



It's a matter of confidence. Institutions, government stability and economic outcomes

Luca Bettarelli¹ · Michela Cella² · Giovanna Iannantuoni² · Elena Manzoni³

Received: 2 July 2019 / Accepted: 25 November 2020 / Published online: 24 December 2020
© The Author(s) 2020

Abstract

In this paper we analyse the effect of constitutional structures over policy outcomes. In particular, we exploit the heterogeneity in parliamentary systems deriving from the presence and the use of the confidence vote to investigate whether stable and unstable parliamentary systems behave differently in terms of the policies they implement. This finer partition of parliamentary systems allows us to identify effects that are more robust than the ones previously discussed in the literature. We show that the difference between presidential and parliamentary systems documented in previous works is driven by a difference between presidential and stable parliamentary systems. We suggest that possible transmission channels are legislative cohesion and (the absence of) selection.

Keywords Presidential system · Parliamentary system · Confidence vote · Government stability

JEL Classification C72 · D72

The authors wish to thank two anonymous reviewers, Alisher Aldashev, Lorenzo Cappellari, Piergiorgio Carapella, Laura Pagani, participants to the 2016 Journées Louis-André Gérard-Varet, and participants to the 2017 EPCS Annual Meeting (Budapest) for useful comments and suggestions. They also express gratitude to Lorenz Blume, Jens Mueller, Stefan Voigt and Carsten Wolf for sharing their dataset.

✉ Luca Bettarelli
luca.bettarelli@ulb.ac.be.it

Michela Cella
<https://sites.google.com/site/mcella/home>

Elena Manzoni
<https://sites.google.com/site/elmanzoni/>

¹ Department of Political Science (Cevipol), Université Libre de Bruxelles (ULB), Avenue F.D. Roosevelt 50-CP 124, 1050 Brussels, Belgium

² Department of Economics, Management and Statistics, University of Milan-Bicocca, Milan, Italy

³ Department of Economics, University of Verona, Verona, Italy

1 Introduction

Over the last decade the political economy literature has focused on the impact that political institutions have on national policies (Persson 2002) and the seminal work of Persson and Tabellini (2003) has shown that institutions, namely political regimes, do matter in shaping the size and the composition of government spending.

While their partition of institutional frameworks in parliamentary and presidential is the most intuitive, it does not lead to results that are robust to changes such as, for example, the set of countries or the times span as Blume et al. (2009) have highlighted.

We suggest that policy decisions, and government spending among them, depend on how political actors respond to the incentives put in place by the institutional system in which they operate. In other words, the equilibria that emerge from the political process depend on the interaction of the institutional mechanisms defined by the constitutions with the underlying characteristics of the political environment.

We therefore focus on the presence of the confidence vote,¹ traditionally the main element of distinction between parliamentary and presidential regimes, and exploit its effective use as a way to account for the interaction between institutions and other characteristics that may influence policy implementation. In some cases, the confidence vote does indeed generate frequent changes of government, thus replacing possibly bad politicians and generating a different government composition (*selection effect*). In some others instead, the confidence vote acts as a credible threat and may induce either the executive to behave better (*disciplining effect*) or the parliament to accept more frequently the executive's misbehaviour (*legislative cohesion*). Hence, the performance of parliamentary systems may depend on politicians' characteristics such as, for example, the quality of the information available and/or the alignment of their interests with the citizens.

The *de facto* implementation of a constitutional feature - the confidence vote—allows us to highlight the greater heterogeneity of parliamentary systems. Given this complexity, we investigate more deeply the characteristics of countries that adopt a parliamentary constitution by considering the stability of the government as a proxy to distinguish different parliamentary systems (Lijphart 2004). We measure stability as inversely related to the frequency of government changes, which is clearly correlated with the effective use of the confidence vote.

Dividing parliamentary systems in stable and unstable ones responds to the suggestion made by some authors (e.g., Acemoglu 2005; Voigt 2011) that the distinction between parliamentary and presidential systems may simply be too coarse to robustly explain the differences in policy outcomes that we observe.

We contribute to the literature that started with Persson and Tabellini (2003). In their paper, the authors compare constitutional systems—presidentialism vs. parliamentarism—and electoral rules—majoritarian vs. proportional—in order to identify

¹ For a detailed review of the relevance of the confidence vote, see Lijphart (1999).

the differences, if any, in a number of relevant social and economic indicators.² Without doubting their seminal role, a large number of follow-up articles challenged their results. Blume et al. (2009) found that, while the results on the effect of the electoral rules are robust, the effects of the parliamentary *vs.* presidential constitutional choice are sensitive to an enlargement of the dataset and the updating of the economic indicators used as regressors.

Moreover, Acemoglu (2005) and Voigt (2011), among others, questioned the presidential/parliamentary classification of constitutional structures in a twofold manner. On the one hand, they claim that the constitutional form of government is endogenous, being an equilibrium outcome rather than an exogenous characteristic. On the other hand, they advocate a finer partition of countries, taking into account the heterogeneity within each group.

The aim of this paper is to verify whether our partition, which is based both on formal institutions (presidential *vs.* parliamentary) and their actual functioning (stable *vs.* unstable), may be one of such finer partitions missing in the existing literature. As a matter of fact, our partition stems from the equilibria emerging from the interaction between institutional and other political elements and therefore could shed light on potential transmission channels generating them, thus further contributing to the existing literature. Our main result is that our classification of constitutional systems (presidential, stable parliamentary, unstable parliamentary) delivers results that are more robust than those in the literature. More precisely, we find that stable parliamentary systems are significantly different both from presidential and unstable parliamentary ones. Unstable parliamentary systems and presidential systems, instead, behave alike in terms of the policy they implement. This result is robust to changes in the set of countries included in the dataset and in the definition of stability. Moreover, our analysis tries to address the endogeneity issue raised by Acemoglu (2005) and Voigt (2011), by using a novel instrument.

In Sect. 2 we suggest how our findings may be explained by two effects that are consistent with previous theoretical literature: a selection effect, as in Cella et al. (2017), and a legislative cohesion effect, as in Baron (1998), Diermeier and Feddersen (1998a, 1998b) and Diermeier and Vlaicu (2011). Section 3 presents the data and the empirical strategy, Sect. 4 discusses the results and lists robustness checks and Sect. 5 concludes.

2 Theoretical background

The idea that further partitioning parliamentary systems may provide insights about the link between institutions and economic policy finds its support in several theoretical papers. As a matter of fact, while the interplay of uncertainty and incentives

² Persson and Tabellini identify presidential/parliamentary regimes according to the legal existence of the confidence vote, so that presidential countries where the government is subject to a confidence vote—as for instance France—are classified as parliamentary. The authors find that presidential systems systematically spend the five percent less than parliamentary systems.

operates unambiguously in presidential systems and generally gives rise to unique equilibria, parliamentary systems are more heterogeneous and can support multiple equilibrium strategies and induce multiple equilibrium outcomes (see among others Cella et al. 2017). For example, outcomes of parliamentary systems may differ depending on whether the confidence vote is mostly used as a threat or whether it effectively replaces executives. This refers to the *de facto* behaviour of politicians that in turn depends on their (short- or long-term) incentive structure. In this context, the confidence vote may affect government duration since a legislative defeat would lead to new elections that will replace the executive and the legislative body with positive probability. On the contrary, in presidential systems both bodies have fixed terms and government stability is not affected by the working of the policy process.³ In other words, under presidentialism, politicians face undistorted incentives and vote according to policy preferences.

In particular, the literature identifies three channels through which the confidence vote may affect the policy-making process in parliamentary systems. First of all, if the confidence vote is actively used to replace politicians, it improves their expected quality (*selection effect*, see Cella et al. 2017; Huber and Gallardo 2008). If instead the confidence vote acts as a threat, it may either reduce the distortions to the executive's behaviour (*disciplining effect*, see Cella et al. 2017; Huber 1996), or induce the voting cohesion in parliament (*legislative cohesion*, see Baron 1998; Diermeier and Feddersen 1998b; Diermeier and Vlaicu 2011).

Cella et al. (2017) highlight this twofold nature of the confidence vote. They model an executive and a legislative body in a parliamentary system where politicians may face early elections if the parliament does not approve the executive's proposed policy. Politicians can be of two types, they either care about implementing the efficient policy or they only care about being in office. The authors show that in such a setting two equilibria may arise, depending on the parameters that describe politicians' quality (type distribution) and information. The confidence vote may act as a threat and induce an office oriented executive to propose the efficient policy in order to prevent early termination of the legislature. In presence of this *disciplining effect*, stable systems are characterised by a low level of inefficiency. Alternatively, the confidence vote may be used in equilibrium to replace possibly bad politicians. As, on average, office oriented politicians are replaced more often, the expected quality of the executive improves. Hence, the selection effect operates more efficiently in unstable systems in which we should observe a better alignment of the executive's and voters preferences.

