Political Institutions as Robust Control: Theory and Application to Economic Growth*

Timothy Besley
LSE and CIFAR

Hannes Mueller
IAE (CSIC), MOVE and Barcelona GSE

July 15, 2015

Abstract

This paper develops a model where an institutional constraint limits incumbent discretion to prevent adverse policy outcomes. We show that, in this framework, executive constraints have an impact on the mean and variance of policy. This allows us to interpret the empirical observation that growth volatility is lower in countries with strong executive constraints. We fit the model to growth data and use our estimates to describe the heterogeneity in performance of weak and strong executive constraints across countries. This is used to illustrate the heterogeneous output response to the adoption of strong executive constraints. A novelty of this paper is to adopt the idea of robust control to institutional comparisons. This supports the idea that strong executive constraints are particularly favored as a means of protecting against bad political outcomes.

*The authors are grateful to audiences at CIFAR meeting, LACEA and the 6th Australian Public Choice Society for comments and suggestions. We thank Heather Sarsons for her help. Mueller acknowledges financial support from the Ramon y Cajal programme and Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2011-0075). All errors are ours.

1 Introduction

How and whether democratic institutions improve policy making and economic performance has been a central debating point in the political economy literature. To that end, the focus has been on the performance of economies cross-sectionally and before and after transitions. There are two dominant features of the literature to date. First, it has tended to focus on aggregated measures of democracy and/or the consequences of elections. Second, it has tended to look at average performance rather than implications for volatility.

This paper's approach is motivated by a combination of theory and facts. From a theoretical perspective, democratic institutions have two main components. The first, and the primary focus in most models, is the contestability of political office. For example, whether a country uses free and fair elections. The second, is a range of institutions which constrain the use of power once acquired. These include the power of the legislature in passing legislation and the codification of laws that need to be obeyed.[1] Both of these dimensions go into the overall democracy score of a country which has been the variable which has dominated empirical work. Our focus here will be on constraints which limit executive discretion. We will develop a model where such constraints affect both the mean and volatility of policy outcomes.

Our main empirical focus is on economic growth. This is motivated by Figure 1 which gives the kernel densities of economic growth when we divide the sample between those countries that have strong and weak executive constraints. The pattern shown here is striking with more upside and downside "risk" associated with low executive constraints. Whether or not this is due to policy rather than other factors, is debatable. However, we will show that a reduction in growth volatility is a feature of countries which make transitions between strong and weak executive constraints where it is more difficult to argue that this will lead to a change in the structure of the economy.

Figure 1 here

The numbers behind Figure 1 are not without consequence when it comes to looking at the downside. They imply that countries with weak executive constraints have a 50% higher chance of negative growth compared to those with strong constraints, are more than twice as likely to have a growth rate below -5% and are over three times more likely to experience a growth rate below -10%. This will be particularly important if institutions are evaluated with more weight on downside risks as with a robust control approach of the kind

[1] The separation of power goes back to Montesquieu (1752). John Stuart Mill described the adoption of executive constraints as follows: *"[...] establishment of constitutional checks, by which the consent of the community, or of a body of some sort, supposed to represent its interests, was made a necessary condition to some of the more important acts of the governing power."* [Mill (1859)]

adopted here.

The theoretical approach developed here offers a natural interpretation of Figure 1 where the effect of executive constraints on average growth and its volatility is driven entirely by changes in the policy making process. Our model also offers a way of thinking about both the cross-sectional variation in growth performance and the within-country volatility. Moreover, we show that there is a straightforward way of fitting the model to the data using non-linear least squares. The parameter estimates that come out of this exercise give us a quantitative assessment of the heterogeneity which is due to politics.

A novelty of this paper is that we adopt the idea of robust control when making institutional comparisons (see, for example, Hansen and Sargent, 2007). We motivate this by invoking the claim that there is a great deal of model uncertainty in understanding the political process making it difficult to think of a known probability distribution over behavior. We therefore employ the maximin expected utility approach of Gilboa and Schmeidler (1989) to making institutional comparisons. To do this we use both data and theory to calculate the maximin of ten years of output growth under both strong and weak executive constraints. Two main findings arise. First, the economic benefits of adopting strong executive constraints crucially depend on whether the choice is made with the supposed knowledge of performance under weak executive constraints or behind a veil of ignorance. Second, in the latter scenario the gain from adopting strong constraints is up to 25% of national income over ten years. We argue that the scale of this gain is driven by an increased risk of sliding into prolonged periods of economic contraction under weak executive constraints.

The ideas in this paper are therefore particularly relevant to the observation that prolonged episodes of negative GDP growth tend to be a feature of countries with weak executive constraints. A good example is Zimbabwe whose economy shrank by over 3% per year on average after constitutional reforms which concentrated powers in the hands of Robert Mugabe. Maintaining strong constraints provide a form of robust control of politicians which limits the frequency of such episodes. The paper gives a theoretical foundation to this argument and supports it with evidence.

The remainder of the paper is organized as follows. In the next section, we discuss some background issues including relevant literature. Section three introduces the model while section four uses it to interpret cross-country growth experiences. In section five, we formulate a way of thinking about institutional comparisons and extend the growth application to this. Section six concludes.

2 Background

Relationship to the Literature There is now a large empirical literature on the link between democracy and growth such as Barro (1996), Papaioannou and Siourounis (2008), Persson and Tabellini (2009a,b), and Przeworski and Limongi (1993). The empirical findings are somewhat equivocal. Moreover, the links between the empirical approach and theoretical models is limited. By now, it is recognized that there is no simple empirical story to be told and that there could be considerable heterogeneity as discussed in Persson and Tabellini (2009b).

Also relevant to this paper is the literature on economic volatility in emerging market economies. Aguiar and Gopinath (2007), for example, observe that shocks to trend growth—rather than transitory fluctuations around a stable trend—are the primary source of fluctuations in emerging markets. This is consistent with the idea developed here that political factors affect growth trends. Koren and Tenreyro (2007) separate growth volatility at a the country level from sector-specific volatility and find that, as countries develop, changes in the sector composition reduce volatility. The paper is also related to the observation by Calvo (1998) that "sudden stops" in capital flows occur in countries because there is policy flexibility; local governments are more constrained in their policy choices creating less policy risk. The role of institutional factors in volatility in emerging markets is emphasized in Rodrik (1999).

This is not the first paper to note that there is a difference in variance of performance in countries which are democratic – see, for example, Acemoglu et al (2003), Almeida and Ferraira, (2002), Henisz (2004), Moborak (2005), Rodrik (1997) and Weede (1996). However, we believe that this is the first paper which tries to tie this fact to an underlying theory of how policy making varies with political institutions.[2]

Theoretical approaches to the relationship between democracy and outcomes follow two main lines. The influential work of Acemoglu and Robinson (2006) focuses on shifts in political control between income groups and how this impacts the economy through redistribution policies. Such models do not have a role for executive constraints with the focus being on how elections affect access to power. This paper uses an agency model of politics of the kind introduced in Barro (1973) and Ferejohn (1986). We supplement the traditional focus of agency models on the external control imposed by voters by considering the behavior of a legislature which can act as an internal control on policy makers.[3] Interestingly, our study suggests that this external control adds stability even within open regimes.

Our theoretical approach is related to Persson et al (1997). They provide an explanation for why the separation of powers improves the accountability of elected officials. A key condition to make separation of powers work in favor of voters is that no policy can be implemented unilaterally, i.e., without the consent of both the executive and the legislature. This relates more generally to the differences between parliamentary and presidential systems as discussed both theoretically and empirically in Persson and Tabellini (2000).

[2] Overbye (1996) discusses some possible theoretical underpinnings for this based on the predictability of policies but does not relate this back to the data.

[3] See Besley (2006) for a discussion of the literature and Besley and Kudamatsu (2009) for a model which focuses on a different model of internal controls in the absence of elections.

Executive Constraints The standard approach in much of the empirical literature on democracy is to work on with some kind of aggregate democracy score.[4] This aggregates components which includes constraints on the executive and the process for acquiring power, i.e. whether there are elections. Here we will focus on a specific dimension of democratic institutions – the imposition of constraints on the executive. Such constraints operate by changing the amount of discretion enjoyed by policy makers while holding office. Thus, some policy decisions might be prevented or overruled.

As a measure of strong executive constraints we use the variable "xconst" in the Polity IV dataset which is on a seven point scale. The manual explains the variable's construction as follows:

> "Operationally, this variable refers to the extent of institutionalized constraints on the decision making powers of chief executives, whether individuals or collectivities. Such limitations may be imposed by any "accountability groups." In Western democracies these are usually legislatures. Other kinds of accountability groups are the ruling party in a one-party state; councils of nobles or powerful advisors in monarchies; the military in coup-prone polities; and in many states a strong, independent judiciary. The concern is therefore with the checks and balances between the various parts of the decision-making process." [p. 24, Polity IV Dataset Users' Manual 2010]

We will create a dummy variable which is equal to one when the score is seven which, in a nutshell, corresponds to a political system where there is "Executive Parity or Subordination: Accountability groups have effective authority equal to or greater than the executive in most areas of activity."[5] The checklist for coders in the Polity IV manual states that the highest score of the variable "xconst" is only allocated if important legislation can be initiated by a parliament which holds the executive to account. Indeed, our reading of the country reports is that those coding countries pay a lot of attention to whether the executive relied on support from another organization (this could be, parliament, independent courts or the military) to conduct policy.

We believe that there are good reasons for disaggregating democracy and focusing on executive constraints separately. The first argument is theoretical and is developed below – the logic of using internal constraints and external political control through elections is somewhat different. Greater openness and competitiveness through free and fair elections allows citizens to remove poorly performing incumbents and create performance-related retention decisions. Executive constraints can affect policy choices apart from

[4] That said, there are a number of papers which look at executive constraints separately.

[5] See http://www.systemicpeace.org/ inscr/p4manualv2010.pdf. Examples of evidence used to assign a score of 7 are (i) A legislature, ruling party, or council of nobles initiates much or most important legislation (ii) The executive (president, premier, king, cabinet, council) is chosen by the accountability group and is dependent on its continued support to remain in office (as in most parliamentary systems). (iii) In multi-party democracies, there is chronic "cabinet instability."

the electoral cycle.

But there is a second empirical argument for the focus. There are a number of countries in the data that are fully open according to polity but do not have strong executive constraints.[6] For example, of the 220 countries in the PolityIV data in 2000, 122 countries have the highest score for openness but only 50 have the highest score for executive constraints. All of these 50 countries have the highest score for openness. So looking at these dimensions separately leads us to classify countries somewhat differently.[7] The fact that all countries which have strong constraints are also open according to PolityIV means that we will build a model where citizens can remove their rulers from office and focus on the consequences of varying executive constraints.

As a final motivation for the focus on executive constraints, it is worth looking at a version of Figure 1 where we look instead at the way that having open executive recruitment affects the distribution of growth rates. This is given in Figure 2 for the sample of country years where executive constraints are weak. The distribution with full openness shifts to the right and has a higher mean.[8] The reduction in the variance is much less pronounced in this case.

Figure 2 here

To further challenge our core definition of executive constraints, we consider two different robustness checks. First, we include only observations that have the highest score of openness and study changes in the growth paths with the adoption of strong executive constraints. In this way we hold other institutions fixed to show that the effect of executive constraints remains.[9] Second, we use an alternative measure of executive constraints which measures the degree to which the chief executive faces checks in the legislature. This was first operationalized by Beck et al (2001) and updated by Keefer (2012). We define strong constraints if checks are larger than 3. In parliamentary systems this level of the score is reached, for example, if the chief executive is elected and has a coalition partner which is needed to maintain a majority. We choose this level so that the share of country/years observations under strong constraints is as close as possible to our main definition.[10] Appendix Table A1 shows that the two resulting measures of strong constraints are correlated but far from identical. We have almost a quarter of country/year observations in which the two measures deviate. For example, the chief executive in Russia is coded as facing strong checks by Keefer (2012) after 1995 but coded as never facing strong executive constraints by Polity IV. Below, we report robustness checks

[6] Full openness is defined as the variable "xropen" in the PolityIV data taking its maximum possible value of 4.

[7] In 2000, there are 114 countries in the PolityIV data which are classified as democratic based on their aggregate democracy score but which do not have have the highest score for executive constraints.

[8] This difference is statistically significant and robust to controlling for country and year fixed effects.

[9] We could also focus on democracies, i.e. observations with a polity score higher than 0.

[10] Our measure is based on the checks_lax variable in Keefer (2012). Details are discussed in the appendix.

using these two alternative definitions.

Disaggregating Growth Performance We now disaggregate growth variation into its between country and within-country variation differentiated by the strength of executive constraints. As well as providing further understanding of the numbers behind Figure 1, it also allows us to think about how we might interpret the variation that we see in the data using a political economy model. Growth is calculated from real GDP per capita provided by the World Penn Tables 7.1.[11]

Countries spend different amounts of time with high or low executive constraints as measured in the PolityIV data. Let $\rho_{\delta c}$ be the proportion of observations in the data where country c has executive constraints $\delta \in \{S, W\}$ where 'S' is strong and 'W' is weak. Let $\mu_c(\delta)$ be country c's average growth in when it has institutions δ. Then the average growth rate for $\delta \in \{S, W\}$ is:

$$\bar{\mu}(\delta) = \sum_c \rho_{\delta c} \mu_c(\delta)$$

and we denote mean growth in the sample as a whole by $\bar{\mu}$.

Now let $\sigma_c^2(\delta)$ be the variance of growth in country c when it has institutions δ. The average difference in the variance of growth between $\delta = W$ and $\delta = S$ is then given by

$$\bar{\sigma}^2(S) - \bar{\sigma}^2(W) = \sum_c \left[\rho_{Sc}\sigma_c^2(S) - \rho_{Wc}\sigma_c^2(W)\right] + \sum_c \left[\rho_{Sc}\left[\mu_c(S) - \bar{\mu}\right]^2 - \rho_{Wc}\left[\mu_c(W) - \bar{\mu}\right]^2\right].$$

This decomposes growth into changes in within-country volatility in growth and a between-country component to the extent that there are differences in the dispersion of growth rates across countries. Figure 1 showed that the overall variance falls when strong executive constraints are adopted but not whether this is primarily a between or within-country change.

In fact both within and between-country dispersion are lower under strong executive constraints. The easiest way to see is this is by running the following growth regression where the dependent variable is the growth rate in country c with institutions δ in year t:

$$g_{\delta ct} = \alpha_{\delta c} + \alpha_t + \eta_{\delta ct} \tag{1}$$

where α_t are year dummy variables and $\alpha_{\delta c}$ is a country/institution fixed effect that captures mean growth

[11] See the Appendix for more details on the data and sample.

for country c under institution δ. This specification also controls for common global shocks that affect growth. The variance of the residuals $\eta_{\delta ct}$ by country for a period when institutions are δ gives us a measure of within-country growth volatility.[12] In other words, our decomposition allows us to look at the across country variances as well as the within country variances.

Figures 3 and 4 plot the kernel densities of the estimates of $\alpha_{\delta c}$ and $\eta_{\delta ct}$ from running this on our entire sample of countries. Figure 3 plots our estimates of $\eta_{\delta ct}$ for both strong and weak executive constraints. The reduction in the variance of growth within country is once again striking. Figure 4 plots the estimated $\alpha_{\delta c}$. Due to the lack of a time dimension there is a lot less data on which to estimate the between-country patterns. But the reduction in between-country variation is also apparent.

Figures 3 and 4 here

Table 1 gives the mean and variance of growth computed from equation (1) for two samples of countries and our three definitions of strong executive constraints. Panel A provides data for all countries in the sample and shows that there is a fall in mean growth with weak executive constraints. However, this fall is not very large, i.e. between 0.1 and 0.5 percentage points. The within country standard deviation, i.e. the standard deviation of $\eta_{\delta ct}$, increases markedly by around 3 percentage points when a country has weak executive constraints. The between country standard deviation, i.e. the standard deviation of $\alpha_{\delta c}$, also rises but by less than 1 percentage point. Given these standard deviations the mean growth differences across regimes are not significant.

We can also look at this for countries that switched into or away from strong executive constraints during our sample period. For the purposes of comparing their performance, we look at those which have spent a minimum of three years with strong and weak executive constraints. In this sample, it is more plausible to argue that changes in volatility are driven, in significant measure, by political factors rather than structural changes in the economy.[13] Results are in Panel B of Table 1. The pattern of changes in the mean and standard deviation of growth is similar to what we found for the full sample. The mean growth rate falls to a similar extent. The within-country standard deviation increases in all three samples by about 1.5 percentage points and the between country standard deviation by about half a percentage point. The take-away from Table 1 is that a reduction of between and within variances can be observed in six different

[12] We could also use a ARCH or GARCH model to capture persistences in the within variance. Henisz (2004) uses these models to provide estimates of policy uncertainty. However, our theory suggests that variance is regime-specific which is not captured by these models.

