Constitutional Rules and Fiscal Policy Outcomes

By TORSTEN PERSSON AND GUIDO TABELLINI*

We investigate the effect of electoral rules and forms of government on fiscal policy outcomes in a large sample of democracies. We rely on different estimation methods to address prospective problems of statistical inference, due to nonrandom selection of these constitutional rules. The findings are consistent with our theoretical priors: presidential regimes induce smaller governments than parliamentary democracies, while majoritarian elections lead to smaller governments and smaller welfare programs than proportional elections. (*JEL* D72, E60, H00)

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, Persson and Tabellini (1999, 2000), Alessandro Lizzeri and Nicola Persico (2001), and Gian Maria Milesi-Ferretti et al. (2002) 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 toward programs 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 vote shares are converted to seat shares in the legislature). 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 just about 25 percent of the national vote: 50 percent in 50 percent of the districts. Under full proportional representation (PR) it needs 50 percent of the national vote; politicians are thus again induced to internalize the policy benefits for larger segments of the population, which leads them to emphasize broad programs.[1]

* Persson: Institute for International Economic Studies, Stockholm University, S-106 91 Stockholm, Sweden, London School of Economics, CEPR, and NBER (e-mail: torsten.persson@iies.su.se); Tabellini: IGIER, Università Bocconi, Via Salasco 3/5, Milan, Italy, CEPR, and Ces-Ifo (e-mail: guido.tabellini@uni-bocconi.it). A longer version of the paper (Persson and Tabellini, 2001) has been circulating under the title: "Political Institutions and Policy Outcomes: What Are the Stylized Facts?" 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 many 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, Andrea Mascotto, José Mauricio Prado Jr., Alessandro Riboni, Davide Sala, and Francesco Trebbi for research assistance at various stages of the project. The research has been supported by grants from the European Commission, Bocconi University, MURST, and the Swedish Research Council.

[1] Perotti et al. (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).

These theoretical papers take the number of parties as given (and often equal to two) and not endogenously dependent on the electoral system. But, as emphasized by political scientists (e.g., Douglas Rae, 1967; Rein Taagepera and Matthew Shugart, 1989; Arend Lijphart, 1990), majoritarian elections are strongly associated with fewer parties. As a result, majoritarian parliamentary systems are more likely to produce single-party majority governments, whereas coalition and minority governments become more likely under proportional elections (Taagepera and Shugart, 1989; Kaare 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, David 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. Yianos Kontopolous 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 et al. (2002) also predict that proportional rule leads to larger overall spending.

How the form of government might influence fiscal policy has not been as extensively studied. In one of few formal analyses, Persson et al. (2000) distinguish between presidential and parliamentary regimes on the basis of whether the executive is accountable to the legislature through a confidence requirement. They build on Daniel Diermeier and Timothy Feddersen (1998), that show how the confidence requirements induce more "legislative cohesion": a stable majority of legislators supporting the executive in place also votes together on legislation, pursuing the joint interest of its voters. For this reason, spending in parliamentary regimes provides benefits to a majority of voters, as in the case of broad social security and welfare programs. Moreover, the stable majority of legislators and their voters becomes a residual claimant on additional revenue and therefore prefers 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. None of these minorities is a residual claimant on revenue and therefore a majority resists high spending, exploiting stronger checks and balances and the greater dispersion of veto rights in presidential regimes than in parliamentary regimes. These forces lead to the prediction that presidential regimes are associated with smaller governments and smaller social transfer programs than parliamentary regimes.

An interesting empirical literature examines how constitutional features in state and local governments correlate with fiscal policy outcomes, particularly in the United States (see the excellent survey by Timothy Besley and Anne Case, 2003). But when it comes to electoral rules and forms of government, the most interesting institutional variation must be sought across different countries. Little research has exploited that variation. A few political scientists, like Evelyne Huber et al. (1993) and Francis 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. Among economists, Milesi-Ferretti et al. (2002) show that proportional elections seem to induce larger governments and larger transfer payments in the OECD data, while Persson and Tabellini (1999) find strong support for the prediction that presidential regimes have lower spending in a cross section of 50 democracies in the early 1990's, but less robust effects of the electoral rule.

In this paper, we rely on new and more extensive data to estimate the effect of electoral rules and forms of government on the size and composition of government spending. We mainly use a cross section of 80 democracies in the 1990's, but also report a few results from an unbalanced panel of 60 democracies for the years 1960–1998. Section I describes these data in more detail.

Fundamental constitutional reforms are very rare. Hence, our inference about the effect of constitutions on policy outcomes must be identified from the cross-country variation in constitutions. This raises a number of statistical issues. A main challenge is that constitution selection is not random: 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 nonobservable 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. The empirical strategy is described in Section II.

Our empirical results on the size (Section III) and composition (Section IV) of government strongly indicate that the political constitution has a causal effect on fiscal policy. One central finding is that the electoral rule exerts a strong influence, in line with the priors from the theory. According to the cross-country evidence, a switch from proportional to majoritarian elections in a country chosen at random reduces total government spending by about 5 percent of GDP and welfare spending by about 2 percent of GDP. These effects are particularly pronounced in better and older democracies. Results from panel-data analysis suggest that a large part of the differences observed in the 1990's derive from different growth of measured government spending across different electoral systems in the last 30 years.

The data also support the predictions concerning the form of government. The estimates indicate that presidentialism reduces the overall size of government by as much as majoritarian elections, roughly 5 percent of GDP. Again, much of the effect can be traced back to a more rapid growth of government spending in parliamentary regimes than in presidential regimes during the 1970's and 1980's. Parliamentary democracies also have larger welfare spending than do presidential democracies, in line with our theoretical prior. But here the estimated constitutional effects are statistically robust only among the older and better democracies.

I. Data

This section discusses the key variables in our empirical analysis. The data have been collected as part of a larger research program on economic policy and comparative politics. Persson and Tabellini (2003) provide a more comprehensive discussion. Succinct descriptions of the sources, as well as the data themselves, are available at the AER Web site (http://www.aeaweb.org/aer/contents/).

Sample Selection.—Our goal is to compare policy outcomes in democracies ruled by different constitutions. We have collected data for 80 democracies, averaging yearly outcomes over the 1990–1998 period. We also report some results in a subset of 60 democracies, where data are available for a longer period.

How do we define a democracy? In the 1990's 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. We include a country in the sample if the average of these two indexes in the 1990–1998 period does not exceed 5. This generous definition of democracy includes countries such as Zimbabwe and Belarus (note that 5 refers to the average score; both countries' scores have deteriorated considerably after 1998). While generosity raises sample size, we also report results based on a split between good and bad democracies at an average score of 3.5 in 1990–1998. The countries in our sample also differ greatly 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 analysis, we always control for the quality (measured by *gastil*) and age (measured by *age*) of democracy.[2]

Constitutional Rules.—We classify electoral rules and regime types by means of two indicator

[2] For the 1960–1998 panel, we mainly rely on the Polity IV data set, which goes farther back and is more comparable over time than the Freedom House data. Specifically, we use the encompassing *polity* index, which assigns to each country and year an integer score ranging from -10 to $+10$ (higher values associated with better democracies), restricting the panel to countries and years with positive values of *polity* (censored observations are treated as randomly missing).

