is highly educated. Thus, it does appear as if it is the current institutions that matter.

Column (7) of Table 2 tries a slightly different approach to the same issue including as a regressor instead, the average experience of democracy since the country entered the sample (lagged five years). This should proxy for any kind of evolving democratic tradition which could be driving the process of who comes forward and is selected as leader. However, this variable is not significant. In column (8), we also try to control for longer-term economic trends by including the average GDP level of the past five years instead of contemporaneous GDP in case high-frequency changes in GDP contain very little signal. However, the reader will see that this variable is not significant. In both columns (7) and (8), the core finding from Table 1 remains both in magnitude and significance.

Education at the Country Level: The literature on the prerequisites for democracy beginning with Lipset (1959) has emphasized the importance of education in the population as a whole for the sustainability of democracy. These ideas have been further developed in Glaeser et al (2007). However, this literature refers to the education level of the citizens as a whole rather than of the leaders in democracies. But it does lead us to expect that we should find more educated populations in democratic countries making it potentially important to control for this in the empirical analysis. The main issue is data availability which will either restrict our sample of countries or the years of our study. However, we are able to use two different data sets on education to look at this.

The first data that we use is from a study by Morrisson and Murtin (2009) who assembled data on educational attainment in 74 countries for the period

1870-2010. For the entire sample, we use the average number of years of education of the population aged over 15.²³ Column (1) of Table 3 repeats the specification of column (3) of Table 1 but including the average years of education of the population older than 15. There is, however, no significant relationship between the general level of education in the population and having a more educated leader. Moreover, the coefficient on democracy at the time of selection remains significant and is slightly larger in magnitude than in the core sample.

In column (2) we construct a different dependent variable to try to capture the leader's education *relative to* the population as a whole. We capture this by taking the years of education of the leader *minus* the average years of education of the population. To construct this, however, we need to impute a number of years of education to correspond to the eight categories of educational attainment in our data.²⁴ Column (2) of Table 3 indicates that democracies select leaders who are, on average around 1.5 years more educated than the average citizen. The effect of democracy does, however, appear to be non-linear as illustrated in column (3) where the dependent variable is a dummy that has value 1 if the leader has at least 14 more years of education compared to the population as whole. This shows that democracies select leaders who are 22% more likely to be more educated according

²³They provide information every 10 years so we interpolate the data in order to have annual data.

²⁴Morrisson and Murtin (2009) consider six years of schooling as primary school completed; six more years of schooling as secondary school completed; and 4 more years of schooling as higher education completed. On this basis, we compute the years of education of our leaders as follows: illiterate(no formal education) – 0 years; literate (no formal education) – 2 years ; grade/elementary/primary school or tutors – 6 years; high/finishing/secondary/trade school – 12 years (+6); special training (beyond high school), such as mechanical, nursing, art, music or military school – 16 (+4) years; college – 16 (+4) years ; graduate or professional school (e.g. master's degree) – 18 years (+2) ; doctorate (e.g. PhD) – 20 years (+2).

to this measure.

Our second effort to control for education levels focuses on the post-1960 sample and uses the well-known Barro-Lee education variables covering a larger group of countries compared to our earlier data and the Penn World tables data for GDP. In columns (4) and (5) of Table 3 we first show that the core Table 1 results hold up when we confine ourselves to the post 1960 sample. In column (6) we include the Barro-Lee variable on the average education in years in the total population aged over 25. As in columns (1) through (3), there is no significant relationship between the general level of education in the population and having a more educated leader. In order to check how the two educational variables compare we include in the post-1960 sample the education variable from Morrisson and Murtin (2009) in column (7). Given the reduced coverage of countries, this reduces the number of observations. However, the coefficient on the democracy variable is the same as when using the full sample. In column (8) and (9), we repeat the specification of column (2) and (3) but using the Barro-Lee data. The results are very similar. The bottom line from Table 3 is that our core results are robust to controlling directly for the educational attainment in the population as a whole.

Measuring Leader’s Education: Table 4 explores robustness in a different dimension – whether the specific threshold for highly educated leaders affects the results. The various columns of Table 4 use different thresholds for measuring whether a leader is educated.

For the sake of comparison, column (1) replicates the core finding for having a graduate degree. Columns (2) through (5) successively relaxes this threshold picking a lower and lower cut-off for classifying whether a leader is educated – the label in the column gives the lowest cut-off level of education in the measure that we use. While the results remain significant across all columns, the size of the effect gets considerably smaller for lower cut-off thresholds. Thus being highly educated does appear to be the key factor

driving the result. In column (6), we include a different measure of education – an indicator variable running from 1-8 for all eight educational achievement categories for leaders. There is a positive and significant threshold in this case too.

The results in this table, we also pick up on a theme from Spilimbergo (2009) who shows that students who study abroad in democracies appear to promote democracy in their countries of origin. Perhaps one feature of democracies is also that their leaders are more likely to have received a foreign education? Column (7) shows that this is indeed the case. The dependent variable is now a dummy variable indicating whether the leader has studied abroad. The result suggests that this is 12% more likely in a democracy.

Finally, we look at whether the results are driven entirely by the importance of lawyers serving as leaders.²⁵ This is important since the exact educational attainment associated with being a lawyer in each country is quite complex and it is arguably an area where we may have mis-measured our left hand side variable. The results in column (8) of Table 4 are encouraging on this front as the size and significance of the democracy effect remains when we exclude the 256 leaders whose are classified as having been professional lawyers.