[13] As an additional test we also derive Table 1 controlling for the level of GDPpc in the regression (1). Results are reported in Table A2 which shows extremely similar patterns.

samples and using two different definitions of strong executive constraints.

3 The Model

This section lays out a framework for thinking about how executive constraints affects policy incentives. The model will show why both the mean and variance of policy outcomes vary with executive constraints.[14]

3.1 Framework

Policy-Making We use a simple model which citizens delegate policy-making power to a policy maker whom we refer as "the executive". The question is whether such delegated authority is used in the interest of citizens and how this is influenced by institutions.

Time is infinite and is denoted by $t = 1, 2, ...$ In each period, the executive faces a policy challenge and must pick from among three possible actions $e \in \{0, 1, q\}$ where q stands for "status quo". The payoff to voters is:

$$u_t = \begin{cases} \Delta_H & \text{if } e_t = s_t \\ \Delta_0 & \text{if } e_t = q \\ \Delta_L & \text{if } e_t \neq s_t. \end{cases}$$

where $\{\Delta_H, \Delta_0, \Delta_L\}$ are public information and $s_t \in \{0, 1\}$ is observed only by the policy maker. Ex-ante $s_t = 1$ with probability $\frac{1}{2}$. These payoffs satisfy the following restrictions:

$$\Delta_H \geq \Delta_0 \ \geq \Delta_L \text{ and } \frac{\Delta_H + \Delta_L}{2} < \Delta_0.$$

Thus, there is a good policy choice and a bad policy choice with the status quo policy lying between the two. The latter condition says that it is never worthwhile for an uninformed policy maker to randomize over $e \in \{0, 1\}$ rather than choosing q.

The model can be interpreted as representing two broad kinds of policy-making. With *policy activism*, the outcome is $e_t \in \{0, 1\}$. With constrained policy-making $e_t = q$, i.e. the incumbent is compelled to stick with the status quo. Whether policy activism benefits voters depends on matching the action to the state.

The Executive We assume that there is large set of potential policy makers who can be picked to be the executive. These policy makers are of two types. A small fraction π of policy makers always chooses $e = s$. We will refer to this type as good. The remaining $1 - \pi$ of policy makers are susceptible to

[14] It is based on Besley (2006, Chapter 3).

misbehaving, being tempted to pick $e \neq s$ and we model this temptation as drawn each period on $[R_L, R_H]$ with distribution function $G(r)$. We refer to them as opportunistic. We assume that $R_H > 0$ but allow for the possibility that $R_L < 0$ so that temptation can sometimes reinforce doing what voters want. Let μ denote the mean of r and assume that $\mu > 0$ so that on average temptation is counterproductive from a voter perspective. We assume that r is drawn each period and is *iid*. We have assumed for convenience that r is earned only with policy activism. One interpretation is to think of r as earning a rent relative to the status quo.

All agents discount the future with the same discount factor β. Let $e_t \in \{0, 1, q\}$ be the action taken by the incumbent. In what follows, we will focus on the case where π is very small. This corresponds to a somewhat pessimistic view.[15] This will give the least incentive for the legislature to give discretion to the incumbent.

Executive Constraints Every period, voters can choose whether to retain the incumbent executive or replace her with a randomly selected alternative who is good with probability π. This is the standard external control on politicians emphasized throughout the political agency literature as in Barro (1973) and Ferejohn (1986).

We add to this the possibility of an internal control in the form of an executive constraint imposed by a legislature which acts in the interests of voters and hence can help to solve the political agency problem. The role of executive constraints is to curtail some instances of bad policy making in the spirit of the veto players model of Tsebelis (2002). This constraint works ex ante, i.e. before the policy decision has been taken. We measure the strength of executive constraints by a parameter $\xi \in [0, 1]$ and suppose that such constraints are active with probability $1 - \xi$. Hence, a higher value of ξ represents weaker constraints. Let $x_t \in \{0, 1\}$ denote active executive constraints, i.e. whether the legislature can reduce incumbent discretion in period t.

We assume that the legislature can observe r_t but not s_t. The inability of the legislature to observe the state determining optimal policies will mean that it can only work as an imperfect disciplining device. Having observed r_t, we allow the legislature to remove the discretion of the incumbent and impose policy q. An executive constraint is binding whenever $x_t = 1$ and $e_t = q$.

We will allow the voter and legislature to update their beliefs about the quality of the executive over time and let Π_t be the "reputation" of the executive at date t, i.e. the belief that the incumbent is good.

[15] It follows a long tradition which goes back at least to Hume (1742). It is also a postulate which Buchanan (1989) has argued strongly in favor of.

Timing Within each period t, the timing is as follows:

1. An incumbent is in place with reputation Π_t.
2. Nature determines $\{r_t, s_t, x_t\}$.
3. If $x_t = 0$ then the incumbent chooses $e_t \in \{0, 1, q\}$
4. If $x_t = 1$ then the legislature can choose between imposing $e_t = q$ and letting the incumbent choose from $e_t \in \{0, 1, q\}$.
5. Payoffs $u_t \in \{\Delta_0, \Delta_L, \Delta_H\}$ are realized and citizens form beliefs about the quality of their incumbent denoted by Π_{t+1}.
6. Citizens choose whether to retain the existing incumbent ($\phi(u_t) = 1$) or pick a new one from the pool ($\phi(u_t) = 0$). The new candidate is good with probability π.

We will look for a stationary perfect Bayesian equilibrium of the game in which citizens optimally make their retention decisions and the executive optimizes its policy choice. If executive constraints are in force, then the legislature optimizes its decision whether to grant discretion to the executive.

We focus on the case where $\pi \to 0$, i.e. almost all policy makers are opportunistic. Having the possibility of good policy makers in the model still plays a role in the equilibrium since it allows us to use Bayes rule on types conditional on seeing a particular outcome. Most importantly, it gives the voters a strict preference for re-appointing any policy maker who has made the right choice if they are granted policy discretion.

3.2 Equilibrium

The equilibrium of the model has three parts which we solve for working backwards. First, we solve for the optimal retention decision made by voters at stage 6. We then solve for the optimal action of the legislature if $x_t = 1$, which is stage 4 above and finally we solve for the behavior of the incumbent if he is granted discretion.

The stationary equilibrium of the model has the following form. There is a value $\hat{r}(\xi)$ such that incumbents who are given discretion choose $e_t = s_t$ for $r_t \leq \hat{r}(\xi)$ and $e_t = 1 - s_t$ otherwise. If executive constraints are in place, the legislature will remove incumbent discretion, i.e. impose policy q, if $r_t > \hat{r}(\xi)$ and allow it otherwise. If voters are given the opportunity, then they remove any incumbent who has misused their discretion, i.e. generated payoff Δ_L. Otherwise the incumbent is retained.

The key behavior in the equilibrium of the model is therefore summarized in

$$\hat{\lambda}(\xi) = G(\hat{r}(\xi))$$

which gives the probability that the bad type incumbents choose the policy which voters want if they have the discretion to do so.

The Behavior of Voters Voters condition their behavior on the observed payoff generated by an incumbent $u_t \in \{\Delta_0, \Delta_L, \Delta_H\}$. In making their decisions, they use Bayes rules to update their beliefs about the incumbent's type. Thus beliefs evolve according to:

$$\Pi_{t+1}(u_t, \Pi_t) = \begin{cases} \frac{\Pi_t}{\Pi_t + (1-\Pi_t)\lambda(\xi)} & \text{if } u_t = \Delta_H \\ \Pi_t & \text{if } u_t = \Delta_0 \\ 0 & \text{if } u_t = \Delta_L. \end{cases}$$

In a voting equilibrium, voters choose to retain an incumbent if the future value of re-election is greater than the value of removing the incumbent based on the expected future stream of policy benefits. The Appendix shows that the voting equilibrium is as follows:

Proposition 1 *In a voting equilibrium,* $\phi(\Delta_L) = 0$ *and* $\phi(\Delta_0) = \phi(\Delta_H) = 1$.

This result makes intuitive sense. Any incumbent that has produced Δ_L must be bad and hence it is worthwhile for voters to return to the pool even if the probability that the new incumbent is good is small. Any incumbent that has produced Δ_H is more likely to be good than a randomly selected incumbent and hence is worth retaining. Nothing is learned about the incumbent's type when Δ_0 is chosen. If the incumbent has generated Δ_H in the past it is strictly better to keep her than return to the pool and, if this is her first term in office, then the voters are indifferent between retaining her and going back to the pool so $\phi(\Delta_0) = 1$ is (weakly) optimal in such cases.

The Behavior of the Legislature We now study the behavior of the legislature and their decision to impose policy q. Suppose that the legislature conjectures that there is a critical value $\hat{r}(\xi) \in [R_L, R_H]$ such that the behavior of the executive is as follows:

$$e_t = \begin{cases} s_t & \text{if } r_t \leq \hat{r}(\xi) \\ 1 - s_t & \text{otherwise.} \end{cases}$$

We will show in the next section that this is indeed the case and characterize $\hat{r}(\xi)$. This implies that an incumbent with discretion chooses the action which gives the voters their largest payoff only if the realization of r_t is low enough.

If $x_t = 0$, then the legislature has no decision to make. If $x_t = 1$, then the legislature can decide after observing r_t whether to impose q or allow the incumbent to choose their preferred policy. Given the conjectured behavior of the executive, the legislature knows that the executive will choose the voter's optimal discretionary policy if r_t is low enough. Otherwise, the wrong policy will be chosen (as $\pi \to 0$). Hence, the status quo policy will be chosen whenever r_t is high. This result, which is proven in the Appendix, is stated as:

Proposition 2 *Suppose that $x_t = 1$, then as $\pi \to 0$, the legislature imposes $e_t = q$ if and only if $r_t > \hat{r}(\xi)$.*

This result shows that, even if constraining the executive is possible, it is not optimal to use that constraint in every case. It makes sense only when the legislature believes that the executive would stray away from the best policy in the event that discretion is granted.

The Behavior of the Executive Finally, we turn to the behavior of the executive given the voting equilibrium in Proposition 1 and behavior of the legislature in Proposition 2. We will characterize the threshold $\hat{r}$ below which the incumbent chooses $e_t = s_t$ if they are granted discretion.

The executive observes s_t and the realization of r_t. If it is given discretion over policy, and succumbs to temptation, it earns r_t but is then removed from office. If it decides to use the discretion to generate Δ_H, then it survives. The value of being retained is given by the expected future rents that it might earn given the likelihood that executive constraints are effective and that it is given discretion. Solving for this value, we can compute the threshold below which discretion is used in the interest of voters. This is given in the following result:

Proposition 3 *There exists a threshold value of temptation denoted by $\hat{r}(\xi) \in [R_L, R_H]$ which is increasing in ξ and solves*

$$\hat{r}(\xi) = \frac{\beta\xi (1 - G(\hat{r}(\xi))) E(r : r \geq \hat{r}(\xi))}{1 - [1 - \xi (1 - G(\hat{r}))] \beta}$$

below which an incumbent with discretion chooses $e_t = s_t$.

The threshold is increasing in ξ. Thus a greater prospect of discretion (a lower likelihood of executive constraints being imposed) increases the ex ante probability that the executive will choose the voter's preferred policy if she is granted discretion. This is due to the fact that discretion yields the possibility of future rents which increase the incentive for good behavior in the present.

Proposition 3 highlights an interesting side-effect of executive constraints. Executive constraints attempt to improve policy outcomes by restricting the choice of the executive. However, through their impact on future payoffs, executive constraints make the executive behave worse in the present.

Discussion The model delivers an insight into how executive constraints affect policy. By allowing the possibility of imposing Δ_0 when Δ_L would have been chosen by the bad incumbent reduces some bad outcomes compared to full discretion. However, because $\hat{r}(\xi)$ is lower when there are executive constrains, there is a deterioration in behavior ex ante. Thus, whether executive constraints improves policy making is unclear a priori. It depends on the relevant parameters and the objective function.

The model predicts that institutions affect the mean and volatility of citizens' payoffs. To see this observe that average voter welfare is

$$G\left(\hat{r}\left(\xi\right)\right)\Delta_H+\left[1-G\left(\hat{r}\left(\xi\right)\right)\right]\tilde{\Delta}_\xi \tag{2}$$

where $\tilde{\Delta}_\xi=\left[1-\xi\right]\Delta_0+\xi\Delta_L$ and the variance of policy outcomes is

$$G\left(\hat{r}\left(\xi\right)\right)\left[1-G\left(\hat{r}\left(\xi\right)\right)\right]\left(\Delta_H-\tilde{\Delta}_\xi\right)^2. \tag{3}$$

Both are functions of ξ. We can get some insight into the trade-offs predicted by the model by differentiating (2) and (3) with respect to ξ. There are basically two effects to consider in each case: the first direct effect comes from the change in $\tilde{\Delta}_\xi$ and the second from the change in $\hat{r}\left(\xi\right)$. For both the mean and standard deviation these effects can work against each other making it an empirical question whether the mean and variance of performance increase or decrease with executive constraints.

The effect on the mean from a change in ξ is given by

$$g\left(\hat{r}\left(\xi\right)\right)\left[\Delta_H-\tilde{\Delta}_\xi\right]\frac{\partial\hat{r}\left(\xi\right)}{\partial\xi}-\left[1-G\left(\hat{r}\left(\xi\right)\right)\right]\left[\Delta_0-\Delta_L\right]. \tag{4}$$

The first term represents the fact that there is better incumbent behavior if there is more discretion, i.e., weaker constraints. This tends to increase the mean outcome. The second term is negative and is due the increased downside risk that having greater discretion imposes.

The effect of an increase in ξ on the variance is given by:

$$\left[g\left(\hat{r}\left(\xi\right)\right)\left[1-2G\left(\hat{r}\left(\xi\right)\right)\right]\left(\Delta_H-\tilde{\Delta}_\xi\right)^2\right]\frac{\partial\hat{r}\left(\xi\right)}{\partial\xi}-G\left(\hat{r}\left(\xi\right)\right)\left[1-G\left(\hat{r}\left(\xi\right)\right)\right]2\left[\Delta_0-\Delta_L\right]. \tag{5}$$

The effect of an increase in discipline on the variance is ambiguous in sign. The second term is always negative. It captures the fact that executive constraints impose a lower downside policy risk, i.e. due to Δ_L being less than Δ_0. Overall, if this difference is large enough or $G\left(\hat{r}\left(\xi\right)\right)\leq 1/2$, then the variance is unambiguously lower when executive constraints are strengthened. We will show that this is the empirically relevant case.

The theory that we have presented has focused on the effect of executive constraints. It would be straightforward to extend the model to allow for varying degrees of contestability for office. This affects how far the threat of removal of an executive who produces Δ_L is real. If producing bad performance does not lead to removal then this will tend to lower $\hat{\lambda}\left(\xi\right)$ for any level of executive constraints. This too will affect the mean and variance of policy outcomes. Changing contestability will generate terms analogous to the first of the two terms in (4) and (5). Reducing contestability will definitely reduce mean performance but has an ambiguous effect on variance depending on whether $\hat{\lambda}(\xi)$ is greater than or less than a half. Hence if we think of an overall democracy index combining executive constraints and contestability, we would not necessarily expect different predictions from contestability. However, the prediction on the mean level of performance should be more clear cut in that case. This is in line with what we found in Figure 2.

4 Application to Economic Growth

In this section, we apply the model to the facts about growth that we discussed above. We will discuss how to fit the model parameters to the data in the case where the policy decisions made by government affect economic growth. In the next section, we will use this to set up a robust comparison of economic performance with and without strong executive constraints.

We work with a growth model where policy affects labor productivity growth. This allows us to forge a link between the theory and data on the mean and variance of realized growth rates. For the purposes of fitting the model we will focus on two extreme cases: either $\xi = 1$ which correspond to weak executive constraints denoted by $\delta = W$ or $\xi = 0$ which correspond to strong executive constraints, $\delta = S$.

Model Structure Consider an open economy where the aggregate production function in country c at date t is

$$Y_{ct} = (\Gamma_{ct} L_c)^{\alpha} (K_{ct})^{1-\alpha}$$

where Γ_{ct} is labor productivity. The latter is assumed to depend on country-level economic policies along the lines articulated by Aghion and Howitt (2006) and evolves stochastically over time according to the following equation:

$$\Gamma_{ct} = \Gamma_{ct-1} e^{\rho_{ct}(\delta)}$$

where $\rho_{ct}(\delta) = \kappa_c(\delta) + \varepsilon_{ct}$ and $\varepsilon_{ct} \sim N\left(-\frac{1}{2}\sigma^2_{\varepsilon c}(\delta), \sigma^2_{\omega c}(\delta)\right)$ with $\delta \in \{S, W\}$ denoting whether executive constraints are strong or weak. This formulation implies that $E\left(e^{\rho_{ct}(\delta)}\right) = e^{\kappa(\delta)}$, i.e. $E\left(e^{\varepsilon}\right) = 1$.

