NBER WORKING PAPER SERIES

ON THE OPTIMAL ALLOCATION OF POLICY-MAKING

Alessandro Dovis
Rishabh Kirpalani
Guillaume M. Sublet

Working Paper 33034
http://www.nber.org/papers/w33034

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2024

On the Optimal Allocation of Policy-Making
Alessandro Dovis, Rishabh Kirpalani, and Guillaume M. Sublet
NBER Working Paper No. 33034
October 2024
JEL No. E0, E60, P0

ABSTRACT

How should society allocate policy-making between the legislative and the executive branches of government?We analyze a model in which biased and polarized policymakers set policy in response to shocks. We show that policy issues for which the policy-maker bias is small relative to the degree of polarization should be delegated to the legislature, while policy issues where the bias is large should be delegated to the executive. Moreover, when executive delegation is preferred, it is optimal to leave little discretion and impose a narrow mandate. This finding contrasts with conventional wisdom that executive delegation allows for greater flexibility. The main difference between the two institutional settings is the ability to restrict ex post bargaining under executive delegation. Thus, when the bias is large, executive delegation is preferred because it can effectively constrain policymakers' choices. In contrast, when the bias is small, the ability to bargain ex post allows for flexible responses to severe shocks while limiting political risk. We also study the credibility of these institutions and show that while delegating to the legislature is typically credible, executive delegation is typically not when the bias is exogenous but can be when the bias arises from time inconsistency problems.

Alessandro Dovis
Department of Economics
University of Pennsylvania
The Ronald O. Perelman Center
for Political Science and Economics
133 South 36th Street
Philadelphia, PA 19104
and NBER
adovis@upenn.edu

Guillaume M. Sublet
University of Montreal
guillaume.sublet@umontreal.ca

Rishabh Kirpalani
Department of Economics
University of Wisconsin-Madison
Madison, WI 53706
rishabh.kirpalani@wisc.edu

> *In the legislature, promptitude of decision is oftener an evil than a benefit. The differences of opinion, and the jarrings of parties in that department of the government, though they may sometimes obstruct salutary plans, yet often promote deliberation and circumspection, and serve to check excesses in the majority....They constantly counteract those qualities in the Executive which are the most necessary ingredients in its composition — vigor and expedition, and this without any counterbalancing good.*
> Alexander Hamilton (Federalist No. 70)

1 Introduction

The allocation of policy-making authority is a fundamental issue in the design of political institutions. Institutions such as the legislature and executive agencies provide mechanisms, such as supermajority requirements and mandates, that help align policymakers' incentives with those of society. The goal of this paper is to understand how policy-making should be allocated between the legislature and executive agencies. The conventional view is that delegation to an agency such as a central bank allows for fast and responsive policy-making; in contrast, legislative bargaining in the presence of political polarization can be slow and lead to political gridlock, even on desirable policies. However, a major downside of delegation is the lack of direct accountability and representation.[1]

Building on this institutional context, this paper studies the optimal allocation of decision-making authority in a setting with three key features: policy-maker *bias* relative to societal preferences, shocks which create a benefit for *flexibility*, and political *polarization* between different societal factions who disagree on the optimal policy. Our main result is that delegating to the legislature is preferred if policy-maker bias is small relative to the degree of polarization, while delegating to an executive agency is preferred if policy-maker bias is large.

The main difference between the two institutional settings is the ability to restrict ex post bargaining under executive delegation. When the bias is large, society would like to constrain the choices of policymakers, which can be achieved by delegating to an agency with a narrow mandate. This contrasts with the aforementioned conventional view that society should delegate policy-making to executive agencies to increase flexibility. In contrast, when the bias is small, the ability to bargain ex post allows for flexible responses to severe shocks while imposing discipline on the faction in power because of the inability to act unilaterally and generating outcomes that are close to those preferred by society. We also show that these narrow mandates are credible if the policy-maker bias arises from a time inconsistency problem but not if there are exogenous preference biases.

[1] See for example Tucker (2019).

We consider a model where society must decide on a policy as a function of an aggregate shock. The population is divided into two factions who differ in their preferred policies, capturing political polarization. Factions are in power with an exogenous probability. Society's welfare is a weighted average of the factions' utilities plus a constant bias term, which can capture either differences in preferences between the factions and their representatives or time inconsistency issues. We compare two institutions for policy-making. The first, *executive delegation*, grants the executive representing the faction in power the authority to unilaterally choose a policy from a pre-specified set, which we term the delegation set. The second, *legislative bargaining*, allows the executive in power to unilaterally choose a policy from a given delegation set, but the executive can also implement a policy outside this set if the other faction agrees. This ability to renegotiate ex post in legislative bargaining is the critical feature that distinguishes the two institutions.

We first characterize the equilibrium outcomes under both institutions. We show that under executive delegation, the optimal delegation set is an interval which results in discretion for moderate shocks but imposes a cap and a floor for extreme shocks. Within this set, policy varies both because of the underlying state and the identity of the faction in power. The latter, which we term *political risk*, is undesirable from society's perspective. The size of the delegation set is chosen to optimally trade off this desire for flexibility with the costs associated with policy-maker bias and political risk. For example, if there are no bias and political risk, it is optimal to grant full flexibility to the policy-maker. On the other hand, if either policy-maker bias or political risk gets large enough, it is optimal to delegate only a single point to the executive. In contrast, under legislative bargaining, the optimal delegation set is discrete and has no intervals.[2] In other words, the faction in charge has almost no unilateral discretion.

Our main result provides conditions under which each of these two institutional settings is preferred from society's perspective. Fixing a level of polarization, we find that legislative bargaining is preferred if policy-maker bias is small enough, while executive delegation is preferred if the bias is large enough. To understand this result, first note that for a fixed delegation set, legislative bargaining gives policymakers more discretion to respond to shocks. Clearly, factions will agree only to policy that is in between the ideal policy of each faction. If the policy-maker bias is small, then this extra discretion through renegotiation is valuable to society. Thus, legislative bargaining is preferred when the bias is small. However, when the bias is large, delegation is preferred to legislative bargaining. The reason for this is that while bargaining increases policy flexibility, it also limits society's ability to correct the policy-maker bias. If the bias is large, this increased flexibility is detrimental to society and thus delegation to an executive is preferred.

[2]One can interpret this set as a generalized notion of the status quo in legislative bargaining games. For example, this set can represent mandatory spending levels as in Bowen et al. (2017).

We next show that policy-maker bias can also arise from time inconsistency problems. We make this point within the context of a Barro-Gordon model (Barro and Gordon, 1983) and show that our results extend to this setting. The key difference between the two environments concerns the credibility of narrow mandates. With a large exogenous bias, the gains from constraining policy through a narrow mandate are realized only by society. Therefore, narrow mandates are not credible, since both factions find it optimal to renegotiate ex post, thus bringing the chosen policy closer to their own preferences and away from society's. In contrast, in the Barro-Gordon model, each policy-maker would like to commit ex ante to not best respond to private actions in order to avoid the cost of expected inflation. Therefore, policymakers also value the ability to constrain policy choices, which makes the narrow mandate attractive if the time consistency problem is severe. In contrast, we show that legislative bargaining is always credible.

The key takeaway from our analysis is that allocating policy-making to executive agencies is desirable when the policy-maker bias is large not because of these agencies' ability to nimbly respond to shocks but rather the society's ability to impose mandates on these agencies which are enforced by the judicial branch. As we discuss in the paper, this finding can help interpret the recent discussions on the Chevron doctrine (U.S. Supreme Court, 1984) and its recent overturning. In contrast, delegating policy-making to the legislature is desirable when the polarization is large relative to the bias, because it allows society to respond to large shocks while simultaneously limiting political risk. In this sense, inaction by the legislature can be a feature and not a problem.

Our results can help inform decisions about which types of policies should be allocated to an executive agency versus assigned to a legislature. In our model, policies that suffer from severe time inconsistency problems should be delegated to an executive with a narrow mandate. This prediction is consistent with the delegation of monetary policy, arguably subject to time inconsistency problems, to central banks with inflation targeting. Our model also suggests that policies that suffer from significant polarization should be delegated to the legislature. This prediction is consistent with redistributive policies, which are arguably highly polarized, being chosen within the legislature and the observed responsiveness of such policies to large shocks but not to small shocks. For example, transfer programs do not vary much for small business cycle shocks, but there is often bipartisan agreement to expand these programs after large shocks such as the 2008 recession and the 2020 Covid pandemic.

Related Literature

Our paper builds on two distinct literature strands: one that studies legislative bargaining and another that studies delegation to a biased policy-maker. The main contribution of

this paper is to identify when one institutional setting is preferred over the other.

Our modeling of legislative bargaining builds on Baron and Ferejohn (1989). One of the key contributions of this and many other subsequent papers is to highlight the power of the proposer in bargaining; the proposer can always guarantee herself a disproportionate share of the surplus. The size of this share is determined by the outside option of the non-proposers, which in Baron and Ferejohn (1989) corresponds to a new round of voting, with a proposer chosen at random. Ali et al. (2019) show how even a little knowledge about future proposers can result in the proposer getting almost all the surplus. A related literature studies the legislative bargaining problem with an endogenous status quo. See, for example, Baron (1996), Kalandrakis (2004), Diermeier and Fong (2011), Bowen et al. (2014), Dziuda and Loeper (2016), Piguillem and Riboni (2015), and Ali et al. (2023). In these papers, the outcome of a previous round of negotiation becomes the status quo for the next round.[3] In contrast to these papers, in our paper, the status quo is part of the institutional design problem and not chosen by the policymakers.[4] Moreover, the status quo is potentially a set and not just one policy.

A large literature studies the trade-off between discipline and discretion in the delegation of policy-making authority in the context of monetary policy (Athey et al. (2005)) and fiscal policy (Halac and Yared, 2014, 2020, 2022a; Sublet, 2023). A main insight from the literature is that the optimal delegation set is an interval (Amador et al., 2006; Amador and Bagwell, 2013a). Sublet (2023) discusses when caps are optimal in an environment with money burning. The specific delegation problem we consider differs from the problems examined in these papers in that we model heterogeneity among policymakers, which maps to a multidimensional screening problem without transfers (see, for example Laffont et al. (1987), Rochet and Choné (1998), Moser and Olea de Souza e Silva (2019), and Boerma et al. (2022) for problems with transfers and Amador et al. (2003) for a problem without transfers).

Our paper shares some of its conclusions with a literature suggesting that delegating monetary policy to an external authority can help resolve the time inconsistency problems associated with it. Rogoff (1985) argues that delegating monetary policy to a central banker who places a larger weight on inflation stabilization can overcome the time inconsistency problem. This delegation is optimal because the executive in charge has the correct preferences to generate optimal policy outcomes from society's perspective. Instead, we assume that policy-maker preferences do not change across the legislature and the executive, and we show how mandates can be designed to alleviate the time inconsistency problem.

[3] Bowen et al. (2017) allow for a separation between current policy choices and future status quo.

[4] In this sense, our analysis is closer to Piguillem and Riboni (2021) study of the role of fiscal rules in legislative bargaining.

Our paper is also related to other principal-agent models of policy-making. We share our motivation with Alesina and Tabellini (2007, 2008), who highlight normative criteria for delegating policy-making to politicians versus bureaucrats. Unlike these authors, we assume that policymakers have identical preferences within the legislature and executive but that the different institutions differ in how policy is chosen. One of those institutions involves decision-making by a single agent (the executive), while the other involves joint decision-making by multiple agents (the legislature). In doing so we also study the optimal design of these institutions. Additionally, our paper has some similarities with Aghion et al. (2004), who study a constitutional design problem to understand how much unilateral power should be granted to factions in power, modeled as the size of the supermajority needed to pass a policy reform. Their problem is similar to our legislative bargaining problem in which society chooses the set of outside options for the faction in power. However, we aim to compare policy outcomes from such institutions that involve joint decision-making with one in which policies are chosen by a single executive.

The problem of separation of powers between the executive and legislature has been widely studied in the political science and legal studies literatures. Epstein and O'halloran (1999) provide an overview of the literature on the incentives of the legislature to delegate policy-making. The authors highlight the role of "political transaction costs" in determining when the legislature will decide to delegate. They argue that these delegation decisions are primarily driven by legislators' political goals. Aranson et al. (1982) and Rao (2015) argue that legislators have private incentives to delegate policy-making in order to garner benefits from interest groups and constituents, even if this weakens Congress institutionally. Thus the authors argue for stricter enforcement of the non-delegation doctrine by the courts. McCubbins et al. (1987) argue that the administrative procedures help align the incentives of executives with those of the legislature. Callander and Krehbiel (2014) show that under supermajoritarianism, the legislature may delegate to an executive agency to break the political gridlock. Volden (2002) studies how the discretion given to executive agencies changes under unified and divided government. We take a constitutional design perspective on delegation and argue that there are economic benefits to allocating certain types of policies to an executive agency.

The rest of the paper is organized as follows. Section 2 presents the main model. Section 3 and Section 4 study the optimal outcome under executive delegation and legislative bargaining, respectively. Section 5 delivers the paper's main result and establishes conditions when each institutional setting is preferable. Section 6 discusses the implications of our main results. Section 7 shows that our main results extend to a Barro-Gordon model. Section 8 studies the credibility of the institutional setting chosen at the constitutional stage. Finally, Section 9 concludes the paper.

2 Model

Consider a static model in which society must decide on a policy π as a function of an aggregate shock $z \in [\underline{z}, \bar{z}]$. The distribution of shocks is F with density f. The population is divided into two *factions*, or parties. Each faction has a type θ that is a preference parameter over the policy. Assume that $\theta \in \{\theta_L, \theta_H\}$ and the share of faction θ_i is α_i for $i = L, H$. We refer to $\Delta = \theta_H - \theta_L$ as the *polarization* between factions. The preferences of faction θ are $u(\pi, z, \theta)$. Societal welfare is a weighted average of the factions' utility plus a constant *policy bias* $\bar{\nu}$:

$$v(\pi, z) = \sum_i \alpha_i u(\pi, z, \theta_i) + \bar{\nu}\pi. \tag{1}$$

The bias $\bar{\nu}$ can capture the difference in preferences between the factions and their representative politicians. For example, politicians may be biased relative to society because they are prone to be captured by special interest groups or to having empire-building motives. In Section 7 we show that the bias can capture in a reduced form time inconsistency problems that arise due to a policy-maker that lacks the ability to commit to a policy plan.[5]

For simplicity, we assume that preferences are quadratic.

Assumption 1. *We assume* $u(\pi, z, \theta) = (z + \theta)\pi + b(\pi)$, *where* $b(\pi) = -\frac{1}{2}\pi^2$.

Note that the preferences can be alternatively represented as $u(\pi, z, \theta) = -\frac{1}{2}(\pi - \theta - z)^2$. All of our results go through for any $b(\pi)$ that is differentiable and strictly concave so that u is single peaked. Moreover, the assumption that θ and z enter additively is not essential to our analysis. For instance, our analysis applies with minor changes to environments in which the type and the state enter multiplicatively, as in $u(\pi, z, \theta) = z\pi + \frac{1}{\theta}b(\pi)$.

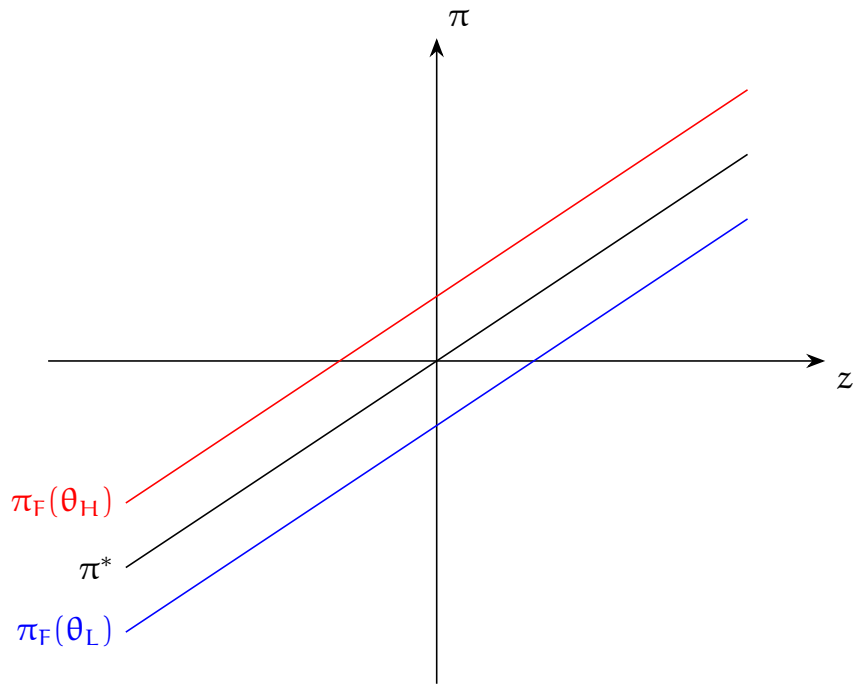
Type θ policymaker's preferred policy in response to a shock z is denoted by π_f, and it solves $-b'(\pi_f(z, \theta)) = z + \theta$. The preferred policy for society is $\pi^*(z)$, and it solves $-b'(\pi^*(z)) = z + \bar{\theta} + \bar{\nu}$, with $\bar{\theta} = \alpha_L\theta_L + \alpha_H\theta_H$. These preferred policies are illustrated in Figure 1 for $\bar{\nu} = 0$ and $\alpha_H = \alpha_L = 1/2$.

Both factions agree that the policy must be increasing in the aggregate state z. However, faction θ_H prefers a higher policy than does faction θ_L for all z. The distance between these preferred policies is the degree of polarization. Society's preferred policy is a convex combination of the factions' preferred policies if the bias is small (as in Figure 1), and it is lower than $\pi_f(\theta_L)$ (resp., higher than $\pi_f(\theta_H)$) if the bias is sufficiently negative (resp., positive).

It is useful to define as a benchmark the *best constant policy* that does not vary with the

[5] While we focus on a constant bias in this section, our results extend to environments with a non-constant bias, as in Amador and Bagwell (2013b).

Figure 1: Preferred policies for $\bar{\nu} = 0$



shock z.[6] This policy solves

$$\bar{\pi}^* = \arg\max_{\pi} \int \nu(\pi, z)\, f(z)\, dz,$$

and it satisfies $\bar{\pi}^* = \pi_f(E(z), E(\theta) + \bar{\nu})$.

Two institutional settings We next consider the problem of how society should delegate and structure the policy-making process. We consider two institutional settings that differ in how society delegates policy-making authority to the factions.