(dummy) variables: *maj* and *pres*. Majoritarian countries (*maj* = 1) are those relying exclusively on plurality rule in the most recent election to the legislature (lower house). Mixed and PR electoral systems are lumped together as proportional (*maj* = 0). Due to the correlated features of electoral systems noted in the introduction, using district magnitudes rather than electoral formulas would produce a similar but not identical classification.

For the form of government, we follow the theory discussed in the introduction and classify as presidential (*pres* = 1) countries where the chief executive/cabinet (in the sphere of economic 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, which has no popularly elected president but a permanent-coalition executive that cannot be brought down by the legislative assembly.

These classifications change very few times in the last 40 years (*pres* does not vary at all from 1960 onwards, whereas *maj* changes only in a few countries, Cyprus, Fiji, France, Japan, New Zealand, the Philippines, and Ukraine and mainly during the 1990's). In the 1990's cross section, we treat the variable *maj* as binary (0 or 1) and if there was a reform in this decade we code its value *before* the reform, as it should take some time before electoral reform affects such slowly moving variables as the size of government or welfare spending. Table A1 in the Appendix lists the countries in the cross section and their classification according to *pres* and *maj*.

The constitutional inertia, sometimes called an "iron law" by political scientists, suggests that we can use history to explain cross-country variation in constitutional rules. We thus construct three indicator variables dating the origin of the current constitution to the periods before 1920, 1921–1950, and 1951–1980 (with the period after 1981 as the default). These indicators (called *con*20, *con*2150, *con*5180 respectively) take a value of 1 if the current constitution (either the regime or the electoral rule) dates back to the period, and 0 otherwise.[3] About one-third of our entire sample has majoritarian elections, but this proportion is much lower (one-seventh) for those with constitutions originating in the 1921–1950 period, and higher (one-half) if originating in 1951–1980. The frequency of forms of government instead varies monotonically over history, presidential regimes being relatively more often associated with younger constitutions (or newer democracies). Even though we do not have a universal explanation for these specific patterns, the forces shaping constitutional rules—such as experience by other democracies, prevalent political and judicial doctrines, and academic thinking—have evidently shifted systematically over time.

To explain the variation in constitutional rules, we also rely on other cultural and geographic variables, namely distance from the equator (*lat*01), the percentage of the population with English (*engfrac*) or a European language (*eurfrac*) as the mother tongue, ethno-linguistic fractionalization (*avelf*), and population size (*lpop*). The correlation between these variables and the constitution varies with the specification and the estimation method (perhaps reflecting collinearity among some of the regressors). But countries with a larger fraction of English speakers are more likely to have majoritarian elections and a parliamentary form of government, plausibly reflecting the British influence, while 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 lower probability of adopting the dominant form of government in Europe.[4]

[3] The origin of the current constitution is defined as the year when the current electoral rule or the current form of government was first established, given that the country was a democracy and an independent nation. Absent reforms since becoming a democracy, the birth dates of the constitution and democracy coincide. 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.

[4] This corresponds to the idea in Robert Hall and Chad Jones (1999) and Daron Acemoglu et al. (2001), who argue—and exploit empirically—that countries close to the

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 (including social security), expressed as a percentage of GDP (called *cgexp*). But we also use central government revenues (*cgrev*) and—as a diagnostic—the government deficit (*dft*), both as percentages of GDP. For the composition of government spending, we measure social security and welfare spending (by central government) as a percentage of GDP (*ssw*). The presumption is that the broad transfer programs included in this measure, like pensions and unemployment insurance, are much harder to target towards narrow geographic constituencies compared to other government outlays. For the size of government, we rely mainly on IFS data, while the welfare spending measure is extracted from the GFS database (both from the IMF).

These policy measures vary a great deal, both across countries and time. In the 1990's cross section the mean value of expenditures is 29.8 percent of GDP with a standard deviation of 10.4 percent, a minimum of 9.7 percent (in Guatemala) and a maximum of 51.2 percent (in the Netherlands). In the 1960–1998 panel, the distribution drifts upwards over time, reflecting growth in the average size of government of about 8 percent of GDP from the 1960's to the mid-1990's. Most of this growth takes place in the 1970's and 1980's. Our measure of welfare spending also shows a wide distribution at any point in time, with an uninterrupted upward drift.

A natural concern is whether our measurement of central (rather than general) government biases 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 available for only about 40 countries. In the 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 the analysis.

Other Covariates.—We obviously want to hold constant a number of variables likely to shape government outlays. Specifically, we always include in our regressions measures of the level of development, log of real per capita income (*lyp*), openness (*trade*), exports plus imports over GDP, population size (*lpop*) and two variables measuring demographic composition, the percentages between 15 and 64 years of age (*prop*1564), and above 65 years of age (*prop*65), respectively. These variables have been shown to correlate with measures of fiscal policy in previous studies, such as David Cameron (1978), Dani Rodrik (1998), and Persson and Tabellini (1999).

To control for nonobservable influences on fiscal policy related to geographic location or economic development, we also rely on indicator (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 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. Since the influence of colonial heritage is likely to fade with time, we weigh these (0,1) indicators by the time since independence, giving more weight to colonial history in young independent states and no weight at all to colonial rule more than 250 years ago. The resulting colonial history variables are called *col_uka, col_espa,* and *col_otha.*

A Preliminary Look at the Data.—Table 1 displays the means of several variables in our 1990's cross section, broken down by constitutional groups. Clearly, both overall government size and welfare-state spending are much smaller in presidential than parliamentary countries; they are also smaller in proportional than majoritarian countries. It is tempting to interpret these patterns in the data as support for 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

equator have less growth-friendly institutions, due to their harsher conditions for Western European settlers.

TABLE 1—CONSTITUTIONS, POLICY OUTCOMES, AND COVARIATES: CROSS-SECTIONAL DATA FOR 85 COUNTRIES, 1990–1998

	(1)	(2)	(3)	(4)	(5)	(6)
	$maj = 1$	$maj = 0$	$p(1,2)$	$pres = 1$	$pres = 0$	$p(4,5)$
cgexp	25.6	30.8	0.03	22.2	33.3	0.00
	(8.2)	(11.3)		(7.2)	(10.0)	
ssw	4.7	10.1	0.00	4.8	9.9	0.00
	(5.4)	(6.6)		(4.6)	(7.0)	
lyp	8.1	8.6	0.04	7.9	8.7	0.00
	(1.2)	(0.8)		(0.9)	(0.9)	
trade	83.7	75.6	0.44	62.5	89.1	0.01
	(59.9)	(37.5)		(27.5)	(54.2)	
*prop*65	6.7	9.6	0.01	5.6	10.3	0.00
	(4.4)	(4.9)		(3.5)	(4.8)	
age	0.22	0.20	0.77	0.16	0.24	0.09
	(0.25)	(0.20)		(0.23)	(0.21)	
gastil	2.7	2.3	0.08	3.1	2.0	0.00
	(1.4)	(1.1)		(1.2)	(1.1)	

Notes: Mean values by constitutional rules; standard deviations are in parentheses. $p(x, y)$ is the probability of falsely rejecting equal means across groups corresponding to columns x and y, under the assumption of equal variances.

random across the two constitutional classifications. Majoritarian and presidential countries tend to be less economically advanced and have worse democratic institutions and younger populations than their proportional and parliamentary counterparts. Further, presidential regimes are found in more closed economies and younger democracies than parliamentary regimes. The geographic distribution also appears nonrandom: presidential regimes are largely located in the Americas, while continental Europe is predominantly proportional and parliamentary. Colonial history is also correlated with current constitutions: former British colonies tend to be parliamentary and majoritarian, while former Spanish colonies tend to be presidential. These differences among constitutional groups might fully explain the observed differences in fiscal policy, with no causal effect left for the constitution. Moreover, these constitutional groups almost certainly differ also in other dimensions, which we cannot identify or observe empirically. Causal inference about the effect of constitutions on policy outcomes requires precise identifying assumptions and statistical methods beyond cross tabulation. We turn to these in the next section.