In a model of parliamentary democracy, where the government controls the legislative agenda, Baron (1998) shows instead that government members may change policies regarding government spending to preserve the government and may also

³ As noted by Diermeier and Vlaicu (2011): "Under presidentialism the policy process is driven by short-term issue-by-issue incentives because there are only short-term consequences of an unsuccessful proposal. In parliamentary systems, on the other hand, the failure of a policy proposal can lead to a change in the composition of the governing coalition. This injects political incentives in the policy process whereby coalition members consider both their short-term policy interests and their long-term political interest".

seek support from the minority in parliament. In other words, the legislative cohesion effect contributes to the stability of parliamentary systems through the approval of a larger fraction of policies aimed at keeping current politicians in power.

The co-existence of these effects implies that parliamentary systems are more heterogeneous than presidential ones, given that their response to the policy implementation process and their degree of stability may vary depending on their characteristics even for a given set of constitutional rules. Moreover, stability and policy response are theoretically correlated, so that it is meaningful to use stability as a proxy to refine the classification of parliamentary systems. The focus of our paper is to exploit empirically this *de facto* heterogeneity.

Starting from these considerations, our paper tries to shed light on the complex mechanisms that link constitutional features to economic outcomes. We compare the effects that constitutional structures have on the policy-making process, adopting a partition of parliamentary systems that takes into account their degree of stability. In particular, the disciplining effect would imply that the difference between presidential and parliamentary structures should derive from unstable parliamentary systems. On the opposite view, both legislative cohesion and selection effects would induce stable systems to be the ones that differ more from presidential ones. Hence our approach will allow a better understanding not only of the existence of the link between institutions and policies but also of the underlying mechanisms that generate it.

3 Empirical strategy

We begin our analysis by introducing our refined classification of countries, which, for parliamentary systems, takes into account the level of stability. To do so we create a stability index, *gov life*, which is the average length of any electoral cycle, computed for each country i , normalized by the legal length of the electoral cycle, over the period 1975–1989.⁴ We limit the time span to 1989 in order not to overlap with the dependent variables, which have been originally built by Persson and Tabellini

⁴ More precisely, the index *gov life* is built using the indicator *yrcurnt* from the World Bank Database of Political Institutions (DPI 2012). The indicator is coded as follows: zero in the election year, $L_t - k$ years after the election, where L_t is the legal length of any electoral term according to country-specific constitutional rules. From the indicator *yrcurnt*, for every legislature l , we build the variable D_l , which characterises the number of years a government has been in office between two elections. The stability index *gov life* is then defined as:

$$gov\ life = \frac{\sum_l D_l / \sum_l E_l}{L_t}, \quad (1)$$

where E_l is a dummy which indicates elections.

as averages over the period 1990–1998.⁵ The index *gov life* ranges from zero to one, with higher values corresponding to higher stability.⁶

Thus, we introduce our finer classification of countries that further partitions parliamentary systems into stable and unstable ones. We classify parliamentary countries in *parl stab* if their value of *gov life* is above the median of the stability distribution, and *parl unstab* if their value of *gov life* is below the median.⁷

In the empirical analysis, we focus on two separate datasets. The first one is the dataset used in Persson and Tabellini (2003) (hereafter PT dataset). The dataset is composed of 85 countries, and it includes data on a set of economic and social indicators.⁸ We do not update the variables in the PT dataset as we use it to evaluate how our specification compares to the original one.

The second one is our general dataset (BCIM dataset). It is obtained from the dataset used in Blume et al. (2009), which extends the PT dataset to include 116 countries.

The descriptive statistics of the dependent and independent variables that we consider are in Table 1.

Recall that our analysis focuses on the role of stability on the policy differences between presidential and parliamentary systems. Specifically, we investigate whether the differences found by Persson and Tabellini (2003) are homogeneously driven by stable and unstable parliamentary countries. We therefore first run a set of regressions described by Eq. (2), estimated through OLS:

$$Y_i = \alpha + \beta_1 \text{parl stab}_i + \beta_2 \text{parl unstab}_i + \delta X_i + \varepsilon_i, \quad (2)$$

where *pres* is the baseline category that represents the control group in our setting.

Dependent variables The main dependent variables are central government expenditure (*cgexp*) and central government revenues (*cgrev*), computed as percentage of the GDP, and averaged over the period 1990–1998.⁹

Independent variables The main explanatory variables are the dummies that classify countries according to the stability index, as discussed above (*pres*, *parl stab* and *parl unstab*). The controls include: *maj*, which is equal to one for majoritarian electoral rules and zero for proportional ones; indicators for the continental location, such as OECD (*oecd*), Central, Latin America and Caribbeans (*laam*), Africa

⁵ Our strategy does not take into account that structural changes in stability due to constitutional changes may have happened during the time span in which we compute the dependent variable, i.e. 1990–1998, and therefore after the last year we consider for the independent variables (i.e. 1989). In order to take this into account, we measured the (in)stability of parliamentary systems using the index *gov life* over the period 1975–1998, instead of 1975–1989, thus including all the years that are relevant for the dependent variable. Results, which are available upon request, are qualitatively identical to those we report in this paper, thus suggesting that the constitutional changes that may have happened in the period 1990–1998, on average, do not affect our analysis.

⁶ As a robustness check, we introduce different measures of stability in Sect. 4.2.

⁷ More precisely, we first drop three countries which are within 0.05 points from the median of the stability distribution in order to avoid a random assignment of countries due to measurement errors.

⁸ For a detailed list of variables and sources, see Persson and Tabellini (2003).

⁹ We consider additional dependent variables in Sect. 4.2 to provide robustness checks.

(*africa*), South and Central Asia (*asiae*); indicators for the colonial history, such as *col_espa* (the country is a former colony of Spain or Portugal), *col_uka* (the country is a former English colony) and *col_otha* (the country is a former colony of a country other than England, Spain and Portugal);¹⁰ democracy level (*gastil*); per-capita income (*lyp*); proportion of people between the age 15–64 (*prop1564*) and over 65 (*prop65*); population size (*lpop*); and a dummy indicating a federalist system (*federal*).

For comparison reasons, we also replicate the following specification used by Persson and Tabellini (2003) and Blume et al. (2009):

$$Y_i = \alpha + \beta pres_i + \delta X_i + \varepsilon_i, \quad (3)$$

where the only difference with specification (2) is that countries are partitioned in presidential and parliamentary countries only, and *parl* is the omitted category.

IV approach The OLS analysis may be biased due to the possible endogeneity of our classification of constitutional systems (Acemoglu 2005). As a matter of fact, political institutions may be shaped by the same factors that determine policies, thus making them endogenous. Therefore, we also perform our analysis using an instrumental variables (IV) approach (Sect. 4.1).

4 Results

We begin our analysis with a preliminary look at Table 1, which reports the descriptive statistics of the dependent and independent variables, split by constitutional groups. We note that both government size and government revenues are significantly smaller in presidential than parliamentary countries. This is also true for the share of welfare spending in the total government size. Instead, the variation within parliamentary systems is less striking, and the only conclusion we can draw at this stage is that the average government revenue is significantly higher in stable parliamentary countries than in unstable ones.

We then compare our refined classification with the standard analysis by Persson and Tabellini (2003) and Blume et al. (2009). In order to do so, we first run regressions (2) and (3) on the PT dataset using central government expenditure and central government revenues as dependent variables (Table 2). Columns (1) and (3) replicate the standard results in the literature, namely that presidential systems spend systematically less than parliamentary ones, regardless of the chosen measure of the government size.

Columns (2) and (4), instead, show that the difference between constitutional systems is driven by the subgroup of stable parliamentary countries. Indeed, the coefficient of *parl stab*, β_1 , is statistically significant in both regressions. On the

¹⁰ All the variables related to the colonial history are weighted for the years of independence as follows: $col_uka = col_uk * (250 - t_indep)/250$, where $col_uk = 1$ is a dummy indicating a former English colony, $t_indep \in [0, 250]$ are the years of independence and 250 is used as the standard value for all non-colonized countries. The same holds for *col_espa* and *col_otha*.