Measuring Democracy: Table 5 considers some different thresholds for being democratic and different classifications of democracy. It is arguable that using a POLITY2 score greater than zero is too permissive and would allow some countries to appear democratic who perhaps ought not to qualify. We therefore consider a number of other thresholds to explore empirically

²⁵Ludwig codifies 307 leaders of our list of leaders for which we have education data as being lawyers. We also construct the variable "professional lawyer" using the same sources that we used to collect the education data and find that 421 leaders in our list of leaders with education data have practiced as lawyers (369 if we restricted to high quality sources for education). Among these, 307 of them coincide with those coded by Ludwig (2002) while the remaining were missing in Ludwig's data.

where the data tells us the key cutoff lies.

Again, for the sake of comparison, we replicate our core result in column (1). Column (2) uses a much more demanding threshold of POLITY2 being bigger than 5 and still finds a positive correlation between a leader being highly educated and democracy. However, the magnitude of the coefficient is somewhat smaller. Columns (3) and (4) show that the result holds up when the cutoff chosen is POLITY2 bigger than 6 or 7. In column (5), we create a series of dummy variables for the POLITY2 measure being in different ranges to assess exactly where the action lies. The message from this exercise is fairly clear, indicating that a measure above minus 1 is the key cutoff with an increasing effect being found for the ranges above that level compared to the omitted category (less than minus 1).

We also check the robustness of our results to using a different data set on democracy, that of Przeworski et al. (2000) in its updated form due to Boix and Rosato (2001). They code a country as democratic if their elections are free and competitive, the executive is accountable to citizens, and at least 50 percent of the male electorate is enfranchised. This data set covers the period 1800-1999. Compare to Polity IV, the Boix and Rosato (2001) measure heavily depends on political contestation, putting less weight on political participation and executive constraints.²⁶ Column (6) confirms that the results are robust to using this alternative measure.

In column (7), we use the POLITY2 variable as a continuous measure of democracy. Again, the core result is robust. Moreover, this result is also robust to measuring education in a more continuous way.

One discouragement to leaders standing for office is the prospect of forcible removal from office. This could affect R_j in the theory. Thus, a history of

²⁶This definition of democracy is less permissive than the definition from Polity IV. In our core sample, Boix and Rosato (2001) classify 200 country-year observations in which leaders are selected as autocratic when the core Polity2 definition classifies them as democratic. There are only 25 observations where Polity2 denotes a country as autocratic when according to Boix and Rosato (2001) the country is democratic.

political instability and violence could well act as a deterrent to higher quality leaders. It is possible also that this history is correlated with being democratic since autocratic leaders are more susceptible to violent removal. To examine this issue, we construct a variable reflecting political history measuring how frequently leaders have left power in coups, revolutions or assassinations – specifically the percentage of past leaders that have left power by such irregular means. The result is presented in column (9). As conjectured, this variable is negative and significant, i.e. instability acts as a deterrent to educated leaders taking office. However, the sign and significance of democracy at the time of selection remains as in the core results.²⁷

Taken together, the results in Tables 2 through 5 underline the robustness of the core findings in Table 1²⁸.

4.4 The Role of the Military

In this section, we explore how far the results that we have found are shaped by the importance of political selection through the military in autocracies. Plausibly this affects the pool of candidates who are available. Ludwig codifies 290 leaders of our list of leaders as being military. We also construct the variable “military professional” using the same sources that we used to collect the education data and find that 314 leaders in our list have served in the military, among which 290 of them coincide with those coded by Ludwig (2002). The remaining 24 were missing in Ludwig. In general, military leaders are more prevalent in autocracies with 40% of leaders who are selected in

²⁷Another potential concern is that leaders who take office as part of a *coup d’etat* are driving the results. We address this issue in the last section.

²⁸There could also be other potential stories that may affect the choice of the leader. For example, a sudden drop in income per capita. We have checked whether shocks to income per capita affect the education of the leader and we find no results.

autocracies having had careers in the military compared to 10% in democracies. If military recruitment tends to come from a less educated population, then this could be driving our results. As anticipated, 36% of non-military leaders have high education compared to 4% of military leaders. However, 8% of military leaders in democracy have high education compared to 2% of military leaders in autocracy.

We look at this issue in more detail in Table 6. For these specifications, in column (1) to (3), we control on the right for whether the previous profession of the leader was in the military. We then include an autocracy dummy and an interaction term between being in the military and an autocrat. Across the board in Table 6, we find that leaders with a military background are around 25% less likely to be highly educated. However, the effect of being selected in an autocracy remains. More interesting, however, is the finding that there is no significant difference between autocracies and democracies in terms of the types of military leaders that they recruit as leaders. Returning to the theory, this suggests that the key difference between autocracies and democracies may well lie in the available candidate set rather selection of a given candidate.

Finally column (4) shows that it is less likely that democratically elected leaders list their prior profession as being in the military. This further confirms that there is a difference in political networks in democratic and autocratic settings leading to a different candidate set being available.

4.5 Other Characteristics

The analysis so far has focused on education. Here, we briefly discuss some other characteristics that may be different between leaders who are selected in democracies and autocracies. Here, we use data on a smaller sample of leaders from Ludwig (2002). The results are reported in Table 7.