Firms hire capital and labor in competitive factor markets. The labor market is closed with a fixed stock of labor L_c while capital is available on a global capital market at price r. We assume that capital is chosen before ε_{ct} is realized so that risk matters to firms.

Then firms choose their optimal capital stock as follows:

$$\begin{aligned} K^*_{ct}(\delta) &= \arg\max \left\{ E\left(e^{\kappa(\delta)+\varepsilon_{ct}}\right)^{\alpha} (\Gamma_{ct-1} L_c)^{\alpha} K^{1-\alpha} - rK \right\} \\ &= \Gamma_{ct-1} L_c \left(\frac{(1-\alpha)}{r}\right)^{\frac{1}{\alpha}} e^{\kappa(\delta) - \frac{1-\alpha}{2}\sigma^2_{\varepsilon c}(\delta)}. \end{aligned}$$

This implies that that the (log of) income per capita, y_{ct}, is given by:

$$\log y_{ct} = [\kappa_c(\delta) + \alpha\varepsilon_{ct} + \log(\Gamma_{ct-1})] - \frac{(1-\alpha)^2}{2}\sigma^2_{\varepsilon c}(\delta) \;\; + B_c$$

where B_c is a time-invariant constant. Using this, the growth rate at date t in country c is $g_{ct} = \kappa_c(\delta) + \alpha\varepsilon_{ct} + (1-\alpha)\varepsilon_{ct-1}$.[16] Note that this depends on executive constraints through $\kappa(\delta)$. The implied mean and variance of growth are given by:

$$\mu_c(\delta) = \kappa_c(\delta) - \frac{1}{2}\sigma^2_{\varepsilon c}(\delta) \tag{6}$$

and

$$\sigma^2_{gc}(\delta) = \sigma^2_{\varepsilon c}(\delta). \tag{7}$$

This allows us to map between moments in the growth data and the parameters of the economic model. In particular, we will use the relationship in equations (6) and (7) to calculate country-specific estimates of the productivity growth trend $\kappa_c(\delta)$ and variance $\sigma^2_{\varepsilon c}(\delta)$ from the moments in the growth data. All we need to do is to use mean growth in an episode as $\mu_c(\delta)$ and the variance of growth in a country/institution episode as $\sigma^2_{gc}(\delta)$. We can then calculate $\kappa_c(\delta)$ and variance $\sigma^2_{\varepsilon c}(\delta)$ from equations (6) and (7).

To map this onto the political model we assume that the productivity growth trend is $\kappa_c(\delta) = \lambda_{\delta c}\Delta_H + (1-\lambda_{\delta c})\tilde{\Delta}_\delta$ where

$$\tilde{\Delta}_\delta = \begin{cases} \Delta_L & \text{if } \delta = W \\ \Delta_0 & \text{if } \delta = S, \end{cases}$$

$\lambda_{Sc} = G(\hat{r}(0))$ and $\lambda_{Wc} = G(\hat{r}(1))$. Then if $\Delta_{ct} \in \{\Delta_H, \Delta_L, \Delta_0\}$ is the realized value of Δ in country c at date t, assume $\varepsilon_{ct} = \Delta_{ct} - \kappa_c(\delta) + \omega_{ct}$ where $\omega_{ct} \sim N\left(-\frac{\sigma^2_{\varepsilon c}(\delta)}{2}, \sigma^2_{\omega c}\right)$ which implies that $\sigma^2_{c\varepsilon}(\delta) = \lambda_{\delta c}(1-\lambda_{\delta c})\left[\Delta_H - \tilde{\Delta}_\xi\right]^2 + \sigma^2_{\omega c}$. Thus the trend productivity growth rate and the standard deviation around the trend vary with executive constraints.

Fitting the Model In order to fit the model to the data we will take the following approach. First, we assume a common set of parameters Δ_H, Δ_0 and Δ_L across countries. We then calculate a set of country/institution-specific values of $\lambda_{\delta c}$. Thus the model is applied under the assumption that differences in productivity growth between countries are due to differences in politics as represented by $\lambda_{\delta c}$.

[16] Growth is given by $\log y_{ct} - \log y_{ct-1} = [\kappa_c(\delta) + \alpha\varepsilon_{ct} + \log(\Gamma_{ct-1})] - [\kappa_c(\delta) + \alpha\varepsilon_{ct-1} + \log(\Gamma_{ct-2})]$

First observe that by the definition of $\kappa_c(\delta)$ we can write

$$\lambda_{\delta c} = \frac{\kappa_c(\delta) - \tilde{\Delta}_\delta}{\Delta_H - \tilde{\Delta}_\delta}. \tag{8}$$

Substituting this into the expression for the variance of growth, we have that

$$\sigma_{gc}^2(\delta) = \left(\kappa_c(\delta) - \tilde{\Delta}_\delta\right)(\Delta_H - \kappa_c(\delta)) + \sigma_{\omega c}^2 \tag{9}$$

where $\kappa_c(\delta) = \mu_c(\delta) + \frac{1}{2}\sigma_{gc}^2(\delta)$. Thus, the variance is a non-linear function of the trend growth rate and the common parameters $\{\Delta_H, \Delta_L, \Delta_0\}$. We can use this to fit the model to the data using non-linear least squares to estimate (9).[17] We run the regression jointly for strong and weak executive constraints imposing the condition from the theory that Δ_H is common across the two samples.[18]

We first estimate $\left\{\hat{\Delta}_H, \hat{\Delta}_L, \hat{\Delta}_0\right\}$ for all countries in the data. Second, we look at only those countries that switched in or out of strong executive constraints during our sample period and which spent at least three years in each regime. The results are shown in Table 2. The point estimates line up as we would expect with $\hat{\Delta}_H > \hat{\Delta}_0 > \hat{\Delta}_L$. This is a key implication of the theory and is a key part of the mechanism behind the lower variance in strong executive constraints. In the restricted sample of switchers we find a significant estimate with $\hat{\Delta}_H = 8.4\%$, a value of $\hat{\Delta}_0 = -1.1\%$ and $\hat{\Delta}_L = -3.2\%$. The estimated magnitude of $\Delta_0 - \Delta_L = 2.1\%$ ties into the predictions of our theoretical model as it underpins the "insurance" effect of strong executive constraints in equation (5). In all our estimates this key magnitude is both economically and statistically significant.[19]

Using equation (8), we can now back out country/regime-specific estimates $\hat{\lambda}_{\delta c}$ using our estimates $\left\{\hat{\Delta}_H, \hat{\Delta}_L, \hat{\Delta}_0\right\}$.[20] We interpret heterogeneity in $\hat{\lambda}_{\delta c}$ as a country-specific political equilibrium depending upon whether executive constraints are strong or weak. Returning to the model, this heterogeneity in country performance could be thought of as being due to different distributions $G(\cdot)$ for each country. While this is a stylized way of fitting the model, given that we have imposed common values for $\left\{\hat{\Delta}_H, \hat{\Delta}_L, \hat{\Delta}_0\right\}$, it gives an interpretation of the moments in the growth data entirely through variations in politics. This pattern is remarkably robust to the sample and definitions we use.

We can now summarize the fruits of this estimation by looking at the distribution of our estimates of

[17] In this we treat $\sigma_{\omega c}^2$ as a country and institution specific error term.

[18] We also used a more flexible specification in which we allowed the Δ_H to be a function of the regime. This exercise confirms that the there are no significant differences in Δ_H between strong and weak executive constraints.

[19] See appendix Table A3 in which we use different samples, definitions of strong constraints or controlling for GDP per capita and year fixed effects. We always find the ordering $\Delta_H > \Delta_0 > \Delta_L$ and a significant difference $\Delta_0 - \Delta_L$ of similar magnitude.

[20] For details and an example see the Appendix B.

$\hat{\lambda}_{\delta c}$. One way to assess the fit to the model is to see whether the prediction that $\hat{\lambda}_{Wc} > \hat{\lambda}_{Sc}$ (Proposition 3) holds in the data. Figure 5 gives the estimated cumulative distribution function for $\hat{\lambda}_{\delta c}$ for $\delta \in \{S, W\}$. In line with the theory, the distribution of $\hat{\lambda}_{Wc}$ first order stochastically dominates that of $\hat{\lambda}_{Sc}$.[21] In other words, it does look like giving greater incumbent discretion under weak executive constraints does improve behavior of incumbents in line with the theory.

Figure 5 here

While this does not prove that the model provides a valid way of thinking about differences across countries, it is a non-trivial finding from the data which has a bearing on the model.

More generally, these estimates provide a useful way of thinking about the consequences of adopting strong executive constraints for growth and income levels in terms of how politics responds to institutions. It also provides a method for considering heterogeneity across countries. We will explore this crucial point further in the next section.

5 Institutional Comparisons

In this section, we use the model to think about how institutions can be compared, suggesting an approach based on the literature on robust control in macro-economics.[22] We begin by discussing the framework that we have in mind and then return to the application that we developed in the last section.

5.1 Executive Constraints as Robust Control

Our next setup is to use our framework to analyze the choice of institutions. To be concrete, consider a set of countries $c = 1, ..., C$ that we wish to compare at different levels of δ. As in our application to growth we will focus on a discrete choice $\delta_c \in \{W, S\}$. We will allow for uncertainty in the political performance of a country under either institutional arrangement which we index by $\omega_j \in \{\omega_1, ..., \omega_J\}$, i.e. which can take on one of J possible values. Following the literature on robust control in macro-economics, we interpret this as model uncertainty. Specifically we let $\hat{r}\left(\delta_c : \omega_j\right)$ so that λ will depend on ω_j. We will consider the case where the distribution of ω_j is ambiguous, i.e. there may be multiple views about the distribution of political outcomes. Given limited knowledge of the world, this is a reasonable outcome even if it departs

[21] This is broadly the pattern for the countries that switched as shown in Table A4. However, there are some countries with $\lambda_{Sc} > \lambda_{Wc}$. The way that such anomalous cases could be reconciled with the theory is by supposing that the $G(\cdot)$ also changes due to greater transparency with strong constraints.

[22] Robust control ideas originated in macroeconomics to think about model uncertainty – see, for example, Hansen and Sargent (2001).

from the standard approach taken in most of economics. That said, there is increasing interest in studying economic policy in a framework of ambiguity.[23]

Let $\omega_j(\delta_c)$ be the realization of ω_j in country c with executive constraints δ_c. We will suppose $\omega_j(\delta_c)$ is determined at the point a new set of institutions is put in place and is heterogeneous across countries. Now

$$\lambda(\delta_c : \omega_j(\delta_c)) = (1 - G(\hat{r}(\delta_c : \omega_j(\delta_c))))$$

is the level of discipline that we expect to see on average in country c with institutions δ_c.

Also relevant to making institutional choice are the parameters $\{\Delta_L, \Delta_H, \Delta_0\}$ which we could allow to be country specific. However, to home in on politics, we will for the moment assume that these are the same across all countries.

To apply the framework, we need to define an evaluation metric. We will introduce a specific form of this for our application below. But to fix ideas, consider a general function according to which institutions will be compared:

$$B(\lambda(\delta_c : \omega_j), \Delta_L, \Delta_H, \Delta_0 : \delta_c)$$

which will typically be increasing in its first four arguments. A range of possible evaluation of criteria could be used depending in the application at hand. We will develop an approach which acknowledges the limits on our knowledge about how institutions affect policy outcomes. We believe that this is particularly reasonable for the application to political models. We propose two ways of making institutional comparisons reflecting this, which we refer to as full uncertainty and partial uncertainty.

As there is uncertainty (ambiguity) rather than risk, we will follow the decision criterion proposed by Gilboa and Schmeidler (1989) by supposing that a decision maker chooses strong executive constraints on the basis of a maximin criterion. Given the binary choice of institutions that we consider, the decision is based on:

$$\max_{\delta \in \{W,S\}} \min_{\omega_j \in \{\omega_1,\ldots,\omega_J\}} B(\lambda(\delta : \omega_j), \Delta_L, \Delta_H, \Delta_0 : \delta).$$

Let

$$\lambda_\delta^{\min} = \arg \min_{\omega_j \in \{\omega_1,\ldots,\omega_J\}} \{(1 - G(\hat{r}(\delta : \omega_j)))\}$$

be the average probability of producing Δ_H in the worst case. When we turn to the data below, we will look at a theoretical lower bound where $\lambda_\delta^{\min} = 0$. However, we will also consider what the data can tell us about the lower bound from country-level experiences.

[23] See, for example, Barlevy (2011), Hansen and Sargent (2007) and Manski (2011).

Partial Uncertainty The first possibility that we consider recognizes that no country is choosing institutions de novo. It begins with a particular realization $\hat{\omega}_c$ in one regime. Given enough time in that regime, it is reasonable to think that an estimate can be made of this realization by observing the performance of that country's policy making; we will give a specific example of this below.

The main issue is what happens if a switch is made. Here, we will suppose that the maximin criterion is used. Suppose that country c has the current realization of $\hat{\omega}_c(\delta)$ when institutional constraints are δ and let $\lambda_c = 1 - G(\hat{r}(\delta : \hat{\omega}_c(\delta)))$. Then it is optimal to switch to strong executive constraints under partial uncertainty if and only if

$$\min_{\omega_j \in \{\omega_1,\ldots,\omega_J\}} B(\lambda(S : \omega_j), \Delta_L, \Delta_H, \Delta_0 : S) > B(\lambda(W : \hat{\omega}_c(1)), \Delta_L, \Delta_H, \Delta_0 : W). \tag{10}$$

In this case, the status quo is compared to the worst outcome under strong constraints to create a robust case for strong constraints.

Full Uncertainty Full uncertainty corresponds to a "behind the veil of ignorance" comparison where choosing any particular institutional arrangement is associated with an ambiguous realization of ω_c with either institutional arrangement. Then strong constraints are preferred to weak constraints if and only if:

$$\min_{\omega_j \in \{\omega_1,\ldots,\omega_J\}} B(\lambda(S : \omega_j), \Delta_L, \Delta_H, \Delta_0 : S) > \min_{\omega_j \in \{\omega_1,\ldots,\omega_J\}} B(\lambda(W : \omega_j), \Delta_L, \Delta_H, \Delta_0 : W). \tag{11}$$

This is like saying that the decision maker picks the worst performance under strong constraints and compares it to the worst performance under weak constraints. If we assume that $\{\Delta_L, \Delta_H, \Delta_0\}$ are constant across countries, then this choice is not country specific. Equation (11) then gives us the robustly optimal institution.[24]

Empirical Implementation The model developed here is very simple. In practice there are many policies that need to be determined across multiple dimensions. To illustrate the approach, we will develop a specific application in the next section. Before turning to this, it is worth thinking about how data can play a role in informing institutional comparisons. To learn about political uncertainty, we can draw on data across a range of countries in particular institutional regimes. These can be used to form a view of the range of experience associated with realizations of ω_j. Working with an ambiguous outcome on this, means that the analyst does not know the probability distribution over λ. However, as we shall see, the data may

[24] Another way of thinking about this is in terms of Rawls (1971) who suggested a maximin approach motivated by concerns about distributive justice rather than ambiguity.

be able to tell us something about the distribution, e.g. the quartiles.

The maximin approach can be applied by looking at the range of experience rather than the frequency of particular experiences. This has a further robustness property since selection into and out of institutions is almost certainly an issue. Suppose that when a country is considering entering a regime it has private information that we as researchers do not have about $\hat{\omega}_c(\delta)$. Take the limiting case where $\delta_c = 1$ if and only if

$$B\left(\lambda\left(S : \hat{\omega}_c(0)\right), \Delta_L, \Delta_H, \Delta_0 : S\right) > B\left(\lambda\left(W : \hat{\omega}_c(1)\right), \Delta_L, \Delta_H, \Delta_0 : W\right).$$

Now we would have a classic case where estimating the conditional mean of λ by regime would be contaminated by selection. So using average performances in a regime would be problematic. The way to think about our decision criterion is that we are comparing the worst performance among those who selected a regime and compare it to the current performance (partial uncertainty) or the worst performance in the other regime (full uncertainty). This retains an attractive robustness property as we are not stacking the deck in either direction. We are using the worst performance among the *observed* examples of countries in either institutional regime to evaluate the gains from changing institutions. Our model also allows us to address remaining selection issues by using a theoretical lower bound which is where $\lambda_c = 0$. This assumption gives a lot of weight to the relative size of Δ_0 and Δ_L - more than is observed in the actual growth data of adopters. However, there is still a question of how big the gains are in specific cases.

5.2 Evaluation Based on Income Levels

We now apply the framework above the growth framework in section (4) using our estimates of $\left\{\hat{\Delta}_L, \hat{\Delta}_H, \hat{\Delta}_0\right\}$ and $\hat{\lambda}_{\delta c}$ to compare strong and weak constraints.

