The Political Economy of Compensatory Federalism*

Sanford C. Gordon  Dimitri Landa
Department of Politics  Department of Politics
New York University  New York University
sanford.gordon@nyu.edu  dimitri.landa@nyu.edu

Abstract

To what extent does the federal structure of policymaking in the United States mitigate or exacerbate national political conflict? We develop a model of two-level governance in a federal system in the presence of interstate preference heterogeneity and cross-state externalities. The key underlying intuition is that states with high demand for public spending or regulation are better positioned to adjust state-level policies to compensate for perceived inadequacies in national policy than corresponding states with low demand. We explore the normative and behavioral implications of this asymmetry under majoritarian and supermajoritarian national policymaking institutions, and use the model to account for a number of empirical regularities in U.S. politics and policymaking.

*Word count: 9,897.
On June 1, 2017, President Donald J. Trump announced his intention to withdraw the United States from the Paris Climate Accord. Within a week, representatives from nine states and 125 cities (a number that, as of writing, has grown to 150) signed an “open letter to the international community,” declaring its signatories’ intention to “support climate action to meet the Paris Agreement.” Governor Jerry Brown declared, “Trump is AWOL but California is on the field, ready for battle.”

State and local reaction to Trump’s announcement provides an example of what observers of intergovernmental relations term *compensatory federalism*, a situation in which “governments at one level are able to compensate for weaknesses or defects at another level” (Derthick 2010). Specific applications of the concept include environmental policy, health care, old age pensions, and others (Derthick 2010; Thompson 1998; Pandey 2002). Partial preemption, the practice whereby the national government sets regulatory “floors” that the states are permitted to exceed but not fall below (O’Reilly 2006; Scicchitano and Hedge 1993), opens the possibility of compensatory federalism by a subset of states. And left-leaning scholars have something close to compensatory federalism in mind when describing “progressive federalism” (e.g., Gerken 2004; Freeman and Rogers 2010; see also Young 2003) as a broader set of decentralized policy responses to Republican dominance in national politics.

Together with the supermajoritarian constraints created by the separation of powers at the national level, federalism is a core component of the United States’ constitutional design. In this paper, we develop a model to explore how the mechanisms of compensatory federalism interact with such supermajoritarian constraints to produce heretofore neglected implications for policy and politics that shed light on the governability and stability of federal democracies. While overlapping authority over policymaking between state and national governments and supermajoritarian lawmaking requirements at the national level are relatively invariant, they define an environment in which other institutional features, such as the status quo authority of the national government and the extent of cross-state externalities, vary over
time and across specific policy or issue dimensions. Our primary focus is on how that variation, anchored to the more permanent features of the constitution, affects political failure – understood as departures of equilibrium outcomes from the first-best – and the incentives of political “losers” to resist status quo policies. While the main substantive focus is the U.S., our results generalize to other political systems sharing these features.

In the model, incentives in the national policymaking process derive from the following predicament facing the states: a state’s ability to compensate at the state level for perceived deficiencies in federal policy – whether it be a regulatory mandate or the provision of a (possibly local) public good – is itself contingent on the magnitude of the federal government’s policymaking presence. When federal regulation is lax or federal provision low, states with a demand for more extensive policy may offset, albeit imperfectly, the perceived inadequacy via compensating state-level policies. A sufficiently extensive national policy, however, will crowd out states with low demand for spending or regulation. Those states will have a harder time adjusting state-level policy to meet local demand in response to what they would perceive as excessive federal-level action. In other words, at a given level of national policy, states with still higher demand can “top up,” but those with lower demand will not be able to “top down.”

An illustrative example of this dynamic in action is minimum wage policy. Because the federal minimum wage is not adjusted for inflation, its real value erodes over time. When the real value is high, the diversity of effective state minimum wages associated with compensatory federalism will be low because of the inability of states to top down. Over time, as inflation erodes the real federal value, topping up by states with preferences for higher minima drives up interstate policy diversity. Figure 1 shows this pattern, plotting the coefficient of variation for quarterly state-level minimum wage against the real federal minimum wage. Each connected line corresponds to a single nominal value in effect for a particular time period (for clarity we label the four longest periods); inflation erosion corresponds to a leftward shift in its real value. The pattern is unmistakable: decreases over
time in the real federal minimum associated with a single nominal value are matched with increases in compensatory federalism.

Our core results develop normative, institutional, and behavioral implications of the asymmetry inherent in compensatory federalism and contrast them with those that would emerge under a unitary system of government. First, we derive a normative benchmark for federal policy, and show that this policy is decreasing in preference heterogeneity across states. Second, we demonstrate that when the national policy that would be arrived at via simple majority rule leaves room for topping up, that policy generally exceeds the normative benchmark. Third, we show that under gridlock-inducing supermajoritarian national institutions, if state-level demand for government action is sufficiently heterogeneous then decreases from far-reaching federal policies improve aggregate welfare more than increases from limited ones. We further draw out some of the implications of these results for the contingency of aggregate welfare comparisons on national policymaking institutions.

One of our main behavioral result concerns the extent of anticipated resistance to different gridlocked policies. We demonstrate that states with high demand for government provision or regulation will be less motivated to resist policies gridlocked at a low level than states with low demand to resist those gridlocked at a high level. In a dynamic extension of our model, we further show that the shadow of the future exacerbates political polarization, but does so asymmetrically, by leading low-demanders to adopt disproportionately more extreme policy positions.

In the most general terms, then, one of the main contributions of the model is to show that a compensatory federal structure upsets a strong intuition deriving from canonical spatial models of politics and policymaking: that the polarity of a policy space (i.e., which direction corresponds to “left wing” and which to “right”) is simply a normalization, with no substantive consequences for the aggregate characteristics of political conflict. Contrary to this intuition, we show that under the institutions of compensatory federalism, whether a federal policy is more or less extensive has fundamentally asymmetric effects on social
The figure plots the coefficient of variation for effective state minimum wages against the real federal minimum wage (2009 US dollars), by quarter. Connected points correspond to unique nominal federal minima. Data from Vaghul and Zipperer (2016).
welfare and potential political conflict.

Our formal analysis sheds light on empirical regularities in U.S. politics and policymaking that have been the subject of recent discussion and analysis: the tendency of politically conservative states to be net recipients of federal programs; the absence of substantial inter-jurisdictional sorting by citizens for political reasons; and asymmetric, pervasive polarization at the national level. We show that these patterns conform to behavioral implications derived from our model.

Background

The Object of Study

Although it may be applied comparatively, the model we present seeks to capture three aspects of the federal structure of the U.S. political system. The first is the existence of heterogeneity in demand for public policy. While of course not unique to federal systems, such heterogeneity is central to justifications of federalism (Riker 1964; De Figueiredo and Weingast 2005). The second is the existence of supermajoritarian institutions for lawmaking at the national level: e.g., bicameralism, provisions for overriding a presidential veto, the filibuster, and any legislative rules that limit proposal rights to a restricted group of public officials. The third critical aspect is the presence of de facto shared sovereignty between the national and state governments with permeable boundaries (Rose-Ackerman 1981). We follow much of the qualitative literature on federalism since Grodzins (1966) in adopting the view that these boundaries do not represent a clean partition, but are better described as a “marble cake” (Grodzins 1966; see also Riker 1975). The permeability of boundaries between national and state governance raises the possibility that the former may crowd out the latter (Bradford and Oates 1971; Volden 2005; Hafer and Landa 2007).

In focusing on these three features of federalism, we necessarily abstract away from several others. To focus on across-state preference heterogeneity, we abstract away from within-state heterogeneity, and, consequently, also do not consider representation failure at the state level.
The interaction between within- and across-state heterogeneity is clearly important, but beyond the scope of the current inquiry.\footnote{See Beramendi and Jensen (2015) for a discussion of malapportionment and diverse preferences within states. We return to the issue of intrastate heterogeneity in the conclusion.} Further, we adopt the approach, common in the literature, in which the national government imposes a uniform (floor) policy across all states. While in reality, provision by the national government need not be completely homogeneous, one can think of the assumption of a uniform policy as a reduced form representation of an expectation that a policy implemented exclusively by the national government across states will be more homogeneous than one arrived at if states implemented policy exclusively within their own borders in the absence of a national government.

Related Research

In addition to the literature on compensatory federalism cited above, our research relates to the literature on fiscal federalism dating to Oates (1972), an important branch of which focuses on the role of central government in mitigating distortions induced by local taxation in the presence of spillovers (e.g., Gordon 1983; Myers 1990; Krelove 1992; for reviews see Inman and Rubinfeld 1997 and Oates 1999). In a more political vein, scholars have analyzed the effects of spillovers and status quo policies at the state level on demand for, and feasibility of, national policies (Rose-Ackerman 1981); the relationship between fiscal externalities between the states, commitment problems at the national level, and the ability of federal systems to maintain fiscal discipline (Rodden 2006); and institutional features of federalism that minimize encroachments by one level of government on the prerogatives of another Bednar (2009). Besley and Coate (2003) demonstrate that even without assuming uniform national policy, owing to misallocation, political uncertainty, and voters’ incentives to elect high-demanders to the legislature (see also Inman and Rubinfeld 1997). Relatedly, Volden (2005) describes a model in which both national and subnational governments can provide public goods and services, but credit-claiming by politicians can yield inefficient overprovision.
Recent work on the political economy of federations has focused on the strategic analysis of the implications of interstate spillovers. Most closely related to the current paper are papers by Crémer and Palfrey (2000, 2006), who analyze federal systems in which the national government sets mandates for state provision that generate positive spillovers; and Alesina et al. (2005) and Hafer and Landa (2007), who analyze “dual provision” models of federalism in which public goods provision with spillovers across states takes place both at state and federal levels. Both Crémer and Palfrey and Hafer and Landa study the effects of externalities and redistributive tensions on coalition formation and policy choice at the national level, while Alesina et al. focus on the determinants of the composition of political unions. In contrast, our model takes the union as a given and is chiefly concerned with the interaction between decision-making at multiple levels of government and supermajoritarian national institutions, and the consequences of that interaction for government failure.

The Model

Primitives

We model policy making as corresponding to government choices taking place at the national and state levels. There is a continuum of states with measure one. Each state \( i \) is characterized by a preference parameter whose support is a compact, convex subset of the positive real line, \( \alpha_i \in [\alpha, \bar{\alpha}] \), with log-concave probability density function \( p(\cdot) \) and cumulative distribution function \( P(\cdot) \). Higher values of \( \alpha \) correspond to higher valuations of the government program.

