Political Institutions and Policy Outcomes: What are the Stylized Facts?*

Torsten Persson† Guido Tabellini‡

This version: August, 2002
First Draft: July 2000.

Abstract

We investigate the effect of electoral rules and political regimes on fiscal policy outcomes in a sample of democracies, exploiting both cross country and time series variation in the data. Presidential regimes lead to smaller governments, while majoritarian elections lead to smaller governments and smaller welfare programs. Different constitutions are also associated with different spending patterns over time and different cyclical response to income shocks. Some of these empirical regularities are in line with recent theoretical work; others still await a theoretical explanation.

*We are grateful for useful comments from Alberto Alesina, Tim Besley, Per-Anders Edin, Felix Oberholzer-Gee, David Strömberg, Jakob Svensson, two anonymous referees and participants in several seminars and conferences. We would also like to thank Christina Lönnblad for editorial assistance and Gani Aldashev, Alessia Amighini, Alessandra Bonfiglioli, Agostino Cosnolo, Thomas Eisensee, Giovanni Favara, Alessandro Riboni, Davide Sala and Francesco Trebbi for research assistance at various stages of the project. This research is supported by a TMR-grant from the European Commission, and by grants from Bocconi University, MURST, and the Swedish Research Council.

†IIES, Stockholm University; London School of Economics; CEPR; NBER.

‡IGIER, Bocconi University; CEPR; Ces-Ifo.

1. Introduction

How do electoral rules and forms of government influence fiscal policy? Despite a recent wave of theoretical research, empirical work on this topic is still scant. In this paper we try to fill this gap: we estimate the effect of electoral rules and forms of government on the size and composition of government spending.

Even though the contribution of this paper is empirical, it is firmly motivated by theory. A recent line of theoretical research contrasts fiscal policy outcomes under proportional and majoritarian elections, or presidential and parliamentary forms of government. Its general predictions are that proportional electoral systems and parliamentary regimes should be associated with more public goods, larger and more universalistic welfare programs, and a larger overall size of government.

Specifically, Lizzeri and Persico (2001), Milesi-Ferretti, Perotti and Rostagno (2002), and Persson and Tabellini (1999, 2000, Ch. 8 and 9) all formally model how electoral rules influence the *composition* of government spending. Though emphasizing somewhat different ideas, these models all predict that proportional elections tilt the composition of public spending towards program benefiting large groups in the population, such as public goods or universalistic welfare programs. One reason is district magnitude (how large a share of the legislature is elected in a typical district). With proportional elections, legislators are elected in large – often national – districts, giving parties strong incentives to seek support from broad coalitions in the population. Majoritarian elections are conducted in smaller districts, inducing politicians to target smaller, but pivotal, geographical constituencies. Another reason is the electoral formula (how votes are converted to legislative seats). The size of the minimal coalition of voters needed to win the election is smaller under winner-takes all, plurality (or majority) rule, because a party can win with only 25 % of the national vote: 50 % in 50 % of the districts. Under full proportional representation (PR) it needs 50% of the national vote; politicians are thus again induced to internalize the policy benefits for larger (or cross-district) segments of the population, which leads them to emphasize broad programs.[1]

These theoretical papers take the number of parties as given (and often equal to two) and independent of the electoral system. But, as emphasized by political

[1] Perotti, Milesi-Ferretti and Rostagno (2002) make a slightly different distinction, namely between programs targeted towards social groups and programs targeted to geographic groups (with proportional elections tilting spending towards the former type).

scientists (e.g., Rae, 1967, Taagepera and Shugart, 1989, Lijphart, 1990), empirically majoritarian elections are associated with fewer parties. Related to this, majoritarian parliamentary systems are more likely to produce single-party majority governments, whereas coalition and minority governments are more likely under proportional elections (Taagepera and Shugart, 1989, and Strom, 1990). Combined with recent theoretical work – which still takes party structure as given – these regularities imply larger *size* of overall government spending under proportional elections. Thus, Austen-Smith (2000) shows that the interaction between elections, redistributive taxation, and the formation of economic groups is likely to produce politico-economic equilibria with higher taxation and overall spending under PR than under plurality. Second, Kontopoulos and Perotti (1999) emphasize that the common-pool problem in fiscal policy might be more pervasive under coalition governments, and that this is likely to lead to larger government spending. Perotti, Milesi-Ferretti and Rostagno (2002) also predict that proportional rule leads to larger overall spending.

How the form of government influences fiscal policy has not been studied as extensively. In their formal model, Persson, Roland and Tabellini (2000) make the distinction between presidential and parliamentary regimes on the basis of whether the executive is accountable to the legislature through a confidence requirement (we use the terms "forms of government" and "regimes" interchangeably throughout the paper). Building on the work of Diermeier and Feddersen (1998), they exploit that this requirement induces more "legislative cohesion": a stable majority of legislators tends to vote together on legislation, pursuing the joint interest of its voters. For this reason, spending in parliamentary regimes provides benefits to a majority of voters, such as with broad social security and welfare programs. Moreover, this stable majority becomes a residual claimant on additional revenue and benefits from high taxes and spending. In presidential regimes, by contrast, legislative coalitions are more unstable and different minorities fight over different issues on the legislative agenda. The resulting allocation of spending targets powerful minorities, typically the constituency of the powerful officeholders such as heads of congressional committees. These minorities are not residual claimants on revenue and resist high spending, exploiting stronger checks and balances and the greater dispersion of veto rights in the presidential regimes, compared to parliamentary regimes. These forces produce smaller governments and smaller social transfer programs in presidential regimes.

A large and interesting empirical literature examines how a number of constitutional features in state and local governments correlate with fiscal policy outcomes, particularly in the US (see the excellent survey by Besley and Case, 2002). But when it comes to electoral rules and forms of government, the most interesting institutional variation must be sought across different countries. Very little research has exploited that variation. A few political scientists, like Huber, Ragin and Stephens (1993) and Castles (1998), have studied the relation between these constitutional features and broad measures of fiscal policy, although indirectly and in data sets encompassing about 20 developed democracies, without obtaining robust results.[2] Among economists, Milesi-Ferretti, Perotti and Rostagno (2002) do ask whether electoral rules influence the size and composition of government spending as theory predicts, relying on a sample with 20 OECD countries (going back to 1960) and 20 Latin American countries (for a shorter period). The support for proportional elections inducing larger governments and larger transfer payments is strong in the OECD data, but much weaker in Latin American data. Persson and Tabellini (1999) study the influence of forms of government, as well as electoral rules, on the size and composition of government in a cross-section of about 50 countries from the early 1990s. They find strong support for the prediction that presidential regimes have lower spending, but less robust support for other hypotheses.

In this paper, we use two new and more extensive data sets to estimate the effect of electoral rules and forms of government on the size and composition of government spending. One is a cross-sectional data set based on information from 85 democracies in the nineties. The other is a panel of 60 democracies for which we could collect annual data for (most of) the period 1960-98. Section 2 describes these data sets in more detail. Here, we explain the overall sample-selection criteria, the measurement of the constitutional rules and policy variables of interest, as well as a number of historical, geographical, cultural, social and economic variables used in our study. While these data sets are extensive, they

[2] Huber at al (1993) argue that presidentialism, as well as majoritarian elections, produce dispersed political power and multiple points of influence on policy and that this might hamper welfare-state expansion, an argument similar to that in the formal models discussed above. Their empirical work uses a constitutional index mixing five different constitutional provisions, including the rules for elections and the form of government. This index has a strong negative influence on different measures of welfare-state expenditures, when a number of other economic and social variables are held constant, in annual data for 17 developed democracies over 30+ years.

Castles (1998) asks how a number of economic, social, and political variables shape economic policy, including the size of government and the welfare state, in 21 developed OECD democracies. One of these variables is a modification of the Huber at al constitutional indicator. Castles finds little effect of this indicator on outcomes, mostly on the basis of bivariate analysis.

do not include information on the kind of political outcomes discussed above – like party structures or types of government – through which the constitution may help shape policy outcomes. Our estimates of the constitutional effects on fiscal policy are thus necessarily on reduced form: we cannot identify whether they run directly through the incentives of politicians or voters, or through some other channel.

The data reveal that fundamental constitutional reforms are very rare. Hence, our inference about the effect of constitutions on policy outcomes must necessarily be identified from the cross-country variation in constitutions. This raises a number of statistical issues. The main challenge is that constitution selection is not random, and countries with different constitutions also differ in many other respects. How can one separately identify the effect of the constitution from that of other observable and non-observable policy determinants? To cope with this fundamental problem, we exploit information on constitutional history. We also use a variety of econometric techniques developed by labor economists to estimate the effect of policy programs on individual performance. Our empirical strategy is discussed and described in Section 3 before we embark on cross-sectional estimation.

In Section 4, we address the main question in a different fashion, exploiting also the time variation in fiscal policy and other observable policy determinants. Here, we use another empirical strategy, holding constant country-specific characteristics on policy – including any direct effect of time-invariant constitutions – by fixed-effects estimation. We then ask whether the time variation in fiscal-policy variables, and their response to economic shocks, is systematically related to the electoral rule and the form of government. This exercise is informative, but the indirect constitutional effects estimated in this section are less clearly related to the theory than the direct effects estimated in Section 3.

Our results strongly indicate that the political constitution has a causal effect on fiscal policy. A central finding is that the electoral rule exerts a very strong influence, in line with the priors from the theory. According to the cross-country evidence, a switch from proportional to majoritarian elections reduces total government spending by almost 5% of GDP and welfare spending by 2-3% of GDP. This finding is also supported by the panel-data analysis. The increase in overall and welfare-state spending taking place in the 1970s and 1980s was much more pronounced in proportional than in majoritarian countries; the cumulative difference across electoral rules amounts to about 5% of GDP for total government spending, and about 2% of GDP for welfare spending (remarkably similar to the cross-sectional effect).

The data also strongly support some predictions concerning the form of government. The cross-country evidence indicates that presidentialism reduces the overall size of government by as much as majoritarian elections, about 5% of GDP. The bulk of this difference can be traced back to a lower growth of government spending in presidential regimes during the 1970s and 1980s. Compared to parliamentary regimes, government spending in presidential democracies is also much less persistent and has a much more dampened response to common unobserved events. Presidential regimes have smaller welfare spending, in line with our theoretical priors. But here the estimated constitutional effects are less robust and it is more difficult to identify an effect of the constitution that is separate from that of other policy determinants.

2. Data

This section discusses the key variables used in the empirical analysis and our basic specification. These data have been collected as part of a larger research program on economic policy and comparative politics. The measurement of political institutions is motivated by the theories summarized in the previous section. A data appendix to this paper gives a precise description of the data sources, while Persson and Tabellini (2003) provide a more comprehensive discussion.

Which countries and years? Our goal is to compare policy outcomes in democratic countries, either in pure cross-sections or in a panel. We study cross-sectional variation in a data set of 85 democracies in the 1990s. For these countries and a large number of variables, we take an average of the yearly outcomes over the 1990-1998 period. We refer to this data set as the nineties cross section, or the 85-country cross section. To study the variation over time, we take a subset of 60 countries for which data are available for a sufficiently long period. Here, the annual observations are kept in panel format, covering each of the years 1960-98, though for many variables and countries data are missing for some of these years. We refer to this data set as the 1960-1998 panel or the 60-country panel.

How do we define a democracy? In the nineties cross section, we rely on the surveys published by Freedom House. The so-called Gastil indexes of political rights and civil liberties (*gastil*) vary on a discrete scale from 1 to 7, with low values associated with better democratic institutions. For the countries included in our default sample, the average of these two indexes in the period 1990-98 does not exceed 5. This is a generous definition of democracy, that permits countries such as Zimbabwe and Belarus (note that we refer to the average score; both countries' scores have deteriorated considerably even after 1998). This generosity maximizes the number of countries, but we also report results for a narrower sample of better democracies, with an average score less than 3.5 in the period 1990-98. Since the countries in our sample also differ in how long they have been democracies, we record the age of each democracy (*age*), defined as the fraction of the last 200 years of uninterrupted democratic rule going back in time from the current date. In the cross-sectional analysis, we always control for the influence on policy of both the quality (as measured by *gastil*) and age (as measured by *age*) of democracy.

