University of Leicester

DEPARTMENT OF ECONOMICS

Which Democracies Pay Higher Wages?*

Miltiadis Makris, University of Southampton, UK
James Rockey, University of Leicester, UK

Working Paper No. 11/09
November 2010

Which Democracies Pay Higher Wages?*

Miltiadis Makris

Department of Economics
University of Southampton

James Rockey

Department of Economics
University of Leicester

August 2010

Abstract

The labor share of income varies markedly across the set of democracies. A model of the political process, situated in a simple macroeconomic environment is analyzed in which the cause of this variation is linked to differences in the form of democracy - in particular the adoption of a presidential or parliamentary system. Presidential regimes are associated with lower taxation but lower wages. Robust evidence for the negative impact of a presidential system on the labor share is obtained using a Bayesian Model Averaging approach. Evidence is also provided that this is due to lower taxation.

*We are grateful to seminar participants at Leicester and the 2^{nd} European Workshop in Political Economy, Dresden. Rockey thanks Jonathan Temple for invaluable guidance and advice. The usual disclaimer applies. Email: Makris-m.makris@soton.ac.uk, Rockey-jcr12@le.ac.uk

1 Introduction

Economic policy in a democracy is driven by the demands of voters and macroeconomic conditions, which are mediated through the political process. Each archetype of democracy reconciles these two forces in different ways, potentially leading to variation in a range of societal outcomes. The focus of this paper is to analyze one such outcome: the labor share of income. This paper considers two broad forms of democracy, Presidential and Parliamentary. As such it builds on the key work of Persson, Roland and Tabellini (2000). By situating a model of these different political processes within a simple macroeconomic framework, this paper argues that as well being associated with lower levels of taxation (and government expenditure), presidential democracies should be expected to have, *ceteris paribus*, a lower wage level. The second part of the paper provides empirical evidence for this claim. It employs a Bayesian Model Averaging appproach to identify the causal effect associated with having a presidential democratic system despite many potential confounding variables. The evidence obtained using this methodology coincides with the prediction of the model: the labor share of income is around 12 percentage points lower in presidential democracies. Dynamic panel data results provide evidence that this variation is due to the mechanism suggested by the model: variation in the level of taxation.

This paper draws on our emerging understanding of how institutions determine societal outcomes. Of particular importance is the work of Persson, Roland and Tabellini (2000). They argue that parliamentary democracies lead to greater redistribution, greater rents for politicians and higher levels of public good provision. A corollary of this is that taxation is lower in presidential democracies. No survey of their more general contribution is attempted here, as the model presented in the next section is a simple variation of their model, and as such provides a more detailed discussion.

Related empirical work includes that of Persson and Tabellini (2003) and Persson and Tabellini (2004) who estimate the effect of presidential democracy and find that it is associated with a six percentage points smaller government share of GDP.[1] Acemoglu (2005) critiqued the methodology employed, arguing that the majority of the explanatory power

[1] These estimates are for the early 1990s

of the instrumental variables used in the first stage was due to variables unable to predict differences in constitutional type.[2] This criticism, in part, motivates the use in this paper of an alternative methodology.

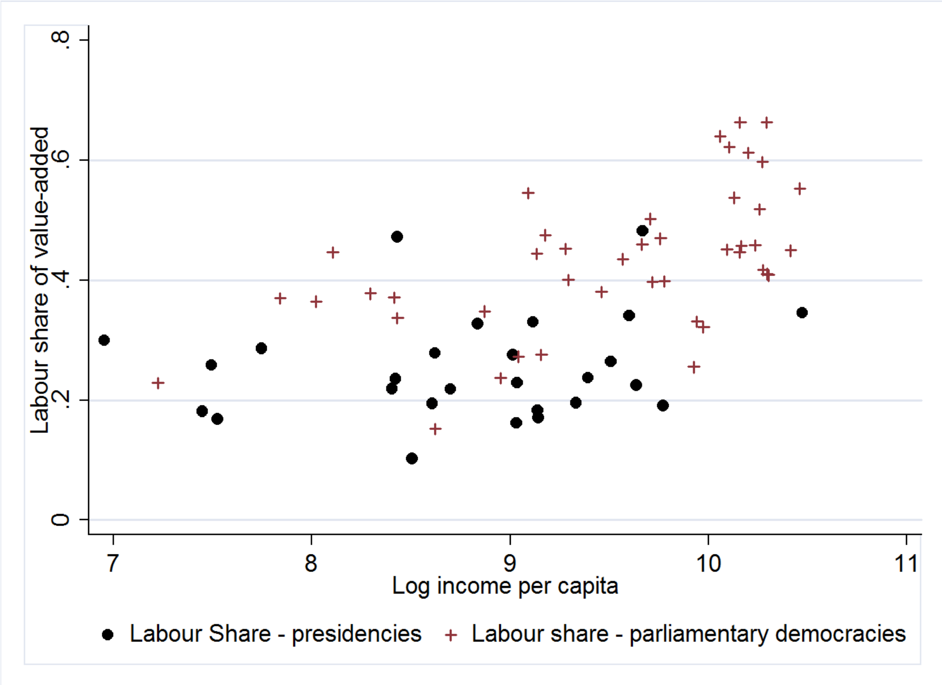
Also important for this paper is the work of Rodrik (1999) who argues that 'Democracies pay higher wages'. In particular his results are that the labor share of income in manufacturing, conditional on income per capita, is higher in democracies than non-democracies. He suggests that this may be because the (Nash) bargaining power of workers is greater in a democracy, due to their greater political and economic freedoms. No argument is advanced in this paper that such rights vary between types of democracy. The claim is that different democratic systems, and in particular presidential democracies, lead, on average, to different policies given similar societal preferences. As such, this paper builds on Rodrik (1999)'s idea that how a country is governed can impact upon factor shares. However, the mechanism that drives difference between autocracies and democracies is very different. This paper shows that the labor share depends on taxation, which in democracies, depends on the legislative process.

As a first step, Figure 1 shows the relationship between income per capita and data on labor's share of value added in manufacturing from the UNIDO (2005) database. Perhaps most notable is the degree of variation in the labor share, from around 10 percent to about 70 percent. It would also seem on casual inspection that workers in richer democracies receive a larger share of income. Perhaps more readily apparent is that the labor share seems higher in parliamentary democracies. This will be confirmed by results obtained from the Bayesian Model Averaging (henceforth, BMA) analysis presented in section 4.3 Specifically, these results suggest that presidential democracies are associated with a labor share 12 percentage points lower than in equivalent parliamentary regimes.

The hypothesis that variation in the labor share is related to the size of government is motivated, in part, by figure 1. This shows that whilst within the OECD the labor share in presidential and parliamentary democracies was initially quite similar, over time a sizeable discrepancy has emerged. As discussed by Boix (2001) and Pickering and Rockey (Forthcoming) the size of

[2] Rockey (2010) suggests that using an alternative set of instruments, and improved instrumental variable estimators, a quantitatively similar causal effect to that identified by Persson and Tabellini (2003) may still be associated with presidential democracy

Figure 1: Scatterplot of log income per capita and the labor share of value added



government has increased dramatically over the period, in large part due to increases in income per capita, and this increase is smaller in presidential democracies, as also suggested by Persson and Tabellini (2003). That the labor share has fallen markedly in OECD presidential democracies, and only slightly in OECD parliamentary democracies, does not conflict with this interpretation. It is easy to imagine that some secular trend, such as skill-biased technological change, has led to a decrease in the labor share, except in parliamentary democracies where this has been offset by more pronounced growth in the size of government. There is, at most, a small increase in the large difference between parliamentary and presidential regimes in the full sample. Both the difference and the increase are consistent with the model studied in the next section, where the labor share is increasing in the tax rate, which is itself higher throughout and grows more in both parliamentary democracies and where labor productivity is greater. Throughout the period, presidential democracies are poorer throughout the period, start with a lower tax rate, which grows at a lower rate. According to the model below, these difference would lead to such a differential, and its (small) increase. Moreover, the empirical work suggests that the difference in conditional means between parliamentary and presidential democracies remains large and statistically significant despite allowing for many possible confounding variables.

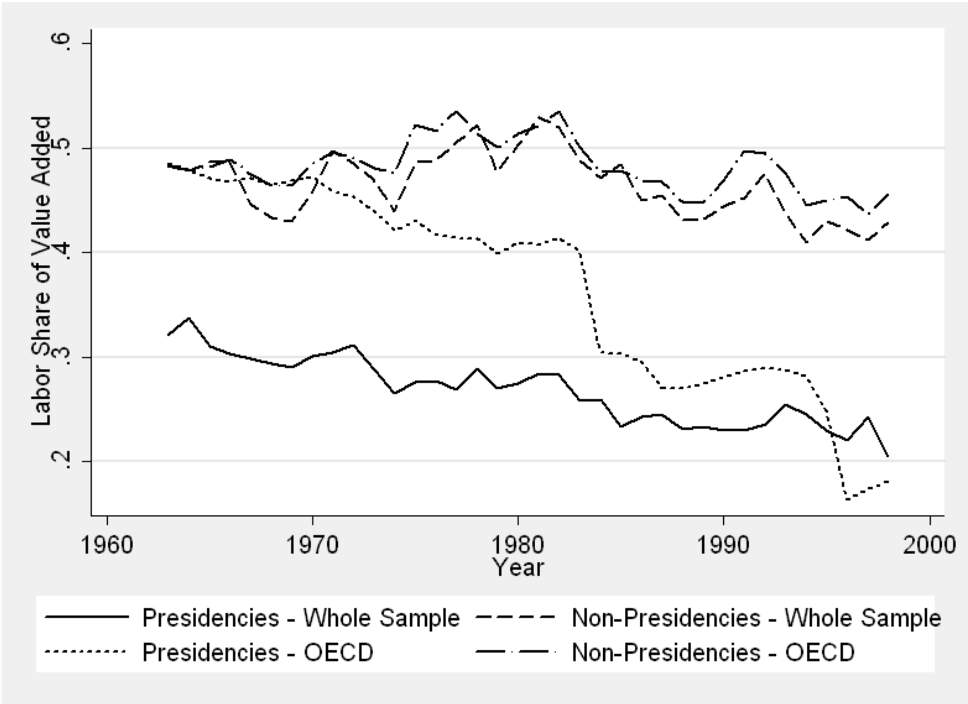


Figure 2: Trends in the average labor Share for Presidential and Parliamentary Democracies

Botero, Djankov, La Porta, Lopez-De-Silanes and Shleifer (2004) address a related question. They provide evidence that a key determinant of labor market regulation is the type of legal system; Common law, Napoleonic, etc. Moreover, whilst richer countries have more generous welfare systems, once legal origin is accounted for, the political power of the left has little impact on the regulation of labor. Section 4.3 provides evidence that whilst a legal system based on the Scandanavian system is associated with a greater labor share, this doesn't reduce the effect of presidential democracy, and other systems have little predictive power for the labor share of income.

However, the complex way in which the legal and democratic institutional causes of societal outcomes interact is highlighted by the work of Pagano and Volpin (2005). They argue that legal origin will have limited predictive power for the extent of investor and employment protection as in both civil and common law systems reforms of corporate governance have often been enacted. They employ a model which builds on that of Persson and Tabellini (1999), and suggest that a majoritarian electoral system tend to enact laws providing strong investor, and weak employee, protection. However, they do find that legal origin retains some additional explanatory power once the form of electoral system is controlled for. The model presented here focuses on a different aspect of constitutional variation, the separation of powers.The results presented in Section 4.3 suggest that unlike the separation of powers, variation in electoral systems has little explanatory power.