One particular issue is whether leaders who are selected in democracies are more honest or trustworthy. Column (1) considers spousal fidelity of leaders

using a dummy variable which is equal to one if the leader is coded by Ludwig as “probably faithful” or “probability unfaithful” and equal to zero if the leader is “definitely unfaithful”. Although the results are significant when year dummy variables are excluded (not reported), there is no significant effect of being selected in a democracy in the core specification.

Next, we look at trust. Here, we use a dummy variable which is equal to one if the leader is coded as being “seemingly trusting, open, above-board, frank”, or “equivocal (neither specially trusting not overly suspicious)”, and zero if the leader is “seemingly distrustful, suspicious, secretive, deceptive”. The results reported in column (2) shows that leaders selected in democracies are more likely to be viewed as trustworthy using the definition in Ludwig (2002). Thus democracies also seem to select better leaders in this dimension which is arguably important given the trust vested in leaders.

Column (3) looks at how far leaders have a wide circle of friends as coded by Ludwig. Here, we use a dummy variable that is equal to one if the leader is coded as having “close long-term friends”, and zero if leader is coded as “having social friendship with people such as business partners but without sharing sensitive, personal feelings” or “not having any close friends”. The results suggest that democratically selected leaders also seem to have more friends according to this measure.

Column (4) looks at oratory skills where, following Ludwig, we code each leader’s skills as equal to one if he is a demagogue, hypnotic speaker, eloquent, spell-binder, good debater, or effective speaker, and zero if the leader is coded as being dull, boring, absence of oratory skills. The results in column (7) suggest no significant difference between democratically and autocratically selected leaders.

Column (5) looks at church attendance which is measured using a dummy variable which is equal to one if the leader attends church regularly, and zero otherwise. There is evidence that democratically selected leaders tend to be more religious. Column (6) looks at age where we find that leaders are around two years older when first selected in democracies.

In the last column (7), we look at whether leaders are from middle and upper class backgrounds. This may tell us something about network selection in political systems. Here we use a dummy variable based on Ludwig (2002)’s classification of whether the most successful parent of the leader is middle- or upper- social stratum. Consistent with our results on education, we find that leaders in democracies appear to come from such relatively privileged backgrounds. Thus, in addition to selecting on education, democracies also seem to be more inclined to select from social elites. This provides interesting food for thought in reflecting on the nature of political networks that underpin democracies. This suggests that, while democracy may give citizens a greater say in who governs, it may not guarantee that citizens who rise to the top come from less privileged backgrounds.

4.6 Reverse Causation

Finally, we address whether reverse causation could explain our results. This would be the case if poorly educated autocratic leaders tended to precede democratizations while highly educated leaders tended to occur before democracy collapses. To address this concern, we first consider an “event study” representation of the data by lining up all the transitions to/from democracy and looking at the educational attainment of the leaders who come before and after the transition. We do this separately for transitions out of autocracy and out of democracy. This allows us to look for “pre-trends” in leaders’ education. The results are portrayed in Figures 2 and 3. Both sets of results do not suggest an association between transitions and the preceding history of the leader’s education.²⁹

We now look at the possibility that, in the year in which the leader comes to power, that there is also an autocratic or democratic transition. In such

²⁹ This result is corroborated by looking at the correlation between political transitions from autocracy to democracy and vice versa which are uncorrelated with leader’s education (results available on request).

cases it may be difficult to distinguish whether the leader in question was causing the transition. We have a total of 44 leaders that come to power in a year in which there was an autocratic transition, and 52 leaders that come to power in a year in which there was a democratic transition³⁰ In similar vein, we can also use data on how leaders came to power to differentiate more finely than our democracy variable allows. Ludwig (2002) provides information on this, including information on whether the leader was (i) elected; (ii) appointed/selected by committee or oligarchy; (iii) selected by a junta, military or revolutionary council; (iv) rose to power through succession; or (v) was the leader of a successful coup.

In column (2) of Table 8 we first present the results dropping the 127 leaders who came to power via a coup. The core result on democracy and education remains in both sign, size and significance. Column (3) looks at what happens when we drop the 44 leaders that rose to power in the year of an autocratic transition. Again, the results are robust. The regression reported in column (4) drops the 52 leaders that came to power in the year of a democratic transition, showing that the results are robust. And in column (5), we drop the leaders that rose to power in the year of any transition (96 leaders). Taken together, these results show that the relationship between democracy and leaders' education is robust to excluding the few observations where reverse causation might be an issue.

Next we construct a variable indicating whether or not a leader was elected rather than using the democracy measure from `polity2`. We use this as it also excludes leaders who oversaw a democratic transition after they came to power. Although the variable "elected" and democracy are highly correlated, there are some leaders who came to power in years in which `polity2` is above zero (a democratic year on our definition), but they did not

³⁰These are the number of observations for which we also have information on per capita income to include in the regression. Without including income per capita as a control, we would have 52 leaders coming to power the same year of an autocratic transition and 68 in the year of a democratic transition.

do so in an election. We also have a small number of leaders who rose to power in an election that polity2 does not regard as democratic. In column (6) of Table 8, we use the variable “elected” instead of our democracy variable to show that our main result is robust. In column (7) we construct a variable “elected & democracy” which takes on a value of one if the variable “elected” is one for this leader and the value for democracy in that year is also one. Again the result remains and now the coefficient of interest is even larger.