For the 1960-98 panel, we mainly rely on the Polity IV data set covering independent nations with a population exceeding half a million people (both criteria refer to 1998) – this index goes far back in time, while the Freedom House index is available only from the early 1970s. Specifically, we use the encompassing *polity* index, which assigns to each country and year an integer score ranging from -10 to +10, with higher values associated with better democracies. The index is based on the competitiveness and openness in selecting the executive, political participation, and constraints on the chief executive.[3] We restrict our panel to those countries and years with positive values of *polity.* Given the constraints on data availability for the other variables in the early part of the sample, this leaves us with a subset of 60 countries in the panel (all of them are also included in the nineties cross section), but some of them enter only in some years. For example, the rule temporarily excludes countries like Turkey (intermittently in the 70s and 80s), Argentina (until 1972 and between 1976 and 1982) and Chile (between 1974 and 1988). As in the cross section, we check the results for a smaller sample of years and countries with a stricter definition of democracy. Throughout, we treat censored observations as randomly missing and do not attempt to model this aspect of sample selection.

Which constitutional rule? Following theoretical work, we classify electoral rules and regime types by means of two indicator (dummy) variables: *maj* and *pres.* Majoritarian countries ($maj = 1$) are those relying exclusively on plurality rule in its previous most recent election to the legislature (lower house). Mixed and

[3]For a few (small) countries where the *polity* index is missing, we use the *gastil* scores (specifically, we regress the two scores on each other and use predicted values from this regression to replace missing observations).

proportional electoral systems are lumped together and classified as proportional ($maj = 0$). Only a few countries have mixed electoral systems, and it is difficult to tell them apart from either strictly majoritarian or strictly proportional systems. Due to the correlated features of electoral systems noted in the introduction, using district magnitude rather than the electoral formula would produce a similar but not identical classification.

With regard to the form of government, we follow the theory and classify as presidential ($pres = 1$) countries where the chief executive/cabinet (or whoever holds executive powers in fiscal policy) is *not* accountable to the legislature through a vote of confidence, and those where it is as parliamentary ($pres = 0$). Despite their directly elected presidents, France and Finland are therefore classified as parliamentary, because economic policy is controlled by a government that can be brought down by a legislative vote of no confidence. Conversely, the presidential regimes include Switzerland, where there is no popularly elected president but the permanent-coalition executive cannot be brought down by the legislative assembly. In evaluating the powers of the executives, we relied on Shugart and Carey (1992) and on local constitutional information.

There are very few changes over time in these classifications ($pres$ does not vary at all from 1960 and onwards, whereas maj changes only in a few countries, Cyprus, Fiji, France, Japan, New Zealand, and mainly during the nineties). In the nineties cross-section, we continue to treat the variable maj as binary (0 or 1) and if there was a reform we code its value *before* the reform, on the argument that it could take some time before electoral reform will impact on such slowly moving variables as the size of government or welfare spending. In the panel data, instead, the variable maj is allowed to change over time in the event of a reform (dated by the first election held under new rules).

This stability of the constitution reflects an inertia of political institutions sometimes called an "iron law" by political scientists. It implies that we can use history to explain the cross-country variation in the constitution. We construct three indicator variables to date the origin of the current constitution with reference to the periods before 1920, 1921-1950, and 1951-80 (called $con20, con2150$, $con5180$ respectively); the period after 1981 is thus the default. These indicators take a value of 1 if the origin of the current constitution (either the regime or the electoral rule) dates from one of these periods, and 0 otherwise.[4] This periodiza-

[4] The date of origin of the current constitution is defined as the year when the current electoral rule or the current form of government was first acquired, given that the country was a democracy and an indpendent nation. If there was no constitutional or electoral reform since becoming a

tion contributes to explain the distribution of constitutions in our sample. While slightly above one third of our sample has majoritarian elections, this proportion is much lower (one seventh) if the current constitution originated in the 1921-50 period, but much higher (one half) if it originated in 1951-80. The form of government instead varies monotonically over time, with presidential regimes having younger constitutions (or being younger democracies) We do not have a universal explanation for this specific pattern. But it suggests that the forces shaping constitutional rules – experience by other democracies, prevalent political and judicial doctrines, academic thinking, etc. – have shifted systematically over time.

To explain the cross country variation in the constitution, we also rely on three other variables Persson and Tabellini (2003) find useful for this purpose, namely distance from the equator ($lat01$), and the percentage of the population speaking English ($engfrac$) or a European language ($eurfrac$) as a mother tongue.[5] Countries with a larger fraction of English speakers are more likely to have majoritarian elections and parliamentary form of government, probably reflecting the influence of British culture and traditions. According to the data, European language speakers (when entered jointly with $engfrac$), are instead associated with proportional elections, possibly reflecting a more general cultural influence of Europe. Finally, countries closer to the equator are less likely to be parliamentary, perhaps reflecting a wave of colonization by the West with a more shallow influence than in other regions, and hence a weaker influence of the form of government becoming dominant in Europe.[6]

Which fiscal policy outcomes? We include fiscal-policy outcomes as suggested by the theory. Thus, we measure the size of government mainly by the ratio of central government spending (inclusive of social security) to GDP, expressed as a percentage (called $cgexp$). But we also use central government revenues as a percentage of GDP ($cgrev$). For the composition of government spending we measure social security and welfare spending (by central government) as a percentage of

democracy, the origin of the current constitution coincides with the birth date of the democracy. For six countries, the electoral rule and the form of government originate in different periods, and for these countries the indicator variables for both periods take a value of 1. See Persson and Tabellini (2003) for more details.

[5] The source for these three variables is Hall and Jones (1999), who show that they contribute to explaining growth promoting structural policies.

[6] This would correspond to the idea in Acemoglu, Johnson and Robinson (2001), who argue and exploit that countries close to the equator might have less growth-friendly institutions, due to their harsher conditions for Western colonizers.

GDP (ssw). The presumption is that broad transfer programs, like pensions and unemployment insurance, are much harder to target towards narrow geographic constituencies compared to other government outlays.

The measures of size are available for most OECD countries and some countries in Latin America for the entire period 1960-98. For many developing countries, availability is limited to the period from the 1970s and onward. Similarly, the measure of welfare-state spending does not become available until the early 1970s. The statistical source for all these variables is the IMF. For the size of government we rely on IFS data, while the welfare spending measure is extracted from the GFS database.

These policy measures vary a great deal, both across time and countries. In the 85-country cross-section the mean value of expenditures is 29.8% of GDP with a standard deviation of 10.4%, a minimum of 9.7% (in Guatemala) and a maximum of 51.2% (in the Netherlands). In the 60-country panel, too, government expenditure in a typical year ranges from below 10 percent of GDP to above 50 percent. The distribution drifts upwards over time, reflecting growth in the average size of government by about 8 percent of GDP from the 1960s to the mid 1990s. Most of this growth takes place in the 1970s and 80s. Our measure of welfare spending also shows a wide distribution at any point in time with an upward drift until at least the mid 1980s in the 60-country panel.

A natural concern is whether our measurement of central (rather than general) government will bias our inference, due to correlation between centralization and the constitutional features of interest. Unfortunately, data on general government are much less reliable than those for central government, and are available on the GFS database for only about 40 countries. But in these countries where both measures of government activity are available, centralization of spending is not systematically correlated with electoral rules or forms of government. To be on the safe side, however, we always include an indicator variable for federal states (called $federal$) in our cross-section analysis.

Which socio-economic controls? We will wish to hold constant a number of variables likely to shape government outlays and revenues. Specifically, we always include in our regressions the level of development, measured by the log of real per capita income (lyp), a measure of openness ($trade$), defined as exports plus imports over GDP, and two variables measuring the demographic composition, defined as the percentages of the population between 15 and 64 years of age (*prop1564*), and above 65 years of age (*prop65*), respectively. These variables have been shown to correlate with measures of fiscal policy in previous studies, such as Cameron (1978), Rodrik (1998), and Persson and Tabellini (1999). All these variables change through time and across countries, and thus appear both in the panel and in the cross-sectional analysis.

To control for non-observable influences on fiscal policy related to the country's geographic location or to its degree of economic development, in the cross sectional analysis we also rely on indictor (0,1) variables for OECD countries (*oecd*), and for continental location of non-OECD countries, in Africa (*africa*), in eastern and southern Asia (*asiae*), and in southern and central America including the Caribbean (*laam*). Finally, to measure the influence of colonial history, we partition all former colonies in our sample into three groups: British, Spanish-Portuguese, and other colonial origin, creating three binary (0,1) indicators for each group (called *col_uk*, *col_esp*, *col_oth*). Since the influence of colonial heritage is likely to fade with time, we weigh these (0,1) indicators by the fraction of time elapsed since independence, giving more weight to colonial history in young independent states. Colonial history dating to more than 250 years ago receives no weight at all.[7] These colonial history variables also do not vary over time, and are only exploited in the cross sectional analysis.

Finally, in some specifications of the panel regressions we use two measures of income shocks. The price of oil in US dollars (*oil*) measures shocks common to all countries; but the coefficient of this variable is allowed to vary between the groups of oil exporting and importing countries in our sample (the definition of this group changes over time). To measure country-specific cyclical shocks, we rely on the output gap, defined as the log difference between real GDP and its trend computed with the Hodrick-Prescott filter (*yshock*).

All these variables are defined more precisely in the Data Appendix.

A preliminary look at the data *Table 1* displays the means of several variables in our nineties cross section of 85 countries, broken down by constitutional groups. Clearly, both overall government size and welfare-state spending are much smaller in presidential than parliamentary countries, and also in proportional than majoritarian countries. It is tempting to infer from these patterns that the data support the theoretical predictions discussed in the introduction. That temptation should be strongly resisted, however. As revealed by the rest of *Table 1*, constitution selection is far from random across these two constitutional classifications,

[7] The resulting variables are called: *col_uka*, *col_espa* and *col_otha*. Thus, for instance, *col_uka* is defined as: *col_uk* * (250 - years of independence)/250.

and there are many other systematic patterns. Both majoritarian and presidential countries tend to be less economically advanced and have worse democratic institutions and younger populations than the proportional and parliamentary counterparts. Presidential regimes are also found in more closed economies and younger democracies than parliamentary regimes. The geographic distribution also appears non-random: presidential regimes are largely located in the Americas, while continental Europe is predominantly proportional and parliamentary. Moreover, colonial history is still a good predictor of the current constitutional form: former UK colonies tend to be parliamentary and majoritarian democracies, while former Spanish colonies tend to be presidential. All these other differences among constitutional groups might fully explain the observed differences in fiscal policy, with no causal effect left for the constitution. Moreover, these groups of countries almost certainly differ also in other dimensions, which we could not identify or observe empirically. Causal inference about the effect of constitutions on policy outcomes requires additional identifying assumptions and more sophisticated statistical techniques. This is what we turn to in the next section.

3. Policy variation across countries

In this section, we exploit the cross-country variation in fiscal policy. A first subsection outlines our empirical strategy. The following subsections exploit this strategy in estimating the constitutional effects on the size of government and welfare spending.

3.1. Estimation strategy

Our empirical model can be thought of as consisting of two equations. One is a stochastic process that assigns a constitution to each country i. To simplify the exposition, suppose there is just one constitutional dimension that can take two possible values, $S_i = 0, 1$. Then we can write the stochastic process for the constitution as:

$$\begin{aligned} S_i &= 1 \quad \text{if } \; G(\mathbf{X}_i) + e_i \; > 0 \; , \\ S_i &= 0 \quad \text{otherwise} \; , \end{aligned} \tag{3.1}$$

where $\mathbf{X}$ is a vector of observables, such as colonial origin or geographic location, while e is an additive unobserved error term. The second equation determines a fiscal policy outcome (Y) in each country, as a function of the constitution, of a vector of other observable controls ($\mathbf{Z}$) possibly overlapping with $\mathbf{X}$, and of an additive unobserved error term u:

$$Y_i = F(S_i, \mathbf{Z}_i) + u_i \ . \tag{3.2}$$

Our goal is to estimate the effect of a constitutional reform – a shift from $S = 0$ to $S = 1$ – on the fiscal policy outcome.