First, we consider *delegating to executives (DE)*. In DE, society delegates policy-making authority to an executive representing the faction in power. We assume that one faction is in power with probability α^i. In general, we allow $\alpha^i \neq \alpha_i$ and say that policy-making is *representative* if $\alpha^i = \alpha_i$. The executive chooses any policy within a delegated set of options D. Without polarization, $\Delta = 0$, our model nests the standard model studied in the literature on delegation without money burning.

Second, we consider *delegating to a legislature (LB)*. In this case, the two factions can *bargain* over the chosen policy, with the status quo being a policy chosen from a set D by the faction in charge. In other words, the faction in power can unilaterally choose a policy from a prescribed set D or choose another policy outside of the set D if the other faction agrees. This institutional setup is a simple version of legislative bargaining that has been

[6] Athey et al. (2005) refers to this policy as the expected Ramsey policy.

studied in the political economy literature. Typically in this literature the executive (or agenda setter) can unilaterally choose one policy only—D is a singleton. Here we study the optimal design of the set D.

While there are of course other possible institutional designs, we argue that these two capture the essential trade-offs inherent in the allocation of policy-making power. Delegation to an executive represents the transfer of the decision-making authority to a single agency, with the aim of promoting flexibility and responsiveness to changing circumstances. The head of the agency is appointed by the faction in power at the time and can choose any policy within the set given to the agency, its mandate. For example, this institutional setting captures the case of a central bank that has freedom to choose a policy within its mandate, and the central bank's governor is appointed by the faction in power that will appoint a governor with the same preferences.

Legislative bargaining, on the other hand, keeps authority in the hands of elected representatives, ensuring a degree of democratic accountability and control. The party in charge has some flexibility to choose among a given set of options (i.e., mandatory spending) but can choose outside the set if the opposition agrees (i.e., discretionary spending).

3 Delegation to executives

We characterize the optimal delegation to executive agencies. To do so, we extend the theory of optimal delegation to environments with two dimensions of private information. We show that under appropriate sufficient conditions, the optimal delegation set is an interval with a cap and a floor. Moreover, we show that for high levels of bias or high levels of polarization it is optimal to leave no discretion to policymakers, and the best outcome is the best constant policy defined in the previous section.

Formally, we can write the problem of society finding limits for executive discretion as a mechanism design problem without transfers, where society chooses a set of feasible policies D and the policy chosen by the executive of faction θ in state z must satisfy the incentive compatibility constraint

$$\pi(z, \theta) \in \arg\max_{\pi \in D} u(\pi, z, \theta). \tag{2}$$

This constraint highlights that under this setting the faction in power unilaterally chooses the policy independently of the other faction. Letting

$$w(\pi_L, \pi_H, z) = \alpha^L v(\pi_L, z) + \alpha^H v(\pi_H, z) \tag{3}$$

be the expected value for society in state z given that faction θ_i chooses policy π_i, society chooses a set D to solve

$$\max_{D\subseteq\mathbb{R},\pi(z,\theta)} \int_{\underline{z}}^{\bar{z}} w(\pi(z,\theta_L),\pi(z,\theta_H),z)f(z)dz \tag{4}$$

subject to (2).

The problem (4) can be equivalently written as

$$\max_{\pi(z,\theta)} \int_{\underline{z}}^{\bar{z}} w(\pi(z,\theta_L),\pi(z,\theta_H),z)f(z)dz \tag{5}$$

subject to

$$u(\pi(z,\theta),z,\theta) \geqslant u(\pi(\hat{z},\hat{\theta}),z,\theta) \quad \text{for } z,\hat{z}\in Z,\ \theta,\hat{\theta}\in\{\theta_L,\theta_H\}. \tag{6}$$

The main proposition in this section shows that under some sufficient conditions on the distribution of z, the optimal delegation set is an interval with a potentially binding cap *and* floor. To state these sufficient conditions, it is helpful to consider a Ramsey problem in which society is restricted to choosing an interval $[\pi_f(z_l,\theta_L),\pi_f(z_h,\theta_L)]$, which is parametrized by two thresholds, z_l and z_h. This problem is formally stated and characterized in Appendix A.

Assumption 2. *Define the distribution* $K(z) \equiv \alpha^L F(z) + \alpha^H F(z-\Delta)$, *with* $\kappa(z)$ *being the associated density and* E_κ, *the expectation with respect to this distribution. Let* $d(z) = \alpha^L\alpha_H f(z) - \alpha^H\alpha_L f(z-\Delta)$, $d_h(z) = -\frac{\int_z^{\bar{z}+\Delta} d(\hat{z})d\hat{z}}{1-K(z)}$ *and* $d_l(z) = \frac{\int_{\underline{z}}^{z} d(\hat{z})d\hat{z}}{K(z)}$. *We assume that*

1. $E_\kappa[\hat{z}|\hat{z}\geqslant z] - z + \bar{\nu} \leqslant (E_\kappa[\hat{z}|\hat{z}\geqslant z_h] - z_h + \bar{\nu})\frac{d_h(z)}{d_h(z_h)}$ *for all* $z\geqslant z_h$,
2. $z - E_\kappa[\hat{z}|\hat{z}\leqslant z] - \bar{\nu} \leqslant (z_l - E_\kappa[\hat{z}|\hat{z}\leqslant z_l] - \bar{\nu})\frac{d_l(z)}{d_l(z_l)}$ *for all* $z\leqslant z_l$,
3. $\Delta\frac{d'(z)}{\kappa(z)} + \bar{\nu}\frac{\kappa'(z)}{\kappa(z)} \leqslant 1$ *for all* $z\in(z_l,z_h)$.

The following proposition shows that the solution to the problem in (4) is an interval.

Proposition 1. *Suppose Assumptions 1 and 2 are satisfied. Then the solution to* (4) *is an interval delegation set* $D = [\pi_f(z_l,\theta_L),\pi_f(z_h,\theta_L)]$.

All proofs not in the main text are in the Appendix. We prove this proposition in two steps. First, we consider a Ramsey problem in which society is restricted to choosing an interval of the form in the proposition, and we characterize the optimal bounds (Ramsey step). Second, we show that no other set can improve upon this interval (verification step). The Ramsey step is straightforward. However, in contrast to the standard delegation problem, the verification step is complicated by the fact that there is uncertainty about both the state z and the type of the policymaker θ. First, in general, we cannot use the

first-order approach to simplify the constraints (6). We show, however, that under the functional form assumption in Assumption 1, the problem is equivalent to a model with only one dimension of uncertainty, $\theta + z$; so we can apply the tools developed by the literature to solve this problem.

There are two key substeps in the verification step. First, we need to prove that granting the executive discretion between the Ramsey bounds is indeed optimal. This is implied by the third condition in Assumption 2. Second, in order to show that interval delegation is optimal, we need to show that it is suboptimal to have two disconnected intervals in the delegation set. The usual argument for this is that disconnecting the interval will result in some types bunching at policies below their preferred one and others bunching at policies above their preferred one. If the policy-maker type is known, then there is certainty about which of these two bunching points generates benefits and costs. Therefore, a single bunching condition (condition 1 or 2 with $\Delta = 0$ in Assumption 2) needs to be imposed, in addition to condition 3, in order to guarantee that the costs outweigh the benefits. However, with multiple types, whether a particular bunching point imposes a benefit or cost depends on the realization of the type. Consequently, for $\Delta > 0$, we need to modify condition 3 in Assumption 2, and to impose an additional bunching condition under which interval delegation continues to be optimal in our setting.

Given that we have established that the optimal delegation set is an interval, we now show how uncertainty about the type of the policy-maker affects the choice of bounds of the interval. To do so, it is illustrative to consider the first-order condition of the Ramsey problem with respect to z_h:

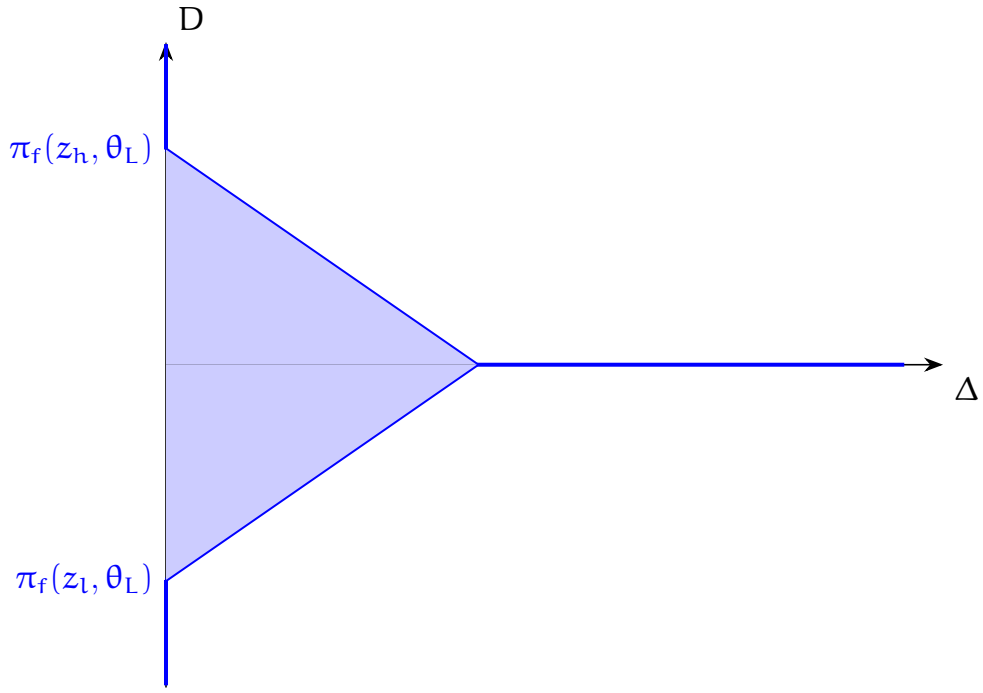
$$\Delta d_h(z_h) - \bar{\nu} = E_\kappa[z|z \geqslant z_h] - z_h.$$

In the absence of heterogeneity in θ, i.e., $\Delta = 0$, an optimal cap, if it is interior, equates the bias on the left-hand side to the loss of discretion on the right-hand side, $-\bar{\nu} = E_\kappa[z|z \geqslant z_h] - z_h$. Thus, a cap is binding only if the bias is negative, i.e., policymakers prefer a higher policy than society does.

With heterogeneity, i.e., $\Delta > 0$, the distribution K captures the higher loss of discretion for θ_H than that for θ_L. Heterogeneity adds a second component to the benefit of a cap, and d_h captures the conditional expectation of the net disagreement over the cap. The measure d_h captures the intensity and the sign of the expected disagreement. A specular logic applies for z_l.

The aforementioned first-order condition highlights a trade-off between *political risk* and flexibility that is absent in the case in which there is certainty about the policy-maker type. By political risk, we mean that the chosen policy varies depending on the type of the policy-maker in charge. To understand this trade-off, suppose that $\bar{\nu} = 0$. If $\Delta = 0$,

Figure 2: Optimal delegation sets as a function of polarization, Δ



then it is optimal for society to grant policymakers complete discretion in how they choose policy, since preferences are completely aligned. Suppose instead that $\Delta > 0$. Then, if society continues to grant complete discretion to policymakers, the chosen policy will on average align with the one preferred by society, but the presence of political risk will lead to undesirable movements in policy as a function of θ. Consequently, it is no longer optimal to grant complete discretion to policymakers. Similarly, for any level of Δ, a larger absolute value of $\bar{\nu}$ implies greater misalignment in preferences between society and policymakers, and consequently it is optimal to grant policymakers less discretion.

The next lemma shows that under a mild distributional condition, increases in both Δ and $|\bar{\nu}|$ reduce society's incentives to grant discretion to policymakers.

Lemma 1. *For some fixed $\Delta \geqslant 0$, there is a threshold $\bar{\nu}(\Delta)$ such that for $|\bar{\nu}| \geqslant \bar{\nu}(\Delta)$, the delegated set is a single point. Similarly, for some fixed $\bar{\nu}$, there is a threshold $\Delta(\bar{\nu})$ such that for $\Delta \geqslant \Delta(\bar{\nu})$, the delegated set is a single point. If the densities f and κ are log-concave, the size of the delegated interval decreases as the bias $\bar{\nu}$ gets more severe.*

Figure 2 illustrates the lemma for $\bar{\nu} = 0$: for a sufficiently large level of polarization, there is no discretion and the set D contains only the best constant policy.[7] We will refer to this situation as a *narrow mandate*.

[7] In the figure, the size of the delegation set is also monotonically decreasing in Δ, which is true in the numerical example.

4 Legislative bargaining

We now turn to study the best outcome under legislative bargaining. Relative to the previous case, the party in charge has the option to unilaterally choose a policy from the set D, but it may also choose $\pi \notin D$ if the other faction agrees. In this section we characterize the optimal set D. We show that, contrary to the case of executive delegation, the set D contains a discrete set of points and is not a continuous interval. The typical outcome has no responses to a small realization of shocks z, but it allows for some flexibility for a large realization of the shocks, in sharp contrast with the typical outcome under delegation.

To set up the problem, we define $\pi_D(z, \theta)$ as the policy that would be unilaterally chosen by the faction in charge, θ, from the set D. This policy choice is the status quo in bargaining and must satisfy (2). Given a shock realization z, if type θ is in power, it can implement any policy in the set $\mathcal{R}(z, \theta, D)$ where

$$\mathcal{R}(z, \theta, D) = \{\pi | u(\pi, z, \hat{\theta}) \geqslant u(\pi_D(z, \theta), z, \hat{\theta}),\ \hat{\theta} \neq \theta\}.$$

In other words, the faction of type θ can implement any policy which is preferred by the other type over the status quo.

We can then write the problem for society when it delegates policy-making to a legislature as

$$\max_{D \subseteq \mathbb{R}, \pi(z,\theta)} \int_{\underline{z}}^{\bar{z}} w(\pi(z, \theta_L), \pi(z, \theta_H), z) f(z) dz \tag{7}$$

subject to

$$\pi(z, \theta) \in \arg \max_{\pi \in D \cup \mathcal{R}(z,\theta,D)} u(\pi, z, \theta). \tag{8}$$

As described in the previous section, the presence of political risk creates undesirable movements in policy. We showed that under executive delegation, this risk is mitigated by limiting the discretion granted to executive agencies by constraining the bounds of an interval. In contrast, in the presence of bargaining, political risk can be mitigated by limiting the ability of the faction in charge to unilaterally choose policies. The way to do this is to restrict the set of unilateral policies available to the faction in charge, D. In particular, we show that when society's preferences are aligned with the policymakers' on average, it is optimal to grant discretion only over a discrete set of points. This finding is in sharp contrast with the previous section.

We say that a delegated set of policies D contains an interval if there is an interval $[z_l, z_h]$, with $z_l < z_h$, such that $\pi_f([z_l, z_h], \theta_L) \subseteq D$ or $\pi_f([z_l, z_h], \theta_H) \subseteq D$. The following lemma provides a condition that guarantees that it is never optimal to delegate an interval over a subset of z.

Lemma 2. *Suppose that for all* $z \in [z_l, z_h]$ *we have*

$$(\bar{v} + \alpha_H \Delta)\, \alpha^L f(z) + (\alpha_L \Delta - \bar{v})\, \alpha^H f(z - \Delta) > 0. \tag{9}$$

Then, the delegated set contains only a discrete subset of points from the interval $[\pi_f(z_l, \theta_L), \pi_f(z_h, \theta_L)]$.

The proofs of the results in this section are in Appendix B.

We can then prove our main result for this section:

Proposition 2. *Suppose that* $\bar{v} \in [-\alpha_H \Delta, \alpha_L \Delta]$. *Then, the set* D *that solves* (7) *does not contain intervals. If* $\bar{v} > \alpha_L \Delta$ *and for* $z \in [\underline{z} + \Delta, \bar{z}]$ *condition* (9) *holds, then the set* D *contains a discrete subset of points over* $[\pi_f(\underline{z} + \Delta, \theta_L), \pi_f(\bar{z}, \theta_L)]$ *and grants full discretion for* $\pi > \pi_f(\tilde{z}_1, \theta_L)$ *where* $\tilde{z}_1 \in [\bar{z}, \bar{z}, +\Delta)$ *and no discretion for* $\pi < \pi_f(\underline{z} + \Delta, \theta_L)$. *Similarly, if* $\bar{v} < -\alpha_H \Delta$ *and for* $z \in [\underline{z}, \bar{z} - \Delta]$ *condition* (9) *holds, then the set* D *contains a discrete subset of points over* $[\pi_f(\underline{z} + \Delta, \theta_L), \pi_f(\bar{z}, \theta_L)]$ *and grants full discretion for* $\pi < \pi_f(\tilde{z}_2, \theta_L)$, *where* $\tilde{z}_2 \in (\underline{z}, \underline{z} + \Delta]$, *and no discretion for* $\pi > \pi_f(\bar{z}, \theta_L)$.

To understand the difference with the pure delegation case and Proposition 1, note that when society allocates policy-making to a legislature, the choice of the set D determines the policy chosen when the two factions cannot find a mutually agreeable policy—similarly to DE—but it also affects the bargaining power of the faction in charge. Allowing for greater discretion in the status quo set increases the bargaining power of the faction in charge and reduces the other faction's ability to discipline policy.

Suppose that $\bar{v} \in [-\alpha_H \Delta, \alpha_L \Delta]$ and the set D contains an interval $[\pi_f(z_l, \theta_L), \pi_f(z_H, \theta_L)]$. Suppose also that the θ_L faction is in power. In this case, for any $z \in [z_l, z_h]$, the low type chooses exactly its preferred policy, so bargaining imposes no discipline on the choice of policy. Now suppose we break the interval into two disconnected intervals ("drilling a hole") by removing interval $(\pi_f(z_{l1}, \theta_L), \pi_f(z_{H1}, \theta_L))$ from the status quo set. The low type cannot unilaterally implement its preferred policy for any z in this range. In particular, in this case, the outside option for the low type is $\{\pi_f(z_{l1}, \theta_L), \pi_f(z_{H1}, \theta_L)\}$, which in turn implies a lower bargaining power and induces greater moderation in the policy choice. That is, the policy does not vary too much in the direction preferred by the θ_L type. Since $\bar{v} \in [-\alpha_H \Delta, \alpha_L \Delta]$ and society's preferences are a convex combination of policymakers' preferences, this moderation is welfare-improving. Note that the status quo set D can contain multiple such points if the range of z is large enough. However, if the range of z is small enough, it may be optimal to have a unique status quo. This is particularly appealing because we can interpret this case as one where there is a given level of mandatory spending (the singleton D) that can be changed by doing some discretionary spending only if there is agreement in the legislature.