The remainder of this paper will proceed as follows. The next section presents the model. Section 3 discusses the equilibrium outcomes under different forms of democracy, and as such explains in more detail why the form of constitution might be expected to impact on the labor share of income. The next two sections describe the empirical analyses. Section 4 outlines the Bayesian Model Averaging approach used for the cross-sectional data. It also presents the results and argues they are robust to outliers and the inclusion of other variables. Section 5 provides evidence for the hypothesis that the observed variation in the labor share is due to differences in taxation, as suggested by the model. This is followed by a brief conclusion.

2 The Model

Our model is a simple extension of the Persson, Roland and Tabellini (2000) model (PRT hereafter) with endogenous labor and wages. Thus, in our case income taxation is distortionary and affects the remuneration of labor. This difference between the two models does not change (conditional on similar parameter restrictions) the predictions offered in PRT regarding the relationship bewteen political regimes and size of taxes. However, the endogeneity of labor and wages in our model enables us to establish also a relationship between political constittions and remuneration of labor - a relationship that will be taken onto the data in the second part of our paper. Given that the only difference between our model and that in PRT is the endogeneity of labor and wages, the description of the model will be brief. For more details, the interested readers could consult PRT.

Time periods are indexed by $t = 1, 2, ..., \infty$ There are 3 identical regions, indexed by $j = 1, 2, 3$, each with a representative citizen. There is a perfectly traded single good and an economy-wide public good $g_t \geq 0$ that has a per-unit cost, in terms of the single good, which is common knowledge and normalized to one.

Policies are chosen by an elected decision-making body, which consists of a representative from each region, according to the constitutional rules in place. These rules will be outlined shortly. Elections for the representatives take place at the end of each period, right after that period's output is determined and policy is implemented but before next period's economic activities commence. Assume partial policy commitment: elected policy-makers can commit only for the policies of next period (i.e. for when they are in office) but they can credibly announce once they are elected, i.e. at the beginning of next period.

Fiscal revenues are raised by means of an income tax rate τ_t across all regions. Denoting region-j citizen's income with y_t^j, total fiscal revenues are thus equal to $\tau_t \sum_{j=1}^{3} y_t^j$. The per-unit cost of public good is non-contractible. This enables the policy-makers to appropriate fiscal resources in the form of rents (claiming a higher per-unit cost). Denote total political rents with $s_t \geq 0$ and $s_t^j \geq 0$ the political rents appropriated by the representative of region j. Fiscal revenues finance the public good, political rents and region-specific (per-capita) transfers, $r_t^j \geq 0$. We turn to the political environment.

2.1 The Political Environment

In each region there are two exogenously given and identical politicians who compete for office. Thus, one of these politicians is the incumbent representative and the other is the challenger. Majority voting determines the region's representative from these two politicians. Voting for the election of the representative takes place according to a retrospective voting rule (for a discussion of retrospective voting, see, for instance, PRT). Thus, voters condition their vote on the past performance of their representative. The winner of the election in region j derives a benefit of W_t^j. This simply denotes the continuation payoff from being elected for period-$t+1$ office in the equilibria of the game. Let δ be the common discounting factor of voters and politicians. The period-t representative of region j has payoff $s_t^j + p_t^j \delta W_t^j$,where p_t^j is the probability of being re-elected for office in period $t+1$.

The timing of events is the following. Fix any period t. Given period-t economic decisions and policies, we have that the period$-t$ net income $y_t^j(1-\tau_t)$ of citizens is determined. Immediately after, at the end of period-t, majority-voting elections take place simultaneously and independently in every region to elect/choose their representatives. Elected representatives comprise the decision-making body - the legislature. Then, agenda-setters are chosen from the set of regional representatives according to the constitutional rules, outlined in the next paragraph. Given the identity of agenda-setters, voters in each region j set simultaneously and independently their cutoff voting rules to be used in the forthcoming elections. These rules are of the form:[3] "if political agreement over a policy with non-zero spending takes place and attained welfare is at least equal to ϖ_t^j then vote for the incumbent representative of my region (j); otherwise vote for the challenger". Given the performance standards of voters, $\{\varpi_t^1, \varpi_t^2, \varpi_t^3\}$, the legislature chooses (and announces) period-$t+1$ policy, $\tau_{t+1}, g_{t+1}, (s_{t+1}^j, r_{t+1}^j)_{j=1}^3$, according to the constitutional rules. Given the period-t net income and rationally anticipated period-$t+1$ policies and prices, period-$t+1$ private consumption and production decisions take place, markets clear and prices are determined. Subsequently, period-$t+1$ policies are implemented and period-$t+1$ net incomes are thus determined, and so on.

The three political constitutions, for the determination of policies, we investigate are the simple-legislature, the presidential system and the parliamentary system. For a discussion on the modeling of these constitutions see PRT. Following PRT, under the **simple-legislature system**, there is only one agenda-setter who proposes all policies. Representatives vote this proposal against the status-quo policy. If the agenda-setter secures a majority then the proposal becomes an actual policy; otherwise the status-quo policy is set: $\tau = \bar{\tau}$, $g = r^j = 0$, for any j, and total fiscal revenues are uniformly shared between representatives. Under the **presidential system**, there is separation of powers. Specifically, there is one agenda-setter responsible for the budget who chooses the tax, and one agenda-setter responsible for the spending allocation who chooses regional transfers/rents and the level of public good. Agenda-setting is as follows. First, the tax-setter proposes a tax. If she secures a majority, the proposal becomes the actual tax-policy; otherwise the status-quo tax is set, $\tau = \bar{\tau}$. Given the tax set at the budget-setting stage, the spending-setter proposes the allocation of the budget into regional transfers and public good and the sharing of political rents between representatives. If she secures a majority, the proposal becomes the actual spending-policy; otherwise the status-quo spending-policy is set, $g = r^j = 0$, for any j, and total fiscal revenues are uniformly shared between representatives. Under the **parliamentary constitution**,

[3] This rule is consistent with the one investigated in Persson, Roland and Tabellini (1997, 1998) and Persson and Tabellini (2000).

there is governmental discipline/cohesion. In more detail, now, there are two ministers (government partners). The senior partner (the prime minister) proposes a public finance policy. Next, the junior member of the government can veto the proposal. If there is no veto, then the proposal becomes actual policy (it is supported by the two ministers). If there is a veto then the government breaks down. When the government breaks down, the legislature becomes a "caretaking" simple-legislature. In this case, the outcome is the equilibrium outcome of a simple-legislature.[4] Moreover, in any case agenda-setters are randomly chosen. Finally, if a policy-setter is indifferent between the other two legislators, we assume that these legislators have the same probability of being included in the winning coalition.

Up to now the model follows that in PRT, where income is exogenously given (and equal to 1). In our case, however, it is endogenous. The economic environment is thus important for our purposes.

2.2 The Economic Environment

In each period t, the representative citizen in region j decides on her consumption c_t^j and labor supply l_t^j, having per-period preferences[5]

$$c_t^j + v[1 - l_t^j] + r_t^j + H[g_t],$$

where $v[0] = 0$, $v' > 0, v'' < 0, \lim_{\lambda \to 0} v'[\lambda] = \infty$, and constraints

$$0 \leq c_t^j = (w_t^j l_t^j + \pi_t^j)(1 - \tau_t),$$
$$0 \leq l_t^j \leq 1,$$

where w_t^j is the real wage rate and π_t^j are the real profits of the region's firm. She chooses optimal consumption and labor supplies to maximize her intertemporal expected discounted utility taking as given current and future prices and profits, and current and future policies.

[4] This version of Parliamentary system is according to Persson and Tabellini (2000). Results would not change if, as in Persson, Roland and Tabellini (1998), (a) nature only draws the spending-setter who, in turn, chooses the tax-setter, (b) the tax proposal is made before the spending proposal, and (c) either of the partners can veto the policy.

[5] Note that we will be using square brackets for functions and brackets for collected terms.

She also takes into account that future consumption and labor decisions will be taken optimally. It follows that consumers' optimal labor supply in every period t is given by

$$l_t^j = \max\{0, 1 - v'^{-1}[w_t^j(1-\tau_t)]\},$$

where v'^{-1} denotes the inverse of v'. Clearly, labor supply is increasing in the wage and decreasing the tax, if $v'^{-1}[w_t^j(1-\tau)] \leq 1$.

Assume that public good preferences are such that $H[0] = 0, H' > 0, H'' < 0$ and $H'[0] > 1$. The latter ensures that the Utilitarian level of public good is positive. Let $H'^{-1}[\frac{1}{3}] \equiv \hat{g}$ be the Samuelson-rule level of public good.

In each region, there is also an identical production technology represented by $\theta q[l_t^j]$, $q' > 0, q'' < 0$, $q[0] = 0$, where θ is a technology parameter. Assume the Inada conditions that $\lim_{l\to 0} q'[l] = \infty$ and $\lim_{l\to\infty} q'[l] = 0$. Assuming that labor is paid its marginal productivity, we have that

$$w_t^j = \theta q'[l_t^j],$$
$$\pi_t^j = \theta(q[l_t^j] - l_t^j q'[l_t^j]).$$

Clearly, then, labor decisions and wages will be the same across regions, due to the assumption that regions are identical. Moreover, observe that $\lim_{l\to 0} q'[l] = \infty$ and $\lim_{\lambda\to 0} v'[\lambda] = \infty$, in conjunction with consumer's optimal labor supply and the remuneration of labor, imply that in equilibrium labor must be positive. Therefore, dropping the superscript j hereafter whenever there is no risk of confusion, we have that the wage and labor supply in each and every region are given by the solution of $l_t = 1 - v'^{-1}[w_t(1-\tau_t)]$ and $w_t = \theta q'[l_t]$, which is denoted by

$$w_t = \theta q'[L[\tau_t, \theta]]$$

and

$$l_t = L[\tau_t, \theta].$$

Let also $w[\tau_t, \theta] \equiv q'[L[\tau_t, \theta]]$ and note that $1 - v'^{-1}[\theta w[\tau_t, \theta](1-\tau_t)] = L[\tau_t, \theta]$.

It follows directly that a **higher income tax increaseses the real wage**: in fact, we have $\partial L[\tau_t,\theta]/\partial\tau_t = \frac{\theta w[\tau_t,\theta]}{u''[L[\tau_t,\theta]]+q''[L[\tau_t,\theta]](1-\tau_t)\theta} < 0$ and hence $\partial w[\tau_t,\theta]/\partial\tau_t = \theta q''[L[\tau_t,\theta]]\{\partial L[\tau_t,\theta]/\partial\tau_t\} > 0$). Note also that $\partial L[\tau_t,\theta]/\partial\theta = -\frac{(1-\tau_t)w[\tau_t,\theta]}{u''[L[\tau_t,\theta]]+q''[L[\tau_t,\theta]](1-\tau_t)} > 0$.

Private income before tax in any region, y_t, equals the region's output: $w_t l_t + \pi_t = \theta q[l_t]$. It follows that the equilibrium fiscal revenues are

$$3\tau_t\theta q[L[\tau_t,\theta]] \equiv R[\tau_t,\theta].$$

Accordingly,

$$R[\tau_t,\theta] = g_t + \sum_{j=1}^{3} r^j + s.$$

Let also

$$\theta q[L[\tau_t,\theta]] \equiv y[\tau_t,\theta]$$

be the private income expressed as a function of the tax rate and productivity. Given $\partial L[\tau_t,\theta]/\partial\tau_t < 0$, we thus have that an increase in the tax reduces the tax base (i.e. income $y[\tau_t,\theta_t]$). Note also that

$$(\partial R[\tau_t,\theta]/\partial\tau_t)/3 = y[\tau_t,\theta] + \tau_t\theta w[\tau_t,\theta](\partial L[\tau_t,\theta]/\partial\tau_t),$$

which is **independent of past outcomes**. Furthermore, after some straightforward differentiation, we have $\partial R[\tau_t,\theta]/\partial\theta = 3\tau_t(\partial y[\tau_t,\theta]/\partial\theta) > 0$.