II. Empirical Strategy

Basic Ordinary Least-Squares (*OLS*).—Our empirical model can be thought of as having two parts. One is a stochastic process for the constitution. To simplify the exposition, suppose there is just one constitutional dimension that can take two possible values in country i, $S_i = 0, 1$. Then we can write the process for constitution selection as:

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

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

$$Y_i = F(S_i, \mathbf{Z}_i) + u_i. \tag{2}$$

Our goal is to estimate the average effect of constitutional reform—a shift from $S = 0$ to $S = 1$—on fiscal policy in our sample: the so-called average treatment effect.

OLS imposes two commonplace assumptions: (i) recursivity: the error term e in the constitution-selection equation (1) is uncorrelated with the error term u in the policy-outcome equation (2), an assumption also known as "conditional independence," or "selection on observables;" (ii) linearity: the function F in (2) is linear with constant coefficients, so the only effect of the constitution is on the intercept of F. By (ii), the constitutional effect on policy is fully captured by the coefficient of the constitutional indicator, S, and by (i), this coefficient can be consistently estimated by OLS. 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 (*lpop, prop*1654 and *prop*65), the age and quality of democracy (*age* and *gastil*), and the indicators for federal and OECD countries (*federal* and *oecd*).

Relaxing Conditional Independence.—Given the nonrandom distribution of the constitution, 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 appear in the regression (and the model is linear). For this reason, when estimating by OLS we always add to the regressors also the indicators for continental location and colonial history introduced in Section I.

But how do we know that we have included enough common determinants of policies and constitutions to satisfy the conditional independence assumption? If some omitted determinant of policy outcomes is correlated with the constitution, conditional independence is violated and the OLS estimates 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 (1) and (2). Of the possible sources of simultaneity bias, we believe that such prospective "omitted variables" are much more important than "reverse causality." A direct feedback from policies to constitutions is hard to reconcile with two features of the data discussed in Section I: considerable policy changes but very few constitutional reforms during the last 40 years.

To relax conditional independence, we use the so-called Heckman correction and instrumental variables. Both methods entail explicit estimation of the constitution-selection equation (the first stage) and the policy-outcome equation (the second stage). In the Heckman correction, we estimate the first-stage (1) with a probit model. This gives an estimate of the correlation coefficient between the error terms e and u of (1) and (2), which is used to correct the OLS estimates. Identification relies on an exclusion restriction (discussed below) plus a strong functional form assumption: (1) and (2) are linear, and the error terms u and e jointly normal.

With instrumental variables, we estimate the constitution selection equation with the linear probability model (perhaps more robust than probit when estimating by instrumental variables, see Joshua D. Angrist and Alan B. Krueger, 2001). The identification assumption is an exclusion restriction. We exploit the historical variables correlated with constitutions and assume them to be uncorrelated with the unobservable determinants of policies, u. Throughout, the policy outcome equation has the same baseline specification as in the OLS estimation, with or without dummy variables for continental location, as noted below. The detailed specification for constitution selection is discussed in context.

Relaxing Linearity.—OLS, instrumental variables, and the Heckman procedure, all exploit the assumption that F in (2) is linear with constant coefficients. Usually, linearity is taken as a convenient local approximation of a more general model. But here we are interested in comparing very different groups of countries. As shown in Table 1, most variables differ considerably across constitutional groups. Suppose, as is plausible, that the constitutional effect on policy outcomes is stronger in older or better democracies. As these features differ systematically across constitutional groups, the local approximation may no longer be tenable and the linear estimates become biased.

How do we address this prospective problem? One way is to directly amend a linear specification (OLS or otherwise) with various interaction terms. We use this direct method for a few variables of immediate interest. As the number of possible interactions and other nonlinearities is close to infinite, however, we also address the problem in a more general way. Specifically, we relax linearity and estimate the effect of the constitution on fiscal policy with nonparametric matching methods, based on the propensity score. In doing so, however, we once again have to rely on the conditional-independence assumption.[5]

The gist of these nonparametric estimators is that they give more weight to comparisons of similar countries, to reduce the effect of any nonlinearities. Countries are ranked on the basis of their "propensity score." In our context, the propensity score is defined, following (1), as the conditional probability that country i is in constitutional state $S_i = 1$, given a vector of observable determinants $\mathbf{X}$. Some countries in this ranking actually have $S = 1$, others do not. The main idea is that the actual assignment of constitutions to countries with similar propensity scores is largely random, which makes it appropriate to compare policy outcomes across different constitutions. The matching underlying this comparison of similar countries can be performed in alternative ways, and each way corresponds to a specific matching estimator. We use three such estimators. With the *stratification* estimator, countries are grouped into different strata on the basis of similarity of their estimated propensity scores. Within 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 those two countries that are closest in terms of their estimated propensity scores. Finally, the *kernel* estimator combines the logic of the previous two estimators. Each $S = 1$ country is matched against a weighted average of all $S = 0$ countries within a certain propensity-score distance, with weights declining in that distance, and the same procedure is followed when matching the $S = 0$ countries.

Multiple Constitutional States.—In the following sections, we measure constitutional rules by the two binary indicators, *pres* and *maj*. In the OLS estimation we include both indicators to our regressions. The estimated coefficient of, say, the presidential indicator *pres* then measures the expected effect of switching from a proportional-parliamentary to a proportional-presidential system, under the constraint that the effect coincides with that of switching from a majoritarian-parliamentary to a majoritarian-presidential system. We test whether these additivity assumptions are fulfilled. When estimating by instrumental variables, we allow for 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 equation 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 neglecting the form of government. We have too few observations to reliably implement these procedures for multiple constitutional states. (How to handle multiple treatments is discussed in Jeffrey Wooldridge, 2002, for the Heckman procedure, and in Michael Lechner, 2002, for the propensity-score methods.)

III. 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 section we ask whether these predictions are consistent with the evidence, using the empirical methods discussed in Section II. The STATA programs generating the estimates discussed below are available at the AER Web site.

OLS Estimates.—We start by estimating equation (2) by OLS, under the assumption of conditional independence and linearity. Unless noted otherwise, the sample of countries is the 1990's cross section, and we hold constant the standard controls in $\mathbf{Z}$ plus our indicator variables for geographic location (Africa, Asia, and Latin America) and colonial origin (United Kingdom, Spain, and other). Among these covariates, countries more open to international trade, with older populations, and former British colonies, tend to have a larger size of government; worse democracies and countries located in Latin America tend to have a smaller

[5] These methods were introduced into economics as tools for evaluating labor market and education programs (see for instance James J. Heckman et al., 1997; Rajeev H. Dehejia and Sadek Wahba, 1999; and Heckman et al., 1999).