Another way to address the possibility of reverse causation is to see whether democracy in the year prior to the leader being elected predicts whether a leader is educated. Leaders who take office in a transition year would be classified as autocrats in such cases. The result is in column (8) of Table 8. Although the effect of democracy on education falls is smaller in magnitude, as we would expect if this exercise introduced pure attenuation bias, it remains positive and significant.

The evidence in table 8 increases our confidence that the correlation between democracy and leader’s education is capturing the effect of the political system on the leader’s education rather than the other way around.

5 Concluding Comments

This paper presents evidence that political selection with respect to education differs between autocracies and democracies. The evidence is drawn from a wide range of countries over more than 150 years and is identified from within country variation in political institutions.

Writing on American state politics some time ago, the great political scientist V.O. Key suggested that:

“The nature of the workings of government depends ultimately on the men who run it. The men we elect to office and the circumstances we create that affect their work determine the nature

of popular government. Let there be emphasis on those we elect to office.” V.O. Key (1956), page 10.

The results suggest that democratically elected leaders are around 25% more likely to be highly educated than those who are picked in autocracies. The results provide convincing evidence that there is a difference between political institutions in the characteristics of those selected to be leader. Thus, a focus on selection to office is warranted.

The paper does not, however, establish whether have more educated leaders is better of that changing the leader’s characteristics matters for policy and other outcomes. Hence, a major challenge in this research agenda is to see how far this difference in selection translates into policy differences and better policy outcomes for citizens.³¹

³¹A first step in this direction is undertaken in Besley, Montalvo and Reynal-Querol (2009) who show that highly educated leaders are correlated with enhanced growth performance while they are in office.

References

- [1] Acemoglu, Daron, Georgy Egorov and Konstantin Sonin, [2008], “Political Selection and the Persistence of Bad Governments,” unpublished typescript.
- [2] Acemoglu, Daron, and James Robinson, [2005], *Economic Origins of Dictatorship and Democracy*. Cambridge: Cambridge University Press.
- [3] Aghion, Philippe, Alberto Alesina, and Francesco Trebbi, [2008], “Democracy, Technology, and Growth,” in Elhanan Helpman (ed), *Institutions and Economic Performance*, Cambridge MA: Harvard University Press.
- [4] Barro, Robert, [1996], “Democracy and Economic Growth,” *Journal of Growth*, 1(1), 1-27.
- [5] Bennedsen, Morton, Francisco Perez-Gonzales and Daniel Wolfenzon, [2007], “Do CEOs matter?” Copenhagen Business School, Working Paper No. 13-2007.
- [6] Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. *The Logic of Political Survival*. Cambridge, MA: MIT Press, 2003.
- [7] Besley, Timothy, [2005], “Political Selection.” *Journal of Economic Perspectives*, 19(3), 43-60.
- [8] Besley, Timothy and Stephen Coate, [1997], “An Economic Model of Representative Democracy,” *Quarterly Journal of Economics*, 112(1), 85-114.
- [9] Besley, Timothy and Andrea Prat, [2006], “Handcuffs for the Grabbing Hand? Media Capture and Government Accountability” *American Economic Review*, 96(3), 720-736.

- [10] Besley, Timothy, and Masayuki Kudamatsu, [2008], "Making Autocracy Work." in Elhanan Helpman (ed), *Institutions and Economic Performance*, Cambridge MA: Harvard University Press.
- [11] Besley, Timothy, Jose Montalvo and Marta Reynal-Querol, [2009], "Do Educated Leaders Matter for Growth?", unpublished typescript.
- [12] Boix, Carles, and Sebastián Rosato, [2001], "A Complete Data Set of Political Regimes, 1800-1999." Unpublished Data set.
- [13] Caselli, Francesco and Massimo Morelli, [2004], "Bad Politicians," *Journal of Public Economics*, 88, 759-782.
- [14] Chattopadhyay, Raghavendra and Esther Duflo, [2004], "Women as Policy Makers: Evidence from a India-Wide Randomized Policy Experiment," *Econometrica*, 72(5), 1409-1444.
- [15] Ferraz, Claudio and Frederico Finan, [2009], "Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance", NBER working paper 14906.
- [16] Galasso, Vincenzo and Tommaso Nannicini, [2009], "Competing on Good Politicians", IZA DP No. 4282
- [17] Glaeser, Edward L, Giacomo A.M. Ponzetto and Andrei Shleifer, [2007], "Why does democracy need education?" *Journal of Economic Growth*, 12(1), 77-99.
- [18] Johnson, W. Bruce, Robert P. Magee, Nandu Nagarajan and Harry A. Newman, [1985], "An analysis of the stock price reaction to sudden executive deaths: Implications for the managerial labor market", *Journal of Accounting and Economics*, 7(1-3), 151-174.
- [19] Jones, Benjamin F., and Benjamin A. Olken, [2005], "Do Leaders Matter? National Leadership and Growth Since World War II." *Quarterly Journal of Economics*, 120(3), 835-864.