The Criterion Our empirical model focused on growth but it makes more sense for evaluating institutions to focus on the impact on income *levels* that follow from different growth paths. Specifically, we look at expected income levels over a ten year time horizon as the criterion for evaluation. With trend production growth $\kappa_c(\delta)$ and variance $\sigma^2_{\varepsilon c}(\delta)$ income after T years will be

$$b\left(\delta : \kappa_c(\delta), \sigma^2_{\varepsilon c}(\delta)\right) = e^{\left\{\left[\hat{\kappa}(\delta) - \frac{1}{2}\hat{\sigma}^2_{\varepsilon c}(\delta)\right]T\right\}}$$

for $\delta \in \{W, S\}$. Thus the mean and variance of productivity shocks drive expected income levels. We use our estimates $\left\{\hat{\Delta}_H, \hat{\Delta}_L, \hat{\Delta}_0\right\}$ together with specific values of $\hat{\lambda}_{\delta c}$ to form estimates of $\left\{\kappa_c(\delta), \sigma^2_{\delta c}\right\}$ using equations

$$\hat{\kappa}(\delta) = \hat{\lambda}_{\delta c}\hat{\Delta}_H + \left(1-\hat{\lambda}_{\delta c}\right)\hat{\Delta}_\delta \tag{12}$$

and

$$\hat{\sigma}^2_{\varepsilon c}(\delta) = \hat{\lambda}_{\delta c}\left(1-\hat{\lambda}_{\delta c}\right)\left[\hat{\Delta}_H - \hat{\Delta}_\delta\right]^2 \tag{13}$$

where $\hat{\Delta}_S = \hat{\Delta}_0$ and $\hat{\Delta}_W = \hat{\Delta}_L$. The choice of $\hat{\lambda}_{\delta c}$ embodies our assumption about how politics affects income comparisons.

This approach allows us to make comparisons which involve countries which have never adopted strong executive constraints by hypothesizing a value λ_{Sc}. We can then hold the economic environment fixed by fixing $\left\{\hat{\Delta}_H, \hat{\Delta}_L, \hat{\Delta}_0\right\}$ for the purposes of comparisons. The values of $\lambda_{\delta c}$ then do all the work in our estimates of expected income. We can use the specific estimates of $\lambda_{\delta c}$ in all countries to have a full range of heterogeneous effects. When we look at the range of experience in Figure 5, we can treat this as giving us a sense of the range of outcomes due to realization of the model uncertainty shocks, ω_j, which we discussed above. For example, the quartiles of performance provide a sense of this.

Partial Uncertainty In the partial uncertainty case, we focus only on those countries which have always had weak executive constraints throughout our time period and suppose that their past experience reflects the realization of $\hat{\omega}_c$ which will remain fixed in future as long as they maintain weak constraints. We use mean productivity growth to derive an estimate of λ_{Wc}. Using (10), we can then estimate the income gains (or losses) from adopting strong executive constraints.

To make the comparison concrete, we need to specify λ_{Sc}. Our maximin framework tells us to look at worst-case scenarios and we suggest a couple of ways of doing this. The first takes an empirical approach by taking the weighted mean of λ_{Sc} in the bottom quartile of the distribution of countries with strong executive constraints. The reason for doing this rather than taking the worse performer is to acknowledge the potential lack of precision in our estimates of λ_{Sc}.[25] The second approach eschews the data completely and supposes that $\lambda_{Sc} = 0$, the theoretically-driven lower bound on performance.

In each case, we explore whether a country would gain from a switch to strong executive constraints, and by how much using the following measure of gains/losses:

$$\frac{b\left(S:\hat{\kappa}(S), \hat{\sigma}^2(S)\right) - b\left(W:\kappa_c(W), \sigma^2_{\varepsilon c}(W)\right)}{b\left(W:\kappa_c(W), \sigma^2_{\varepsilon c}(W)\right)} = e^{\left(\left[\hat{\kappa}(S)-\frac{1}{2}\hat{\sigma}^2(S)\right]-\left[\kappa_c(W)-\frac{1}{2}\sigma^2_{\varepsilon c}(W)\right]\right)T} - 1 \tag{14}$$

[25] If countries are selecting their institutions at least in part on the basis of average growth performance, we may though be concerned that could lead to an upward bias, i.e. a bias towards strong constraints.

supposing that countries would keep their current estimated value of λ_{Wc} if they remain with weak constraints. In our calculations we assume $T = 10$.

Figure 6 gives the estimates of (14) for our two scenarios. The blue dots give the estimate of gain using the mean value of λ_{Sc} in the lowest quartile of the distribution, i.e. we assume that countries that switch would reach a level of discipline similar to the worst performers. More than one third of the countries which currently have weak constraints would gain from this switch. Not surprisingly, the countries that would lose are those with a strong growth performance under weak constraints such as China. Given the latter's sustained growth performance and our estimate of $\sigma^2_{\delta c}$, this approach predicts that China will continue to perform better with weak constraints assuming that λ_{Wc} does not change.

The case where we posit $\lambda_{Sc} = 0$ generates a more cautious assessment with very few countries in the sample being predicted to gain from strong constraints. Figure 6 puts country labels on a small number of countries in the middle of the distribution of λ_{Wc} to illustrate. Venezuela and Cameroon, for example, would gain significantly if they had the average λ_{Sc} in the lowest quartile after switching to strong executive constraints but would lose if they drew the worst possible λ_{Sc}. However, in some cases, the ten year income gains remain positive even with this extremely conservative approach. This is the case for the Democratic Republic of Congo, the Central African Republic, Somalia and Guinea-Bissau.[26]

Figure 6 here

Full Uncertainty We now compare institutions based on the full uncertainty criterion (11). This can be motivated by thinking of a "veil of ignorance" comparison or by considering the possibility that any country could be subject to a bad political shock in future which would change $\lambda_{\delta c}$ within a regime. We view this as an important aspect of assessing robustness questions in relation to the performance of institutions. Even China, with its impressive growth performance, is a case in point. If it does not reform its institutions, then there is always a risk of a fracture in the communist party which diminishes effective governance or for corruption to become an insurmountable problem beyond the control of existing institutions. Thus, it is arguable that it would be misleading to put too much weight on the current experience with λ_{Wc} and to look at what would happen if there were political shocks, interpreted here as changes in ω_c.

[26] Figure 6 does not change significantly when we look only at country years with open access to power. Appendix Figure A1 shows that the same broad message carries over if we use the measure of checks on the chief executive to define strong constraints.

To look at this, we focus on the following measure of proportionate gain:

$$\frac{b\left(S:\hat{\kappa}\left(S\right),\hat{\sigma}^{2}\left(S\right)\right)-b\left(W:\hat{\kappa}\left(W\right),\hat{\sigma}^{2}\left(W\right)\right)}{b\left(W:\hat{\kappa}\left(W\right),\hat{\sigma}^{2}\left(W\right)\right)}=e^{\left(\left[\hat{\kappa}(S)-\frac{1}{2}\hat{\sigma}^{2}(S)\right]-\left[\hat{\kappa}(W)-\frac{1}{2}\hat{\sigma}^{2}(W)\right]\right)T}-1, \tag{15}$$

i.e. replacing $\left\{\kappa_c\left(W\right),\sigma^2_{\varepsilon c}\left(W\right)\right\}$ in (14) with $\left\{\hat{\kappa}_c\left(W\right),\hat{\sigma}^2_{\varepsilon c}\left(W\right)\right\}$ where the latter are derived from some assumption about future draws of λ_{Wc}. Again, we set $T = 10$.

As above, we consider two possibilities for the worst case with or without strong executive constraints. The first calculates the mean value of $\lambda_{\delta c}$ for the lowest quartile under both strong and weak executive constraints. We then use these values along with $\left\{\hat{\Delta}_H,\hat{\Delta}_L,\hat{\Delta}_0\right\}$ to calculate values of $\left\{\hat{\kappa}\left(S\right),\hat{\sigma}^2\left(S\right)\right\}$ and $\left\{\hat{\kappa}\left(W\right),\hat{\sigma}^2\left(W\right)\right\}$. In the second case, we set $\lambda_{Sc} = \lambda_{Wc} = 0$ and calculate the gain or loss from (15). Table 3 reports the results for A) countries which have actually switched and B) the full sample.

We know already from Figure 5 that the average estimated value of λ_{Wc} is higher in each quartile. Even so, trend productivity growth is still lower in the bottom quartile due to the fact that we have estimated that $\hat{\Delta}_0 > \hat{\Delta}_L$. The implied variance of productivity shocks is also higher with weak constraints. This reinforces the idea that it is protection from the downside that matters empirically when adopting strong constraints.

The gain (15) is 4.64% if we use the bottom quartile of performance as the worst case. However, using $\lambda_{Sc} = \lambda_{Wc} = 0$, the gain is 23.88%. This is coming not only from the lower trend growth but also the larger variance in growth under weak executive constraints. Whichever way we look at it, strong constraints do appear like a form of robust control on policy. Moreover, the estimated gains of nearly a quarter in income levels over ten years on the most extreme version of this are quite sizeable. Thus looking through the lens of robust control does yield a clear-cut and quantitatively non-trivial estimate of the gains from strong institutions. The estimates are even large when we look at the full sample of countries: 9.16% and 34.85% respectively.[27]

Finally Table 4, pushes the robustness theme further still by replacing the point estimates of $\left\{\hat{\Delta}_H,\hat{\Delta}_L,\hat{\Delta}_0\right\}$ by the lower bound of the 95% confidence interval. The qualitative story is unchanged but the sizes of the gains increase further. We now estimate a gain from the adoption of strong executive constraints of 9.51% with bottom quartile of the $\lambda_{\delta c}$ distribution and 36.30% when $\lambda_{Sc} = \lambda_{Wc} = 0$. Thus, our core conclusion is reinforced further by this.

[27] Table A6 shows that this pattern is the same if we impose openness. This is in line with raw data moments which we showed in Table 1. Results are also similar if we use the checks on the executive as our measure of strong executive constraints but become larger in size (see Table A7). The estimated gain in the worst case scenario now rises to 32% in the sample of switchers.

The Importance of Within-Country Volatility The comparisons above focus on differences between country. However, the theoretical model also gives a role to within-country variance in level comparisons as shown in both (14) and (15). This ties in with the observation in Aguiar and Gopinath (2007) and Koren and Tenreyro (2007) that volatility is important in developing and emerging-market economies. However, these tend also to be countries with weak executive constraints.

To focus on how politics contributes to volatility, note that the larger within-country variance in growth rates under weak executive constraints is driven by the fact that in our framework $\Delta_L > \Delta_0$. This is brought into particularly sharp relief when $\lambda_{Sc} = \lambda_{Wc} = 0$ since within-country variation in the variance then plays a key role when we estimate (15). This is an appealing feature of the approach used since, as we have already observed, larger within-country variance is a key feature of the data with weak executive constraints.

Our findings are consistent with the observation that prolonged episodes of negative GDP growth are a feature of countries with weak executive constraints. A case in point is Zimbabwe whose economy shrank by 3.17% on average between 1992 and 2002. This example is particularly striking since the decline was preceded by constitutional changes that concentrated powers on the president. However, this is by no means an isolated case; Uganda's economy shrank by 3.48 % between 1971 and 1981 and Nicaragua's economy shrank by 3.72% between 1978 and 1988. Such periods of economic decline tend to be shorter and less frequent with strong executive constraints.[28] Indeed, none of these three countries had a long spell of negative growth during their time under strong executive constraints. Our model offers an insight into these patterns and provides a political economy link to debates about emerging market volatility.

Summing Up The approach to comparing institutions presented here is novel in three respects. First, we have looked at both volatility and trend growth in a unified way based on changes in political institutions. Second, we have combined theory and empirical estimates to construct an explicit counter-factual for income growth. Third, we have insisted on robust comparisons to reflect the reasonable assumption that we lack any precise sense of the probability distribution over political outcomes when different institutions are adopted.

As a corollary of this approach, we get away from a focus on average effects which have been the main focus of the regression-based evidence. Authors such as Persson and Tabellini (2009b) have rightly stressed that what we care about in assessing institutions is not what happens on average but the specific comparison based on reasonable views about heterogeneity. Here, we have been explicit about one source of heterogeneity as suggested by the theory – the specific $\{\lambda_{Wc}, \lambda_{Sc}\}$ that a country faces. The approach using full uncertainty deflects some concerns about endogenous switching between institutions as we are no

[28] We used a Markov Chain model of growth to check that the persistence of good growth states is lower under weak executive constraints. Results are available from the authors on request.

longer using country-specific experiences to generate mean performance.

When making comparisons an analyst can use their preferred beliefs about $\{\lambda_{Wc}, \lambda_{Sc}\}$. However, we have argued that if there is ambiguity then there is a fairly strong case for opting for strong executive constraints. The question of whether a country like China will continue to thrive is really a judgement on the persistence of λ_{Wc}. Weak executive constraints may appear favorable if currently successful countries are used as examples. But weak executive constraints expose a country to more extreme risk consistent with the facts about within-country volatility. In the end, there is no substitute for trying to get better empirical knowledge about the factors that determine $\{\lambda_{Wc}, \lambda_{Sc}\}$.

6 Concluding Comments

This paper is motivated by the need to work with models that allow us to think about the mean and volatility of economic performance across political regimes. It is motivated in part by the frequently observed fact that countries with weak political institutions generally have more variable economic performance, a finding which is also apparent when comparing transitions between strong and weak executive constraints. Moreover, the fall in variance consists of a fall of within-country variance over time as well as lower between-country variance.

To account for these facts, we have suggested an extremely simple model where executive constraints affect both the mean and variance of performance and have applied this to economic growth. By attributing variability in growth performance to politics, we can explain the decline in variance as the result of changes in policy-making which are broadly in line with the prediction of the model. Our model also gives us a way of thinking about the heterogeneous growth performance of countries with strong and weak executive constraints getting beyond average effects following Persson and Tabellini (2009a,b). The approach that we have taken is quite simple. However, this has enabled us to unpack the heterogeneous effects of institutional differences. This remains an important research program – the average effects are often of little policy relevance.

We have argued that making institutional comparisons is fraught with ambiguity – we lack precise models which allow us to predict which countries will perform well and badly in each regime. And this is confounded further when countries switch endogenously. In response to this, we suggest an approach to institutional comparisons based on robust control where the distribution of political outcomes is ambiguous and the analyst uses maximin expected utility. We have two main findings when applying this to economic growth. First, under the assumption of partial uncertainty, a robust comparison between institutions warrants caution towards institutional experiments. Most countries under weak executive constraints would not benefit from institutional reform when compared to the worst case in strong executive constraints. Second, from the standpoint of full uncertainty, there is a strong argument for institutional reform from a robustness perspective. The latter is compelling when the comparisons that we make are robust to political shocks from an unknown distribution in countries which start out with weak executive constraints. The analysis highlights the need for better knowledge that we currently have in understanding what draws of $\lambda_{\delta c}$ are plausible in any specific country.

While our focus here has been on growth, in future work it is important to study a wider range of outcomes and policies. An initial look at the data on life expectancy and infant mortality strongly suggests that the effect on the second moment is not confined to economic growth. In Figures 7 and 8, we plot the within residuals from a regression similar to equation (1) which controls for country fixed effect and year dummy variables. Once again weak executive constraints show a higher within episode standard deviation compared to strong constraints. This could, to some extent, also be a reflection of general macro-economic conditions in the economy but they could also reflect a wider set of policy factors. This merits a closer investigation in future work. However, these Figures are in line with the thrust of the argument developed here.

Figures 7 and 8 about here

Another potentially fruitful line of enquiry involves studying private sector responses to institutions in greater detail. In Besley and Mueller (2015), we investigate how investment flows by foreign firms respond to reductions in macro-economic risk following a switch to strong executive constraints. We argue that this might explain why countries with strong executive constraints appear to attract more investment inflows.

References

[1] Acemoglu, Daron, Simon Johnson, James Robinson, and Yunyong Thaicharoen, [2003], "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth," *Journal of Monetary Economics*, 50, 49-123.

[2] Aghion, Philippe and Peter Howitt, [2006], "Appropriate Growth Policy:A Unifying Framework," *Journal of the European Economic Association*, 4(2–3):269–314

[3] Aguiar, Mark and Gita Gopinath, [2007], "Emerging Market Business Cycles: The Cycle Is the Trend," *Journal of Political Economy*, 115 (1), 69-102.

[4] Almeida, Heitor and Daniel Ferraira, [2002], "Democracy and the Variability of Economic Performance," *Economics and Politics*, 14(3), 225-257.

[5] Barlevy, Gardi, [2011], "Robustness and Macroeconomic Policy," *Annual Review of Economics*, 3, 1-24.

[6] Barro, Robert J., [1973], "The Control of Politicians: An Economic Model", *Public Choice*, 14, 19-42.