Assume that $\partial^2 R[\tau_t,\theta]/\partial\tau_t^2 < 0$ and $\lim_{\tau\to 1}\partial R[\tau,\theta]/\partial\tau_t < 0$. Noting that $y[0,\theta] > 0$ and hence $\partial R[0,\theta]/\partial\tau_t > 0$, we therefore have that total revenues have a unique maximum for any given θ. Denote $\hat{s}[\theta] > 0$ this maximal tax revenues and $\hat{\tau}[\theta] > 0$ the corresponding (revenue maximizing) tax rate (in PRT these were exogenously given to be equal to 3 and 1, respectively). Let now $\bar{s}[\theta] \equiv R[\bar{\tau},\theta]$ and, echoing similar assumption in PRT, focus on the case where $\bar{s}[\theta] < \hat{s}[\theta]$ for any θ. We will often refer to $\hat{s}[\theta]$ as maximum possible rents and to $\hat{\tau}[\theta]$ as revenue-maximizing or Leviathan tax. In our setup, policy-makers would never choose a tax higher than the Leviathan tax as this would lead to lower public good and/or transfers (and hence lower welfare of voters) and/or political rents than the revenue-maximizing tax. We therefore assume hereafter, without loss of generality, that admissible taxes are $\tau \in [0, \hat{\tau}[\theta]]$ for given θ. Clearly, for any such tax we have that total tax revenues are **increasing** in the tax.

We finish this section with the consumers' per-period value functions. Let

$$V[\tau_t, \theta] \equiv y[\tau_t, \theta](1 - \tau_t) + v[1 - L[\tau_t, \theta]].$$

The per-period value function of the citizen in the typical region is then given by

$$V[\tau_t, \theta] + r_t^j + H[g_t].$$

Note, after using the envelope theorem, that:

$$\partial V[\tau_t, \theta]/\partial \tau_t = -y[\tau_t, \theta],$$

which is **independent of past outcomes** and **negative**. Let us focus hereafter on the case where, $V[\tau; \theta] + H[R[\tau, \theta] - T]$ is strictly concave with respect to τ for any $\tau \geq 0$ and $T \in [0, R[\tau, \theta]]$. This ensures that the Utilitarian policy is well-defined. It also implies that $\frac{\partial V[\tau,\theta]}{\partial \tau} + H'[R[\tau] - T]\frac{\partial R[\tau,\theta]}{\partial \tau}$, $\frac{\partial V[\tau,\theta]}{\partial \tau} + \frac{\partial R[\tau,\theta]}{\partial \tau}$ and $2\frac{\partial V[\tau,\theta]}{\partial \tau} + \frac{\partial R[\tau,\theta]}{\partial \tau}$ are[6] strictly decreasing with respect to τ for any $\tau \geq 0$ and $T \in [0, R[\tau, \theta]$. Finally, assume that $H'[0]\frac{\partial R[0,\theta]}{\partial \tau} > -\frac{\partial V[0,\theta]}{\partial \tau}$. This ensures that the Utilitarian tax is positive.[7]

To complete the model, it remains to define the type of equilibria that we characterize in Section 3.

[6]For the second one needs to set $T = R[\tau, \theta] - H'^{-1}[1]$ and for the third $T = R[\tau, \theta] - H'^{-1}[\frac{1}{2}]$.

[7]At an interior solution, the Utilitarian optimum is given by zero, as expected, political rents, zero regional transfers and tax τ^o such that

$$H'[R[\tau^o, \theta]] = -\frac{\partial V[\tau^o, \theta]/\partial \tau}{\partial R[\tau^o, \theta]/\partial \tau}.$$

2.3 Equilibrium Concept

We focus on sequentially rational equilibria in symmetric (pure) Markov policy-strategies.[8] Markov strategies imply that actions in any given period depend on past history only through the 'state'. The 'state' is a (possibly multi-dimensional) variable which summarizes the influence of past interactions on the current strategic environment. In other words, the state is the *minimal* information in the history of a game which is relevant for the strategic interaction between players. In our context, the past has no direct effect on the actions of consumers, legislators and voters; that is, the state is empty. Bearing this in mind, in a Markov Perfect Equilibrium (MPE) of the model here, period-t policy decisions are optimal from each and every legislator's point of view given the constitutional rules of political interaction and voting performance standards, while the latter are a Nash equilibrium given the constitution and the rationally anticipated equilibrium period$-t$ policy decisions. In any period, legislators and voters take into account the effects of their actions in the yet to be determined competitive equilibria. In any period-t, consumers and firms make optimal decisions taking as given prices, profits and policies over time.

Given that policy-makers face the same environment in any given period in office, we focus on stationary equilibria. Accordingly, we drop the time-superscript t hereafter to lighten the notation. Moreover, we drop, until further notice, the dependence on productivity θ of the various endogenous variables.

We turn to the characterization of equilibria and how these depend on the constitutional rules. As in PRT, we restrict attention to the case where $\bar{s}/3 \geq \delta W$ and assume that the status quo-policy is inefficient in that, in equilibrium, voters prefer the equilibrium policy instead of the status quo-policy. Moreover, we will restrict attention to the case of sufficiently low discount factor; in particular, we restrict parameters so that (in equilibrium) $\hat{s} > 6\delta W$.

[8] It is well-known that in dynamic non-cooperative games multiplicity of equilibria arises. For a discussion of the advantages of Markov strategies in dynamic games see Fudenberg and Tirole (1991, ch13) Our focus on symmetric equilibria is driven by the fact that regions are ex ante identical.

3 Equilibrium

That citizens and politicians are assumed to be identical implies that the problem of how first-period representatives are selected is orthogonal to our analysis, and hence skipped. Given the stationarity of equilibria, we can thus focus on the strategic interactions within some arbitrary period from the point where the legislature is formed until the stage when policy is chosen (and credibly announced).

Importantly, the structure of our model is isomorphic to PRT with $\hat{\tau}$ here being the counterpart of the maximum tax (set to 1) there. Namely, for any $\tau < \hat{\tau}$, a marginal increase in the tax lowers, all other things equal, voters' utility (by y utils here and by 1 util in PRT) and raises total fiscal revenues (by $3\theta(q + \tau q' \frac{\partial L}{\partial \tau})$ units here and 3 units in PRT). It follows directly that the fundamental properties of the equilibrium characterized in PRT (for the parameter restrictions mentioned above) carry through unchanged in our set up. These properties are well-known and understood by now and, hence, their duscussion will be brief here.[9] Moreover, anticipating our empirical work we focus only on the discussion of policies in the Parliamentary and Presidential systems.

Under all systems the region which is not included in the (minimum winning) policy-making coalition does not receive any transfer and its representative does not receive any political rents. The reason is that this region's political support is not needed and transfers and political rents channeled to this region would leave the coalition with less resources to satisfy their voters and appropriate as political rents. Under the Presidential system, competition between voters for their regions to be included in the winning coalition formed by the spending-setter leads to zero transfers to all regions but the one of the spending-setter. Because under the Parliamentary system the policy-making coalition is predetermined, such competition does not take place. Moreover, because voters set their performance standards independently and simultaneously, there is a continuum of equilibria in terms of the division of total equilibrium transfers between the regions whose representatives form the government.

In the Parliamentary system, we also have the following. First, to ensure that they do not receive the lowest possible welfare and that they do not leave excessive rents to politicians,

[9]The details of equilibrium characterization are as in PRT and available upon request.

voters must leave the agenda-setting winning coalition indifferent between agreeing on the 'Leviathan' policy of maximizing revenues and providing zero public good and regional transfers (attaining a joint payoff of $\hat{s}$). And agreeing on a policy that ensures their re-election (attaining a joint payoff $s + 2\delta W$). Thus, in equilibrium, $s = \hat{s} - 2\delta W$. From these rents, the spending-setter will give just enough rents to his partner-legislator (and keep the rest for himself) to gain the latter's support for his proposal. Second, the equilibrium policy must be jointly optimal for the voters represented in the policy-setting coalition, conditional on satisfying the constraint that $s \geq \hat{s} - 2\delta W$. This follows from the fact that here voters in the region of the junior-partner are not threatened to be excluded from the governing coalition as long as the latter attains enough political rents. Therefore, voters in the jurisdictions whose representatives are in the government maximize, in effect, through their choice of their performance standards their total utility $2(V[\tau] + H[g] + r)$ subject to $s = \hat{s} - 2\delta W$ and $s + r + g = R[\tau]$, where r denotes total regional transfers. This ensures that they attain the maximum possible payoff conditional on the government not breaking down and being re-elected and the fiscal resource constraint is satisfied.

Under the Presidential constitution, spending-setting is constrained by the tax-outcome in the tax-setting stage. We thus have that voters must ensure at the lowest cost for them that the agenda-setting winning coalition is (weakly) worse off by agreeing on providing zero public good and regional transfers (attaining a joint payoff of $R[\tau]$) instead of agreeing on a policy that ensures their re-election (attaining a joint payoff $s + 2\delta W$). Thus, in equilibrium, $s = \max\{0, R[\tau] - 2\delta W\}$. From these rents, the spending-setter will give just enough rents to the partner-legislator to gain the latter's support for his proposal. In addition, at the tax-setting stage, all legislators are residual claimants in expected terms. This implies that in equilibrium (where legislators are reelected), the tax-setter proposes a tax that ensures revenues such that $\max\{0, R[\tau] - 3\delta W\} \geq \hat{s} - 6\delta W$. Any lower revenues would imply that he has an incentive to deviate by offering the maximum admissible tax.[10] Let τ^C denote the

[10]To see this, note first that, in equilibrium, the tax-setter will be included in the spending-setter's coalition with probability 1/2. Hence, his expected payoff is $\frac{1}{2}(\max\{0, \frac{1}{3}R[\tau] - W\}) + W$. In equilibrium, it must be that the former is at least as high as the expected payoff from deviating to a tax τ' different than τ. If this payoff is $\frac{1}{2}\frac{\hat{s}}{3}$, this implies the inequality in the main text. The latter payoff corresponds to a deviation of $\tau' = \hat{\tau}$ and the spending-setter proposing the full expropriation policy with the tax-setter being the partner with probability 1/2. The fact that $\tau' = \hat{\tau}$ is the best deviation for the tax-setter follows

minimum admissible tax that satisfies this inequality. Following the corresponding arguments in PRT, one can see, due to $\bar{\tau} < \hat{\tau}$ and legislators being residual claimants (in expected terms) that in equilibrium $\tau = \tau^C = R^{-1}[\hat{s} - 3\delta W]$, where R^{-1} is the inverse if R. Finally, voters in the spending-setting jurisdiction maximize, in effect, through their choice of the performance standard, their utility $V[\tau] + H[g] + r$ subject to $s = \max\{0, R[\tau] - 2\delta W\}$ and $s + r + g = R[\tau]$ *taking as given* the tax $\tau = \tau^C$. This ensures that they attain the maximum possible payoff conditional on the minimum winning coalition being re-elected and the fiscal resource constraint is satisfied given the tax determined in the tax-setting stage.

We can now turn to the characterization of equilibrium policy.