TABLE 2—SIZE OF GOVERNMENT AND CONSTITUTIONS: OLS ESTIMATES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	*cgexp*	*cgexp*	*cgrev*	*dft*	*cgexp*	*cgexp*	*cgexp*
pres	−5.18		−5.00	0.16	−2.65	−7.75	−6.46
	(1.93)***		(2.47)**	(1.15)	(2.70)	(2.70)***	(2.98)**
maj	−6.32		−3.68	−3.15	−1.45	−7.94	−6.33
	(2.11)***		(2.15)*	(0.87)***	(2.32)	(3.74)**	(2.48)**
propres		−6.56					
		(3.01)**					
majpar		−6.96					
		(3.72)*					
majpres		−10.37					
		(3.03)***					
pres_newdem						3.50	
						(2.72)	
maj_newdem						3.58	
						(4.03)	
newdem						−4.08	
						(2.23)*	
pres_baddem							2.42
							(4.16)
maj_baddem							2.06
							(5.97)
baddem							−5.73
							(3.46)
F-test (*pres*)		0.43				4.01**	1.40
F-test (*maj*)						3.18*	0.66
Sample	1990's	1990's	1990's	1990's	1960–1973	1990's	1990's
Observations	80	80	76	72	42	80	80
R^2	0.71	0.70	0.68	0.50	0.79	0.72	0.70

Notes: Robust standard errors are in parentheses. All regressions include our standard controls, *lyp, lpop, gastil, age, trade, prop*65, *prop*1564, *federal,* and *oecd,* plus a set of indicator variables for continental location and colonial origin, except that *age* is missing in column (5)–(6), while *gastil* is missing in column (7) and replaced by *polity* in column (5). *F*-test (*pres*) refers to tests of the hypotheses that the coefficient for *propres* is equal to the difference between the coefficients for *majpres* and *majpar* [column (2)], the sum of the coefficients for *pres* and *pres_newdem* is zero [column (6)], and the sum of the coefficients for *pres* and *pres_baddem* is zero [column (7)]. *F*-test (*maj*) refers to the corresponding tests with regard to *maj* [columns (6) and (7)].

* Significant at the 10-percent level.
** Significant at the 5-percent level.
*** Significant at the 1-percent level.

size of government. The other covariates do not generally have coefficients significantly different from zero, but since some of them are correlated with the constitution, they are still included to minimize the risk of violating conditional independence.

Column (1) of Table 2 indicates that effect on government size of presidential rather than parliamentary democracy is 5 percent of GDP, and 6 percent of GDP of majoritarian rather than proportional elections. These point estimates are not only highly statistically significant, but also economically and politically relevant. In column (2), we break down the constitutional variables into a finer partition (*majpres, majpar,* and *propres*). The effects of the two constitutional features indeed appear additive: an *F*-test does not reject the null that the estimated coefficient of *majpres* equals the sum of the estimated coefficients of *propres* and *majpar.* According to these estimates, 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 percent of GDP. Since the additive specification is not rejected by the data, we focus on this more parsimonious specification in what follows.

The estimated constitutional effects, particularly the one of presidential government, are remarkably robust to the specification of the control vector **Z**. Omitting some of the less influential controls, the continental dummy variables, or the colonial origin variables, adding other controls such as income inequality, ethnic and linguistic fractionalization, the size of the mining sector, or a dummy variable for former socialist countries, the estimated coefficient of presidentialism stays stable and statistically different from zero. The estimated coefficient of majoritarian elections is always negative and almost always statistically significant, although in some specifications it drops towards -3 percent of GDP. The additional controls generally have no additional explanatory power, with the exception of mining (for which data are available only for a smaller set of countries, however).[6]

In column (3), we measure the size of government by tax revenue instead of spending (*cgrev* rather than *cgexp*). The effect of presidential regimes is the same as before, while the effect of majoritarian elections is now cut in half. This is consistent with the results in column (4), which show presidentialism to have no effect on the size of the budget deficit (*dft*), whereas majoritarian elections are associated with a smaller deficit of about 3 percent of GDP. Thus, a switch from presidential to parliamentary government would increase both spending and taxation by the same amount, about 5 percent of GDP. A reform of the electoral rule from majoritarian to proportional, instead, would increase spending by about 6 percent of GDP, financed by higher taxes and deficits in similar proportions. Naturally, these cross-sectional estimates pick up the long-run effects of constitutional reforms under the assumption that fiscal policy is in the steady state (recall that deep constitutional reforms are very rare in our sample), and say nothing about transitional dynamics. The theories reviewed in the introduction are generally static, and have no predictions about the constitutional effects on budget deficit. One important exception is the idea that proportional electoral rule leads to more spending because of a common-pool problem in coalition governments. According to the theory, this mechanism would also lead to larger public debts and, temporarily, to larger budget deficits, because the common-pool problem also induces myopic fiscal policy (see Persson and Tabellini, 2000). Thus, the estimated effect of the electoral rule on the budget deficit indirectly supports the common-pool interpretation of why proportional elections might induce larger governments.[7]

The size of government has not remained constant over time. Were these constitutional differences of the 1990's already present in the late 1960's, before the big postwar expansion in the welfare state? To answer this question, column (5) runs the same cross-sectional regression on a smaller sample of 42 democracies for which data are available in a much earlier period, namely 1960–1973. For each country, observations are averaged over 1960–1973, though for several countries data are actually available only in the early 1970's.[8] The estimated coefficients of *pres* and *maj* are now much smaller in absolute value and no longer significantly different from zero. This is due to the earlier time period, not to sample selection: restricting the sample to the same 42 countries, but using data averaged over the 1990's, as in the rest of Table 2, the estimated coefficient of *pres* and *maj* are -10.34 and -4.52, both significantly different from zero. To gain a better understanding of these issues, a previous version of this paper (Persson and Tabellini, 2001) and Persson and Tabellini (2003) use an unbalanced panel of 60 countries over the period 1960–1998. Two significant differences emerge between proportional-parliamentary countries and the other constitutional groups. First, the

[6] The failure to find an association between income inequality and the size of government may appear surprising in light of the celebrated prediction of median voter models, but it is common to several other empirical studies (see the review by Persson and Tabellini, 2000).

[7] For this interpretation to make sense, however, we have to assume that the steady-state level of debt has not yet been reached, at least in the proportional countries.

[8] Thus, the average does not refer to the same period for all countries. The end point is chosen because 1973 is the year when IMF fiscal-policy data become available for several countries. Similar results are obtained if the data are averaged over a shorter subperiod, such as 1965–1973.

size of government spending has a steeper time trend in proportional and parliamentary countries in the early part of the sample.[9] Second, the constitution also affects the cyclical response of fiscal policy. In proportional and parliamentary countries government spending displays a "ratchet effect": it goes up as a fraction of GDP during cyclical downturns, but it does not come down during cyclical upturns. This asymmetry is not apparent in the other constitutional groups. We cannot offer a theoretical explanation of these interactions between the constitution and the dynamics of fiscal policy, but they are consistent with the cross-sectional results reported in Table 2.

This evidence suggests that, over the 1970's and 1980's, all countries adjusted their spending levels towards a new steady state with higher spending but the adjustment was particularly marked in proportional and parliamentary democracies. An additional explanation may be that in these democracies decisions on more generous entitlement programs—such as pensions and unemployment insurance—were made already in the 1960's and early 1970's. But the full effects on spending may not have been revealed until a couple of decades later, in pace with older populations and higher unemployment.