If instead $\bar{v} > \alpha_L \Delta$, then society's preferences are always closer to those of faction θ_H, which implies that this moderation is no longer desirable. In this situation, the best case for society is if type θ_H always implemented its flexible policy. However, society cannot just give full bargaining power to the faction in charge, since the low type would also choose its flexible policy when it is in charge. We show that for intermediate ranges of policy, it is optimal to limit bargaining power by delegating only a discrete set of points, while for policies large enough so that they would never be chosen by the θ_L type, it is optimal to grant full bargaining power to the faction in charge. A symmetric logic holds for the case in which $\bar{v} < -\alpha_H \Delta$.

Finally, we show that if the degree of polarization in society is sufficiently large, then it is not possible to reach any agreement between factions and the best that society can do is to set the status quo to the best constant policy, $D = \{\overline{\pi}^*\}$, that is not changed in the bargaining phase:

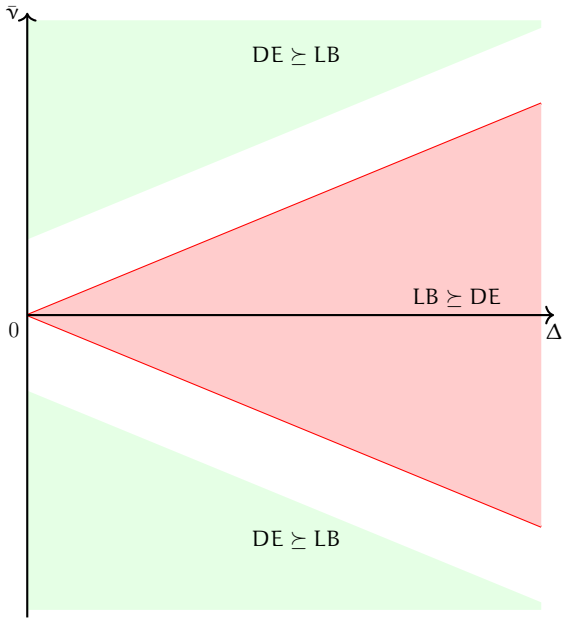
Lemma 3. *Fix some $|\bar{v}| > 0$. Then, there exists a $\Delta(|\bar{v}|) > 0$ such that for all $\Delta > \Delta(|\bar{v}|)$, the set $D = \{\overline{\pi}^*\}$, where $\overline{\pi}^*$ is the best constant policy, and the allocation implemented by LB is $\pi_{LB}(z, \theta) = \overline{\pi}^*$.*

Proof. For any given $\bar{v}$ suppose Δ is large enough so that $\pi_f(\bar{z}, \theta_L) < \pi^*(\underline{z})$ and $\pi^*(\bar{z}) < \pi_f(\underline{z}, \theta_H)$. We first show that in this case implementing the best constant policy as an outcome is feasible. To see this, suppose that $D = \{\overline{\pi}^*\}$. Then, for any z, $\mathcal{R}(z, \theta, D) = \{\overline{\pi}^*\}$ for any z, θ. This is because for any z, $u(\pi, z, \theta_H) > u(\overline{\pi}^*, z, \theta_H)$ for any $\pi > \overline{\pi}^*$, and $u(z, \theta_L, \pi) < u(z, \theta_L, \overline{\pi}^*)$ for any $\pi < \overline{\pi}^*$. We now show that this set D is optimal from society's perspective. We know from Proposition 2 that the set D is a discrete set. Given our assumption on Δ, it must be that $D \subset [\pi^*(\underline{z}), \pi^*(\bar{z})]$. Now suppose the set $D \neq \{\overline{\pi}^*\}$. In particular, suppose that $D = \{\pi_1, .., \pi_N\}$, with $\pi_1 < ... < \pi_N$. Then, the outcome under LB is $\pi_{LB}(z, \theta_H) = \pi_N$ and $\pi_{LB}(z, \theta_L) = \pi_1$, for all z. Clearly, this outcome is dominated by the best constant policy from society's perspective. Q.E.D.

5 How should society allocate policy-making?

In this section, we analyze the conditions under which delegation to an executive agency (DE) or legislative bargaining (LB) is optimal for societal welfare. Our main result is that for any level of polarization Δ, if the bias $\bar{v}$ is small enough, then society prefers LB to DE. If instead the bias is large enough, then society prefers DE to LB. To this end, we compare the solutions to (2) and (8) for any choice of D and show that this comparison alone provides conditions for which one of the two institutional settings dominates in terms of welfare, independently of the choice of the delegation set.

Figure 3: Preferred institutional setting as a function of polarization, Δ, and policy-maker bias, $\bar{\nu}$



Proposition 3. *For any $\Delta > 0$, there exist thresholds $\bar{\nu}_L(\Delta), \bar{\nu}_H(\Delta)$ such that society weakly prefers LB to DE if $|\bar{\nu}| < \bar{\nu}_L(\Delta)$, with the preference being strict if Δ is small enough, and society strictly prefers DE to LB if $|\bar{\nu}| > \bar{\nu}_H(\Delta)$. If $\Delta = 0$, society prefers DE to LB.*

Proof. Suppose that $|\bar{\nu}| < \max\{\alpha_H\Delta, \alpha_L\Delta\}$. Then, we can write society's utility as a convex combination of the two policy-maker types' utility. In fact, let γ be defined by $\bar{\nu} = -\gamma\alpha_H\Delta + (1-\gamma)\alpha_L\Delta$. Note that for $\bar{\nu} \in [-\alpha_H\Delta, \alpha_L\Delta]$, $\gamma \in [0,1]$. With this definition, we can write

$$\begin{aligned} v(\pi, z) &= \left(z + \sum_i \alpha_i\theta_i + \bar{\nu}\right)\pi + b(\pi) \\ &= ((1-\gamma)\theta_H + \gamma\theta_L)\pi + b(\pi) \\ &= (1-\gamma)u(\pi, z, \theta_H) + \gamma u(\pi, z, \theta_L), \end{aligned} \tag{10}$$

which is indeed a convex combination of policy-maker preferences.

Next, consider an arbitrary delegation set D. Under DE, the equilibrium outcome is

$$\pi_D(z, \theta; D) \in \arg\max_{\pi \in D} u(\pi, z, \theta).$$

Under LB, the outcome solves

$$\pi_{LB}(z, \theta; D) \in \arg\max_{\pi} u(\pi, z, \theta)$$

subject to

$$u\left(\pi, z, \hat{\theta}\right) \geqslant u\left(\pi_D\left(z, \theta; D\right), z, \hat{\theta}\right) \quad \hat{\theta} \neq \theta.$$

Thus, it must be that for all z

$$u\left(\pi_{LB}\left(z, \theta; D\right), z, \theta\right) \geqslant u\left(\pi_D\left(z, \theta; D\right), z, \theta\right), \tag{11}$$

$$u\left(\pi_{LB}\left(z, \theta; D\right), z, \hat{\theta}\right) \geqslant u\left(\pi_D\left(z, \theta; D\right), z, \hat{\theta}\right) \quad \hat{\theta} \neq \theta, \tag{12}$$

where the first inequality is a strict inequality whenever there is an agreement to go outside of the set D. We can then write the welfare for an arbitrary delegation set D as

$$\begin{aligned}\sum_i \alpha^i \int v\left(\pi_D\left(z, \theta_i; D\right), z\right) f\left(z\right) dz &\leqslant \sum_i \alpha^i \int v\left(\pi_{LB}\left(z, \theta_i; D\right), z\right) f\left(z\right) dz, \\ &\leqslant W_{LB} = \max_D \sum_i \alpha^i \int v\left(\pi_{LB}\left(z, \theta_i; D\right), z\right) f\left(z\right) dz,\end{aligned}$$

where the first inequality follows from the fact that we can write v as as convex combination of policymakers' utility and (11)–(12), and the second inequality follows from optimality. Since the second inequality holds for any D, it must also hold for the delegation set that attains the maximum under DE. Thus, we have $W_{LB} \geqslant W_{DE}$. Note that from Proposition 1 and Lemma 1, we know that for small Δ, the optimal delegated set is a strict interval. From Proposition 2, we know that the optimal set D under LB contains no interval and strictly dominates any interval. Therefore, it must be that $W_{LB} > W_{DE}$.

Next, assume that $|\bar{v}| \geqslant \text{sd}\left(z\right) + \max\{\alpha_H\Delta, \alpha_L\Delta\}$. We will show that delegating a single point to the executive achieves a higher welfare than any allocation that can be implemented via LB. There are two cases to consider. Suppose first that $\bar{v} \leqslant -\text{sd}\left(z\right) - \alpha_H\Delta$. Consider the best constant policy, $\bar{\pi}^* = \pi_f\left(E\left(z\right), E\left(\theta\right) + \bar{v}\right)$. Under DE it is always feasible to implement the best constant policy so

$$W_{DE} \geqslant \int_{\underline{z}}^{\bar{z}} v\left(\pi_f\left(E(z), E\theta + \bar{v}\right), z\right) f(z) dz.$$

Moreover, when $\bar{v} \leqslant -\alpha_H\Delta$, $v\left(\pi_f\left(z, \theta_L\right), z\right)$ is an upper bound for social welfare under legislative bargaining, since the equilibrium policy is in $[\pi_f\left(z, \theta_L\right), \pi_f\left(z, \theta_H\right)]$ and society's preferences are more aligned to those of type θ_L factions. Thus,

$$\int_{\underline{z}}^{\bar{z}} v\left(\pi_f\left(z, \theta_L\right), z\right) f(z) dz \geqslant W_{LB}.$$

When $\bar{v} \leqslant -\text{sd}\left(z\right) - \alpha_H\Delta$ and u is quadratic, direct calculations show that society's value of following the constant policy is higher than the one of following type θ_L's preferred

policy, or

$$\begin{aligned}\int_{\underline{z}}^{\bar{z}} v\left(\pi_f\left(E(z), E\theta+\bar{v}\right), z\right) f(z) dz &= \frac{1}{2}(E(z)+E\theta+\bar{v})^2 \\ &> \int\left((z+E\theta+\bar{v})(z+\theta_L)-\frac{1}{2}(z+\theta_L)^2\right) f(z)\, dz \\ &= \int_{\underline{z}}^{\bar{z}} v\left(\pi_f\left(z, \theta_L\right), z\right) f(z) dz.\end{aligned}$$

Thus, $W_{DE} > W_{LB}$. A symmetric argument holds for the case when $\bar{v} > \text{sd}(z) + \alpha_H \Delta$.

Finally, suppose that there is no bias, $\Delta = 0$. Under LB, factions can always implement their preferred policy, $\pi_{LB}(z, \theta; D) = \pi_f(z, \theta)$. This outcome is always feasible under DE since society can just grant full discretion to the policy-maker. Then we have $W_{LB} \leqslant W_{DE}$. Q.E.D.

Fixing the level of political polarization $\Delta > 0$, if the bias $\bar{v}$ is small enough, then society prefers LB to DE. This is true irrespective of the distribution of shocks z, because for a moderate bias, for any delegation set, the policy implemented by a legislature dominates the policy implemented by the executive in terms of welfare for any realization of the shock.

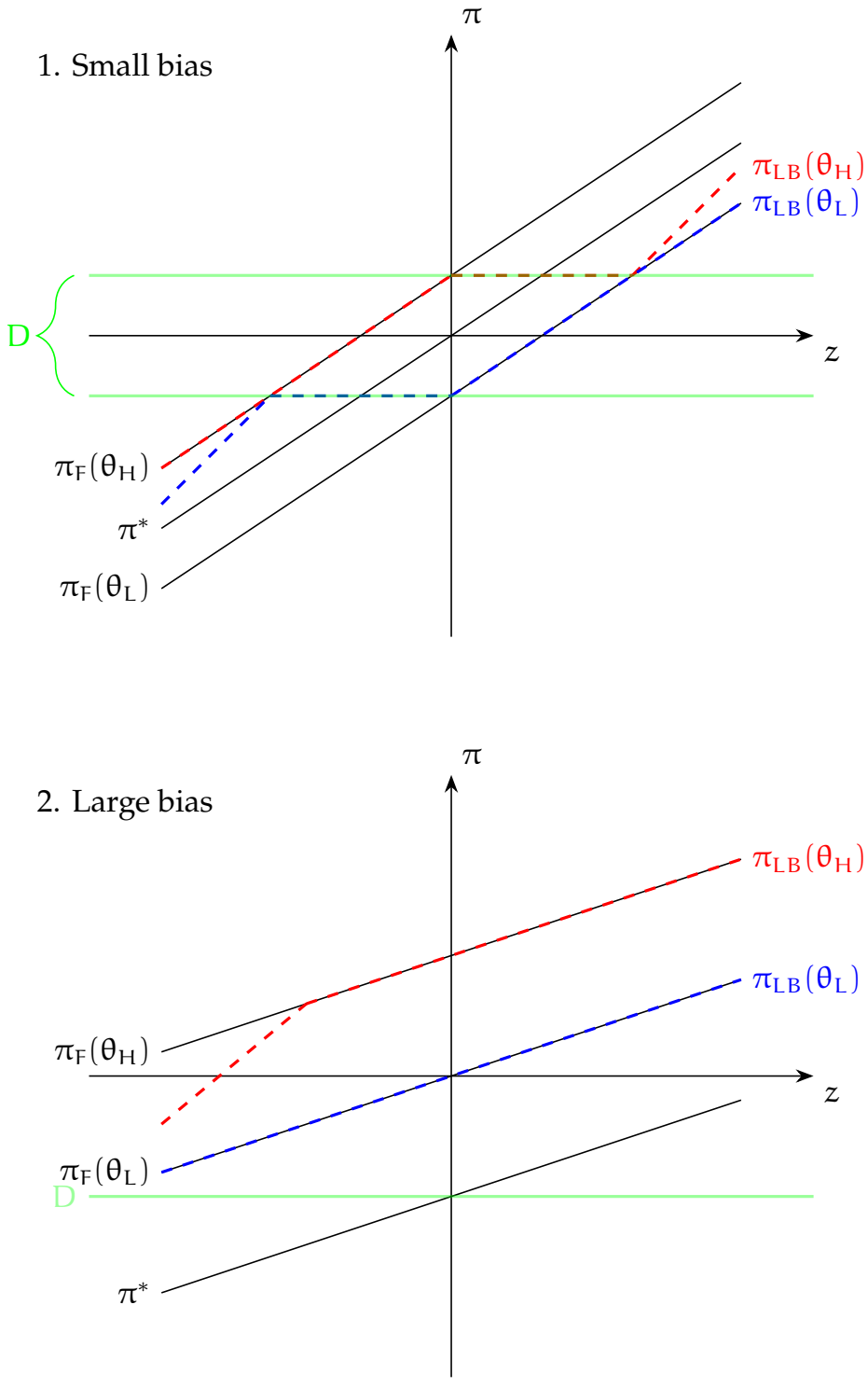
To understand this result, first note that for a fixed delegation set D, bargaining gives policymakers more discretion to respond to shocks. In fact, through bargaining, factions can agree to implement policies beyond those that are at the discretion of the executive. Clearly, factions will agree only to policies that are in between the ideal policies of each factions. This is beneficial if the bias $\bar{v}$ is small (in absolute value) because

$$\pi_f(z, \theta_L) \leqslant \pi_{LB}(z, \theta; D) \leqslant \pi_f(z, \theta_H).$$

Thus, the renegotiated policy is closer to society's preferences than what can be unilaterally chosen by each agent, as illustrated in the first panel of Figure 3.

Intuitively, when $|\bar{v}|$ is small, the flexible policy chosen by the average type in response to z is close to society's optimal policy; therefore, granting flexibility to the average type is desirable. However, under DE, allowing for more discretion with respect to z also implies greater political risk, since π will also respond to θ. Thus, there is a trade-off between flexibility and political risk. In fact, as Lemma 1 shows, as Δ gets large, it becomes optimal to eliminate this political risk even at the cost of eliminating any flexibility with respect to z. In the case of LB, however, bargaining between factions moderates this political risk and generates outcomes close to what the average type would have chosen, which is valuable to society. This scenario manifests itself in political gridlock and inaction for intermediate realizations of z but flexibility for large realizations of z, where both factions prefer to

Figure 4: Outcomes for DE and LB given a delegation set, D



renegotiate. Note that in this case, gridlock is a *feature* of the optimal institutional choice because it avoids costly political risk. Moreover, the ability to respond to these large shocks generates benefits relative to the case of DE, which does not allow such renegotiation and limits response to large shocks with a cap and floor.

In contrast, when $|\bar{v}|$ is large, the preferences of society and those of the average type are severely misaligned. In this case, society would like to limit the discretion granted to the average type. As already discussed, limiting discretion under LB for large shocks is difficult since policymakers can always choose to renegotiate policies outside D. However, through DE society can limit policymakers' discretion by granting a narrow mandate. In fact, under the bound provided in the proposition, delegating a single point to the executive achieves a higher welfare than any allocation that can be implemented via LB. This is because factions can ex post agree to a policy that is farther away from society's preference for any z than the policy with no discretion. This scenario is illustrated in the second panel of Figure 4.

Eventually, as the level of polarization increases, the outcomes from the two institutions coincide. This is because under LB, renegotiation between the two factions becomes harder, and under DE, it is better to implement a narrow mandate in order to avoid political risk.

Proposition 4. *Fix some $|\bar{v}| > 0$. Then for Δ large enough, society is indifferent between DE and LB because they both implement the best constant policy.*

This follows as a corollary from Lemmas 1 and 3.

6 Implications

We now discuss some takeaways from our analysis. First, our results emphasize that delegating policy-making to executive agencies is desirable when the policy-maker bias is large not because of the executive agencies' ability to nimbly respond to shocks but rather because of society's ability to impose mandates on these agencies that are enforced by the judicial branch. In contrast, delegating to the legislature is desirable when the polarization is large relative to the policy-maker bias, because it allows society to respond to large shocks while simultaneously limiting political risk. In this sense, Congress inaction can be a feature and not a problem.

The case of environmental regulation is a good illustration of the contrast between the conventional view of policy-making and the one argued for in this paper. As noted by Klyza and Sousa (2013), the increased role for the executive in environmental policy was seen as a solution to the political gridlock on environmental issues in Congress; however, this increase in executive discretion has led to significant political risk in the enacted

policies. For example, the authors highlight work which shows that fine and violation notices to polluters dropped significantly during the George W. Bush administration as compared with the Clinton administration. Our paper argues that political gridlock in Congress is a feature of the system that can limit this political risk.

In summary, our results caution strongly against allocating policy-making to an agency because of political gridlock within the legislature. As emphasized previously, this gridlock is part of the optimal mechanism and is used to limit undesirable political risk in policy outcomes. For example, broadening central bank mandates to include objectives such as climate change or inequality to get around Congress inaction might lead to unwanted political risk and excessive volatility in these outcomes.