Table 1 Descriptive statistics

	min	max	pres	parl unstab	parl stab	p(1,2)	p(1,3)	p(2,3)	p(pt,bcim)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
cgexp	10.7	51.4	22.9	32.1	32.5	0.00	0.00	0.32	0.43
cgrev	10.9	50.8	20.7	27.6	30.7	0.00	0.00	0.08	0.91
ssw	1.13	21.1	4.84	9.25	9.89	0.00	0.02	0.60	-
socprot	3.24	18.4	8.64	12.4	12.8	0.01	0.00	0.82	0.27
edspend	1.98	9.15	3.92	4.86	4.94	0.01	0.00	0.78	0.35
cpi	1.27	8.25	6.19	4.24	2.83	0.00	0.00	0.06	0.05
logyl	6.95	10.5	8.75	9.24	9.44	0.05	0.01	0.47	0.00
ggexp	14.2	59.2	26.7	35.3	36.0	0.00	0.00	0.84	0.69
maj	0	1	0.32	0.36	0.61	0.78	0.02	0.08	0.29
lpop	-4.34	6.81	1.95	2.13	1.35	0.78	0.22	0.18	0.00
federal	0	1	0.21	0.12	0.11	0.39	0.29	0.89	0.20
lyp	6.27	9.94	7.88	8.53	8.77	0.01	0.00	0.31	0.22
trade	17.6	343.4	68.8	79.5	92.5	0.27	0.05	0.31	0.05
gastil	1	4.89	2.92	2.49	2.12	0.09	0.07	0.08	0.06
prop1564	48	72	59.2	62.5	62.5	0.03	0.02	0.97	0.18
prop65	2	17.4	5.83	9.78	8.87	0.00	0.01	0.95	0.11
age	0.03	1	0.14	0.21	0.23	0.16	0.08	0.97	0.00
oecd	0	1	0.05	0.34	0.36	0.00	0.00	0.98	0.08
africa	0	1	0.20	0.17	0.08	0.78	0.14	0.15	0.39
asia	0	1	0.10	0.14	0.21	0.63	0.19	0.58	0.08
laam	0	1	0.52	0.07	0.18	0.00	0.00	0.25	0.89
col_uka	0	1	0.24	0.25	0.45	0.96	0.04	0.08	0.90
col_espa	0	1	0.14	0.04	0.02	0.03	9.02	0.29	0.58
col_otha	0	1	0.16	0.21	0.24	0.62	0.37	0.71	0.36
magn	0.01	1	0.46	0.45	0.59	0.86	0.19	0.18	0.18
right_left	0	1	2.02	2.02	2.01	0.75	0.64	0.82	0.71
n. obs.			40	35	41				

Entries in columns (3), (4) and (5) are mean values for constitutional categories when using the BCIM extended dataset. Entries in columns (6), (7) and (8) are p values. $p(x,y)$ is the *t-test* on the probability of falsely rejecting equal means across groups corresponding to columns x and y , under the assumption of equal variances. Column (9) is the probability of falsely rejecting equal means across the original PT dataset and the BCIM extended dataset, under the assumption of equal variances. *n. obs.* at the bottom of the Table reports the number of countries for each constitutional group

contrary, the coefficient of *parl unstab*, β_2 , is never significantly different from zero, that is, we cannot reject the hypothesis that unstable parliamentary systems behave like presidential ones. A test for $\beta_1 = \beta_2$ shows that we can reject the null hypothesis in both cases (p-values of 0.013 and 0.002 respectively); thus, stable parliamentary systems spend significantly more than unstable ones, regardless of the chosen measure of expenditure.

Table 2 Constitutions, Central Government Expenditure and Central Government Revenues. OLS estimations. PT dataset

Dep.Var.	cgexp (1)	cgexp (2)	cgrev (3)	cgrev (4)
pres	-5.181*** (1.93)		-5.001** (2.02)	
parl stab		6.298*** (2.28)		7.541*** (3.20)
parl unstab		1.383 (1.87)		-0.104 (1.05)
maj	-6.315*** (2.11)	-7.566*** (1.99)	-4.681** (2.15)	-5.567*** (1.98)
gastil	-3.060** (1.25)	-3.179** (1.22)	-2.439* (1.33)	-2.532** (1.25)
age	-3.622 (4.157)	-4.438 (4.015)	-2.450 (4.142)	-3.053 (3.740)
lyp	-0.658 (2.09)	-0.365 (2.28)	2.249 (1.73)	2.368 (2.03)
lpop	0.957 (0.80)	1.730** (0.83)	1.020 (0.78)	2.130** (0.83)
trade	0.042* (0.02)	0.062** (0.02)	0.074*** (0.02)	0.101*** (0.02)
prop1564	-0.331 (0.36)	-0.330 (0.38)	-0.392 (0.32)	-0.360 (0.32)
prop65	0.932** (0.44)	0.953** (0.47)	0.914** (0.45)	0.917* (0.46)
pederal	-4.450* (2.46)	-5.200** (2.50)	-4.719** (2.28)	-5.822** (2.38)
oecd	-0.683 (4.04)	-0.778 (3.97)	-2.265 (4.40)	-2.613 (3.75)
africa	1.690 (4.88)	2.213 (4.72)	7.382 (6.44)	8.137 (5.30)
asiae	-4.763 (3.30)	-3.660 (3.47)	-1.989 (3.81)	-1.244 (3.68)
llam	-5.983* (3.01)	-6.812** (3.07)	-2.273 (3.90)	-3.704 (3.55)
col_espa	5.063 (5.58)	2.651 (4.70)	5.233 (5.39)	3.118 (4.09)
col_uka	6.194* (3.36)	6.290* (3.34)	2.306 (3.33)	2.364 (2.73)
col_otha	-2.239 (2.87)	-3.581 (2.57)	-1.372 (3.25)	-3.528 (2.79)
F test		6.97**		10.82***
Observations	80	80	76	76
Adjusted R^2	0.63	0.64	0.58	0.64

White heteroskedasticity-consistent standard errors in parentheses

F test (columns (2) and (4)) for $\beta_1 = \beta_2$.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Let us now comment on the effects of the other control variables. First, we note that countries with majoritarian electoral system (*maj*) spend systematically less than those with a proportional one, and this is consistent with the previous literature (see Persson 2002). The same effect holds for federalist countries. The variable *gastil*, which is inversely related to the quality of the democracy, has a negative effect on both government size and revenues, as expected. On the contrary, costs and revenues increase with population (even though the level of significance varies depending on the econometrics specification) and with the share of citizens over 65. Geographical dummies, and those indicating colonial history, are only weakly correlated with government expenditure and revenues. Note that covariates' magnitude, sign and significance are only marginally affected by our specification of the model, thus ensuring that we do not increase correlation between independent variables when we introduce our finer classification of constitutional systems.

We then run the same set of regressions on the extended BCIM dataset to check the robustness of our approach (Table 3). As shown by Blume et al. (2009), and replicated in columns (1) and (3), the difference between constitutional systems in the traditional classification is no longer significant, even though the coefficients retain the same sign.

Columns (2) and (4) display the results of our specification. Results of Table 2 prove to be robust to changes in the dataset. First, the coefficient β_1 is still significantly different from zero, showing that stable parliamentary systems spend more than presidential ones. Moreover, β_2 is not significantly different from zero, that is, we cannot distinguish the performance of unstable parliamentary systems from presidential ones. Finally, we can reject the hypothesis that $\beta_1 = \beta_2$ (p values are 0.049 and 0.009, respectively). In other words, our analysis strongly suggests that parliamentary systems are not a homogeneous group, so that refining their classification is a modelling improvement.

Moreover, as in Blume et al. (2009), the effect of the electoral system remains robust to the change in the dataset.¹¹ The same is true for the dummy indicating a federalist system and the proportion of people over 65. Again, we highlight that our classification of countries does not interfere with other covariates.

4.1 Instrumental variables

As discussed in Sect. 3, our model may give rise to issues of endogeneity, due to omitted variables. This may be particularly relevant given the cross-section nature of

¹¹ Previous literature has analysed the interplay between constitution and voting system, showing that countries adopting presidential constitutions and majoritarian electoral systems spend, on average, 10% less than parliamentary countries with proportional voting (Persson and Tabellini 2004). We have replicated the same analysis in order to verify if our setting provides novel insights also with respect to the interplay between constitution and voting system. Results, which are available upon request, confirm the findings of the previous literature. In particular, the electoral rule may either mitigate or exaggerate the difference between constitutional systems, depending on whether a parliamentary country adopts a majoritarian or a proportional rule, respectively. As a result, the highest government expenditure is observed in stable parliamentary countries adopting a proportional voting system.

our dataset, as we are not able to include fixed effects at country-level to control for time-invariant country-specific characteristics.¹² Hence, we replicate our baseline analysis with an IV approach, as previously done in the relevant literature (Persson and Tabellini 2004; Voigt 2011).

We introduce a selection equation where we let the constitutional choice C depend on a set of observable characteristics:

$$C_{ij} = X_i\delta + Z_i\gamma + \eta_{ij}. \quad (4)$$

Subscripts i and j refer to individual observations and outcome alternatives respectively (where $j = 0, 1, 2$ indicates presidential, stable parliamentary and unstable parliamentary systems), X_i is the same set of controls entering equation (2), and Z_i is the vector of instruments.