3.1 The Parliamentary System

Denote the policy under the Parliamentary system with the superscript P. We thus have that rents are given by:

$$s^P = \hat{s} - 2\delta W^P.$$

From the above, the tax, the level of public good and total transfers to the governing jurisdictions are given by the solution to:

$$\max_{\substack{r \geq 0, \tau \in [0,\hat{\tau}] \\ R[\tau] - r \geq s^P}} 2H[R[\tau] - r - s^P] + 2V[\tau] + r.$$

Note that $R[\tau] - s^P \geq r \geq 0$ and $\hat{s} > 6\delta W^P$ (by assumption) implies that $R[\tau^P] \geq s^P > 0$ and hence $\tau^P > 0$. Let $\tilde{\tau}^P$ be the solution of $2V'[\tau] + R'[\tau] = 0$. Note that $\tilde{\tau}^P < \hat{\tau}$. It turns out that we can ignore the constraints $\tau \in [0, \hat{\tau}]$. These are satisfied by the solution to the relaxed problem. We then have (after a trivial inspection of the first order conditions with respect to τ and r of the relaxed problem[11]) that the solution of the above problem is such that:

from an argument which is identical to the corresponding one in PRT, after noting that $V[\tau]$ is decreasing and $R[\tau] - \max\{0, \frac{1}{3}R[\tau] - \delta W\}$ is increasing in τ.

[11]Note that if $R[\tau] - r = s^P$ then the first order condition with respect to r and our assumption that $H'[0] > 1$ implies that $r \geq 0$ is also binding.

$g^P = \min\{H'^{-1}[\frac{1}{2}], R[\tau^P] - s^P\}$ and

$r^P = \max\{0, R[\tilde{\tau}^P] - s^P - H'^{-1}[\frac{1}{2}]\}$ with

(a) $\tau^P = \tilde{\tau}^P$ if $H'^{-1}[\frac{1}{2}] \leq R[\tilde{\tau}^P] - s^P$,

(b) $0 = V'[\tau^P] + H'[R[\tau^P] - s^P]R'[\tau^P]$

if $H'^{-1}[\frac{1}{2}] > R[\tilde{\tau}^P] - s^P$ and $V'[R^{-1}[s^P]] + H'[0]R'[R^{-1}[s^P]] > 0,$ and

(c) $\tau^P = R^{-1}[s^P]$

if $H'^{-1}[\frac{1}{2}] > R[\tilde{\tau}^P] - s^P$ and $V'[R^{-1}[s^P]] + H'[0]R'[R^{-1}[s^P]] \leq 0.$

3.2 The Presidential System Policies

Let superscript C denote policy under the Presidential system. Recall then that the tax is given by

$$\tau^C = R^{-1}[\hat{s} - 3\delta W^C]$$

Turning to spending policy for given tax equal to τ^C, given that $R[\tau^C] = \hat{s} - 3\delta W^C > 2\delta W^C$ (by assumption), we have that

$$s^C = \hat{s} - 5\delta W^C$$

and transfers and public good are given by the solution of:

$$\max_{R[\tau^C] - s^C \geq r \geq 0} H[R[\tau^C] - r - s^C] + r.$$

Due to our assumption that $H'[0] > 1$, we can ignore the constraint $R[\tau^C] - s^C \geq r$ (it is satisfied by the solution of the relaxed problem). In fact, after a trivial inspection of the first order condition with respect to r of the relaxed problem, the solution is:

$$g^C = \min\{H'^{-1}[1], R[\tau^C] - s^C\},$$
$$r^C = \max\{0, R[\tau^C] - s^C - H'^{-1}[1]\}.$$

3.3 Comparing Taxes

To compare the taxes of the Parliamentary and Presidential systems we need to characterize the equilibrium contnuation payoffs across regimes. We have that in equilibrium:

$$W = \frac{s}{3} + \delta W,$$

given that every legislator has 1/3 chance to be the spending-setter and 1/3 chance of being in the spending-setting coalition. Recall also that $s^P = \hat{s} - 2\delta W^P$, while $s^C = \hat{s} - 5\delta W^C$. Using these, we have

$$W^P = \frac{\hat{s}}{(3-\delta)},$$
$$W^C = \frac{\hat{s}}{(3+2\delta)}.$$

Clearly, $W^C < W^P$. Note also that $\hat{s} > 6\delta W^P$ and $\hat{s} > 6\delta W^C$ if $\delta < 3/7$. Thus, restricting attention to the case of $\delta < 3/7$ (to be consistent with the case we focus on of $\hat{s} > 6\delta W$ under any political system) we have that $\hat{s} - 3\delta W^C < \hat{s} - 2\delta W^L$.

We turn to the comparison of taxes between the parliamentary and the presidential systems.[12] Clearly, we have from the above that $R[\tau^P] \geq s^P$. Therefore, $\tau^P \geq R^{-1}[\hat{s} - 2\delta W^P]$

[12]From $\hat{s} - 5\delta W^C < \hat{s} - 2\delta W^p$ we also have that political rents are lower under the presidential system. The reason is, as in PRT, that the spending-setter in this system does not have tax-setting powers as well. Recall also that under the presidential system we have $g^C = \min\{H'^{-1}[1], R[\tau^C] - s^C\} = \min\{H'^{-1}[1], 2\delta W^C\}$, while under the parliamentary system we have $g^P = \min\{H'^{-1}[\frac{1}{2}], R[\tau^P] - s^P\}$. Note that $H'^{-1}[\frac{1}{2}] > H'^{-1}[1]$. Moreover, note that $R[\tau^C] - s^C = 2\delta W^C = 2\delta\frac{\hat{s}}{3+2\delta}$ and $R[\tau^P] - s^P = R[\tau^P] - \hat{s} + 2\delta W^P = R[\tau^p] - 3\frac{\hat{s}(1-\delta)}{3-\delta}$. Using these observations, one can very easily see that if $2\delta\frac{\hat{s}}{3+2\delta} < R[\tau^p] - 3\frac{\hat{s}(1-\delta)}{3-\delta}$ then $g^C < g^P$ as in PRT. Similarly, if $2\delta\frac{\hat{s}}{3+2\delta} \geq R[\tau^p] - 3\frac{\hat{s}(1-\delta)}{3-\delta}$ and $H'^{-1}[1] < R[\tau^p] - 3\frac{\hat{s}(1-\delta)}{3-\delta}$. However, in the remaining case, we have, in contrast to PRT, that $g^P \leq g^C$: in this case the tax distortions are high enough to make public good provision lower under the Parliamentary system. Note that this case is relevant if, for instance, $H'^{-1}[\frac{1}{2}] > R[\tilde{\tau}^P] - s^P$ and $V'[R^{-1}[s^P]] + H'[0]R'[R^{-1}[s^P]] \leq 0$, in which case we have (after recalling the characterization of Parliamentary policy) that $R[\tau^p] = s^P = 3\frac{\hat{s}(1-\delta)}{3-\delta}$ (note also that the latter inequality

$> R^{-1}[\hat{s} - 3\delta W^C] = \tau^C$. Thus, as in PRT, the presidential system raises **less taxes** than the parliamentary system. The reason is similar: the distortions in the tax-setting are sufficiently low to not deter the agenda-setter of a simple legislator from setting a very high tax, and voters in the tax-setting jurisdiction under a presidential system prefer the lowest possible tax consistent with equilibrium as they receive no transfers.

Anticipating the empirical part of our paper, we summarize the key empirical predictions of the model. Recall, that an increase in the tax-rate leads to an increase in the labor share, all other things equal. Note here that, for any given tax, the equilibrium wage depends on the productivity level. This effect however cannot be signed (due to $q' > 0$, $q'' < 0$ and $\partial L[\tau_t, \theta]/\partial\theta > 0$) without imposing further restrictions on the fundamentals of the model. Note also that, for any given tax, the equilibrium income per capita is increasing in productivity (due to $q > 0$, $q' > 0$ and $\partial L[\tau_t, \theta]/\partial\theta > 0$). This implies, in turn, that total revenues are also increasing in productivity for any given tax. Taking into account that optimal taxes depend also on productivity, complicates the relationships in question, as discussed above. Specifically, to sign them also requires further parameter restrictions. Using the labor share (which in terms of our model is $q'L/q$) as a proxy for wages, our data analysis shows that we should expect wages to be positively related with productivity (ie. that the direct effect of productivity - that is, for given labor - might dominate in reality). Moreover, private income per capita and the size of the government as measured by revenues per GDP, which in terms of our model is the income tax rate (note that $R/3y = \tau$), are positively related to productivity.

The key prediction of the model is that Parliamentary systems other things equal will have higher wages solely due to the political system. Of course, wages will also be influenced over time by changes in productivity, with the net effect depending on the strength of the effect of the latter on taxes (for given political system). The first of the next two sections provide evidence of the effect of Parliamentary systems, and the second that this effect is due to variation in the tax level.

---

cannot be satisfied in PRT where taxation is lump-sum taxation because in that case we have $V' = -R'$). Nevertheless, distinguishing (pure) public good provision from redistribution in the data is very hard and therefore we do not pursue this issue further in what follows.

4 Cross-Sectional Estimates

4.1 Methodology

This section will provide a brief overview of the first econometric approach employed, and how it provides for causal inference. Isolating the effects of constitutions from other potential determinants of the labor share is intrinsically complicated by the interactions between market and state. The approach taken is to conceive of the choice of constitution as a treatment, and to estimate the effect of that treatment. However, consistent estimation requires that the choice of constitution must be independent of any other factor determining the labor share. Accordingly, we first outline what is meant here by a causal effect of presidential democracy and the conditions necessary to estimate it consistently. We then use consider how BMA can be used to maximize the chances of satisfying these conditions, given the available data.

Formally, let Y_C, Y_P be the outcomes associated with a congressional/presidential system or a parliamentary constitution respectively. X is the set of variables which may partially determine the choice of constitution, and $S \in \{0, 1\}$ is the choice of constitution, with $S = 1$ denoting a presidential system. It is unlikely that S is independent of Y_C, Y_P i.e:

(1) $$S \perp Y_C, Y_P$$

But it is potentially true that conditional on X, S is independent of Y_C, Y_P:

(2) $$S \perp (Y_C, Y_P) \mid X$$

As is standard, the relationship between the outcomes Y and the treatment S can be written as follows:

(3) $$Y = (1 - S)Y_C + SY_P = Y_C + S(Y_P - Y_C)$$

If estimates of (3) using OLS are to be unbiased then (2) must hold and as such it is necessary to include the confounding variables X, whilst Y_P is subsumed into the constant term which is denoted α. Then by including a binary variable, S, to denote whether or not a particular country has received the treatment (in this case a presidential constitution) the associated coefficient β is an estimate of the treatment effect. Such an OLS model can be written as follows:

$$Y = \alpha + \beta S + \gamma X + \varepsilon \tag{4}$$

such that:

$$E(\varepsilon) = E(\varepsilon S) = E(\varepsilon X) = E(\varepsilon \mid S) = 0 \tag{5}$$

For estimation to be consistent both (2) and (5) must be true. This requires that there are no relevant variables missing from X, and that those variables included are uncorrelated with the error term, i.e. exogenous.

