ESSAYS IN POLITICAL ECONOMY:

Incentive And Efficiency Implications Of Institutional Rules In Three Political Settings

GIRIDHAR PARAMESWARAN

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF ECONOMICS
ADVISERS: JOHN B. LONDREGAN & STEPHEN MORRIS

JUNE 2013

Abstract

This collection of essays employs techniques in dynamic game theory to study applied questions in political economy and associated fields. In each essay, I investigate the consequences and efficiency implications of the various rules, structures and biases associated with a given political institution. The first essay studies rule making by common law courts that are bound by the doctrine of *stare decisis*. The court is imperfectly informed about the ideal legal rule but can learn about it through the cases it hears. Since agents are rationally responsive to the court's decisions, the court's learning is limited by its ability to make experimentation incentive compatible. The model provides a systematic explanation for why courts write narrow rules in some cases, but broad rules in others. The constrained court will never perfectly discover the ideal rule, and so the common law will entrench inefficient rules with positive probability. The second essay studies the properties of fiscal policies that are the consequence of political competition when voters exhibit a projection bias, whereby agents are unduly optimistic during booms (about the boom's persistence) and unduly pessimistic during recessions. In this environment, I show that the equilibrium fiscal policy will feature taxes that are inefficiently volatile and debt that ceases to efficiently smooth the deadweight cost of taxes through time. This mechanism explains both the public's complacency towards rising debt during booms, when debt should ideally fall, and the public's significant debt aversion during recessions, when debt should ideally increase. An implication for public policy is that the government should implement mechanisms - such as strongly progressive tax codes - that amplify deficits during recessions and amplify surpluses during booms. The third essay studies the nature of optimal coalitions in bicameral legislatures when the preferences of certain agents in the two chambers are correlated. Contrary to the received wisdom, I show that it is possible for bicameralism to privilege large states, by skewing the composition of the coalitions that will optimally form in equilibrium.

Acknowledgements

I am very grateful to my supervisors - John Londregan and Stephen Morris for all their assistance during my Ph.D candidacy. I owe them both a significant debt of gratitude for their patience, time, care and generosity.

I am also grateful to Charles Cameron and Jack Fanning and to participants at the 2012 EconCon conference for helpful comments on the first chapter of this dissertation, *Ruling Narrowly and Broadly*; to Marco Battaglini and Edoardo Grillo for all of their assistance with the second chapter - *Psychological Belief Distortions and Debt*; and to Wolfgang Pesendorfer for very helpful comments on the third chapter. Each of these chapters also benefited from helpful comments from participants at the microeconomic theory student seminar and political economy student seminar at Princeton University.

Contents

List of Figures

1 Introduction

This collection of essays employs techniques in dynamic game theory to study applied questions in political economy and associated fields. In each essay, I investigate the consequences and efficiency implications of the various rules, structures and biases associated with a given political institution. The first essay, *Ruling Narrowly or Broadly? Learning, Experimentation and Law Creation*, studies rule making by common law courts - in environments where the court is imperfectly informed about the ideal legal rule. The court must trade-off the benefits or ruling broadly, to reduce uncertainty about the law, against the costs of entrenching inefficient rules and forgoing the opportunity to learn about the law through future cases. Courts are bound by the doctrine of *stare decisis* - or respect for precedent. This requires courts to implement prior rules that have been found to be inefficient, *ex post*, but also provides a commitment device for the court to affect agent behavior, as information is revealed.

This essay extends previous inquiries in this field by considering agents who are rationally responsive to the court's rules. I show that, in this environment, the court will not always be able to provide incentives for agents to experiment; the court's ability to learn is constrained by its ability to make experimentation incentive compatible. In contrast to other models, I show that the constrained court will never perfectly discover the ideal rule, and so the common law will entrench inefficient rules with positive probability. The model provides a systematic explanation for why courts write narrow rules in some cases, but broad rules in others. As opportunities to learn disappear, the court's incentives to write broad rules increases. Surprisingly, this incentive is asymmetric - the court will tend to write broader rules when encouraging behavior by the agent - but narrower rules when restricting that behavior.

The second essay, *Psychological Belief Distortions and Debt*, studies the properties of fiscal policies that result from political competition when voters exhibit a projection bias - during booms, voters are unduly optimistic (i.e. overweight the probability) about the persistence of the boom, whilst during recessions, they become unduly pessimistic about the persistence of the recession. In this environment, I show that the government will over-accumulate debt during booms and under-accumulate debt during recessions, resulting in fiscal policy with taxes that are inefficiently volatile and debt that ceases to efficiently smooth the deadweight cost of taxes through time. This mechanism explains both the public's complacency towards rising debt during booms, when debt should ideally fall, and the public's significant debt aversion during recessions, when debt should ideally increase. In contrast to conventional solutions to inefficient fiscal policy choices by the government (such as balanced budget amendments or statutory debt ceilings), the model prescribes mechanisms that amplify deficits during recessions and amplify surpluses during booms - such as a strongly progressive tax code.

The third essay, *Less Representation is Better: How Bicameralism Can Benefit Large States*, extends a standard model of legislative bargaining a bicameral legislature that must allocate funds amongst districts. In each bargaining round, a legislator is selected to propose an allocation, which is implemented only if it is supported by majorities in both chambers. To ensure the proposal's passage, the proposer must allocate enough funds to certain districts to 'buy' the support of a minimum winning coalition of legislators.

The bicameral setting introduces complementarities between the preferences of agents in the different chambers - in particular, by buying the support of a senator from a particular state, the proposer will automatically receive the support from lower house members from that state. This complementarity changes the nature of the optimal coalition - although the proposer seeks to purchase the cheapest winning coalition - this may involve the votes of large state legislators (since this implies the support of many more lower house members) - even if large states are more 'expensive'. This insight challenges the received wisdom that bicameralism necessarily privileges small states in a federal system. I show that the relative gains or losses to small states from bicameralism depends on the size of the majority requirement in each chamber, and the recognition probabilities of legislators from large and small states.

2 Ruling Narrowly or Broadly? Learning, Experimentation and Law Creation

2.1 Introduction

Courts play a central role in settling disputes between litigants. In resolving most disputes, the role of the court is to uncover the relevant facts in order to correctly apply an existing legal rule. In these cases, the court's behavior is quite mechanical - it simply applies existing law to the case before it. However, in some cases, the existing law may be silent or unclear about how the dispute ought to be settled. In such cases the common law tradition envisions an additional, and often controversial, role for the court - to extend the law in a way that allows it to resolve the issue at hand. In doing so, courts sometimes write narrow opinions - extending the law only insofar as is necessary to resolve the dispute at hand - whilst at other times writing broad opinions, that extend the law further.

This paper seeks to address the question of how the court ought to optimally create law. These questions are clearly important. The public's interest in maintaining the separation of powers in a democracy, for example, creates perennial concerns about judges 'legislating from the bench' when they write broad opinions. These questions are also important given the increasing popularity of minimalism as a judicial philosophy.

Minimalists cite many and varied reasons for their particular approach to law creation. One important motivation for the minimalists' cautious approach is to minimize the likelihood that the court will make mistakes that it cannot easily undo. For example, in *City of Ontario v Quon*[1], Justice Kennedy's opinion asserted that: "Prudence counsels caution before the facts in the instant case are used to establish far-reaching premises... A broad holding ... might have implications for future cases that cannot be predicted." Similarly, at his Senate confirmation hearing, Justice Alito stated:

[1] *City of Ontario v Quon, 560 U.S. _____ (2010)*

"I think that my philosophy of the way I approached issues is to try to make sure that I get right what I decide. And that counsels in favor of not trying to do too much, not trying to decide questions that are too broad, not trying to decide questions that don't have to be decided, and not going to broader grounds for a decision when a narrower ground is available."

The minimalist approach to adjudication is, of course, not without its own problems. Whilst the minimalist court's 'epistemic humility' may reduce the costs associated with adopting sub-optimal rules, it perpetuates the existing uncertainty surrounding the law, which is itself costly. Even justices who typically advocate for the minimalist approach are cognizant of the significant costs that uncertainty places on future decision making. For example, in *Blakely v Washington*[2], the majority held that the State of Washington's criminal sentencing system, which gave judges the ability to increase sentences based on their own determination of facts, violated the Sixth Amendment right to trial by jury. However, the majority opinion did not address the constitutionality of the Federal Sentencing Guidelines, which had many similarities to the Washington law. In a dissenting opinion, Justice Breyer argued: "But this case affects tens of thousands of criminal prosecutions... Federal prosecutors will proceed with those prosecutions subject to the risk that all defendants in those cases will have to be sentenced, perhaps tried, anew. Given this consequence and the need for certainty, I would not proceed further piecemeal." The narrow approach "all but invites further challenges" Smith (2010). Indeed, the very first case that the Court heard in the following term - *United States v Booker*[3] - dramatically changed the legal framework within which legal sentencing takes place at the Federal level (Siegel, 2005).