Our analysis differs from the IV strategy adopted by Persson and Tabellini (2004) in two ways. First, we estimate the entire model using the Conditional Mixed Process (CMP), developed by Roodman (2011). In our framework, it is a limited information maximum likelihood (LIML) estimator, which replicates standard IV intuitions but considers both the headline (2) and the selection (4) models as a joint system of equations. This method may potentially generate a gain in efficiency because (i) it directly accounts for the non-continuous nature of the endogenous variables (i.e., the type of constitution), and (ii) it considers potential linkages between the error processes of headline and selection equations.¹³ In detail, we estimate the selection equation by a multinomial probit model, and the headline regression through OLS. In both equations, our constitutional setups enter as a single categorical variable, as defined above. However, we recognize that little is known empirically about the actual performance of the CMP estimator with a multinomial first stage; as a result, its behaviour may be unpredictable. Thus, we replicate the IV analysis using alternative approaches: a standard 2SLS, and H-FUL estimator by Hausman et al. (2012) that allows for heteroskedasticity.¹⁴ In these models, we include two binary endogenous variables, *parl stab* and *parl unstab*, in the first stage. This exercise allows us

¹² We acknowledge that the use of panel data techniques would make the empirical analysis more robust. However, there is in general very little variation of constitutional systems at country level, and our data are no exception to this. According to the Database of Political Institutions, issued by the World Bank, among the 116 countries we include in our analysis, in the period 1975–1989, there are only 24 cases in which countries have changed their constitutional system, out of 1,326 observations available, thus representing only 1.81% of total observations. If we extend the time span to the entire period potentially under investigation in our paper (till 1998), we have 35 constitutional changes out of 2,229 observations, 1.57% of total observations. This limits the possibility to run panel data models, as already stressed by several authors (Gurr 1975; Persson and Tabellini 2004) who have defined the constitution as the *iron law*, due to the typical constitutional inertia (Persson and Tabellini 2004, p. 28).

¹³ For more information about the econometrics of the CMP approach, see Roodman (2011). Bound et al. (1995), Baum (2014), Rivers and Vuong (1988) and Wooldridge (2010) discuss in detail why ML estimators may be more efficient than standard two-stages techniques when the IV approach involves not-continuous variables. This approach has been extensively adopted in the empirical literature (see for example Perez and Sanz 2005; Petreski et al. 2014; Suarez-Varela et al. 2015).

¹⁴ H-FUL has been proposed by Hausman et al. (2012) to deal with issues raised by the presence of many instruments in heteroskedastic data. The proposed solution is a Fuller (1977) like estimator with standard errors that are robust to heteroskedasticity and many instruments. Recently, Anatolyev and Skolkova (2019) have made available a Stata command that performs the H-FUL estimator: *mivreg*.

Table 3 Constitutions, Central Government Expenditure, Central Government Revenues. OLS estimations, BCIM dataset

Dep. var.	cgexp (1)	cgexp (2)	cgrev (3)	cgrev (4)
pres	-3.755 (2.42)		-2.701 (2.36)	
parl stab		5.206** (2.11)		6.882*** (2.32)
parl unstab		1.734 (1.98)		0.302 (1.95)
maj	-5.681*** (2.02)	-6.229*** (2.05)	-2.969** (1.15)	-4.352** (1.85)
gastil	-1.937 (1.35)	-1.948 (1.31)	-1.549 (1.37)	-1.487 (1.18)
age	-4.347 (4.33)	-5.187 (4.14)	-2.913 (3.82)	-3.271 (3.66)
lyp	-0.347 (2.11)	-0.155 (2.11)	3.104* (1.79)	3.446* (1.90)
lpop	0.491 (0.82)	0.881 (0.98)	0.716 (0.61)	1.435* (0.77)
trade	0.043 (0.03)	0.054 (0.03)	0.071*** (0.02)	0.086*** (0.02)
prop1564	-0.176 (0.32)	-0.234 (0.32)	-0.286 (0.29)	-0.326 (0.28)
prop65	0.862** (0.43)	0.877* (0.45)	1.060** (0.41)	1.057** (0.43)
federal	-5.177** (2.13)	-5.137** (2.50)	-5.904*** (2.16)	-5.976** (2.45)
oecd	-1.699 (3.87)	-1.954 (3.77)	-4.058 (4.12)	-4.599 (3.62)
africa	-2.700 (5.05)	-2.222 (5.12)	3.370 (5.58)	4.387 (5.06)
asiae	-7.744** (3.56)	-7.244* (3.77)	-3.159 (3.37)	-3.063 (3.52)
laam	-8.407* (4.38)	-8.761* (4.56)	-3.885 (4.04)	-4.301 (4.01)
col_espa	2.282 (6.29)	2.272 (5.24)	3.221 (5.18)	2.242 (4.17)
col_uka	5.901* (3.14)	5.718* (3.20)	3.340 (2.37)	3.470 (2.15)
col_otha	-4.046 (2.91)	-3.942 (2.67)	-3.040 (3.18)	-3.458 (2.78)
F test		8.30***		9.82***
Observations	91	89	87	85
Adjusted R^2	0.67	0.69	0.66	0.71

White heteroskedasticity-consistent standard errors in parentheses

F test (columns (2) and (4)) for $\beta_1 = \beta_2$.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

to test the robustness of our IV analysis, and to provide standard diagnostic tests on the validity of the instruments. Results are reported in Table 6, in Appendix A.

Second, we modify the set of instruments to make the analysis more robust. We take into account both Persson and Tabellini (2004) justification of their instruments, and the critiques discussed in later works (see in particular Acemoglu 2005). Thus, we maintain in our set of instruments only some of those adopted by Persson and Tabellini (2004). Specifically, we include in Z_i the variables *con2150*, *con5180* and *con81*, respectively dating the origin of the constitution between 1921–1950, between 1951–1980, and after 1981, with before 1921 as the omitted category. These instruments are motivated by the fact that waves of adoptions of specific types of constitution have historically been observed, and this guarantees the correlation with the constitutional choice. The exclusion restriction is instead ensured by the time lag that separates the constitutional dating variables from more recent policy outcomes. Instead, we do not include the Hall and Jones (1999) variables as instruments (i.e., *engfrac*, *eurfrac*, *lat01*, which indicate the fraction of the population speaking major European languages as native tongues and the distance from the equator, respectively), as Persson and Tabellini (2004) did, as their validity has been widely questioned in the literature (see Acemoglu 2005; Rockey 2012), and by Persson and Tabellini themselves (Persson and Tabellini 2004, p. 37).

Moreover, we include a novel instrument in the analysis. We do so as the predictive power of the constitutional dating variables is somewhat weak, even though the F test on their joint validity significantly rejects the null hypothesis. Our novel instrument, *confl mean*, is computed as the sum of years a country has been involved in violent conflicts between 1800 and the year the country adopted a constitution for the first time, over the reference period.¹⁵ Violent conflicts include both intra and inter state conflicts.¹⁶ The variable *confl mean* significantly affects the probability of a country to fall into a particular category of our classification.

First, the degree of conflict may strongly impact the choice of the constitutional system itself. High values of the index may indicate deeply divided societies in which political decision making needs to rely on power-sharing rules. Parliamentary systems offer the ideal political environment for a broad power-sharing executive, given that the cabinet is a collegial decision-making body (Lijphart 2004). On the contrary, presidential systems introduce rules that favour a winner-takes-all outcome, also by facilitating the adoption of a majoritarian electoral rule (Linz 1994). As additional evidence in favour of this correlation, Jung and Deering (2015) show that unstable conditions at the time of the constitutional choice increase the likelihood of adoption of a parliamentary system. Second, this effect should be larger in the case of unstable parliamentary systems given that the degree of stability of a

¹⁵ Source: Correlate of War Project.

¹⁶ For a detailed review of the definition and categorization of conflicts, see Sarkees (2010).

country's political environment may be persistent, i.e., a more unstable environment is more likely to arise in presence of past political instability (Alesina et al. 1996). Therefore, we argue that our new instrument has a robust predictive power for the endogenous regressor, i.e., the constitution type in our classification.

As far as the exclusion restriction is concerned, we recognize that some issues may arise, given that countries may increase government expenditure to finance conflicts. However, we argue that this is not a crucial issue in our analysis, given that the instrument has been computed for a period that long precedes policy choices. As a matter of fact, we compute the instrument for a period before constitution adoption, and we note that the average time between constitution adoption and policy choices is long (81.2 years). Therefore, the level of conflict that we consider as an instrument cannot have a direct effect on the dependent variable, given the variety of policy relevant reforms, events and shocks that may have occurred in that period (e.g., political reforms, changes in the political environment, ideological orientation of governments).¹⁷

To sum up, the vector Z_i includes the three constitutional dating variables, already adopted by Persson and Tabellini (2004), plus the new instrument *confl mean*. For the IV analysis, we restrict the focus on our main research goal, thus considering Eq. (2) as headline regression and the BCIM dataset.

Before discussing IV results, we comment on the validity of the instruments. The bottom part of Table 6, in Appendix A, reports the standard diagnostic tests. The performance of the instrumenting strategy is satisfactory, as it significantly rejects the null-hypothesis for under-identification, while it does not reject the null for over-identification. The Stock–Yogo test for weak-instruments indicates that instruments do not satisfy the 5% critical value, but they do satisfy the 10%. This represents further evidence in favour of the adoption of the ML approach. The endogeneity test, which under the null indicates that the endogenous variables should be treated as exogenous, rejects the null-hypothesis at the 5% level, thus justifying the adoption of the IV strategy.