The reasons why different nations have chosen different constitutional rules are complex and varied. As discussed in Persson and Tabellini (2003) and Acemoglu (2005) intellectual fashion and also potential colonial influence have been of particular importance. But there are many other, potentially complimentary, possible explanations and the number of variables required to describe these is large. The small sample available prohibits including them all in a regression analysis and hence leads to concerns about model uncertainty since it is not known *a priori* what are the constituents of X.[13] Many traditional econometric approaches to this problem, such as stepwise regression, suffer from path-dependence, that is they are sensitive to the order in which variables are included. Moreover identifying the constituents of X via any attempt to test down to a parsimonious specification from a large set of variables will lead to the inferential problems associated with data-mining as detailed by Miller (2002, Ch.6). In contrast, a Bayesian Model Averaging (BMA) approach may be preferable since it will provide an estimate of the likelihood of different choices of X

[13]Whilst our model describes the relationship between taxation, government size, and the labor share, it does not make any predictions as to which form of democracy a country will adopt.

and also a posterior distribution for β obtained from each of the different possible models weighted by their respective posterior model probabilities.

The remainder of this section will provide a brief overview of BMA which is described in more detail in Hoeting, Raftery, Madigan and Volinsky (1999) and Malik and Temple (2009). BMA is premised on the basis that since there are sometimes multiple, similarly likely, statistical models which imply different inferences, it is sometimes helpful to consider a wide range of possible models, and the overall likelihood of a variable being important.[14] A BMA analysis starts with a set of prior beliefs about which models are expected to be more likely, and prior beliefs about the distribution of the coefficients on particular variables. For example, if one had a strong theoretical justification for believing, or previous results suggested, that a certain variable was likely to be statistically important then models which included that variable could be given a higher prior probability. Similarly, if it was believed that this variable was very likely to be negatively associated with the dependent variable then its prior distribution could be chosen such that the majority of the probability mass was where the coefficient was negative. In the analysis here, few assumptions are made as to the prior distribution. Instead a "diffuse" prior is used: in particular it is assumed that every possible model has an equal prior probability, that is if there are 2^{25} possible models then each model has a prior probability of $\frac{1}{2^{25}}$. This assumption implies that every variable, including the treatment, is assumed to have an equal chance of 0.5 of inclusion in any given model. The prior distributions of the coefficients associated with each variable are chosen to have zero mean, and variance proportional to the sample variance of the explanatory variable.

Given these choices, the posterior model distribution (the probability of each model given the data) is calculated. Following Hoeting, Raftery, Madigan and Volinsky (1999) let Λ be the quantity of interest, such as the effect of a presidential constitution, and D the dataset. There are $N = 2^K$ possible models M_i where K is the number of explanatory variables. The posterior distribution of Λ is given by:

[14] The motivation for this approach is slightly different to that used in the Growth Determinants literature (c.f. Sala-I-Martin, Doppelhofer and Miller (2004)) where the focus is often on identifying which variables explain variation in Y. Here, as discussed, the focus is on ensuring consistent estimation of β.

$$pr(\Lambda \mid D) = \sum_{i=1}^{N} pr(\Lambda \mid M_i, D) pr(M_i \mid D) \tag{6}$$

where the posterior probability of any given model, M_i is:

$$pr(M_i \mid D) = \frac{pr(D \mid M_i) pr(M_i)}{\sum_{j=1}^{N} pr(D \mid M_j) pr(M_j)} \tag{7}$$

where:

$$pr(D \mid M_i) = \int pr(D \mid \phi_i, M_i) pr(\phi_i \mid M_i)\, d\phi_i \tag{8}$$

The vector ϕ_i represents the parameters for model i, i.e. $\phi_i = \{\alpha_i, \beta_{i1}, ..., \beta_{iK}, \sigma_i\}$. Where $\sigma_i = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{T-k}$ denotes the error variance of model i, where T is the number of observations and k is the number of variables in model i. The exact interpretation of (6), (7), and (8) are discussed more thoroughly in Kass and Raftery (1995), Raftery, Madigan and Hoeting (1997), and Hoeting, Raftery, Madigan and Volinsky (1999). In essence (8) describes the chance of observing the data if that particular model was the model believed before the data were observed. The posterior probability of a particular model given by (7) describes the probability of that model once the data have been observed and (6) describes the calculation of the conditional distribution of Λ, given the data D that is the probability of Λ taking a given value for each model multiplied by the posterior probability of that model.

Once the posterior model probability (PMP) of each model has been calculated, several related quantities can be obtained. The posterior inclusion probability (PIP) is the sum of the PMPs of those models which include that variable (i.e. those in which its coefficient is non-zero). Also, the posterior mean and standard deviation of a given variable can be calculated by computing the weighted average of the mean or standard deviation across all models weighted by the PMPs.

4.2 Data

This section will discuss the data used to measure the labor share, the type of democracy, and the set of candidate control variables. Following Rodrik (1999) the labor share in value added in manufacturing is used as the measure of the labor share. And similarly the data are taken from the UNIDO Industrial Statistics database. The labor share was calculated as average labor costs divided by the mean value added per worker, and a five year average was then created. Recall, that in the notation of the model this is:

$$\frac{q'[l_t^j]L[\tau_t, \theta_t]}{q[l_t]} \tag{9}$$

The only difference with Rodrik's approach is that the data were calculated for each year in the period 1990-94; this period represented the years for which there was greatest data availability and corresponds to the data used by Persson and Tabellini which is also for the early 1990s.[15] There has been some criticism of the use of factor-share data. In particular, Gollin (2002) claims that previous work using data on factor shares overstates the variation between countries, as a consequence of failing to take into account the income of entrepreneurs and more generally the self-employed. However, these criticisms seem less applicable to manufacturing industry data which is used for this reason and because it is available for a large set of democracies. Moreover, we are unaware of any alternative measure that represents comparable quantities across countries and time. One advantage of using the labor share is that since it is constructed as a ratio it alleviates potential concerns about inadequate adjustments for inflation or exchange rate movements.

Persson and Tabellini (2003) define six variables that describe different aspects of constitutional type. These variables all measure aspects of the differences between parliamentary and presidential democracies. The first, *pres,* is a dummy variable which takes a value of one if the executive is not accountable to the legislature via no-confidence

[15] Further details, and the data used are available on request.

votes. This measure corresponds to the distinction between congressional and parliamentary regimes discussed by Persson, Roland and Tabellini (1997), and the model presented above in Section 2. This variable will be used to measure the 'treatment' associated with adopting a presidential democracy. In order to establish that it is indeed a presidential system that is the key source of variation rather than a more general majoritarian consensual dichotomy five other measures of constitutional form, describing electoral systems, proposed by Persson and Tabellini (2003) were used in a principal components analysis to create a single measure of other differences in electoral system. The principal components were created using Persson and Tabellini's variables *Pind*, *Magn*, *Sdm*, *Spropn*, and *maj*.[16] The five principal components will be denoted by *g1* , *g2*, ... , *g5* . An analysis of the loadings of the principal components suggests that indeed the first component broadly measures countries on a First-Past-The-Post to Proportional-Representation spectrum.[17]

The predictions of the theoretical analysis are that the labour share will be higher in parliamentary democracies than presidential democracies, conditional on the productivity level (θ). As a measure of θ, we use *loga*, the logarithm of Total Factor Productivity (TFP). TFP is not necessarily predetermined and maybe partly determined by the choice of constitution and the pre-treatment control variables that generate constitutional selection.[18] If constitutional choice partly determines productivity and this has an effect on the labor share, then there will be an indirect effect of constitutional choice on the labor share due to its effects on income levels. In this case, to maintain the assumption that the coefficients on the constitutional variables identify a causal treatment effect it is required that *loga*is independent

[16] *Pind* describes the proportion of the legislature not elected on the basis of party lists. In bicameral democracies it refers to elections to the lower house. *Magn* is the "inverse of district magnitude", the number of electoral districts per seat in the lower chamber. *Sdm* is analogous to *magn*, but where there are electoral districts of different sizes it calculates the inverse of district magnitude as the weighted average of the different district sizes, where the weighting for each district size is the percentage of seats in the legislature elected from districts of that size. *Spropn* describes the proportion of electoral seats elected from national electoral districts rather than sub-national districts. In this respect it captures something akin to *pind*. The final variable is *maj* which takes a value of one if elections to the lower-house of the legislature are by plurality (first-past-the-post) rule, or zero otherwise.

[17] The first principal component explains 70% of the total variance and positively weights all of the constitutional variables except *spropn* which has a very small negative value. The second principal component acounts for a further 20% of the variance and places most weight on *spropn.*

[18] Persson and Tabellini (2004, ch7) provide evidence that Presidential regimes are associated with lower productivity (and income per capita) and that this effect is due to government policy less supportive of productive activity.

of the error term conditional on X (the predetermined controls), that is, $loga \perp \varepsilon \mid X$. Following Lee (2005, Ch.2) then the inclusion of *loga* as a candidate independent variable will mean that the indirect effects on the labor share of constitutional choice through *loga* will be partialled out. Therefore, the coefficients associated with the treatment (constitutional choice) will describe solely the direct effects of the treatment, which is the quantity of interest.

It might be expected that impact of both democracy in general and its type would vary with the quality of government. Furthermore, the work of Acemoglu and Robinson (2001) and Dulleck and Frijters (2004) suggests that part of this variation may be due to reasons not captured by our existing control variables. In particular, resource wealth may, other things being equal, lead to lower quality government. However, including a measure of the degree of democracy, *PolityIV* in X suggested it had little predictive power.[19] But, it is not plausible to make the same assumption about the conditional exogeneity of *polityiv* as it is for *loga*. Therefore a plausibly exogenous proxy variable, or instrument, is needed. The instrumental variable *partitioned* is from the data created by Alesina, Easterly, and Matuszeski (2006) in which they investigate the extent to which states are often "artificial", created by previous colonialists rather than representing underlying ethnic groups. *Partitioned* describes the proportion of a state's population who are members of an ethnic group which is present in one or more adjacent countries. They find that *partitioned* is correlated with measures of good governance. Using both OLS and BMA analysis, *partitioned* is found to be a good predictor of *polityiv* and is considered plausibly exogenous. However, results not reported in the interests of brevity suggest that our conclusions are quantitatively and qualitatively unaffected by the inclusion of either variable, and the different sample sizes this entails.

The other candidate control variables are largely from Persson and Tabellini (2003). *engfrac* describes the proportion of the population speaking English as a first language, *eurfrac* is the same but for the major European languages English, French, German, Portuguese, or Spanish. *engfrac* and *eurfrac* are included based upon the work of Hall and Jones (1999) and, to a lesser extent, Acemoglu, Johnson and Robinson (2001). *ilat01* measures absolute distance from the equator. A variety of age measures were employed, *age* is Persson and Tabellini

[19]The Polity IV project codes regimes from 10 - fully institutionalized democracy to -10 - fully institutionalized autocracy.

(2003)'s measure. Also included are the variables, proposed by Rockey (2010): *mthconstit* and *mthelect* which are new measures of when a country first promulgated a democratic constitution, and when it held its first democratic election respectively. These variables are argued to represent a useful alternative to *age*. The variables $con2150$,$con5180$,$con81$ are indicator variables which describe whether the current constitution was promulgated between 1921 and 1950, 1951-1980 or post-1981 with 1920 or earlier the omitted category. The inclusion of these variables is designed to represent the well-documented notion of different waves of democratization. These waves coincided with systematic variations in what constitutions were chosen, as discussed in more detail in Persson and Tabellini (2003) and Rockey (2010). Finally indicator variables are included for whether a country has a federal government (*federal*) or was colonized by the UK, Spain, or another European nation discounted by time since independence (*coluka, colespa, colotha*) and finally which continent it is part of (*africa, asiae, laam*). Summary statistics for these variables are reported in Table 1.

4.3 Results

This section comprises four parts. Firstly, it will discuss the results and implications of the benchmark specification presented in Table 2. Then it will discuss the implications of the relative lack of predictive power of the legal origin variables developed by Botero et al. (2004) as shown in Table 3. It will argue for the robustness of these results using a Markov-Chain-Monte-Carlo (MC^3) estimation approach that simultaneously performs BMA and the identification of outliers, the results of which are presented in Table 4. Related OLS estimates of all these analyses are contained in 5.