In this paper, I develop a model that captures the trade-off that the court faces between reducing uncertainty about the law on the one hand, and forgoing the opportunity to learn and thereby entrenching potentially sub-optimal rules, on the other . In so doing, I provide a framework with which to assess the merits of the minimalist approach. This paper focuses on the informational aspects of decision making. I ignore the democracy and heterogeneity

[2] *Blakely v Washington*, 542 U.S. 296 (2004)
[3] United States v. Booker, 543 U.S. 220 (2005)

based arguments for minimalism by considering a common value problem, where all similarly informed judges would make the same choice. Importantly, I show that when the incentives of the agents (whose behaviour is being regulated) are properly modeled, the court's ability to learn about the ideal legal rule will be limited - and so that even in the long run, the court will implement inefficient policies with high probability. This result stands in contrast to existing models (e.g. Baker and Mezzetti (2012)) that show that learning under the common law is consistent with long run efficiency.

I model decision making within the context of a profit maximizing firm whose production generates a negative externality (such as water pollution) that imposes a harm on third parties. In the absence of regulation, the firm will generate more output than is socially efficient, since it fails to internalize the external cost it imposes upon others. The size of the external cost is unknown - although the court and firm share common beliefs about its value. As the court hears different cases, it observes whether the firm's output (in that case) was inefficiently large or not, and updates its beliefs about the efficient output level. In this way, the court 'learns' about what the ideal policy ought to be. (The learning process described has many parallels with the literature on price setting by a monopolist facing an unknown demand. For example, see Rothschild (1974) and Aghion et al. (1988).)

The court's role is to regulate the firm's behavior in order to implement the socially efficient outcome. The court does this by announcing a partial legal rule - which specifies thresholds for production above (below) which the firm will be definitely (definitely not) held liable - and a penalty for over-production. For example, the partial legal rule may definitely hold the firm liable whenever it produces a quantity in excess of 60, and definitely hold the firm not liable if it produces a quantity below 50. The rule is silent as to the liability status of a firm producing between 50 and 60 units of output - it will be held liable or not depending on the outcome of the court's investigation, as explained above. In the uncertain region, the firm's beliefs about the size of the efficient allocation will determine its assessment of the probability of being penalized.

In writing its opinion, the court can commit to holding the firm liable (or not) even at output levels where it is uncertain about what the ideal outcome ought to be. For example, the court's belief may be that the efficient output level lies between 45 and 70. Then, in the above example, the announced rule commits the court to holding a firm that produces 50 units of output not liable even though the court believes that, with some probability, the efficient level is actually below 50. Similarly, it commits the court to find a firm producing 61 units liable, even though its beliefs suggest that the efficient level may actually be above this level. The lower threshold is permissive, since it establishes the region over which the firm may produce with impunity. By contrast, the upper threshold is restrictive, since it establishes the region over which the firm will be punished for sure.

A rule is narrow if it commits to holding the firm liable (not liable) only in cases where it puts probability one on the firm's output being inefficiently high (low). A rule is broad if it finds the firm definitely liable or definitely not liable even in cases where the common beliefs imply some uncertainty. (Such a rule is broad, because it establishes a rule even for cases of the sort that the court has not heard and had an opportunity to learn about, and, as such, potentially entrenches costly errors.) The effect of broad rules is to create a wedge between the probability that the firm expects that it ought to be penalized - as implied by its beliefs about the efficient output level - and the probability that it will actually be penalized. In so doing, the court can create incentives for the firm to produce more or less than it might otherwise do.

The model's focus on the regulation of a polluting firm is purely to give context to the analysis. The model could be applied equally well to any other situation in which the court seeks to regulate the behavior of agents whose incentives are misaligned. For example, the framework presented could accommodate a model of search and seizure jurisprudence, in which the court seeks to regulate the conduct of the police - who have an incentive to conduct more invasive searches than is socially desirable, given individuals' privacy interests. Similarly, the framework could accommodate a model of separation of powers - in which the court seeks to regulate the activity of the executive branch (for example) or the federal government - who may have an incentive to extend their influence into matters that are properly the realm of other branches or tiers of government.

An important ingredient in my model is that the firm is assumed to rationally respond to the court's prior rulings and to other available information. (This assumption stands in contrast to other models of law creation, including Baker and Mezzetti (2012) and Niblett (2010), in which agent behavior is unmodeled and often assumed to be drawn randomly from some legal-rule-invariant distribution.) The assumption that agents are responsive to the court's rulings becomes especially important given that many common law countries require that courts only adjudicate actual controversies. The court cannot simply expound on the law if there is no actual dispute between the parties that requires resolution. It cannot adjudicate hypothetical cases. This procedural constraint, along with the fact that the firm chooses its policy rationally, has two important implications. First, the court may not be able to learn about the validity of a particular type of behavior if agents never choose to behave in that way. The court can learn about the ideal rule only insofar as it can provide incentives for the firm to experiment by choosing an output level in the uncertain region. The court does not learn by experimenting - rather it *learns by having others experiment* - and to this extent, its learning is limited by its ability to make experimentation incentive compatible. Hence, in adjudicating a case, the court must be cognizant of the effect of its current decision on its future ability to learn. Second, the ability of the court to modify the law in the future is limited to the extent that a case arises in the future that requires the court to modify its rule. This may create an incentive for the court to preemptively write broader opinions than it otherwise might, just in case the opportunity to revise the rule does not arise soon enough in the future.

A second important feature of the model that has consequences related to those above, is the assumption that common law courts are bound by the doctrine of *stare decisis* - or respect for precedent. Adherence to precedent implies that the court will mechanically dispose of cases that are already governed by an existing rule. This further prevents learning. The law only progresses when the court can provide incentives for the firm to experiment in the ambiguous region of the law.

The model makes several predictions about the nature of law creation. First, the model demonstrates that there is an asymmetry in the efficacy of permissive rules *vis a vis* restrictive rules, in affecting the firm's output choice. This reflects the inherent asymmetry in the model, in which the unregulated firm deviates systematically from the efficient level, by over-producing. Making the rule more permissive, then, will cause the firm to weakly increase its output. By contrast - making the rule more restrictive can cause the firm's choice to both increase or decrease. (If the penalty is not too high, as the rule becomes more restrictive, the firm may prefer to simply over-produce and bear the penalty for sure,mathnormalan produce a lower level of output at which it may still be penalized.) The model predicts that the court ought to write broad permissive opinions and narrow restrictive opinions. For example, when ruling on the legality of police action, the court ought to write broader opinions in cases where it finds the police action to be appropriate, but narrow opinions when it seeks to limit police power. The implications for minimalism as a doctrine are thus divided. Narrow opinions are ideal when the court seeks to restrict the behavior of the agent, but broad opinions are preferred when the court seeks to encourage that behavior.

Second, the model predicts that the court will use its opinion to entice the firm to experiment - but that the experimentation that it induces generates an inefficiently low level of learning. (One way to think about the court inducing experimentation is that courts often signal to future potential litigants about the sorts of cases that they would like to hear.) Efficient learning occurs when the court experiments at a level that minimizes the expected future cost of uncertainty. (I will define this more precisely in Section 2.5.) In my model, efficient learning requires the firm to experiment at the mean of the uncertain region. This makes it equally likely that the experimental level will be found to be acceptable or not, and hence equally likely that the court will be able to revise its lower and upper opinions. However, since the court's policy tools are not equally efficacious, the court would rather decide a case that allowed it to revise its lower rule than its upper rule. This creates an incentive for the court to entice the firm to experiment at a level that is more likely to be found acceptable, than not. The court would rather learn less (in the sense that the expected future costs of uncertainty are larger), but acquire information that it can use more effectively to implement the efficient allocation in the future.

This model contributes to an emerging literature on judicial decision making and law creation. The two papers most similar to this are Fox and Vanberg (2011) and Baker and Mezzetti (2012). Fox and Vanberg (2011) consider a court that has determined that the legislature's current policy is unconstitutional, but is unsure of what the ideal law ought to be. They show that by writing a broad opinion, the court can force the legislature to experiment in a region of the policy space where the court can learn more efficiently about the ideal law. Fox and Vanberg (2011) consider a one-sided model of law creation (in which the court either determines a policy to be definitely unconstitutional or not) - and so their model does not display the asymmetries that will become important in this model. Moreover, they consider a context in which the court's opinion is the only tool that it can use to affect the legislature's policy - there is no other penalty for non-compliance (such as fine) - and this implies a greater need for the court to write broad opinions. By providing the court with additional policy tools, my model 'stacks the deck' in favor of minimalism, but nevertheless finds that the court ought to write broad opinions in certain cases.