Our first estimation method imposes two standard assumptions. (i) Recursivity: the error term e of the constitution selection equation (3.1) is uncorrelated with the error term u of the policy outcome equation (3.2). This assumption is also known as "conditional independence", or "selection on observables". (ii) Linearity: the function F in (3.2) is linear and with constant coefficients, so that the only effect of the constitution is on the intercept of the function F. By this assumption, the effect of the constitution on policy is fully captured by the coefficient of the constitutional indicator, S; and by conditional independence, this coefficient can be consistently estimated by OLS, which is our first estimation method. To make the conditional independence assumption more credible, we use a rich baseline specification, where the vector of observables $\mathbf{Z}$ always includes per capita income (lyp), openness ($trade$), the demographic variables ($prop1654$ and $prop65$), the age and quality of democracy (age and $gastil$) and dummy variables for federal and OECD countries ($federal$ and $oecd$).

Given the non random distribution of constitutional features in our sample, conditional independence is a strong assumption. Historical variables determining the current constitution could also influence policy outcomes. This is not a problem if all the common historical determinants of policy outcomes and constitution are included in the regression (and if the model is linear). For this reason, when estimating by OLS we typically add among the regressors the sets of indicators for continental location and colonial history introduced in Section 2.

But how do we know that we have included enough common determinants of policy outcomes and of the constitution to really satisfy the conditional independence assumption? If some omitted determinant of policy outcomes are correlated with the constitution, conditional independence is violated and the OLS estimates of the constitutional effect are biased. The sign of the bias is hard to pin down precisely in a multivariate context, but is likely to reflect the sign of the correlation coefficient between the error terms u and e of (3.2) and (3.1). Of the possible sources of simultaneity bias, we believe that this prospective instance of "omitted variables" is a much more important problem than that of "reverse causality". A direct feedback from policies to constitutions seems hard to reconcile with two features discussed in Section 2: changing policies over the last forty years and very few registered constitutional reforms during the same period of time.

To relax the conditional-independence assumption, we use two estimation methods: the so-called Heckman correction and instrumental variables. Both of these entail an explicit estimation of the constitution selection equation (the first stage) as well as the policy outcome equation (the second stage). In the Heckman correction, we estimate the first-stage equation (3.2) with a probit model. This gives us an estimate of the correlation coefficient between the error terms e and u of (3.2) and (3.1), which can be used to correct for the bias in the OLS estimates. Identification is made possible by an exclusion restrictions discussed below, plus a strong functional form assumption: (3.2) and (3.1) are linear, and the error terms u and e are jointly normal.

With instrumental variables, we estimate the constitution selection equation with the linear probability model (which is more robust compared to probit when estimating by instrumental variables, see Angrist and Krueger, 2001). The identification assumption is an exclusion restriction. We exploit historical variables correlated with the constitution and assume that they are uncorrelated with the unobservable determinants of policy outcomes, u. Throughout, the policy outcome equation has the same baseline specification as in the OLS estimation, with or without the dummy variables for colonial origin, as noted below. The detailed specification of the stochastic process for the constitution is discussed in context, when implementing these estimation strategies.

OLS, instrumental variables and the Heckman procedure, all exploit the assumption that F in (3.2) is linear and with constant coefficients. Usually, linearity is taken to be a convenient local approximation of a more general model. But here we are interested in the comparison of very different groups of countries. As shown in *Table 1*, almost all variables differ considerably across constitutional groups. Suppose as is plausible that the effect of the constitution on policy outcomes is stronger in richer countries or in better democracies. As these features differ systematically across constitutional groups, the local approximation is no longer tenable and the linear estimates are biased. Rather than trying out specific cases (out of an infinite possibility) of such interactions, we address the problem in a more general way. Specifically, we relax linearity and estimate the effect of the constitution on fiscal policy with non-parametric matching methods, based on the propensity score. In doing so, however, we once again have to rely on the conditional-independence assumption.[8]

[8] These methods were introduced into economics as tools for evaluating labor market and

The gist of these non-parametric estimators is that they give more weight to comparisons of similar countries, to reduce the effect of any non-linearities. Countries are ranked on the basis of their "propensity score". In our context, the propensity score can be defined on the basis of (3.1), as the conditional probability that country i is in constitutional state $S_i = 1$, given the vector of observable constitutional determinants $\mathbf{X}$. Some countries in this ranking actually have $S = 1$, others don't. The main idea is that the actual assignment of constitutions to countries with similar propensity scores is largely random. It is therefore appropriate to compare the policy outcomes across different constitutions. There are many possible ways of performing the matching underlying this comparison of similar countries, and each method corresponds to a specific matching estimator.

We rely on three alternative matching estimators. All three estimate the effect of the constitution on fiscal policy for a country drawn at random from our sample.[9] With the *stratification* estimator, countries are ranked on the basis of their estimated propensity scores and grouped into different strata, each stratum made up of similar countries. Inside each stratum, we compute the average difference in policy outcomes between countries with different constitutions. We then weight each stratum by the number of countries it contains, to produce an overall estimated difference in policy outcomes. The *nearest neighbor* estimator only compares countries that are closest in ranking. For each country with $S = 1$, we find its closest twin in the opposite constitutional state and compute the difference in policy outcomes. Then the procedure is reversed, and for each country with $S = 0$ we find its closest twin with the opposite constitutional state, and compute the difference in policy outcomes. Finally, the *kernel* estimator combines the logic of the previous two estimators. Each $S = 1$ country is a matched against a weighted average of all $S = 0$ countries within a certain propensity-score distance, with weights declining in that distance. And conversely when we match the $S = 0$

education programs (see for instance Dehejia and Wahba, 1999, and Heckman, Ichimura and Todd, 1997). More recently, they have been applied to cross country-comparisons in a variety of studies -see Persson and Tabellini (2003) for additional references. A useful and accessible survey, which puts the methodology in context, can be found in Blundell and Costa Dias (2000). More general discussions about matching vs. other evaluation methods can be found in Angrist and Kreuger (1999), Heckman, Lalonde and Smith (1999), and Ichino (2001).

[9] That is, we estimate what is also known as the *average* treatment effect. This corresponds to what was estimated in the two previuos subsections. Sometimes the literature on program evaluation is interested in other effects, such as the effect of treatment on the treated. See Heckman et al. (1999) for alternative estimators and definitions.

countries.[10] As for the other estimators, we discuss the detailed implementation in context.

A few additional points. Throughout the section we measure the constitution by the two indicators, *pres* and *maj*, defined in the previous section. In the OLS estimation we add both these indicators to our egressions. The estimated coefficient of the presidential indicator *pres* measures the effect of switching from a proportional-parliamentary to a proportional-presidential system, under the constraint that it coincides with the switch from a majoritarian-parliamentary to a majoritarian-presidential system. A similar assumption is made for the coefficient on *maj*. When estimating with OLS estimation we check whether these assumptions of additivity seem to be fulfilled. When estimating by instrumental variables, we allow for the joint endogeneity of both *pres* and *maj*. But when we implement the Heckman procedure and the matching estimation based on the propensity score, we do it one constitutional dimension at a time, first estimating the selection process for the form of government while neglecting the electoral rule (or treating it as randomly assigned), then repeating the same procedure for the electoral rule while neglecting the form of government. We have too few observations to reliably implement these procedures for multiple constitutional states (see Lechner, 2000 discusses how to generalize the propensity-score methods to multiple program treatments).

3.2. Size of Government

The theory reviewed in the introduction predicts that presidential regimes cause smaller governments. Some models also predict the same causal effect of majoritarian electoral rules. In this subsection we ask whether these predictions are consistent with the evidence.

OLS estimates We start by imposing the conditional independence and linearity assumptions, estimating equation (3.2) by OLS. The most parsimonious specification, in column 1 of *Table 2*, relies on our nineties cross-section. It holds constant the standard controls in $\mathbf{Z}$, but not colonial origin or geographic location. Presidential countries have smaller governments by 6% of GDP. The point estimate is not only highly statistically significant, but also economically and

[10] These estimators are quite common in the applied microeconometric literature. See Ichino (2001) or Persson and Tabellini (2003) for more details.

politically significant. Majoritarian elections appear also to produce smaller governments, but here the effect is smaller, about 3% of GDP, and less precisely estimated.

The next column adds our indicator variables for geographical location (Africa, Asia and Latin America) and colonial origin (UK, Spanish, and other). Because these variables are correlated with constitution selection, the conditional-independence assumption becomes more credible in this more comprehensive specification. Some of the continental and colonial-origin dummy variables do indeed have a statistically significant impact on policy. But the estimated constitutional effect of presidential regimes is remarkably stable, dropping just slightly and maintaining about the same level of precision. The estimated effect of majoritarian elections now approaches 5% of GDP. These results are quite robust to more parsimonious specifications of the continental dummy variables and the colonial origin variables, dropping one set of dummies but not the other, and to adding other controls such as income inequality, ethnic and linguistic fractionalization, population size, or a dummy variable for former socialist countries.

In column 3, we break down the constitutional variables into a finer partition ($majpres$, $majpar$ and *propres*). The effects of the two constitutional features indeed appear additive, so that introducing both a presidential form of government and majoritarian electoral rules in a proportional-parliamentary country would reduce the size of government by a whopping 10% of GDP.

In column 4, we measure the size of government by revenue instead of spending (the variable *cgrev* rather than *cgexp*). The effect of presidential regimes is the same as before, while the effect of majoritarian elections is now weaker.

The 80 countries in our broad sample for the nineties (where we have data for all variables) include some dubious democracies. In weak democracies, the formal constitution might play a less important role as compared to other informal practices and norms. Moreover, some of the weaker democracies tend to be presidential regimes, which might introduce systematic biases. Column 5 thus restricts the estimation to the better democracies in a narrower sample (62 countries for which *gastil* is smaller than 3.5). The effect of presidential regimes now appears to be even stronger, whereas the effect of majoritarian elections remains stable and (borderline) significant.

What happens when the average size of government is computed for a longer time period than the nineties? Column 6 reports on the same specification as column 2, when the dependent variable is the average (of *cgexp*) taken over the whole time span in our 60 country panel (i.e., since 1960. The effect of both presidentialism and majoritarian elections is still negative, but neither estimate is significantly different from zero. The weaker results do reflect the different time period, rather than the different sampling of countries. To show this, column 7 in the table returns to the nineties cross section, restricting the sample to the countries included in the longer panel. These results strongly suggest that the differences observed in the nineties data largely result from faster *growth* of government over the last forty years in countries with parliamentary regimes and proportional elections. We will return to this important theme in Section 4.

In summary, when we impose the assumptions of conditional independence and linearity, the negative constitutional effect of presidential regimes is large and robust to the specification. Majoritarian electoral rules also cut the size of government. Both these effects conform to prior expectations from theory. They are stronger in the later period, suggesting that the constitution has influenced post-war growth in the size of government.

Relaxing conditional independence: Heckman and IV estimates How robust are the previous result when we try to relax conditional independence? This is the question we now address, starting with the Heckman procedure and then turning to instrumental variables estimation.