Policy made at the national level is denoted by \( F \in \mathbb{R}_+ \), while the policy selected by each state \( i \) over and above \( F \) is denoted by \( S_i \in \mathbb{R}_+ \). Define the effective policy in state \( i \) as \( Z_i = F + S_i \), with \( Z \in [0, Z] \). For example, suppose the policy in question were the minimum wage. As of writing, the federal minimum wage is $7.25/hr. California’s is $10.50/hr. So given our formalism, \( F = 7.25 \), \( S_{CA} = 3.25 \), and \( Z_{CA} = 10.50 \). For reasons that will become

\(^2\)The assumption of a continuum is a mathematical convenience.
immediate below, we will assume that \( Z > \bar{\alpha} \).

State \( i \)'s utility is expressed as

\[
\alpha_i (Z_i + \beta \Psi(Z)) - \frac{Z_i^2}{2}, \tag{1}
\]

where \( Z \) is the vector of \( Z_i \)'s and

\[
\Psi(Z) = \int Z_i(\alpha_i)dP(\alpha_i)
\]

represents the total spillovers from all states. The parameter \( \beta > 0 \) captures how much a state cares about policy in other states relative to policy within its own borders.

Let \( B \in \mathcal{B} \) represent the federal bargaining protocol, which maps the states’ preference profile and the (exogenously given) status quo federal level of provision into a level of federal provision \( F \).\(^3\) Rather than assume a specific bargaining protocol, we will assume that \( \mathcal{B} \) contains q-rules, as well as legislative bargaining institutions with gatekeepers or veto players such as Pivotal Politics (Krehbiel 1996, 1998) or Negative Agenda Control (Cox and McCubbins 2005). With the exception of simple majority rule, these institutions will often yield a (non-degenerate) compact and convex interval of policies that are gridlocked, that is, that cannot be beaten by another policy even if a majority favors policy change.

The game unfolds as follows:

1. The Federation decides on federal policy \( F \) via bargaining protocol \( B \);
2. Each state decides on its own state policy, \( S_i \);
3. Payoffs are realized.\(^4\)

\(^3\)Formally, let \( F^0 \in \mathbb{R}_+ \) be the status quo federal policy; \( \mathcal{U}(p(\cdot|\alpha, \bar{\alpha})) \) be the preference profile within the federation given the distribution of the \( \alpha_i \)s, \( p(\cdot|\alpha, \bar{\alpha}) \), and \( \mathcal{U} \) be the set of all preference profiles. Then \( B := \mathcal{U} \times \mathbb{R}_+ \to \mathbb{R}_+ \).

\(^4\)One could envision a variant of the model in which states move first, followed by federal action, followed by state adjustment. This would produce equivalent results.
Interpreting the Setting

Before proceeding to a description of the equilibrium, we provide several interpretive comments. First, the primary goal of our inquiry is to examine the relationship among the three essential features of American federalism described above, abstracting away from others. To do so, we write down the simplest possible model consistent with a conception of the building blocks of federalism and supermajoritarianism outlined above, and draw out novel political implications of those features in a public goods provision environment. Key departures of this environment from the standard model of public goods provision (utility maximization of a mix of public and private consumption subject to Inada conditions and a budget constraint) are purely instrumental. The symmetry of the quadratic functional form, for example, allows us to isolate asymmetries in outcomes brought about by the structure of federal institutions, which may be obscured with other functional forms that themselves generate asymmetric indirect utilities even in the absence of federal institutions (e.g. log-linear).

Second, we model effective policy in a state as the sum of a federal component common to all states and a state-specific component. This formalization of compensatory federalism is consistent with two interpretations. The first, “dual provision” interpretation, is government spending on public goods in states, with the federal government providing a floor level of expenditure above which the states may spend. The second, “regulatory” interpretation is regulatory federalism with partial preemption: the federal policy may be interpreted as a minimum standard that the states must comply with, though they are permitted to exceed that standard.\(^5\)

Finally, similar to Alesina et. al. (2005), we model states’ demand for government-funded projects as primitive in order to focus on specific properties of collective choices in a federal setting. In practice, state demand is a function of a variety of economic antecedents (i.e., derived from comparisons of the marginal values of public project and private consumption

\(^5\)In the absence of total remediation of the harms a regulatory regime seeks to address, modeling cross-state spillovers as positive and increasing in the intensity of regulation is equivalent to modeling them as negative and decreasing in the harms themselves.
in a redistributive setting [e.g., Hafer and Landa 2007]) and socio-cultural factors (e.g., prior immigrant group experiences with oppressive governments [Fischer 1989], or modernization-induced liberalism [Ingelhart and Welzel 2005]; see also Elazar 1966). The relationship between a state’s preference parameter $\alpha_i$ and, e.g., its average income must, thus, depend on an underlying preference-generating mechanism: for an economics-focused redistributive mechanism, high-demanders may be relatively poor states; for a socio-cultural mechanism, relatively rich ones.\footnote{In fact, as modeled, the state’s utility does not include a term for state income. As long as the federal government does not completely expropriate state income, this exclusion is immaterial for our results.} Recognizing these possibilities, we do not commit ourselves to a particular unmodeled mechanism generating demand for public projects.\footnote{A similar approach is taken in Alesina, Angeloni, and Etro 2005.}

**Induced Preferences Over Federal Policy**

The equilibrium concept is subgame perfect Nash. We proceed by backward induction and begin by considering the state policymaking subgame.

*State-level policymaking.* In the last stage of the game, federal policy $F$ has been set, and the states condition their choices on $F$. A state’s utility is globally concave in $S_i$. Solving state $i$’s first-order condition yields the optimal state policy

$$S_i^*(F) \equiv \max\{0, \alpha_i - F\}. \quad (2)$$

This expression immediately gives rise to the following remark:

**Remark 1 (Crowding Out)** A state’s own policy is weakly decreasing in the level of federal policy.

A prima facie interpretation of Remark 1 is that the model rules out the possibility of “flypaper effects,” wherein program-specific state and local expenditures appear to increase in response to intergovernmental aid (e.g., Courant, Gramlich, and Rubinfeld 1979; but see Knight 2002). A straightforward tweak of the model would be to introduce complementarities
between \( F \) and \( S_i \), which would generate flypaper effects. To the extent that complementarities between federal and state policy do exist, they would only affect the results that follow if they are sufficiently large to offset the crowding-out effect.\(^8\)

Given each state’s best response, for a given federal policy \( F \) the states will be partitioned into two groups: those with \( \alpha_i \leq F \) (“low-demanders”) will implement \( S^*_i = 0 \), making the effective policy in those states the federal floor, \( F \). Those states with \( \alpha_i > F \) (“high-demanders”) will “top up” with \( S^*_i > 0 \) to offset the perceived deficiency in the federal floor, thus achieving effective policy \( \alpha_i \). Substituting \( S^*_i(F) \) into \( \Psi \) and simplifying (see Appendix for details) gives the equilibrium value of \( \Psi \) as a function of \( F \),

\[
\Psi^*(F) = \int_\alpha^F Fp(\alpha)d\alpha + \int_F^\infty \alpha p(\alpha)d\alpha = E[\alpha] + \hat{P}(F), \tag{3}
\]

where \( \hat{P}(F) = \int_\alpha^F P(\alpha)d\alpha \) is the integral of the cdf of \( \alpha \).

**Federal policymaking.** Anticipating the state policies in the second stage, each state seeks to maximize

\[
u_i(F; \alpha_i, \beta, P(\cdot)) = \begin{cases} \frac{\alpha^2_i}{2} + \alpha_i \beta (E[\alpha] + \hat{P}(F)) & \text{if } F < \alpha_i \\ \alpha_i F + \alpha_i \beta (E[\alpha] + \hat{P}(F)) - \frac{F^2}{2} & \text{otherwise.} \end{cases} \tag{4}
\]

The first line represents the payoffs to states that will perceive the federal policy to be too low, and implement a positive offsetting state policy. This portion of the state’s utility is increasing and strictly convex in \( F \). The second line represents payoffs to states that have been fully crowded out. The two lines of (4) are equal at \( F = \alpha_i \), implying that the state’s induced utility is continuous. A state \( i \)'s preferences can be either single- or double-peaked. Lemma A.1 provides necessary and sufficient conditions for single-peakedness.

Our first proposition describes several key features of a state’s induced preferences over

\(^{8}\)Alternatively, one may interpret the model as averaging over multiple policies, only a subset of which have these complementarities and the attendant flypaper effects.
federal policy. Let \( \hat{F}_i \equiv \hat{F}(\alpha_i, \beta, P(\cdot)) \) denote the most-preferred federal policy of state \( i \). Then:

**Proposition 1 (Induced Preferences over Federal Policy)**

1. A state’s ideal federal policy, \( \hat{F}_i \), is

   (a) strictly greater than \( \alpha_i \) for all \( \alpha_i > \bar{\alpha} \); and

   (b) strictly increasing in the state’s preference parameter, \( \alpha_i \) and the scale of cross-state spillovers, \( \beta \) if and only if \( \hat{F}_i \) is interior.

2. State \( i \)'s disutility from a feasible marginal decrease in federal policy away from \( \hat{F}_i - \Delta \) is lower than its disutility from a feasible marginal increase in federal policy away from \( \hat{F}_i + \Delta \) for \( \Delta \) sufficiently large.

(All proofs of formal results appear in the Appendix.)

Part 1(a) of the proposition establishes that a state would prefer a federal policy higher than the one it would impose upon itself in the absence of a binding federal mandate. In essence, this reflects the desirability of a centralized solution to the collective action problem induced by the cross-state externalities. To understand the logic, note that an increase in any federal policy less than \( \alpha_i \) is all upside for state \( i \), increasing the positive spillovers flowing from other states while leaving the effective policy in state \( i \) (and its cost) unaffected. By contrast, a federal policy higher than \( \alpha_i \) is binding on effective policy in the state, so incremental increases in that policy come with costs as well as benefits. On the margin, benefits outweigh costs at \( F = \alpha_i \), and so the state would prefer still larger federal policies due to the accompanying spillovers. Part 1(b) is intuitive: given the equilibrium level of the state’s own policy, the benefits to a more extensive federal policy are strictly increasing in \( \alpha_i \) and \( \beta \), so an increase in either will yield an increase in \( \hat{F}_i \).