Baker and Mezzetti (2012) model a resource constrained court that must, in each period, determine how broadly to construe the outcomes of previous cases. They show in an infinite horizon model that the law will always converge, and provide conditions under which it converges to the efficient level. Whilst their analysis is two-sided, the decisions of the agents whose behavior is being regulated by the court are unresponsive to the court's policy - cases are simply drawn from a fixed distribution. This papert demonstrates the important limitations on court learning when agents are assumed to respond rationally to the court's decisions. Indeed, contrary to Baker and Mezzetti (2012), this paper shows as that as the amount of uncertainty diminishes, the court is increasingly limited in its ability to provide incentives for the firm to continue to experiment. When the degree of uncertainty is small enough, experimentation stops entirely.

Both these papers also make reduced form assumptions about the nature of judicial preferences. To my knowledge, this is the first model that fully endogenizes the court's preferences, and which models the costs of uncertainty in a structural, rather than reduced form, way. As will become clear, this has important consequences for the sorts of conclusions that follow from the model.

These papers are at the intersection of two strands of the literature on judicial politics. Previous studies of law creation have typically investigated the implications of heterogeneity (bias) in judicial preferences on the evolution of the common law. (See, for example, Gennaioli and Shleifer (2007), Niblett (2010), Ponzetto and Fernandez (2008).) Gennaioli and Shleifer (2007), for example, provide foundations for the "Cardozo Theorem", which states that the individual biases of judges tend to wash out as law is created piece-meal. These models typically do not involve any uncertainty about the ideal legal rule - there is no role for learning. A separate literature - most notably Clark and Kastellec (2010) and Beim (2012) - consider models of judicial learning. These papers focus on the Supreme Court's decision to grant *certiorari* (or not) in a given case, based on signals from lower court decisions in related cases. A similar literature considers on the informational aspects of decision making, in situations where a superior court cannot perfectly convey its ideal policy to inferior courts, or where its ability to monitor inferior courts is imperfect. (See, for example, De Mesquita and Stephenson (2002) and Cameron et al. (2000).)

The remainder of this paper proceeds as follows: In subsection 2.2 I present the formal model. Subsection 2.3 analyzes the full information case and shows that the court implements the efficient allocation. Subsection 2.4 considers the firm's decision under uncertainty, and Subsection 2.5 analyzes the court's optimal choice. Subsection 2.6 presents some extensions.

2.2 Model

There is a profit maximizing firm that must choose the quantity to produce of a single good. The firm's profit from producing q units of output is $\pi(q) = \alpha q - \frac{1}{2}\beta q^2$, where $\alpha, \beta > 0$. The profit maximizing level of output is $q_H = \frac{\alpha}{\beta}$. The concavity of the profit function implies that the firm is averse to risk.

The firm's production generates an externality that harms a third party. The external cost of q units of output is $C(q) = \theta q$. In the presence of the negative externality, the unregulated firm will produce more output than is socially optimal, because it does not internalize the external cost of its production. The socially efficient level of output is $q^* = \frac{\alpha-\theta}{\beta}$. To ease notation, I let $S(\theta) = \frac{\alpha-\theta}{\beta}$ denote the efficient quantity when the marginal external cost is θ, and let $T(q) = \alpha - \beta q$ denote the marginal external cost for which q is the efficient output level. Obviously, $S(\theta) = T^{-1}(\theta)$.

The size of the marginal external cost, θ, is unknown. I assume that all agents share common beliefs about θ. For simplicity, suppose prior beliefs are uniformly distributed on the interval $[l_0, u_0]$, where $0 < l_0 < u_0 < \alpha$. These bounds imply that the efficient output level is positive ($q^* \geq \frac{\alpha-u_0}{\beta} > 0$) but less than the level that maximizes the firm's profit ($q^* \leq \frac{\alpha-l_0}{\beta} < \frac{\alpha}{\beta}$). This implies a role for an efficiency minded court to regulate the firm's behavior, and implement the socially efficient outcome.

The external party may bring a case before the court to seek compensation if it experiences harm. The court uses the *liable if over-produce* rule, whereby it holds the firm liable for damages if it finds that the firm has over-produced relative to the socially efficient output level. This is analogous to the Hand Rule[4] that determines the standard of care in negligence cases.[5] Since the external party is always harmed by the firm's production, the court will

[4] United States v. Carroll Towing Co. 159 F.2d 169 (2d. Cir. 1947).

[5] Indeed, Judge Learned Hand proposed (in quite explicit mathematical terms) that the defendant be held liable for damages if the expected reduction in the injury from taking extra precaution outweighed the additional cost - i.e. if less than efficient precaution was taken. Judge Hand wrote: "Possibly it serves to bring this notion into relief to state it in algebraic terms: if the probability be called P; the injury, L; and the burden, B; liability depends upon whether B is less than L multiplied by P: i.e., whether B < PL."

always be adjudicating a genuine controversy between the parties. For simplicity of exposition, I assume that the external party brings a case before the court whenever it believes that there is a positive probability of being compensated. If the firm is held liable, it must pay a flat penalty F. (The assumption that the fine is invariant to the level of output simplifies the analysis considerably, however I acknowledge that it is quite restrictive. In Section 2.6 I demonstrate that a proportional penalty schedule $F = fq$ has similar features to the flat penalty, and argue that the qualitative implications of the simple model should extend to a general class of continuous and increasing penalty schedules.)

The law is described by a partial legal rule (λ, μ) that partitions the case-space into three equivalence classes. For any output $q \leq \lambda$, the rule prescribes that the firm will not be held liable (even if beliefs assign a positive probability to the chosen output being inefficiently large), and for any $q > \mu$, the rule prescribes that the firm will definitely be held liable (even if beliefs assign a positive probability to the chosen output being below the efficient level). Whenever the court encounters a case that is governed by the partial rule, it mechanically applies the rule and determines the firm to be liable or not, as required. (Alternatively, in a hierarchical court structure, lower courts are assumed to always mechanically apply the higher court's rule once set, and the higher court will never agree to review such cases.) The partial rule is silent about the validity of choices in the region $\lambda < q \leq \mu$. If a case arises in this region, the court fully investigates the case - it hears expert testimony, receives amicus briefs etc. - and perfectly learns whether the chosen output level was above or below the socially efficient level. Note that learning is only possible if a case arises in the ambiguous region, since outside this region, the court mechanically disposes of the case without investigation.

A latent assumption is that the court chooses rules that are monotonic - i.e. whenever the rule finds q to be acceptable, then it will also find any $q' < q$ acceptable, and whenever it finds q unacceptable, it will also find any $q' > q$ unacceptable. Given that the case space is unidimensional and ordered, it is natural to focus on monotonic rules of this sort. Moreover, the unidimensionality of the case space provides an ordering over policies that allows a natural way to classify rules as narrow or broad.

The court's investigation is public, and so all players update their beliefs about θ in the same way after observing the outcome of a particular case. Hence, beliefs are common. The posterior distribution of beliefs remains uniform, since the prior was uniform and the new information simply truncates the support of beliefs. As an illustration, suppose the court finds output q to be inefficiently large, which implies that $T(q) = \alpha - \beta q < \theta$. Then, the posterior beliefs are given by:

$$\begin{aligned} \Pr[\theta \leq x | \theta > T(q)] &= \frac{\Pr[T(q) < \theta < x]}{\Pr[\theta > T(q)]} \\ &= \frac{x - T(q)}{u_0 - T(q)} \end{aligned}$$

and so $\theta^{post} \sim U[T(q), u_0]$. If q were found to be acceptable, then $\theta^{post} \sim U[l_0, T(q)]$. Since beliefs are always uniform, they are completely summarized by the extreme values of the support of the distribution. Hence, from herein, I will denote beliefs by a pair (l, u) where $l \leq u$.