- [20] Jones, Benjamin F., and Benjamin A. Olken, [2007], "Hit or Miss? The Effect of Assassinations on Institutions and War." BREAD Working Paper: No. 150.
- [21] Key, V.O., [1956], *American State Politics: An Introduction*, New York: Alfred A. Knopf.
- [22] Lentz, Harris M, [1994], *Heads of States and Governments: A Worldwide Encyclopedia of Over 2,300 Leaders, 1945 through 1992*. Jefferson, NC: McFarland.
- [23] Lipset, Seymour Martin, [1959], "Some social requisites for democracy: Economic Development and Political Legitimacy", *American Political Science Review*, 53, 69-105.
- [24] Ludwig, Arnold M., [2002], *King of the Mountain: The Nature of Political Leadership*, The University Press Kentucky.
- [25] Marshall, Monty G., and Keith Jagers, [2005], "POLITY IV Project: Political Regime Characteristics and Transitions, 1800-2004." <www.cidcm.umd.edu/polity>.
- [26] Majumdar, Sumon and Sharun Mukand, [2007], "The Leader as Catalyst: On Leadership and the Mechanics of Institutional Change," available at http://www.econ.queensu.ca/working_papers/papers/qed_wp_1128.pdf.
- [27] Mulligan, Casey B., Richard Gil and Xavier Sala-i-Martin, [2004], "Do Democracies have Different Public Policies than Non-Democracies?" *Journal of Economic Perspectives*, 18(1), pp. 51-74.
- [28] Morrisson, Christian and Fabrice Murtin, [2008], "The Century of Education", PSE Working Paper No. 2008-22.

- [29] Osborne, Martin J. and Al Slivinski, [1996], "A Model of Political Competition with Citizen Candidates," *Quarterly Journal of Economics*, 111(1), 65-96.
- [30] Papaioannou, Elias and Gregorios Siourounis, [2008] "Democratization and Growth", *Economic Journal*, 118(10), 1520-15551.
- [31] Pande, Rohini, [2003], "Minority Representation and Policy Choices: The Significance of Legislator Identity," *American Economic Review*; 93(4), 1132-1151.
- [32] Persson, Torsten, and Guido Tabellini, [2000], *Political Economics - Explaining Economic Policy*. Cambridge, MA: MIT Press.
- [33] Persson, Torsten, and Guido Tabellini, [2003], *The Economic Effect of Constitutions*. Cambridge, MA: MIT Press.
- [34] Persson, Torsten, and Guido Tabellini, [2006], "Democratic Capital: The Nexus of Political and Economic Change." NBER Working Paper: No.12175.
- [35] Persson, Torsten, and Guido Tabellini, [2008], "Political Regimes and Economic Growth." in Elhanan Helpman (ed), *Institutions and Economic Performance*, Cambridge MA: Harvard University Press.
- [36] Przeworski, Adam, Michael E. Alvarez, Jose Antonio Cheibub and Fernando Limongi, [2000], *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. New York: Cambridge University Press.
- [37] Przeworski, Adam, and Fernando Limongi, [1993], "Political Regimes and Economic Growth." *Journal of Economic Perspectives*, 7(3), pp. 51-69.
- [38] Spilimbergo, Antonio, [2009], "Democracy and Foreign Education," *American Economic Review*, 99(1), 528-543.

Table 1: Democracy and Education (Core Results).

Method:	OLS	OLS	OLS	OLS	OLS	OLS	Conditional Logit	Conditional Logit	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Democracy	0.23*** (0.03)	0.26*** (0.04)	0.22*** (0.5)	0.24*** (0.06)	0.20*** (0.04)	0.24*** (0.05)	1.53*** (0.22)	1.62*** (0.28)	0.20*** (0.04)	0.25*** (0.04)
Log (GDP per capita)	0.04*** (0.01)	0.09*** (0.03)	-0.00 (0.06)	-0.14 (0.09)			0.58*** (0.14)	-0.13 (-0.36)		
Country dummies	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	No	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Country specific time trends	No	No	No	Yes	No	Yes	No	No	No	Yes
Observations	1133	1133	1133	1133	1329	1329	941	941	1513	1513
R-squared	0.0908	0.2955	0.3800	0.4674	0.3626	0.4474			0.3528	0.4395

Notes: All OLS regressions are reported with robust standard-errors clustered at the country level. Standard Errors in parentheses. * significant at 10%; **significant at 5%; ***significant at 1%. The dependent variable is a dummy that has value 1 if the leader has a graduate degree and zero otherwise. Democracy is a dummy variable that has value 1 if the polity2 score is larger than 0, and zero otherwise. The sample is a panel of 198 countries: 1848-2004. Each observation is for the first year the leader is selected. The democracy and per capita income variables are measured in the first year the leader is selected. Columns (9) and (10) are for the lower quality sample.