Let us now discuss IV results of the headline equation, obtained using the LIML estimates.¹⁸ Table 4 tests the model with central government expenditure and central government revenues as dependent variables. Coefficients are similar to those obtained through OLS in terms of direction and statistical significance, but they differ in size. In particular, coefficients associated with our variables of interest, i.e., *parl stab* and *parl unstab*, are now larger in magnitude than OLS ones, thus indicating that the OLS estimates are likely to be downward biased, a result already stressed by Persson and Tabellini in the comparison between presidential and parliamentary

¹⁷ To further ensure that we do not have issues of reverse causality between government size and the degree of conflict, we run Eq. (2) including *confl mean* as additional regressor. Coefficients associated with *confl mean* are far from being significant at the conventional level, both when the dependent variable is central government expenditure and central government revenues (p-values are 0.966 and 0.803, respectively). Results are available upon request.

¹⁸ Results from the instrumenting equation are reported in Appendix A, Table 5.

systems. When we perform the IV strategy using other approaches (Table 6, Appendix A), results do not qualitatively change.

Therefore, we can conclude that the IV analysis, regardless of the econometric strategy adopted, corroborates our modelling choice and further confirms that taking into account the level of stability of parliamentary systems provides novel and interesting insights into the existing literature on the comparison of constitutional systems.

4.2 Robustness checks and possible transmission channels

The empirical results presented above show that parliamentary systems perform differently depending on their level of stability. However, this interpretation can be questioned in several ways. First of all, is stability what we really capture with our measure? Moreover, does stability matter only for parliamentary systems, as our theoretical considerations suggest? Finally, are there other possible interpretations for the result?

This section addresses these three lines of concerns. First, we test the robustness of our results to different measures of stability, to support our claim that stability is what matters. Second, we run a placebo test refining also the classification of presidential systems according to their stability level. Third, we introduce additional controls to take into account alternative effects that may be relevant, namely the possibility that distribution of leftist governments, large electoral districts or income inequality may drive variations in government spending.¹⁹ At the end of the section, we also discuss our results in light of the theoretical models, and investigate possible transmission channels.

Different measures of stability. So far, in the analysis, countries were partitioned according to the index *gov life*. As discussed in Sect. 3, *gov life* measures stability as the average duration of governments, normalized by the legal duration of a legislature. As a first robustness check, we avoid the normalization and include the legal duration of the legislature as a control variable to make the relationship between (in)stability and the outcome variable more flexible.

A different approach to stability could define stable countries as those in which governments reach the natural end of the legislature, while defining all others as equally unstable. In other words, stability could have a non-linear effect. For this reason, we introduce an alternative index, *gov end*, defined as the fraction of governments that are successful in reaching the legal term of the mandate. The index is built using the indicator *yrcurnt* from the DPI dataset. Higher values of the index correspond to higher stability.

Another possible concern is that *gov life* measures the life of a government between two elections. However, changes in the executive do not map one to one to election years. There is the question of how to classify countries in which the

¹⁹ Additional robustness checks on the nature of the spending (dependent) variables that we choose are included in Appendix B.

Table 4 Constitutions, Central Government Expenditure and Central Government Revenues. IV estimations. BCIM dataset. ML estimation approach

Dep. var.	cgexp (1)	cgrev (2)
parl stab	7.674*** (2.496)	7.263*** (2.754)
parl unstab	1.203 (2.385)	0.119 (2.219)
maj	-7.140*** (1.823)	-4.504** (1.855)
gastil	-2.434** (1.294)	-1.689 (1.406)
age	-2.468 (4.639)	-0.590 (4.222)
lyp	-2.571 (2.212)	0.849 (2.054)
lpop	1.825* (0.951)	1.620** (0.120)
trade	0.067** (0.028)	0.091*** (0.020)
prop1564	-0.301 (0.286)	-0.280 (0.283)
pProp65	1.330*** (0.367)	1.260*** (0.438)
federal	-5.460* (2.815)	-5.245** (2.247)
oecd	-1.342 (3.971)	-1.354 (3.487)
africa	-1.027 (3.440)	3.783 (4.737)
asiae	-5.998** (2.729)	-3.326 (3.360)
laam	-4.295 (3.064)	-2.012 (3.306)
col_espa	5.781 (4.972)	6.859 (4.784)
col_otha	-4.379* (2.413)	-4.925 (3.064)
col_uka	7.476*** (3.037)	5.978** (2.725)
Observations	86	82
Dataset	BCIM	BCIM

White heteroskedasticity-consistent standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

executive changes frequently, but elections are held at regular intervals. To tackle this issue we introduce a third measure of stability, *year exec*, defined as the average tenure of the head of the executive, weighted by the legal length of any electoral term. This index is built using the indicator *yearoff* from the DPI dataset, which collects information about the number of years the head of the executive has been in office. Higher values of the index correspond to higher stability.

Moreover, parliamentary countries in which we have several changes of government, and early elections, but always the same party in power (such as Italy with Democrazia Cristiana from the 1950s to the 1970s) have a low index of stability when we use *gov life*. One could argue, however, that those countries are stable, as the leadership is strongly held by the same political group. To take this concern into account, we introduce a fourth index of stability, *year party*, defined as the average number of years the governing party has been in office weighted by the legal length of the electoral term. This index is built using the indicator *prtyin* from the DPI dataset, which counts the number of years the chief executive party has been in office. Higher values of the index correspond to higher stability.

Finally, an alternative way to measure (in)stability is the number of excess elections, computed as the number of observed elections minus the number of expected elections by law, in the period under investigation. This last indicator can be squared in order to further penalize very unstable parliamentary systems.

Using these alternative stability indexes, we perform our main regression, as in Eq. (2), on *cgexp* as dependent variable and, as Table 7 in Appendix A shows, our results do not qualitatively change.

Placebo test Table 8 in Appendix A shows the descriptive statistics of stability by constitutional group. We notice that the average stability is not significantly different between groups. This could raise the concern that we pick up the effect of stability in general, and not the effect of different equilibria in the legislative process of parliamentary systems. To address this issue, we run a placebo test by splitting presidential countries according to their stability, with the same methodology used for parliamentary countries in the main analysis. Table 9 in Appendix A shows that no effect is found when we adopt this classification, that is, stable presidential systems are not significantly different from unstable ones.²⁰

Additional controls One may wonder how our results may be affected by an uneven distribution of leftist governments or large electoral districts, which may inflate government spending (Milesi-Ferretti et al. 2002). To address this concern we estimate the model including the executive's ideological position (*right_left*), and the district magnitude (*magn*) as additional regressors.²¹ Table 10 in Appendix A shows that results do not change.

Moreover, we may be concerned by the effect of income inequality on government spending. The theoretical literature on the size of government (Meltzer

²⁰ Results do not change if we partition presidential countries according to the same threshold of stability that we use to partition parliamentary countries, instead of using the median level for presidential ones.

²¹ Source: DPI, 2012.

and Richard 1981) has shown that public expenditure may depend on the relation between the mean income and the income of the median (decisive) voter. Thus, failing to control for that may lead to an omitted variable bias. As a robustness check, we attempt to include information on the share of poor voters in our analysis. Given the lack of direct data, we consider two variables that together may capture the presence of a poor electorate at country level: (i) a measure of income inequality; (ii) an indicator of the average turnout. We believe that these two variables may efficiently account for the link between poor voters and the size of government. In fact, income inequality indicates the concentration of wealth and, as such, it is expected to be positively correlated with a feeling of injustice among the citizens, who may use the ballot box to ask for redistributive policies (Alesina and Angeletos 2005). Hence, income inequality captures the effect of the presence of a poor population on the size of government. Moreover, even if we do not have direct information on the share of poor people among voters, we include the average turnout at country level as additional variable, so as to control for differences among countries that may be driven by differences in the propensity to vote. Data on income inequality are taken from the Global Consumption and Income Project (GDIP) that provides comprehensive information about income inequality for a large number of countries, over the period 1960–2015. First, we restrict the sample to countries that are present in the BCIM dataset and to the period under investigation, that is, up to 1989. Then, for each country i , we extract from the GDIP dataset two measures of inequality: the GINI coefficient, i.e., cumulative proportions of population compared to cumulative proportions of income, and the P80/P20 ratio, that is, the ratio of income owned by the 20% richest to the 20% poorest. As for the turnout, we use the IDEA Voter Turnout Dataset that registers the turnout for each election in a wide number of countries, since 1945. Again, we restrict the dataset to match the same sample of countries and time span, i.e., 1960–1989, as for income inequality. In light of the cross-section nature of our dataset, we average the two variable over the considered period.

Table 11, in Appendix A, shows the results: columns (1) and (2) use the GINI coefficient as measure of income inequality, while columns (3) and (4) use the ratio P80/P20. We test this exercise on both the PT (columns 1 and 3) and BCIM (columns 2 and 4) dataset. Results corroborate theoretical intuitions: coefficients associated with income inequality are always positive, statistically significant and large in magnitude, thus indicating that there exists a positive correlation between inequality at country level and the size of government. Moreover, the turnout positively affects the size of government too, even though to a lesser extent. However, our main regressors, i.e., *parl stab* and *parl unstab*, retain their original relevance both in terms of statistical power and magnitude.