In addition to updating beliefs, the court must also extend the partial rule in a way that is consistent with the outcome of the case. This is the process of law creation and it is asserted through the court's opinion in a given case. For example, if the existing partial rule is (λ_0, μ_0) and the court finds output level q to be unacceptable, then the court must write an opinion that extends the partial rule such that $\mu_1 \leq q$. This ensures that if an identical case were to arise in the future, the future court would adjudicate the case in the same way. (Note importantly that the disposition of the case restricts the sorts of opinions that the court may write. The court cannot simply act as a legislator choosing its ideal policy - it is constrained by the facts of the case that it is adjudicating. In this respect, my model implicitly uses the case-space approach developed in Kornhauser (1992b) and Kornhauser

(1992a).) I assume that courts are always bound by the doctrine of *stare decisis* - so that the court cannot undo a rule, once set. The court can only create new law in the ambiguous region. In principle, when writing an opinion, the court can extend both sides of the partial legal rule. However, I restrict the court to only amending the side of the law that is necessary to justify the outcome in the case before it. In the above example, the court must revise μ so that it is at a level that is consistent with q being unacceptable. However, it may not revise λ, since doing so can never explain the outcome of the case. (Obviously the court cannot revise up λ such that $\lambda \geq q$, since this would make the ruling inconsistent. Whilst it is consistent to set $\lambda < q$, this merely says that output level q is not definitely acceptable, which does not explain that it is, in fact, unacceptable.)

It is without loss of generality to assume $S(u) \leq \lambda \leq \mu \leq S(l)$. To see why, suppose the court sought to write a rule with $\lambda < S(u)$. Such a rule would guarantee that any output $q \leq \lambda$ will be found acceptable. But all agents know that the court will never discover a case $q < S(u)$ to be inefficiently large - since $S(u)$ is the efficient output level associated with the highest possible value of the external cost θ, given beliefs . The firm could behave as if $\lambda = S(u)$, and this would not affect its decision in any way. Hence, it is without loss of generality to assume $\lambda \geq S(u)$. (A similar argument holds for $\mu \leq S(l)$). The legal rule only begins to have bite when it creates a wedge between the expected penalty implied by the law and the expected penalty implied by beliefs. For example, if the court writes an opinion with $\lambda > S(u)$, then for $q \in (S(u), \lambda]$ the court commits to finding the firm not liable as a matter of law, even though it believes with positive probability that the firm is over-producing. I say that an opinion is *narrow* if the court's opinion sets $\lambda = S(u)$ or $\mu = S(l)$. On the other hand, an opinion is *broad* if the court's opinion sets $\lambda > S(u)$ or $\mu < S(l)$. Broad opinions potentially enable the court to affect the firm's choice by distorting the probability that it expects to be penalized.

To capture the dynamic process of learning, I analyze the choices of both the court and firm, over two periods. The timing of the game is as follows. At period 1, a case q_1 exogenously arises. The court learns and writes an opinion. This specifies the environment for the next period, which is a pair of beliefs (l_1, u_1), a partial rule (λ_1, μ_1) and a penalty F_1. (I assume that this is the court's first opportunity to create law on this issue. Given the above discussion, this is equivalent to assuming that the period 0 legal rule was purely narrow.) The second period environment must satisfy the following: If q_1 was deemed acceptable, then $l_1 = l_0$ and $\mu_1 = \mu_0$, whilst $u_1 = T(q_1)$ and $\lambda_1 \geq q_1$. On the other hand, if q_1 is deemed unacceptable, then $u_1 = u_0$ and $\lambda_1 = \lambda_0$, whilst $l_1 = T(q_1)$ and $\mu_1 \leq q_1$. At period 2, the firm optimally chooses its output level q_2, given this new environment. If $q_2 \notin (\lambda_1, \mu_1)$, then the court mechanically disposes of the case according to the existing law. The court does not investigate the case, and there is no learning. The environment for the following period remains unchanged. If $q_2 \in (\lambda_1, \mu_1)$, then the court investigates the case, learns and writes a new opinion (which satisfy the aforementioned requirements). This generates a new environment, which again is a pair of beliefs (l_2, u_2), a partial rule (λ_2, μ_2), and a penalty F_2. In period 3, the firm chooses its output and the game ends.

The court's objective is to implement the output that minimizes the expected (utilitarian) social deadweight loss. If θ were known, the social utility from producing the efficient level of output is $\frac{1}{2\beta}(\alpha - \theta)^2$. With uncertainty, the per-period net social loss from a given output choice q is:

$$\begin{aligned} & E_\theta\left[\frac{1}{2\beta}(\alpha-\theta)^2\right] - E_\theta\left[(\alpha-\theta)q - \frac{1}{2}\beta q^2\right] \\ = \; & \frac{1}{2\beta}Var[\theta] + \frac{1}{2\beta}(T(q) - E[\theta])^2 \end{aligned} \tag{2.1}$$

In each period, the court seeks to provide incentives for the firm to choose the output level that minimizes the sum of current and future expected social losses. Equation (2.1) demonstrates that two factors contribute to the social deadweight loss. The first factor is the presence of uncertainty, which causes the variance term to be positive. The second factor is any deviation of the chosen output from the *ex ante* efficient level. Since the court issues its ruling after the firm has made its choice, it cannot reduce uncertainty in the current period - although by inducing the firm to experiment, it may cause uncertainty to decrease in the future. In the final period, since there is no further benefit from learning, the court has a strict incentive to implement the *ex ante* efficient output. In the first period, however, it must trade off the benefit from choosing the efficient first period outcome against the benefits of learning in the optimal way for the future.

2.3 Full Information Benchmark

The model of law creation that I described in the previous section embeds a standard externality problem. A well known solution to the problem is to force the firm to internalize the externality by imposing a Pigovian tax equal to the marginal external cost. However, the court does not use this mechanism, and in this paper, has only a blunt tool (in the form of a flat penalty) to provide incentives to the firm. In this section, I verify that a court using the *liable if over-produce* rule can implement the efficient allocation in a full information environment.

Suppose θ is commonly known. The optimal output level is $S(\theta) = \frac{\alpha-\theta}{\beta}$. The court sets a rule whereby the firm is liable if and only if $q > S(\theta)$. The firm chooses output q to maximize its profit:

$$\pi^f = \begin{cases} \alpha q - \frac{1}{2}\beta q^2 & q \leq S(\theta) \\ \alpha q - \frac{1}{2}\beta q^2 - F & q > S(\theta) \end{cases}$$

Proposition 1. *Suppose θ is known and the legal rule holds the firm liable whenever $q > S(\theta)$. Then the firm will choose the efficient output level $q = \frac{\alpha-\theta}{\beta}$ whenever $F \geq \frac{\theta^2}{2\beta}$. If $F < \frac{\theta^2}{2\beta}$, then the firm will produce $q_H = \frac{\alpha}{\beta}$, which is inefficiently large.*

Proposition 1 shows that in the complete information case, the court can induce the efficient output level using the *liable if over-produce* rule, as long as the penalty for over-production is not too low. Moreover, the size of this penalty need not be unreasonably large. Indeed, the 'natural' penalty - the penalty that fully compensates the external party in the event that the firm over-produces - is sufficient to entice the firm to choose the efficient output level. To see this, suppose the firm ignored the fine and over-produced anyway. The best deviation for the firm is to produce $\frac{\alpha}{\beta}$, since this maximizes its pre-penalty profit. The harm to the external party is $\frac{\alpha}{\beta}\theta$. Letting $F_N = \frac{\alpha}{\beta}\theta$, we have:

$$F_N = \frac{\alpha}{\beta}\theta = \frac{\theta^2}{2\beta} \cdot \frac{2\alpha}{\theta} > \frac{\theta^2}{2\beta}$$

since $\alpha > \theta$.

The efficiency of the *liable if over-produce* rule mirrors the efficiency of the negligence rule in the absence of contributory negligence (see Brown (1973) and Cooter et al. (1979), amongst others).

2.4 Firm's Decision

In this section, I consider the firm's optimal decision at each stage in the game with uncertainty. I assume that, in each period, the firm simply chooses the output that maximizes its expected profit in that period - ignoring the effect of its current decision on the environment it may face in the future. This assumption implies that the firm's policy function is time independent - given identical beliefs and rules, the firm will make the same choice in both periods - which simplifies the analysis considerably. (To this extent, I omit time subscripts on the variables in this section.) The assumption is natural in situations where there is a long lived court and a sequence of short-lived firms, each of which makes a decision in only one period. In section 2.6, I extend the analysis to the case of a long lived firm that is strategic in its first period choice.