Table 2: Democracy and Education: Robustness to Adding Lags

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Democracy	0.23*** (0.06)	0.23*** (0.05)	0.24*** (0.06)	0.22*** (0.06)	0.22*** (0.06)	0.24*** (0.05)	0.22*** (0.05)	0.23*** (0.05)
Log (GDP per capita)	-0.04 (0.05)	-0.04 (0.06)	-0.05 (0.06)	-0.07 (0.06)	-0.08 (0.06)	-0.07 (0.05)	-0.09 (0.06)	
Democracy t-1	-0.003 (0.05)							
Democracy t-2		-0.003 (0.05)						
Democracy t-3			-0.02 (0.05)					
Democracyt-4				0.01 (0.06)				
Democracy t-5					-0.01 (0.06)			
Democracy t-6						-0.03 (0.05)		
Average GDP in the last 5 years								-0.06 (0.06)
Average democracy (lagged by 5 years)							0.02 (0.20)	0.03 (0.19)
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1071	1061	1053	1040	1034	1029	1055	1026
R-squared	0.3955	0.3948	0.3939	0.3946	0.3984	0.3959	0.3878	0.3866

Note: The estimation method is OLS. Standard errors clustered at the country level in parentheses. * significant at 10%; **significant at 5%; ***significant at 1%. The sample is a panel of 198 countries: 1848-2004. Each observation is for the first year a new leader is selected. The democracy and per capita income variables are measured in the first year that the leader is selected. Average Democracy (lagged 5 years) is the average democracy variable of the history of the country lagged 5 years.

Table 3: Democracy and Education: Controlling for education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Variable	Graduate Degree	Education al Distance	Education al Distance Dummy	Graduate Degree	Graduate Degree	Graduate Degree	Graduate Degree	Education al Distance	Education al Distance Dummy
Democracy	0.26*** (0.05)	1.49*** (0.37)	0.22*** (0.05)	0.28*** (0.06)	0.20*** (0.07)	0.20*** (0.07)	0.27*** (0.08)	1.25*** (0.48)	0.16*** (0.06)
Log (GDP per capita)	-0.03 (0.08)	0.19 (0.58)	0.14 (0.08)	0.11** (0.05)	0.00 (0.06)	-0.03 (0.07)	-0.06 (0.12)	0.25 (0.81)	0.008 (0.08)
Average Years of Education	-0.01 (0.03)	-1.31*** (0.23)	-0.10** (0.04)				0.02 (0.07)		
Average Years of Education (population over age 25)						0.01 (0.04)		-0.97*** (0.33)	-0.05*** (0.03)
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Sample				1960 onwards	1960 onwards	1960 onwards	1960 onwards	1960 onwards	1960 onwards
Observations	811	811	811	654	654	602	422	602	602
R-squared	0.3776	0.5599	0.5014	0.3760	0.4437	0.4441	0.4403	0.6000	0.5980

Note: The estimation method is OLS. Standard errors clustered at the country level in parentheses. * significant at 10%; **significant at 5%; ***significant at 1%. The sample is a panel of 198 countries: 1848-2004. Each observation is for the first year a new leader is selected. The democracy, per capita income variables and average education variables are measured in the first year that the leader is selected. Average Years of Education is the average years of schooling in the total population over 15, interpolated, from Morriison and Murtin (2009). The Average Years of Education (population over age 25) is the average years of schooling in the total population over 25, interpolated, from Barro-Lee (original variable is tyr25). For the sample 1960 on. Educational Distance and Educational Distance Dummy are explained in the text.

Table 4 Democracy and Education: Robustness to the use of different education variables

Dep. Variable	Graduate degree	College	special training, beyond high school	high-secondary school	elementary-primary School	Education continuous (from 1 to 8)	Studied abroad	Graduate degree (excluding lawyers)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Democracy	0.22*** (0.5)	0.11*** (0.04)	0.08** (0.04)	0.05** (0.028)	0.01** (0.006)	0.57*** (0.12)	0.12*** (0.04)	0.20*** (0.04)
Log (GDP per capita)	-0.00 (0.06)	-0.04 (0.05)	-0.007 (0.05)	-0.01 (0.04)	0.01 (0.02)	-0.06 (0.20)	0.11** (0.05)	-0.06 (0.06)
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1133	1133	1133	1133	1133	1133	1133	857
R-squared	0.3800	0.3680	0.3441	0.3908	0.4630	0.4037	0.3876	0.411

Note: The estimation method is OLS. Standard errors clustered at the country level in parentheses. * significant at 10%; **significant at 5%; ***significant at 1%. The sample is a panel of 198 countries: 1848-2004. Each observation is for the first year a new leader is selected. The democracy and per capita income variables are measured in the first year that the leader is selected. Graduate degree is the original Education variable describe above. College is a dummy that has value 1 if the minimum education level of the leader is college, and zero otherwise. Special training, beyond high school is a dummy that has value 1 if the minimum education level of the leader is special training beyond high school, and zero otherwise. High-secondary school is a dummy that has value 1 if the minimum education level of the leader is high-secondary school, and zero otherwise. Elementary-primary school is dummy that has value 1 if the minimum education level of the leader is elementary-primary school, and zero otherwise. Study Abroad is a dummy variable that has value 1 if the leader studied abroad and zero otherwise.