Part 2 of the proposition suggests a fundamental asymmetry in the shape of a state’s induced preferences over \( F \). Recall from above that up to \( \alpha_i \), the state’s utility is increasing
and convex in $F$, because the state can take compensating action to “top up” in the presence of a federal policy perceived as too low. Above $\alpha_i$, states have no such recourse – they cannot “top down.” For sufficiently high values of $F$, a state’s utility is decreasing and concave, as the effects of additional increases in federal policy become increasingly onerous. This asymmetry will play a key role in several of the results that follow.

Figure 2 displays the relationship between state decision-making and preferences over federal policy when the latter are single-peaked and assuming a uniform $p(\alpha)$. The horizontal axis depicts the level of federal policy $F$. The thick black kinked line represents state $i$’s effective policy in equilibrium. When $F$ is low, the state implements its own policy sufficient to bring its effective policy to $\alpha_i$ (the flat portion). For higher values of $F$, the federal policy is the effective policy, as the state’s own policy has been fully crowded out (the diagonal portion). The dashed curve represents the state’s induced utility over federal policy when the state can implement its own policy to compensate for what it perceives as an inadequate federal policy. This corresponds to the first line of (4); for reference we have included a portion of the curve above $\alpha_i$ (which would correspond to infeasible negative values of $S_i$). The thin black solid curve corresponds to the second line of (4): induced utility over federal policy when $S_i$ is constrained to zero. Overall, the state’s induced utility over $F$ – shaded gray – is defined by the dashed curve below $\alpha_i$ and the solid curve above it. As the figure makes clear, the asymmetry around $\hat{F}_i$ is the direct consequence of the federal structure of decision-making, and not the underlying functional form of the utility function, which, given uniform $p(\alpha)$, is itself symmetric.

National Politics in the Federal System

Having characterized induced preferences over federal policy for individual states, we proceed to a discussion of the political implications of these preferences in different institutional settings.

The Normative Benchmark. We begin by discussing features of the federal policy that maximizes states’ aggregate welfare (given anticipated state-level optimal adjustments) at
Effective policy (the thick black kinked line) is constrained by sufficiently high federal floors \((F > \alpha_i)\). This constraint generates an asymmetry in the state’s induced utility over federal policy (the shaded gray curve).
different levels of preference heterogeneity. Note that preference heterogeneity could be modeled with two related but distinct features of the distribution of state preferences: the range of the support, $[\alpha, \overline{\alpha}]$, with broader supports corresponding, ceteris paribus, to more heterogeneous preferences; and the shape of the distribution holding the support constant, with “flatter” (unimodal, given log-concavity) distributions corresponding to greater preference heterogeneity. To adopt a general notion of the relative heterogeneity of two distributions that captures both of these notions while minimizing the need for parametric assumptions, we use the language of Hopkins and Kornienko (2004), and say that distribution $P_A(\alpha)$ dominates $P_B(\alpha)$ in the unimodal likelihood ratio (ULR) order ($P_A(\cdot) \succ_{ULR} P_B(\cdot)$) if and only if the ratio of the associated densities is strictly unimodal and $E_A[\alpha] \geq E_B[\alpha]$. ULR-dominance of distribution $A$ over $B$ implies second order stochastic dominance; further, if $E_A[\alpha] = E_B[\alpha]$, then ULR-dominance implies that $B$ is a mean-preserving spread of $A$ (and, thus, that $B$ will have higher variance and mean absolute deviation than $A$). We have the following result:

**Proposition 2 (Optimal Federal Policy)** There exists a unique federal policy, $F^*(\beta, P(\cdot)) \in (\alpha, \overline{\alpha}]$, that maximizes the aggregate welfare of the states. Further:

1. $F^*$ is weakly increasing in the importance of externalities $\beta$ and average demand $E[\alpha]$, and strictly increasing if and only if $F^* < \overline{\alpha}$.

2. Let $P_A(\cdot) \succ_{ULR} P_B(\cdot)$. Then $F^*(\beta, P_B(\cdot))$ is weakly lower than $F^*(\beta, P_A(\cdot))$, and strictly lower if and only if $F^*(\beta, P_B(\cdot)) < \overline{\alpha}$.

That the optimal federal policy is increasing in the importance of externalities is unsurprising: as externalities increase, the underprovision associated with exclusively state-level policymaking becomes increasingly acute, warranting a more substantial federal presence. Likewise, it is unsurprising that as the typical demand for public policy increases, the optimal policy increases correspondingly. The more noteworthy part of the result concerns the relationship between the optimal federal policy and the extent of preference heterogeneity.
within the federation. The result emerges because of the asymmetry in states’ induced preferences over federal policy that exists when $F < \bar{\alpha}$. Holding fixed the federal policy, when the distribution of types becomes more heterogeneous, an increasing fraction of low-demand states are fully crowded out; this raises the marginal benefit of a reduction in the federal policy relative to the marginal cost borne by high-demand states, which are in a position to accommodate the reduction via compensating state policy.\(^9\)

Note that when $F^* > \bar{\alpha}$, no state wishes to top up (i.e., $S_i = 0$ for all $i$). It is straightforward to demonstrate that in this case, which is equivalent to a fully centralized political system, the welfare maximizing policy is equal to the mean state’s ideal point, $E[\alpha](1 + \beta)$. While this policy would respond predictably to the magnitude of $\beta$ and $E[\alpha]$, it would be unresponsive to changes in heterogeneity. In contrast, in the presence of compensatory federalism ($F^* < \bar{\alpha}$), heterogeneity matters, because of the asymmetries in preferences that are absent when all states are fully crowded out.

**Some Properties of Federal Policies and Institutions.** The next step in our analysis is to examine the relationship between national policymaking institutions and departures from the normative benchmark policy. In conducting this analysis, it is helpful to distinguish two types of failure or inefficiency that may result from collective choice. The first occurs when national policymaking fails to take into account the asymmetric burdens of high- and low-demanders. The second is associated with supermajoritarian institutions in particular, and stems from the inability to realize gains from trade when a non-empty interval of policies is gridlocked. This second type of inefficiency, which could arise even if the majoritarian policy were itself efficient, may either mitigate or exacerbate the first type. Given the asymmetry of burdens described above, however, the inefficiency associated with gridlock at high and low levels of federal policy may differ considerably.

To examine the first kind of failure, we examine policymaking under simple majority rule with an open agenda. The following lemma establishes a majority rule voting equilibrium:

\(^9\)It is immediate that if $P_A(\cdot)$ first order stochastically dominates $P_B(\cdot)$, then $F^*(\beta, P_A(\cdot)) \geq F^*(\beta, P_B(\cdot))$, with the inequality holding strictly if $F^*(\beta, P_B(\cdot)) < \bar{Z}$.\(^{16}\)
Lemma 1 (Majority Voting Equilibrium) States’ induced preferences over federal policies are single-crossing; thus, under the simple majority rule, the equilibrium federal policy is $\hat{F}_m$, the most preferred federal policy of the state with the median preference parameter, $\alpha = \alpha_m$.

How does the federal policy arrived at via majority rule, $\hat{F}_m$, compare to the normative benchmark, $F^*$? This question is particularly relevant given the point of comparison with the centralized polity discussed above: if the density of types, $p(\cdot)$, is symmetric, so that the median and mean coincide, the two policies are the same under full centralization. Thus, to isolate the effect of the states’ strategic incentives brought about by compensatory federalism on any divergence between these two policies, we will assume for the next result that the distribution of types is symmetric, with the median/mean state’s type given by $\alpha_m$.

Proposition 3 (Federal Policymaking under Simple Majority Rule) Suppose $p(\cdot)$ is symmetric. Then the federal policy arrived at under simple majority rule is weakly higher than the normative benchmark, and strictly higher if and only if $\hat{F}_m < \overline{\alpha}$.

The basic intuition underlying this result stems from the asymmetry of burdens associated with federal policies. Suppose the median state’s ideal policy leaves room for topping up among high-demanders. In that situation, majority rule weights the benefits of internalizing positive spillovers higher than the disutility of crowded-out low-demanders. However, this weighting is not consistent with aggregate welfare maximization for any $\beta$ such that some states continue to top up. By contrast, when the median state’s policy crowds out all states, the system is effectively completely centralized, and (as described above), the median state’s ideal point and the aggregate welfare maximizing policies coincide.

Given the foregoing, it is natural to consider when features of the political environment moderate or exacerbate the extent of the discrepancy between the majoritarian policy and the normative benchmark. The next remark addresses this question under additional conditions on the distribution of the $\alpha$s.
Remark 2 (Relative Inefficiency Under Uniform Distribution of Types) Suppose that \( \alpha \) is distributed uniformly on the interval from \( \alpha_m - R \) to \( \alpha_m + R \), and that \( \hat{F}_m < \alpha_m + R \). Then the disparity between the majority rule equilibrium policy and the aggregate welfare-maximizing policy, \( \hat{F}_m - F^* \), is increasing in \( R \).

This remark implies that if the majoritarian policy admits topping up by at least some states, an increase in preference heterogeneity will increase the discrepancy between the majoritarian and normative benchmark policies. While an increase in heterogeneity decreases both policies, it does so faster for the socially efficient one, as the median state’s preferred policy does not fully reflect the asymmetric burdens induced by the federal structure. Again, we note a contrast with the institutions of full centralization: given a symmetric distribution of tastes, the difference between the majoritarian and normative benchmark policies would be zero irrespective of the degree of preference heterogeneity.

Next, we shift to the second variety of inefficiency described above: political failures arising as a consequence of policy gridlock. Particularly in the United States, the existence of interstate preference heterogeneity serves as a primary justification for both federalism and supermajoritarian national institutions, which tend to perpetuate the status quo when policy change does not enjoy broad-based support. Our focus, then, is on inefficiencies that arise from the interaction of these two constitutional features.