Suppose the agents have beliefs (l, u) and the court has issued prior opinions (λ, μ) that are consistent (in the sense that $l \leq T(\mu) \leq T(\lambda) \leq u$). Then, the firm's profit function is:

$$\pi(q) = \begin{cases} \alpha q - \frac{1}{2}\beta q^2 & q \leq \lambda \\ \alpha q - \frac{1}{2}\beta q^2 - \Pr[\theta > T(q)] F & q \in (\lambda, \mu] \\ \alpha q - \frac{1}{2}\beta q^2 - F & q > \mu \end{cases}$$

where $\Pr[\theta > T(q)] = \frac{u - T(q)}{u - l}$. The profit function exhibits discontinuities at $q = \lambda$ whenever $\lambda > S(u)$ and at $q = u$ whenever $\mu < S(l)$. The discontinuities reflect the wedge that broad opinions create between the expected penalty implied by the law and the expected penalty implied by beliefs. If $\lambda > S(u)$, the court will not hold the firm liable for producing $q \in (S(u), \lambda]$ even though there is a positive probability that the firm is producing above the efficient level. However, as soon as the firm produces slightly beyond λ, it is no longer immune to the penalty. In fact, the probability of receiving the penalty jumps discontinuously from 0 to the level implied by beliefs at $q = \lambda$. The probability of being penalized similarly jumps discontinuously to one, when the firm produces slightly above μ.

The firm's marginal profit is:

$$\pi_2'(q) = \begin{cases} \alpha - \beta q & q < \lambda \\ \alpha - \beta q - \frac{\beta}{u - l} F & q \in (\lambda, \mu) \\ \alpha - \beta q & q > \mu \end{cases}$$

The marginal expected profit function is piece-wise linear and also has discontinuities at $q = \lambda$ and $q = \mu$. If the firm produces $q \leq \lambda$ or $q > \mu$, then it is either penalized for sure, or not at all - and so the marginal profit is unaffected by the penalty. However, when the firm produces $q \in (\lambda, \mu]$, then a small increase in output increases the probability of being penalized. Since beliefs are uniform, and the penalty is constant, this reduces the expected marginal profit by a constant amount.

Figure 2.1: The firm's optimal decision

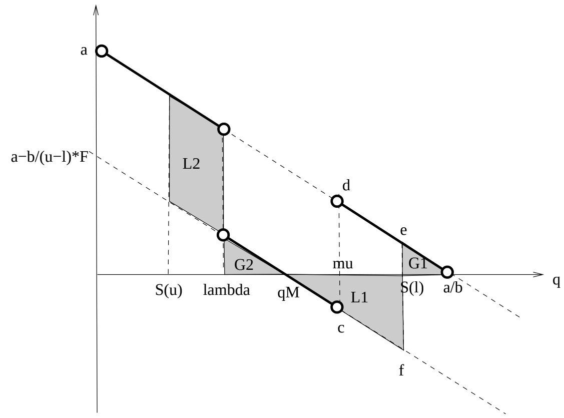


Figure 2.1 illustrates the nature of the marginal profit function for arbitrarily chosen beliefs (l, u), partial rule (λ, μ) and penalty F. The area L_2 indicates the amount that the firm's expected profit falls when output increases beyond λ due to the discontinuous jump in the probability of being penalized. Similarly, the area $cdef$ is the amount that the firm's expected profit falls when output increases beyond μ (due to the probability of being penalized increasing discontinuously to one).

The diagram indicates that there are three candidate solutions for the optimal output - $q = \lambda$, $q = q_M$ and $q = \frac{\alpha}{\beta}$. (Note that even though $\pi'(\lambda) > 0$, the firm may not wish to increase output beyond λ since expected profits fall discontinuously at this output level. Also, since $\pi'(\mu) < 0$, it cannot be optimal to produce at $q = \mu$.) If the firm increases output from $q = \lambda$ to $q = q_M$, then it loses L_2 in expected profits, but gains G_2. Similarly, if the firm increases output from $q = q_M$ to $q = \frac{\alpha}{\beta}$, it gains G_1 in profit but loses L_1. (More precisely, the firm loses parallelogram $cdef$ and $\Delta q_M \mu c$, and then gains $\Delta \mu d \frac{\alpha}{\beta}$. The net effect is summarized by the gain G_1 and loss L_1.) For the scenario illustrated in figure 2.1, the firm prefers q_M to $\frac{\alpha}{\beta}$, since $L_1 > G_1$, and the firm prefers λ to q_M, since $L_2 > G_2$. Hence the firm's optimal choice is $q^* = \lambda$.

The following proposition summarizes the nature of the firm's optimal output decision, given beliefs (l, u), the partial rule (λ, μ), and the penalty F. (The full statement of the firm's equilibrium policy is provided in Section 2.8.1, in the Appendix.)

Proposition 2. *Let the environment be given by beliefs* (l, u)*, a partial rule* (λ, μ) *and a penalty* F*. There exist thresholds* $\overline{\lambda}$*,* $\hat{\mu}(\lambda)$*,* $\underline{F} < F'(\lambda) < \overline{F(\lambda)}$ *, and* $\hat{F}(\mu)$ *such that:*

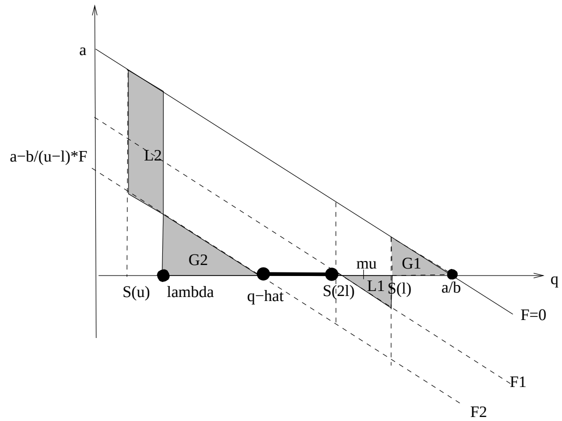
1. *If* $\lambda > \overline{\lambda}$ *or* $u < 2l$*, then the firm will produce* $q_H = \frac{\alpha}{\beta}$ *whenever* $F < F'(\lambda)$ *and* $q = \lambda$ *whenever* $F \geq F'(\lambda)$*.*

2. *If* $\lambda \leq \overline{\lambda}$ *and* $u \geq 2l$*, then the firm will produce* $q_H = \frac{\alpha}{\beta}$ *whenever* $F < \underline{F}$ *and* $q = \lambda$ *whenever* $F > \overline{F(\lambda)}$*. If* $F \in \left[\underline{F}, \overline{F(\lambda)}\right]$*, the firm will produce* $q_M(F) = \frac{\alpha}{\beta} - \frac{F}{u-l}$ *if* $q_M(F) \in (\lambda, \mu)$*. Else it will produce* $q = \mu$*.*

Proposition 2 states the firm will choose its ideal output $q_H = \frac{\alpha}{\beta}$ whenever the penalty is low enough, and will produce the safe output $q = \lambda$ (which is guaranteed to not attract penalty) when the penalty is high enough. For moderate penalties, the firm will experiment by choosing $q_M = \frac{\alpha}{\beta} - \frac{F}{u-l}$ if this quantity is indeed in the ambiguous region (λ, μ), and choose $q = \mu$ otherwise. The intuition behind this result can be seen in figure 2.2 which graphically represents the case of $\lambda \leq S\left(2\sqrt{l(u-l)}\right)$ and $\mu \geq S(2l) > S(u)$. When the penalty is F_1 the firm is indifferent between producing $q_M(F_1) = S(2l)$ and $q = \frac{\alpha}{\beta}$ (since the gains G_1 from increasing output from q_M to $\frac{\alpha}{\beta}$ are exactly balanced by the losses L_1). For any $F < F_1$, the gains would be larger than the losses, so the firm would strictly prefer to produce $\frac{\alpha}{\beta}$. Similarly, the firm prefers $q_M(F)$ to $\frac{\alpha}{\beta}$ whenever $F > F_1$. By a similar argument, the firm is indifferent between producing λ and $q_M(F_2)$ when the penalty is F_2 - since the gains G_2 from producing q_M are exactly balanced by the losses L_2. The firm strictly prefers q_M to λ when $F < F_2$ and strictly prefers λ when $F > F_2$. Hence, the firm produces $\frac{\alpha}{\beta}$ when the penalty is low ($F < F_1$), it produces λ when the penalty is high ($F > F_2$) and it produces $q_M(F)$ when the penalty is moderate ($F \in [F_1, F_2]$).

An important observation is that the firm's supply function is discontinuous in the size of the penalty F. For example, as F increases from zero, the firm's output choice remains

Figure 2.2: The firm's supply function



constant at $\frac{\alpha}{\beta}$ until $F = F_1$, and then it jumps down to $q = S(2l)$. In fact, the firm's output choice is unresponsive to the penalty whenever $F < F_1$ or $F > F_2$; it is only responsive when $F \in [F_1, F_2]$. It is also only in this region that the firm 'experiments' - by choosing an output in the ambiguous region, where the probability of being penalized is uncertain.[6] Experimentation is important, since learning is only possible when the firm experiments.