If constitutional effects have become stronger over time, could they also be affected by the age of democracy? All regressions include the age of democracy (*age*) as a control variable. But if there is an interaction between the age and form of democracy, this would not be enough. To allow for it, we create two new variables, *maj_newdem* and *pres_newdem*, defined as the product of *maj* and *pres* (respectively) and *newdem*, a dummy variable taking a value of 1 if the country became a democracy after 1959 and 0 otherwise (28 democracies qualify as old according to this definition). In this specification, the estimated coefficients of *pres* and *maj* give us the constitutional effect in the old (pre-1960) democracies, while the estimated coefficient of *maj_newdem* and *pres_newdem* give us the difference between new and old democracies. Column (6) of Table 2 reports the results (also replacing *age* with *newdem*). For both majoritarian elections and presidentialism, the constitutional effects do seem a bit stronger in the old democracies (the point estimates of *maj* and *pres* rise by about 2 points in absolute value). But the constitutional effects are not significantly different in new and old democracies, while both constitutional effects are significantly different from zero also in the new democracies [cf. the *F*-tests in column (6)].[10]

The 80 countries in our sample include some dubious democracies. In weak democracies, the formal constitution might play a relatively less important role than informal practice or social norm. Moreover, as seen in Section I, weaker democracies tend to be presidential, which might introduce systematic bias. For this reason, we also interact our constitutional indicators with a new dummy variable, *baddem*, taking a value of 1 in the democracies with a *gastil* score higher than 3.5—18 bad democracies appear in our sample. The results are displayed in column (7) of Table 2 (where we also replace *gastil* with *baddem*). The interpretation of the estimated coefficients is analogous to that in column (6), except that we focus on the quality rather than age of democracy.[11] The effect of presidential regimes appears to be stronger in the better democracies, whereas the effect of majoritarian elections remains stable. Neither constitutional effect appears to be significantly different in the good and bad democracies (the estimated coefficients of *pres_baddem* and *maj_baddem* are not significantly different from zero). Here, however, we cannot reject the hypothesis that both constitutional effects are zero among the 18 bad democracies (cf. the *F*-tests reported in Table 2).

In summary, imposing the assumptions of conditional independence and linearity (or

[9] Since the regressors include demographic variables, these different time trends cannot simply be interpreted as the effects of more rapidly aging populations. They reflect the reaction of government spending to unobserved events common to all countries, such as ideological shifts in the voters' preferences.

[10] These results are confirmed also in more parsimonious regressions estimated for the smaller sample of old democracies.

[11] Quality and age of democracy are indeed different criteria. Many of the new democracies qualify as good (i.e., have *gastil* scores less than 3.5) according to our definition, and the correlation coefficient between the variables *baddem* and *newdem* is only 0.25.

restricting nonlinearity to the interaction effects just discussed), the negative constitutional effect of presidential regimes and majoritarian elections are large and robust. Both effects conform to prior expectations from theory. They are much stronger in the 1990's, however, suggesting that the constitution influenced postwar growth in the size of government.

Relaxing Conditional Independence: Heckman and IV Estimates.—How robust are the previous results 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 methods is to specify the determinants of constitution selection. We impose a similar, albeit not identical, specification in the two cases (to rely on alternative identification assumptions). Consider first the Heckman procedure. As noted in Section I, the current constitution is well explained by cultural and historical variables, such as the cultural influence of the West (Great Britain in particular), by geographic location, and by the size and heterogeneity of population. Our first-stage specification therefore includes the fraction of the population whose mother tongue is English (*engfrac*) or a European language (*eurfrac*), distance from the equator (*lat*01), a dummy variable for Latin America (*laam*), population size (*lpop*), and a measure of ethno-linguistic fractionalization (*avelf*). These variables have considerable explanatory power and are generally statistically significant: the pseudo-R^2 of the probit equation for the Heckman procedure is 0.43 for the form of government, 0.47 for the electoral rule (results not shown).

The constitutional effects estimated with the Heckman procedure are reported in columns (1) and (2) of Table 3.[12] 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 (1)], 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 positive and considerably higher, namely 0.62. Thus, the OLS estimates are likely to be upward-biased, and the Heckman correction produces an even 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 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 believe the restriction is fulfilled for the three historical indicators for the origin of the currently observed constitutional rule (form of government or electoral rule), *con*20, *con*2150, *con*5180—recall that the age of democracy (*age*) is always included among the second-stage regressors. Thus, we assume that, holding constant the age of democracy and other second-stage variables, fiscal policy is not correlated with the historical period when the current constitution was established. These three instruments are only weakly correlated with the constitution, however. To increase the explanatory power of the first-stage regression, we also include some of the constitutional determinants already exploited in the Heckman procedure, as well as the continuously measured age of democracy, but only if they appear to be statistically significant. This results in a slightly different specification for *maj* and *pres*, reported in Table A2 of the Appendix. Note that the constitutional dating variables are jointly significant in both regressions, and in one case also individually significant.[13] Since

[12] As noted above, we apply the Heckman correction to one constitutional dimension at a time, treating the other dimension as random. Estimation is by maximum likelihood.

[13] This parsimonious first-stage specification is chosen to avoid excessively weak instruments. Thus, we estimate the first stage by OLS as reported in Table A2 of the Appendix, run the second stage on the predicted values of *maj* and *pres*, and correct the second-stage residuals as discussed by G. S. Maddala (1977, Ch. 11). The point estimates are very

TABLE 3—SIZE OF GOVERNMENT AND CONSTITUTIONS: INSTRUMENTAL VARIABLES, HECKMAN, AND MATCHING ESTIMATES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	*cgexp*	*cgexp*	*cgexp*	*cgexp*	*cgexp*	*cgexp*	*cgexp*
pres	−5.29	−11.52	−6.51	−4.22	−5.86	−2.54	−7.30
	(2.18)**	(4.54)**	(3.71)*	(3.99)	(4.53)	(2.26)	(2.36)***
maj	−6.21	−6.77	−4.83	−4.18	−4.86	−6.59	−5.76
	(2.82)**	(1.98)***	(3.19)	(3.17)	(3.57)	(3.40)*	(2.59)**
Conts & Cols	Yes	Yes	*col_uka*	*col_uka, laam*			
Sample	1990's	1990's	1990's	1990's	1990's	1990's	1990's
Endogenous selection	*maj*	*pres*	*pres* *maj*	*pres* *maj*	*pres* *maj*	*pres* *maj*	*pres* *maj*
Estimation	Heckman ML	Heckman ML	2SLS	2SLS	Stratification	Nearest neighbor	Kernel
Rho	0.05	0.62					
	(0.29)	(0.33)					
Chi-2			3.29	2.23			
Adjusted R^2			0.59	0.59			
Observations	75	75	75	75	65(*pres*) 67(*maj*)	65(*pres*) 67(*maj*)	65(*pres*) 67(*maj*)

Notes: Standard errors are in parentheses. Always included in second-stage specification in columns (1)–(4): *age, lyp, trade, prop*1564, *prop*65, *gastil, federal, oecd, lpop*; Conts & Cols refer to indicator variables for continental location and colonial history. Specification of constitution selection in Heckman procedure in columns (1)–(2) includes: *engfrac, eurfrac, lat*01, *avelf, lpop, laam*; Rho is the estimated correlation coefficient between the error terms in the first and second stage. Estimation is by maximum likelihood. First-stage specification of 2SLS in columns (3)–(4) includes (see Table A2, Appendix): for *maj*: *con*2150, *con*5180, *con*81, *engfrac, eurfrac, lpop, avelf,* for *pres*: *con*2150, *con*5180, *con*81, *engfrac, eurfrac, lat*01, *age*; Chi-2 is the test statistic for rejecting the overidentifying restrictions implied by exogenous (additional) instruments; critical value Chi-2 (5,0.05) = 11.07. Propensity-score logit estimation underlying columns (5)–(7) includes: *lyp, prop*65, *gastil, federal, col_uka, laam*; estimates of the constitutional effects in these columns are carried out separately rather than jointly; numbers at the bottom indicate observations used in estimation (observations outside the common support for the propensity score of each constitutional feature are deleted).