Methodologically, we proceed as follows. First, note that if preferences over federal policy are single-peaked, a supermajoritarian federal bargaining protocol \( B \) will induce a nonempty gridlock interval (Krehbiel 1996, 1998).\(^{10}\) This implies, in turn, the potential for a gap to exist between the equilibrium federal policy arrived at under federal bargaining protocol and both the ideal policy of the median state (which, recall, would be decisive under majority rule) and the normative benchmark. Rather than analyzing specific bargaining protocols, we focus instead on the consequences of different federal policies, with the understanding

\(^{10}\)Lemma A.1 shows that violations of single-peakedness occur only under special situations.
that such policies would emerge as the outcome of a mapping into the equilibrium federal policy from the conjunction of a status quo policy, preference profile, and some bargaining protocol.\textsuperscript{11}

Second, the potential gap between the equilibrium and aggregate welfare-maximizing federal policies suggests the possibility of inter-state transfers, which can be thought to represent compromises that states may be willing to make on other policy dimensions to help break the gridlock on federal policy. While it is straightforward that, in a Nash Bargaining framework with transfers, the states will be able to reach a socially efficient bargaining agreement that realizes the social welfare maximum, there are strong reasons to believe that such transfers will frequently be politically infeasible.\textsuperscript{12} This points to the value of taking the gridlock induced by supermajoritarian national institutions as given, and exploring in that context the consequences of the federalism-induced asymmetric preferences described above for national politics and, in particular, the social costs of gridlock.

Our measure of \textit{gridlock inefficiency} (GI) is simply the absolute value of states’ aggregate marginal utility. The intuition is straightforward: large GI suggests that an incremental shift in the direction of the efficient common policy would help the beneficiaries of that shift much more than it would hurt the losers, suggesting large unrealized gains from trade associated with gridlock. When GI is zero, the aggregate welfare of states is maximized and there would be no room for mutually beneficial trades even if they were feasible.

\textsuperscript{11}For any Pareto efficient federal policy, there exists a triple of status quo policy, preference profile, and bargaining protocol for which the policy is an equilibrium.

\textsuperscript{12}Most critically, the agreements they imply are radical idealizations that in the messy politics outside our model run into severe political transaction costs. Acemoglu (2003) and Bednar (2009) point to two closely related sources of such costs that are particularly relevant to our substantive setting: the commitment problems due to the incentive to renege on the part of the interests that control the government at a given time, and opportunistic burden-shifting and shirking by some states at the expense of others. Further, a natural interpretation of transfers that states could make to each other is as policy concessions trading off gains and losses across different policy dimensions. This would, of course, entail the policy space being non-trivially multidimensional. But, with limited exceptions, policy conflict at the national level in the U.S. has been (consistent with our model) largely unidimensional, particularly since the early 1980s (Poole and Rosenthal 2000).
Our next result points to an asymmetry in GI between policies gridlocked above the optimal policy, \( F^* \), and those gridlocked below that policy.

**Proposition 4 (Gridlock Inefficiency)** If \( p(\alpha) \) is sufficiently dispersed, then gridlock inefficiency at \( F = F^* + \Delta \) is strictly greater than at \( F = F^* - \Delta \) for \( \Delta > 0 \).

Informally, Proposition 4 says that, in sufficiently heterogeneous federations, federal policies gridlocked at a high level suffer, on the margin, from more unrealized gains from trade than federal policies gridlocked at a low level. A federal policy gridlocked at a low level may be suboptimal from the perspective of the high-demander state; however, that state can partially compensate for the deficiency with its own within-state policy. Because of crowding out, low-demander states have no such recourse when the policy is gridlocked at a high level.

In the presence of compensatory federalism, the existence of these two types of inefficiencies provides clues regarding the relationship between the status quo policy and the relative merits of different national policymaking institutions. Suppose the federal bargaining protocol yields a gridlock interval \([\hat{F}_l, \hat{F}_h]\) and that the majoritarian policy would (1) fall within the gridlock interval and (2) admit topping up by some states, so \( F^* < \hat{F}_m \). Then depending on the locations of these critical values, situations may emerge in which either the super-majoritarian bargaining protocol welfare-dominates a purely majoritarian one, or in which the majoritarian one dominates. Figure 3 delineates these conditions for the case in which \( \hat{F}_l < F^* < \hat{F}_m < \hat{F}_h \).

First, no policy gridlocked above \( \hat{F}_m \) dominates the majoritarian policy, because given \( F^* < \hat{F}_m \), the majoritarian policy occupies a point on the downhill portion of the aggregate welfare function. Second, policies gridlocked under a supermajoritarian bargaining protocol dominate the majoritarian equilibrium policy only for a (nonempty) subset of those policies on the interval from \( \hat{F}_l \) to \( \hat{F}_m \). The figure illustrates two cases. For the value \( \hat{F}_l' \), that subset is proper, and there are some low gridlocked policies dominated by the majoritarian equilibrium policy. For the value \( \hat{F}_m'' \), however, the entire range of gridlocked policies below \( \hat{F}_m \) dominates the majoritarian policy (in addition to some ungridlocked federal policies). It
Figure 3: Welfare Properties of Majoritarian and Supermajoritarian Federal Bargaining Protocols

The figure shows the range of federal policies (shaded gray) that welfare-dominate the policy arrived at by simple majority rule.
is immediate that an identical logic to this case holds when $F^* < \hat{F}_l$. The following remark summarizes the foregoing:

**Remark 3** Suppose $\hat{F}_m \in (\hat{F}_l, \hat{F}_h)$, and that $\hat{F}_m < \bar{F}$. Then, there exists a nonempty subset of federal policies contained in $[\hat{F}_l, \hat{F}_m]$ s.t., for a given status quo policy, a supermajoritarian bargaining protocol welfare-dominates the majoritarian one if and only if it induces an equilibrium outcome in that subset.

**Gridlocked Policies and the Potential for National Conflict.** We next explore how the asymmetries brought about by compensatory federalism affect the extent to which a gridlocked policy $F$ would be likely to provoke greater resistance toward the policy or, alternatively, a greater investment in effort to change it. What the “right” model of explicit conflict over policy is unclear, and any attempt at such a model would require imposing substantial additional structure, which we prefer to avoid here in the interests of parsimony. With that in mind, we focus our analysis on a measure of resistance potential (RP) that has the following intuition. Conflict over federal policy is embedded in a broader (unmodeled) political setting that incorporates electoral competition, advocacy, legal challenges, etc. A political actor with a stake in the federal policy might be motivated to pursue policy change through these channels, with an intensity proportional to the stakes involved. Any policy $F$ creates natural coalitions of high-demanders favoring higher $F$ and low-demanders favoring lower $F$. We wish to construct a measure that captures the extent of dissatisfaction among political “losers” – informally, the high-demanders when $F$ is relatively low, and the low-demanders when it is relatively high. We will then be in a position to assess how resistance to a federal policy varies with that policy, and the effect of the federal structure on that relationship.

The absolute value of the marginal utility of a state provides an index of its willingness to invest in an incremental shift in $F$: $\frac{\partial u_i(\alpha)}{\partial F}$ for high-demanders states and $-\frac{\partial u_i(\alpha)}{\partial F}$ for low-demanders. Let $RP^+(F) \equiv \int_{A(F)} \frac{\partial u_i(\alpha)}{\partial \alpha} p(\alpha)d\alpha$ be the resistance potential of high-demanders states given policy $F$, where $A(F)$ denotes the $\alpha < F$ whose corresponding ideal point
is \( F \); and \( RP^-(F) \equiv -\int_{\alpha} A(\alpha) \frac{\partial u(\alpha)}{\partial \alpha} p(\alpha) d\alpha \) the resistance potential of low-demander states given policy \( F \). For relatively low \( F \), the resistance potential of high-demanders will be larger than that for low-demanders \( (RP^+(F) > RP^-(F)) \): in essence, the high-demanders are the bigger “losers”. By contrast, for high levels of \( F \), the low-demanders are more aggrieved \( (RP^-(F) > RP^+(F)) \). We then define overall resistance potential given policy \( F \) as \( RP(F) \equiv \max\{RP^+(F), RP^-(F)\} \).

Our next result considers how \( RP(F) \) varies with a (potentially gridlocked) federal policy. Given our notation, let \( \hat{F}(\alpha) \) denote the most preferred federal policy of a state with the lowest feasible demand for policy, and \( \tilde{F}(\overline{\alpha}) \) the most preferred federal policy of a state with the highest feasible demand. Then:

**Proposition 5 (Gridlock and Resistance Potential)**  
1. \( RP^-(F) \) is strictly increasing in \( F \), and \( RP^+(F) \) is strictly increasing in \( F \) for sufficiently low values of \( F \).

2. There exists a pair of federal policies \( \{F, \overline{F}\} \), with \( F > \hat{F}(\alpha) \) and \( \overline{F} < \tilde{F}(\overline{\alpha}) \), such that for all policies \( F' < F \) and policies \( F'' > \overline{F} \), \( RP(F') < RP(F'') \).

To understand the first part of the result, note that low-demanders, by definition, have no recourse to compensating state action. As \( F \) increases, the size of the coalition of fully-crowded out low-demanders increases, with the lowest-demand states particularly hard hit and thus most willing to invest in incremental decreases in the policy. Both the size and intensity effects point in the same direction. The resistance potential of high-demanders is more nuanced. At low values of federal policy, many states are in the high-demand coalition, but the marginal effect of an increase in cross-state spillovers is small; hence, high-demanders have little incentive to invest in increases to the policy. As policy increases, the marginal benefit of increasing the externalities increases, and so resistance potential increases concurrently. At high values of the policy, the reduction in the size of the high-demand coalition outweighs the spillover-related benefits of further increases to those remaining in the coalition, driving down their resistance potential. The left panel of Figure 4 displays this
intuition graphically.

The core intuition for the second part of the result again comes down to the asymmetries in states’ induced utilities. As noted above, when the policy is high, the resistance potential is induced by the marginal utilities of low-demand states, and when the policy is low, by those of the high-demand states. However, the low-demand states are not in a position to top down the federal policy and so cannot mitigate the hit to their utility from the policy that is, for them, too high. Consequently, when the federal policy is sufficiently high, the value to low-demand states of avoiding a further increase or of bringing about a decrease is high as well—higher than the value to the high-demand states of avoiding a drop or of bringing about a marginal increase in federal policy when that policy is low.

To isolate the effect of federal institutional structure, the right panel of Figure 4 displays resistance potential under centralization ($S_i = 0$ for all states). At very high levels of federal policy under compensatory federalism, the preponderance of states will be fully crowded out under compensatory federalism, and so resistance potential under the two systems (left and right panels) are comparable. At low levels of policy, however, the pictures diverge sharply. Under a unitary system, high-demanders are as willing to fight to increase a low policy as low-demanders are willing to fight to increase a high one. By contrast, under the compensatory federal structure the high-demanders have considerably less incentive to contest an insufficiently extensive federal policy.