Second, our model implies that when it is optimal to delegate policy-making to an executive agency, a relatively narrow mandate should be imposed since delegation is valuable precisely because of the ability of the judicial system to enforce these mandates. Our results can help interpret the recent discussions on the Chevron doctrine and its recent overturning. The Chevron doctrine (U.S. Supreme Court, 1984), established in 1984, held that courts should defer to an agency's interpretation of an ambiguous statute if the interpretation was reasonable. This principle granted executive agencies substantial discretion within the framework provided by Congress and led to political risk in enacted policies, as in the case of the aforementioned environmental regulation. However, the recent *Chevron vs. Loper Bright* (U.S. Supreme Court, 2024) decision marks a significant shift in that courts now take a more active role in interpreting statutes, reducing the deference previously given to agencies. Our results imply that narrow mandates under delegation are optimal, and this finding aligns with the recent ruling that can be interpreted as a tightening of mandates. The *Chevron vs. Loper Bright* decision reflects a judicial preference for more precise legislative guidelines and reduced agency discretion, which is consistent with our result that when policy bias is significant and it is optimal to delegate, tighter control over executive discretion is beneficial. Of course, this argument relies on judges who are immune from bias and political risk; otherwise, policies delegated to the executive can continue to suffer from undesirable political risk.

A classic example of a policy handled by the legislature is income tax policy, while an example of policy that is delegated to an executive agency is monetary policy. Through the lens of our model, allowing Congress to decide on personal income tax policy is optimal if the degree of polarization on this issue is large relative to the bias. Arguably, there is significant disagreement between income tax policies between the different factions and parties in Congress. Our results also provide predictions on how tax policy should respond to shocks under LB. Indeed, consistent with observation, taxes are typically not responsive to small shocks (for example, business cycle shocks) but are responsive to large shocks (for example, Covid checks) where all factions agree on how to change these

policies.

In contrast, in our model, delegating to a central bank is optimal if monetary policy suffers from a large bias relative to polarization. This bias can arise from time inconsistency, a link that we make more precise in Section 7. Observationally, the fact that we see relatively high-frequency movements in monetary policy might seem at odds with a narrow mandate. However, the movements in instruments (for example, interest rates) might still be consistent with a narrow mandate on outcomes, as is indeed the case for most central banks that have inflation-targeting mandates. Within the inflation-targeting framework, the central bank has full flexibility on how to achieve this mandate, but it has no flexibility on choosing its target. While our baseline model is silent on the discussion of instruments versus targets (see, for example, Atkeson and Kehoe (2001); Atkeson et al. (2007); Halac and Yared (2022b)), we could easily extend our model to incorporate this distinction. In particular, suppose that π corresponds to inflation, and $\pi = g(i, z)$, where i is the nominal interest rate controlled by the central bank. All our results would go through in this framework and would call for a narrow mandate on π but full flexibility on the nominal interest rates.

7 Time inconsistency and source of the bias

So far, we have assumed that the policy-maker bias, $\bar{\nu}$, is an exogenous preference parameter. For example, politicians may be biased relative to society because they are prone to be captured by special interest groups. In this section, we show that bias can also arise from time inconsistency problems. We make this point within the context of a Barro-Gordon model and show that our results extend to this setting.

We follow Athey et al. (2005) and let x denote the nominal wage inflation, u_n denote the unemployment rate, and z denote the state of the economy. Suppose that z is distributed with f over a support $Z = [\underline{z}, \bar{z}]$ and with mean $E(z) = 0$. The shock affects the desired level of inflation, and there is *heterogeneity* in this desired level. The payoff of type θ when the state of the economy is z is $-\frac{1}{2}\left[u_n^2 + (\pi - z - \theta)^2\right]$, where π denotes the growth rate of money (which also pins down the price inflation). Unemployment is determined by a static Phillips curve $u_n = U + x - \pi$, where U denotes the natural rate of unemployment. Substituting the Phillips curve in the payoff of type θ gives

$$R(x, \pi, z, \theta) = -\frac{1}{2}\left[(U + x - \pi)^2 + (\pi - z - \theta)^2\right].$$

A *monetary authority* of type θ sets the growth rate of money $\pi(z, \theta)$ as a function of the shock. Firms have rational expectations and set wage growth to equate the expected

money growth rate

$$x = \sum_i \alpha^i \int \pi(z, \theta_i) f(z) dz, \tag{13}$$

where α^i denotes the probability that the policy authority is of type θ_i. We refer to (13) as the implementability constraint. The timing of events is the following: firms set nominal wage growth x, the state z is realized, and with probability α^i the monetary authority θ_i sets the money growth rate.

Under commitment, society's preferred allocation solves

$$\max_{\pi(\cdot)} \sum_i \alpha_i \int R(x, \pi(z), z, \theta_i) f(z) dz$$

subject to the implementability constraint (13). An alternative interpretation of the aforementioned preferences is that they can arise from a Nash-bargaining game between the different factions with bargaining weights α_i. The solution, which we call the full information Ramsey allocation, is $\pi^*(z) = \bar{\theta} + \frac{z}{2}$, and price setters rationally expect $x = \bar{\theta}$, where $\bar{\theta} = \sum_i \alpha_i \theta_i$.The constant Ramsey policy is $\bar{\pi}^* = \bar{\theta}$.

Without commitment, the monetary authority solves $\max_\pi R(x, \pi, z, \theta)$ taking x as given. The best response is then $\pi_f(z, \theta; x) = \frac{U+x+\theta+z}{2}$. If private agents believe that the monetary authority is acting with full discretion, using (13), the expected nominal wage inflation is $x = U + \bar{\theta}$ and $\pi(z, \theta) = \bar{\theta} + \frac{z}{2} + U + \frac{\theta - \bar{\theta}}{2}$. The static Nash outcome differs from society's preferred allocation in two respects. There is an inflation bias U due to the monetary authority's attempt to exploit the Phillips curve as well as due to political risk $(\theta - \bar{\theta})/2$. These differences motivate the institutional choice problem described in the previous section. In particular, we characterize when it is optimal to use an executive agency (DE) versus the legislature (LB) to conduct monetary policy.

We can then write the problem for DE as

$$\max_{D, \pi(z,\theta), x} \sum_i \sum_j \alpha_i \alpha^j \int R(x, \pi(z, \theta_j), z, \theta_i) f(z) dz \tag{14}$$

subject to the incentive compatibility constraint for all θ

$$\pi(z, \theta) = \arg\max_{\pi \in D} R(x, \pi, z, \theta)$$

and the implementability constraint (13).

The problem for LB can be written as

$$\max_{D \subseteq \mathbb{R}, \pi(z,\theta), x} \sum_i \sum_j \alpha_i \alpha^j \int R(x, \pi(z, \theta_j), z, \theta_i) f(z) dz \tag{15}$$

subject to

$$\pi(z,\theta) \in \arg \max_{\pi \in D \cup \mathcal{R}(z,\theta,D)} R(x,\pi,z,\theta) \tag{16}$$

and the implementability constraint (13).

To see the connection with our baseline model, consider the Lagrangian for the two problems mentioned earlier, and letting λ be the Lagrange multiplier on the implementability constraint, it is as if society had a bias with preferences

$$R\left(x,\pi(z),z,\bar{\theta}\right) - \lambda\pi(z).$$

Thus, the term $\lambda\pi(z,\theta)$ captures the bias term $\bar{\nu}\pi$ in the case analyzed so far. The main difference is that now the bias is endogenous to the allocation and dependent on the severity of the time-inconsistency problem as captured by the term U.

The next proposition shows that the conclusion of Proposition 3 extends to this economy where the severity of the time-inconsistency problem U plays the role of the exogenous bias $\bar{\nu}$:

Proposition 5. *For any $\Delta > 0$, there exist thresholds $U_L(\Delta), U_H(\Delta)$ such that society weakly prefers LB to DE if $U < U_L(\Delta)$, with the preference being strict if Δ is small enough, and society strictly prefers DE to LB if $U > U_H(\Delta)$. If $\Delta = 0$, society prefers DE to LB.*

8 Credibility

As shown in the previous sections, executive delegation is valuable if mandates can be enforced and factions can be prevented from renegotiating ex post. However, in practice, since the decision to delegate is typically passed by Congress, it can choose to either change the mandate to allow for greater discretion or choose the policy within the legislature. For example, Congress has the authority to change the mandate of the central bank or decide monetary policy itself. Therefore, it is important to understand if the institutional arrangements described earlier are *credible*.

We say that an institutional arrangement is credible if the legislature cannot agree to change it. Formally, let $V^i_{\hat{\theta}}(D,\theta)$ denote the expected value to faction $\hat{\theta}$ of having faction θ set the policy with unilateral discretion within D in institutional setting $i \in \{DE, LB\}$. An institutional setting consisting of $i \in \{DE, LB\}$ and a delegated set D is *credible* if for all possible factions in power, such faction, θ, prefers (i, D) to any alternative institution $(j, \hat{D})$, that the other faction $\hat{\theta} \neq \theta$ would agree to. That is,

$$V^i_\theta(D,\theta) \geqslant \max_{\hat{D}} \left\{ V^j_\theta(\hat{D},\theta) \;\middle|\; V^j_{\hat{\theta}}(\hat{D},\theta) \geqslant V^i_{\hat{\theta}}(D,\theta) \right\} \quad \text{for all } \theta.$$

Proposition 6. *Suppose* $|\bar{v}| > \bar{v}_H(\Delta)$ *so that DE is preferred to LB. Then DE is not credible.*

Proof. If $|\bar{v}| > \bar{v}_H(\Delta)$ we know from the proof of Proposition 3 that society prefers the best constant policy to the preferred policy of type θ_H (assuming $\bar{v} > 0$). Then the DE outcome cannot be type θ_H's preferred policy. Assume now that type θ_H is in charge. Type θ_H can obtain full flexibility since

$$\int u\left(\pi_{DE}(\theta_H, z), \bar{\theta} + \bar{v}, z\right) f(z) > \int u\left(\pi_f(\theta_H, z), \bar{\theta} + \bar{v}, z\right) f(z),$$
$$\int u(\pi_{DE}(\theta_H, z), \theta_H, z) f(z) < \int u(\pi_f(\theta_H, z), \theta_H, z) f(z)$$

imply that

$$\int u(\pi_{DE}(\theta_H, z), \theta_L, z) f(z) < \int u(\pi_f(\theta_H, z), \theta_L, z) f(z)$$

under our assumption on preferences. Q.E.D.

In contrast, for a bias resulting from a strong degree of time inconsistency, then a well-designed narrow mandate delegated to the executive is credible.

Proposition 7. *For all* Δ, *there exists* $\bar{U}_H \geqslant U_H(\Delta)$ *such that for all* $U \geqslant \bar{U}_H$ *DE is preferred to LB and the optimal DE outcome is a credible narrow mandate at the constant Ramsey policy.*

With a large exogenous bias, the gains from preventing ex post renegotiation are realized only by society. Therefore, narrow mandates are not credible, since both factions find it optimal to renegotiate ex post, thus bringing the chosen policy closer to their own preferences and away from society's. In contrast, in the Barro-Gordon model, each policymaker would like to commit ex ante to not best respond ex post to x, to avoid the cost of expected inflation. Therefore, policymakers value the ability to not renegotiate ex post, which makes the narrow mandate attractive if the time-consistency problem is severe. One way of attaining this benefit is by delegating policy-making to an agency that follows the constant Ramsey policy. However, this is costly for two reasons. First, policy can no longer respond to the shock z, and, second, the inflation target differs from the one preferred by any individual type θ (since it corresponds to the average type $\bar{\theta}$). We show that the benefits outweigh the costs for all types if the degree of the time inconsistency, U, is large relative to the volatility of z and the degree of polarization.

Note that it is critical that the mandate be chosen before x is realized, since commitment is valuable only if x can be manipulated. If not, the narrow mandate would no longer be credible, since it would always be optimal to best respond to x. However, in a dynamic environment, the narrow mandate can continue to be credible by standard reputational arguments (Abreu (1988)). Interestingly, political polarization makes it easier to support good outcomes. This is because the faction in charge cannot unilaterally choose its best

response, as it must obtain the approval from the opposition. In turn, this limitation reduces the static gains of deviating and makes deviations less attractive. This mechanism is similar to the one in Chari et al. (2020) for a monetary union.

In contrast to Proposition 6, LB is always credible, as we show next. To fix ideas, suppose that the set D in LB contains a single point π_o.

Proposition 8. *Suppose we are in the environment with exogenous bias and that* $D = \{\pi_o\}$. *Then* LB *is always credible.*

Proof. Suppose the high type is in charge. First notice that it is not feasible for type θ_H to change D (with the approval of θ_L), since increasing π_o is beneficial for θ_H but detrimental for θ_L. Consider the problem of delegating policy-making to a decisionmaker who has identical preferences so long as the other faction receives at least the payoff received under LB. This problem boils down to choosing an interval $D = \left[\pi^f(\underline{z}), \pi^f(z^*)\right]$. Notice that if $\pi^f(z^*) = \pi_o$, then $V^{LB}_{\theta_H}(\pi_o, \theta_H) > V^{DE}_{\theta_H}(D, \theta_H)$, since both have identical payoffs for $z < \left(\pi^f\right)^{-1}(\pi_o)$ and LB has strictly higher payoffs for z larger than this threshold owing to the greater flexibility. Therefore, in order for delegation to be strictly preferred, it must be that $\pi^f(z^*) > \pi_o$.

Next, consider the LB problem. Given that we assumed quadratic utility, the solution to the LB problem is

$$\pi^{LB}(z, \theta_H) = \begin{cases} z + \theta_H & \text{if } z \leqslant \pi^o - \frac{\Delta}{2} \\ \pi^o & \text{if } z \in \left(\pi^o - \frac{\Delta}{2}, \pi^o + \frac{\Delta}{2}\right). \\ \min\{z + \theta_H, 2(z - \theta_H) - \pi^o\} & \text{if } z \geqslant \pi^o + \frac{\Delta}{2} \end{cases}$$

Now suppose that $\bar{z} \leqslant 3\theta_H + \pi_o$ so the participation constraint for the low type is always binding. Then $V^{LB}_L(\pi_o, \theta_H) = V^{DE}_L(D, \theta_H)$, where D has a cap equal to π_o. Therefore, in this range we have that $V^{DE}_L\left(\tilde{D}, \theta_H\right) < V^{LB}_L(\pi_o, \theta_H)$ for any $\tilde{D}$ with a larger cap. As a result, the best the H type can do in delegation is to choose $\pi^f(z^*) = \pi_o$, and delegation is not optimal. Now consider the case in which $\bar{z} > 3\theta + \pi_o$. For $3\theta + \pi_o < \bar{z}$, we have that $V^{LB}_L(\pi_o, \theta_H) > V^{DE}_L(D, \theta_H)$, where D has a cap equal to π_o. This implies that the H type can delegate only with a lower cap than π_o. Therefore, the H type strictly prefers LB. Q.E.D.

This result suggests that Congress will never choose to delegate policy-making so long as the faction in power needs approval from the opposing faction. This interpretation seems in contrast to the observed expansion of delegation to executive agencies. We view this overdelegation as arising from the executive branch, where it may be possible for the factions in power to use executive agencies to implement these factions' desired policies without much approval from the opposition. In the context of the model, this scenario

would correspond to the case in which the faction in power was no longer subject to the participation constraint; thus, it would prefer to delegate.

9 Conclusion

This paper studies the optimal delegation of decision-making authority in an economy with three key features: policy *bias* of the decision-makers relative to societal preferences, shocks which create a benefit for *flexibility*, and political *polarization* between different factions in society who disagree on the optimal policy. We show that allocating policy-making to the legislature is preferred when political risk is large relative to the bias, while allocating policy-making to an executive is preferred in the opposite case. We argue that, in contrast to conventional wisdom, the legislature is the institution through which flexibility is granted, and not the executive, because when it is optimal to delegate to the latter, it is optimal to do so with a relatively narrow mandate. We also study the credibility of these institutions and show that while legislative bargaining is always credible, executive delegation is credible only when the bias arises from time-inconsistency problems.

Our work provides a normative theory of these institutions. One interesting avenue for future work is to develop a positive theory for the prevalence of these institutions and, in particular, account for the rise of executive agencies in policy-making. Our model abstracts from many features that would make executive delegation attractive for Congress. For example, as argued by Epstein and O'halloran (1999), politicians might choose to delegate policies that might hurt their re-election chances, such as the closure of military bases in the 1980s.

Finally, given the narrow mandates that are optimal in the context of executive delegation, it would be interesting to study the design of these mandates in a dynamic economy in the presence of bias and political risk while accounting for their credibility. See, for example, Dovis and Kirpalani (2021).