* Significant at the 10-percent level.
** Significant at the 5-percent level.
*** Significant at the 1-percent level.

population size and age of democracy also appear among our second-stage regressors, the identifying assumption is that the constitutional dating variables (*con*21, *con*2150, *con*5180), the language variables (*engfrac* and *eurfrac*), latitude (*lat*01) and fractionalization (*avelf*) are all uncorrelated with the remaining unobserved determinants of fiscal policy. We think this is fulfilled in the case of the three constitutional dating variables, while we are less certain about the remaining four instruments. Assuming that the first three instruments are valid, however, the validity of the remaining four can be tested via the implied overidentifying restrictions.

Columns (3) and (4) of Table 3 report the estimated constitutional effects on the size of government, for a second-stage specification that includes our standard controls, the indicator for British colonial origin and, in column (4), the indicator for Latin America (the other colonial and continental indicators are not significant). The estimated constitutional effects remain negative and not too different from the OLS estimates. But the standard errors are larger and only the estimated coefficient of *pres* is significantly different from zero, and only if we omit the Latin American indicator. Note that the overidentifying restrictions for the validity of the instruments are never rejected, even in column (3) where the dummy variable for Latin America is omitted.[14] The fact that the

similar (or more negative) if all second-stage controls are added to the first-stage regression.

[14] Estimating by instrumental variables allows us to have a more parsimonious second-stage specification. The effect of any omitted variable, say being in Latin America, would show up in the residual of the second-stage equation. This

estimated effects remain negative, large, and not too distant from the OLS estimates, reassures us of the validity of our inference, despite the large standard errors.

Relaxing Linearity: Matching Estimates.—How robust are the results when we relax the assumption of linearity? Estimating 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. Paul Rosenbaum and Donald Rubin, 1983). The first one is a version of conditional independence: once we have conditioned upon **X**, the unobserved determinants of the constitution and policy outcomes, i.e., u and e in (1) and (2), are uncorrelated. In the specification of constitution selection, we should thus not omit any variables driving fiscal policy outcomes. This speaks in favor of an inclusive specification.[15] 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, and this means 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 because 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 six potentially important determinants of the size of government: per capita income (*lyp*), the share of old people (*prop*65), the quality of democracy (*gastil*), the presence of a federal system (*federal*), and the indicators for previous British colonies and Latin American location (*col_uka* and *laam*).[16]

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 propensity 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 noncomparable. 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 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, thus replicating within each of these strata the equal-means tests for the whole sample reported in Table 1. In the first and second strata we reject (at the 5-percent level) the null of equal means for only one variable (different variables in the two strata); in the third stratum we can never reject the null. Given the striking mean differences for the whole

would not bias the IV estimates, however, as long as our instruments are not correlated with the omitted variable. Adding all the colonial origin and continental variables to the second stage, the standard errors grow even further.

[15] This was not a concern in the Heckman correction or instrumental variables. On the contrary, in the instrumental variable estimation we deliberately chose a parsimonious first and second stage, to avoid 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 equations.

[16] As noted in the previous section, 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.

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 display the results for the alternative matching estimators described in Section II, with standard errors estimated by a bootstrapping procedure. Notice also that the restriction to the common support means that we are typically discarding five to ten observations. Since we use separate estimation procedures for *maj* and *pres*, the resulting sample size is different for the two sets of estimates.

The matching estimates confirm that our earlier results hold up. Admittedly, 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 nonparametric 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 nonlinearities do not seem to be a major problem plaguing our earlier estimates.

Summary.—The three sets of results 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 1990's cross section. Presidential regimes and majoritarian elections each cut the size of government by about 5 percent of GDP. These results are robust to relaxing the linearity and conditional independence assumptions, and they conform with our theoretical prior.

IV. Composition of Government

Do the constitutional effects extend to other aspects of fiscal policy? In this section, we ask whether majoritarian electoral rules and presidential forms of government indeed cut welfare-state spending, as theory predicts. We rely on the same battery of methods as for the size of government.

OLS Estimates.—Table 4 reports on a variety of linear regression estimates. Our measure of social transfers (*ssw*) is available for a dozen fewer countries than the size of government (*cgexp*). We hold constant the same variables as for the size of government, except that we drop three regressors that are never significant: two demographic variables (*lpop* and *prop*1564) and openness to international trade (*trade*). Column (1) shows that both presidential regimes and majoritarian elections appear to reduce welfare-state spending by 2 percent of GDP, a large and statistically significant effect. Results are similar in other (nonreported) specifications, but they are less robust than those for the overall size of government. One reason for this greater fragility could be collinearity with the proportion of elderly in the population (*prop*65), since both presidential and majoritarian countries have younger populations (recall Table 1).

In column (2), the constitution is 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 also statistically significant. The other estimated coefficients also rise in absolute value, but so do their standard errors, and they are not statistically significant. Since we cannot reject the additivity assumption imposed in column (1) [cf. the *F*-test in column (2)], we focus on this more parsimonious specification.

Column (3) asks whether the constitutional effects were present in the early part of the sample, here defined as the period 1972–1977 (1972 is the first year for which data are available). The answer is a clear and resounding no. The panel analysis (described in detail in Persson and Tabellini, 2001 and 2003) reveals that proportional and parliamentary democracies had a more rapid growth of social security and welfare spending in the 1970's and 1980's, compared to the other constitutional groups. As with the overall size of government, this seems to be due both to a faster time trend as well as to an asymmetric response to cyclical fluctuations.

Repeating the analysis of Table 2, columns (4) and (5) of Table 4 interact the two constitutional indicators with the dummy variables for new (*newdem*) and bad (*baddem*) democracies. Both constitutional effects appear much stronger in the older and better democracies. The effect of presidentialism is particularly large in

TABLE 4—COMPOSITION OF GOVERNMENT AND CONSTITUTIONS: OLS ESTIMATES

	(1)	(2)	(3)	(4)	(5)
Dependent variable	*ssw*	*ssw*	*ssw*	*ssw*	*ssw*
pres	−2.24		−0.25	−5.47	−4.28
	(1.11)**		(2.06)	(1.19)***	(1.30)***
maj	−2.25		−1.02	−2.66	−3.03
	(1.25)*		(1.36)	(1.52)*	(1.50)**
propres		−3.22			
		(1.74)*			
majpar		−3.14			
		(2.18)			
majpres		−3.91			
		(2.41)			
pres_newdem				4.97	
				(1.65)***	
maj_newdem				1.74	
				(1.77)	
newdem				−5.36	
				(1.69)***	
pres_baddem					5.61
					(2.00)***
maj_baddem					3.67
					(1.62)**
baddem					−4.24
					(1.75)**
F-test (*pres*)		0.83		0.17	0.83
F-test (*maj*)				0.65	0.19
Sample	1990's	1990's	1972–1977	1990's	1990's
Observations	69	69	42	69	69
R^2	0.81	0.81	0.77	0.84	0.82

Notes: Robust standard errors are in parentheses. All regressions include our standard controls, *lyp, gastil, age, prop*65, *federal,* and *oecd,* plus a set of indicator variables for continental location and colonial origin, except that *age* is missing in columns (3)–(4), while *gastil* is missing in column (5) and replaced by *polity_gt* in column (3). *F*-test(*pres*) refers to tests of the hypotheses that the coefficient for *propres* is equal to the difference between the coefficients for *majpres* and *majpar* [column (2)], the sum of the coefficients for *pres* and *pres_newdem* is zero [column (4)], and the sum of the coefficients for *pres* and *pres_baddem* is zero [column (5)]. *F*-test(*maj*) refers to the corresponding tests with regard to *maj* [columns (4) and (5)].