**The Shadow of the Future.** In the previous section, we demonstrated how changes in the status quo policy can yield variation in the potential for political conflict even holding the preference profile, and by extension political polarization, constant. In this section, we show how the shadow of the future can directly increase polarization, even holding the preference profile constant, with asymmetric implications for political conflict that echo those described above. With this in mind, we consider a two-period extension of our model with a preference shock occurring between the two periods, and analyze the effect of the entailed shadow of the future on the width and position of the gridlock interval induced by a supermajoritarian
Figure 4: Resistance Potential for Different Levels of Federal Policy under Compensatory Federalism and Centralized Government

The left panel shows how resistance potential varies with federal policy in a federal system, while the right panel shows this relationship in a centralized one. Resistance potential (the shaded gray curve) is the maximum of the Resistance Potential of High-Demanders ($R_P^+$, the solid black line) and of Low-Demanders ($R_P^-$, the dashed black line).
bargaining protocol.

The basic intuition for the incentives in the dynamic setting is threefold. First, and most obviously, if a state anticipates that future shocks may make a candidate federal policy less attractive to it, then it is less likely to accede to that policy in the current period. Second, anticipating the possibility of future opposition from other states to a federal policy it may turn out to favor even more than at present, a state should be more willing to incur a current-period hit to give itself a greater opportunity to implement that policy in the future. For states with low demand for the policy, this means preferring an even lower federal policy than they would prefer in the one-shot environment, because in the future, they may be pivotal in adjusting it upward if necessary. For high-demand states, it means preferring a still higher federal policy because they are pivotal in adjusting it downward, if necessary. The third part of the intuition relates more closely to the specific details of the politics of federalism. Because the low- and high-demand states are asymmetric in their abilities to adjust their state-level provision, they may respond asymmetrically to the possibility of future shocks.

To capture these intuitions in a simple way, suppose that at the beginning of the second period, all states’ $\alpha$ parameters are shocked by some value $\sigma$ symmetrically distributed around 0, with pdf $g(\sigma)$. Thus, state $i$’s second-period demand, $\alpha^2_i$, is equal to $\alpha_i + \sigma$. We will show that, in expectation of these shocks and equilibrium behavior, the existence of the second period leads the right bound of the first-period gridlock interval to move to the right, and the left bound to move to the left, and, when externalities $\beta$ are not too high, by a still greater amount, asymmetrically stretching the gridlock interval in the first period relative to the baseline one-period game. Put differently, the weight of the future will increase the present-day disagreement and will do so by making the low-demand states disproportionately less willing to compromise.

---

13 Part 1 of Proposition 6, which captures these two intuitions, is in the spirit of the recent work by Dziuda and Loeper (2016), of which we became aware after independently deriving our result. Dziuda and Loeper show that the logic underlying these intuitions persists in a number of different environments, underscoring their broader applicability.
First, note that in each period, each state $i$ will choose its own state level of provision optimally, dependent only on the current period’s values of $\alpha_i^t$ and $F^t$. Let $\hat{F}^2(\alpha_i^2)$ be $i$’s optimal federal policy in $t = 2$ anticipating optimal second-period state-level provision $S^2(\alpha_i + \sigma, F^2)$; and let $\hat{F}^1(\alpha_i)$ be $i$’s optimal federal policy in $t = 1$, anticipating optimal subsequent play.

Recall that $\alpha_L$ and $\alpha_H$ are types whose optimal federal policies define the lower and upper bounds of the gridlock interval, respectively. For mathematical convenience, we will assume that if the inherited status quo is below (above) the induced optimal preference of the state defining the left (right) bound of the second period’s gridlock interval, then the policy will be pulled up (down) to that lower (upper) bound. If the second period’s inherited status quo $F^1$ is in the second period’s gridlock interval, then it will persist as the federal policy in that period. Let $\hat{F}(\alpha_i)$ (without superscript) refer to $i$’s ideal level of federal provision in the one-shot environment. Our result in this section is the following proposition:

**Proposition 6** The following properties describe the relationship between the gridlock intervals in the first and the second periods of the two-period environment:

1. In the first period, the left boundary of the gridlock interval is lower, and the right boundary higher, than their respective boundaries in the one-shot game (i.e., $\hat{F}^1(\alpha_L) < \hat{F}(\alpha_L)$ and $\hat{F}^1(\alpha_H) > \hat{F}(\alpha_H)$); and

2. If externalities $\beta$ are not too great, then the left boundary moves left farther than the right boundary moves right (i.e., $|\hat{F}^1(\alpha_L) - \hat{F}(\alpha_L)| \geq |\hat{F}^1(\alpha_H) - \hat{F}(\alpha_H)|$).

The condition on $\beta$ in part 2 emerges because when the importance of across-states externalities is very high, the incentive for states with relatively low demand to impose a higher federal policy on states with still lower demand dominates the incentive to reduce the federal policy as insurance against future shocks. When externalities are not too great, the low-demand states’ lower flexibility in the state-level policy adjustment in the event of a negative demand shock will lead them to become relatively more extreme than the high-demand states in their induced preferred federal policy in expectation of demand shocks.
Discussion

Our model of compensatory federalism is deliberately parsimonious, omitting key political complexities to focus on the functioning and implications of a specific causal mechanism. That being said, it permits a straightforward broader framing in a more comprehensive political environment in which a number of its empirical implications can be readily evaluated. In this section, we briefly discuss several widely recognized empirical observations about U.S. politics and policymaking illuminated by our analysis.

**A Paradox of Red and Blue States.** First, consider a pattern initially pointed out by Lacy (for a recent restatement, see Lacy 2014): citizens of states that are net beneficiaries of federal largesse (relative to taxes) tend to vote conservatively, a pattern that has led some critics to decry hypocritical “red state moochers” (e.g., Pearlstein 2016). Though our model permits only a uniform federal policy, a simple extension would account for this pattern. Suppose instead of setting $F$ constant across states, we allowed it to vary, while holding the cost constant across states. (Equivalently, we could hold policy constant and vary costs.) Even a low-demand state that is a net revenue recipient in a federal policy may, nonetheless, oppose it because the return on its contribution is lower than what it could obtain on a different expenditure (including private consumption) that has been crowded out by the high federal policy. Of course, senators and representatives from states with low overall demand for a policy might support a deal that provided their states with sufficiently disproportionally high benefits. In either case, the result would be the ostensible paradox.

**Constraints on Tiebout Sorting.** In his model of interjurisdictional sorting, Tiebout (1956) demonstrates that in the absence of moving costs and when public goods are provided locally, citizens with heterogeneous levels of demand for government services will cluster according to their demand for those services. The empirical record on residential sorting is mixed, however: Rohde and Strumpf (2003), for example, show that despite substantial decreases in mobility costs since the second half of the 19th century, interjurisdictional policy heterogeneity has declined. Mummolo and Nall (2017), in developing a test of homophily-
based (rather than Tiebout) sorting, find that individuals do not tend to migrate to politically like-minded communities. And a straightforward reading of the data described in Levendusky and Pope (2010) suggests that intrastate ideological heterogeneity significantly exceeds interstate heterogeneity.

Our model of compensatory federalism squares these findings. Over time, as the scope of federal programs has increased, the extent of interjurisdictional heterogeneity with respect to effective policy has declined. This reduces the incentive to migrate in search of more favorable policies, which may further constrain policy heterogeneity.

Asymmetric and Encompassing Polarization. One of the most robust findings in political science is the increase in partisan polarization among elites since the early 1980s. Recently, scholars have concluded that the increasing gap between Democrats and Republicans is due primarily to a pronounced conservative shift among Republicans (Barber and McCarty 2015; Mann and Ornstein 2012), which cannot be wholly accounted for by the transition of southern conservatives to the Republican Party. Asymmetric polarization is consistent with our findings concerning resistance potential, as well as our result on dynamic polarization: specifically, in an environment of compensatory federalism, a large federal presence is likely to activate a high degree of opposition from conservatives relative to the degree of opposition from liberals of a small presence.

A related point emerges from the comparison of contemporary political polarization with polarization during the Progressive Era (1900-1916). Although standard measures suggest comparably high degrees of legislative polarization in both periods (McCarty, Poole, and Rosenthal 2006), the incidence of successful bipartisan lawmaking could not be more different. The earlier period is remembered for landmark legislation expanding the national administrative state, much of which passed with large bipartisan majorities. By contrast, the contemporary era has seen few (if any) major legislative accomplishments passing in the absence of strenuous minority party opposition. If we accept polarization as a sufficient statistic for political discord, this difference in consequences is puzzling. Our analysis of
resistance potential suggests an alternative measure of discord and points to a resolution of the puzzle. Larger resistance to marginal changes from limited federal policies than from extensive ones accounts for the difference in the policymaking in the two eras and opens a new avenue for future empirical work on political conflict that leverages our measure of discord.

Finally, our model suggests the following implication. Suppose state-level policy in one domain is fully crowded out. As noted above, in such a situation an increase in the federal policy will produce greater opposition from that state. But a second (not explicitly modeled) effect of such an increase may be to compel the state to reallocate resources away from other policy areas to meet the commitments entailed by the first one. This will, in turn, increase state-level resistance to a federal presence in those other policy areas. The overall effect, then, is to potentially exacerbate conflict across multiple policy dimensions. Interpreted in a broader political context, we should therefore expect political conflict and gridlock in the federal system to be encompassing phenomena, comprehensively paralyzing national policymaking across issue domains.