A crucial step in both estimation methods is to specify the determinants of constitution selection in the first stage. We impose a similar, albeit not identical, first-stage specification in the two cases. Consider first the Heckman procedure. As noted in Section 2, the current constitution is well explained by historical variables such as its date of origin and the cultural influence of the West and of Great Britain in particular. We measure the date of origin of the constitution by our three indicator variables $(con20, con2150, con5180)$, together with the age of the democracy (age) – recall that the three dating variables capture the origin of the current constitution *or* the date of becoming a democracy, whatever came last. We measure the cultural influence of Great Britain and of Western Europe by the fraction of the population whose mother tongue is English $(engfrac)$ or a European language $(eurfrac)$ and by the distance from the equator $(lat01)$. All these variables are thus always included in $\mathbf{X}$, the vector of observable constitutional determinants. Since many countries in Latin America tend to be presidential systems with proportional legislative elections, we also include in the vector $\mathbf{X}$ a dummy variable for Latin America $(laam)$. This set of variables has considerable explanatory power: the pseudo R^2 of the probit equation for the Heckman procedure is about 51% for the form of government, about 47% for the electoral rule (results not shown here).

The second-stage estimates for the Heckman procedure are reported in columns 1 and 2 of *Table 3*.[11] The policy outcome equation) is specified with the usual set of regressors. To minimize the necessary adjustment for the correlation between unobserved determinants of constitution selection and performance, we also include dummy variables for colonial origin and continental location. The estimated constitutional effects remain negative and strongly significant. Allowing for endogenous selection of majoritarian elections (column 2), the estimated correlation coefficient between the random parts of constitution selection and performance (rho in the table) is practically zero. Thus, the estimate is similar to the OLS estimates. When we allow for endogenous selection of presidential regimes (column 1), the correlation coefficient is instead 0.43, positive and considerably higher. Thus, the OLS estimates are likely to be upward-biased, and the Heckman correction produces a larger negative estimate of the constitutional effect. These results are quite robust to alternative specifications of the first-stage equation for constitution selection.

Next, we turn to instrumental-variable estimation. Here, we exploit the crucial exclusion restriction that some variables entering the first stage do not influence fiscal policy, except through their effect on the constitution, once we control for other regressors.

We start with a parsimonious specification for both the first and second-stage regression. The second-stage regressors include our standard controls, but not continental and colonial indicator variables. The first stage is kept as in the Heckman estimation, except that we drop the dummy variable for Latin America ($laam$). Thus, the identifying assumption is that the constitutional dating variables ($con21$, $con2150$, $con5180$), the language variables ($engfrac$ and $eurfrac$) and the latitude ($lat01$) are all uncorrelated with the remaining unobserved determinants of fiscal policy. The constitutional effects on the size of government are reported in column 3 of *Table 3*. They are similar to and – if anything – larger in absolute value than the OLS estimates of *Table 2*.[12]

[11] As noted above, we apply the Heckman correction to one constitutional dimension at the time, treating the otehr dimension as random.

[12] Thus, among the second stage regressors, only age also enters the first stage. This parsimonious first-stage specification is chosen to avoid excessively weak instruments. Imposing the restriction that only age plus the six instruments enter the first stage, we estimate the first stage by OLS, run the second stage on the predicted values of maj and $pres$, and correct the second stage residuals as discussed by Maddala (1977, ch 11) and Wiggins (2000). The point estimates are very similar (or stronger) if all second-stage controls are added to the first-stage regression.

Our identifying assumption says that any omitted variable is not correlated with our instruments. For instance, if colonial origin or being in Latin America influences the size of government, their effect would show up in the residual of the second-stage equation (because they are omitted in column 3). This would not bias the IV estimates, however, as long as our instruments are not correlated with colonial origin or continental location. We think this is a reasonable assumption in the case of the three continental dating variables, while we are less certain about the remaining three instruments. If we assume that the first three instruments are valid, however, the validity of the remaining three can be tested via the implied over-identifying restrictions. As shown in column 3, we cannot reject the over-identifying assumptions, which reassures us that the estimates are consistent, despite the omission of colonial origin and continental location. We are also reassured by the fact that the estimates in column 3 correspond closely to those obtained with the Heckman correction in columns 1 and 2.

Nevertheless, the power of the over-identification test could be low, since the dating variables are only weakly correlated with constitution selection. Indeed, if we re-specify the first stage by omitting the more dubious instruments ($lat01$, $engfrac$ and $eurfrac$), the fit of the first stage becomes weak enough that the estimated constitutional effects become statistically insignificant (though still negative and, in the case of majoritarian elections even larger). For this reason, column 4 reports on the results when we add the most likely culprits to the second stage, namely the dummies for British colonial origin and Latin American location. The constitutional effect of presidential regimes now drops down towards its OLS estimate with a larger standard error, while the estimate for majoritarian elections becomes more pronounced.[13] An interpretation of this behavior of the estimates runs as follows. A parsimonious first stage leaves only a small share of the variation in constitutional arrangements explained by the first-stage regressors. The remaining variation is insufficient for exerting a significant influence on the size of government, once we have also included all the dummy variables to the second stage (since adding auxiliary controls keeps removing variation from the size of government).

[13]The results are very similar if the first stage regression associated with the estimates in column 4 is expanded to also include the dummy variable for Latin America. Adding all the colonial origin and continental variables to the second stage, the standard errors grow even further.

Relaxing linearity: matching estimates How robust are the results when we relax the assumption of linearity? The estimation of the propensity score associated with each constitutional dimension is a crucial step in the matching methods we use. These methods are based on two assumptions (cf. Rosenbaum and Rubin 1983). The first one is a version of conditional independence: once we have conditioned upon $\mathbf{X}$, the unobserved determinants of the constitution and policy outcomes, i.e., u and e in (3.2) and (3.1), are uncorrelated. In the specification a probit or logit model for constitution selection, we should thus not omit any variables really driving fiscal policy outcomes. This speaks in favor of an inclusive specification.[14] The second assumption is that the propensity score is strictly between 0 and 1 (the so called *common support* condition). To satisfy this assumption, we must obviously preserve some randomness in constitution selection. Preserving enough randomness peaks for a parsimonious logit/probit specification. If we explain constitutional choice "too well", we shrink the region of overlapping propensity scores between countries having different constitutions. For extreme observations, with probabilities close to 0 or 1, matching becomes difficult it is hard to find comparable countries in the opposite constitutional state.

We have experimented with different estimation methods for the propensity scores: probit vs. logit. As the differences are minor, we only display the results for the logit estimates. We have also tried different specifications of the set of variables entering these logits. The final results are similar, but to save space we only report results for a logit formulation, which includes four potentially important determinants of the size of government, per-capita income (lyp), the share of old people ($prop65$), the quality of democracy ($gastil$), and the presence of a federal system ($federal$), plus the indicators for previous British colonies and Latin American location (col_uka and $laam$).[15]

All estimated propensity scores lie strictly in between 0 and 1. Nevertheless, to be on the safe side with regard to the common support condition, we define the *estimated* common support as the interval between the *minimum* estimated propensity score among the $S = 1$ countries, and the *maximum* estimated propen-

[14] Note that this was not a concern in the specification of the first stage for either the Heckman correction or instrumental variables. On the contrary, in the instrumental variable estimation we deliberately chose a parsimonious first and second stage, since our concern was avoiding correlation between the variables included in the first stage and the error term of the second stage. Here instead we want to avoid correlation between the error terms of the two equation.

[15] As noted in the previous subsection, we proceed one constitutional dimension at a time, estimating a propensity score for the electoral rule and one for the form of government. The specification of the logit equation is always the same, however.

sity score among the $S = 0$ countries, doing it separately for the electoral rule and for the form of government. All observations outside this support are discarded as non-comparable. This procedure reduces our sample size, but it has the advantage of excluding outliers. It reinforces the idea that matching estimation relies on inference from local comparisons among similar countries.

A natural question is whether countries that end up are close in the ranking of propensity scores are indeed more similar when it comes to the distribution of observable covariates, irrespective of their constitution. To check this, we group the countries inside the estimated common support in three strata, corresponding to values of the propensity scores below 1/3, between 1/3 and 2/3, and above 2/3 (we do it separately for the form of government and the electoral rule). We then test whether the means of the controls used in the simple regressions of *Table 2* are equal in the constitutional groups of majoritarian vs. proportional and presidential vs. parliamentary, replicating the same kind of means tests as those reported in *Table 1* for the whole population, but now we do it within each stratum. In the first and second stratum we reject (at 5% level) the nul of equal means for only one variable (different variables in the two strata); in the third stratum we can never reject the nul. Given the striking mean differences for the whole sample reported in *Table 1* and the parsimonious specification of our logit, the strata define groups of countries that are remarkably similar.

Based on this metric of similarity, we compare the policy outcomes of similar countries under different constitutions. The last three columns of *Table 3* displays the results for the alternative matching estimators described in the subsection 3.1. The underlying standard errors have been estimated by a bootstrapping procedure. Notice also that the restriction to the common support means that we are typically discarding 5 to 10 observations.

Given the sample, the results are most directly comparable to those in column 2 of *Table 2* and the previous columns of *Table 3.* They confirm that our earlier results hold up: presidential regimes and majoritarian elections are both estimated to cut the size of government by about 5% of GDP. The standard errors of these estimates are larger than those of the OLS estimates, but that is to be expected as we are trading off less specification bias against higher standard errors in this non-parametric estimation. The most precise estimates are found by the Kernel estimator, which is also intuitive because this method is less sensitive to individual observations than the other two.

All in all, selection on observables problems due to non-linearities do not seem to be a major problem plaguing our earlier estimates.

Summary The three sets of results in this subsection paint a very consistent picture. If we are willing to assume conditional independence, given a large set of covariates, both constitutional effects are negative for the nineties cross section. Presidential regimes and majoritarian elections each cut the size of government by about 5% of GDP. These results are robust to relaxing the linearity assumption. Relaxing conditional independence does not change the estimated effect of majoritarian elections, whereas the effect of presidential regimes appears to be even larger. The results conform with our theoretical prior.

3.3. Composition of government

Do the constitutional effects extend to other aspects of fiscal policy? In this subsection we ask whether electoral rules and forms of government (as measured by the indicators *pres* and *maj*) shape welfare-state spending, relying on the same battery of methods as for the size of government.

OLS estimates *Table 4* reports on a variety of linear regression estimates. We hold constant the same variables as for the size of government. Here, the results are weaker than in *Table 2* and we typically do not find any significant constitutional effect. Column 1 refers to the full sample of countries in the nineties cross-section. Our measure of social transfers (*ssw*), is available for a dozen less countries than our measure of the size of government (*cgexp*). Both presidential regimes and majoritarian elections appear to reduce welfare-state spending by 1-2% of GDP, which are large numbers. But neither effect is statistically significant (p-values of 0.12 and 0.17). Results are similar in other (non-reported) specifications, such as when we drop the dummies for continents and colonial origin.

The absence of a clear constitutional effect may seem puzzling, as *Table 1* showed the size of welfare-state spending to be (unconditionally) much smaller in presidential and majoritarian countries. The key socio-economic covariate driving the result is the proportion of elderly in the population (as measured by $prop65$). When this variable is included, the estimated constitutional effects are statistically insignificant; when it is omitted, they are negative and significant. In other words, presidential and majoritarian countries do have smaller welfare spending on average, but this largely seems to reflect their younger populations.

In column 2, the constitution is further subdivided into four separate groups. Reform from a parliamentary to a presidential regime, maintaining proportional elections, is now estimated to have a larger negative effect on the welfare state, which is close to statistically significant. Columns 3 and 4 restrict the sample to better democracies and to those countries belonging to our longer panel. Imposing these two restrictions, we are left with about 55 observations (not the same for the two criteria). The estimated effect of a presidential regime is negative and significant, as predicted, in the better democracies. The effect of majoritarian elections, albeit negative, remains imprecisely estimated. Column 5 relies on data from our 60-country panel (the social-transfers data run from 1972 to the mid 1990s for most countries). The results for this longer time average are similar to those in column 1. They are weaker than the findings by Milesi-Ferretti, Perotti and Rostagno (2002), who consistently estimate a negative and significant effect of less proportional electoral rules on social transfers in the OECD countries from 1960 to 1995.[16]

IV and Heckman estimates Next, we relax conditional independence in the broad sample of the nineties cross section, where we have the most countries, using the Heckman procedure and instrumental variables. The first-stage specification is identical to that for the size of government for both the Heckman and 2SLS estimates.