Table 5 Democracy and Education: Robustness to Measurement of Democracy

Dep. Variable	Graduate degree	Graduate degree	Graduate degree	Graduate degree	Graduate degree	Graduate degree	Graduate degree	Education (years)	Graduate degree
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Democracy	0.22***(0.5)								0.21***(0.05)
Democracy >5		0.17***(0.06)							
Democracy >6			0.16***(0.05)						
Democracy >7				0.15***(0.06)					
Democ - 5 to -1					0.06(0.06)				
Democ 0 to 5					0.17***(0.07)				
Democ 6 to 10					0.26***(0.07)				
Democracy BOIX-ROSATO						0.21***(0.05)			
Democracy (continuous measure)							0.02*** (0.004)	0.04*** (0.001)	
Average past political instability.									-0.37*** (0.13)
Log (GDP per capita)	-0.00 (0.06)	-0.01 (0.06)	-0.02 (0.06)	-0.03 (0.06)	0.004 (0.06)	-0.008 (0.07)	-0.008 (0.06)	-0.07 (0.21)	-0.06 (0.06)
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1133	1133	1133	1133	1132	1076	1133	1133	1074
R-squared	0.3800	0.3698	0.3671	0.3644	0.3757	0.3468	0.3777	0.4025	0.3940

Note: The estimation method is OLS. Standard errors clustered at the country level in parentheses. * significant at 10%; **significant at 5%; ***significant at 1%. The sample is a panel of 198 countries: 1848-2004. Each observation is for the first year a new leader is selected. The democracy and per capita income variables are measured in the first year that the leader is selected. Average political instability is the percentage of past leaders that loose power by irregular means (coups, revolutions or assassinations). Following Archigos definition of exit by irregular means.

Table 6 Autocracy, military, and education

Dependent variable	Graduate Degree	Graduate Degree	Graduate Degree	Military Professional
	(1)	(2)	(3)	(4)
Autocracy	-0.19*** (0.04)	-0.19*** (0.05)	-0.14*** (0.05)	
Autocracy*military(profession)	0.18** (0.05)	0.10 (0.08)	0.08 (0.09)	
Military (profession)	-0.18*** (0.03)	-0.22*** (0.05)	-0.25*** (0.06)	
Log (GDP per capita)	0.03 (0.02)	0.08** (0.03)	-0.04 (0.06)	-0.10 (0.05)
Democracy				-0.34*** (0.05)
Country dummies	No	Yes	Yes	Yes
Year dummies	No	No	Yes	Yes
Observations	1118	1118	1118	1305
R-squared	0.1366	0.3369	0.4218	0.4754

Note: The estimation method is OLS. Standard errors clustered at the country level in parentheses. * significant at 10%; **significant at 5%; ***significant at 1%. The sample is a panel of 198 countries: 1848-2004. Each observation is for the first year a new leader is selected. The democracy and per capita income variables are measured in the first year that the leader is selected. Military Professional is a dummy that is equal to 1 if the leader was in the military immediately before holding office.

Table 7 Democracy, honesty and popularity

Dependent variable	Fidelity	Trust	Friends	Oratory	Church Attendance	Age	Middle and upper class
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Democracy	0.32 (0.29)	0.25*** (0.09)	0.23** (0.11)	0.18 (0.14)	0.06*** (0.02)	1.89* (1.03)	0.27** (0.13)
Log (GDP per capita)	0.00 (0.26)	0.15 (0.10)	0.29 (0.22)	0.15 (0.17)	0.02 (0.03)	-0.84 (1.21)	-0.08 (0.20)
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	235	337	337	337	1447	1775	341
R-squared	0.7813	0.6153	0.6292	0.5886	0.3181	0.3308	0.6116

Note: The estimation method is OLS. Standard errors clustered at the country level in parentheses. * significant at 10%; **significant at 5%; ***significant at 1%. The sample is a panel of 198 countries: 1848-2004. Each observation is for the first year a new leader is selected. The democracy and per capita income variables are measured in the first year that the leader is selected. Fidelity is a dummy that is equal to 1 if the leader is coded as “probably faithful” or “probably unfaithful”, and zero if the leader is coded as “definitely unfaithful”. Trust is a dummy variable that is equal to 1 if the leader is coded as being “seemingly trusting, open, above-board, frank”, or “equivocal (neither specially trusting not overly suspicious)”, and zero if the leader is “seemingly distrustful, suspicious, secretive, deceptive”. Friends is a dummy variable that is equal to one 1 if the leader is coded as having “close long-term friends”, , and zero if leader is coded as “having social friendship with people such as business partners but without sharing sensitive, personal feelings” or “not having any close friends”. Oratory is a dummy that is equal 1 if the leader is coded as having oratory skills as demagogue, hypnotic speaker, eloquent, spell-binder, or good debater, effective speaker, and zero if the leader is coded as being dull, boring, absence of oratory skills. Church attendance is a dummy that is equal to 1 if the leader attends church regularly, and zero otherwise. Age is age of the leader when they first take office. Middle Upper class is a Dummy that has value 1 if the Social status of the most successful parent is middle-stratum or upper-stratum.

Table 8: Democracy and Education: Robustness

Method:	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	Core result from Table 1	Omitting the 169 leaders who come to power following a coup	Omitting the 44 leaders who come to power in the year of a transition to autocracy	Omitting the 52 leaders who come to power in the year of a transition to democracy	Omitting the 96 leaders who come to power in the year of a transition to autocracy or democracy			
Democracy	0.22*** (0.5)	0.18*** (0.006)	0.20*** (0.05)	0.24*** (0.05)	0.22*** (0.05)			
Democracy t-1								0.14*** (0.05)
Elected						0.19*** (0.04)		
Elected & Democracy							0.24*** (0.05)	
Log (GDP per capita)	-0.00 (0.06)	-0.03 (0.06)	-0.01 (0.06)	0.002 (0.05)	-0.01 (0.06)	-0.00 (0.06)	0.01 (0.06)	-0.05 (0.05)
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1133	1006	1089	1081	1037	1074	1074	1079
R-squared	0.3800	0.3974	0.3901	0.3951	0.4057	0.3695	0.3779	0.3822

Note: The estimation method is OLS. Standard errors clustered at the country level in parentheses. * significant at 10%; **significant at 5%; ***significant at 1%. The dependent variable is a dummy that has value 1 if the leader has a graduate degree and zero otherwise.