Conclusion

Our analysis has focused on an important feature of many federal systems: just as national policies can compensate for inadequate state-level action, compensating actions by states can mitigate perceived deficiencies in national-level policies. Most importantly, we demonstrate that due to the structure of federalism, there is a fundamental asymmetry in the ability of states with different levels of demand for a policy to engage in some kind of “self-help” to mitigate what they perceive are the adverse consequences of the national one. One way to think of our paper is in comparison with the canonical argument that decentralization of policymaking responsibilities to subnational governments is particularly valuable in societies with heterogeneous preferences across jurisdictions (cf., Oates’ [1972] “decentralization theorem”). While that argument concerns economic efficiency, it is silent on distributive effects and what might be thought of as political efficiency, the focus of this paper.
While one might be tempted to view our model as a critique of “big government” at the national level, this would be too simplistic a view. Rather, our analysis illuminates a mechanism wherein the political institutions of federalism mediate the relationship between “big” national government and political dysfunction. In doing so, it uncovers critical biases in the operation of federal systems. One of these biases is a particularly favorable status for states with a high demand for public policy: the asymmetric effect of crowding out implies that states with low-demand for a particular policy are more constrained in their ability to compensate for undesirably high levels of federal provision than high-demand states are able to compensate for undesirably low levels. Another bias is in the particularly reactionary behavior of states with low demand: while (the fear of future) gridlock at an undesirable level of federal provision tends to make conservative/low-demand states more conservative and liberal/high demand states more liberal, (the fear of future) state-level crowding out tends to make the conservative ones particularly recalcitrant.

This second bias also relates to an important element of federal politics that we have not explicitly addressed in our analysis, but that certainly must be a significant feature of a general theory of federalism as well as of the account of federalism in American political history: preference heterogeneity and the quality of representation within states. While not addressing these concerns explicitly is a limitation of our analysis, it does, nonetheless, also suggest a path for their future consideration. The variation in resistance potential over the range of national policies arguably tracks the incentives on the part of the powerful intra-state actors to suppress voting rights of groups whose votes could shift the policies within the state and the state’s induced/expressed preference over national policies, as well as the incentives of actors from other states to champion the voting rights of those groups with robust national-level protections. Analyzing the net effects of those incentives in detail, however, requires additional structure, including on the technology of conflict between state-level interests, that we have sought to avoid imposing in the present paper. We leave that analysis for future work.
References


O’Reilly, James T. 2006. *Federal Preemption of State and Local Law: Legislation, Regu-

Pearlstein, Steven. 2016. “Under Trump, Red States are Finally Going to be Able to Turn Themselves into Poor, Unhealthy Paradises.” Blog Post on *themonkeycage.org*, December 4, 2016.


Vaghul, Kavya, and Ben Zipperer. 2016. “Historical State and Sub-State Minimum


A Supplemental Appendix for "The Political Economy of Compensatory Federalism"

Derivation of $\Psi^*(\cdot)$

From the text,

$$\Psi^*(F) = \int_\alpha^F Fp(\alpha)d\alpha + \int_F^\bar{\alpha} \alpha p(\alpha)d\alpha$$

$$= FP(F) + \int_F^\bar{\alpha} \alpha p(\alpha)d\alpha$$

(A.1)

Integrating the second expression in the second line of (A.1) by parts,

$$\int_F^\bar{\alpha} \alpha p(\alpha)d\alpha = \left[ \alpha P(\alpha) - \int P(\alpha)d\alpha \right]_{\bar{\alpha}}^F$$

$$= \bar{\alpha}P(\bar{\alpha}) - \hat{P}(\bar{\alpha}) - FP(F) + \hat{P}(F),$$

where $\hat{P}(x) = \int_\alpha^x P(\alpha)d\alpha$, i.e., the integral of the cdf. Substituting into (A.1) and noting that $P(\bar{\alpha}) = 1$, we have

$$\Psi^*(F) = \bar{\alpha} - \hat{P}(\bar{\alpha}) + \hat{P}(F).$$

(A.2)

By definition, $E[\alpha] = \int_\alpha^{\bar{\alpha}} \alpha p(\alpha)d\alpha$. Integrating by parts as above gives

$$\int_\alpha^{\bar{\alpha}} \alpha p(\alpha)d\alpha = \bar{\alpha}P(\bar{\alpha}) - \hat{P}(\bar{\alpha}) - \alpha P(\alpha) + \hat{P}(\alpha).$$

Noting that $P(\bar{\alpha}) = 1$, and $P(\alpha) = \hat{P}(\alpha) = 0$, we have $E[\alpha] = \bar{\alpha} - \hat{P}(\bar{\alpha})$. Substituting into (A.2), $\Psi^*(F) = E[\alpha] + \hat{P}(F)$.

Lemma A.1 Let $\bar{\alpha}$ be the modal value of $\alpha_i$. State i’s preferences are single-peaked on $F \in [0, \bar{\alpha}]$ if and only if one of the following four conditions holds:

1. $\bar{\alpha}_i \geq \alpha_i$ and $p(F) \leq \frac{1}{\alpha_i^2}$ for all $F > \alpha_i$;
2. \( \tilde{\alpha}_i \geq \alpha_i \) and \( p(\alpha_i) \geq \frac{1}{\alpha_i \beta} \);

3. \( \tilde{\alpha} < \alpha_i \); or

4. Conditions (a) through (c) are violated, but \( Z < \tilde{F}_i \), where \( \tilde{F}_i \) is the value of \( F \) corresponding to a local minimum in \( u_i(F) \) for \( F > \alpha_i \).

Otherwise, state \( i \) has double-peaked preferences.

**Proof.** First, note that for \( F > \alpha_i \), \( \frac{\partial u_i}{\partial F} = \alpha_i + \alpha_i \beta p(F) - F \). At \( F = \alpha_i \), this quantity is equal to \( \alpha_i \beta p(\alpha_i) \), which is strictly positive. For sufficiently large \( F > \alpha_i \), this quantity is strictly negative. Next, observe that \( \frac{\partial^2 u_i}{\partial F^2} = \alpha_i \beta p(F) - 1 \). Rearranging, this quantity is negative if and only if

\[
p(F) < \frac{1}{\alpha_i \beta}.
\]

(Necessity). Violation of (a), (b), and (c) imply \( \tilde{\alpha} > \alpha_i \), \( p(\alpha_i) < \frac{1}{\alpha_i \beta} \), and \( p(\tilde{\alpha}) \geq \frac{1}{\alpha_i \beta} \); in this case, \( u_i(F) \) is first concave, then convex, then concave in \( F \) for \( F > \alpha_i \). Given \( \frac{\partial u_i}{\partial F} \big|_{F=\alpha_i} > 0 \) and \( \lim_{F \to \infty} \frac{\partial u_i}{\partial F} < 0 \), this implies double-peakedness if \( F \) is not constrained as it is in (d).

(Sufficiency).

1. If inequality (A.3) holds for all \( F > \alpha_i \), then \( u_i(F) \) is strictly concave in that range. Given \( \frac{\partial u_i}{\partial F} \big|_{F=\alpha_i} > 0 \) and \( \lim_{F \to \infty} \frac{\partial u_i}{\partial F} < 0 \), this implies single-peakedness.

2. Log-concavity of \( p(\alpha) \) implies unimodality. This condition therefore implies that \( u_i(F) \) is first convex, and then concave in \( F \) for \( F > \alpha_i \). Given \( \frac{\partial u_i}{\partial F} \big|_{F=\alpha_i} > 0 \) and \( \lim_{F \to \infty} \frac{\partial u_i}{\partial F} < 0 \), this implies single-peakedness.

3. If \( \tilde{\alpha} < \alpha_i \), then unimodality implies \( p(F) \) is either first convex and then concave in \( F \) for \( F > \alpha_i \), or concave for all \( F > \alpha_i \). In either case, given \( \frac{\partial u_i}{\partial F} \big|_{F=\alpha_i} > 0 \) and \( \lim_{F \to \infty} \frac{\partial u_i}{\partial F} < 0 \), this implies single-peakedness.
4. As noted above, when conditions (a) through (c) do not hold but (d) does, single-peakedness is established by construction.

\[\partial u_i(F; \alpha_i, \beta) = \begin{cases} 
\alpha_i \beta P(F) & \text{if } F < \alpha_i \\
\alpha_i - F + \alpha_i \beta P(F) & \text{otherwise.}
\end{cases} \quad \text{(A.4)}\]

(a) Note that \(\frac{\partial u_i}{\partial F} > 0\) for all \(F < \alpha_i\). Thus, \(\hat{F}_i\) solves the second line of (A.4). The second line of (A.4) evaluated at \(F = \alpha_i\) simplifies to \(\alpha_i \beta P(\alpha_i) > 0\). Thus \(\hat{F}_i > \alpha_i\).

(b) From the foregoing, \(\hat{F}_i\) is defined implicitly by equating the second line of (A.4) to zero. Differentiating \(i\)'s marginal utility with respect to \(\alpha_i\) yields \(\frac{\partial^2 u_i(\cdot; \cdot)}{\partial F \partial \alpha_i} = 1 + \beta P(F) > 0\), and with respect to \(\beta\) yields \(\frac{\partial^2 u_i(\cdot; \cdot)}{\partial F \partial \beta} = \alpha_i P(F) > 0\). Thus, \(\frac{\partial u_i}{\partial F}\) is strictly increasing in both \(\alpha_i\) and \(\beta\), implying \(\hat{F}_i\) is strictly increasing in both parameters (Edlin and Shannon 1998).

2. The statement in the proposition is equivalent to
\[\lim_{\Delta \to \infty} \left( \left. \frac{\partial u_i}{\partial F} \right|_{F = \hat{F}_i + \Delta} + \left. \frac{\partial u_i}{\partial F} \right|_{F = \hat{F}_i - \Delta} \right) > 0. \quad \text{(A.5)}\]
\(\frac{\partial u_i}{\partial F}\) evaluated at \(\hat{F}_i + \Delta\) is given by the second line of (A.4). For sufficiently large \(\Delta\), \(\frac{\partial u_i}{\partial F}\) evaluated at \(\hat{F}_i - \Delta\) is given by the first line of (A.4). Substituting, (A.5) is equivalent to
\[\lim_{\Delta \to \infty} \left( \alpha_i - \hat{F}_i - \Delta + \alpha_i \beta P(\hat{F}_i + \Delta) + \alpha_i \beta P(\hat{F}_i - \Delta) \right) > 0.\]
From the first order condition for an interior optimum, \(\hat{F}_i = \alpha_i + \alpha_i \beta P(\hat{F}_i)\). Substiti-
tution and rearranging yields

\[
\lim_{\Delta \to \infty} \left( \alpha_i \beta \left( P(\hat{F}_i - \Delta) + P(\hat{F}_i + \Delta) - P(\hat{F}_i) \right) - \Delta \right) < 0,
\]

which holds because the term multiplying \( \alpha_i \beta \) is bounded between zero and one.

\[\square\]

**Proof of Proposition 2**

Integrating (4) over \( p(\alpha) \), aggregate welfare is given by

\[
W \equiv \int_{\alpha}^{\alpha} \alpha \beta (E[\alpha] + \hat{P}(\alpha)) p(\alpha) d\alpha + \int_{\alpha}^{F} \left( \alpha F - \frac{F^2}{2} \right) p(\alpha) d\alpha + \int_{\alpha}^{\alpha} \frac{\alpha^2}{2} p(\alpha) d\alpha. 
\tag{A.6}
\]

Via the Leibniz integral rule, marginal aggregate welfare is

\[
\frac{\partial W}{\partial F} = (\beta E[\alpha] - \delta(F)) P(F), 
\tag{A.7}
\]

where where \( \delta(F) \equiv \frac{\dot{P}(F)}{P(F)} = F - E[\alpha | \alpha < F] \) is the mean advantage over inferiors function from reliability theory. From part 1 of Proposition 1, any \( F < \alpha \) is Pareto dominated. Lemma 1 of Bagnoli and Bergstrom (2005) shows for log-concave \( p(\cdot) \) that \( \delta(F) \) is strictly increasing in \( F \) (from zero at \( F = \alpha \)). Therefore \( F^\ast \) is unique and defined implicitly by the first order condition \( \delta(F^\ast) = \beta \) (or by the corner \( Z \) when \( \delta(Z) < \beta E[\alpha] \)).