Transmission channels The overall empirical analysis is consistent with the theoretical intuitions stated in Sect. 2. In particular, in a model as in Cella et al. (2017), the performance of parliamentary systems closely approaches the performance of presidential ones if the confidence vote entails a better ability of legislators to reject bad policy proposals, thus creating a more unstable political environment. This mechanism drives stable parliamentary systems far from the performance of presidential ones, but also from the one of unstable parliamentary

systems. The same effect is consistent with the presence of the *legislative cohesion* effect in models as in Baron (1998), according to which the executive and the parliament coordinate to keep politicians in power and to avoid a no confidence motion. Coordination leads either the executive to formulate policy proposals that please the majority of the veto-players or legislators to accept a larger fraction of executive's proposals to avoid early elections. To sum up, the selection effect makes unstable parliamentary systems very similar to presidential ones and legislative cohesion makes stable parliamentary systems very different from presidential ones. We have shown how robust this difference is in explaining an array of implemented policies but our analysis does not allow us to separate the two effects just described.

The main findings described so far do not bring evidence in favour of the *disciplining effect* that would predict a similar performance by presidential and stable parliamentary systems, by favouring congruent behaviour in both the executive and the parliament. However, this may depend on the fact that we do not empirically isolate the subgroup of fully stable parliamentary countries that should display such effect. This leads us to suppose that the difference between presidential and parliamentary systems is not monotonically increasing in the level of stability of the latter. To test this insight, we further split parliamentary countries into four categories according to their stability distribution. Even if the analysis is sensitive to the small number of countries included in each category, we do find that the difference between constitutional systems is increasing in the stability of the parliamentary constitutional design, but it drops when we consider fully stable parliamentary countries that belong to the top quartile. Results are reported in Appendix A, Table 12.²²

5 Conclusions

Our analysis has been driven by the belief that institutional features shape the legislative process interacting with other underlying characteristics of the political environment. The presence of the confidence vote in parliamentary systems and its different effect across countries has suggested to classify parliamentary systems according to their degree of stability.

This finer partition of constitutional systems allows us to improve previous results in the literature that studied the effect of constitutions in terms of implemented policies. We show in fact that the difference in the performance of

²² Further evidence in support of the disciplining effect is that results of our main specification hold even when we make the threshold move along the stability distribution. In detail, results remain significant until stable parliamentary countries are in the 75th percentile of the stability distribution. After that, results are no longer significant. This is consistent with the disciplining effect according to which fully stable parliamentary countries should not be significantly different from any other constitutional category.

presidential and parliamentary systems is driven by stable parliamentary ones. Our results hold for many policy variables, including government expenditure and revenues, and are robust to changes in the data and the definition of stability.

To address concerns regarding the endogeneity of the constitutional choice, we perform an IV analysis using as instruments the time of adoption of the constitution as already done by previous works. To improve our identification strategy, we introduce a new instrument. We build an index that summarizes the degree of conflict in a country from 1800 up to the year of the constitutional choice. The idea is that deeply polarized societies may prefer to rely on power-sharing rules as those offered by parliamentary constitutions. We provide further evidence for this correlation and claim that the instrument is a robust predictor for our (possibly) endogenous regressor.

We also provide some novel insights on the transmission channels that may generate our empirical results. One of the reasons that may make unstable parliamentary systems more similar, in terms of policy choice, to presidential ones is the selection effect that improves the quality of the executive whenever the legislature does not reach the legal term limit. On the other side, our results may be driven by legislative cohesion that makes stable parliamentary systems more different from presidential ones. In other words, the coordination of the executive and the parliament to stay in office till the end of the term produces very different policy choices. We also find some mild evidence of a disciplining effect for the very stable parliamentary systems. In countries at the top of the stability distribution the executive is very disciplined by the threat of the confidence vote and the legislative process will lead to the same equilibria of presidential systems.

Funding Open access funding provided by Università degli Studi di Milano - Bicocca within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A: Additional tables

Table 5 IV: instrumenting equation. ML approach. BCIM dataset

Model	Multinomial (base=pres)	
	stable	unstable
dep-var	(1)	(2)
confl mean	0.65*** (0.12)	0.87*** (0.31)
con2150	0.64* (0.31)	0.68* (0.29)
cCon5180	0.79** (0.32)	0.77* (0.33)
con81	-0.97*** (0.33)	-0.03 (0.26)

White heteroskedasticity-consistent standard errors in parentheses. Columns (1) and (2) are estimated by multinomial probit. Entries are marginal effects. The following controls are included: *maj, age, gastil, lyp, lpop, trade, prop1564, prop65, federal, oecd, asiae, laam, africa, col_uka, col_espa, col_otha*
 * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6 Constitutions, Central Government Expenditure and Central Government Revenues. IV estimations. BCIM dataset. Estimators: 2SLS and H-FUL

Estimators	2SLS	2SLS	H-FUL	H-FUL
Dep. var.	cgexp	cgrev	cgexp	cgrev
	(1)	(2)	(3)	(4)
parl stab	7.398*** (2.310)	8.084*** (2.509)	6.751** (3.304)	7.288** (3.112)
parl unstab	3.889 (2.414)	0.291 (2.198)	3.295 (3.406)	3.112 (4.098)
maj	-8.129*** (1.931)	-5.062*** (1.840)	-4.751** (1.951)	-5.122** (2.712)
gastil	-2.867** (1.214)	-2.065* (1.096)	-0.271 (0.893)	-0.726 (0.819)
age	-4.615 (3.793)	-3.409 (3.289)	-4.696 (2.901)	-2.862 (2.227)
lyp	-1.189 (2.032)	1.660 (1.704)	-1.012 (2.001)	3.565** (1.091)
lpop	2.084*** (0.780)	2.269*** (0.729)	0.476 (0.610)	1.437 (0.878)

Table 6 (continued)

Estimators	2SLS	2SLS	H-FUL	H-FUL
Dep. var.	cgexp	cgrev	cgexp	cgrev
	(1)	(2)	(3)	(4)
trade	0.083*** (0.028)	0.110*** (0.020)	0.456** (0.019)	0.084*** (0.014)
prop1564	-0.436 (0.295)	-0.405 (0.270)	-0.127 (0.227)	-0.510** (0.192)
prop65	1.362*** (0.354)	1.298*** (0.411)	0.867** (0.353)	1.301*** (0.339)
federal	-4.416* (2.385)	-4.625** (2.227)	-4.416** (1.801)	4.923** (1.856)
oecd	-1.498 (3.544)	-4.436 (2.975)	-2.212 (2.644)	-4.692* (2.766)
africa	1.172 (4.408)	5.111 (4.926)	4.780 (3.819)	6.730* (3.973)
asiae	-3.873 (3.184)	-3.346 (3.320)	-3.059 (3.227)	-0.988 (2.586)
laam	-4.431 (3.390)	-3.455 (3.381)	-4.622 (3.065)	-1.392 (2.288)
col_espa	2.079 (4.327)	1.874 (3.517)	2.211 (3.910)	5.073 (3.798)
col_otha	-2.292 (2.623)	-3.057 (2.837)	-3.012 (3.911)	-3.622 (3.011)
col_uka	8.179*** (3.004)	4.157** (2.049)	6.011* (3.142)	3.841* (2.033)
Observations	86	82	86	82
Adjusted R^2	0.71	0.72	0.61	0.64
Dataset	BCIM	BCIM	BCIM	BCIM
Instruments' diagnostic test:				
K-P rk Wald F-stat	10.634	10.854		
5% max IV relative bias	16.85	16.85		
10% max IV relative bias	10.27	10.27		
Under-identification test	31.2***	30.9***		
Over-identification test	2.461	4.568	0.734	0.812
Endog. test	6.544**	7.661**		

White heteroskedasticity-consistent standard errors in parentheses. Columns (3) and (4) report the CHNSW test of overidentifying restrictions (see Chao et al. 2014), which is robust to many instruments and heteroskedasticity

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7 Robustness checks with different stability indexes. OLS estimations

Dep. var.	cgexp (1)	cgexp (2)	cgexp (3)	cgexp (4)	cgexp (5)	cgexp (6)	cgexp (7)
Dataset	PT						
parl stab	6.298*** (1.93)	5.932** (2.62)	4.816* (2.78)	6.064** (2.03)	6.959*** (2.25)	4.596** (2.18)	3.615* (2.07)
parl unstab	1.383 (1.87)	-0.316 (2.57)	2.560 (1.63)	1.852 (1.94)	3.846 (2.69)	1.573 (2.26)	2.995 (2.35)
legal_length	2.411*** (0.82)						
Dataset	BCIM						
parl stab	5.206** (2.11)	4.510** (2.08)	3.656* (1.98)	4.590** (2.17)	5.275** (2.55)	4.864** (2.01)	3.927* (2.03)
parl unstab	1.734 (1.98)	-1.700 (1.81)	2.522 (2.07)	2.827 (1.50)	-0.879 (3.38)	1.427 (2.14)	0.452 (2.24)
legal_length	2.062** (0.85)						