As in the section on size, *Table 5* reports two specifications for the second-stage instrumental-variable estimates, one exclusive of the dummy variables for British colonial origin and Latin America (column 3), the other inclusive of these variables (column 3). Our previous concerns about the validity of the instruments remain, but are not repeated. Now, the over-identifying restrictions can be rejected at the 5% level for the less parsimonious second-stage specifications (column 4).

Despite these concerns, the pattern of the constitutional effects is consistent across the estimates reported in *Table 5*, yet different from the OLS estimates in *Table 4.* The presidential effect is practically zero: this is apparent in columns 1 and 3-4 (column 2 does not allow for endogenous selection into presidential regimes, so it is similar to the OLS estimates). As shown in column 1, the estimated correlation coefficient between the unobserved determinants of constitution selection and performance is weakly negative (*rho* is -0.12). If correct, this implies that the OLS estimate for presidential regimes in *Table 4* is slightly biased downwards; adjusting for this bias produces a negative estimate, but one that is not significantly different from zero. This result is confirmed by the instrumental-variable

[16] Restricting the regressions for the longer cross section to the 25 OECD countries in our sample (including the same covariates except the continental and colonial-origin dummies), however, we obtain an insignificant effect close to zero.

estimates, which are also closer to zero than the OLS estimates.

The effect of majoritarian elections, on the other hand, is now negative and statistically significant according to both procedures. Column 2 in *Table 5* suggests errors with a strong positive correlation (rho is $+0.50$), implying an upward bias in the OLS estimate of the constitutional effect in *Table 4.* When the bias is corrected, the constitutional effect of majoritarian elections becomes negative and statistically significant (column 2). This result is confirmed by our instrumental-variable estimates (columns 3-4).

The consistency of these results is an indication that accounting for deviations from conditional independence might be important. Once this is done, there is stronger evidence that majoritarian elections do induce a smaller welfare state, whereas the form of government appears to be unimportant.

Matching estimates Finally, we turn to the matching methods, thereby relaxing the assumption that the welfare-state relation is linear in the covariates. The results are based on the same logit specification for the propensity score as the one we used for the size of government. The last three columns of *Table 5* display the results for our three matching methods. As the estimates refer to the nineties cross section, they should be compared to the estimated constitutional effects in Column 1 of *Table 4* and the four first columns of *Table 5.* They are quite similar: both presidential regimes and majoritarian elections have a negative effect on welfare-state spending, but the effects are imprecisely estimated and rarely statistically significant. As higher standard errors are to be expected, the consistently negative estimates still strengthen our belief that the constitutional effect of majoritarian elections is indeed negative.

Summary Our findings suggest that majoritarian elections cut welfare spending, as predicted by theory, and by as much as 2-3% of GDP. For presidential regimes, there is less evidence of a stable overall constitutional effect. In this case, selection bias seems to be a more severe problem than for the size of government: relaxing conditional independence reinforces the negative constitutional effect of majoritarian elections, but weakens the negative effect of presidential regimes.

4. Policy variation across time

We now turn to the time variation in the data by help of our 60-country panel, where data go back to the 1960s or 1970s. As noted in Section 3, the constitutional effects estimated from the cross sectional variation are stronger in cross sections from the nineties than from the whole period. The constitution might thus have influenced the dynamics of fiscal policy. As deep constitutional reforms are very rare, we cannot exploit time variation in constitutional rules. Instead, we exploit the interaction between the constitution and other time-varying variables.

We focus on two sets of interactions. Do electoral rules and forms of government modify the influence of unobserved determinants of policy common to all countries? We address this question in subsection 4.2, capturing the effect of the common determinants by means of time-dummy variables and asking whether their estimated coefficients differ across constitutional groups. Do income shocks, deviations of GDP from its long run trend, impact on fiscal policy in a way that depends on the constitution? In subsection 4.2, we focus such cyclical fluctuations. We also ask whether the response to positive and negative income shocks are any different. When it comes to these questions we do not have firm priors. We are thus not testing specific hypotheses, but aim to establish some stylized facts and, if possible, gain a better understanding of the mechanisms behind the cross-country differences uncovered in the previous section.

Compared to the previous section, we pay less attention to endogeneity (selection) of the constitution. The reason is twofold. For one, it is difficult. As we are estimating dynamic interaction effects between the constitution and other variables, allowing for endogenous constitution selection would raise a number of new econometric subtleties. For another, selection bias is less of a concern here. We always allow for country fixed effects that pick up any time-invariant and country-specific unobserved determinants of fiscal policy, such as any *direct* effects of the constitution itself, or of history, geography or culture. We focus instead on *indirect* constitutional effects, captured by interaction terms. The possibility that historical or cultural determinants of the constitution would also influence these interactions seems more remote, compared to the likelihood of a direct influence on fiscal policy.

4.1. Unobserved common events

Our fiscal-policy measures have a similarly shaped time path in most countries. A plausible conjecture is that these trends reflect some common economic and political events, such as the worldwide rise of left-wing ideologies in the late 1960s and 70s, the turn to the right in the mid 1980s, or the productivity slowdown and oil shocks in the 1970s and 80s. Our goal in this section is to find out whether and how the impact of such common events on our policy measures depends on electoral rules and forms of government. Since our main interest is in these constitutional features, we do not seek to identify and measure the common events. Instead, we treat them as unobserved and proxy for them by a set of year-specific indicator variables, asking whether these interact with the constitution. This method was suggested by Blanchard and Wolfers (2000) to study how labor-market institutions influence the reaction of unemployment to common unobservable shocks, and used also by Milesi-Ferretti, Perotti and Rostagno (2002) to compare the reaction of fiscal policy under different electoral systems in the OECD countries.

Specifically, we estimate an equation of the following form:

$$Y_{it} = \lambda^0 Y_{it-1} + S_i(\lambda^1 - \lambda^0)Y_{it-1} + \boldsymbol{\beta}\mathbf{X}_{it} + (1 + \gamma S_i)\boldsymbol{\delta}\mathbf{Q}_t + \alpha_i^* + u_{it} \ , \qquad (4.1)$$

where the coefficient on the lagged dependent variable, λ^S, is allowed to depend on the constitutional state, $S = 0, 1$. All observable controls $_{it}$ are instead assumed to have the same vector of coefficients irrespective of the constitution. $\mathbf{Q}_t$ is the time-t value of a vector $\mathbf{Q}$ of year indicators (i.e., a set of dummies one of which takes a value of 1 in year t, while the others take a value of 0).[17] Our interest is primarily in the coefficients γ (one per constitutional rule). If they are zero, the unobserved common events have the same impact on policy in all countries irrespective of the constitution; if some component γ is different from zero, this impact depends on the constitution. A positive value implies that the constitutional feature measured by $S_i = 1$ inflates the impact of common events (relative to the default constitutional feature $S_i = 0$), while a negative value of γ captures a dampening effect. To the extent that $\lambda^0 \neq \lambda^1$, the adjustment over time to this impact effect (and any changes in $\mathbf{X}$) will also differ across constitutions.

The unobserved error term is decomposed in two terms: a country-specific and fixed component, α_i^*, and a remaining error term, u_{it} , which is assumed to be identically distributed across countries and time. As suggested by (4.1), we estimate the parameters of interest by non-linear least squares, including country fixed effects.[18]

[17] In the estimation, we use a set of time dummies from 1961 to 1998 plus an intercept.

[18] If $\lambda^1, \lambda^0 > 0$, an asymptotic bias remains in our estimate of λ^S even as the number of

Throughout this section, the vector of controls $\mathbf{X}_{it}$ always includes the variables introduced in section 2, namely per capita income (lyp), demographics ($prop65$ and $prop1564$) and openness ($trade$). All these variables vary both across countries and time. But we omit time-invariant variables such as the indicators for federalism, OECD-membership, geography or colonial origin, because their effects on policy are already subsumed in the country fixed effect.

Size of government The first two columns of *Table 6* reports the estimates of the coefficient γ, for presidential regimes and majoritarian elections for the size of government spending. As both constitutional dummy variables are included in the same regression, the default group consists of proportional and parliamentary countries. The vector of estimated coefficients $\boldsymbol{\delta} = (\delta_t)$ (one per year, not reported in the table), reflects the impact of the vector of unobserved common events $\mathbf{Q}$ in this default group. The coefficient on presidential regimes ($pres$) in *Table 6* captures the difference between presidential-proportional and parliamentary-proportional countries (alternatively, between majoritarian-presidential and majoritarian-parliamentary). Similarly, the coefficient on majoritarian elections (maj) captures the difference between majoritarian-parliamentary and proportional-parliamentary countries.

In column 1, we impose the restriction that $\lambda^0 = \lambda^1 = 0$, excluding the lagged dependent variable from the regression. This specification thus forces all sources of dynamics to be captured either by the included controls or the time dummies. Since the controls included in $\mathbf{X}_{it}$ exhibit limited time variation, we are attributing a large fraction of the dynamics in government spending to the unobserved common events. The estimated values of γ reported in column 1 are negative and highly significant for both constitutional features. To interpret it consider an unobserved event in period t that raises government spending by 1% of GDP in proportional-parliamentary countries (formally, a year when $\delta_t - \delta_{t-1} = 1$). A coefficient of – 0.58 for presidential regimes means that this event only raises spending by about 0.4% of GDP (≈1 – 0.58) in presidential-proportional countries, a very large difference. The dampening effect of majoritarian elections is smaller,

countries tends to infinity; the sign of the bias has the opposite sign of the true λ^S, so that if $\lambda^S > 0$ we under-estimate persistence. The bias, which could also spill over to our estimate of γ,shrinks as the length of the panel increases, however (Hsiao 1986). In the size of government regressions, the average length of the panel is 26 years, and the bias is thus negligible. In the case of welfare spending, we have on average 16 years per country, however, and the bias could be more relevant.

with a coefficient of –0.38, but also highly relevant.

Figure 1 depicts the estimated effect of unobserved common events in three groups of countries. The uppermost line marked with diamonds refers to the default group of proportional-parliamentary countries (in each year, the line depicts the estimated coefficients δ_t pre-multiplying the dummy variable of each year in the regression of column 1). Squared and triangular shapes indicate presidential-proportional and majoritarian-parliamentary countries (i.e., each point on these lines depicts the same coefficient δ_t multiplied by the relevant $(1+\gamma)$ term). In all three groups, the size of government rose almost without interruption until the early 1980s, and then remained roughly constant. But the upward trend in the early part of the sample is much more pronounced in the proportional-parliamentary countries, compared to the other two constitutional groups. From the early 1980s onwards, instead, the time path of spending looks much more similar in all democracies. To a large extent, differences in the size of government across constitutional groups observed today – and documented in the previous section – seem to be due to events the 1960-80 period.

These time patterns across constitutional groups might appear surprising. The early 1980s coincide with the rise of conservative governments in several Western democracies. It is natural to conjecture that different constitutions imply different reactions of spending to this ideological swing to the right. But this is not what we observe. The time trend of government spending stabilizes in all countries to about the same extent and at about the same time. We have, of course, imposed the constraint it in our specification that the differences across constitutional groups stay constant over time. To relax this constraint, we re-specify equation (4.1) allowing the γ-coefficients interacted with maj and $pres$ to take on different values before 1982 (the period of average upward trend) and after 1982 (without average upward trend). But the estimates for the coefficients γ are very similar across the two sub-periods, both for the electoral rule and the form of government. The interaction between the constitution and the common unobserved events is thus the same before and after the 1980s. In the early half of the sample, however, something drove up government spending everywhere, although much more in proportional-parliamentary countries.