[7] Barro, Robert J., [1996], "Democracy and growth," *Journal of Economic Growth,* 1(1), 1–27.

[8] Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh, [2001], "New tools and new tests in comparative political economy: the database of political institutions," *World Bank Economic Review* 15, 165–176.

[9] Besley, Timothy, [2006], *Principled Agents? The Political Economy of Good Government*, The Lindahl Lectures, Oxford: Oxford University Press.

[10] Besley, Timothy and Masa Kudamatsu [2009], "Making Autocracy Work" in Elhanan Helpman (ed) *Institutions and Economic Performance*, Cambridge: Harvard University Press.

[11] Besley, Timothy and Hannes Mueller, [2015], "Volatility, Institutions and Investment", CEPR Discussion Papers 10373

[12] Buchanan, James M., [1989], "The Public-Choice Perspective," in *Essays on the Political Economy*, Honolulu: University of Hawaii Press.

[13] Calvo, Guillermo, [1998], "Capital Flows and Capital-Market Crises: The Simple Economics of Sudden Stops", *Journal of Applied Economics*, 1, 35-53.

[14] Ferejohn, John [1986], "Incumbent Performance and Electoral Control", *Public Choice*, 50, 5-25.

[15] Gilboa, Itzhak and David Schmeidler, [1989], "Maxmin Expected Utility with Nonunique Prior," *Journal of Mathematical Economics*, 18(2), 141-53.

[16] Glaeser, Edward, Giacomo Ponzetto and Andrei Shleifer, [2007], "Why does democracy need education?," *Journal of Economic Growth*, Springer, 12(2), 77-99.

[17] Keefer, Philip [2012], "DPI2012 Database of Political Institutions: Changes and Variable Definitions"

[18] Hansen, Lars-Peter and Thomas Sargent, [2001], "Robust Control and Model Uncertainty," *The American Economic Review* (Papers and Proceedings), 91(2), 60-66.

[19] Hansen, Lars-Peter and Thomas Sargent, [2007], *Robustness*, Princeton: Princeton University Press.

[20] Henisz, Witold, [2004], "Political Institutions and Policy Volatility," *Economics and Politics*, 16 (1), 1 - 27.

[21] Hume, David, [1742], Of the Independency of Parliament in Essays, Moral Political and Literary, Liberty Fund, Inc. 1987. Ed. Eugene F. Miller. Library of Economics and Liberty. 5 June 2005df. <http://www.econlib.org/library/LFBooks/Hume/hmMPL6.html>.

[22] Mill, John Stuart, [1859], *On Liberty*, New York: Bartleby.com, 1999.

[23] Koren, Miklós and Silvana Tenreyro, [2007], "Volatility and Development." *Quarterly Journal of Economics*, 122(1), 243-287.

[24] Manski, Charles, [2011], "Choosing Treatment Policies Under Ambiguity," *Annual Reviews of Economics*, 3: 25–49.

[25] Mobarak, Ahmed Mushfiq, [2005], "Democracy, Volatility and Economic Development," *The Review of Economics and Statistics*, 87(2): 348–361.

[26] Montesquieu, Charles de, [1752], *The Spirit of the Laws*, Cambridge: Cambridge University Press, 1989.

[27] Overbye, Einar, [1996], "Democracy as Insurance," *Public Choice* 87: 319-345.

[28] Papaioannou, Elias, and Gregorios Siourounis, [2008], "Democratization and Growth," *Economic Journal*, 118, 1520-1551.

[29] Persson, Torsten, Gérard Roland and Guido Tabellini, [1997], "Separation of Powers and Political Accountability," *Quarterly Journal of Economics*, 112 (4), 1163-1202.

[30] Persson, Torsten and Guido Tabellini [2000], *Political Economics: Explaining Economic Policy*, Cambridge MA, MIT Press.

[31] Persson, Torsten and Guido Tabellini [2009a], "The Growth Effects of Democracy: Is it Heterogenous and How Can it Be Estimated?", in Elhanan Helpman (ed) *Institutions and Economic Performance*, Cambridge: Harvard University Press.

[32] Persson, Torsten and Guido Tabellini [2009b], "Democratic Capital: The Nexus of Economic and Political Change", *American Economic Journal: Macroeconomics*, 1(2), 88-126.

[33] Przeworski, Adam, and Fernando Limongi, [1993], "Political Regimes and Economic Growth," *Journal of Economic Perspectives*, *7*(3), 51-69.

[34] Rawls, John, [1971], *A Theory of Justice*, Cambridge MA: Harvard University Press.

[35] Rodrik, Dani [1997], "Democracy and Economic Performance," unpublished typescript available at http://www.hks.harvard.edu/fs/drodrik/Research%20papers/demoecon.PDF

[36] Rodrik Dani, [1999], "Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses," *Journal of Economic Growth*, 4(4), 385-412.

[37] Tsebelis, George, [2002], *Veto Players: How Political Institutions Work*, New York and Princeton, NJ: Russell Sage Foundation and Princeton University Press.

[38] Weede, Erich, [1996], "Political Regime Type and Variation in Economic Growth Rates," *Constitutional Political Economy*, 7, 167-176.

A Proof of Proposition 1

Proof. Let $u(z) = z\Delta_H + (1-z)\hat{\Delta}(\xi)$ where $\hat{\Delta}(\xi) = \xi\Delta_L + (1-\xi)\Delta_0$. Define voter discounted utility $W(\lambda, \xi, \boldsymbol{\phi}, \pi_n)$ from the recursion:

$$\begin{aligned} W(\lambda,\xi,\boldsymbol{\phi},\pi_n) &= u(\xi\psi^n + (1-\xi)\lambda) + \psi^n\xi\beta\left[\max_{\phi(\Delta_H)}\{\phi W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1}) + (1-\phi)[W(\lambda,\xi,\boldsymbol{\phi},\pi)]\}\right] \\ &\quad + [1-\psi^n]\xi\beta\left[\max_{\phi(\Delta_L)}\{\phi W(\lambda,\xi,\boldsymbol{\phi},0) + (1-\phi)[W(\lambda,\xi,\boldsymbol{\phi},\pi)]\}\right] \\ &\quad + (1-\xi)\beta\left[\max_{\phi(\Delta_0)}\{\phi W(\lambda,\xi,\boldsymbol{\phi},\pi_n) + (1-\phi)[W(\lambda,\xi,\boldsymbol{\phi},\pi)]\}\right] \end{aligned}$$

where $u(\xi\psi^n + (1-\xi)\lambda)$ captures expected utility within the period. Also, we have defined $\pi_{n+1} = \frac{\pi_n}{\pi_n + \lambda(1-\pi_n)}$ and $\psi^n = (\pi_n + (1-\pi_n)\lambda)$ which are both increasing sequences. Here we have used the fact that $\Pi = \pi_n$ if $e_t = q$, i.e. there is no updating of the incumbent's type under executive constraints. We adopt the convention $\pi_1 = \pi$.

We need to show that for all λ, ξ it is optimal to remove the incumbent whenever $\Pi = 0$ and to re-elect her if she has always produced Δ_H if she has had discretion. Note that

$$\begin{aligned} \phi(\Delta_L) &= \arg\max_{\phi}\{\phi W(\lambda,\xi,\boldsymbol{\phi},0) + (1-\phi)[W(\lambda,\xi,\boldsymbol{\phi},\pi)]\} \\ \phi(\Delta_H) &= \phi(\Delta_0) = \arg\max_{\phi}\{\phi W(\lambda,\xi,\boldsymbol{\phi},\pi_n) + (1-\phi)[W(\lambda,\xi,\boldsymbol{\phi},\pi)]\} \text{ for } n \geq 2. \end{aligned}$$

Suppose that $\phi(\Delta_H) = \phi(\Delta_0) = 1$. We will show that $\phi(\Delta_L) = 0$. If $\phi(\Delta_L) = 0$, then $W(\lambda,\xi,\boldsymbol{\phi},0) < W(\lambda,\xi,\boldsymbol{\phi},\pi)$. Suppose not, then $\phi(\Delta_L) = 1$ and:

$$W(\lambda,\xi,\boldsymbol{\phi},0) = \frac{u(\lambda)}{1-\beta}. \tag{16}$$

For this to be a voting equilibrium, we require that $W(\lambda,\xi,\boldsymbol{\phi},0) \geq W(\lambda,\xi,\boldsymbol{\phi},\pi)$. Then,

$$\begin{aligned} W(\lambda,\xi,\boldsymbol{\phi},\pi) &= \delta\left(\xi\psi^1 + (1-\xi)\lambda\right) + \psi^1\xi\beta\left[\max\{W(\lambda,\xi,\boldsymbol{\phi},\pi), W(\lambda,\xi,\boldsymbol{\phi},\pi_2)\}\right] \\ &\quad + \left[1-\psi^1\right]\xi\beta W(\lambda,\xi,\boldsymbol{\phi},0) + (1-\xi)W(\lambda,\xi,\boldsymbol{\phi},\pi) \\ &\geq \delta\left(\xi\psi^1 + (1-\xi)\lambda\right) + \left[\xi\psi^1 + (1-\xi)\right]\beta W(\lambda,\xi,\boldsymbol{\phi},\pi) \\ &\quad + \left(1-\psi^1\right)\beta\xi W(\lambda,\xi,\boldsymbol{\phi},0) \end{aligned}$$

which implies that

$$\begin{aligned} \left(1-\beta\left[\xi\psi^1 + (1-\xi)\right]\right)W(\lambda,\xi,\boldsymbol{\phi},\pi) &\geq \delta\left(\xi\psi^1 + (1-\xi)\lambda\right) + \left(1-\psi^1\right)\xi\beta W(\lambda,\xi,\boldsymbol{\phi},0) \\ &> \delta(\lambda) + \left(1-\psi^1\right)\beta\xi W(\lambda,\xi,\boldsymbol{\phi},0) \\ &= W(\lambda,\xi,\boldsymbol{\phi},0)\left(1-\beta\left[\xi\psi^1 + (1-\xi)\right]\right) \end{aligned}$$

using (16), a contradiction. So $W(\lambda,\xi,\boldsymbol{\phi},0) < W(\lambda,\xi,\boldsymbol{\phi},\pi)$ and $\phi(\Delta_L) = 0$ as claimed.

Suppose then that $\phi(\Delta_L) = 0$ and $\phi(\Delta_0) = 1$. We will show that, beginning with an incumbent with reputation π, $\phi(\Delta_H) = 1$. Suppose not. For $\phi(\Delta_H) = 0$, we require that $W(\lambda,\xi,\boldsymbol{\phi},\pi_2) \leq W(\lambda,\xi,\boldsymbol{\phi},\pi)$. This implies that

$$W(\lambda,\xi,\boldsymbol{\phi},\pi) = \frac{u\left(\xi\psi^1 + (1-\xi)\lambda\right)}{1-\beta}.$$

Now observe that if the incumbent were retained, then

$$\begin{aligned} W(\lambda,\xi,\boldsymbol{\phi},\pi_2) &= u\left(\xi\psi^2 + (1-\xi)\lambda\right) + \psi^2\beta\xi\max\{W(\lambda,\xi,\boldsymbol{\phi},\pi_3), W(\lambda,\xi,\boldsymbol{\phi},\pi)\} \\ &\quad + \left(1-\psi^2\right)\xi\beta W(\lambda,\xi,\boldsymbol{\phi},\pi) + (1-\xi)\beta W(\lambda,\xi,\boldsymbol{\phi},\pi_2). \\ &\geq u\left(\xi\psi^2 + (1-\xi)\lambda\right) + \xi\beta W(\lambda,\xi,\boldsymbol{\phi},\pi) + (1-\xi)\beta W(\lambda,\xi,\boldsymbol{\phi},\pi_2) \end{aligned}$$

which implies that:

$$\begin{aligned} [1-\beta(1-\xi)]\,W(\lambda,\xi,\boldsymbol{\phi},\pi_2) &\geq u\left(\xi\psi^2 + (1-\xi)\lambda\right) + \beta\xi W(\lambda,\xi,\boldsymbol{\phi},\pi) \\ &> u\left(\xi\psi^1 + (1-\xi)\lambda\right) + \beta\xi W(\lambda,\xi,\boldsymbol{\phi},\pi) \end{aligned}$$

which implies that $W(\lambda,\xi,\boldsymbol{\phi},\pi_2) > W(\lambda,\xi,\boldsymbol{\phi},\pi)$ – a contradiction so that $\phi(\Delta_H) = 1$ as claimed.

To complete the proof, we need to show that $W(\lambda,\xi,\boldsymbol{\phi},\pi_n) > W(\lambda,\xi,\boldsymbol{\phi},\pi)$ implies that $W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1}) > W(\lambda,\xi,\boldsymbol{\phi},\pi)$ for all $n \geq 1$. Suppose not. Then $W(\lambda,\xi,\boldsymbol{\phi},\pi_n) > W(\lambda,\xi,\boldsymbol{\phi},\pi) \geq W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1})$. This implies that

$$\begin{aligned} W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1}) &= u\left(\xi\psi^{n+1} + (1-\xi)\lambda\right) + \beta\psi^{n+1}\xi\max\{W(\lambda,\xi,\boldsymbol{\phi},\pi), W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+2})\} \\ &\quad + \left[1-\psi^{n+1}\right]\xi\beta W(\lambda,\xi,\boldsymbol{\phi},\pi) + (1-\xi)\beta W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1}) \\ &\geq u\left(\xi\psi^{n+1} + (1-\xi)\lambda\right) + \beta\xi W(\lambda,\xi,\boldsymbol{\phi},\pi) + (1-\xi)\beta W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1}). \end{aligned}$$

or

$$[1-(1-\xi)\beta]\,W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1}) \geq \delta\left(\xi\psi^{n+1} + (1-\xi)\lambda\right) + \beta\xi W(\lambda,\xi,\boldsymbol{\phi},\pi).$$

It also implies that

$$\begin{aligned} W(\lambda,\xi,\boldsymbol{\phi},\pi_n) &= u(\xi\psi^n + (1-\xi)\lambda) + \beta\xi\psi^n\max\{W(\lambda,\xi,\boldsymbol{\phi},\pi), W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1})\} \\ &\quad + (1-\psi^n)\xi\beta W(\lambda,\xi,\boldsymbol{\phi},\pi) + (1-\xi)\beta W(\lambda,\xi,\boldsymbol{\phi},\pi_n) \\ &= u(\xi\psi^n + (1-\xi)\lambda) + \beta\xi W(\lambda,\xi,\boldsymbol{\phi},\pi) + (1-\xi)\beta W(\lambda,\xi,\boldsymbol{\phi},\pi_n) \end{aligned}$$

or

$$[1-(1-\xi)\beta]\,W(\lambda,\xi,\boldsymbol{\phi},\pi_n)=u(\xi\psi^n+(1-\xi)\lambda)+\beta\xi W(\lambda,\xi,\boldsymbol{\phi},\pi).$$

Hence:

$$\begin{aligned} u(\xi\psi^n+(1-\xi)\lambda)+\beta\xi W(\lambda,\xi,\boldsymbol{\phi},\pi) &> W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1}) \\ &\geq u\left(\xi\psi^{n+1}+(1-\xi)\lambda\right)+\beta\xi W(\lambda,\xi,\boldsymbol{\phi},\pi) \end{aligned}$$

which is a contradiction since

$$u\left(\xi\psi^{n+1}+(1-\xi)\lambda\right)>u(\xi\psi^n+(1-\xi)\lambda).$$

using Bayes rule. So we must have $W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1})>W(\lambda,\xi,\boldsymbol{\phi},\pi)$. Then the result holds by induction for all $n>1$. Finally note that the last step also implies that $\phi(\Delta_0)=1$. For this to be the case, we need that $W(\lambda,\xi,\boldsymbol{\phi},\pi_{n+1})\geq W(\lambda,\xi,\boldsymbol{\phi},\pi)$ which we have already shown holds. ■

B Proof of Proposition 2

Proof. Suppose now that $x_t=1$. There are two cases, suppose that $r_t\leq\hat{r}(\xi)$, then all incumbents will generate Δ_H and the given a reputation Π, this yields

$$\Delta_H+\beta W(\lambda,\xi,\boldsymbol{\phi},\Pi)$$

since nothing is learned about the type. Imposing the status quo yields a payoff of

$$\Delta_0+\beta W(\lambda,\xi,\boldsymbol{\phi},\Pi)$$

since nothing is learned about the incumbent's type in this case either. Hence allowing discretion is optimal.