1. Follows immediately from the first order condition \( \delta(F^\ast) = \beta E[\alpha] \) and \( \delta(\cdot) \) increasing.

2. Let \( F_A^\ast \) be the aggregate welfare maximizing policy under distribution \( P_A(\cdot) \) and correspondingly \( F_B^\ast \) under \( P_B(\cdot) \). From the first order condition and the definition of \( \delta(F) \), \( \beta E[\alpha] = F - E[\alpha | \alpha < F] \). From Lemmas 2 and 4 of Hopkins and Kornienko (2003), \( P_A(\cdot) \succ_{ULR} P_B(\cdot) \) implies \( E_{P_A(\cdot)}[\alpha | \alpha < F] > E_{P_B(\cdot)}[\alpha | \alpha < F] \). Therefore \( \delta_A(F) < \delta_B(F) \). Since both \( \delta_A(\cdot) \) and \( \delta_B(\cdot) \) are strictly increasing in \( F \), the latter must cross \( \beta E[\alpha] \) to the right of where the former does.
Proof of Lemma 1

Differentiating (4) with respect to $F$ and again with respect to $\alpha_i$ yields

\[
\frac{\partial^2 u_i(F; \alpha_i, \cdot)}{\partial F \partial \alpha_i} = \begin{cases} 
\beta P(F) & \text{if } F < \alpha_i \\
1 + \beta P(F) & \text{otherwise.}
\end{cases}
\]  

(A.8)

Both the first and second lines of (A.8) are strictly positive, implying increasing differences, which are sufficient for single-crossing. Given single-crossing preferences, a majority rule voting equilibrium exists, and the median state will be decisive (Gans and Smart 1996).

Proof of Proposition 3

From Proposition 2 and given $E[\alpha] = \alpha_m$, $\delta(F^*) = \beta \alpha_m$. Since $\delta(\cdot)$ is monotone increasing for log-concave densities, its inverse exists and is also monotone increasing. Therefore $F^* = \delta^{-1}(\beta \alpha_m)$, and $F^* < \hat{F}_m$ if and only if

\[
\delta(\hat{F}_m) > \beta \alpha_m.
\]  

(A.9)

Recalling that $\delta(F) = \frac{\hat{P}(F)}{P(F)}$, substituting into (A.9) and rearranging yields

\[
\hat{P}(\hat{F}_m) > \alpha_m \beta P(\hat{F}_m).
\]  

(A.10)

From Proposition 1, $\hat{F}_m$ is defined implicitly by the first order condition $\alpha_m \beta P(\hat{F}_m) = \hat{F}_m - \alpha_m$. Substituting into (A.10) yields the condition

\[
\hat{P}(\hat{F}_m) > \hat{F}_m - \alpha_m.
\]  

(A.11)

Note that at $\beta = 0$, $\hat{F}_m = \alpha_m$ and (A.11) holds trivially. Recall from the derivation of $\Psi^*(F)$ above (and given symmetry) that $\alpha_m = \bar{\alpha} - \hat{P}(\bar{\alpha})$, or $\hat{P}(\bar{\alpha}) = \bar{\alpha} - \alpha_m$. Also note that
for all $\hat{F}_m > \alpha$, $\frac{\partial \hat{P}(\hat{F}_m)}{\partial \hat{F}_m} = P(\hat{F}_m) = 1$. Therefore for all $\hat{F}_m \geq \alpha$, (A.11) holds at equality, which in turn implies $\hat{F}_m = F^*$. 

Next, assume $\hat{F}_m < \alpha$. Then the derivative of the left side of (A.11), $P(\hat{F}_m)$, is strictly less than one, while the derivative of the right side is equal to one. Suppose there exists some $\hat{F}_m' < \alpha$ such that (A.11) does not hold. Given convexity of $\hat{P}(\cdot)$, $\hat{P}(\alpha_m) > 0$, $\hat{F}_m \geq \alpha_m$, and $\hat{P}(\alpha) = \alpha - \alpha_m$, it must then be the case that there exists some $\hat{F}_m'' \in (\hat{F}_m', \alpha)$ such that $\frac{\partial \hat{P}}{\partial \hat{F}_m} \bigg|_{\hat{F}_m=\hat{F}_m''} > 1$, a contradiction. Therefore (A.11) holds for all $\hat{F}_m < \alpha$. ■

Proof of Remark 2

From the proof of Proposition 1, the first order condition for $m$’s ideal point is given by $\hat{F}_m = \alpha_m(1 + \beta P(\hat{F}_m))$. Substituting using the uniform cdf and solving yields

$$\hat{F}_m = \alpha_m \left(1 + \frac{\beta R}{2R - \beta \alpha_m}\right).$$

Substituting into the condition $\hat{F}_m < \alpha_m + R$ and rearranging yields the condition $R > \beta \alpha_m$.

From the proof of Proposition 2, $F^*$ is defined implicitly by $\delta(F^*) = \beta$. For the uniform distribution, some algebra reveals $\delta(F) = \frac{F - \alpha_m + R}{2}$, and so

$$F^* = \alpha_m + 2\beta \alpha_m - R.$$

Then

$$\frac{\partial}{\partial R} \left(\hat{F}_m - F^*\right) = \frac{4R(R - \beta \alpha_m)}{(2R - \beta \alpha_m)^2}.$$ 

This quantity is strictly positive if and only if $R > \beta \alpha_m$, which, from above, is the condition for $\hat{F}_m < \alpha_m + R$. ■

Proof of Proposition 4

From the definition of gridlock inefficiency, $\frac{\partial W}{\partial F} \bigg|_{F=F^*-\Delta} > 0$ and $\frac{\partial W}{\partial F} \bigg|_{F=F^*+\Delta} < 0$. Therefore it is sufficient to demonstrate that $\frac{\partial W}{\partial F} \bigg|_{F=F^*-\Delta} < -\left(\frac{\partial W}{\partial F} \bigg|_{F=F^*+\Delta}\right)$. Integrating by parts and
rearranging, $\delta(F) = \frac{\hat{p}(F)}{P(F)}$. Substituting from (A.7) and rearranging, this is equivalent to

$$\beta < \frac{\hat{P}(F^* + \Delta) + \hat{P}(F^* - \Delta)}{P(F^* + \Delta) + P(F^* - \Delta)}, \quad (A.12)$$

or, given the first order condition,

$$\frac{\hat{P}(F^*)}{P(F^*)} < \frac{\hat{P}(F^* + \Delta) + \hat{P}(F^* - \Delta)}{P(F^* + \Delta) + P(F^* - \Delta)}. \quad \text{(A.12)}$$

Rearranging, this expression is equivalent to

$$\hat{P}(F^*) \left( \frac{P(F^* + \Delta) + P(F^* - \Delta)}{2} \right) < \left( \frac{\hat{P}(F^* + \Delta) + \hat{P}(F^* - \Delta)}{2} \right) P(F^*).$$

From the convexity of $\hat{P}(\cdot)$, $\hat{P}(F^*) < \frac{\hat{P}(F^* + \Delta) + \hat{P}(F^* - \Delta)}{2}$. Assume a sequence of densities $\{p_1(\alpha), p_2(\alpha), \ldots\}$ such that $p_k(\cdot) \succ_{ULR} p_{k+1}(\cdot)$ for all $k$. Then $\lim_{k \to \infty} \left( \frac{P(F^* + \Delta) + P(F^* - \Delta)}{2P(F^*)} \right) = 1$, and (A.12) holds for sufficiently large $k$. ■

**Proof of Proposition 5**

Let $A(F; \beta, P(\cdot)) \equiv \frac{F}{1 + \beta P(F)}$ denote the value of $\alpha$ that would yield $F$ as an ideal point. From the expressions for states’ marginal utilities in (A.4),

$$RP^-(F; \beta, P(\cdot)) \equiv - \int_0^{A(F^*)} (\alpha(1 + \beta P(F)) - F)p(\alpha)d\alpha \quad \text{and}$$

$$RP^+(F; \beta, P(\cdot)) \equiv \int_0^{F} (\alpha(1 + \beta P(F)) - F)p(\alpha)d\alpha + \beta P(F) \int_{F}^{a} \alpha p(\alpha)d\alpha.$$

Substituting for $A(F)$, integrating by parts (see derivation of $\Psi^*(\cdot)$ above for details) and rearranging yields
\[ RP^-(F; \beta, P(\cdot)) = (1 + \beta P(F)) \hat{P}(A(F; \cdot)) \quad \text{and} \]
\[ RP^+(F; \beta, P(\cdot)) = (1 + \beta P(F)) \hat{P}(A(F; \cdot)) + \beta P(F) E[\alpha] - \hat{P}(F). \] (A.13)

\[ \frac{\partial RP^-}{\partial F} = (1 + \beta P(F)) P'(A(F)) \frac{\partial A(F)}{\partial F} + \beta p(F) \hat{P}(A(F)) > 0 \text{ for all } F > \alpha \text{ (noting that } \frac{\partial A(F)}{\partial F} > 0), \text{ and} \]
\[ \frac{\partial RP^+}{\partial F} = \frac{\partial RP^-}{\partial F} + \beta p(F) E[\alpha] - P(F). \]

Having established \( \frac{\partial RP^-}{\partial F} > 0 \), it is sufficient to demonstrate that \( \beta p(F) E[\alpha] - P(F) > 0 \) for sufficiently small values of \( F \). Rearranging, the sufficient condition is \( \frac{p(F)}{P(F)} > (\beta E[\alpha])^{-1} \). From the definition of log-concavity, \( \frac{p(F)}{P(F)} \) is strictly decreasing. Further, \( \lim_{F \to \pi^-} = \infty \).