References

Abreu, D. (1988): "On the theory of infinitely repeated games with discounting," *Econometrica: Journal of the Econometric Society*, 383–396. 24

Aghion, P., A. Alesina, and F. Trebbi (2004): "Endogenous political institutions," *The Quarterly Journal of Economics*, 119, 565–611. 5

Alesina, A. and G. Tabellini (2007): "Bureaucrats or Politicians? Part I: A Single Policy Task," *American Economic Review*, 97, 169–179. 5

——— (2008): "Bureaucrats or politicians? Part II: Multiple policy tasks," *Journal of Public Economics*, 92, 426–447. 5

Ali, S. N., B. D. Bernheim, A. W. Bloedel, and S. C. Battilana (2023): "Who Controls the Agenda Controls the Legislature," *American Economic Review*, 113, 3090–3128. 4

Ali, S. N., B. D. Bernheim, and X. Fan (2019): "Predictability and power in legislative bargaining," *The Review of Economic Studies*, 86, 500–525. 4

Amador, M. and K. Bagwell (2013a): "The theory of optimal delegation with an application to tariff caps," *Econometrica*, 81, 1541–1599. 4

——— (2013b): "The Theory of Optimal Delegation with an Application to Tariff Caps," *Econometrica*, 81, 1541–1599. 6, 34

Amador, M., I. Werning, and G.-M. Angeletos (2003): "Commitment vs. Flexibility," Tech. Rep. 10151, NBER. 4

——— (2006): "Commitment vs. flexibility," *Econometrica*, 74, 365–396. 4

Aranson, P. H., E. Gellhorn, and G. O. Robinson (1982): "Theory of legislative delegation," *Cornell L. Rev.*, 68, 1. 5

Athey, S., A. Atkeson, and P. J. Kehoe (2005): "The optimal degree of discretion in monetary policy," *Econometrica*, 73, 1431–1475. 4, 7, 21

Atkeson, A., V. V. Chari, and P. J. Kehoe (2007): "On the optimal choice of a monetary policy instrument," . 21

Atkeson, A. and P. J. Kehoe (2001): "The advantage of transparent instruments of monetary policy," Tech. rep., National Bureau of Economic Research. 21

Baron, D. P. (1996): "A dynamic theory of collective goods programs," *American Political Science Review*, 90, 316–330. 4

Baron, D. P. and J. A. Ferejohn (1989): "Bargaining in legislatures," *American political science review*, 83, 1181–1206. 4

Barro, R. J. and D. B. Gordon (1983): "Rules, discretion and reputation in a model of monetary policy," *Journal of monetary economics*, 12, 101–121. 3

Boerma, J., A. Tsyvinski, and A. P. Zimin (2022): "Bunching and taxing multidimensional skills," Tech. rep., National Bureau of Economic Research. 4

Bowen, T. R., Y. Chen, and H. Eraslan (2014): "Mandatory versus discretionary spending: The status quo effect," *American Economic Review*, 104, 2941–2974. 4

Bowen, T. R., Y. Chen, H. Eraslan, and J. Zápal (2017): "Efficiency of flexible budgetary institutions," *Journal of Economic Theory*, 167, 148–176. 2, 4

Callander, S. and K. Krehbiel (2014): "Gridlock and Delegation in a Changing World," *American Journal of Political Science*, 58, 819–834. 5

Chari, V., A. Dovis, and P. J. Kehoe (2020): "Rethinking Optimal Currency Areas," *Journal of Monetary Economics*, 111, 80–94. 25

Diermeier, D. and P. Fong (2011): "Legislative bargaining with reconsideration," *The Quarterly Journal of Economics*, 126, 947–985. 4

Dovis, A. and R. Kirpalani (2021): "Rules without commitment: Reputation and incentives," *The Review of Economic Studies*, 88, 2833–2856. 26

Dziuda, W. and A. Loeper (2016): "Dynamic collective choice with endogenous status quo," *Journal of Political Economy*, 124, 1148–1186. 4

Epstein, D. and S. O'halloran (1999): *A transaction cost politics approach to policy making under separate powers*, Cambridge: Cambridge university press. 5, 26

Halac, M. and P. Yared (2014): "Fiscal rules and discretion under persistent shocks," *Econometrica*, 82, 1557–1614. 4

——— (2020): "Commitment versus flexibility with costly verification," *Journal of Political Economy*, 128, 4523–4573. 4

——— (2022a): "Fiscal rules and discretion under limited enforcement," *Econometrica*, 90, 2093–2127. 4

——— (2022b): "Instrument-based versus target-based rules," *The Review of Economic Studies*, 89, 312–345. 21

Kalandrakis, A. (2004): "A three-player dynamic majoritarian bargaining game," *Journal of Economic Theory*, 116, 294–14. 4

Klyza, C. M. and D. J. Sousa (2013): *American Environmental Policy: Beyond Gridlock, Updated and Expanded Edition*, Cambridge, MA: MIT Press. 19

Laffont, J.-J., E. Maskin, and J.-C. Rochet (1987): "Optimal nonlinear pricing with two-dimensional characteristics," *Information, Incentives and Economic Mechanisms*, 256–266. 4, 32

Luenberger, D. G. (1969): *Optimization by Vector Space Methods*, Series in Decision and Control, Wiley. 34

McCubbins, M. D., R. G. Noll, and B. R. Weingast (1987): "Administrative procedures as instruments of political control," *The Journal of Law, Economics, and Organization*, 3, 243–277. 5

Moser, C. and P. Olea de Souza e Silva (2019): "Optimal paternalistic savings policies," *Columbia Business School Research Paper*. 4

Piguillem, F. and A. Riboni (2015): "Spending-biased legislators: Discipline through disagreement," *The Quarterly Journal of Economics*, 130, 901–949. 4

——— (2021): "Fiscal rules as bargaining chips," *The Review of Economic Studies*, 88, 2439–2478. 4

Rao, N. (2015): "Administrative Collusion: How Delegation Diminishes the Collective Congress," *NYUL rev.*, 90, 1463. 5

Rochet, J.-C. and P. Choné (1998): "Ironing, sweeping, and multidimensional screening," *Econometrica*, 783–826. 4

Rogoff, K. (1985): "The optimal degree of commitment to an intermediate monetary target," *The quarterly journal of economics*, 100, 1169–1189. 4

Sublet, G. (2023): "The Optimal Degree of Discretion in Fiscal Policy," . 4

Tucker, P. (2019): *Unelected power: The quest for legitimacy in central banking and the regulatory state*, Princeton University Press. 1

U.S. Supreme Court (1984): "Chevron U.S.A., Inc. v. Natural Resources Defense Council, Inc." Supreme Court of the United States, 468 U.S. 837. 3, 20

——— (2024): "Loper Bright Enterprises v. Raimondo," Supreme Court of the United States, no. 22-451. 20

Volden, C. (2002): "A formal model of the politics of delegation in a separation of powers system," *American Journal of Political Science*, 111–133. 5

Appendix

A Proofs of results in Section 3

A.1 Proof of Proposition 3

The Ramsey step

As described in the main text, we prove this Proposition 3 in two steps. First, we consider a Ramsey problem in which the society is restricted to delegating a continuous interval and chooses the bounds of this interval. Second, we show that no other incentive feasible allocation dominates this.

We now consider the Ramsey problem in which society chooses an interval $D = [\underline{\pi}, \bar{\pi}]$. Given this interval, and our assumption on policy-maker preferences, the choice of a policy-maker of type θ_L is

$$\pi_D(z, \theta_L) = \min\{\max\{\pi_f(z, \theta_L), \underline{\pi}\}, \bar{\pi}\}$$

and

$$\pi_D(z, \theta_H) = \pi_D(z + \Delta, \theta_L).$$

The latter equality implies that we can write the policy choice of θ_H in terms of θ_L. To do so we need to extend the range of states to $[\underline{z}, \bar{z} + \Delta]$. Then, we can parametrize the bounds of the interval by threshold state realizations $z_l, z_h \in [\underline{z}, \bar{z} + \Delta]$ so that $D = [\underline{\pi}, \bar{\pi}] = [\pi_f(z_l, \theta_L), \pi_f(z_h, \theta_L)]$. We say that a cap is binding if $\pi_f(z, \theta_L) > \bar{\pi}$ for some $z \in [z_h, \bar{z} + \Delta]$. Similarly, a floor is binding if $\pi_f(z, \theta_L) < \underline{\pi}$ for some $z \in [\underline{z}, z_l]$.

To simplify the notation, let $K(z) = \alpha^L F(z) + \alpha^H F(z - \Delta)$, $\kappa(z) = \alpha^L f(z) + \alpha^H f(z - \Delta)$, and E_κ denote the expectation with respect to the distribution K. Let $d(z) = \alpha_H \alpha^L f(z) - \alpha_L \alpha^H f(z - \Delta)$ denote a measure of society's *net disagreement* absent any bias, i.e., for $\bar{\nu} = 0$. To understand this, fix some policy π and let z be the state such that the low type optimally chooses π. This policy choice is costly for the fraction α_H of society that has preferences aligned with θ_H if the faction in power is θ_L, which happens with probability α_L, and the realization of the state is z. This is the first term in $d(z)$. A specular logic holds for the fraction of society aligned with θ_L, which corresponds to the second term of $d(z)$.

The next lemma shows that the optimal cap is set such that there is no distortion–on average–over the bunching at the top.

Lemma 4 (Ramsey: optimal cap)**.** *If the optimal cap is binding and* $z_h \in (z_l, \bar{z} + \Delta)$*, then*

$$\Delta d_h(z_h) - \bar{\nu} = E_\kappa[z | z \geqslant z_h] - z_h, \tag{17}$$

where $d_h(z_h) = -\frac{\int_{z_h}^{\bar{z}+\Delta} d(z)dz}{1-K(z_h)}$. *If the cap does not bind, that is* $z_h = \bar{z} + \Delta$, *then* $\Delta\alpha_L - \bar{\nu} < 0$.

Proof. The Ramsey problem of optimally setting an interior cap $z_h \in (z_l, \bar{z} + \Delta)$ is

$$\max_{\hat{z}_h} \int_{z_l}^{\hat{z}_h} \left[\alpha^L v(\pi_f(z, \theta_L), z) + \alpha^H v(\pi_f(z, \theta_L), z - \Delta)\frac{f(z-\Delta)}{f(z)}\right] f(z)dz$$
$$+ \int_{\hat{z}_h}^{\bar{z}+\Delta} \left[\alpha^L v(\pi_f(\hat{z}_h, \theta_L), z) + \alpha^H v(\pi_f(\hat{z}_h, \theta_L), z - \Delta)\frac{f(z-\Delta)}{f(z)}\right] f(z)dz.$$

If z_h is interior (i.e., $z_h \in (z_l, \bar{z} + \Delta)$), then the first-order condition with respect to z_h gives

$$0 = \int_{z_h}^{\bar{z}+\Delta} \left[\alpha^L v_1(\pi_f(z_h, \theta_L), z) + \alpha^H v_1(\pi_f(z_h, \theta_L), z - \Delta)\frac{f(z-\Delta)}{f(z)}\right] dz. \tag{18}$$

Using

$$v_1(\pi_f(z_h, \theta_L), z) = z - z_h + \alpha_H\Delta + \bar{\nu},$$

and

$$v_1(\pi_f(z_h, \theta_L), z - \Delta) = z - z_h - \alpha_L\Delta + \bar{\nu},$$

the first-order condition (18) reads

$$0 = \int_{z_h}^{\bar{z}+\Delta} [z - z_h + \bar{\nu}](\alpha^L f(z) + \alpha^H f(z-\Delta)) + \Delta\left[\alpha^L\alpha_H f(z) - \alpha^H\alpha_L f(z-\Delta)\right] dz.$$

Dividing both sides by $\alpha^L(1 - F(z_h)) + \alpha^H(1 - F(z_h - \Delta))$ gives

$$E_\kappa[z|z \geqslant z_h] - z_h + \bar{\nu} = \Delta\left[\frac{\alpha^H\alpha_L(1 - F(z_h - \Delta)) - \alpha^L\alpha_H(1 - F(z_h))}{\alpha^L(1 - F(z_h)) + \alpha^H(1 - F(z_h - \Delta))}\right].$$

□

Without heterogeneity in θ, i.e., $\Delta = 0$, an optimal cap, if it is interior, equates the bias on the left-hand side to the loss of discretion on the right-hand side, $-\bar{\nu} = E_\kappa[z|z \geqslant z_h] - z_h$. Thus, a cap is binding only if the bias is negative i.e. policymakers prefer a higher policy than society.

With heterogeneity in θ, i.e., $\Delta > 0$, the distribution K captures the higher loss of discretion for θ_H than that for θ_L. Heterogeneity adds a second component to the benefit of a cap and d_h captures the conditional expectation of the net disagreement over the cap. It aggregates the net disagreement for all policies greater than the cap. For example, suppose that $\alpha_L = \alpha^L$ and the distribution is uniform. Then absent the bias societal preferences are a convex combination of those of the low and high type. If $\Delta > 0$ the cost of to society of θ_H choosing its flexible policy depends on the weights of this convex combination.

Note that the cap binds if $\Delta\alpha_L - \bar{\nu} > 0$ because, for shocks distributed on a compact support, the mean residual life is zero at the upper bound $\bar{z}$. The analysis for an optimal floor mirrors the analysis for the cap.

Lemma 5 (Ramsey: optimal floor)**.** *If the optimal floor is binding then* $\underline{\pi} = \pi_f(z_l, \theta_L)$ *and* $z_l \in (\underline{z}, z_h)$ *solves*

$$\Delta d_l(z_l) + \bar{\nu} = z_l - E_k[z|z \leqslant z_l] \tag{19}$$

where $d_l(z_l) = \frac{\int_{\underline{z}}^{z_l} d(z)dz}{K(z_l)}$*. If the floor does not bind, that is* $z_l = \underline{z}$*, then* $\Delta\alpha_H + \bar{\nu} \leqslant 0$*.*

The proof mirrors that of Lemma (4) and it is omitted.

In the standard delegation setting with one-dimensional private information, the optimal delegation set is an interval with either a binding floor or cap. In this case both the cap and floor can be binding depending on which faction is in charge.

The verification step

With one dimension of private information, say the shocks z, the global incentive compatibility constraints are equivalent to a local incentive compatibility constraint—the envelope condition—and a monotonicity condition. In our setting, incentives, must also be compatible across the other dimension of private information. The next lemma shows that under our assumptions about preferences, global incentive compatibility is equivalent to the two conditions previously described as well as a simple condition that connects the allocations of the two types. Following Laffont et al. (1987), we call the latter the *integrability condition*.

Lemma 6. *An allocation* $\pi(z, \theta)$ *satisfies (6) if only if,*

$$u(\pi(z, \theta_L), z, \theta_L) = u(\pi(\underline{z}, \theta_L), \underline{z}, \theta_L) + \int_{\underline{z}}^{z} \pi(x, \theta_L)dx, \text{ for } z \in [\underline{z}, \bar{z} + \Delta], \tag{20}$$

$$\pi(z, \theta_L) \text{ is non-decreasing in } z, \tag{21}$$

$$\pi(z, \theta_H) = \pi(z + \Delta, \theta_L). \tag{22}$$

Proof. Let $U(z, \theta) = u(\pi(z, \theta), z, \theta)$. For $z > z'$,

$$\frac{U(z, \theta) - U(z', \theta)}{z - z'} \geqslant \pi(z', \theta),$$

and

$$\pi(z, \theta) \geqslant \frac{U(z, \theta) - U(z', \theta)}{z - z'}.$$

Combining the two inequalities,

$$\pi(z,\theta) \geqslant \frac{U(z,\theta) - U(z',\theta)}{z - z'} \geqslant \pi(z',\theta).$$

Hence, $\pi(z,\theta)$ is increasing in z. A monotonic function is almost everywhere differentiable so it is almost everywhere continuous. Taking the limit $z \to z'$ gives 20.

Because $u(\pi, z, \theta) = u(\pi, z + \theta - \hat{\theta}, \hat{\theta})$, it follows that $\pi(z, \theta_H) = \pi(z + \theta_H - \theta_L, \theta_L)$, which gives 22. Incentive compatibility for θ_H follows from the integrability condition and incentive compatibility for θ_L with respect to z; formally,

$$\begin{aligned}
u(\pi(z,\theta_H), z, \theta_H) &= u(\pi(z+\Delta,\theta_L), z, \theta_H) \\
&= u(\pi_L(z+\Delta,\theta_L), z+\Delta, \theta_L) \\
&= u(\pi(\underline{z},\theta_L), \underline{z}, \theta_L) + \int_{\underline{z}}^{z+\Delta} \pi(x,\theta_L)\,dx \\
&= u(\pi(\underline{z},\theta_L), \underline{z}, \theta_L) + \int_{\underline{z}}^{\underline{z}+\Delta} \pi(x,\theta_L)\,dx + \int_{\underline{z}+\Delta}^{z+\Delta} \pi(x,\theta_L)\,dx \\
&= u(\pi(\underline{z}+\Delta,\theta_L), \underline{z}+\Delta, \theta_L) + \int_{\underline{z}}^{\underline{z}+\Delta} \pi(x,\theta_L)\,dx + \int_{\underline{z}+\Delta}^{z+\Delta} \pi(x,\theta_L)\,dx \\
&= u(\pi(\underline{z},\theta_H), \underline{z}, \theta_H) + \int_{\underline{z}}^{z} \pi(x,\theta_H)\,dx.
\end{aligned}$$

□

Conditions (20) and (21) are the standard envelope and monotonicity conditions for incentive compatibility for one dimension of private information. The integrability condition (22) ensures incentive compatibility across the other dimension of private information. The intuition for the integrability condition follows from the linearity of preferences in both dimensions of private information, which implies $u(\pi, z+\Delta, \theta_L) = u(\pi, z, \theta_H)$.[8] Note that incentive compatibility for type θ_H obtains from the integrability condition and incentive compatibility for type θ_L over an extended range of shocks $[\underline{z}, \bar{z} + \Delta]$ to capture incentives over policies above θ_L's preferred policy.

The proof for the verification step has three parts. First, we map the two-dimensional delegation problem (4) into a standard one-dimensional delegation problem. Second, we use global Lagrangian methods to account for the incentive compatibility constraints and derive a set of sufficient optimality conditions. Third, we guess a Lagrange multiplier function and verify that the solution satisfies the optimality conditions.

Part I. The upshot of Lemma 6 is that the two-dimensional delegation problem (4)

[8] If the types enter multiplicatively in the utility index, i.e., $u(\pi, z, \theta) = z\pi + \theta b(\pi)$, the integrability condition is $\pi_H(z/\theta_H) = \pi_L(z/\theta_L)$.

maps to a one-dimensional delegation problem. Note that the domain of definition of π is extended by Δ so that the integrability condition characterizes $\pi(z, \theta_H)$ for $z \in [\underline{z}, \bar{z}]$. We can then express problem (5) as:

$$\max_{\pi(\cdot)} \int_{\underline{z}}^{\bar{z}} w(\pi(z), \pi(z+\Delta), z) f(z) dz \tag{23}$$

subject to the envelope condition (20) and the monotonicity condition (21).