Now consider the case where $r_t>\hat{r}(\boldsymbol{\xi})$. Then the expected payoff from allowing discretion is

$$\Pi\Delta_H+(1-\Pi)\,\Delta_L+\beta\left[\Pi W(\lambda,\xi,\boldsymbol{\phi},1)+(1-\Pi)\,W(\lambda,\xi,\boldsymbol{\phi},\pi)\right].$$

If no discretion is offered, then the payoff is

$$\Delta_0+\beta W(\lambda,\xi,\boldsymbol{\phi},\Pi).$$

Thus discretion is offered if and only if:

$$\begin{aligned} & \Pi\Delta_H + (1-\Pi)\,\Delta_L - \Delta_0 \\ \geq\; & \beta\left[W(\lambda,\xi,\boldsymbol{\phi},\Pi) - \Pi W(\lambda,\xi,\boldsymbol{\phi},1) - (1-\Pi)\,W(\lambda,\xi,\boldsymbol{\phi},\pi)\right] \end{aligned} \tag{17}$$

Now note that $W(\lambda,\xi,\boldsymbol{\phi},\Pi) \geq W(\lambda,\xi,\boldsymbol{\phi},\pi)$ and as $\pi \to 0$ we have $\Pi \to 0$ so that the right-hand-side is positive. Moreover $\Delta_L - \Delta_0 < 0$. Hence (17) cannot hold and discretion is never offered in this case, as claimed. ■

C Proof of Proposition 3

Proof. Given ξ let

$$V(\hat{r};\xi) = \xi \int_{\hat{r}}^{R_H} r dG(r) + [1 - \xi(1 - G(\hat{r}))]\,\beta V(\hat{r};\boldsymbol{\phi},\xi)$$

where $1-\xi(1-G(\hat{r}))$ is the probability that the incumbent stays in office. Then the incumbent is indifferent between taking r and continuing at

$$\hat{r} = \beta V(\hat{r};\xi)$$

and solving this yields (**??**). Now let:

$$H(x,\xi) = \frac{\beta\xi\int_x^R r dG(r)}{1-\beta(1-\xi+\xi G(x))}.$$

Observe that

$$H_x(x,\xi) = \frac{\beta\xi\,[H(x,\xi)-x]\,g(x)}{1-\beta(1-\xi+\xi G(x))}.$$

Note that $H(R_L,\xi) > 0$ by assumption and $H(R_H,\xi) = 0 < R_H$. Thus there exists $\hat{r}(\xi) \in (R_L,R_H)$ such that $H(\hat{r}(\xi),\xi) = \hat{r}(\xi)$. Moreover $H_x(\hat{r}(\xi),\xi) = 0$ so that $H_x(x,\xi) < 0$ for all $x > \hat{r}$. Finally, note that

$$H_\xi(x,\xi) = \frac{[1-\beta]\,\beta\int_x^R r dG(r)}{1-\beta(1-\xi+\xi G(x))} > 0.$$

Observe that for $\xi' > \xi$.

$$H(r,\xi) < H(r,\xi').$$

Now let

$$H(\hat{r}(\xi),\xi) = \hat{r}(\xi)$$

and suppose that $\hat{r}(\xi) < \hat{r}(\xi')$. Now this implies that $H(\hat{r}(\xi),\xi') < H(\hat{r}(\xi),\xi)$ since $H(\hat{r}(\xi),\xi')$ is increasing for $\hat{r}(\xi) < \hat{r}(\xi')$. This is a contradiction so $\hat{r}(\xi) \geq \hat{r}(\xi')$ as claimed. ■

Table 1: Across and Within Regime Performance

Panel A: Growth in percent (entire sample)

	measure based on executive constraints		measure based on executive constraints, open access to power		measure based on checks on the chief executive	
	strong executive constraints	weak executive constraints	strong executive constraints	weak executive constraints	strong executive constraints	weak executive constraints
mean	2.41	2	2.41	2.3	2.15	1.6
within sd	3.52	6.86	3.49	5.96	3.83	6.73
across sd	1.01	1.82	1.03	1.94	1.39	1.8

Panel B: Growth in percent (sample of switchers)

	measure based on executive constraints		measure based on executive constraints, open access to power		measure based on checks on the chief executive	
	strong executive constraints	weak executive constraints	strong executive constraints	weak executive constraints	strong executive constraints	weak executive constraints
mean	2.34	2.13	2.34	2.31	2.2	1.48
within sd	3.74	5.12	3.74	4.77	4.07	5.88
across sd	1.3	1.99	1.35	2.1	1.57	1.72

Note: Growth is real GDP pc growth. Sample is from 1952-2010 if we us executive constraints and from 1975-2010 if we define executive constraints through the checks on the chief executive. The entire sample is 130 countries in the first column, 120 countries in the second column and 127 countries in the third column. The switchers are 38 countries with at least three years under both strong and weak executive constraints. Within sd is the standard deviation of growth within country/institution episodes. Across sd is the standard deviation of growth across countries. For details see the discussion of equation (1) in the main text.

Table 2: Parameter Estimates of Δ

	(1)	(2)
	entire sample	sample of switchers
VARIABLES		
Δ_0	0.0118***	-0.0107
	(0.00425)	(0.0101)
Δ_L	-0.0181	-0.0321**
	(0.0113)	(0.0150)
Δ_H	0.154***	0.0842***
	(0.0424)	(0.0182)
Observatioı	172	76

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Table shows the fit of a non-linear least squares regression of productivity growth variances on productivity growth means at the country/institution level (equation (9) in the main text). Observations are weighted by episode length. The entire sample is 130 countries as described in the appendix. The switchers are 38 countries with at least three years under both strong and weak executive constraints.

Table 3: Robustness Comparisson of Weak and Strong Exceutive Constraints

Panel A: sample of switchers

	weak executive constraints			*strong executive constraints*			
Assumption on λ	λ	implied mean productivity growth	implied productivity variance	λ	implied mean productivity growth	implied productivity variance	output gain from switching
lower quartile λ	0.2766	0.0001	0.0027	0.1531	0.0038	0.0012	4.64%
λ=0	0.0000	-0.0321	0.0000	0.0000	-0.0107	0.0000	23.88%

Panel B: entire sample

	weak executive constraints			*strong executive constraints*			
Assumption on λ	λ	implied mean productivity growth	implied productivity variance	λ	implied mean productivity growth	implied productivity variance	output gain from switching
lower quartile λ	0.2936	0.0020	0.0028	0.2199	0.0102	0.0015	9.16%
λ=0	0.0000	-0.0181	0.0000	0.0000	0.0118	0.0000	34.85%

Note: The entire sample are 130 countries as described in the appendix. Switchers are 38 countries with at least three years under both strong and weak executive constraints. Lambdas are calculated from equation (8) in the main text and are reported in appendix Tables A4 and A5. The "lower quartile λ mean" is the mean in the lower quartile weighted by length of the country/institution episode. Mean and variance of productivity growth are calculated using Δ estimates from the respective column in Table (2). For details see the main text. We set λ=0 if λ<0. This affects 2 countries in the low quartile in Panel A under strong executive constraints.

Table 4: Robustness Comparisson of Weak and Strong Exceutive Constraints (with estimation error in Δ estimates)

Panel A: sample of switchers

	weak executive constraints			*strong executive constraints*			
Assumption on λ	λ	implied mean productivity growth at 95% lower bound	implied productivity variance	λ	implied mean productivity growth at 95% lower bound	implied productivity variance at 95% lower bound	output gain from switching
lower quartile λ	0.2766	-0.0094	0.0030	0.1531	-0.0011	0.0015	9.51%
λ=0	0.0000	-0.0614	0.0000	0.0000	-0.0304	0.0000	36.30%

Panel B: entire sample

	weak executive constraints			*strong executive constraints*			
Assumption on λ	λ	implied mean productivity growth	implied productivity variance	λ	implied mean productivity growth	implied productivity variance at 95% lower bound	output gain from switching
lower quartile λ	0.2936	-0.0100	0.0030	0.2199	-0.0007	0.0015	10.51%
λ=0	0.0000	-0.0401	0.0000	0.0000	0.0035	0.0000	54.72%

Note: The entire sample are 130 countries as described in the appendix. Switchers are 38 countries with at least three years under both strong and weak executive constraints. The λ's are calculated from equation (8) in the main text and are reported in appendix Tables A4 and A5. The lower quartile λ is the mean of the lower quartile weighted by the length of the country/institution episode. Mean and variance of productivity growth are calculated using Δ estimates and their standard error from the respective column in Table (2). Instead of the point estimate we use the lower bound of the 95% confidence intervall for each Δ. We set λ=0 if λ<0. This affects 2 countries in quartile 1 in Panel A under strong executive constraints.

Table A1: Comparison of Strong Executive Constraint Measures

		measure based on checks and balances		
		weak	strong	Total
measure based on executive constraints	weak	2,367	474	2,841
	strong	510	810	1,320
	Total	2,877	1,284	4,161

Table A2: Performance Controlling for GDPpc

Panel A: Growth in percent (entire sample)

	measure based on executive constraints		measure based on executive constraints, open access to power		measure based on checks on the chief executive	
	strong executive constraints	weak executive constraints	strong executive constraints	weak executive constraints	strong executive constraints	weak executive constraints
mean	2.41	2	2.41	2.3	2.15	1.6
within sd	3.51	6.86	3.47	5.95	3.81	6.73
across sd	1.18	1.94	1.33	2.09	1.53	1.92

Panel B: Growth in percent (sample of switchers)

	measure based on executive constraints		measure based on executive constraints, open access to power		measure based on checks on the chief executive	
	strong executive constraints	weak executive constraints	strong executive constraints	weak executive constraints	strong executive constraints	weak executive constraints
mean	2.34	2.13	2.34	2.31	2.2	1.48
within sd	3.72	5.12	3.72	4.77	4.06	5.88
across sd	1.81	2.29	1.7	2.29	1.76	1.92

Note: Growth is real GDP pc growth. Sample is from 1952-2010 if we us executive constraints and from 1975-2010 if we define executive constraints through the checks on the chief executive. The entire sample is 130 countries in the first column, 120 countries in the second column and 127 countries in the third column. The switchers are 38 countries with at least three years under both strong and weak executive constraints. Within sd is the standard deviation of growth within country/institution episodes. Across sd is the standard deviation of growth across countries. For details see the discussion of equation (1) in the main text. Regression adds level of GDP per capita on the right hand side.

Table A3: Robustness Delta Parameter Estimates

VARIABLES	(1) 5 years in both regimes	(2) only years with open access to power	(3) controlling for GDP per capita and year fixed effects	(4) measure based on checks on the chief executive
Δ_0	-0.0106	-0.00526	-0.0313***	-0.0273
	(0.0111)	(0.00807)	(0.00887)	(0.0213)
Δ_L	-0.0357*	-0.0196*	-0.0557***	-0.0550***
	(0.0194)	(0.00987)	(0.0139)	(0.0121)
Δ_H	0.0841***	0.0964***	0.0575***	0.0734***
	(0.0196)	(0.0200)	(0.0158)	(0.0183)
Observations	66	74	76	134

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Table shows the fit of non-linear least squares regressions of productivity growth variances on productivity growth means at the country/regime level. In columns (2)-(4) the sample is countries with at least three years under both strong and weak executive constraints. In column (1) the sample is countries with at least five years under both strong and weak executive constraints. Column (4) uses the alternative definition of checks on the chief executive to define strong constraints.

Table A4: Lambda Estimates in Detail (Switchers)

country	number of years under strong executive constraints	number of years under weak executive constraints	estimated λ under strong executive constraints	estimated λ under weak executive constraints
Albania	6	34	0.81	0.44
Argentina	6	53	0.13	0.43
Bolivia	27	32	0.17	0.29
Botswana	14	31	0.43	0.99
Bulgaria	21	19	0.34	0.72
Chile	22	37	0.58	0.41
Colombia	4	55	0.37	0.44
Croatia	11	9	0.41	0.21
Ecuador	17	42	0.11	0.52
France	6	53	0.53	0.48
Greece	40	19	0.46	0.58
Haiti	3	47	0.01	0.29
Hungary	21	19	0.26	0.51
Kenya	3	45	0.26	0.32
Lesotho	18	27	0.47	0.49
Macedonia	9	11	0.43	0.22
Madagascar	5	45	0.00	0.23
Malaysia	12	42	0.44	0.68
Mongolia	19	21	0.36	0.45
Nicaragua	16	43	0.29	0.32
Niger	4	46	0.00	0.23
Nigeria	11	40	0.04	0.42
Pakistan	15	44	0.32	0.49
Paraguay	18	41	0.20	0.44
Peru	12	47	0.47	0.40
Philippines	5	54	0.29	0.46
Poland	16	24	0.59	0.42
Portugal	29	30	0.33	0.69
Romania	7	43	0.60	0.64
Slovak Republic	13	5	0.54	0.82
Spain	33	26	0.30	0.72
Sri Lanka	26	33	0.34	0.65
Taiwan	7	52	0.58	0.79
Thailand	14	45	0.44	0.64
Turkey	33	26	0.38	0.50
Uganda	4	45	0.43	0.38
Uruguay	26	33	0.47	0.30
Zimbabwe	9	32	0.54	0.27

Note: The table reports λ estimates from country level growth data. See equation (8) in the main text. The Δ estimates are from Table (2), column 2. We set $\lambda=0$ if $\lambda<0$. This affects two countries under strong executive constraints. Switchers are countries with at least three years under both strong and weak executive constraints. For details see Appendix B.

Table A5: λ Estimates in Detail (Non-Switchers)

strong executive constraints			weak executive constraints					
country	number of years under strong executive constraints	estimated λ under strong executive constraints	country	number of years under weak executive constraints	estimated λ under weak executive constraints	country	number of years under weak executive constraints	estimated λ under weak executive constraints
Jamaica	52	0.1938	Zaire	51	0.1580	Russia	19	0.4481
Germany	21	0.2656	Central African Repu	50	0.2144	Cambodia	40	0.4516
New Zealand	59	0.2858	Somalia	40	0.2196	Afghanistan	40	0.4542
Switzerland	59	0.2867	Guinea-Bissau	37	0.2533	Mexico	59	0.4549
United States	59	0.3127	Guinea	51	0.2907	Malawi	47	0.4550
Sweden	59	0.3274	United Arab Emirates	24	0.2951	Congo	50	0.4636
Canada	59	0.3279	Senegal	50	0.2985	Mozambique	36	0.4660
Australia	59	0.3288	Togo	50	0.3088	Sierra Leone	49	0.4684
Costa Rica	59	0.3453	Papua New Guinea	36	0.3313	Jordan	56	0.4722
United Kingdom	59	0.3516	Cameroon	50	0.3364	Tanzania	50	0.4853
Netherlands	59	0.3641	Ivory Coast	50	0.3453	Syria	50	0.5029
Denmark	59	0.3666	Honduras	59	0.3556	Cuba	40	0.5081
Belgium	59	0.3839	Burundi	49	0.3575	Tunisia	49	0.5083
Slovenia	20	0.3918	Zambia	47	0.3621	Brazil	59	0.5211
Italy	59	0.4131	Venezuela	59	0.3682	Angola	36	0.5247
Norway	59	0.4148	Benin	51	0.3828	Yemen	21	0.5443
Finland	59	0.4185	Ghana	51	0.3896	Rwanda	50	0.5499
Israel	59	0.4233	Saudi Arabia	24	0.3935	Kuwait	24	0.5554
Austria	59	0.4397	Burkina Faso	51	0.3940	Iran	55	0.5589
Ireland	59	0.4401	El Salvador	59	0.3971	Egypt	59	0.5598
Czech Republic	18	0.4409	Guatemala	59	0.3980	Mauritania	50	0.5617
Trinidad & Tobago	49	0.4508	Libya	24	0.4034	Morocco	55	0.5638
Estonia	20	0.5095	Liberia	40	0.4129	Panama	59	0.5748
Mauritius	43	0.5118	Nepal	50	0.4165	Indonesia	50	0.6040
Serbia	5	0.5136	Mali	50	0.4172	Oman	40	0.6255
Japan	59	0.5524	Ethiopia	59	0.4181	Laos	40	0.6279
			Algeria	49	0.4195	Vietnam	35	0.6962
			Namibia	21	0.4224	Iraq	40	0.7786
			Chad	50	0.4245	China	58	0.8477
			Uzbekistan	20	0.4467	Bosnia and H.	19	1.0000

Note: Table reports estimates of λ from country level growth data. See equation (8) in the main text. Estimates of Δ are from Table (2), column 2. We set λ=1 if λ>1. This affects one country under weak executive constraints. For details see Appendix B.