As government spending is highly persistent over time, it could be a mistake to attribute all unexplained variation in spending during a particular year to unobserved common events in that same year. Some of the observed variation could simply reflect a delayed response to previous events. To allow for persistence in government spending, the specification in column 2 adds to the regression the lagged dependent variable (*lcgexp*), allowing its coefficient to differ across constitutional groups. The direct effect of common events captured by time-dummy variables now play a smaller role: the estimated coefficients $\boldsymbol{\delta}$ are much smaller than in the regression of column 1 and less often statistically different from zero. Presidential regimes display significantly less persistence compared to parliamentary regimes.[19] But the evidence that the constitution interacts with the unobserved common events is now weaker: the estimates of the γ coefficient for presidentialism is smaller and statistically different from zero only at the 10% confidence level. Notice, however, that the implied cumulative differences across different regimes are still similar to those in column 1. While the estimated common events and the difference in their impact (as γ is smaller in absolute value), a sequence of positive common spending shocks still produce a larger spending binge in parliamentary regimes due to the higher inertia in spending (as λ^1 is smaller than λ^0).[20] Majoritarian elections are also associated with less inertia in spending, but now the estimated value of γ is not significantly different from zero.

Overall, these estimates suggest that presidential democracies have less inertia in spending, as well as a more dampened reaction to common unobserved events, compared to parliamentary democracies. Majoritarian countries also have less persistent dynamics, but look much more similar to proportional democracies. In other words, during the postwar period government spending increased in most countries until the mid 1980s. It increased more in parliamentary than presidential countries, both because the (generally upward) movements in spending had a larger permanent component and because spending reacted more strongly to the unobserved common events.

Welfare spending We already know from the results in Section 3 that proportional-parliamentary democracies have larger welfare states than other constitutional groups. As welfare-state spending typically stems from entitlement programs, it is likely to be highly persistent. Thus, it is natural to attribute the higher growth and persistence of total government spending in proportional-parliamentary democracies to their larger welfare states. With this motivation, we now turn to the

[19] The estimated coefficient of *pres$*$ lcgexp* in column 2 corresponds to the difference in persistence between presidential-proportional and parliamentary-proportional countries, $\lambda^1 - \lambda^0$, and similarly for the estimated coefficient of *maj*lcgexp*.

[20] This difference in the persistence of presidential regimes is robust: it remains even if we add the dummy variable for Latin America (*laam*) interacted with lagged spending.

interaction between the constitution and the dynamics of welfare spending, repeating the analysis of the previous subsection. Note that our panel is shorter in this case, as data on welfare spending is available only from the early 1970s for most countries.

The last two columns of *Table 6* contain the results for welfare-state spending in percent of GDP (ssw). In column 3, we estimate with non-linear least squares without the lagged dependent variable. The year-specific indicator variables proxying for common events now span the period from 1973 to 1998. The estimated coefficients on these variables peak in the early 1990s and remain roughly constant thereafter. At the peak, the difference with the estimated coefficient for the1973 year dummy is about 5. This means that, in the default group of proportional-parliamentary countries, the unobserved common events account for a rise of welfare spending of about 5% of GDP throughout this period. But the impact on the other constitutional groups is much smaller, as revealed by the estimated γ coefficients: –0.65 for presidential regimes and –0.35 for majoritarian elections. These estimates are remarkably similar to those for total government spending reported in column 1 of *Table 6*. The unobserved common events that raised welfare spending by 5% of GDP in the default group only raised it by about 3% of GDP in majoritarian-parliamentary countries, by 1.5% of GDP in proportional-presidential countries, and not at all in majoritarian-presidential countries.

When we add the lagged dependent variable ($lssw$) in column 4, the estimated interaction between the time dummy variables and presidentialism drops to –0.38 remaining significantly different from zero, while the interaction term with majoritarian elections goes up to –0.49. Contrary to the findings for total government spending, we can not reject the hypothesis that the lagged dependent variable has the same coefficient irrespective of the constitution.[21]

The data thus reveal important indirect constitutional effects on welfare spending. These effects are similar to those uncovered for total government spending, with some subtle differences. The dynamics of both total spending and welfare spending are more dampened in presidential than in parliamentary regimes. While total spending is less persistent in presidential countries, the constitution does not affect the persistence of welfare spending. Instead, the constitutional effect on welfare spending stems from a different reaction to common unobserved events: common events that increased welfare spending in parliamentary countries have had a smaller impact in presidential regimes. The electoral rule also shapes

[21] This remains true even if add an interaction between the lagged dependent variable and dummy variables for Latin America and UK colonial origin.

the dynamics of welfare spending, with majoritarian countries reacting less to unobserved common events. Naturally, the differences across policy measures could, to some degree, reflect different sampling periods: 1973-1998 for welfare spending, and 1961-98 for overall spending.

4.2. Income shocks

Part of the time variation in fiscal policy reflects the response to changes in other economic variables, especially shocks to aggregate income. Such responses could be the result of automatic stabilizers – for given tax schedules or remuneration rates in entitlements programs – or deliberate policy decisions triggered by the shock. In this section, we focus on such cyclical fluctuations in government spending and their interaction with the constitution. As in the previous section, our main purpose is to describe the systematic patterns in the data.

Throughout this section we estimate the following equation:

$$Y_{it} = \lambda^0 Y_{it-1} + S_i(\lambda^1 - \lambda^0) Y_{it-1} + \phi^0 yshock_{it} + S_i(\phi^1 - \phi^0) yshock_{it} + \boldsymbol{\beta}\mathbf{X}_{it} + \alpha_i^* + u_{it} \ . \tag{4.2}$$

The variable *yshock* is the percentage deviation of income from a country specific trend, as defined in Section 2. We want to know whether the effect of this variable on overall or welfare-state spending depends on the constitutional state (i.e., whether the coefficients ϕ^1 and ϕ^0 are the same or not). The other controls in the vector $\mathbf{X}$ include the same variables as in the previous section (the two population variables, openness to trade and per-capita income). We also include the price of oil (*oil*), as a proxy for economic shocks common to most countries, but we allow its coefficient to vary across oil exporting and oil importing countries. All these variables are constrained to have the same coefficients irrespective of the constitution. The constitution is measured by our two indictors for majoritarian elections and presidential regime (*maj* and *pres*) with proportional-parliamentary countries as the default group.[22]

To allow for a country-specific component of the error term, α_i^*, we estimate equation (4.2) in levels with country fixed effects. We also check that the results are robust to estimating in first differences and allowing for country-specific autocorrelation in the error term (but do not report these alternative estimates).

[22] In principle, all controls in $\mathbf{X}$ could interact with the constitution and their $\boldsymbol{\beta}$ coefficients could vary with the constitutional state. In practice, this does not happen: for most variables and most specifications, we cannot reject the nul hypothesis that the $\boldsymbol{\beta}$ coefficients are the same irrespective of the constitutional state.

Since one of the regressors (the price of oil) is common to all countries, we drop the year fixed effects to avoid collinearity.

The income shocks take on very large values (as large as 10% or more) for some observations. To avoid drawing inferences from a few outlying observations, we restrict the sample to observations where the income shocks are strictly less than 5% in absolute value (including the full sample with the outlier observations for income shocks strengthens the results reported below). Finally, we ignore a possible estimation problem: a component of the income shock could be endogenous and reflect exogenous variation in fiscal policy itself. This could bias the estimated coefficient ϕ upwards, but the bias is unlikely to seriously affect our inferences about constitutional interactions, unless the endogenous component of income shocks varies with the constitution.

Size of government We begin with overall government spending (*cgexp*). Column 1 of *Table 7* estimates equation (4.2) for the full sample of democracies, with the exclusion of income shocks less than 5% in absolute value. As in *Table 6*, government spending is much more persistent in proportional-parliamentary democracies, particularly compared to presidential regimes. The new finding here is that the cyclical response of fiscal policy varies with the constitution. In the default group of proportional-parliamentary countries, the estimated coefficient of income shocks (*yshock)* is consistently negative with a value of about –0.18, meaning that a 5% drop in real income induces a rise in spending of nearly 1% of GDP. Because spending is highly serially correlated, this effect persists over time. By contrast, in presidential countries the spending to GDP ratio is not affected by the income shock. In majoritarian countries, the contemporaneous impact effect of income shocks is also smaller than in proportional countries, but the difference is not statistically significant. Restricting attention to smaller income shocks (below 3 in absolute value) or to better democracies (where the quality indicator *polity* is above 5 rather than 0) yields very similar results.

To gain a better understanding of the adjustment in different political systems, column 2 disaggregates income shocks into positive (*posys*) and negative (*negys),* interacting them with our two constitutional dummy variables. An asymmetry is apparent. In proportional-parliamentary countries, only *negative* income shocks have a statistically significant effect on the spending ratio, and the estimated coefficient is much larger in absolute value. In presidential and majoritarian countries, instead, the absence of a cyclical reaction is confirmed. This asymmetry suggests a constitution-dependent ratchet effect. A negative income shock induces a lasting expansion in the size of government, which is not undone when income grows above potential. But this ratchet effect is not present in presidential countries, and it is smaller under majoritarian elections (though in the latter group the difference with proportional elections is not statistically significant).[23]

How can these constitutional effects be explained? The larger cyclical response of the spending to GDP ratio in proportional-parliamentary democracies could reflect their larger welfare states: the outlays of such entitlement programs are fixed in cash terms, or perhaps even inversely related to income. But the presence of a ratchet effect only among proportional-parliamentary countries is harder to explain, and suggests that the constitution might also have a direct effect on the discretionary policy reaction to exogenous events. One possibility is related to the theory discussed in the introduction. Proportional elections and parliamentary regimes both have a bias towards larger overall spending. Politicians in those systems may then be less prepared to cut spending when the economy is doing badly. Another possibility is coalition governments; as discussed in the introduction, these are more common in proportional-parliamentary countries. Such governments may induce a greater status-quo bias due to the difficulties of bargaining, highlighted by economists such as Alesina and Drazen (1991) and political scientists as Tsebelis (2002). Yet, another possibility is that some democracies are more prone to borrowing constraints. If presidential democracies are more likely to experience political crises, as some political scientists hold, they may also have more frequent debt or currency crises. Borrowing constraints would impart a procyclical bias to fiscal policy: governments must cut spending or raise revenues when hit by a recession or by a financial crisis, since they cannot let the deficit absorb the shock. Indeed, many presidential regimes are located in Latin America or Africa, where financial crisis have been more frequent, and earlier studies have shown that fiscal policy in Latin America tends to be more pro-cyclical than elsewhere – see in particular Gavin and Perotti (1997). Whatever its interpretation, this asymmetric ratcheting upwards of government spending contributes to the differential size of government in different political systems uncovered by the cross-sectional analysis.

[23] The results are robust to estimating in first differences. The result for presidential countries is also robust to interacting the income shock both with our constitutional dummy variables and with two other dummy variables, for Latin America (*laam*) and for British colonial origin (*col_uka*).The result on majoritarian elections is more fragile to the specification, perhaps also because of the high correlation bewteen the electoral rule and UK colonial origin.

Welfare spending Finally, we turn to the response of welfare spending to income shocks, estimating the same type of regression as in the previous subsection. Column 3 of *Table 7* reports the estimated response to income shocks smaller than 5% (the results are similar in the sample of better democracies or for smaller income shocks in absolute value). The results are similar to those for total government spending, with a few differences. As for total spending, welfare spending is most counter-cyclical among proportional-parliamentary democracies, and least counter-cyclical among presidential ones. Now the electoral rule also plays a role, however, with majoritarian countries responding significantly less than proportional countries. Moreover, in contrast to total government spending, inertia in welfare spending is not affected by the constitution (confirming our finding in the non-linear estimation of subsection 4.1). Finally, in proportional-parliamentary countries the cyclical response of welfare spending to income shocks is smaller than that of total government spending – cf. column 1 in *Table 7*. But for presidential countries, the reverse is true: total government spending as a faction of GDP remains constant over the cycle, while welfare spending is somewhat counter-cyclical though not as much as in parliamentary democracies.

Column 4 of *Table 7* decomposes income shocks into positive and negative. Again, we find a ratchet effect in proportional-parliamentary countries: positive income shocks inducing no effect on welfare spending relative to GDP, but negative income shocks expanding the welfare state. As for total government outlays, the ratchet is weaker in presidential democracies, but now the electoral rule also makes a difference.[24] The overall estimated coefficients are somewhat smaller than for the size of government. This suggests that the ratchet effect mainly concerns the welfare state, but that other spending items must also exhibit an asymmetric response to income shocks.