Part II. We use global Lagrangian methods to account for the constraint that there are no transfers in (23). The global theory of constrained optimization shows that the maximizer of a Lagrangian is the solution to the constrained optimization problem of interest. We start by defining the Lagrangian. Denote the Lagrange multiplier function on the continuum of equality constraints (20) by $\Lambda : [\underline{z}, \bar{z}+\Delta] \mapsto \mathbb{R}$ with $1-\Lambda$ integrable. The Lagrangian is a functional on the set of allocations that satisfy the monotonicity condition,

$$\Phi = \{\pi \mid \pi : [\underline{z}, \bar{z}+\Delta] \mapsto \mathbb{R} \text{ is non-decreasing}\}$$

defined as follows: $\mathcal{L} : \Phi \to \mathbb{R}$,

$$\begin{aligned}\mathcal{L}(\pi) = &\int_{\underline{z}}^{\bar{z}} [w(\pi(z), \pi(z+\Delta), z)] f(z) dz \\ &+ \int_{\underline{z}}^{\bar{z}+\Delta} \left[u(\pi(z), z, \theta_L) - u(\pi(\underline{z}), \underline{z}, \theta_L) - \int_{\underline{z}}^{z} \pi(x) dx \right] d\Lambda(z).\end{aligned} \tag{24}$$

Integrating the nested integrals in (24) by parts, the Lagrangian reads,

$$\begin{aligned}\mathcal{L}(\pi) = &\int_{\underline{z}}^{\bar{z}} [w(\pi(z), \pi(z+\Delta), z)] f(z) dz \\ &+ \int_{\underline{z}}^{\bar{z}+\Delta} u(\pi(z), z, \theta_L) d\Lambda(z) - u(\pi(\underline{z}), \underline{z}, \theta_L)(1-\Lambda(\underline{z})) - \int_{\underline{z}}^{\bar{z}+\Delta} \pi(z)(1-\Lambda(z)) dz.\end{aligned}$$

The global theory of constrained optimization outlined in Luenberger (1969) Chapter 8 shows that the maximizer of a concave Lagrangian is the solution to the constrained optimization problem of interest. Theorem 1 in Section 8.3 in Luenberger (1969) (see also Appendix B in Amador and Bagwell (2013b)) implies that a maximizer of the Lagrangian (24) is a solution to (23). Lemma 1 in Section 8.7 in Luenberger (1969) implies that if there exists π in the convex cone Φ and an integrable function Λ such that $\mathcal{L}$ is concave,

$$\partial \mathcal{L}(\pi, \pi) = 0, \tag{25}$$

and

$$\partial\mathcal{L}(\pi, h) \leqslant 0 \text{ for } h \in \Phi, \tag{26}$$

then π maximizes the Lagrangian, and hence is a solution to (23), where the derivatives are Gateaux derivatives $\partial\mathcal{L}(\pi, h) = \lim_{\alpha\downarrow 0} \frac{1}{\alpha}[\mathcal{L}(\pi + \alpha h) - \mathcal{L}(\pi)]$ in the direction $h : [\underline{z}, \bar{z} + \Delta] \to \mathbb{R}$,[9]

$$\begin{aligned}\partial\mathcal{L}(\pi, h) = &\int_{\underline{z}}^{\bar{z}} [w_1(\pi(z), \pi(z+\Delta), z)h(z) + w_2(\pi(z), \pi(z+\Delta), z)h(z+\Delta)]\, f(z)dz \\ &+ \int_{\underline{z}}^{\bar{z}+\Delta} u_1(\pi(z), z, \theta_L)h(z)d\Lambda(z) - u_1(\pi(\underline{z}), \underline{z}, \theta_L)h(\underline{z})(1-\Lambda(\underline{z})) \\ &- \int_{\underline{z}}^{\bar{z}+\Delta} h(z)(1-\Lambda(z))dz.\end{aligned}$$

Part III. We guess that the solution to the delegation problem (23) coincides with the solution to the Ramsey problem parametrized by z_l and z_h in Lemmas 4 and 5, that is

$$\pi(z) = \begin{cases} \pi_f(z_l, \theta_L) & \text{for } z \in [\underline{z}, z_l] \\ \pi_f(z, \theta_L) & \text{for } z \in (z_l, z_h) \\ \pi_f(z_h, \theta_L) & \text{for } z \in [z_h, \bar{z} + \Delta]. \end{cases}$$

We also guess the following Lagrange multiplier function

$$\Lambda(z) = \begin{cases} 1 - K(z) & \text{for } z \in [\underline{z}, z_l] \\ 1 - [(\alpha_H\Delta + \bar{\nu})\alpha^L f(z) + (-\alpha_L\Delta + \bar{\nu})\alpha^H f(z-\Delta)] & \text{for } z \in (z_l, z_h) \\ 1 + (1 - K(z)) & \text{for } z \in [z_h, \bar{z} + \Delta], \end{cases} \tag{27}$$

and verify that the three optimality conditions (which are that $\mathcal{L}(\pi)$ is concave, (25), and (26)) are satisfied.

Concavity of the Lagrangian. We can rewrite the Lagrangian as follows:

$$\begin{aligned}\mathcal{L}(\pi) = &\int_{\underline{z}}^{\bar{z}} [w(\pi(z), \pi(z+\Delta), z)f(z) - \kappa(z)u(\pi(z), z, \theta_L)]\, dz \\ &+ \int_{\underline{z}}^{\bar{z}+\Delta} u(\pi(z), z, \theta_L)(d\Lambda(z)/dz + \kappa(z))dz \\ &- u(\pi(\underline{z}), \underline{z}, \theta_L)(1-\Lambda(\underline{z})) - \int_{\underline{z}}^{\bar{z}+\Delta} \pi(z)(1-\Lambda(z))dz.\end{aligned}$$

The first integral is concave in π by definition of w and κ. The second integral is concave

[9] Amador Werning and Angeletos (2006) show that the Gateaux differential exists in Lemma A.1 p. 390.

if $\Lambda(z) + K(z)$ is non-decreasing because u is concave. The last two terms are linear in the allocation, and hence concave. Hence, it suffices that $\Lambda(z) + K(z)$ is non-decreasing for $\mathcal{L}(\pi)$ to be concave. By construction, $\Lambda(z) + K(z)$is constant for $z \leqslant z_l$and for $z \geqslant z_h$. Assumption 2 part 3 is equivalent to $\Lambda(z) + K(z)$is non-decreasing for $z \in (z_l, z_h)$. Lastly, Assumption 2 parts 1 and 2 imply that $\Lambda(z) + K(z)$ is non-decreasing at z_l and z_h.

Condition (25). Substituting the marginal utility at the allocation implemented by a cap and a floor (recall $u(\pi, z, \theta_i) = (z + \theta)\pi + b(\pi)$)

$$u_1(\pi(z), z, \theta_L) = \begin{cases} z - z_l & \text{for } z \in [\underline{z}, z_l] \\ 0 & \text{for } z \in (z_l, z_h) \\ z - z_h & \text{for } z \in [z_h, \bar{z} + \Delta], \end{cases}$$

in the Gateaux derivative gives

$$\begin{aligned}
\partial\mathcal{L}(\pi, h) = &\int_{\underline{z}}^{z_l} \left[\alpha^L(z - z_l + \alpha_H\Delta + \bar{\nu})f(z) + \alpha^H(z - z_l - \alpha_L\Delta + \bar{\nu})f(z - \Delta)\right] h(z)dz && (28)\\
&+ \int_{\underline{z}}^{z_l} (z - z_l)h(z)d\Lambda(z) - \int_{\underline{z}}^{z_l} h(z)(1 - \Lambda(z))dz && (29)\\
&+ \int_{z_l}^{z_h} \left[\alpha^L(\alpha_H\Delta + \bar{\nu})f(z) + \alpha^H(-\alpha_L\Delta + \bar{\nu})f(z - \Delta) - (1 - \Lambda(z))\right] h(z)dz && (30)\\
&+ \int_{z_h}^{\bar{z}+\Delta} \left[\alpha^L(z - z_h + \alpha_H\Delta + \bar{\nu})f(z) + \alpha^H(z - z_h - \alpha_L\Delta + \bar{\nu})f(z - \Delta)\right] h(z)dz && (31)\\
&+ \int_{z_h}^{\bar{z}+\Delta} (z - z_h)h(z)d\Lambda(z) - \int_{z_h}^{\bar{z}+\Delta} h(z)(1 - \Lambda(z))dz && (32)\\
&- (\underline{z} - z_l)h(\underline{z})(1 - \Lambda(\underline{z})). && (33)
\end{aligned}$$

The following argument shows that for z_l and z_h that solve the Ramsey problem as in Lemmas 4 and 5, and given the Lagrange multiplier function (27), condition (25) is satisfied. Note that the direction of the solution $h(z) = \pi(z)$ is constant for $z \in [\underline{z}, z_l]$ and for $z \in [z_h, \bar{z} + \Delta]$. As a result $h(z)$ can be taken out of the integral in lines (28), (29), (31), and (32). Substituting the Lagrange multiplier function (27) and rearranging terms gives the optimality conditions for z_l and z_h in Lemmas 4 and 5. Line (30) is zero by definition of the Lagrange multiplier function (27). Lastly, line (33) is zero because $\Lambda(\underline{z}) = 1$.

Conditions (26). The following argument shows that, given the Lagrange multiplier (27), the conditions in Assumption 2 parts 1 and 2 imply that the optimality condition (26) is satisfied. Again, line (30) is zero by definition of the Lagrange multiplier function (27), and line (33) is zero because $\Lambda(\underline{z}) = 1$. Integrating (29) by parts and substituting the optimality condition for z_l from Lemma 5 gives the condition in part 2 of Assumption 2.

Similarly, Integrating (32) by parts and substituting the optimality condition for z_h from Lemma 4 gives the condition in part 1 of Assumption 2. Q.E.D.

A.2 Proof of Lemma 1

For a given Δ, the upper bound of the delegated set $\bar{\pi}$ depends on the bias $\bar{\nu}$, as well as on the mean-residual-life $E_\kappa[\hat{z}|\hat{z} \geqslant z] - z$ and the function $d_h(z)$. Lemma (4) shows that if there is $z_h \in (z_l, \bar{z} + \Delta)$ such that $\pi_f(z_h, \theta_L) = \bar{\pi}$, then it solves equation (17). If, instead the bias is positive and large in the sense that $\Delta\alpha_L < \bar{\nu}$, then (17) does not have a solution in $[z_l, \bar{z} + \Delta]$ and the upper bound is not constraining because $z_h = \bar{z} + \Delta$. If the bias is negative or positive and small in the sense that $\Delta d_h(z) - (E_\kappa[z|z \geqslant z] - z) > \bar{\nu}$ for $z \in (z_l, \bar{z} + \Delta)$, then $\bar{\pi} = \underline{\pi} = \bar{\pi}^*$.

For an interior $z_h \in (\underline{z}, \bar{z} + \Delta)$, let $z_h(\bar{\nu})$ denote the implicit function of $\bar{\nu}$ that solves (17). The threshold $z_h(\bar{\nu})$ depends on the slope of the mean-residual life and the function d_h. For a log-concave density κ, the mean-residual life is monotone decreasing in z. Also, as we now show, for a log-concave density f, the function d_h is monotone increasing. To see this, note that for $z \geqslant \bar{z}$, the function

$$d_h(z) = \left[\frac{\alpha_L \alpha^H (1 - F(z - \Delta) - \alpha_H \alpha^L (1 - F(z))}{\alpha^L (1 - F(z)) + \alpha^H (1 - F(z - \Delta))}\right]$$

is constant (it is equal to α_L). For $z < \bar{z}$, because F is differentiable,

$$d_h'(z) = \frac{\alpha^H \alpha^L}{\alpha^H (1 - F(z - \Delta) + \alpha^L (1 - F(z))} \left[f(z)(1 - F(z - \Delta)) - f(z - \Delta)(1 - F(z))\right].$$

Because the hazard rate of a log-concave density f is monotone increasing, $\frac{f(z-\Delta)}{1-F(z-\Delta)} \leqslant \frac{f(z)}{1-F(z)}$ and $d_h'(z) \geqslant 0$ for $z < \bar{z}$. It follows that $\Delta d_h(z) - (E_\kappa[z|z \geqslant z] - z)$ is an increasing function of z and, in turn, $z_h(\bar{\nu})$ is a monotone non-decreasing function of $\bar{\nu}$. A specular logic applies to the lower bound of the delegated set $\pi_f(z_l, \theta_L)$.

For a given $\bar{\nu}$, there exists a $\Delta(\bar{\nu})$ large enough (e.g, larger than $\bar{z} - \underline{z}$ if $\bar{\nu} = 0$) such that conditions (17) and (19) both admit a solution in $[\bar{z}, \bar{z} + \Delta]$ and the optimal delegated set is a single point. Q.E.D.

B Proofs of results in Section 4

B.1 Proof of Lemma 2

Suppose, by way of contradiction, that the delegated set contains an interval $[\pi_f(z_l, \theta_L), \pi_f(z_h, \theta_L)]$. Without loss, suppose $z_h - z_l < \Delta$.

Consider an alternative set $\hat{D}$ obtained by removing the interval $(\pi_f(z_l, \theta_L), \pi_f(z_h, \theta_L))$ from D, i.e., "drilling a hole" in D). Formally,

$$\hat{D} = D \setminus (\pi_f(z_l, \theta_L), \pi_f(z_h, \theta_L)).$$

We prove this lemma in three steps. First, we characterize the relevant outside option associated with $\hat{D}$. Second, we characterize the chosen policies under LB given the outside options from step 1.Finally, in step 3, we show that societal welfare is larger under $\hat{D}$ relative to D.

Step 1: Characterizing the relevant outside option associated with $\hat{D}$

The set of available policies through bargaining $\mathcal{R}(z, \theta, \hat{D})$ depends on the outside option the executive θ_L has in $\hat{D}$. For $z \in [z_l, z_h]$ and θ_L, the relevant set of policies to choose from is the discrete set $\{\pi_f(z_l, \theta_L), \pi_f(z_h, \theta_L)\}$. Hence, there exists $z^* \in (z_l, z_h)$,

$$\pi_{\hat{D}}(z, \theta_L) = \begin{cases} \pi_D(z, \theta_L) & \text{for } z \leqslant z_l, \\ \pi_f(z_l, \theta_L) & \text{for } z \in [z_l, z^*], \\ \pi_f(z_h, \theta_L) & \text{for } z \in (z^*, z_h], \\ \pi_D(z, \theta_L) & \text{for } z \geqslant z_h. \end{cases}$$

The threshold z^* is such that θ_L is indifferent between $\pi_f(z_h, \theta_L)$ and $\pi_f(z_l, \theta_L)$, that is

$$z^*(z_h, z_l) = -\frac{b(\pi_f(z_l, \theta_L)) - b(\pi_f(z_h, \theta_L))}{\pi_f(z_l, \theta_L) - \pi_f(z_h, \theta_L)} - \theta_L. \tag{34}$$

Step 2: Characterizing the chosen policies under LB

We first characterize the chosen policies for type θ_L. We will show that for $z \in (z_l, z_h)$ the solution to the bargaining problem (8) is

$$\begin{cases} \pi_f(z, \theta_L) & \text{for } z \in [z_l, z^*(\theta_L)], \\ \pi_f(z_h, \theta_L) & \text{for } z \in [z^*(\theta_L), z_h]. \end{cases}$$

To see this, first notice that for $z \in (z_l, z^*(\theta_L))$,

$$\pi_{\hat{D}}(z, \theta_L) = \pi_f(z_l, \theta_L) < \pi_f(z, \theta_L),$$

and hence the solution in (8) for $z \in (z_l, z^*(\theta_L))$ is $\pi_f(z, \theta_L)$.

Next, for $z \in (z^*(\theta_L), z_h)$,

$$\pi_{\hat{D}}(z, \theta_L) = \pi_f(z_h, \theta_L) = \pi_f(z_h - \Delta, \theta_H)$$

where the inequality follows from $z_h - z_l < \Delta$. Hence, for $z \in (z^*(\theta_L), z_h)$ and $\pi <$

$\pi_{\hat{D}}(z, \theta_L)$,

$$u(\pi, \theta_H) < u(\pi_{\hat{D}}(z, \theta_L), \theta_H).$$

Therefore, $\pi \notin \mathcal{R}(z, \theta_L, \hat{D})$ and the solution in (8) for $z \in (z^*(\theta_L), z_h)$ is $\pi_{\hat{D}}(z, \theta_L)$.

Next, we characterize the chosen policies for type θ_H. We will show that for $z \in (z_l - \Delta, z_h - \Delta)$ the solution to the bargaining problem (8) is

$$\begin{cases} \pi_f(z_l - \Delta, \theta_H) & \text{for } z \in [z_l - \Delta, z^*(\theta_H) - \Delta], \\ \pi_f(z, \theta_H) & \text{for } z \in [z^*(\theta_H) - \Delta, z_h - \Delta]. \end{cases}$$

To see this, first notice that for $z \in (z^*(\theta_H) - \Delta, z_h - \Delta)$,

$$\pi_{\hat{D}}(z, \theta_H) = \pi_f(z_h - \Delta, \theta_H) > \pi_f(z, \theta_H),$$

and hence the solution to the bargaining problem (8) for $z \in (z^*(\theta_H) - \Delta, z_h - \Delta)$ is $\pi_f(z, \theta_H)$.

Next, for $z \in (z_l - \Delta, z^*(\theta_H) - \Delta)$,

$$\pi_{\hat{D}}(z, \theta_H) = \pi_f(z_l - \Delta, \theta_H) = \pi_f(z_l, \theta_L) > \pi_f(z, \theta_H),$$

where the inequality follows from $z_h - z_l < \Delta$. Hence, for $z \in (z^*(\theta_H), z_h)$ and $\pi < \pi_{\hat{D}}(z, \theta_H)$,

$$u(\pi, \theta_H) < u(\pi_{\hat{D}}(z, \theta_H), \theta_H)$$

and $\pi \notin \mathcal{R}(z, \theta_H, \hat{D})$. The solution to (8) for $z \in (z^*(\theta_H), z_h)$ is $\pi_f(z_h, \theta_H)$.

Step 3: Showing that societal welfare is larger under $\hat{D}$ relative to D.

Notice that the welfare implications of delegating $\hat{D}$ instead of D depend only on the solution to the bargaining problem (8) for $z \in [z_l, z_h]$ for θ_L and for $z \in [z_l - \Delta, z_h - \Delta]$ for θ_H. Then, the previous steps imply that the welfare effect of "drilling a hole" in D is

$$\begin{aligned} \delta(z_h, z_l) = &\int_{z_l}^{z^*(z_h, z_l)} \left[b(\pi_f(z_l - \Delta, \theta_H)) - b(\pi_f(z - \Delta, \theta_H)) \right] \alpha^H f(z - \Delta) dz \\ &+ \int_{z_l}^{z^*(z_h, z_l)} \left[\left(z - \Delta + \bar{\nu} + \sum_i \alpha_i \theta_i \right) (\pi_f(z_l - \Delta, \theta_H) - \pi_f(z - \Delta, \theta_H)) \right] \alpha^H f(z - \Delta) dz \\ &+ \int_{z^*(z_h, z_l)}^{z_h} \left[b(\pi_f(z_h, \theta_L) - b(\pi_f(z, \theta_L)) \right] \alpha^L f(z) dz \\ &+ \int_{z^*(z_h, z_l)}^{z_h} \left[\left(z + \bar{\nu} + \sum_i \alpha_i \theta_i \right) [\pi_f(z_h, \theta_L) - \pi_f(z, \theta_L)] \right] \alpha^L f(z) dz. \end{aligned} \tag{35}$$

We will show that $\delta(z, z_l)$ has a local minimum at $\delta(z_l, z_l) = 0$. This implies for z_h

close enough to z_l, the welfare gain of $\hat{D}$ relative to D is positive.