The results in Table 2 show that *pres* has a PIP of 99.9%. This implies that *pres* is included in almost every likely model. The posterior mean associated with *pres* is -0.12 which implies that workers in countries with a presidential system receive a share of value added 12 percentage points less than their counterparts in parliamentary democracies. This is especially striking given the small range of values of $mean9094$ which has a standard deviation of only 0.14. However, none of the other constitutional variables had high PIPs.[20]

[20] *g5* is excluded from the analysis as whilst it had an intermediate PIP, this was judged to be a statistical artefact given the negligible PIP of the first four principal components.

Taken together, these results provide strong evidence that the relevant dimension of variation that is important is the presence (or not) of a presidential system.

Table 2 also strongly suggests, in line with the predictions of the model, that more productive nations pay their workers a greater share of output. The predicted share of trade in national income (*frankrom*) also has a PIPs of over 90%. That countries likely to trade more than average have a higher labor share is perhaps related to the finding of Rodrik (1998) who suggests that more open economies have bigger governments to offset the risk of terms of trade shocks. Three measures of the age of democracy have intermediate PIPs ($con2150$,$con81$ and *mthconst*) which poses the question of whether they enter together or substitute for each other. Inspection of figure 4.3 which displays the composition of each model shows that they tend to enter models together. In fact, there would seem to be, broadly speaking, two likely types of model. The first, is similar to Model 2 in Table 2 and contains no age variables. The second type is similar to Models 1, 3, 4, and 5, which contain measures of democratic age in different combinations.[21] Furthermore, the main results are robust to the exclusion of all of the age-of-democracy variables from the analysis.

Rodrik (1999) finds that income per capita, is a significant determinant of the labor share. To facillitate comparison with his estimates, and anticipating the next-section where data-availability obliges using income per-capita as a proxy for productivity in the dynamic-panel estimates, estimates using (log) income per capita, denoted *logyl*, were obtained. Recall, that the model predicts that income per capita y is increasing in productivity θ. In the interests of parsimony we do not report the full results. However, they are even stronger in this case: the PIP of the *Pres* remains above 99% and the PIP of *logyl* rises to a similar figure, with a coefficient of around 0.08, suggesting a one standard deviation increase in income per capita leads to a similar increase in the labor share. The results are otherwise largely similar, although the PIP of *ilat01* falls by nearly 0.5.

Table 3 reports BMA results including the measures of legal origin *common*, *french*, *scandi* and *social* taken from Botero et al. (2004). These results are reported separately because only 47 observations are available for all of the relevant variables. The results are

[21]Blue/Darker Grey indicates a negative coefficient

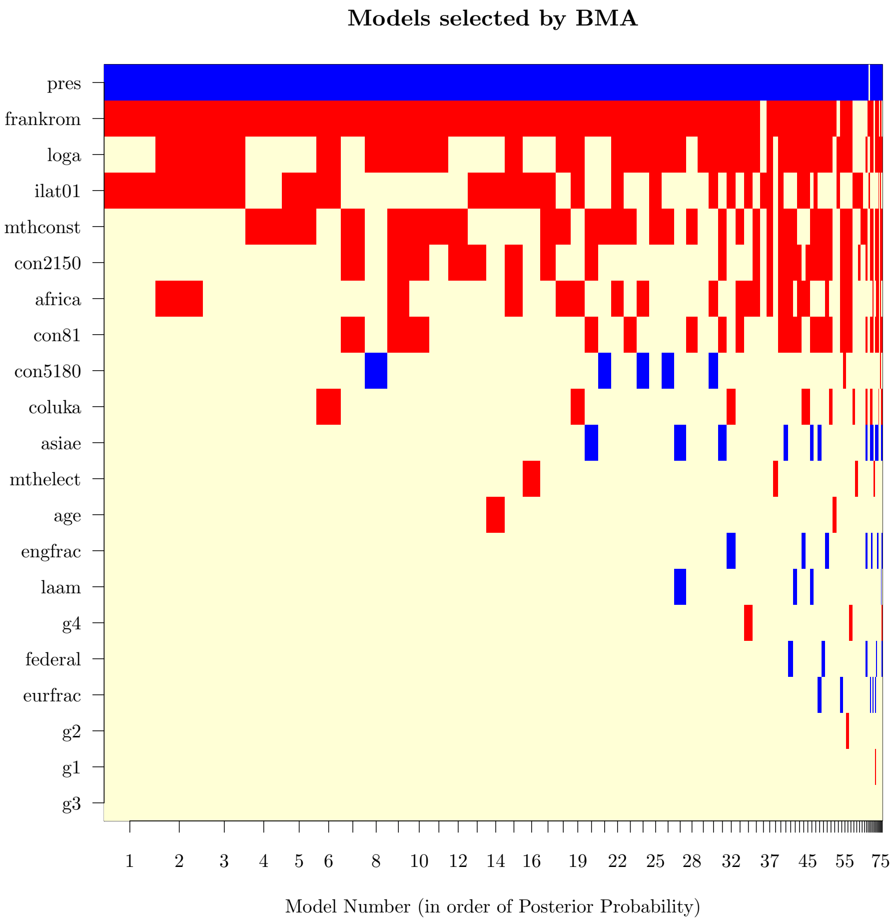


Figure 3: Models Chosen By BMA analysis

largely unchanged, but *scandi* has a high PIP of 0.68 and a positive coefficient. We don't explore here why this might be, but it is notable that there are no presidential democracies that use Scandanavian based law. We do not argue here that the low PIPs of the other legal origin variables cast doubt on Botero et al. (2004)'s findings, rather that if anything, one inference is that the determinants of labor regulation and the labor share are different.

Outliers are a problem that can affect cross-country analyses, as discussed by Temple (1998). Standard post-estimation methods of outlier identification, such as DFITS, are not compatible with the Bayesian approach. Instead, the two-stage estimation method combining outlier detection and BMA proposed by Hoeting et al. (1996) is used here. First, possible outliers are identified by the robust estimator Least Trimmed Squares (LTS) as developed by Rousseeuw (1984). Secondly, a Markov Chain Monte Carlo (MCMC) algorithm is used to estimate the posterior distribution of models. However, unlike the earlier BMA analysis, it estimates (6) and simultaneously the posterior distribution of the outliers identified by the LTS estimation.[22] The results of this estimation method are contained in Table 4. The results confirm that *pres* has a PIP of just under 100%, but also suggest Honduras and Ireland are outliers with 81% and 58% probability respectively. Posterior expected values are not available, but the OLS results in Table 5 discussed below suggest that the signs of the coefficients are unchanged, with the coefficient on *pres* slightly larger at around 14%. Despite its relative high PIP of 68% *loga* only appears in two of the five most likely models, although the cumulative posterior model probability of these five is not large. Again, variables seem to be entering jointly, with a parsimonious models in columns 1,2, and 4, and larger models in the other two columns. The parsimonious model is similar to that identified above for the BMA estimates, although the larger model includes different controls. For the smaller model, the PIPs of *frankrom* and *ilat01* are reversed, but otherwise the models are similar. In the larger models, the effect of democratic age is now largely being captured by *con2150*, and a good model is more likely to contain *africa* and *coluka*. Equivalent results, not reported, suggest that the alternative model including the legal origin variables is also outlier robust.[23]

OLS estimates of the most likely models identified in the BMA and MC^3 analyses with (HC3) robust standard errors are reported in Table 6. Again *pres* has an estimated coefficient of around -0.12 and is significant at the 1% level. However, some caution is necessary when interpreting these results, since when estimating a model identified through extensive model selection, conventional t-ratios, (as discussed by Miller (2002) and Malik and Temple (2009)), are generally biased away from zero. The coefficients are slightly different as there are slightly more observations available for the specific models reported than there are for the overall Model Space.[24] But, taken as a whole the BMA, MC^3 and OLS results all point

[22]The BMA was performed using code written in 'R'. The particular package used, bicreg, was Ian S. Painter's translation from the S^+ code by Adrian Raftery and revised by Chris Volinsky. The MC^3 estimates were arrived at using Ian S. Painter's translation of Jennifer Hoeting's S^+ code. More precisely, the BMA estimates were obtained considering all models that at most were 100 times less likely than the model identified as most likely. The hyperparameters used for the MC^3 estimation were those recommended in Hoeting, Raftery and Madigan (1996).

[23]A variety of other robustness tests were also performed. The results are as expected given the close to zero PIP of the constitutional variables other than *pres.* Tests included using the original variables describing constitutional form rather than those derived from the principal components analysis, using binary variable versions of the principal components, principal components derived from subsets of the constitutional variables, and including a wide range of interaction terms involving *age, polityiv* and the constitutional variables.

[24]BMA requires the same observations to be available for every possible specification, and there are

in the same direction: presidential democracies pay significantly lower wages. Moreover, since BMA helps to circumvent traditional issues concerning model uncertainty, there is little to suggest that the main finding is not unusually robust.

5 Dynamic Estimates

Whilst, the empirical evidence supports the central prediction of our model - that presidential democracies lead *ceteris paribus* to lower wage rates - it is worthwhile assessing whether the evidence also supports the mechanism that gives rise to this prediction. A central prediction of the model is that the labor share and average taxation are positively related. Moreover, as is discussed in Section 2, we expect the level of taxation to be increasing in the productivity level θ in our framework. As such, one test of our model is whether the tax rate, wages, and productivity are indeed all positively related. Unfortunately, to the best of our knowledge only a time-series data on TFP is only available from the mid-1980s. As such we use data on income per capita, since the model, the results of the previous section and standard economic intuition suggest they will be related. However, to verify this, a regression was run of GDP/capita growth ($\Delta logyl_{it}$) on TFP growth ($\Delta loga_{it}$) as well as country and year fixed effects. As expected, the results suggest a close correspondence between *logyl* and *loga*.[25] Using cross-sectional time series data on central government revenue (*cgrev*), the labor share of income (*labshare*) as well as income per capita (*yp*), we find evidence that this is indeed the case.

The presentation of the methodology employed will be necessarily brief here, however it is important to note two key econometric issues. Given the different historical experiences of the countries in the sample, it is reasonable to expect cross-sectional variation in the relationship between taxation, the labor share, and income. Moreover, there may be cross-sectional correlation between observations within a given period either due to the impact of a common shock, or interdependence between countries. To ameliorate such concerns, the data were

---

fewer observations for the variables *g1,...,g4*.

[25] These data are different to those used in Section 4.2 and are taken from the (Groningen Growth Development Centre (2010)) Database. The precise regression run was

$$\Delta logyl_{it} = \alpha_i + \beta \Delta loga_{it} + \lambda_t + \varepsilon_{it}$$

.

transformed such that each country-year observation is now expressed as deviations from the cross-country annual mean.

Panel tests for Unit Roots were conducted on these demeaned data, and the results are reported in Table 6. As is common for short time-series, these tests, which often lack power, are not altogether conclusive.[26] The Im-Pesaran-Shin tests and the ADF-Fisher tests both test the null-hypothesis of individual unit-roots. The tests suggest that we are unable to reject the null for *yp* and narrowly *labshare*the labor share. We fail to find evidence for a unit root of *cgrev*. For all three series we reject the null after differencing the data. Two tests of alternative null-hypotheses were employed, the Breitung test, and the Hadri test. However these are only technically suitable for balanced-panels, whereas our data are unbalanced. As such, whilst the tests are reported, their interpretation merits great caution. The Breitung test suggests we can not reject the null of a common-unit root for all three variables. We are able to reject the null for the first-difference of *labshare* and *cgrev*, but not for *yp*. We disregard this result as an artefact given that we are not aware of any other results in the wider literature that concur with the result that the first difference of income per capita is non-stationary.