These differences across constitutions and types of spending are consistent with several mechanisms might be at work. On the one hand, proportional-parliamentary regimes have larger welfare states, and hence automatic stabilizers might be more important in this constitutional group. But the results may also reflect the stronger incentives, expected from theory, to spend in a discretional way on broad transfer programs in proportional and parliamentary systems; or, again, the propensity for these systems to generate coalition governments with a bias towards the status quo.

[24] When we interact the income shock with the dummy variables for Latin America and British colonial origin, the results survive when we estimate in first differences, but not when we estimate in levels.

5. Conclusion

Do electoral rules and forms of government shape fiscal policy? Our empirical results strongly suggest that the answer is yes. Several of these empirical regularities are in line with the first wave of theory discussed in introduction. As predicted, presidential regimes have smaller governments, while majoritarian elections lead to smaller governments and smaller welfare programs.

Other findings await a satisfactory theoretical explanation. A puzzling but robust feature of the data is the larger counter-cyclical response of spending in proportional-parliamentary countries than in presidential and majoritarian countries. Larger automatic stabilizers built into more generous welfare-state programs in proportional-parliamentary systems could account for part of this finding, but are unlikely to be the whole story. Proportional and parliamentary regimes display a ratchet effects in government spending, in the wake of negative income shocks, that we do not find under other constitutions. Government spending rose everywhere until the mid 1980s, but much more in the proportional and parliamentary groups. The cross-country differences we observe today can largely be attributed to different responses to common political or economic events between the early 1960s and the early 1980s. Why do we observe these different patterns in different political systems?

Much work remains to be done. Understanding the cyclical reaction of fiscal policy, or delayed fiscal adjustments, requires dynamic models of policymaking in different political systems. Such theory does not yet exist, as the existing predictions of comparative politics and economic policy are generally drawn from static models, in which there is no role for state variables such as government debt, and no link between current policy decisions and the future status quo.

On the policy side, we have concentrated on government spending. It would be interesting, and certainly feasible, to study other policy instruments – like the structure of taxation, including trade policy – with similar methods. On the institutional side, it would be valuable to study the effect on the policy mix of more detailed constitutional features; for instance, different types of checks and balances (legislative powers of presidents vs. congresses, or of cabinets vs. parliamentary committees), different types of confidence requirements, different barriers of entry in politics (closed vs. open party lists, electoral thresholds), to mention a few.

Most important of all, perhaps, we have estimated reduced-form effects, mapping the constitution into policy outcomes. This way, we have not been able to identify whether the constitution operates through a direct effect, for given political representation, or through indirect effects via altered political representation. The latter, in turn, may entail effects on party structures, types of government, occurrence of elections or government crises, or representation of different political ideologies. As mentioned already in the introduction, these political outcomes do vary systematically with electoral rules and government regimes. To make further progress, we must open this black box to better distinguish the different channels whereby the constitution exerts its influence on policy outcomes. This is likely to require a close interplay of theoretical and empirical work, including the collection of new data, in a domain right at the borderline between traditional economics and political science. The findings in this paper suggest that it is worth embarking on this ambitious task.

DATA APPENDIX

africa: regional dummy variable, equaling 1 if a country is African, 0 otherwise.

age: age of democracy. Defined as: $age = (2000 - dem\_age)/200$. Varies between 0 and 1, with US being the oldest democracy (value of 1). Source: see *dem_age.*

asiea: regional dummy variable, equaling 1 if a country is East Asian, 0 otherwise.

cgexp: central government expenditures as a percentage of GDP. Constructed using the item Government Finance - Expenditures in the IFS, divided by the GDP at current prices and multiplied by 100. Source: IMF - IFS CD-Rom and IMF - IFS Yearbook.

cgrev: central government revenues as a percentage of GDP. Constructed using the item Government Finance - Revenues in the IFS, divided by the GDP at current prices and multiplied by 100. Source: IMF - IFS CD-Rom and IMF - IFS Yearbook.

col_espa: $col\_espa = col\_es*(250\text{-}t\_indep)/250$. Represents the combined effect of *col_es*, which describes if a country was a colony of Spain or Portugal or not, and *t_indep*, years of independence, ranging from 0 to 250 (the latter value used for all non-colonized countries). Source: Wacziarg (1996).

col_otha: $col\_otha = col\_oth*(250\text{-}t\_indep)/250$. Represents the combined effect of *col_oth*, which describes if a country was a colony of a country other than the UK, Spain or Portugal, and years of independence. Source: Wacziarg (1996).

col_uka: $col\_uka = col\_uk*(250\text{-}t\_indep)/250$. Represents the combined effect of *col_uk*, which describes if a country was a colony of UK or not, and years of independence. Source: Wacziarg (1996).

dem_age: first year of democratic rule. Corresponds to the first year of an uninterrupted string of positive yearly *polity* values (see below) until the end of the sample, given that the country was also an independent nation. Does not count foreign occupation during WWII as an interruption of democracy. Source: See *polity*.

engfrac: measures the fraction of the population speaking English as a native language. Source: Hall and Jones (1999).

eurfrac: measures the fraction of the population speaking one of the major languages of Western Europe: English, French, German, Portoguese, or Spanish. Source: Hall and Jones (1999).

federal: (0,1) indicator variable for federalism. Source: Boix 2000.

gastil: average of indexes for civil liberties and political rights. Measured on one-to-seven scale with one representing the highest degree of freedom and seven the lowest. Countries whose combined averages for political rights and for civil liberties fall between 1.0 and 2.5 are designated "free", between 3.0 and 5.5 "partly free" and between 5.5 and 7.0 "not free". Source: Freedom House, Annual Survey of Freedom Country Ratings.

laam: regional dummy, equaling 1 if a country is Latin American, 0 otherwise.

lat01: rescaled variable for latitude. The absolute value of latitude divided by 90, returning a number between 0 and 1. Source: for absolute latitudes Hall and Jones (1999).

lyp: natural log of the per capita real GDP. Sources: Penn World Tables - mark 5.6 (PW); Easterly's series on www.worldbank.org; The World Bank's World Development Indicators (WDI).

maj: dummy variable for electoral systems. Equals 1 in presence of (exclusively) either a majority or a plurality rule, 0 otherwise. Only legislative elections (lower house) are considered. Source: Persson and Tabellini (2003).

oecd: Dummy variable for OECD member countries, taking the value 1 if a Country is an OECD member, 0 otherwise. Source: Persson and Tabellini (1999).

polity: *t*he score is computed by subtracting the $AUTOC$ score from the $DEMOC$ score; the resulting unified polity scale ranges from +10 (strongly democratic) to -10 (strongly autocratic). Source: Polity IV Project (http://www.cidcm.umd.edu/inscr/p

pres: dummy variable for government regimes. Equals 1 in presence of presidential regimes, 0 otherwise (Parliamentary). Only those regimes where the confidence of the assembly is not necessary for the executive (even if the president is not chief executive, i.e., assembly-independent) are included among presidential regimes. Premier-presidential (semi-presidential like France) and president-parliamentary systems (like Ecuador) are generally classified as parliamentary. Source: Persson and Tabellini (2003).

prop1564: percentage of population between 15 and 64 years old in the total population. Source: World Development Indicators CD-Rom 1999.

prop65: percentage of population over the age of 65 in the total population. Source: World Development Indicators CD-Rom 1999.

trade: sum of exports and imports of goods and services measured as a share of GDP. Source: The World Bank's World Development Indicators CD-Rom 2000.

yshock: deviation of aggregate output from its trend value in percent. Difference between the natural log of the real GDP in the country and its country-specific trend (computed using the Hodrick-Prescott filter).

References

[1] Acemoglu, D., S. Johnson, and J. Robinson (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation", *American Economic Review* 91, 1369-1401.

[2] Alesina, A., and Drazen, A. (1991), "Why are Stabilizations Delayed?" *American Economic Review* 81, 1170-1188.

[3] Angrist, J. and A. Kreuger (1999), "Empirical Strategies in Labor Economics", in Ashenfelter, O. and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3c, North-Holland.

[4] Angrist, J. and A. Krueger (2001), "Instrumental Variables and the Search for Identification", *Journal of Economic Perspectives* 15, 69-85.

[5] Austen-Smith, D. (2000), "Redistributing Income under Proportional Representation", *Journal of Political Economy* 108, 1235-1269.

[6] Besley T. and A. Case (2002), "Political Institutions and Policy Choices: Evidence from the United States", *Journal of Economic Literature*, forthcoming.

[7] Blais, A. and L. Massicotte (1996), "Electoral Systems" in LeDuc, L., R. Niemei and P. Norris (eds.) *Comparing Democracies: Elections and Voting in Global Perspective*, Sage.

[8] Blanchard, O. and J. Wolfers (2000), "The Role of Shocks and Institutions in the Rise of European Unemployment: The Aggregate Evidence, 1999 Harry Johnson Lecture", *Economic Journal* 100, C1-33.

[9] Blundell, R. and M. Costa Dias (2000), "Evaluation Methods for Non-Experimental Data," *Fiscal Studies* 21, 427-468.

[10] Cameron, D. (1978), "The Expansion of the Public Economy: A Comparative Analysis", *American Political Science Review* 72, 1203-1261.

[11] Castels, F. (1998), *Comparative Public Policy. Patterns of Post-war Transformation*, Edward Elgar.

[12] Dehejia, R., and S. Wahba (1999), "Causal Effects in Non-Experimental Studies: Re-evaluating the Evaluation of Training Programs," *Journal of the American Statistical Association* 94, 1053-1062.

[13] Diermeier, D. and T. Feddersen (1998), "Cohesion in Legislatures and the Vote of Confidence Procedure", *American Political Science Review* 92, 611-621.

[14] Gavin, M. and R. Perotti (1997), "Fiscal Policy in Latin America", in Bernanke, B. and J. Rotemberg (eds.), *NBER Macroeconomics Annual 1997*, MIT Press.

[15] Hall, R. and C. Jones (1999), "Why Do Some Countries Produce so Much More Output Per Worker than Others?", *Quarterly Journal of Economics* 114, 83-116.

[16] Hallerberg, M. and J. Von Hagen (1999), "Electoral Institutions, Cabinet Negotiations, and Budge Deficits in the European Union", in J. Poterba and J. Von Hagen (eds.), *Fiscal Institutions and Fiscal Performance,* University of Chicago Press.

[17] Heckman, J., H. Ichimura, and P. Todd (1997), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program", *Review of Economic Studies* 64, 605-654.

[18] Heckman, J. R. Lalonde, and J. Smith (1999), "The Economics and Econometrics of Active Labor Market Programs", in Ashenfelter, O. and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3c, North-Holland.

[19] Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge University Press.

[20] Huber, E., C. Ragin, and J. Stephens (1993), "Social Democracy, Christian Democracy, Constitutional Structure and the Welfare State", *American Journal of Sociology* 99, 711-749.

[21] Ichino, A. (2001), *Lecture Notes*, European University Institute, mimeo.

[22] Kontopopous, Y, and R. Perotti (1999), "Government Fragmentation and Fiscal Policy Outcomes: Evidence from the OECD Countries", in Poterba J. and J. von Hagen (eds.), *Fiscal Institutions and Fiscal Perfomance*, University of Chicago Press.

[23] Lijphart, A. (1990), "The Political Consequences of Electoral Laws, 1945-1985", *American Political Science Review* 84, 481-496.

[24] Lizzeri, A. and N. Persico (2001), "The Provision of Public Goods under Alternative Electoral Incentives", *American Economic Review* 91, 225-245.

[25] Maddala, G. (1977), *Econometrics*, McGraw Hill.

[26] Milesi-Ferretti G-M., Perotti, R. and M. Rostagno (2002), "Electoral Systems and the Composition of Public Spending", *Quarterly Journal of Economics* 117, 609-657.