As a first step let's take the derivative of (35) with respect to z_h and substitute the first-order conditions $b'(\pi_f(z_h, \theta_L)) = -(z_h + \theta_L)$, and $\pi_f(z - \Delta, \theta_H) = \pi_f(z, \theta_L)$ to obtain

$$\begin{aligned}
\frac{\partial \delta(z_h, z_l)}{\partial z_h} = {} & \alpha^H \frac{\partial z^*(z_h, z_l)}{\partial z_h} \left[b(\pi_f(z_l, \theta_L) - b(\pi_f(z^*(z_h, z_l), \theta_L))\right] f(z^*(z_h, z_l) - \Delta) \qquad (36)\\
& + \alpha^H \frac{\partial z^*(z_h, z_l)}{\partial z_h} \left[(z^*(z_h, z_l) + \theta_L + \bar{\nu} - \alpha_L \Delta)\left(\pi_f(z_l, \theta_L) - \pi_f(z^*(z_h, z_l), \theta_L)\right)\right] f(z^*(z_h, z_l) - \Delta)\\
& - \alpha^L \frac{\partial z^*(z_h, z_l)}{\partial z_h} \left[b(\pi_f(z_h, \theta_L)) - b(\pi_f(z^*(z_h, z_l), \theta_L))\right] f(z^*(z_h, z_l))\\
& - \alpha^L \frac{\partial z^*(z_h, z_l)}{\partial z_h} \left[(z^*(z_h, z_l) + \theta_L + \bar{\nu} + \alpha_H \Delta)\left(\pi_f(z_h, \theta_L) - \pi_f(z^*(z_h, z_l), \theta_L)\right)\right] f(z^*(z_h, z_l))\\
& + \alpha^L \frac{\partial \pi_f(z_h, \theta_L)}{\partial z} \int_{z^*(z_h, z_l)}^{z_h} \left[z - z_h + \bar{\nu} + \alpha_H \Delta\right] f(z) dz.
\end{aligned}$$

Since u is quadratic we also have that $\frac{\partial \pi_f(z_h, \theta_L)}{\partial z} = \frac{1}{2}$. The implicit function theorem for the threshold z^* defined in (34) implies that

$$\frac{\partial z^*(z_h, z_l)}{\partial z_h} = \frac{\partial \pi_f(z_h, \theta_L)}{\partial z} \frac{z_h - z^*(z_h, z_l)}{\pi_f(z_h, \theta_L) - \pi_f(z_l, \theta_L)}.$$

Therefore, we can write

$$\frac{\partial \delta(z_h, z_l)}{\partial z_h} = \frac{1}{2} S(z_h, z_l)$$

where

$$S(z_h, z_l) \equiv \alpha^L \int_{z^*(z_h, z_l)}^{z_h} \left[z - z_h + \bar{\nu} + \alpha_H \Delta\right] f(z) dz + \frac{z_h - z^*(z_h, z_l)}{\pi_f(z_h, \theta_L) - \pi_f(z_l, \theta_L)} A(z_h, z_l)$$

where

$$\begin{aligned}
A(z_h, z_l) \equiv {} & \alpha^H \left[b(\pi_f(z_l, \theta_L)) - b(\pi_f(z^*(z_h, z_l), \theta_L))\right] f(z^*(z_h, z_l) - \Delta)\\
& + \alpha^H \left[(z^*(z_h, z_l) + \theta_L + \bar{\nu} - \alpha_L \Delta)\left(\pi_f(z_l, \theta_L) - \pi_f(z^*(z_h, z_l), \theta_L)\right)\right] f(z^*(z_h, z_l) - \Delta)\\
& - \alpha^L \left[b(\pi_f(z_h, \theta_L)) - b(\pi_f(z^*(z_h, z_l), \theta_L))\right] f(z^*(z_h, z_l))\\
& - \alpha^L \left[(z^*(z_h, z_l) + \theta_L + \bar{\nu} + \alpha_H \Delta)\left(\pi_f(z_h, \theta_L) - \pi_f(z^*(z_h, z_l), \theta_L)\right)\right] f(z^*(z_h, z_l)).
\end{aligned}$$

Taking the limit as z_h tends to z_l and using L'Hôpital's rule for the term $\frac{z_h - z^*(z_h, z_l)}{\pi_f(z_h, \theta_L) - \pi_f(z_l, \theta_L)}$, yields $\partial \delta(z_h, z_l)/\partial z_h = 0$. Finally, we show that $\delta(z, z_l)$ is strictly convex at $z = z_l$. We have

$$\frac{\partial S(z_h, z_l)}{\partial z_h}|_{(z_l, z_l)} = \alpha^L \frac{\partial}{\partial z_h} \left[\int_{z^*(z_h, z_l)}^{z_h} \left[z - z_h + \bar{\nu} + \alpha_H \Delta\right] f(z) dz\right] + \frac{\partial}{\partial z_h} \left[\frac{z_h - z^*(z_h, z_l)}{\pi_f(z_h, \theta_L) - \pi_f(z_l, \theta_L)} A(z_h, z_l)\right].$$

Examining the first term, we have

$$\left.\frac{\partial \int_{z^*(z_h,z_l)}^{z_h} [z - z_h + \bar{\nu} + \alpha_H \Delta] \, f(z) dz}{\partial z_h}\right|_{z_h=z_l} = \frac{1}{2}(\bar{\nu} + \alpha_H \Delta) f(z_l).$$

For the second term we have

$$\left.\frac{\partial \frac{z_h - z^*(z_h,z_l)}{\pi_f(z_h,\theta_L) - \pi_f(z_l,\theta_L)} A(z_h, z_l)}{\partial z_h}\right|_{z_h=z_l} = \left.\frac{z_h - z^*(z_h,z_l)}{\pi_f(z_h,\theta_L) - \pi_f(z_l,\theta_L)} A_1(z_l,z_l)\right|_{z_h=z_l} = \frac{1}{2\frac{\partial \pi_f(z_l,\theta_L)}{\partial z}} A_1(z_l,z_l),$$

where we used $\lim_{z_h \to z_l} \frac{\partial \frac{z_h - z^*(z_h,z_l)}{\pi_f(z_h,\theta_L) - \pi_f(z_l,\theta_L)}}{\partial z_h} A(z_h, z_l) = 0$, because $\frac{\partial \frac{z_h - z^*(z_h,z_l)}{\pi_f(z_h,\theta_L) - \pi_f(z_l,\theta_L)}}{\partial z_h}$ converges as $z_h \to z_l$ and $A(z_l) = 0$. We have

$$A_1(z_l, z_l) = -\frac{1}{2} \frac{\partial \pi_f(z_l, \theta_L)}{\partial z} [(\bar{\nu} - \alpha_L \Delta) \, \alpha^H f(z_l - \Delta) + (\bar{\nu} + \alpha_H \Delta) \, \alpha^L f(z_l)].$$

and so

$$\frac{\partial S(z_l, z_l)}{\partial z_h} = \frac{1}{4}(\bar{\nu} + \alpha_H \Delta) \alpha^L f(z_l) + \frac{1}{4} (\alpha_L \Delta - \bar{\nu}) \, \alpha^H f(z_l - \Delta)$$

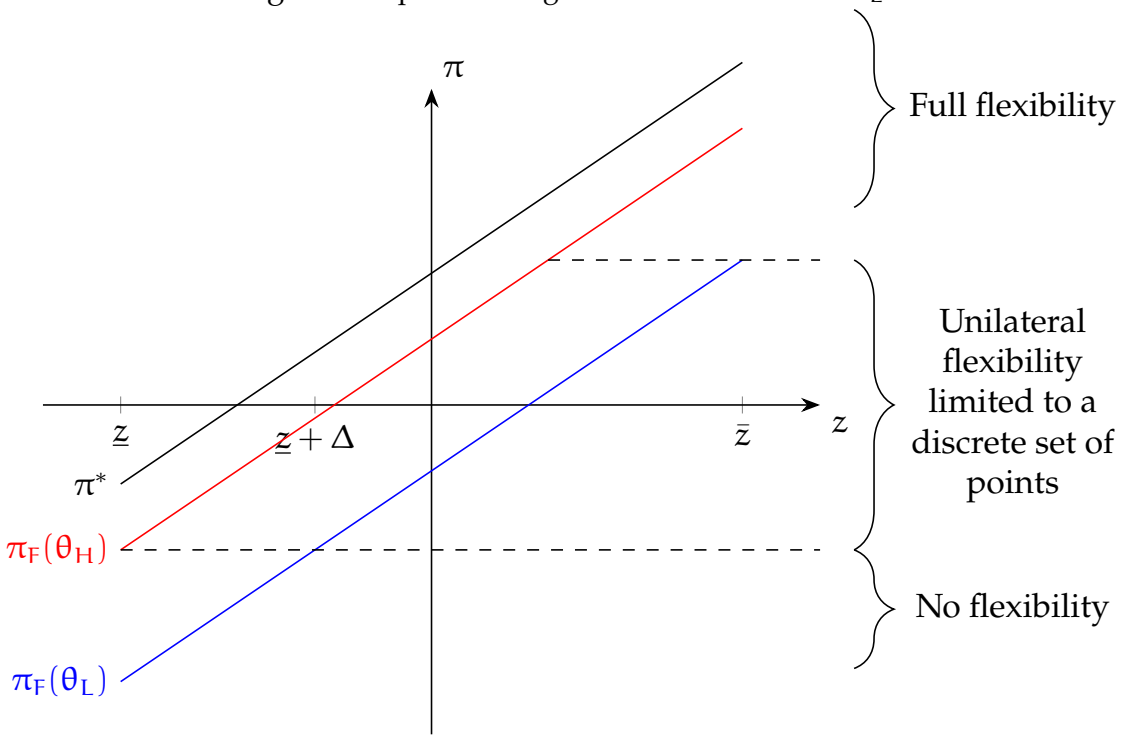
which is positive given condition 9. Q.E.D.

Proof of Proposition 2 Suppose that $\bar{\nu} \in [-\alpha_H \Delta, \alpha_L \Delta]$. Then, (9) holds for all z which implies that the delegated set contains only a discrete set of points.

Next, suppose that $\bar{\nu} > \alpha_L \Delta$. The delegation set D takes the form outlined in Figure 5. If the condition (9) holds for $z \in [\underline{z} + \Delta, \bar{z}]$ then, from the previous lemma the set D contains a discrete subset of points over $[\pi_f(\underline{z} + \Delta, \theta_L), \pi_f(\bar{z}, \theta_L)]$. Suppose now that the set D contains some point $\tilde{\pi}$ from the interval $[\pi_f(\bar{z}, \theta_L), \pi_f(\bar{z} + \Delta, \theta_L)] = [\pi_f(\bar{z} - \Delta, \theta_H), \pi_f(\bar{z}, \theta_H)]$. Then define $\tilde{D}$ to be $D \cup [\tilde{\pi}, \pi_f(\bar{z} + \Delta, \theta_L)]$. We show that societal welfare is higher under $\tilde{D}$. To see why note that if θ_H is in charge and the set is $\tilde{D}$, the chosen policy for will correspond to his flexible best response in this new range. Since the equilibrium choices in LB are always in between the flexible policies of θ_H and θ_L, this change preferable from society's perspective given the level of bias. Now suppose that θ_L is in charge. Since $\tilde{\pi} > \pi_f(\bar{z}, \theta_L)$, we have that $\mathcal{R}(z, \theta_L, \tilde{D}) = \mathcal{R}(z, \theta_L, D)$ for all z thus leaving the chosen policies the same as if the delegation set is D.

Finally, we show that the set D contains some $\tilde{\pi} \in [\pi_f(\bar{z}, \theta_L), \pi_f(\bar{z} + \Delta, \theta_L))$. If not, consider adding any point in that interval to D. If type θ_H is in charge, the status quo set $\mathcal{R}(z, \theta_H, D) \subset \mathcal{R}(z, \theta_H, D \cup \{\tilde{\pi}\})$ which increases the bargaining power of θ_H and moves the

Figure 5: Optimal delegation set in LB for $\bar{v} > \alpha_L \Delta$



policy choice closer to society. If instead θ_L is in charge, there are two cases to consider. Suppose first that $\pi_D(z, \theta_L) \neq \tilde{\pi}$ for all z. Then $\mathcal{R}(z, \theta_H, D) = \mathcal{R}(z, \theta_H, D \cup \{\tilde{\pi}\})$ and so the chosen policy is the same. Second, suppose that $\pi_D(z, \theta_L) = \tilde{\pi}$ for some z. If it is true for some z then it must be true for $\bar{z}$ since the desired policy is increasing in z. Then,

$$\begin{aligned}\tilde{\pi} &= \arg\max_{\pi}\{u(\pi, \bar{z}, \theta_L) \text{ subject to } u(\pi, \bar{z}, \theta_H) \geqslant u(\tilde{\pi}, \bar{z}, \theta_H)\} \\ &> \pi_f(\bar{z}, \theta_L) = \arg\max_{\pi}\{u(\pi, \bar{z}, \theta_L) \text{ subject to } u(\pi, \bar{z}, \theta_H) \geqslant u(\pi_D(z, \theta_L), \bar{z}, \theta_H)\}\end{aligned}$$

where the first equality follows from the observation that $\pi_f(\bar{z}, \theta_H) \geqslant \tilde{\pi} > \pi_f(\bar{z}, \theta_L)$ so there is no policy that the $\theta_L$type prefers to $\tilde{\pi}$ that is also preferred by the θ_H type, the second equality follows from the observation that if $\tilde{\pi}$ is preferred to $\pi_D(z, \theta_L)$ by the θ_L type then it must be that $\pi_D(z, \theta_L) < \pi_f(\bar{z}, \theta_L)$, and so the low type can attain its preferred policy when $\tilde{\pi}$ is not included in the set D. Thus, the chosen policy moves towards society's best response for z close to $\bar{z}$ which increases welfare.

A symmetric logic gives the result for $\bar{v} < -\alpha_H \Delta$. Q.E.D.

C Proofs of results in Section 7

We prove Proposition 5 in two steps. First, we show that there exists $U_L(\Delta)$ such that if $U \leqslant U_L(\Delta)$ then LB is preferred to DE. Second, we show that if the bias is sufficiently large, then delegating to executives dominates legislative bargaining.

LB preferred to DE if U small

Proposition 9. *If* $U \leqslant U_L(\Delta) = \frac{1}{2}\alpha_H\Delta$, *then society prefers LB to DE.*

Proof. Let D denote a delegated set to executives. $\pi_D(z,\theta)$ denotes the allocation implemented by delegating D to executives, and x_D denotes expectations of inflation associated with $\pi_D(z,\theta)$. Let $\pi_{LB}(z,\theta)$ and x_{LB} denote the allocation implemented, and the associated inflation expectations, by delegating $D + \{\frac{x_{LB}-x_D}{2}\}$ to the legislature. (A fixed point argument guarantees that such x_{LB} exists because best responses have slope $1/2 < 1$ with respect to x.)

The proof consists of showing that the welfare is higher for $\pi_{LB}(z,\theta)$ with x_{LB} than that for $\pi_D(z,\theta)$ with x_D. The proof proceeds in three steps. A first step bounds the welfare loss associated with inflation expectations x_{LB} instead of inflation expectations x_D, keeping the discretion granted to executives constant. A second step bounds the welfare gain from renegotiation through legislative bargaining, keeping the inflation expectations constant. The third step shows that the condition $U \leqslant \frac{1}{2}\alpha_H\Delta$ is sufficient for the welfare gain due to renegotiation to outweigh the welfare loss due to the change in inflation expectations.

Step 1: First, we evaluate the welfare change due to inflation expectations changing from x_D to x_{LB}. The shift in inflation expectations induces a change in the allocation implemented by delegating to executives (although inflation expectations x_{LB} are not rational at the allocation considered, this is a useful intermediary step in evaluating the change in welfare in going from DE to LB). Because

$$\pi_f(z,\theta;x_{LB}) = \pi_f(z,\theta;x_D) + \frac{x_{LB}-x_D}{2},$$

consider the delegated set $\hat{D} = D + \{\frac{x_{LB}-x_D}{2}\}$. Intuitively, shifting the delegated set keeps the discretion granted to executives constant at the new inflation expectations. As a result, the allocation implemented by delegating $\hat{D}$ to executives, with inflation expectations x_{LB}, is a vertical shift of the allocation implemented by delegating D with inflation expectations x_D. That is,

$$\pi_{\hat{D},x_{LB}}(z,\theta) = \pi_D(z,\theta) + \frac{x_{LB}-x_D}{2}.$$

We next show that the change in welfare from the allocation $\pi_D(z,\theta)$ with x_D to the

allocation $\pi_{\hat{D},x_{LB}}(z,\theta)$ with x_{LB} is

$$-(U+x_D-\bar{\theta})\frac{x_{LB}-x_D}{2}-\left(\frac{x_{LB}-x_D}{2}\right)^2.$$

To see this, let the path from x_D to x_{LB} be parametrized by $t\in[0,1]$ and

$$x(t)=tx_{LB}+(1-t)x_D.$$

Similarly, let the path from $\pi_D(z,\theta)$ to $\pi_{\hat{D},x_{LB}}(z,\theta)$ be parametrized by $t\in[0,1]$ as follows

$$\pi_t(z,\theta)=t\pi_{\hat{D},x_{LB}}(z,\theta)+(1-t)\pi_D(z,\theta).$$

The change in welfare is the integration of marginal changes along the path parametrized by t,

$$\hat{w}(x(1),\pi_{t=1}(\cdot,\cdot))-\hat{w}(x(0),\pi_{t=0}(\cdot,\cdot))=\int_0^1\frac{d\hat{w}(x(t),\pi(\cdot,\cdot))}{dt}dt,$$

where $\hat{w}$ denotes society's expected welfare. Faction $i\in\{H,L\}$is in power with probability α^i, hence

$$\hat{w}(x,\pi(\cdot,\cdot))=\alpha^L\hat{v}(x,\pi(\cdot,\cdot),\theta_L)+\alpha^H\hat{v}(x,\pi(\cdot,\cdot),\theta_H),$$

where

$$\hat{v}(x,\pi(\cdot,\cdot),\theta)=\int_{\underline{z}}^{\bar{z}}\left[R\left(x,\pi(z,\theta),z,\bar{\theta}\right)\right]dF(z).$$

The marginal change in welfare is the Gateaux derivative in the direction $h(z,\theta)=\pi_{\hat{D},x_{LB}}(z,\theta)-\pi_D(z,\theta)$ and $h_x=x_{LB}-x_D$,

$$\frac{d\hat{w}(x(t),\pi(t,z))}{dt}=\alpha^L\partial\hat{v}(x(t),\pi(t,z,\theta),\theta_L;h_x,h(\cdot,\cdot))+\alpha^H\partial\hat{v}(x(t),\pi(t,z,\theta),\theta_H;h_x,h(\cdot,\cdot)),$$

where

$$\begin{aligned}\partial\hat{v}(x,\pi(\cdot,\cdot),\theta;h_x,h(\cdot,\cdot))=&\int_{\underline{z}}^{\bar{z}}\left[\frac{\partial R(x,\pi(z,\theta),z,\bar{\theta})}{\partial x}h_x\right]dF(z)\\&+\int_{\underline{z}}^{\bar{z}}\left[\frac{\partial R(x,\pi(z,\theta),z,\bar{\theta})}{\partial\pi}h(z,\theta)\right]dF(z).\end{aligned}$$

Using $\frac{\partial R(x,\pi,z,\bar{\theta})}{\partial x}=-(U+x-\pi)$, and $\frac{\partial R(x,\pi(z,\theta),z,\bar{\theta})}{\partial\pi}=(U+x-\pi)-(\pi-z-\bar{\theta})$, it follows

that

$$\partial\hat{v}(x(t),\pi_t(z,\theta),\theta;h_x,h(\cdot,\cdot)) = -\left[\frac{U+x_D-\bar{\theta}}{2}+t\frac{x_{LB}-x_D}{2}\right](x_{LB}-x_D),$$

given that $h(z,\theta)=\frac{x_{LB}-x_D}{2}$ and $h_x = x_{LB}-x_D$.