Figure 1: Average education of Democracies and Autocracies

Average Education of Democracies and Autocracies from 1874 to 2004

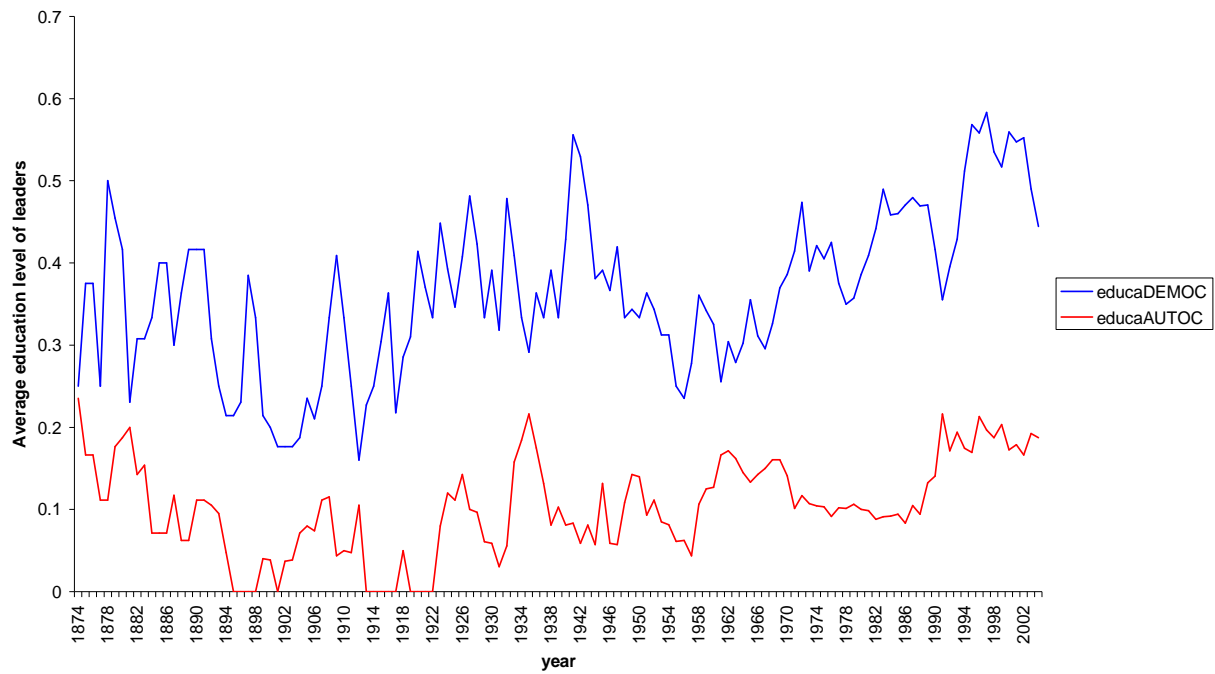


Figure 2: Democratic Transitions

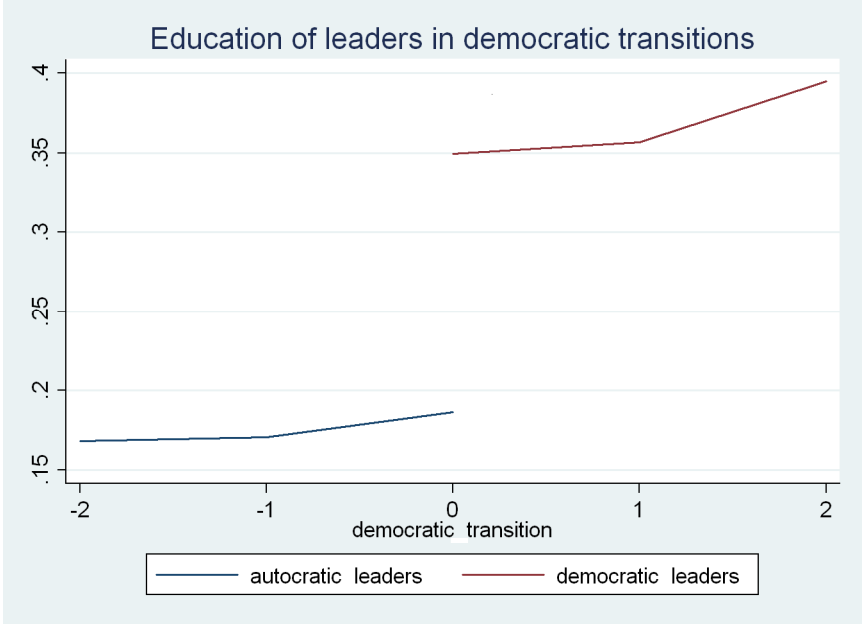


Figure 3: Autocratic Transitions