[27] Persson, T., Roland, G., and G. Tabellini (2000), "Comparative Politics and Public Finance", *Journal of Political Economy* 108, 1121-1141.

[28] Persson, T. and G. Tabellini (1999), "The Size and Scope of Government: Comparative Politics with Rational Politicians, 1998 Alfred Marshall Lecture", *European Economic Review* 43, 699-735.

[29] Persson, T. and G. Tabellini (2000), *Political Economics: Explaining Economic Policy*, MIT Press.

[30] Persson, T. and G. Tabellini (2003), *Economic Policy in Represenative Democracies*, MIT Press, forthcoming.

[31] Rae, D. (1967), *The Political Consequences of Electoral Law*, Yale University Press.

[32] Rodrik, D. (1998), "Why Do More Open Economies Have Bigger Governments?", *Journal of Political Economy* 106, 997-1032.

[33] Rosenbaum, P.and D. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika* 70, 41-55.

[34] Scartascini, C. and M. Crain (2001), "The Size and Composition of Government Spending in Multi-Party Systems", mimeo, George Mason University.

[35] Shugart, M. and J. Carey (1992), *Presidents and Assemblies: Constutional Design and Electoral Dynamics*, Cambridge University Press.

[36] Strom, K. (1990), *Minority Government and Majority Rule*, Cambridge University Press.

[37] Taagepera, R. and M. Shugart (1989), *Seats and Votes: The Effects and Determinants of Electoral Systems*, Yale University Press.

[38] Tanzi, V. and L. Schuknecht (2000), *Public Spending in the 20th Century*, Cambridge University Press.

[39] Tsebelis, G. (2002), *Veto Players: How Political Institutions Work*, Princton University Press, forthcoming.

[40] Wiggins, V. "Two-Stage Least Squares Regression", http://www.stata.com/support/tags/stat/irreg.html

Table 1
Constitutions, policy outcomes and covariates
Cross section of 85 countries 1990-98

	(1)	(2)	(3)	(4)	(5)	(6)
	maj=1	*maj*=0	*p*(1,2)	*pres*=1	*pres*=0	*p*(3,4)
cgexp	25.9	31.3	0.022	22.2	33.3	0.000
	(8.2)	(11.0)		(7.2)	(10.0)	
ssw	4.8	10.5	0.000	4.8	9.9	0.002
	(5.3)	(6.4)		(4.6)	(7.0)	
lyp	8.1	8.6	0.058	7.9	8.7	0.000
	(1.1)	(0.8)		(0.9)	(0.9)	
trade	85.2	78.8	0.580	62.5	89.1	0.011
	(59.6)	(40.2)		(27.5)	(54.2)	
prop65	6.6	9.9	0.003	5.6	10.3	0.000
	(4.4)	(4.8)		(3.5)	(4.8)	
age	0.2	0.2	0.926	0.2	0.3	0.056
	(0.2)	(0.2)		(0.2)	(0.2)	
gastil	2.8	2.1	0.027	3.1	2.0	0.000
	(1.4)	(1.1)		(1.2)	(1.1)	

p(x,y) is the probability of falsely rejecting equal means across groups corresponding to columns x and y, under the assumption of equal variances
Standard errors in brackets

Table 2
Size of government and constitutions
OLS estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. var.	*cgexp*	*cgexp*	*cgexp*	*cgrev*	*cgexp*	*cgexp*	*cgexp*
pres	-6.04	-5.22		-5.07	-8.23	-3.36	-7.87
	(1.97)***	(1.95)***		(2.46)**	(2.81)***	(3.85)	(2.69)***
maj	-3.27	-4.99		-2.42	-4.42	-2.34	-3.79
	(1.73)*	(1.85)***		(1.75)	(2.36)*	(3.04)	(2.60)
propres			-6.47				
			(2.74)**				
majpar			-6.15				
			(2.88)**				
majpres			-9.79				
			(2.69)***				
Continent	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Colonies	No	Yes	Yes	Yes	Yes	Yes	Yes
Sample	90s, default	90s, default	90s, default	90s, default	90s, *gastil*<3.5	60-90s, default	90s,obs as(6)
Obs.	80	80	80	76	62	60	60
Adj.R2	0.58	0.64	0.64	0.59	0.63	0.56	0.65

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
All regressions include standard controls: *lyp*, *gastil*, *age trade*, *prop65*, *prop1564*, *federal*, *oecd*

Table 3
Size of government and constitutions
Instrumental variables, Heckman and Matching Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. var.	*cgexp*	*cgexp*	*cgexp*	*cgexp*	*cgexp*	*cgexp*	*cgexp*
pres	-8.70	-5.31	-8.71	-4.61	-5.89	-3.23	-7.45
	(3.91)**	(2.14)**	(3.59)**	(3.86)	(3.02)*	(2.74)	(2.34)***
maj	-5.06	-4.92	-3.92	-5.14	-4.81	-5.34	-5.59
	(1.91)***	(2.57)*	(3.42)	(3.58)	(3.41)	(2.73)*	(2.61)**
Conts & Cols	Yes	Yes	No	*col_uka, laam*			
Sample	90s, default	90s, default	90s, default	90s, default	90s, default	90s, default	90s, default
Endogenous selection	*Pres*	*maj*	*pres maj*	*pres maj*	*pres maj*	*pres maj*	*pres maj*
Estimation	Heckman 2-step	Heckman 2-step	2SLS	2SLS	Strat	Nearest	Kernel
rho	0.43	0.05					
Chi-2 over-id			4.59	3.64			
Adj. R2			0.59	0.60			
Obs.	75	75	75	75	66(pres) 70(maj)	66(pres) 70(maj)	66(pres) 70(maj)

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Critical value of chi-2(4,0.05) = 9.49
Always included in second-stage specification (cols 1-4): *age, lyp, trade, prop1564, prop65, gastil, federal, oecd*
First-stage specification of Heckman (cols 1-2) includes: *con2150, con5180, con81, age, engfrac, eurfrac, lat01, laam*
First-stage specification of 2SLS (cols 3-4) includes: *con2150, con5180, con81, age, engfrac, eurfrac, lat01*
Propensity-score logit estimation (cols 5-6) includes: *lyp, prop65, gastil, federal, col_uka, laam.*

Table 4
Composition of government and constitutions
OLS estimates

	(1)	(2)	(3)	(4)	(5)
Dep. var.	*ssw*	*ssw*	*ssw*	*ssw*	*ssw*
pres	-1.70		-4.35	-2.30	-1.41
	(1.30)		(1.82)**	(1.50)	(1.97)
maj	-1.64		-1.94	-0.97	-1.18
	(1.21)		(1.70)	(1.35)	(1.29)
propres		-2.26			
		(1.64)			
majpar		-2.14			
		(1.63)			
majpres		-3.08			
		(2.30)			
Continents & Colonies	Yes	Yes	Yes	Yes	Yes
Sample	90s, default	90s, default	90s, narrow	90s, part of 60-panel	60-90s, broad
Obs.	69	69	56	54	59
Adj. R2	0.75	0.75	0.74	0.78	0.80

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
All regressions include standard controls:
lyp, prop65, prop1564, gastil, age, trade, federal, oecd

Table 5
Composition of government and constitutions
Instrumental variables, Heckman and Matching Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. var.	*ssw*	*ssw*	*ssw*	*ssw*	*ssw*	*ssw*	*ssw*
pres	-1.19	-1.34	0.28	-0.39	-3.06	-2.28	-3.79
	(2.09)	(1.35)	(1.93)	(2.30)	(2.67)	(1.79)	(2.36)
maj	-1.42	-2.92	-3.58	-4.03	-1.85	-1.90	-3.46
	(1.13)	(1.66)*	(1.79)*	(2.08)*	(1.91)	(1.67)	(1.84)*
Continents & Colonies	Yes	Yes	No	col_uka laam			
Sample	90s, default	90s, default	90s, default	90s, default	90s, default	90s, default	90s, default
Endogenous Selection	*pres*	*maj*	*pres* *maj*	*pres* *maj*	*pres* *maj*	*pres* *maj*	*pres* *maj*
Estimation	Heckman 2-step	Heckman 2-step	2SLS	2SLS	Strat	Nearest	Kernel
Rho	-0.12	0.50					
Chi-2,over-id			5.69	9.60**			
Adj. R2			0.78	0.78			
Obs.	64	64	64	64	64(*pres*) 70(*maj*)	64(*pres*) 70(*maj*)	64(*pres*) 70(*maj*)

Standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Critical value of chi2(4,0.05) = 9.49
Always included in second-stage regression (columns 1-4): *age*, *lyp*, *trade*, *prop1564*, *prop65*, *gastil*, *federal*, *oecd*
2SLS first-stage specification(columns 3-4)includes: *con20*, *con2150*, *con5180*, *age*, *eurfrac*, *engfrac*, *lat01*
Heckman first-step probit specification(columns 1-2)includes: *con20*, *con2150*, *con5180*, *lat01*, *engfrac*, *eurfrac*, *age*, *laam* (*con2150* omitted from probit for *maj* to avoid perfect predictions)
Propensity score logit estimation includes: *lyp*, *prop65*, *gastil*, *federal*, *col_uka*, *laam*

Table 6
Size and composition of government spending
Adjustment to unobserved common events

	(1)	(2)	(3)	(4)
Dep. var.	*cgexp*	*cgexp*	*ssw*	*ssw*
ldepvar		0.85		0.82
		(0.02)***		(0.02)***
*pres*ldepvar*		-0.19		0.03
		(0.03)***		(0.04)
*maj*ldepvar*		-0.04		-0.03
		(0.02)*		(0.03)
pres	-0.58	-0.31	-0.65	-0.42
	(0.04)***	(0.17)*	(0.07)***	(0.17)**
maj	-0.38	0.04	-0.35	-0.43
	(0.04)***	(0.18)	(0.06)***	(0.14)**
Estimation	NL, FE	NL, FE	NL ,FE	NL, FE
N. obs.	1623	1576	1002	940
Adj. R^2	0.87	0.95	0.96	0.99

Controls always included: *trade, lyp, prop65, prop1564*, country fixed effects
Standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%
ldepvar is *lcgexp* in columns 1-2, *lssw* in 3-4

Table 7
Reaction of spending to income shocks

	(1)	(2)	(3)	(4)
Dep. var.	*cgexp*	*cgexp*	*ssw*	*ssw*
ldepvar	0.84	0.84	0.79	0.80
	(0.02)***	(0.02)***	(0.02)***	(0.02)***
*pres*ldepvar*	-0.21	-0.21	0.03	0.03
	(0.03)***	(0.03)***	(0.04)	(0.04)
*maj*ldepvar*	-0.06	-0.06	-0.03	-0.03
	(0.02)**	(0.02)**	(0.03)	(0.03)
yshock	-0.19		-0.13	
	(0.06)***		(0.02)***	
*pres*yshock*	0.19		0.07	
	(0.08)**		(0.03)**	
*maj*yshock*	0.11		0.07	
	(0.08)		(0.03)**	
posys		-0.08		-0.05
		(0.11)		(0.04)
*pres*posys*		0.02		-0.02
		(0.16)		(0.06)
*maj*posys*		0.07		0.02
		(0.15)		(0.06)
negys		-0.30		-0.20
		(0.11)***		(0.04)***
*pres*negys*		0.35		0.15
		(0.16)**		(0.06)**
*maj*negys*		0.15		0.12
		(0.15)		(0.06)**
Sample	|yshock| < 5	|yshock| < 5	|yshock| < 5	|yshock| < 5
Estimate	ctry FE	ctry FE	ctry FE	ctry FE
Obs.	1474	1474	888	888
N. ctries	60	60	56	56
Adj. R2	0.82	0.82	0.75	0.75

Standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%
Controls: *lyp*, *trade*, *prop1564*, *prop65*, *oil* (interacted with a dummy variable for oil exporter or oil importer)

Figure 1
Unobserved common events and the size of government