As a result,

$$\hat{w}(x(1),\pi(1,\cdot,\cdot))-\hat{w}(x(0),\pi(0,\cdot,\cdot)) = -\frac{U+x_D-\bar{\theta}}{2}(x_{LB}-x_D)-\left(\frac{x_{LB}-x_D}{2}\right)^2. \quad (37)$$

Step 2: The second step evaluates the welfare implications of allowing for renegotiation, while keeping inflation expectations fixed at x_{LB}. That is, the welfare change from $\pi_{\hat{D},x_{LB}}(z,\theta)$ to $\pi_{LB}(z,\theta)$, given inflation expectations x_{LB}.

The Gateaux derivative in the direction $h(z,\theta)$, keeping inflation expectations fixed, that is $h_x=0$, is

$$\partial\hat{v}(x_{LB},\pi(\cdot,\cdot),\theta;h_x=0,h(\cdot,\cdot)) = \int_{\underline{z}}^{\bar{z}}\left[\left(U+x_{LB}+z+\bar{\theta}-2\pi(z,\theta)\right)h(z,\theta)\right]dF(z).$$

Let the path from $\pi_{\hat{D},x_{LB}}(z,\theta)$ to $\pi_{LB}(z,\theta)$ be parametrized by $t\in[0,1]$ as follows

$$\pi^r_t(z,\theta) = t\pi_{LB}(z,\theta)+(1-t)\pi_{\hat{D},x_{LB}}(z,\theta),$$

and let $h^r(z,\theta)=\pi_{LB}(z,\theta)-\pi_{\hat{D},x_{LB}}(z,\theta)$.

We are going to show that: If $h^r(z,\theta_L)>0$ then

$$U+x_{LB}+z+\bar{\theta}-2\pi^r_t(z,\theta) = 2(1-t)h^r(z,\theta_L)+\alpha_H\Delta>0. \quad (38)$$

If $h^r(z,\theta_L)<0$ then

$$U+x_{LB}+z+\bar{\theta}-2\pi^r_t(z,\theta) = (1-2t)h^r(z,\theta_L)-\alpha_L\Delta. \quad (39)$$

Similarly, if $h^r(z,\theta_H)<0$, then

$$U+x_{LB}+z+\bar{\theta}-2\pi^r_t(z,\theta) = 2(1-t)h^r(z,\theta_H)-\alpha_L\Delta<0. \quad (40)$$

Lastly, if $h^r(z,\theta_H)>0$ then

$$U+x_{LB}+z+\bar{\theta}-2\pi^r_t(z,\theta) = (1-2t)h^r(z,\theta_H)+\alpha_H\Delta. \quad (41)$$

To see these results, notice that for $h^r(z,\theta_L)>0$, because $\theta_L\leqslant\theta_H$, then the low type can

attain its preferred policy, $\pi_{LB}(z,\theta_L) = \pi_f(z,\theta_L)$. Because R is concave and the marginal payoff for θ_L at $\pi_f(z,\theta_L)$ is zero, it follows that

$$U + x_{LB} - 2\pi_{\hat{D},x_{LB}}(z,\theta_L) + z + \theta_L = 2h^r(z,\theta_L) > 0.$$

More generally, for $t \in [0,1]$, because the marginal utility is linear in the policy, if $h^r(z,\theta_L) > 0$, then

$$U + x_{LB} - 2(\pi_{\hat{D},x_{LB}}(z,\theta_L) + th^r(z,\theta)) + z + \theta_L = 2(1-t)h^r(z,\theta_L).$$

Adding $\alpha_H\Delta$ on both sides of the equality gives (38).

Similarly, if $h^r(z,\theta_L) < 0$ instead, then $\pi_{\hat{D},x_{LB}}(z,\theta_L) > \pi_f(z,\theta_H)$. Because R is concave and the marginal payoff for θ_H at $\pi_f(z,\theta_H)$ is zero, it follows that,

$$U + x_{LB} - 2\pi_{\hat{D},x_{LB}}(z,\theta_L) + z + \theta_H < 0.$$

Subtracting $\alpha_L\Delta$ on both sides of the inequality gives

$$U + x_{LB} - 2(\pi_{\hat{D},x_{LB}}(z,\theta_L) + th^r(z,\theta_L)) + z + \bar{\theta} = (1-2t)h^r(z,\theta_L) - \alpha_L\Delta$$

which is (39).

The argument for $h^r(z,\theta_H)$ is symmetric.

The change in welfare is the integration of marginal changes along the path parametrized by t,

$$\hat{w}(x_{LB},\pi^r_{t=1}(\cdot,\cdot)) - \hat{w}(x_{LB},\pi^r_{t=0}(\cdot,\cdot)) = \int_0^1 \frac{d\hat{w}(x_{LB},\pi^r_t(\cdot,\cdot))}{dt}dt.$$

Based on (38)–(41), we partition the range of shocks as follows, $Z(+,\theta) \equiv \{z \in Z \mid h^r(z,\theta) > 0\}$, $Z(-,\theta) \equiv \{z \in Z \mid h^r(z,\theta) < 0\}$, and $Z(=,\theta) \equiv \{z \in Z \mid h^r(z,\theta) = 0\}$, which gives

$$\begin{aligned}\partial\hat{v}(x_{LB},\pi^r_t(\cdot,\cdot),\theta_L;h_x=0,h^r(\cdot,\cdot)) = &\int_{Z(+,\theta_L)} \left[(2(1-t)h^r(z,\theta_L) + \alpha_H\Delta)\, h^r(z,\theta_L)\right] dF(z) \\ &+ \int_{Z(-,\theta_L)} \left[((1-2t)h^r(z,\theta_L) - \alpha_L\Delta)\, h^r(z,\theta_L)\right] dF(z).\end{aligned}$$

Integrating over the path from $\pi_{\hat{D},x_{LB}}(z,\theta_L)$ to $\pi_{LB}(z,\theta_L)$,

$$\begin{aligned}&\int_0^1 \partial\hat{v}(x_{LB},\pi^r_t(\cdot,\cdot),\theta_L;h_x=0,h^r(\cdot,\cdot))dt \\ &= \int_{Z(+,\theta_L)} \left[h^r(z,\theta_L)^2\right] dF(z) + \alpha_H\Delta \int_{Z(+,\theta_L)} \left[h^r(z,\theta_L)\right] dF(z) - \alpha_L\Delta \int_{Z(-,\theta_L)} \left[h^r(z,\theta_L)\right] dF(z).\end{aligned}$$

Similarly for θ_H,

$$\partial\hat{v}(x_{LB}, \pi^r_t(\cdot,\cdot), \theta_H; h_x = 0, h^r(\cdot,\cdot)) = \int_{Z(+,\theta_H)} [((1-2t)h^r(z,\theta_H) + \alpha_H\Delta)\, h^r(z,\theta_H)]\, dF(z) + \int_{Z(-,\theta_H)} [(2(1-t)h^r(z,\theta_H) - \alpha_L\Delta)\, h^r(z,\theta_H)]\, dF(z).$$

And, integrating over the path from $\pi_{\hat{D},x_{LB}}(z,\theta_H)$ to $\pi_{LB}(z,\theta_H)$,

$$\int_0^1 \partial\hat{v}(x_{LB}, \pi^r_t(\cdot,\cdot), \theta_H; h_x = 0, h^r(\cdot,\cdot))dt = \int_{Z(-,\theta_H)} \left[h^r(z,\theta_H)^2\right] dF(z) + \alpha_H\Delta \int_{Z(+,\theta_H)} [h^r(z,\theta_H)]\, dF(z) - \alpha_L\Delta \int_{Z(-,\theta_H)} [h^r(z,\theta_H)]\, dF(z).$$

Step 3: Finally, we are going to show that if $2U \leqslant \alpha_H\Delta$, then the benefits outlined in the second step outweigh the cost outlined in the first step and summarized in (37). That is, it suffices to show that,

$$\alpha_H\Delta \sum_i \alpha^i \int_{Z(+,\theta_i)} [h^r(z,\theta_i)]\, dF(z) - \alpha_L\Delta \sum_i \alpha^i \int_{Z(-,\theta_i)} [h^r(z,\theta_i)]\, dF(z) + \sum_i \alpha^i \int_{\underline{z}}^{\bar{z}} [h^r(z,\theta_i)]^2\, dF(z) \geqslant (U + x_D - \bar{\theta})\frac{x_{LB}-x_D}{2} + \left(\frac{x_{LB}-x_D}{2}\right)^2. \tag{42}$$

Using the implementability condition (13) we can write

$$\begin{aligned} x_{LB} - x_D &= \sum_i \alpha^i \int [\pi_{LB}(z,\theta_i) - \pi_D(z,\theta_i)] f(z) dz \\ &= \sum_i \alpha^i \int [\pi_{LB}(z,\theta_i) - \pi_{\hat{D},x_{LB}}(z,\theta_i) + \pi_{\hat{D},x_{LB}}(z,\theta_i) - \pi_D(z,\theta_i)] f(z) dz \\ &= \sum_i \alpha^i \int [h^r(z,\theta_i) + h(z,\theta_i)] f(z) dz. \end{aligned}$$

where the first equality follows from (13), the second from adding and subtracting $\pi_{\hat{D},x_{LB}}(z,\theta_i)$, and the last one uses the definition of h^r and h. Thus, the change in inflation expectations between the two regimes can be expressed in terms of h^r and h.

Because the set $\hat{D}$ is such that $h(z,\theta_i) = \frac{x_{LB}-x_D}{2}$, it follows that

$$\frac{x_{LB}-x_D}{2} = \sum_i \alpha^i \int [h^r(z,\theta_i)] f(z) dz, \tag{43}$$

and, by Jensen's inequality,

$$\left(\frac{x_{LB}-x_D}{2}\right)^2 \leqslant \sum_i \alpha^i \int [h^r(z,\theta_i)]^2 f(z)\,dz.$$

Hence, using the above into (42), to prove our result it suffices to show that

$$\alpha_H \Delta \sum_i \alpha^i \int_{Z(+,\theta_i)} [h^r(z,\theta_i)]\,dF(z) - \alpha_L \Delta \sum_i \alpha^i \int_{Z(-,\theta_i)} [h^r(z,\theta_i)]\,dF(z)$$
$$\geqslant (U + x_D - \bar{\theta})\frac{x_{LB}-x_D}{2}.$$

The second term on the left-hand side of the inequality subtracts a negative term because $h^r(z,\theta_i) < 0$ for $z \in Z(-,\theta_i)$. Thus, for (42) to hold it suffices to show that

$$\alpha_H \Delta \sum_i \alpha^i \int_{Z(+,\theta_i)} [h^r(z,\theta_i)]\,dF(z) \geqslant (U + x_D - \bar{\theta})\frac{x_{LB}-x_D}{2}.$$

Note that

$$\sum_i \alpha^i \int_{Z(+,\theta_i)} [h^r(z,\theta_i)]\,dF(z) \geqslant \sum_i \alpha^i \int_{\underline{z}}^{\bar{z}} [h^r(z,\theta_i)]\,dF(z) = \frac{x_{LB}-x_D}{2},$$

where the last equality follows from (43). Thus, the following inequality is a sufficient condition for the result to hold:

$$\alpha_H \Delta > U + x_D - \bar{\theta}.$$

We assumed that $U \leqslant \frac{1}{2}\alpha_H \Delta$ so for the above inequality to hold it is sufficient to show that $\alpha_H \Delta \geqslant 2U \geqslant U + x_D - \bar{\theta}$ or simply that $x_D \leqslant U + \bar{\theta}$. Suppose by way of contradiction that $x_D > U + \bar{\theta}$. This contradicts that π_D solves the mechanism design problem of delegation to executives because even in the Nash equilibrium (with full flexibility) expected inflation would be $U + \bar{\theta}$ and so the Nash equilibrium outcome would achieve higher welfare than π_D with inflation expectations equal to $U + \bar{\theta}$. But this is not possible since the Nash equilibrium outcome is feasible for the delegation problem. Thus, we have a contradiction and $x_D \leqslant U + \bar{\theta}$, implying that condition (42) holds. Q.E.D.

DE preferred to LB if U large

In contrast, if the bias is sufficiently large, then delegating to executives dominates legislative bargaining.

Proposition 10. *If $\Delta = 0$ or $U \geqslant \alpha_H \Delta + \frac{1}{2}E[z - \underline{z}]$, then society prefers DE to LB.*

Proof. We compare a lower bound for the welfare associated with DE to an upper bound for the welfare associated with LB. The welfare associated with DE at least as high as the welfare resulting from delegating a narrow mandate at the best constant policy $D = \{\bar{\pi}^*\}$, where $\bar{\pi}^* = \bar{\theta}$, to the monetary authority. Thus,

$$V_{DE} \geqslant \bar{V}^* = \int \sum_i \alpha_i R\left(\bar{\pi}^*, \bar{\pi}^*, z, \theta_i\right) f(z)\, dz = -\frac{1}{2}\left[U^2 + \sum_i \alpha_i \left(\bar{\theta} - \theta_i\right)^2 + var(z)\right]$$

To find an upper bound for the LB value, not the the LB outcome is such that

$$\pi_f\left(x_{LB}, \theta_L, z\right) \leqslant \pi_{LB}(\theta, z) \leqslant \pi_f\left(x_{LB}, \theta_H, z\right)$$
$$x_{LB} = \sum_i \alpha^i \int \pi_{LB}\left(\theta_i, z\right) f(z)\, dz$$

so $x_{LB} \leqslant x_M(\theta_L) = U + \theta_L + \frac{z}{2}$ where $x_M(\theta_L)$ is the expected inflation in the Markov equilibrium where the θ_L type chooses policy for sure. Thus,

$$V_{LB} \leqslant \bar{V}_{LB} = \int \sum_i \alpha_i R\left(x_M(\theta_L), \pi_f\left(x_M(\theta_L), z, \bar{\theta}\right), z, \theta_i\right) f(z)\, dz$$

Direct calculations shows that $\bar{V}^* > \bar{V}_{LB}$ if

$$U^2 > \frac{1}{2} var(z) + \frac{\Delta^2}{4} + U\bar{\theta} \tag{44}$$

which implicitly defines $U_H(\Delta)$. Thus, if $U \geqslant U_H(\Delta)$ then $V_{DE} > V_{LB}$. Q.E.D.

D Proofs of results in Section 8

Proof of Proposition 7

First we show that if U is large enough then the best constant policy is credible. Note that the value of the constant Ramsey policy for type θ is

$$\bar{V}^*(\theta) = \int R\left(\bar{\pi}^*, \bar{\pi}^*, z, \theta_i\right) f(z)\, dz = -\frac{1}{2}\left[U^2 + \left(\bar{\theta} - \theta\right)^2 + var(z)\right]$$

and the value of LB for type θ is at most $\bar{V}_{LB}(\theta)$ defined as

$$V_{LB}(\theta) \leqslant \bar{V}_{LB}(\theta) = \int R\left(x_M(\theta_L), \pi_f\left(x_M(\theta_L), z, \theta\right), z, \theta\right) f(z)\, dz$$

where $x_M(\theta_L)$ is the Markov equilibrium expected inflation defined in the proof of Proposition 9 and it is a lower bound for the equilibrium inflation under LB. Direct calculations show that

$$U^2 > \frac{1}{2}var(z) + 2U\Delta + \frac{3}{4}\Delta^2 \tag{45}$$

are a set of sufficient conditions for $\bar{V}_{LB}(\theta) < \bar{V}^*(\theta)$ for all θ. This is turn implies that $\bar{V}^*(\theta) > V_{LB}(\theta)$ for all θ and so factions do not want to switch to LB from the constant Ramsey policy.

We next show that for U sufficiently high the constant Ramsey policy is the optimal policy under DE. Clearly, the constant Ramsey policy is preferred to all the other constant policies i.e. policies that do not vary with z. We next compare the value of the constant Ramsey policy to an upper bound for all policies feasible in DE that allows for some discretion. For such class of policies, the expected inflation must be at least equal to

$$x_{min} = \theta_L + \frac{z_l}{2} + U = \min_{z,\theta} \pi_f(x, z, \theta)$$

This minimal amount of expected inflation is higher than the one under the constant Ramsey policy, $\bar{x}^* = \bar{\theta}$, if U is sufficiently high. Thus, the constant Ramsey policy is beneficial in this respect. With this lower bound, we can write that

$$V_{DE}(\theta) \leqslant \bar{V}_{DE}(\theta) = \int R(x_{min}, \pi_f(x_{min}, z, \theta), z, \theta) f(z)\, dz$$

$$V_{DE} \leqslant \bar{V}_{DE} = \int \sum_i \alpha_i R\left(x_{min}, \pi_f\left(x_{min}, z, \bar{\theta}\right), z, \theta_i\right) f(z)\, dz$$

Direct calculations show that if

$$U^2 > \frac{1}{2}var(z) + \frac{\Delta^2}{4} + 2U\Delta + U|z_l| + 2\left(\frac{1}{2}\Delta + \frac{|z_l|}{4}\right)^2 \tag{46}$$

then $\bar{V}_{DE}(\theta) < \bar{V}^*(\theta)$ for all θ and so the constant Ramsey outcome is the preferred outcome under DE.

Defining $\bar{U}_H$ as the smallest U such that conditions (45) and (46) hold, we have that for all $U \geqslant \bar{U}_H$ DE is preferred to LB and the optimal DE outcome is a credible narrow mandate set at the constant Ramsey policy. Q.E.D.