* Significant at the 10-percent level.
** Significant at the 5-percent level.
*** Significant at the 1-percent level.

this group of countries, and rises to −5 percent and −4 percent of GDP respectively. Unlike what we found for the overall size of government, however, here the difference between new and old (or good and bad) democracies is generally statistically significant, and no constitutional effect shows up among the younger and worse democracies. In this section, we have fewer degrees of freedom, however, and the interaction effects are more demanding on the data.

IV and Heckman Estimates.—Next, we relax conditional independence, 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 two-stage least-squares (2SLS) estimates.[17]

Table 5 reports two specifications for the second-stage instrumental variable estimates, one exclusive of the dummy variables for Latin America [column (3)], the other inclusive [col-

[17] The maximum likelihood procedure often does not converge, so here we report results from the two-step Heckman estimator.

TABLE 5—COMPOSITION OF GOVERNMENT AND CONSTITUTIONS: INSTRUMENTAL VARIABLES, HECKMAN AND MATCHING ESTIMATES

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable	*ssw*	*ssw*	*ssw*	*ssw*	*ssw*	*ssw*	*ssw*
pres	0.20	−2.38*	0.75	0.49	−3.15	−0.45	−3.75
	(3.27)	(1.33)	(2.00)	(2.14)	(2.65)	(1.78)	(2.50)
maj	−2.05*	−4.27	−3.21	−3.21	−1.84	−2.46	−3.53
	(1.12)	(1.79)**	(1.61)*	(1.62)*	(2.01)	(1.71)	(1.79)*
Conts & Cols	Yes	Yes	*col_uka*	*col_uka laam*			
Sample	1990's	1990's	1990's	1990's	1990's	1990's	1990's
Endogenous	*pres*	*maj*	*pres*	*pres*	*pres*	*pres*	*pres*
Selection			*maj*	*maj*	*maj*	*maj*	*maj*
Estimation	Heckman 2-step	Heckman 2-step	2SLS	2SLS	Stratification	Nearest neighbor	Kernel
Rho	−0.46	0.59					
Chi-2			9.53*	9.98*			
Adjusted R^2			0.78	0.78			
Observations	64	64	64	64	56(*pres*) 58(*maj*)	56(*pres*) 58(*maj*)	56(*pres*) 58(*maj*)

Notes: Standard errors are in parentheses. Always included in second-stage specification in columns (1)–(4): *age, lyp, trade, prop*1564, *prop*65, *gastil, federal, oecd, lpop*; Conts & Cols refer to indicator variables for continental location and colonial history. First-stage specification of Heckman procedure in columns (1)–(2) includes: *engfrac, eurfrac, lat*01, *avelf, lpop, laam*; Rho is the estimated correlation coefficient between the error terms in the first and second stage. First-stage specification of 2SLS in columns (3)–(4) includes (see Appendix): for *maj*: *con*2150, *con*5180, *con*81, *engfrac, eurfrac, lpop, avelf*; for *pres*: *con*2150, *con*5180, *con*81, *engfrac, eurfrac, lat*01, *age*; Chi-2 is the test statistic for rejecting the overidentifying restrictions implied by exogenous (additional) instruments; critical value Chi-2 (5,0.05) = 11.07. Propensity-score logit estimation underlying columns (5)–(7) includes: *lyp, prop*65, *gastil, federal, col_uka, laam*; estimates of the constitutional effects in these columns are carried out separately rather than jointly; numbers at the bottom indicate observations used in estimation (observations outside the common support for the propensity score of each constitutional feature are deleted).

* Significant at the 10-percent level.
** Significant at the 5-percent level.
*** Significant at the 1-percent level.

umn (4)]. Our previous concerns about the validity of the instruments remain, but are not repeated. Here they are heightened by the fact that the overidentifying restrictions can now be rejected at the 10-percent level.

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. In columns (1), (3), and (4), the presidential effect is practically zero [column (2) does not allow for endogenous selection into presidential regimes, and 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 negative (Rho is −0.46). This implies that the OLS estimate for presidential regimes in Table 4 is biased downwards; adjusting for this bias produces a positive but insignificant estimate. The result is confirmed by the instrumental variable estimates, which are also small positive numbers.

The effect of majoritarian elections, on the other hand, is confirmed to be negative and statistically significant according to both procedures. Column (2) in Table 5 suggests errors with a strong positive correlation (Rho is +0.59), 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, statistically significant and larger in absolute value than the OLS estimates [column (2)]. Again, the result is confirmed by the instrumental variable estimates [columns (3)–(4)].

The consistency of the results demonstrates 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 unimportant.

Matching Estimates.—Finally, we turn to the matching methods, 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. They are quite similar to the OLS estimates: 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 percent of GDP. The effect is stronger in the older and better democracies. For presidential regimes, there is less evidence of a stable overall constitutional effect, at least in the full set of democracies. In the case of welfare spending, 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.

V. Conclusion

Do electoral rules and forms of government shape fiscal policy? Our empirical estimates strongly suggest the answer to be yes. Most of them are in line with the first wave of theory discussed in the introduction. As predicted, majoritarian elections lead to smaller governments and smaller welfare programs than proportional elections. Presidential democracies are associated with smaller governments than parliamentary democracies, as predicted by theory emphasizing the force of the confidence requirement. But the auxiliary prediction of the theory, namely that parliamentary democracy should foster larger broad programs, does not find robust support in the data on welfare spending (though the prediction appears to fare better in old and good democracies).

Other findings await a satisfactory theoretical explanation. The cross-country differences we observe today can largely be attributed to different responses to common political or economic events in the 1970's and 1980's. Why do we observe these different patterns in different political systems? Since the theories that motivated this analysis are generally static, they cannot help us answer this question.

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 directly into policy outcomes. This way, we have not been able to identify whether constitutional rules operate 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 the black box of political outcomes 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.