White heteroskedasticity-consistent standard errors in parentheses. A detailed explanation of the way in which the stability indexes have been assembled is reported in Sect. 4.2. Columns (1), (2), (3) and (4) report the stability indexes *gov life*, *gov end*, *year exec*, *year party*, respectively. Column (5) includes the legal length of each electoral cycle as a control, instead of using it to normalize the stability indicator. Here, the stability index is simply the average length of any electoral cycle computed for each country *i*. Column (6) considers the number of excess elections as measure of stability, while column (7) accounts for the squared number of excess elections. All the regressions include the following controls: *maj*, *age*, *ltp*, *trade*, *prop1564*, *prop65*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_espa*, *col_otha*
 * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8 Constitution and stability

	Mean (1)	SD (2)	(3)
pres	0.55	0.35	
parl	0.49	0.34	
p(pres,parl)			0.41

Entries in columns (1) and (2) are mean values and standard deviations for the stability index *gov life*. The dataset is the BCIM extended one. Column (3) reports the p-value of the *t-test* [p(pres,parl)] on the probability of falsely rejecting equal means across groups, under the assumption of equal variances. We cannot reject the hypothesis that stability does not vary according to the constitution type

Table 9 Constitutions and size of Government. OLS estimations

Specification	PT	BCIM	PT	BCIM
Dep. var.	cgexp	cgexp	cgrev	cgrev
	(1)	(2)	(3)	(4)
pres stab	-4.431** (2.29)	-3.377 (2.59)	-5.001** (2.36)	-2.744 (2.85)
pres unstab	-5.641*** (2.24)	-4.006 (2.73)	-5.061** (2.42)	-2.767 (2.39)
F test	0.25	0.07	0.00	0.00
Observations	80	91	76	91
Adjusted R^2	0.667	0.707	0.666	0.701

White heteroskedasticity-consistent standard errors in parentheses. All the regressions include the following controls: *maj*, *age*, *lyp*, *trade*, *prop1564*, *prop65*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_esp*, *col_otha*. *F test* refers to the hypothesis that the coefficients for *pres stab* and *pres unstab* are equal

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10 Constitutions, Central Government Expenditure, District Magnitude and Ideology. OLS Estimations. PT and BCIM dataset

Dataset	PT		BCIM	
	cgexp	cgexp	cgexp	cgexp
Dep. Var.	(1)	(2)	(3)	(4)
pres	-3.327 (1.54)		-2.641 (2.05)	
parl stab		5.543** (2.26)		4.502** (2.30)
parl unstab		-0.0559 (2.03)		1.484 (1.64)
right_left	1.266 (1.94)	0.970 (1.72)	1.968 (1.39)	2.055 (1.48)
magn	-0.228 (0.24)	-0.271 (0.24)	-0.333 (0.27)	-0.354 (0.27)
Observations	75	75	85	85
Adjusted R^2	0.671	0.692	0.628	0.645

White heteroskedasticity-consistent standard errors in parentheses. The regressions include: *maj*, *age*, *lyp*, *trade*, *prop1564*, *prop65*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_esp*, *col_otha*. The addition controls *right_left* and *magn* are included. *right_left* reports the average ideological position of the executive from 1970 to 1989. Values are between 1—right-oriented executive—to 3—left-oriented executive. *magn* represents the district magnitude weighted by the country's population. Source: DPI dataset

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11 Constitutions and Central Government Expenditure. OLS estimations. PT and BCIM dataset

Dep. Var.	cgexp (1)	cgexp (2)	cgexp (3)	cgexp (4)
parl stab	6.999*** (2.343)	4.845** (2.212)	6.653*** (2.299)	4.673** (2.125)
parl unstab	0.786 (1.931)	0.387 (1.821)	0.803 (1.611)	0.856 (1.709)
inequality	3.989** (1.837)	4.507*** (1.932)	2.775** (1.058)	3.199*** (1.190)
turnout	0.159** (0.075)	0.224** (0.086)	0.185** (0.075)	0.247*** (0.088)
Dataset	PT	BCIM	PT	BCIM
Observations	77	86	77	86
Adjusted R ²	0.80	0.76	0.81	0.77

White heteroskedasticity-consistent standard errors in parentheses. The regressions include the following controls: *maj*, *age*, *lyp*, *trade*, *prop1564*, *prop65*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_espa*, *col_otha*. Inequality in columns (1) and (2) is measured through the GINI coefficient; in columns (3) and (4) through the ratio P80/P20

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12 Constitutions and Central Government Expenditure. OLS estimations. BCIM dataset

Dep. var.	cgexp (1)	cgexp (2)
parl stab_1	2.67 (3.01)	1.96 (2.93)
parl stab_2	4.91** (2.60)	2.24 (2.83)
parl stab_3	3.53 (3.01)	5.31** (2.99)
parl stab_4		3.26 (3.33)
Observations	91	87
Adjusted R ²	0.67	0.71

White heteroskedasticity-consistent standard errors in parentheses. Parliamentary countries are split into 3 categories (columns (1)) and 4 categories (columns (2)) according to the index *gov life*. Note that in this specification *parl stab(0)=pres*. The regressions include the following controls: *maj*, *age*, *lyp*, *trade*, *prop1564*, *prop65*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_espa*, *col_otha*

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix B: Additional robustness checks

In addition to the robustness checks described in Sect. 4.2, we also test the model on dependent variables different from those analysed in the main text: central government expenditure on social services (*ssw*) for the period 1990–1998, log of the output per worker (*logyl*) in 2000, perception of corruption (*cpi*) for the period 2000–2005, social protection expenditure as percentage of the GDP (*socprot*) for the period 1990–2012, expenditure on education as percentage of the GDP (*edspend*) for the period 1995–2012, and general government expenditure as percentage of the GDP (*ggexp*) for the period 2000–2015.²³ We also modify the interval in which we compute some of the dependent variables so as to introduce additional sources of variation in the dataset. When dependent variables are computed starting from years more recent than the original ones, then the index *gov life* ranges from 1975 to the year preceding the one of the dependent variable.

We run regressions (2) and (3) on both the original and enlarged datasets whenever possible. The only exception is the regression with *ssw* as dependent variable, as *ssw* is only available for a subgroup of countries even in the PT dataset. The change in dependent variables does not alter our results. Moreover, when estimating the effect of the constitutional design on the share of social welfare spending, Persson and Tabellini (2003) slightly modify the original specification by dropping three control variables—*lpop*, *pro1564* and *trade*. Our specification remains significant even when all controls are included. Results are reported in Tables 13, 14, 15, 16, 17, 18 below.

Table 13 Constitutions and Social Welfare Expenditure. OLS estimations. PT dataset

Specification	PT-modified	BCIM-modified	PT	BCIM
Dep. var.	<i>ssw</i>	<i>ssw</i>	<i>ssw</i>	<i>ssw</i>
	(1)	(2)	(3)	(4)
<i>pres</i>	-2.244** (2.03)		-1.927 (1.69)	
<i>parl stab</i>		2.707** (2.03)		3.366*** (2.25)
<i>parl unstab</i>		-0.543 (1.48)		-0.487 (2.46)
F test		5.79**		6.83**
Observations	69	69	69	69
Adjusted R^2	0.632	0.699	0.523	0.683

Notes: White heteroskedasticity-consistent standard errors in parentheses. PT-modified and BCIM-modified refer to Persson and Tabellini (2003) specification of the model where the authors include all the standard controls—*maj*, *age*, *lyp*, *prop65*, *gastil*, *federal*, *oecd*, *africa*, *asiae*, *laam*, *col_uka*, *col_espa*, *col_otha*—except that *lpop*, *prop1564* and *trade* are missing. Then, we re-estimate the model using the same specification as in the previous table. *F test* (columns (2), (4)) refers to the hypothesis that the coefficients for *parl stab* and *parl unstab* are equal ($\beta_1 = \beta_2$)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

²³ Sources: IMF/GFS Yearbook.