Table A6: Summary of λ Estimates

Panel A: sample of switchers

	weak executive constraints			*strong executive constraints*			
Assumption on λ	λ	implied mean productivity growth	implied productivity variance	λ	implied mean productivity growth	implied productivity variance	output gain from switching
lower quartile λ	0.1703	0.0002	0.0019	0.0960	0.0045	0.0009	4.92%
λ=0	0.0000	-0.0196	0.0000	0.0000	-0.0053	0.0000	15.42%

Panel B: entire sample

	weak executive constraints			*strong executive constraints*			
Assumption on λ	λ	implied mean productivity growth	implied productivity variance	λ	implied mean productivity growth	implied productivity variance	output gain from switching
lower quartile λ	0.1440	-0.0029	0.0017	0.1543	0.0104	0.0014	14.39%
λ=0	0.0000	-0.0139	0.0000	0.3379	0.0110	0.0000	28.25%

Note: The entire sample are 120 countries. Switchers are countries with at least three years under both strong and weak executive constraints. The λ's are calculated from equation (8). The lower quartile λ is the mean of the lower quartle weighted by length of the country/institution episode. In the restricted sample mean and variance of productivity growth are calculated using delta estimates from Table A3, column (2). For details see the main text.

Table A7: Summary of λ Estimates

Panel A: sample of switchers

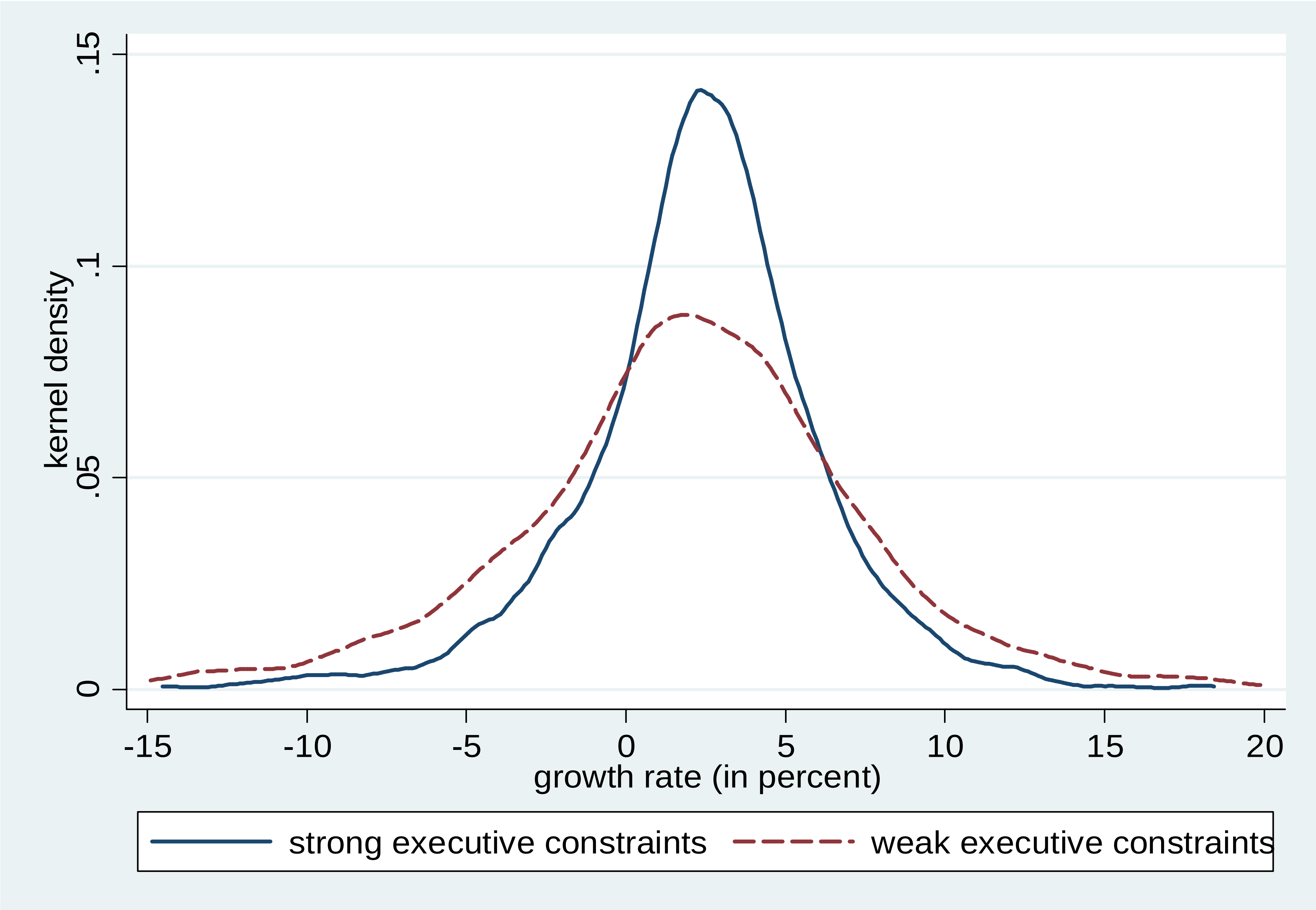
	weak executive constraints			*strong executive constraints*			
Assumption on λ	λ	implied mean productivity growth	implied productivity variance	λ	implied mean productivity growth	implied productivity variance	output gain from switching
lower quartile λ	0.4014	-0.0035	0.0040	0.3415	0.0071	0.0023	12.09%
λ=0	0.0000	-0.0550	0.0000	0.0000	-0.0273	0.0000	31.92%

Panel B: entire sample

	weak executive constraints			*strong executive constraints*			
Assumption on λ	λ	implied mean productivity growth	implied productivity variance	lambda	implied mean productivity growth	implied productivity variance	output gain from switching
lower quartile λ	0.4032	-0.0032	0.0040	0.3524	0.0082	0.0023	13.03%
λ=0	0.0000	-0.0227	0.0000	0.0000	0.0047	0.0000	31.44%

Note: The entire sample are 127 countries. Switchers are countries with at least three years under both strong and weak executive constraints. The λ's are calculated from equation (8). The lower quartile λ is the means of the lower quartile weighted by the length of the country/institution episode. In the restricted sample mean and variance of productivity growth are calculated using delta estimates from Table A3, column (4). For details see the main text.

Figure 1: Growth Rates and Executive Constraints

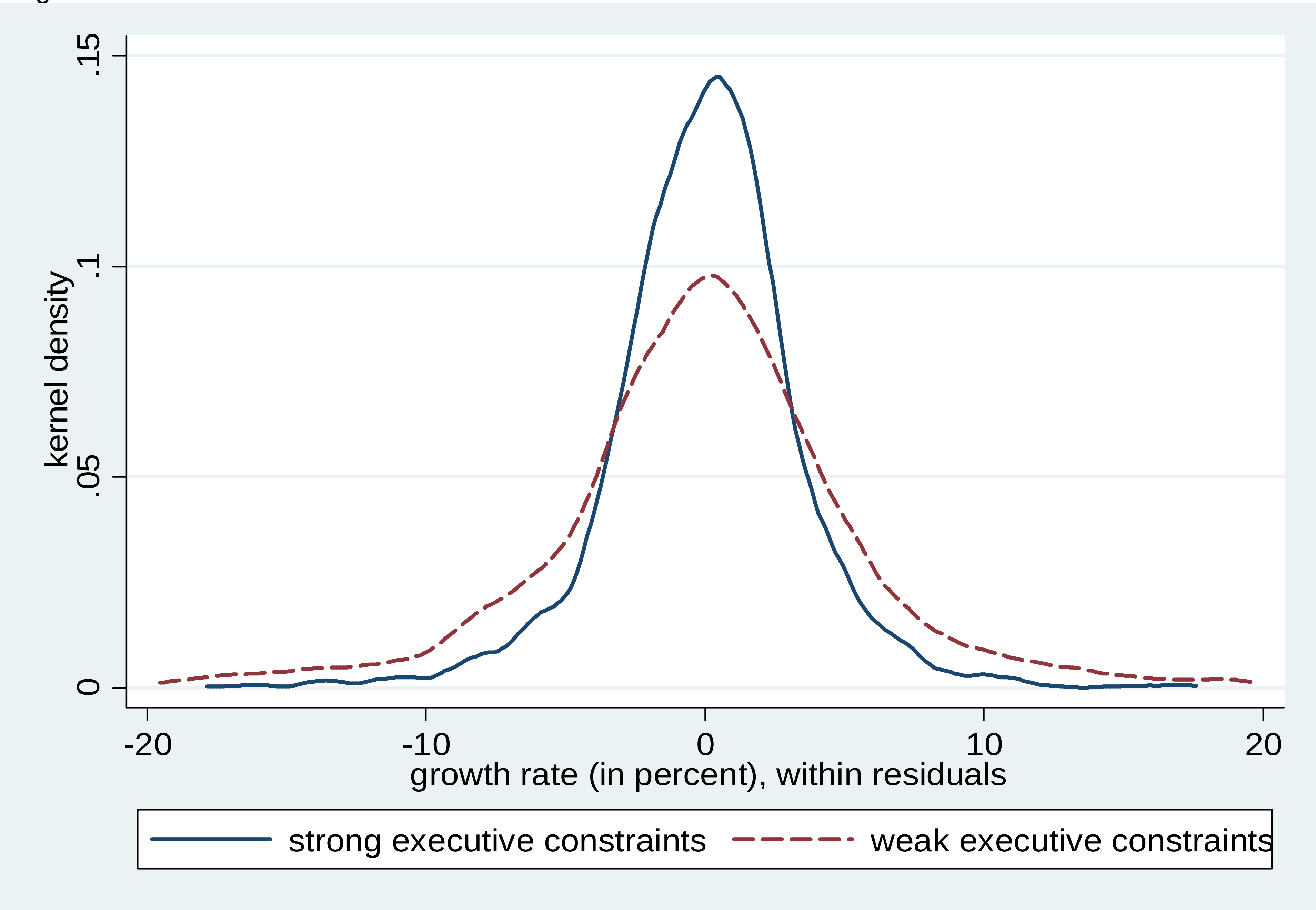


Note: Growth data is from 1952-2010 calculated using real GDP per capita from the Penn World Tables 7.1

Figure 2: Growth Rates and Openness of Executive Recruitment

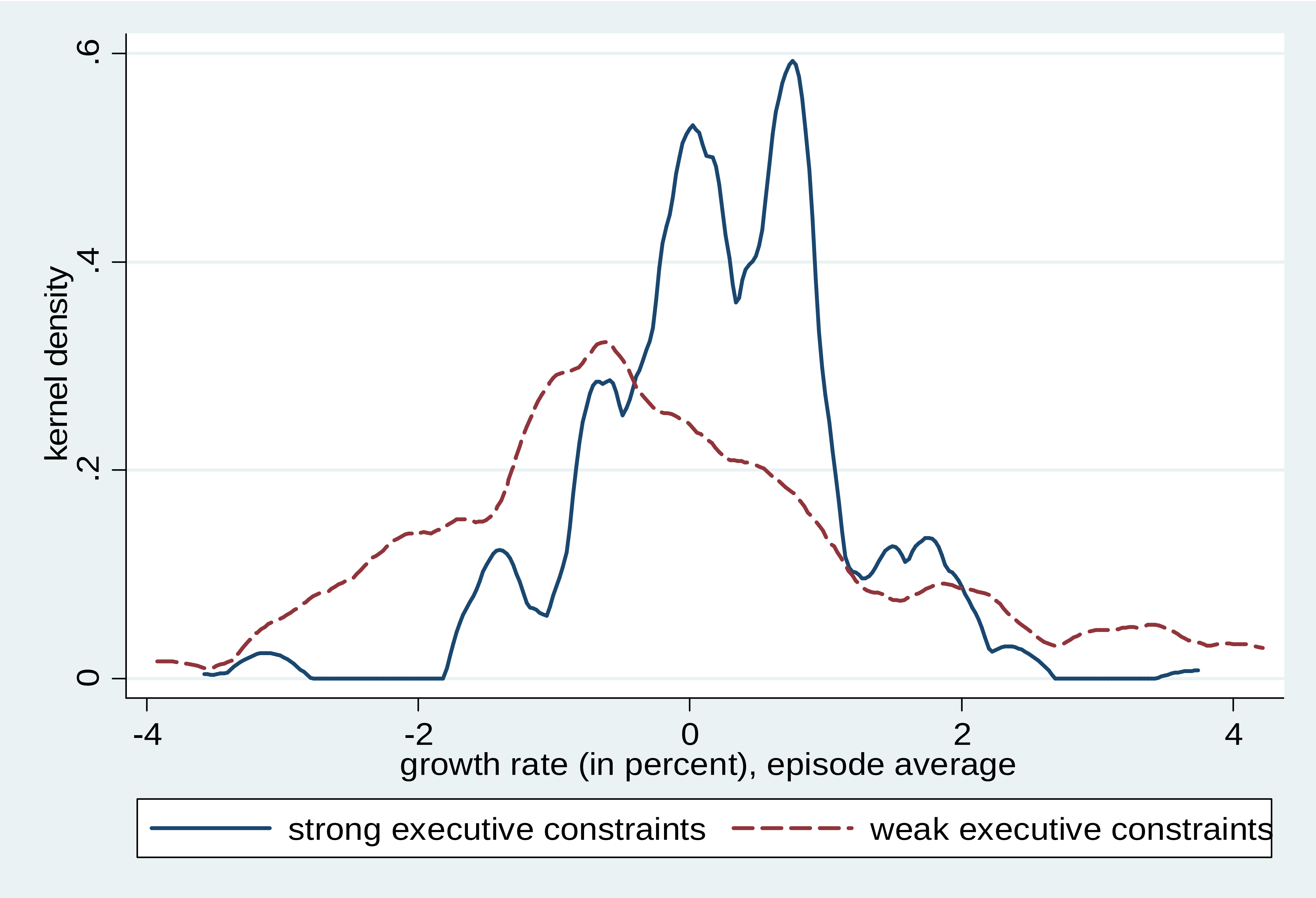
Note: Growth data is from 1952-2010 calculated using real GDP per capita from the Penn World Tables 7.1. The Figure excludes episodes with strong executive constraints.

Figure 3: Within Residuals of the Growth Rate



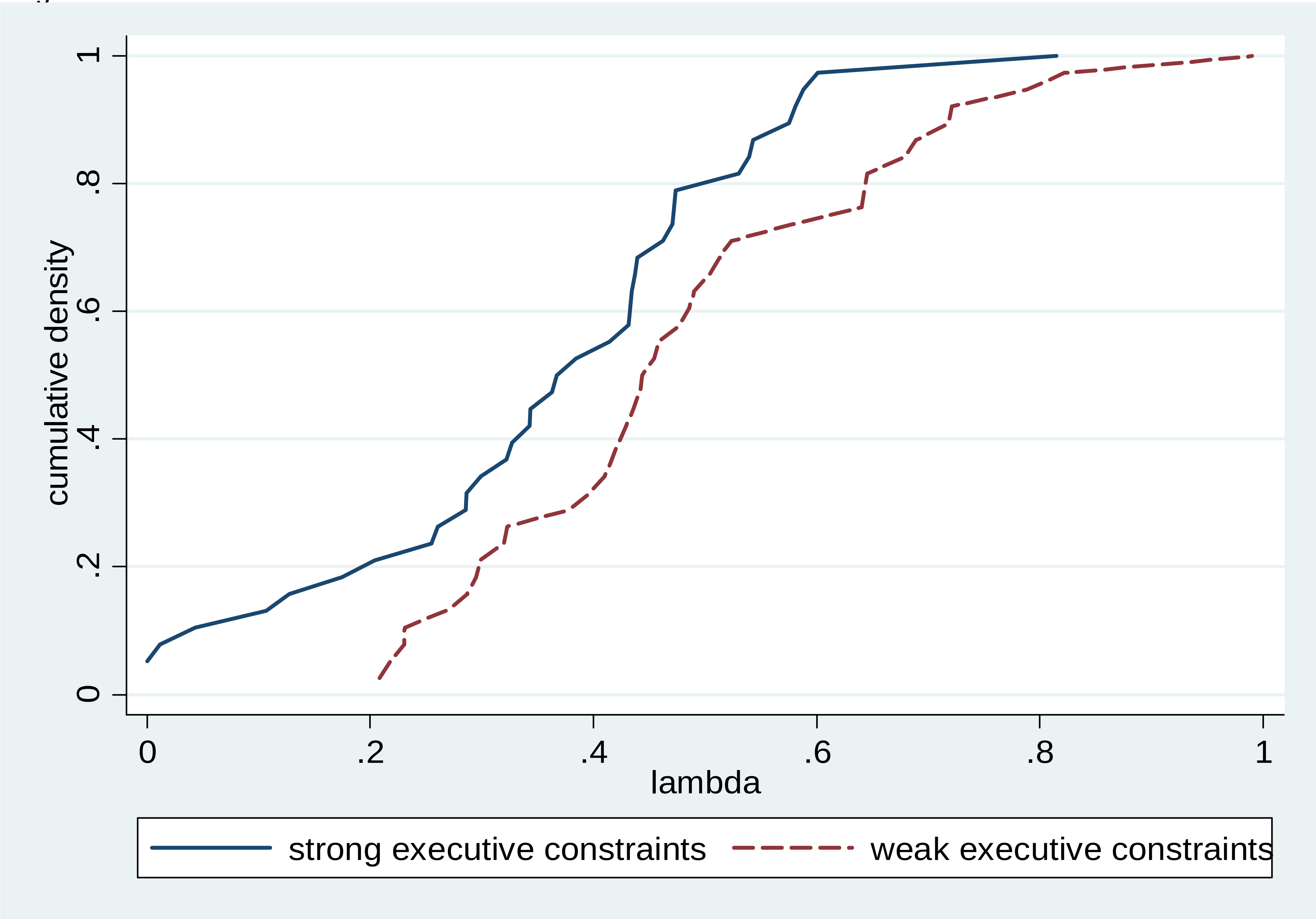
Note: The graph reports within country/institution residuals calculated from equation (1).