TABLE A1—ELECTORAL RULES AND FORMS OF GOVERNMENT IN THE 1990'S

Country	*maj*	*pres*	Country	*maj*	*pres*	Country	*maj*	*pres*	Country	*maj*	*pres*
Argentina	0	1	Finland	0	0	Netherlands	0	0	Trinidad and Tobago	1	0
Australia	1	0	France	1	0	New Zealand	1	0	Turkey	0	0
Austria	0	0	Gambia	1	1	Nicaragua	0	1	Uganda	1	1
Bahamas	1	0	Germany	0	0	Norway	0	0	Ukraine	1	0
Bangladesh	1	0	Ghana	1	1	Papua New Guinea	1	0	United Kingdom	1	0
Barbados	1	0	Greece	0	0	Pakistan	1	1	United States	1	1
Belarus	1	1	Guatemala	0	1	Paraguay	0	1	Uruguay	0	1
Belgium	0	0	Honduras	0	1	Peru	0	1	Venezuela	0	1
Belize	1	0	Hungary	0	0	Philippines	1	1	Zambia	1	1
Bolivia	0	1	Iceland	0	0	Poland	0	0	Zimbabwe	1	1
Botswana	1	0	India	1	0	Portugal	0	0			
Brazil	0	1	Ireland	0	0	Romania	0	0			
Bulgaria	0	0	Israel	0	0	Russia	0	1			
Canada	1	0	Italy	0	0	Senegal	0	0			
Chile	1	1	Jamaica	1	0	Singapore	1	0			
Colombia	0	1	Japan	1	0	Slovak Rep.	0	0			
Costa Rica	0	1	Latvia	0	0	South Africa	0	0			
Cyprus	0	1	Luxembourg	0	0	South Korea	0	1			
Czech Rep.	0	0	Malawi	1	1	Spain	0	0			
Denmark	0	0	Malaysia	1	0	Sri Lanka	0	1			
Dominican Rep.	0	1	Malta	0	0	St. Vincent and Grenada	1	0			
Ecuador	0	1	Mauritius	1	0	Sweden	0	0			
El Salvador	0	1	Mexico	0	1	Switzerland	0	1			
Estonia	0	0	Namibia	0	1	Taiwan	0	0			
Fiji	0	0	Nepal	1	0	Thailand	1	0			

Notes: Classifications follow criteria described in the text: exclusive reliance on plurality rule in (lower house) legislative elections are coded $maj = 1$, other countries $maj = 0$; countries in which the executive is not accountable to the legislature through a confidence procedure are coded $pres = 1$, others $pres = 0$ (see Persson and Tabellini, 2003, for a discussion of borderline cases). For Fiji, Japan, New Zealand, the Philippines, and Ukraine, which all reformed their electoral rules in the mid-1990's leading to a change in *maj*, the prereform classification is used.

TABLE A2—FIRST-STAGE SPECIFICATION OF 2SLS ESTIMATES

	(1)	(2)
Dependent variable	*pres*	*maj*
*con*2150	−0.04	−0.13
	(0.14)	(0.12)
*con*5180	−0.13	0.28
	(0.18)	(0.10)**
*con*81	0.29	0.12
	(0.20)	(0.11)
engfrac	−0.68	1.09
	(0.13)***	(0.13)***
eurfrac	0.39	−0.21
	(0.11)***	(0.13)
lpop		0.07
		(0.02)***
*lat*01	−1.43	
	(0.34)***	
age	0.56	
	(0.31)*	
avelf		0.74
		(0.21)***
F-test	4.26***	3.26**
R^2	0.51	0.51
Observations	75	75

Notes: Robust standard errors are in parentheses. *F*-test refers to joint significance of *con*2150, *con*5180, and *con*81.
* Significant at the 10-percent level.
** Significant at the 5-percent level.
*** Significant at the 1-percent level.

REFERENCES

Acemoglu, Daron; Johnson, Simon and Robinson, James. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review*, December 2001, *91*(5), pp. 1369–401.

Angrist, Joshua D. and Krueger, Alan B. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives*, Fall 2001, *15*(4), pp. 69–85.

Austen-Smith, David. "Redistributing Income under Proportional Representation." *Journal of Political Economy*, December 2000, *108*(6), pp. 1235–69.

Besley, Timothy and Case, Anne. "Political Institutions and Policy Choices: Evidence from the United States." *Journal of Economic Literature*, March 2003, *41*(1), pp. 7–73.

Cameron, David. "The Expansion of the Public Economy: A Comparative Analysis." *American Political Science Review*, December 1978, *72*(4), pp. 1243–61.

Castles, Francis. *Comparative public policy. Patterns of post-war transformation.* Cheltenham: Edward Elgar, 1998.

Dehejia, Rajeev H. and Wahba, Sadek. "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs." *Journal of the American Statistical Association*, December 1999, *94*(448), pp. 1053–62.

Diermeier, Daniel and Feddersen, Timothy. "Cohesion in Legislatures and the Vote of Confidence Procedure." *American Political Science Review*, September 1998, *92*(3), pp. 611–21.

Hall, Robert and Jones, Chad. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics*, February 1999, *114*(1), pp. 83–116.

Heckman, James J.; Ichimura, Hidehiko and Todd, Petra E. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies*, October 1997, *64*(4), pp. 605–54.

Heckman, James J.; Lalonde, Robert J. and Smith, Jeffrey A. "The Economics and Econometrics of Active Labor Market Programs," in Orley Ashenfelter and David Card, eds., *Handbook of labor economics*, Vol. 3A. Amsterdam: North-Holland, 1999, pp. 1865–2097.

Huber, Evelyne; Ragin, Charles and Stephens, John. "Social Democracy, Christian Democracy, Constitutional Structure and the Welfare State." *American Journal of Sociology*, November 1993, *99*(3), pp. 711–49.

Kontopolous, Yianos and Perotti, Roberto. "Government Fragmentation and Fiscal Policy Outcomes: Evidence from the OECD Countries," in J. Poterba and J. von Hagen, eds., *Fiscal institutions and fiscal performance.* Chicago: University of Chicago Press, 1999, pp. 81–102.

Lechner, Michael. "Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies." *Review of Economics and Statistics*, May 2002, *84*(2), pp. 205–20.

Lijphart, Arend. "The Political Consequences of Electoral Laws, 1945–1985." *American Political Science Review*, June 1990, *84*(2), pp. 481–96.

Lizzeri, Alessandro and Persico, Nicola. "The Provision of Public Goods under Alternative Electoral Incentives." *American Economic Review*, March 2001, *91*(1), pp. 225–45.

Maddala, G. S. *Econometrics.* New York: McGraw Hill, 1977.

Milesi-Ferretti, Gian Maria; Perotti, Roberto and Rostagno, Massimo. "Electoral Systems and Public Spending." *Quarterly Journal of Economics*, May 2002, *117*(2), pp. 609–57.

Persson, Torsten; Roland, Gérard and Tabellini, Guido. "Comparative Politics and Public Finance." *Journal of Political Economy*, December 2000, *108*(6), pp. 1121–61.

Persson, Torsten and Tabellini, Guido. "The Size and Scope of Government: Comparative Politics With Rational Politicians." *European Economic Review*, April 1999, *43*(4–6), pp. 699–735.

______ . *Political economics: Explaining economic policy.* Cambridge, MA: MIT Press, 2000.

______ . "Political Institutions and Policy Outcomes: What Are the Stylized Facts?" Mimeo, Stockholm University, 2001.

______ . *The economic effect of constitutions.* Cambridge, MA: MIT Press, 2003.

Rae, Douglas. *The political consequences of electoral law.* New Haven, CT: Yale University Press, 1967.

Rodrik, Dani. "Why Do More Open Economies Have Bigger Governments?" *Journal of Political Economy*, October 1998, *106*(5), pp. 997–1032.

Rosenbaum, Paul and Rubin, Donald. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, April 1983, *70*(1), pp. 41–55.

Strom, Kaare. *Minority government and majority rule.* Cambridge: Cambridge University Press, 1990.

Taagepera, Rein and Shugart, Matthew. *Seats and votes: The effects and determinants of electoral systems.* New Haven, CT: Yale University Press, 1989.

Wooldridge, Jeffrey. *Econometric analysis of cross section and panel data.* Cambridge, MA: MIT Press, 2002.