Finally, we perform a Hadri test, which tests the alternative null-hypothesis of stationarity, whilst rejecting stationarity for the level data, it also rejects it for the differenced data. This would seem suprising, however Hlouskova and Wagner (2006) argue based on a Monte-Carlo study that this test often lacks power, and we interpret these results as a consequence of this and our misapplication of the tests to a non-balanced panel. Our interpretation of the test results is that they suggest it is prudent to use differenced data.

The cointegration tests reported in the bottom part of Table 6 are less ambiguous. The results of the Pedroni (1999, 2004) test suggest that we can reject the null of no cointegration. In sum, the tests suggest that the variables are both integrated and cointegrated. As such we use first differences of the variables and specify the following error-correction model. The specification allows for cross-country heterogeneity in the time-series behavior of the data allowing country specific slope coefficients

[26]For an excellent discussion of the tests used, see Banerjee and Wagner (2009) and references therein.

$$
\begin{aligned}
\Delta labshare_{it} \sim\ & \psi_i(labshare_{i,t-1} - \zeta_{0i} - \zeta_{1i} cgrev_{it} - \zeta_{2i} yp_{it}) \\
& + \alpha_i \Delta labshare_{i,t-1} + \beta_i \Delta cgrev_{i,t-1} + \gamma \Delta yp_{it} \\
& + \lambda_{i1} \Delta trade_{it} + \lambda_{2i} \Delta oil_{it} + \epsilon_{it}
\end{aligned}
$$

However, a Hausman test suggests that we are unable to reject the hypothesis that the there is no systematic variation in ζ_i. That is, we are unable to reject the hypothesis of a consistent cross-country long-run equilibrium. Under the assumption of long-run homogeneity, the Pooled-Mean-Group estimator proposed by Pesaran, Shin and Smith (1999) is our preferred estimator. It constrains the long-run error-correction parameters to be equal across countries, but allows for cross-country variation in the short run variables and the error correction term, and is both efficient and consistent in the case of homogeneity in the error correction terms. In the context of the above model, this implies that $\zeta_i = \zeta$ for all i, and the estimates reported are for the following model:

$$
\begin{aligned}
\Delta labshare_{it} \sim\ & \phi_i(labshare_{t-1} - \zeta_1 cgrev_{it} - \zeta_2 yp_{it}) \\
& + \alpha_i \Delta labshare_{i,t-1} + \beta_i \Delta cgrev_{i,t-1} + \gamma \Delta yp_{it} \\
& + \lambda_{i1} \Delta trade_{it} + \lambda_{2i} \Delta oil_{it} + \epsilon_{it}
\end{aligned}
$$

The results are presented in Table 7. The evidence supports the hypothesis of a long-run relationship between the labor share and taxation. The coefficient on *cgrev*, ζ_1 in the previous equation, is positive and significant at the one percent level. The reported short-run coefficient estimates are the average of the coefficients obtained for each country. There would seem to be little evidence for any short-run relationship between the labor share, taxation or income, although an increase in the share of trade in GDP is associated with a decrease in the labor share.

One further test of the relationship between taxation and the labor share, and hence presidentialism and the labor share was conducted. Columns 2 and 3 report PMG estimates for sub-samples for presidential and non-presidential systems respectively. If, contrary to the BMA results, some ommited factor were driving the results we would expect different estimated relationships between taxation, income, and the labor share. Of course, that we do does not confirm our theory, but is considered to be supportive both of the mechanism described by the model, and the overall effect of presidential democracies on wages. As if mechanism driving our result were different to what we postulate, then we might expect to observe qualitatively different relationships between the variables, in different systems.[27] As discussed above, ideally we would observe random transitions between types of democratic systems so to easily identify the causal effect of presidential system on taxation and the labor share. However, we observe few changes in the form of democracy, and these changes are unlikely to be random. Hence, our empirical strategy has been first, to identify a causal effect of the type of democracy given an assumption of selection on observables, and then to confirm the second prediction of the model that differences in the labor share should be in part due to differences in income and taxation levels. Taken together, we argue that the evidence suggests that presidential democracies pay lower wages, and that this is because of, all else equal, lower levels of taxation.

6 Conclusions

This paper has argued that presidential democracies are associated with lower wages because in equilibrium there are lower levels of taxation. In particular it employs a variant of the Persson, Roland and Tabellini (1998) model with costly taxation, situated in a simple macroeconomic framework to generate these predictions. These predictions were then tested using Bayesian Model Averaging to estimate the treatment effect on the labor share of income associated with Presidential as opposed to Parliamentary democracies. Furthermore, evidence that these differences are indeed due to differences in taxation is provided by results obtained using Pesaran, Shin and Smith (1999)'s Pooled-Mean-Group estimator. However, it would be foolhardy to claim too much on the basis of these results. In particular the effect estimated is the average treatment effect, and does not consider possible heterogeneity.

[27]Results were also obtained using data explicitly describing revenues from different sources of taxation, business, personal, etc. However, these data are only available for some of the OECD and only one presidential democracy (the USA) although the same positive relationship was identified.

Moreover, no welfare analysis is attempted of whether voters are better of with high wages, high taxes, high levels of public goods, and high levels of wastage, or vice-versa. There are also several limitations to our approach, in particular, ideally there would be improved data on the labor share available for a large cross-section of countries, and similarly data on tax revenues for non-OECD countries. However, whilst improved data might make for more precise inferences, we do not believe that the overall result is likely to be affected.

Table 1: Summary Statistics - BMA Variables

Variable	N	Mean	Std. Dev.	Min	Max
mean9094	72	0.366	0.138	0.102	0.664
pres	85	0.388	0.490	0.000	1.000
g1	75	0.000	1.886	−1.880	2.778
g2	75	0.000	1.024	−0.678	4.142
g3	75	0.000	0.497	−1.509	1.361
g4	75	0.000	0.336	−0.948	1.217
g5	75	0.000	0.187	−0.700	0.338
federal	83	0.157	0.366	0.000	1.000
logyl	84	9.176	0.881	6.954	10.476
ilat01	85	30.747	18.456	1.218	67.470
age	85	0.208	0.219	0.030	1.000
mthconst	85	0.542	0.569	0.035	2.577
mthelect	85	0.509	0.479	0.044	2.074
con2150	85	0.106	0.310	0.000	1.000
con5180	85	0.294	0.458	0.000	1.000
con81	85	0.494	0.503	0.000	1.000
coluka	85	0.282	0.393	0.000	0.928
africa	85	0.129	0.338	0.000	1.000
asiae	85	0.153	0.362	0.000	1.000
laam	85	0.271	0.447	0.000	1.000
frankrom	78	2.866	0.840	0.940	5.639
engfrac	78	0.138	0.318	0.000	1.000
eurfrac	78	0.398	0.437	0.000	1.004
common	59	0.339	0.477	0.000	1.000
social	59	0.153	0.363	0.000	1.000
french	59	0.356	0.483	0.000	1.000
german	59	0.085	0.281	0.000	1.000
scandi	59	0.068	0.254	0.000	1.000

Table 2: Benchmark Specification - BMA Estimates

Variable	Posterior Inclusion Probability	Posterior Expected Value	Posterior Standard Deviation	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	1	−0.074	0.304	0.200	−0.370	−0.133	0.245	0.199
pres	0.998	−0.120	0.037	−0.106	−0.100	−0.099	−0.142	−0.112
frankrom	0.964	0.043	0.017	0.044	0.041	0.045	0.048	0.046
loga	0.607	0.037	0.037		0.071	0.042		
ilat01	0.521	0.001	0.001	0.003	0.002	0.002		0.002
mthconst	0.51	0.031	0.037				0.071	0.044
con2150	0.323	0.023	0.040					
africa	0.296	0.024	0.044		0.085			
con81	0.246	0.016	0.032					
con5180	0.095	−0.004	0.017					
coluka	0.089	0.005	0.022					
asiae	0.073	−0.005	0.020					
mthelect	0.034	0.002	0.011					
age	0.028	0.003	0.019					
engfrac	0.027	−0.002	0.013					
laam	0.027	−0.002	0.012					
g4	0.015	0.000	0.005					
federal	0.015	−0.001	0.007					
eurfrac	0.014	−0.001	0.007					
g2	0.004	0.000	0.001					
g1	0.001	0.000	0.000					
g3	0	0.000	0.000					
R^2				0.57	0.623	0.596	0.565	0.593
BIC				−39.150	−38.992	−38.777	−38.453	−38.361

The dependent variable is the mean of the labor share of value-added in manufacturing in the period 1990-1994. The variables are sorted by their posterior inclusion probabilities (PIPs): the sum of the posterior probabilities of those models which include that variable. The posterior expected value is the weighted average of the expected value in each model, weighting using the Posterior Model Probability (PMP) of each model. The posterior standard deviation is calculated using the same approach but using the standard deviations in each model. Models 1-5 are the 5 most likely models as assessed by their PMPs, and describe the variables included in those models and their coefficients. The variables g1-g5 are derived from a principal components analysis of the five measures of constitutional type:- *maj*, *magn*, *pind*, *sdm*, and *spropn*.

Table 3: Results Including Botero et al (2004) Legal Origin Variables - BMA Estimates

Variable	Posterior Inclusion Probability	Posterior Expected Value	Posterior Standard Deviation	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	1.000	−0.145	0.305	−0.386	0.200	−0.159	0.198	−0.370
pres	0.999	−0.127	0.033	−0.137	-0.106	−0.136	−0.146	−0.100
frankrom	0.984	0.043	0.016	0.043	0.044	0.047	0.047	0.041
loga	0.716	0.045	0.037	0.072		0.044		0.071
scandi	0.679	0.073	0.064	0.105		0.104	0.103	
mthconst	0.588	0.036	0.036	0.057		0.059	0.076	
con2150	0.512	0.041	0.048	0.084		0.079	0.086	
con81	0.502	0.040	0.046	0.076		0.079	0.069	
africa	0.347	0.028	0.046	0.082				0.085
ilat01	0.225	0.000	0.001		0.003			0.002
con5180	0.096	0.000	0.018					
coluka	0.089	0.005	0.019					
asiae	0.087	−0.005	0.020					
age	0.074	0.008	0.037					
laam	0.036	−0.002	0.013					
g5	0.032	0.002	0.016					
mthelect	0.026	0.001	0.010					
g1	0.015	0.000	0.001					
g4	0.013	0.000	0.004					
common	0.011	0.000	0.005					
federal	0.007	0.000	0.004					
german	0.007	0.000	0.004					
engfrac	0.006	0.000	0.007					
g2	0.000	0.000	0.000					
g3	0.000	0.000	0.000					
eurfrac	0.000	0.000	0.000					
french	0.000	0.000	0.000					
R^2				0.697	0.570	0.672	0.648	0.623
BIC				−39.967	-39.150	−39.141	−39.016	−38.992

As for Table 2.