Intuitively the scope for experimentation is greatest when the court has written the narrowest opinion - since the ambiguous region is largest. By Proposition 2, the firm experiments only when $\lambda \leq \overline{\lambda}$ and if $\mu \geq \hat{q}(\lambda)$. For opinions (λ, μ) satisfying these conditions, the region of experimentation is $\left[S\left(u - \sqrt{u^2 - T(\lambda)^2}\right), \min\{S(2l), \mu\}\right]$. (See Appendix for a derivation of these terms.) Hence an increase in λ (i.e. a broader lower opinion) strictly decreases the scope for experimentation, whilst a decrease in μ (i.e. a broader upper opinion) weakly decreases the scope - and the effect is strict only if $\mu < S(2l)$. Since a narrow upper opinion sets $\mu = S(l)$, for μ to have a strict effect upon the scope for experimentation, the upper opinion must be sufficiently broad.

The next corollary describes the firm's optimal choice, and the scope for experimentation, if the legal rule is perfectly narrow (i.e. if $\lambda = S(u)$ and $\mu = S(l)$).

[6]In this section, I use the terms 'uncertain' and 'ambiguous' in a particular way. Uncertainty exists when the agents' beliefs imply a non-trivial probability that the output chosen lies above the socially efficient level. Ambiguity exists when the legal rule does not prescribe definitely whether the output chosen will incur a penalty or not. Whenever there is ambiguity, there must also be uncertainty - however, the converse need not be true if the court's opinions are broad enough.

Corollary 1. *Suppose the legal rule is perfectly narrow so that* $\lambda = S(u)$ *and* $\mu = S(l)$*. The firm's optimal output choice satisfies:*

- *If* $u < 2l$*, then* $q^* = \frac{\alpha}{\beta}$ *if* $F < \frac{1}{2\beta}u^2$*, and* $q^* = \lambda$ *otherwise. There is no experimentation.*
- *If* $u \geq 2l$*, then* $q^* = \frac{\alpha}{\beta}$ *if* $F < 2l\frac{u-l}{\beta}$*,* $q^* = \lambda$ *if* $F > u\frac{u-l}{\beta}$ *and* $q^* = q_M(F) = \frac{\alpha}{\beta} - \frac{F}{u-l}$ *otherwise. The experimentation region is* $[S(u), S(2l)]$

Corollary 1 shows that the even with the narrowest possible rule, the court may not be able to induce the firm to experiment. Moreover, the scope for experimentation depends upon the extent of uncertainty. If $u < 2l$ (i.e. if the region of uncertainty is small), then marginally increasing output in the uncertain region causes a relatively large increase in the probability of being penalized. This creates a disincentive for the firm to experiment. If the penalty is low enough, it will produce $q_H = \frac{\alpha}{\beta}$. Otherwise, it will produce $q = \lambda$ and not risk being penalized at all. Hence, as the amount of uncertainty narrows, the scope for the court to learn disappears completely. Once the extent of uncertainty is sufficiently small, learning stops entirely.

2.5 Court's Optimal Choice

In the previous section, I showed that the court cannot always induce the firm to choose its desired level of output simply by varying the penalty. Consequently, court opinions that are purely narrow may result in inefficient social outcomes. The following lemma provides conditions under which narrow opinions and appropriately chosen penalties can induce the firm to choose the *ex ante* optimal output.

Lemma 1. *If the court writes narrow opinions,* $\lambda = S(u)$ *and* $\mu = S(l)$*, it can induce the firm to choose the ex ante optimal output* $q = S\left(\frac{u+l}{2}\right)$ *by setting* $F = \frac{u^2 - l^2}{2\beta}$*, provided that beliefs satisfy:* $u \geq 3l$*. If this condition is not satisfied, then there is no penalty that the court can choose to induce optimal behavior.*

Lemma 1 shows that a court that writes narrow opinions can only implement the *ex ante* optimal output if uncertainty about external costs is sufficiently large. If $u < 2l$, experimentation is not possible, so the firm will choose either $q = S(u)$ which is inefficiently low, or $q = \frac{\alpha}{\beta}$ which is inefficiently high. If $2l \leq u < 3l$, then although experimentation is possible, the region of experimentation does not include the efficient allocation.

The next lemma shows that by writing broad opinions, the court may be able to implement outcomes that it would not have been able to implement using narrow opinions alone.

Lemma 2. *Suppose the court can update its lower opinion, λ. For any beliefs, (l, u), and for any feasible upper opinion $\mu \in [S(u), S(l)]$, there exists a pair (λ, F), with $\lambda \in [S(u), \mu]$, that induces the firm to choose output q if and only if $q \in [S(u), \mu] \cup \left\{\frac{\alpha}{\beta}\right\}$. Moreover, any $q \in (\max\{S(u), S(2l)\}, \mu]$ can only be implemented by writing a broad opinion, $\lambda = q$.*

The court can induce the firm to choose any outcome in the ambiguous region, simply by choosing λ and F appropriately. Moreover, if the upper opinion, μ, is not too broad (i.e. if $\mu > S\left(\frac{u+l}{2}\right)$, then the court can always induce the firm to choose the *ex ante* optimum, by writing a broad lower opinion that targets the efficient level. Hence, as the opportunity to learn disappears, the court is best off writing a broad opinion to ensure that the efficient policy is chosen. The next lemma shows that the court's ability to affect the firm's choice using the upper opinion, μ, is far more restricted. Let $\hat{q}$ and $\hat{\mu}$ be as defined in Section 2.8.1:

Lemma 3. *Suppose the court can update its upper opinion, μ, and consider any beliefs, (l, u). If $\lambda \in \left(S\left(2\sqrt{l(u-l)}\right), S(l)\right)$ or $u < 2l$, then the firm will either choose $q = \lambda$ or $q = \frac{\alpha}{\beta}$, and this choice is independent of μ. If $\lambda \in \left[S(u), S\left(2\sqrt{l(u-l)}\right)\right]$ and $u \geq 2l$, then there exists a pair, (μ, F), that induces the firm to choose output q if and only if $q \in [\hat{\mu}(\lambda), S(2l)] \cup \left\{\frac{\alpha}{\beta}\right\}$. Moreover, any $q \in [\hat{\mu}(\lambda), \hat{q}(\lambda))$ can only be implemented by writing a broad opinion $\mu = q$.*

Lemma 3 demonstrates that the court can affect the firm's choice by manipulating μ. However, the lemma also makes clear that the upper opinion μ is a much blunter instrument than the lower opinion λ. To see this, note that by manipulating λ, the court can entice the firm to choose any output in the range $q \in [S(u), \mu]$. The symmetric statement is not true for μ - the court cannot implement any $q \in [\lambda, S(l)]$ simply by manipulating μ. Indeed output in the interval $(\max\{S(u), S(2l)\}, S(l)]$ is never implementable, and if the existing lower opinion is broad (i.e. $\lambda > S(u)$), then output in the region $(\lambda, \hat{\mu}(\lambda))$ is not implementable either. Finally, if the existing lower opinion is narrow (i.e. $\lambda = S(u)$), then any outcome that can be implemented by writing a broad opinion can also be implemented by writing a narrow opinion, and choosing F appropriately. (This stands in contrast to the range of outcomes that only a broad lower opinion can implement, even when the upper opinion is narrow.)

The effect of opinion writing on the firm's choice is asymmetric, and is a consequence of the incentive for the firm to produce output at a level that deviates from the efficient level in a systematic way. Since the lower regulation, λ, is permissive (in the sense that it provides a region of penalty free production), the firm will always find it desirable to produce at least λ units of output. The court can then ensure that the firm does not produce beyond this level by setting the penalty for over-production arbitrarily high. The upper regulation, μ, on the other hand, is restrictive - it extends the region of guaranteed punishment. Unlike with λ, the firm will not always find it optimal to produce at most μ units of output - if the penalty is small enough, it will choose to produce $\frac{\alpha}{\beta}$ and receive the penalty for sure. Moreover, since the firm's supply function is discontinuous, increasing the penalty may cause the firm's output to jump below μ. Hence, the court cannot always use the upper opinion, μ, to target a desired output level in the way that it can use the lower opinion, λ.