Figure 4: Episode Average Growth Rates



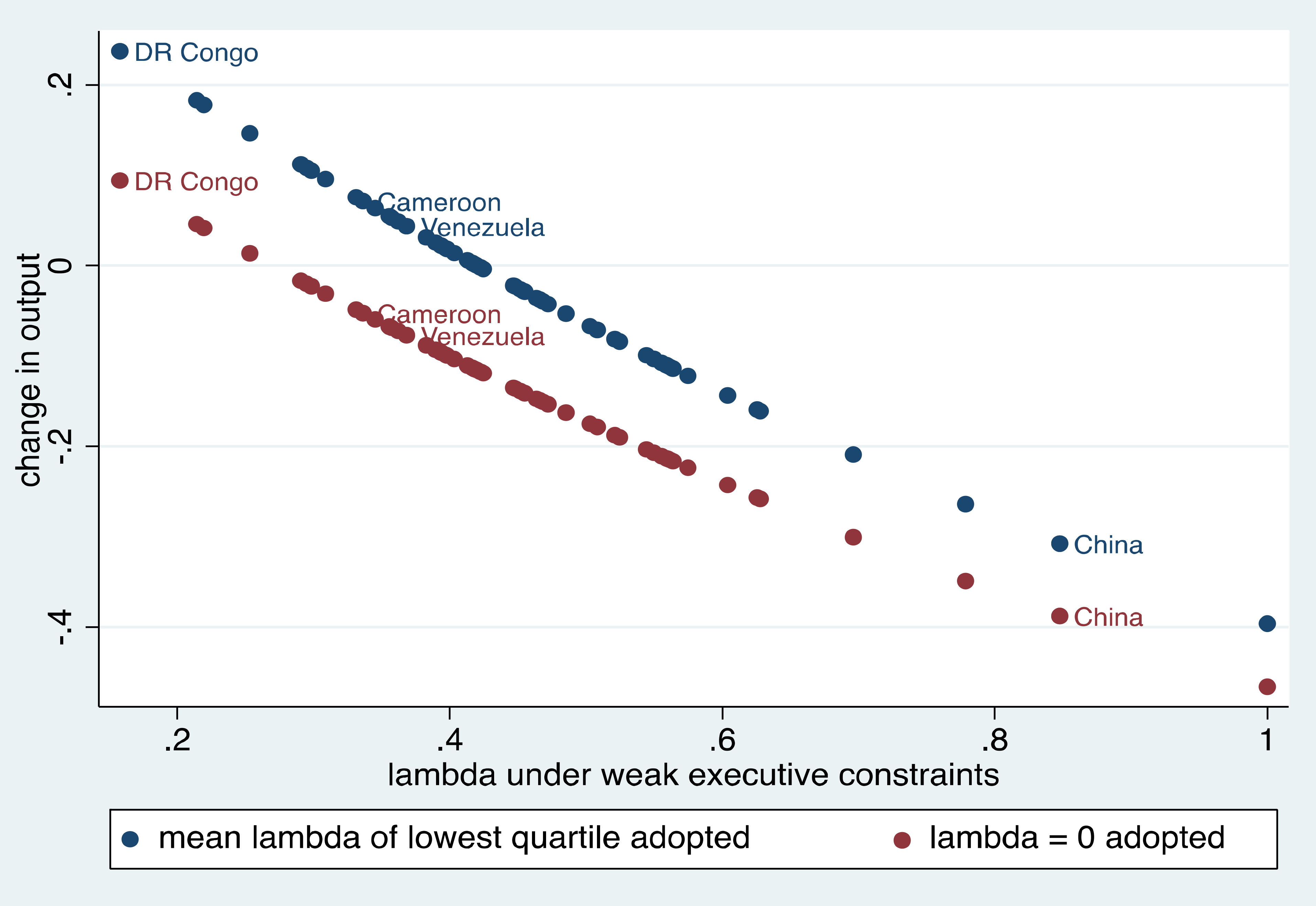
Note: The graph reports across growth averages for each country/institution episode calculated from equation (1).

Figure 5: Estimated Lambda Distributions



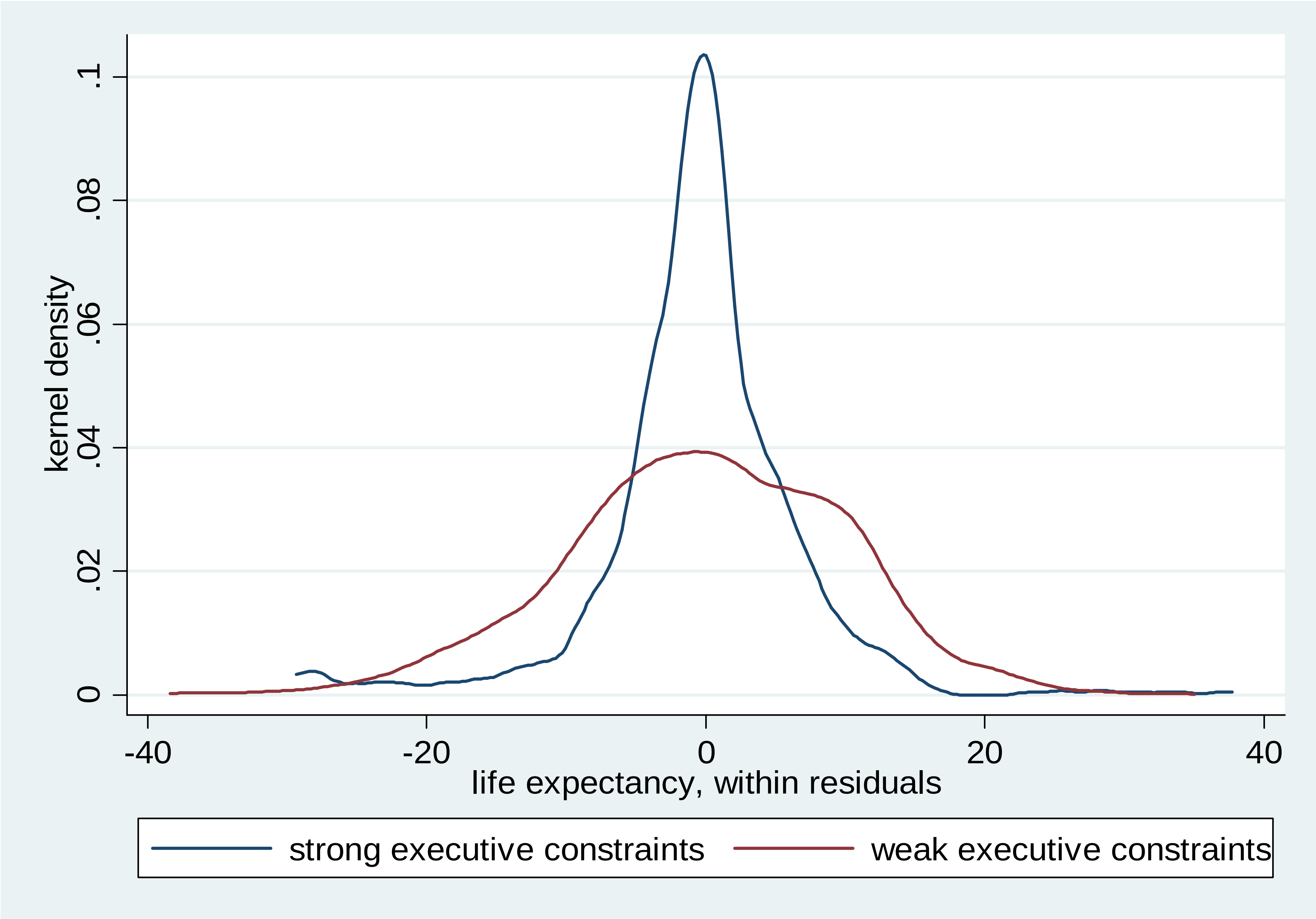
Note: Lambda values are from Table A4, i.e. countries that spent at least 3 years in weak and strong executive constraints.

Figure 6: Policy Experiment – Adoption of Strong Executive Constraints



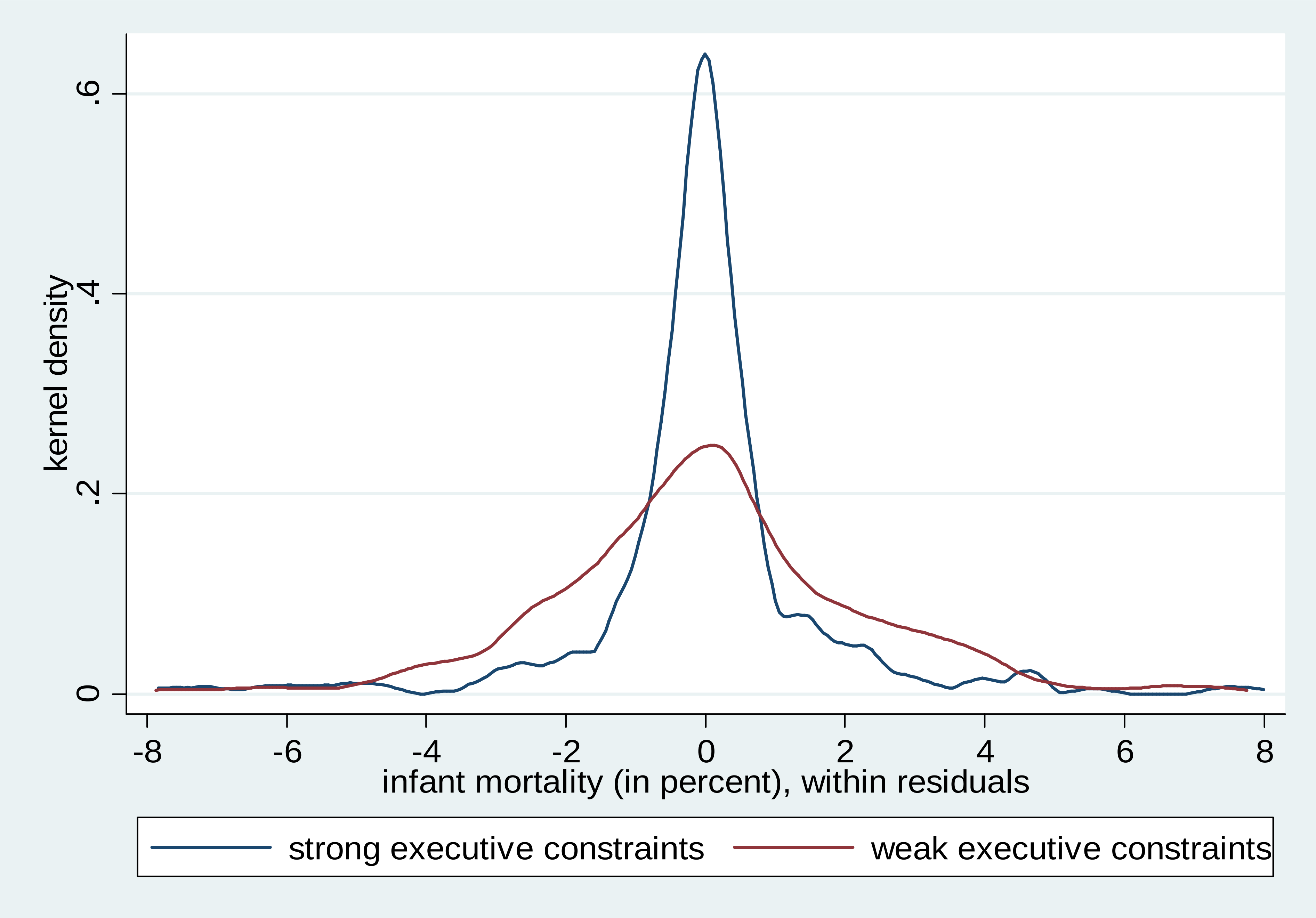
Note: The output gain is calculated using equation (14) and estimates in Table 2, column (2) and Table A5.

Figure 7: Within Residuals of Life Expectancy and Executive Constraints



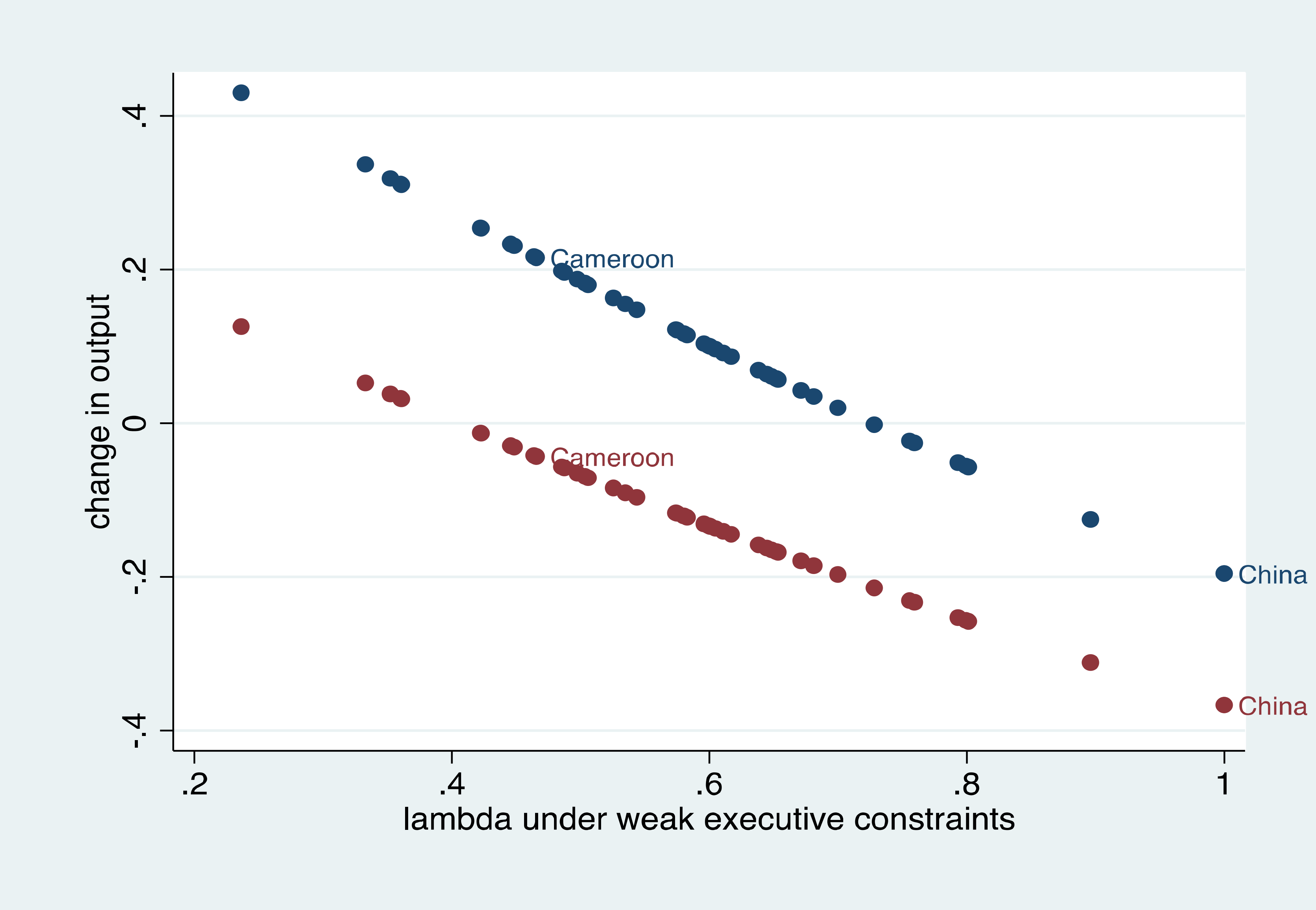
Note: Life expectancy data is from 1960-2010 from the World Development Indicators 2012.

Figure 8: Within Residuals of Infant Mortality and Executive Constraints



Note: Infant mortality data is from 1960-2010 from the World Development Indicators 2012.

Figure A1: Policy Experiment – Adoption of Strong Checks on the Chief Executive



Note: The output gain is calculated using equation (14) and estimates in Table A3, column (4).