Table 4: Benchmark Specification - MC^3 Estimates

Variable	Posterior Inclusion Probability	Model 1	Model 2	Model 3	Model 4	Model 5
loga	0.680			✓		✓
pres	0.967	✓	✓	✓	✓	✓
g1	0.081					
g2	0.022					
g3	0.027					
g4	0.049					
age	0.051					
federal	0.202					
ilat01	0.823	✓	✓	✓	✓	✓
engfrac	0.043					
eurfrac	0.043					
frankrom	0.633	✓	✓	✓	✓	
mthconst	0.267					
mthelect	0.058					
coluka	0.542			✓		✓
africa	0.366			✓		✓
laam	0.064					
asiae	0.099					
con2150	0.586		✓	✓		✓
con5180	0.064					
con81	0.142					
Possible Outliers	Posterior Probability of being an outlier					
Ireland	0.811		✓	✓	✓	✓
Honduras	0.576			✓		✓
Turkey	0.120					
Gambia	0.083					
Sri Lanka	0.070					
France	0.054					
Cyprus	0.036					
Germany	0.035					
Ghana	0.030					
Philippines	0.014					
Posterior Model Probability		0.012	0.010	0.009	0.008	0.008

The dependent variable is the mean of the labor share of value-added in manufacturing in the period 1990-1994. Estimates obtained using the MC^3 estimator. The variables are sorted by their posterior inclusion probabilities (PIPs): the sum of the posterior probabilities of those models which include that variable. Models 1-5 are the 5 most likely models as assessed by their PMPs, and describe the variables included, and observations identified as outliers, in those models.

Table 5: OLS Estimates of the models identified by BMA/MC^3 as being most likely

	(1)	(2)	(3)	(4)	(5)	(6)
pres	-0.106^{***}	-0.098^{***}	-0.150^{***}	-0.128^{***}	-0.106^{***}	-0.163^{***}
	(0.028)	(0.029)	(0.029)	(0.033)	(0.028)	(0.029)
frankrom	0.043^{***}	0.039^{**}	0.014	0.018	0.043^{***}	
	(0.013)	(0.015)	(0.022)	(0.021)	(0.013)	
ilat01	0.003^{***}	0.003^{***}		0.003^{***}	0.003^{***}	
	(0.001)	(0.001)		(0.001)	(0.001)	
loga		0.063^{***}	0.066^{**}			0.089^{***}
		(0.022)	(0.028)			(0.026)
africa		0.076^{*}	0.126^{**}			0.096^{**}
		(0.039)	(0.054)			(0.042)
scandi			0.108^{***}			
			(0.039)			
mthconst			0.064^{***}			
			(0.019)			
con2150			0.089			0.093^{**}
			(0.058)			(0.042)
con81			0.043			
			(0.032)			
coluka						0.008
						(0.042)
Constant	0.201^{***}	-0.303	-0.265	0.262^{***}	0.201^{***}	-0.325
	(0.055)	(0.195)	(0.242)	(0.072)	(0.055)	(0.226)
R^2	0.553	0.621	0.712	0.568	0.553	0.615
N	70	65	49	51	70	63

Huber-White (HC3) Robust Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Columns 1 and 2 correspond to Models 1 and 2 in Table 2. Columns 3 and 4 similarly correspond to Models 1 and 2 in Table 3, and Columns 5 and 6, correspond to the Models 1 and 3 in Table 4 with the 6th column excluding observations for Honduras and Ireland, that were identified as outliers by the MC^3 algorithm.

Table 6: Panel-Time Series Integration and Cointegration Tests

		labshare		cgrev		yp	
Unit Root Tests							
H0: Individual Unit Roots							
Im-Pesaran-Shin							
	Level	-1.514	0.065	−3.222	0.001	9.135	1.000
	First Difference	-20.516	0.000	−19.371	0.000	−11.934	0.000
ADF-Fisher							
	Level	-1.537	0.062	−3.250	0.001	8.406	1.000
	First Difference	-17.750	0.000	−17.514	0.000	−10.312	0.000
H0: Common Unit Root							
Breitung							
	Level	3.165	0.999	2.181	0.985	12.310	1.000
	First Difference	-8.505	0.000	−8.003	0.000	5.697	1.000
H0: No Unit Root							
Hadri							
	Level	9.382	0.000	7.875	0.000	15.590	0.000
	First Difference	5.447	0.000	4.725	0.000	5.483	0.000
Cointegration Tests							
H0: No Cointegration							
Pedroni Panel ADF		-3.298	0.000				
Pedroni Group ADF		-2.645	0.004				

All tests conducted using automatic lag-order selction, and individual intercepts, but no individual trends (except the Breitung test which has both). ADF-Fisher test statistics are the Choi Z-statistic.

Table 7: Pooled Mean Group Estimates of Dynamic Model

	Dependent Variable is the Change in the Labor Share		
	(1) Whole Sample	(2) Presidential Democracies	(3) Parliamentary Democracies
$cgrev_t - \overline{cgrev_t}$	0.008***	0.006***	0.002*
	(0.001)	(0.001)	(0.001)
$yp_t - \overline{yp_t}$	0.006**	0.017***	0.010***
	(0.002)	(0.006)	(0.002)
ψ_i	−0.259***	−0.221***	−0.312***
	(0.038)	(0.076)	(0.050)
$\Delta(labshare_{it-1} - \overline{labshare_{t-1}})$	0.019	0.048	0.022
	(0.038)	(0.065)	(0.043)
$\Delta(cgrev_{it-1} - \overline{cgrev_{t-1}})$	0.000	−0.000	0.001
	(0.001)	(0.002)	(0.002)
$\Delta(yp_{it-1} - \overline{yp_{t-1}})$	−0.017	−0.037	−0.011
	(0.012)	(0.030)	(0.009)
$\Delta(trade_{it-1} - \overline{trade_{t-1}})$	−0.002***	−0.002***	−0.002***
	(0.000)	(0.001)	(0.001)
$\Delta(oil_{it-1} - \overline{oil_{t-1}})$	0.001	0.003	0.000
	(0.003)	(0.005)	(0.003)
Constant	−0.002	−0.023***	0.022***
	(0.006)	(0.009)	(0.007)
N	1029	353	676

Standard errors in parentheses. Estimates are obtained using the Pooled Mean Group estimator of Pesaran et al. (1999). ψ is the Short-Run Error Correction Term. $^{*}(p < 0.10)$, $^{**}(p < 0.05)$, $^{***}(p < 0.01)$

References

Acemoglu, Daron, "Constitutions, politics, and economics: A review essay on Persson and Tabellini's The Economic Effects of Constitutions," *Journal of Economic Literature*, 2005, *43* (4), 1025–1048. 48.

_ **and James A. Robinson**, "A theory of political transitions," *American Economic Review*, 2001, *91* (4), 938–963.

_ **, Simon Johnson, and James A. Robinson**, "The colonial origins of comparative development: An empirical investigation," *American Economic Review*, 2001, *91* (5), 1369–1401. 82.

Banerjee, Anindya and Martin Wagner, "Panel Methods to Test for Unit Roots and Cointegration," in Kerry Patterson and Terence C. Mills, eds., *Palgrave Handbook Of Econometrics: Volume 2: Applied Econometrics*, Vol. 2, Palgrave Macmillan, 2009, chapter 13, pp. 632–726.

Boix, Carles, "Democracy, Development, and the Public Sector," *American Journal of Political Science*, 2001, *45* (1), 1–17.

Botero, Juan C., Simeon Djankov, Rafael La Porta, Florencio Lopez-De-Silanes, and Andrei Shleifer, "The Regulation of Labor*," *Quarterly Journal of Economics*, 2004, *119* (4), 1339–1382.

Dulleck, Uwe and Paul Frijters, "Why the US and not Brazil? Old Elites and the Development of a Modern Economy," May 2004. Vienna Economics Papers.

Fudenberg, Drew and Jean Tirole, *Game Theory*, The MIT Press, December 1991.

Gollin, Douglas, "Getting income shares right," *Journal of Political Economy*, 2002, *110* (2), 458–474.

Groningen Growth Development Centre, "The Conference Board Total Economy Database, http://www.conference-board.org/data/economydatabase/," January 2010.

Hall, Robert E. and Charles I. Jones, "Why do some countries produce so much more output per worker than others?," *Quarterly Journal of Economics*, 1999, *114* (1), 83–116. 58.

Hlouskova, Jaroslava and Martin Wagner, "The Performance of Panel Unit Root and Stationarity Tests: Results from a Large Scale Simulation Study," *Econometric Reviews*, June 2006, *25* (1), 85–116.

Hoeting, Jennifer, Adrian E. Raftery, and David Madigan, "A method for simultaneous variable selection and outlier identification in linear regression," *Computational Statistics & Data Analysis*, 1996, *22* (3), 251–270.

_ **, _ , _ , and Chris T. Volinsky**, "Bayesian model averaging: A tutorial," *Statistical Science*, 1999, *14* (4), 382–401.

Kass, Robert E. and Adrian E. Raftery, "Bayes Factors," *Journal of the American Statistical Association*, 1995, *90* (430), 773–795.

Lee, Myoung-Jae, *Micro-Econometrics for Policy, Program and Treatment Effects* Advanced Texts In Econometrics, 1st ed., Oxford, Uk: Oxford University press, 2005.

Malik, Adeel and Jonathan R.W. Temple, "The geography of output volatility," *Journal of Development Economics*, November 2009, *90* (2), 163–178.

Miller, Alan, *Subset Selection in Regression*, 2nd ed., London: CRC Press Inc, 2002.

Pagano, Marco and Paolo F. Volpin, "The Political Economy of Corporate Governance," *American Economic Review*, September 2005, *95* (4), 1005–1030.

Persson, T., G. Roland, and G. Tabellini, "Towards micropolitical foundations of public finance," *European Economic Review*, 1998, *42* (3-5), 685–694. 8.

Persson, Torsten and Guido Tabellini, "The size and scope of government: Comparative politics with rational politicians," *European Economic Review*, April 1999, *43* (4-6), 699–735.

_ **and** _ , *Political Economics: Explaining Economic Policy* Zeuthen Lecture Books, 1st ed., Cambridge, MA: The MIT Press, 2000.

_ **and** _ , *The Economic Effects of Constitutions* Munich Lectures in Economics, 1st ed., Cambridge, MA: MIT Press, 2003.

_ **and** _ , "Constitutional rules and fiscal policy outcomes," *American Economic Review*, 2004, *94* (1), 25–45. 29.

_ **, Gérard Roland, and Guido Tabellini**, "Separation of powers and political accountability," *Quarterly Journal of Economics*, 1997, *112* (4), 1163–1202. 32.

_ **,** _ **, and** _ , "Comparative politics and public finance," *Journal of Political Economy*, 2000, *108* (6), 1121–1161. 50.

Pesaran, M. Hashem, Yongcheol Shin, and Ron P. Smith, "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels," *Journal of the American Statistical Association*, 1999, *94.*

Pickering, Andrew and James Rockey, "Ideology and the Growth of Government," *Review of Economics and Statistics*, Forthcoming.

Raftery, Adrian E., David Madigan, and Jennifer A. Hoeting, "Bayesian model averaging for linear regression models," *Journal of the American Statistical Association*, 1997, *92* (437), 179–191.

Rockey, James, "Reconsidering the Fiscal Effects of Constitutions," Technical Report 10/16, University of Leicester Economics Discussion Papers May 2010. University of Leicester Discussion Paper in Economics 10/17.

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

_ , "Democracies pay higher wages," *Quarterly Journal of Economics*, 1999, *114* (3), 707–738.

Rousseeuw, Peter J., "Least Median of Squares Regression," *Journal of the American Statistical Association*, 1984, *79* (388), 871.

Sala-I-Martin, X., G. Doppelhofer, and R. I. Miller, "Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach," *American Economic Review*, 2004, *94* (4), 813–835.

Temple, Jonathan R. W., "Robustness tests of the augmented Solow model," *Journal of Applied Econometrics*, 1998, *13* (4), 361–375.