2.5.1 Second Period Opinion

Suppose at the beginning of the second period, the firm chooses output level q_2. If $q_2 \leq \lambda_1$ or $q_2 > \mu_1$, then the court mechanically applies the existing rule. If $q_2 \in (\lambda_1, \mu_1]$, then the court must investigate the case. There are two scenarios to consider. In the first case, the court learns that a case q_2 is acceptable. It updates its beliefs so that $u_2 = T(q_2)$ (with l unaffected so that $l_2 = l_1$) and writes an opinion $\lambda_2 \geq q_2$ (with $\mu_2 = \mu_1$ fixed). In the second scenario, the court learns that q_2 is unacceptable. It updates its beliefs so that $l_2 = T(q_2)$ (with $u_2 = u_1$ unaffected) and writes an opinion $\mu_2 \leq q_2$ (holding $\lambda_2 = \lambda_1$ fixed). The question of interest is where the court locates its opinions, λ_2 and μ_2. Since the game ends after the firm's choice in the next period, the court would ideally induce the firm to choose the *ex ante* optimum ($S\left(\frac{u_2+l_2}{2}\right)$). The following propositions outline the court's optimal second period policy:

Proposition 3. *Suppose the court is able to revise its lower opinion λ in the second period:*

1. *If $\mu_2 \geq S\left(\frac{u_2+l_2}{2}\right)$, then the court will implement the efficient output level, $q^* = S\left(\frac{u_2+l_2}{2}\right)$, by choosing either: (i) $\lambda_2 = S\left(\frac{u_2+l_2}{2}\right)$ and F sufficiently high ($F \geq \frac{(u_2+l_2)^2}{2\beta}$ is usually sufficient); or (ii) $\lambda_2 \in \left[S(u_2), S\left(\sqrt{u_2^2 - \left(\frac{u_2-l_2}{2}\right)^2}\right)\right]$ and $F = \frac{u_2^2-l_2^2}{2\beta}$ provided $u_2 \geq 3l_2$.*

2. *If $\mu_2 < S\left(\frac{u_2+l_2}{2}\right)$, then the efficient output cannot be implemented. The court will implement the second best outcome $q = \mu_2$ by choosing (i) $\lambda_2 = \mu_2$ and choosing $F_2 \geq \frac{1}{2\beta}T(\mu_2)^2$; (ii) $\lambda_2 \leq S\left(\sqrt{u_2^2 - (u_2 - T(\mu_2))^2}\right)$ and $F_2 \in \left[\frac{u_2-l_2}{2\beta}\frac{T(\mu_2)^2}{T(\mu_2)-l_2}, \frac{u_2-l_2}{\beta}T(\mu_2)\right]$ provided that $\mu_2 \leq S(2l_2)$; or (iii) $\lambda_2 \in \left(S\left(\sqrt{u_2^2 - (u_2 - T(\mu_2))^2}\right), S\left(\sqrt{\frac{u_2-l_2}{T(\mu_2)-l_2}}T(\mu_2)\right)\right)$ and $F_2 \in \left[\frac{u_2-l_2}{2\beta}\frac{T(\mu_2)^2}{T(\mu_2)-l_2}, \frac{u_2-l_2}{2\beta}\frac{T(\lambda_2)^2-T(\mu_2)^2}{u_2-T(\mu_2)}\right]$.*

If $\mu_2 \geq S\left(\frac{u_2+l_2}{2}\right)$, then the existing upper opinion $\mu_1 (= \mu_2)$ was not written so broadly as to make it impossible to implement the *ex ante* optimal policy. This policy can be implemented in potentially one of two ways. First, the court can simply write a broad lower opinion, setting $\lambda_2 = S\left(\frac{u_2+l_2}{2}\right)$. Then, if the penalty is high enough, the firm will optimally choose to produce λ_2 units of output. Second, the court can choose the penalty optimally (i.e. $F_2 = \frac{u_2^2-l_2^2}{2\beta}$) so that the firm both experiments and chooses the *ex ante* optimal policy.

This is only possible if the region of experimentation is broad enough. If the lower opinion is narrow, then it is sufficient for beliefs to satisfy $u_2 \geq 3l_2$. In fact, as long as the lower opinion is not too broad (i.e. $\lambda_2 \leq S\left(\sqrt{u_2^2 - \left(\frac{u_2-l_2}{2}\right)^2}\right)$), the condition on beliefs suffices. Note that choosing the latter option (where the firm is forced to experiment) requires a lower actual penalty (relative to the former case where the firm is definitely not penalized along the equilibrium path), since the positive marginal probability of being penalized disciplines the firm from over-producing.

If $\mu_2 < S\left(\frac{u_2+l_2}{2}\right)$, then the existing upper opinion μ_1 is too broad and so the court cannot implement the efficient policy. The court will implement the policy that is closest to the optimum - i.e. $q = \mu_2$. Again it can do this either in a brute-force way, by setting $\lambda_2 = \mu_2$ and choosing a penalty high enough, or by writing a narrower opinion and using the penalty to cause the firm to choose $q = \mu_2 > \lambda_2$. Proposition 3 provides conditions under which each of these options are available.

Proposition 4. *Suppose the court is able to revise its upper opinion, μ, in the second period.*

1. *If $u_2 \geq 3l_2$ and $\lambda_2 \leq S\left(\sqrt{2}\frac{u_2+l_2}{2}\right)$, then the court can implement the efficient output level. If $\lambda_2 \leq S\left(\sqrt{u_2^2 - \left(\frac{u_2-l_2}{2}\right)^2}\right)$, it can do this either by setting $F_2 = \frac{u_2^2-l_2^2}{2\beta}$ and $\mu_2 \in \left(S\left(\frac{u_2+l_2}{2}\right), S(l_2)\right]$, or by setting $\mu_2 = S\left(\frac{u_2+l_2}{2}\right)$ and choosing $F_2 \in \left[\frac{1}{\beta}\left(\frac{u_2+l_2}{2}\right)^2, \frac{u_2^2-l_2^2}{2\beta}\right]$; and if $S\left(\sqrt{u_2^2 - \left(\frac{u_2-l_2}{2}\right)^2}\right) < \lambda_2 \leq S\left(\sqrt{2}\frac{u_2+l_2}{2}\right)$, it can do this by setting $\mu_2 = S\left(\frac{u_2+l_2}{2}\right)$ and choosing $F_2 \in \left[\frac{1}{\beta}\left(\frac{u_2+l_2}{2}\right)^2, \frac{1}{\beta}\left(T(\lambda_2)^2 - \left(\frac{u_2+l_2}{2}\right)^2\right)\right]$.*

2. *If $u_2 < 3l_2$ or $\lambda_2 > S\left(\sqrt{2}\frac{u_2+l_2}{2}\right)$, then the efficient output is not implementable. If $u_2 < 2l_2$ or $\lambda_2 > S\left(2\sqrt{l_2(u_2-l_2)}\right)$, the court will implement $q = \lambda_2$ by choosing $F_2 \geq \frac{T(\lambda_2)^2}{2\beta}$; if $2l_2 \leq u_2 < 3l_2$ and $\lambda_2 < S\left(2\sqrt{l_2(u_2-l_2)}\right)$, the court will implement $q = S(2l_2)$ by choosing $F_2 = \frac{2l_2(u_2-l_2)}{\beta}$ and $\mu_2 \in [S(2l_2), S(l_2)]$; and if $S\left(\sqrt{2}\frac{u_2+l_2}{2}\right) < \lambda_2 < S\left(2\sqrt{l_2(u_2-l_2)}\right)$ and $u_2 \geq 3l_2$, the court will either implement $q = \lambda_2$ by choosing F large enough, or set $\mu = \hat{\mu}(\lambda)$ and $F = F_4(\hat{\mu}(\lambda))$.*

Proposition 4 shows that a broad upper opinion is only necessary for efficiency if the existing lower opinion $\lambda_1 = \lambda_2$ is too broad. If the $\lambda_2 = S(u_2)$ is narrow, then it is never strictly necessary to write a broad upper opinion to generate second period efficiency. By contrast, Proposition 3 demonstrates that a broad lower opinion may be necessary even if the existing upper opinion is narrow.

2.5.2 First Period Decision

The court's second period policy is simply aimed at implementing the *ex ante* efficient outcome in the final period, taking beliefs as given. However, in the first period, the court can potentially learn (and hence reduce costly future uncertainty) by enticing the firm to experiment. Recall, the expected social loss from producing output q is $L = Var[\theta] + \frac{1}{2\beta}(T(q) - E[\theta])$. The first term captures the social loss that results from uncertainty, whilst the second term captures the loss from choosing an inefficient outcome (i.e. one that deviates from the *ex ante* social optimum.) I say that the court learns optimally, if it entices the firm to experiment in such a way that minimizes the expected second period social loss from uncertainty.

Lemma 4. *Efficient learning requires experimentation at* $q = S\left(\frac{u+l}{2}\right)$.