Therefore the condition holds for sufficiently small values of \( F \).

(2) We proceed by showing that there exists an \( F \) such that the result holds for \( F = F^* \). Comparing the expressions from (A.13), \( RP^- > RP^+ \) if and only if \( \delta(F) > \beta E[\alpha] \) (where, as above, \( \delta(F) = \frac{P(F)}{P'(F)} \)). From the proof of Proposition 2, at equality this statement defines \( F^* \) implicitly. Via monotonicity of \( \delta(\cdot) \), therefore, \( RP^- > RP^+ \) if and only if \( F > F^* \). Suppose \( F > F^* \), so \( RP = RP^- \). As \( RP^- \) is strictly increasing, it is minimized at \( RP^-(F^*) > 0 \).

Suppose \( F < F^* \), so \( RP = RP^+ \). From part (1), \( RP^+ \) is increasing for sufficiently small values of \( F \). From the second line of (A.13), \( RP^+(\alpha) = 0 \). As \( RP^+ \) is therefore increasing from zero, there must be some \( F \) such that for all \( F < F, RP^+(F) < RP^-(F^*) \). \( \blacksquare \)

Proof of Proposition 6

(1) Note first that state levels of provision \( S^t(\alpha^t_i; F^t) \) are not sticky, and are chosen optimally given \( \alpha^t_i \) and \( F^t \). It is clear that \( \hat{F}^2(\alpha^2_i; \cdot) = \hat{F}(\alpha^2_i; \cdot) \) from the one-period model. Because \( \hat{F}^2(\cdot; \cdot) \) is monotone in \( \alpha^2 \), its inverse is well-defined. Let \( A^2(F) \) be the inverse of \( \hat{F}^2(\alpha^2_i; \cdot) \), that is, the \( \alpha^2_i \) for whom \( \hat{F}^2(\alpha^2_i; \cdot) = F \) is \( A^2(F) \). Thus, for \( F^1 \) to be in the gridlock interval in \( t = 2 \), it must be that the shock \( \sigma \) is such that \( A^2(F^1) \in [\alpha_L + \sigma, \alpha_H + \sigma] \), i.e., \( \sigma \in [A^2(F^1) - \alpha_H, A^2(F^1) - \alpha_L] \).
The ex ante expected utility from choice $F^1$, given subsequent equilibrium behavior, is

$$
u(F^1, \alpha, \cdot) + \int_{\alpha}^{A^2(F^1) - \alpha_L} g(\sigma)u(F^1, \alpha_i + \sigma, \cdot) d\sigma$$

$$+ \int_{-\infty}^{\alpha} g(\sigma)u(\hat{F}^2(\alpha + \sigma), \alpha_i + \sigma, \cdot) d\sigma$$

$$+ \int_{A^2(F^1) - \alpha_H}^{\infty} g(\sigma)u(\hat{F}^2(\alpha - \sigma), \alpha_i + \sigma, \cdot) d\sigma,$$

where $u(F^1, \alpha, \cdot)$ is the indirect single-period payoff for the first period given that $S^t_i = S(\alpha, F^t)$, the first integral is the expected second-period payoff when the shock lands the system in the gridlock interval, and the second and third integrals are the second-period payoffs when the shock moves the system to the right of the right bound and to the left of the left bound of the gridlock interval, respectively. Note that in the integrand of the first integral, the payoff is evaluated at $F^1$, because the gridlock in the second period implies that $F^2 = F^1$.

The first-order condition that defines the optimal value $F^1$ for $\alpha_i$ is

$$\frac{\partial u(F^1, \alpha_i, \cdot)}{\partial F^1}$$

$$+ g(A^2(F^1) - \alpha_L)u(F^1, \alpha_i + A^2(F^1) - \alpha_L, \cdot) \frac{\partial A^2(F^1)}{\partial F^1}$$

$$- g(A^2(F^1) - \alpha_H)u(F^1, \alpha_i + A^2(F^1) - \alpha_H, \cdot) \frac{\partial A^2(F^1)}{\partial F^1}$$

$$+ \int_{A^2(F^1) - \alpha_H}^{\alpha} p(\sigma) \frac{\partial u(F^1, \alpha_i + \sigma)}{\partial F^1} d\sigma$$

$$+ g(A^2(F^1) - \alpha_H)u(\hat{F}^2(\alpha + A^2(F^1) - \alpha_H), \alpha_i + A^2(F^1) - \alpha_H, \cdot) \frac{\partial A^2(F^1)}{\partial F^1}$$

$$- g(A^2(F^1) - \alpha_L)u(\hat{F}^2(\alpha + A^2(F^1) - \alpha_L), \alpha_i + A^2(F^1) - \alpha_L, \cdot) \frac{\partial A^2(F^1)}{\partial F^1}$$

$$= 0$$
Noting that $\hat{F}^2(A^2(F^1)) = F^1$ and canceling terms, we obtain an equivalent condition

$$\frac{\partial u(F^1, \alpha_i, \cdot)}{\partial F^1} + \frac{A^2(F^1) - \alpha_L}{A^2(F^1) - \alpha_H} \int_{A^2(F^1) - \alpha_H}^{A^2(F^1) - \alpha_L} g(\sigma) \frac{\partial u(F^1, \alpha_i + \sigma, \cdot)}{\partial F^1} d\sigma = 0 \quad (A.14)$$

If $\int_{A^2(F^1) - \alpha_H}^{A^2(F^1) - \alpha_L} g(\sigma) \frac{\partial u(F^1, \alpha_i + \sigma)}{\partial F^1} d\sigma$ is less (greater) than 0, then $\frac{\partial u(F^1, \alpha_i, \cdot)}{\partial F^1}$ must be greater (less) than 0 to compensate. It follows that the value of $F^1$ that solves (A.14) is less (greater) than the value $\hat{F}(\alpha_i, \cdot)$ from the one-shot model, for which $\frac{\partial u(\hat{F}(\alpha_i, \cdot), \alpha_i, \cdot)}{\partial F^1} = 0$.

Recall that $A(F^1)$ is the state for which $\frac{\partial u(F^1, A(F^1), \cdot)}{\partial F^1} = 0$ in the one-period model. If $\alpha_i + \sigma < A(F^1)$, then the integrand in equation (A.14) is less than 0, and so for $\sigma < A^2(F^1) - \alpha_i$, $\frac{\partial u(F^1, \alpha_i + \sigma)}{\partial F^1} < 0$. So, for $A^2(F^1) - \alpha_L \leq A(F^1) - \alpha_i$, the integrand is less or equal to zero, and thus the value of the integral is less than 0. Thus, if $\alpha_i < \alpha_L$, then $\hat{F}^1(\alpha_i) < \hat{F}(\alpha_i) = \hat{F}^2(\alpha_i)$.

If $\alpha_i + \sigma > A(F^1)$, then $\frac{\partial u(F^1, \alpha_i + \sigma, \cdot)}{\partial F^1} > 0$. Thus, if $\sigma > A^2(F^1) - \alpha_i$, $\frac{\partial u(F^1, \alpha_i + \sigma, \cdot)}{\partial F^1} > 0$. So, if $A(F^1) - \alpha_H \geq A(F^1) - \alpha_i$, then the value of the integral is greater than 0. Thus, for $\alpha_i \geq \alpha_H$, $\hat{F}^1(\alpha_i) \geq \hat{F}(\alpha_i) = \hat{F}^2(\alpha_i)$.

It follows that $\hat{F}^1(\alpha_L) < \hat{F}(\alpha_L)$ and $\hat{F}^1(\alpha_H) > \hat{F}(\alpha_H)$.

(2) First, note that for $\alpha_i = \alpha_L$, the integrand from (A.14) evaluated at the upper bound $\sigma = A^2(F^1) - \alpha_L$ is $g(A^2(F^1) - \alpha_L) \frac{\partial u(F^1, A^2(F^1), \cdot)}{\partial F^1}$. From the definition of $A^2(F^1)$, $\frac{\partial u(F^1, A^2(F^1), \cdot)}{\partial F^1} = 0$, and thus $\forall \sigma < A^2(F^1) - \alpha_L$, the integrand is negative. Similarly, for $\alpha_i = \alpha_H$, the integrand evaluated at the lower bound $\sigma = A^2(F^1) - \alpha_H$ is $g(A^2(F^1) - \alpha_H) \frac{\partial u(F^1, A^2(F^1), \cdot)}{\partial F^1} = 0$, and $\forall \sigma > A^2(F^1) - \alpha_H$, the integrand is positive.

Observe next that if the integrals in (A.14) had equal magnitudes for $\alpha_i = \alpha_L$ and $\alpha_i = \alpha_H$, then the asymmetry in $u_i(F, \alpha_i, \cdot)$ around $F = \alpha_i$ would imply that $|\hat{F}(\alpha_L) -$
\( \hat{F}^1(\alpha_L) > |\hat{F}^1(\alpha_H) - \hat{F}(\alpha_H)| \). However, from (4), it can be seen that for \( \beta \) sufficiently small,

\[
\left| \int_{A^2(F^1) - \alpha_L} g(\sigma) \frac{\partial u(F^1, \alpha_L + \sigma, \cdot)}{\partial F} d\sigma \right| > \left| \int_{A^2(F^1) - \alpha_H} g(\sigma) \frac{\partial u(F^1, \alpha_H + \sigma, \cdot)}{\partial F} d\sigma \right|.
\]

Thus, \( \left| \frac{\partial u(\alpha_L, F^1(\alpha_L), \cdot)}{\partial F} \right| > \left| \frac{\partial u(\alpha_H, F^1(\alpha_H), \cdot)}{\partial F} \right| \), reinforcing the asymmetry in the divergence of \( \hat{F}^1(\alpha_i) \) from \( \hat{F}(\alpha_i) \) for low- and high-demanders. ■

References