Table 14 Constitutions and output per worker. OLS estimations

Dataset	PT		BCIM	
	logyl	logyl	logyl	logyl
Dep. var.	(1)	(2)	(3)	(4)
pres	-0.294*		-0.157	
	(1.84)		(1.01)	
parl stab		0.325*		0.392**
		(1.78)		(2.04)
parl unstab		0.115		-0.0364
		(1.55)		(2.20)
Observations	74	73	84	83
Adjusted R^2	0.731	0.695	0.753	0.721

White heteroskedasticity-consistent standard errors in parentheses. *logyl* is the productivity level as in Persson and Tabellini (2003) (Columns (1) and (2)), and in Blume et al. (2009) (Columns (3) and (4)). The regressions include the following controls: *maj*, *age*, *lyp*, *trade*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_esp*, *col_otha*, *avelf*, *prot80*, *catho80*, *confu*

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 15 Constitutions and Perception of corruption. OLS estimations

Dataset	PT		BCIM	
	cpi	cpi	cpi	cpi
Dep. var.	(1)	(2)	(3)	(4)
pres	-0.620*		-0.326	
	(1.76)		(-1.05)	
parl stab		0.627*		0.559*
		(1.80)		(1.72)
parl unstab		0.491		0.362
		(1.40)		(1.19)
avelf	1.274**	1.567**	0.987*	1.432**
	(2.09)	(2.42)	(1.83)	(2.49)
Observations	78	78	88	88
Adjusted R^2	0.829	0.833	0.806	0.820

White heteroskedasticity-consistent standard errors in parentheses. *cpi* is the perception of corruption as in Persson and Tabellini (2003) (Columns (1) and (2)), and in Blume et al. (2009) (Columns (3) and (4)). The regressions include: *maj*, *age*, *lyp*, *trade*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_esp*, *col_otha*, *avelf*, *prot80*, *catho80*, *confu*. The additional control *avelf* is included and reported in the table. *avelf* is the index of ethnolinguistic fractionalization, as in La Porta et al. (1998)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 16 Constitutions and social protection. OLS estimations

Dataset	PT		BCIM	
	socprot	socprot	socprot	socprot
Dep. var.	(1)	(2)	(3)	(4)
pres	-2.184 (2.10)		-2.028 (2.01)	
parl stab		2.924** (2.05)		2.149** (1.09)
parl unstab		-0.346 (2.16)		0.270 (1.71)
Observations	76	76	86	86
Adjusted R^2	0.758	0.761	0.714	0.729

Notes: White heteroskedasticity-consistent standard errors in parentheses. *socprot* is the central government social protection expenditure as defined by the IMF-GFS dataset (averaged over years 1990–2012). The regressions include the following controls: *maj*, *age*, *lyp*, *trade*, *prop1564*, *prop65*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_espa*, *col_otha*

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 17 Constitutions and education expenditure. OLS estimations

Dataset	PT		BCIM	
	edspend	edspend	edspend	edspend
Dep. var.	(1)	(2)	(3)	(4)
pres	-0.413 (1.10)		-0.512 (1.16)	
parl stab		0.832** (0.65)		1.10** (0.89)
parl unstab		0.346 (1.16)		0.270 (1.71)
Observations	75	75	84	84
Adjusted R^2	0.508	0.521	0.514	0.529

Notes: White heteroskedasticity-consistent standard errors in parentheses. *edspend* is the central government expenditure on education as percentage of the GDP as defined by the IMF-GFS (averaged over years 1995–2012). The regressions include the following controls: *maj*, *age*, *lyp*, *trade*, *prop1564*, *prop65*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_espa*, *col_otha*

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 18 Constitutions and General Government Expenditure. OLS estimations

Dataset	PT		BCIM	
	ggexp	ggexp	ggexp	ggexp
Dep. var.	(1)	(2)	(3)	(4)
pres	−5.236*** (1.52)		−3.572 (2.69)	
parl stab		6.386*** (3.24)		4.751** (1.87)
parl unstab		1.403 (1.62)		0.363 (1.37)
Observations	80	81	101	97
Adjusted R^2	0.707	0.818	0.781	0.796

White heteroskedasticity-consistent standard errors in parentheses. *ggexp* is the general government expenditure, averaged over the period 2000–2015. The regressions include the following controls: *maj*, *age*, *lyp*, *trade*, *gastil*, *federal*, *oecd*, *lpop*, *africa*, *asiae*, *laam*, *col_uka*, *col_espa*, *col_otha*

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

References

- Acemoglu, D. (2005). Constitutions, politics, and economics: A review essay on persson and Tabellini's the economic effects of constitutions. *Journal of Economic Literature*, 43(4), 1025–1048.
- Alesina, A., & Angeletos, G.-M. (2005). Fairness and redistribution. *American Economic Review*, 95(4), 960–980.
- Alesina, A., Ozler, S., Roubini, N., & Swagel, P. (1996). Political instability and economic growth. *Journal of Economic Growth*, 1(2), 189–211.
- Anatolyev, S., & Skolkova, A. (2019). “Many instruments: Implementation in Stata”, *The Stata Journal: Promoting communications on statistics and Stata*, Vol. 19, N. 4.
- Baron, P.D. (1998). “Comparative Dynamics of Parliamentary Governments”, *American Political Science Review*, Vol.92, 3(09/1998): 593–609.
- Baum, C. F. (2014). Conditional mixed-process models, Boston College. <http://fmwww.bc.edu/EC-C/S2016/8823/ECON8823.S2016.nn14.slides.pdf>.
- Blume, L., Müeller, J., Voigt, S., & Wolf, C. (2009). The economic effects of constitutions: Replicating—and extending—Persson and Tabellini. *Public Choice*, 139, 197–225.
- Bound, J., Jaeger, A. D., & Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90(430), 443–450.
- Cella, M., Iannantuoni, G., & Manzoni, E. (2017). “Do the right thing. A Comparison of Politicians' Incentives Across Constitutional Systems”. *Economica*, 84, 430–453.
- Diermeier, D., & Feddersen, T. J. (1998a). Comparing constitutions: Cohesion and distribution in legislatures. *European Economic Review*, 42, 665–672.
- Diermeier, D., & Feddersen, T. J. (1998b). Cohesion in legislatures and the Vote of Confidence Procedure. *American Political Science Review*, 92(3), 611–621.
- Diermeier, D., & Vlaicu, R. (2011). Executive control and legislative success. *Review of Economic Studies*, 78, 846–871.
- Fuller, W. A. (1977). Some properties of a modification of the limited information estimator. *Econometrica*, 45, 939–954.
- Gurr, T. R. (1975). Persistence and change in political systems: 1800–1971. *American Political Science Review*, 68, 1482–1504.

- Hall, R. E., & Jones, C. I. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, *114*(1), 83–116.
- Hausman, J. A., Newey, W. K., Woutersen, T., Chao, J. C., & Swanson, N. R. (2012). Intrumental variable estimation with heteroskedasticity and many instruments. *Quantitative Economics*, *3*, 211–255.
- Huber, J. D. (1996). The vote of confidence in parliamentary democracies. *American Political Science Review*, *90*(2), 269–282.
- Huber, J. D., & Gallardo, C. M. (2008). Replacing cabinet ministers: Patterns of ministerial stability in parliamentary democracies. *American Political Science Review*, *102*(2), 169–180.
- Jung, J. K., & Deering, C. J. (2015). Constitutional choices: Uncertainty and institutional design in democratising nations. *International Political Science Review*, *36*(1), 60–77.
- Lijphart, A. (1999). “Patterns of Democracy. Government Forms and Performance in Thirty-Six Countries”, *New Haven, CT: Yale University Press*.
- Lijphart, A. (2004). Constitutional design for divided societies. *Journal of Democracy*, *15*(2), 96–109.
- Linz, J.J. (1994). “Presidential or parliamentary democracy: Does it make a difference?”, In J.J. Linz *The Failure of Presidential Democracy*, Johns Hopkins University Press.
- Milesi-Ferretti, G. M., Perotti, R., & Rostagno, M. (2002). Electoral systems and the composition of public spending. *Quarterly Journal of Economics*, *117*, 609–657.
- Perez, J. I. G., & Sanz, Y. R. (2005). Wage changes through job mobility in Europe: A multinomial endogenous switching approach. *Labour Economics*, *12*(4), 531–555.
- Persson, T. & Tabellini, G. (2004). “The economic effects of constitutions”, *American Economic Review*, Vol. *94*, N.1.
- Persson, T. (2002). Do political institutions shape economic policy? *Econometrica*, *70*(3), 883–905.
- Persson, T., & Tabellini, G. (2003). *The economic effects of constitutions*. Cambridge: The MIT Press.
- Petreski, M., Blazevski, N. M., & Petreski, B. (2014). Gender Wage gap when women are highly inactive: Evidence from repeated imputations with macedonian data. *Journal of Labour Research*, *35*, 393–411.
- Rivers, D., & Vuong, Q. H. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics*, *39*, 347–366.
- Rockey, J. (2012). Reconsidering the Fiscal effects of constitutions. *European Journal of Political Economy*, *28*, 313–323.
- Roodman, D. (2011). Fitting fully observed recursive mixed-process models with cmp. *The Stata Journal*, *11*(2), 159–206.
- Sarkees, M. R. (2010). “Defining and categorizing wars”, In *Resort to war: A data guide to inter-state, extra-state, intra-state, and non-state wars, 1816–2007*, by Meredith Reid Sarkees and Frank Whelon Wayman, 39–73. Washington, DC: CQ Press.
- Suarez-Varela, M., Martinez-Espineira, R., & Gonzalez-Gomez, F. (2015). An analysis of the price escalation of non-linear water tariffs for domestic uses in Spain. *Utilities Policy*, *34*, 82–93.
- Voigt, S. (2011). Empirical constitutional economics: Onward and upward? *Journal of Economic Behavior and Organization*, *80*, 319–330.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge: MIT Press.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.