Lemma 4 prescribes that efficient learning takes place when the firm experiments at the mean of the belief distribution. Since the cost of uncertainty is strictly convex (it is quadratic in the size of the uncertain region $(u - l)$), the court has a strict incentive to smooth (probabilistically) the size of the uncertain region that results after the acceptability of the first period output level is determined. Intuitively, if experimentation occurs at the mean of the distribution, then learning is greatest in expectation, since the uncertain region is reduced by a half, regardless of the outcome. Moreover, Lemma 4 shows that there is no conflict between the court's desire to efficiently learn and implementing the first period socially efficient allocation. However, as we will see below, the asymmetry in second period outcomes induces the court to choose an inefficiently low level of learning in the first period.

I assume that there are no existing opinions prior to the first period - which is equivalent to assuming that the period 0 partial legal rule is narrowly constructed (i.e. $\lambda_0 = S(u_0)$ and $\mu_0 = S(l_0)$). To avoid having to model the firm's time 0 decision, I assume that a case exogenously arises in the first period. (This avoids needing to specify the firm's beliefs about the expected penalty in period 0 - at a time when the court has not indicated what that penalty will be.) As in the second period - two possibilities arise - either the case is found to be acceptable or not.

Revising from above (μ_1) First consider the case where q_1 is found unacceptable. Let (l_1, u_1) be the new beliefs (where $u_1 = u_0$ and $l_1 = T(q_1)$) and $\lambda_1 = \lambda_0 = S(u_1)$. The court can revise its upper opinion by choosing any $\mu_1 \in [\lambda_1, S(l_1)]$. The following proposition characterizes the court's optimal policy in this environment.

Proposition 5. *Suppose the court is able to revise its upper opinion, μ, in the first period. If $u_1 < 2l_1$, then learning is not possible. The court will implement the static second-best outcome, $q = S(u_1)$, by setting $F_1 \geq \frac{u_1^2}{2\beta}$ and choosing any $\mu_1 \in [S(u_1), S(l_1)]$. If $u_1 \geq 2l_1$, then learning is possible. The court will choose a penalty $F_1^* = \frac{u_1 - l_1}{\beta} T(x^*)$ which induces the firm to experiment by producing output*

$$x^* = \begin{cases} S\left(\frac{8-2\sqrt{10}}{3}u_1 + \frac{2\sqrt{10}-5}{3}l_1\right) & u_1 \geq \frac{11-2\sqrt{10}}{8-2\sqrt{10}}l_1 \\ S(2l_1) & u_1 \in \left[2l_1, \frac{11-2\sqrt{10}}{8-2\sqrt{10}}l_1\right] \end{cases}$$

, and will write an opinion $\mu_1 \in \left[\frac{x^* + S(l_1)}{2}, S(l_1)\right]$.

Proposition 5 has several interesting features. First, it shows that there is no strict incentive for the court to write a broad upper opinion. Since the existing lower opinion is originally narrow (and this implies $\hat{\mu}(\lambda_1) = \lambda_1 = \hat{q}(\lambda_1)$), Lemma 3 implies that any output that can be implemented by writing a broad opinion $\mu_1 < S(l_1)$, can also be implemented by writing a narrow opinion $\mu_1 = S(l_1)$ and choosing the penalty that creates the desired level of experimentation. Since experimentation is desirable, the court will prefer a narrow opinion with experimentation and learning, to a broad opinion.

Second, Proposition 5 shows that - whilst the court will seek to generate experimentation whenever possible - it will not induce efficient learning, even though this maximizes its first period payoff. (This follows since $x^* < S\left(\frac{u_1+l_1}{2}\right)$.) Inefficient learning is a direct result of the asymmetry in the efficacy of the court's upper and lower opinions. Efficient learning requires that the court experiment in the middle of the uncertain region, and this implies that the future court will find the experimental level appropriate or not with equal probability. In the former case, the second period court will be able to revise up its lower opinion, λ, and perfectly target the second period efficient allocation. By contrast, in the latter case, the second period court will be able to revise down its upper opinion, μ, but by Proposition 4, it will be unable to target the efficient allocation. (This is true regardless of period one beliefs, since in this case, $l_2 = \frac{u_1+l_1}{2}$, and so $u_2 < 2l_2$.) Since the social loss in the latter case is much larger than in the former, the court has an incentive to experiment in a way that makes the latter case less likely to arise. Hence, the experimental level chosen in equilibrium will be below the mean of the uncertain region. The court faces a strict trade-off between efficient learning and *ex post* second period efficiency. At the optimum, the court learns 'less' - but is able to use the information it learns more effectively.

Proposition 5 shows that there is no strict incentive for the court to write a broad upper opinion. However, somewhat broad opinions are permissible. This follows because the court is indifferent between rules that generate the same outcome. Since writing a slightly broad opinion does not change the firm's incentive to experiment at the desired level, such a broad opinion is permissible. But note, it is not permissible for the court to write an opinion that is so broad as to constrain future policy making. (This is reflected in the restriction $\mu \geq \frac{x^*+S(l)}{2}$.) If the game had a longer horizon, then the region in which the court has scope to write broad policies will shrink, as it is constrained by the desire to keep policy sufficiently flexible at every period in the future.

Revising from below (λ_1) Now consider the case where q_1 is found to be acceptable. Unlike the previous case, there is now a strict incentive for the court to write a broad opinion. To see why, as above, if the court is able to experiment, then there is a strict incentive to experiment and learn. But, as has been shown, if the court learns that the experimental level is too high, then the best it can do is to implement $q = \lambda_1$ in the following period. Since in that period, the court cannot revise up λ to target the efficient level (it only has the option to revise the upper opinion, μ), then the court would benefit from having preemptively written a broad opinion λ_1 in the first period - anticipating the desire to implement an outcome closer to the efficient level, in the second period. This captures the intuition presented in the introduction - that since the court can only revise opinions in the course of settling genuine disputes, it may write broad opinions as a preemptive tool, to hedge against the risk of not having an opportunity to revise its opinion in the future.

Proposition 6. *Suppose the court is able to revise its lower opinion λ in the first period. If $u_1 < 2l_1$, the learning is not possible, and the court will write a broad opinion $\lambda_1 = S\left(\frac{u_1+l_1}{2}\right)$ and implement this by choosing $F_1 \geq \frac{1}{2\beta}\left(\frac{u_1+l_1}{2}\right)^2$. If $u_1 \geq 2l_1$, learning is possible. The court will induce experimentation at y^* with $S(u_1) < y^* < S\left(\frac{u_1+l_1}{2}\right)$ by choosing $F_1 = \frac{u_1-l_1}{\beta}T(y^*)$ and write a broad opinion $\lambda_1 = S\left(\sqrt{u_1^2 - (u_1 - T(y^*))^2}\right)$.*

The optimal policy y^* is determined by the first order condition (2.4), which is presented in the proof, in the Appendix. Propositions 5 and 6 exhibit many similar features. In both cases, the court induces the firm to experiment, and the experimental level is below the efficient first period level. Furthermore, this implies that there is inefficient learning. The court skews its experimentation in way that causes it to learn less, but which allows it use its more effective policy instrument in the second period with greater probability. Note, however, that learning is less inefficient when the court is able to revise its lower opinion λ_1 in the first period. Since the court writes a preemptively broad opinion, the *ex post* social loss that results in the case that the experimental level is found to be too high, is not as large. As such, the incentive for the court to skew the allocation away from this outcome is not as great.

As proposition 6 shows, a strict application of the minimalist approach can result in sub-optimal outcomes, in which there is significant underproduction of the good. By ignoring the effect of its ruling on future firms' decision making, the minimalist approach can fail to adequately perform the court's role of providing incentives to agents in order to generate efficient outcomes.

2.6 Extensions

2.6.1 Strategic Firm

In the previous sections, I assumed that the firm was myopic in its first period output choice - it simply maximized its expected profit in that period, ignoring the effect of its current choice on the future environment and hence on future expected profits. This assumption may be valid in situations where the firm-like agent engages in the regulated activity for only a short period of time. Of course, in other situations, the assumption is less appealing.

Strategic considerations matter only insofar as experimentation today can affect the environment - and hence the profit making opportunities - that the firm will face in the future. Let $\Pi(q; \lambda, \mu, l, u)$ be the firm's continuation value in the second period if it chooses an output level q in the first period. It was noted in the previous section that there are often (but not always) a range of optimal policies that the court can implement in the second period. Importantly, these policies are not payoff equivalent from the perspective of the firm. For example, the court can induce the firm to choose some output level q' by writing a broad opinion $\lambda = q'$ and setting the penalty high enough. In equilibrium the firm always complies, and so it never pays the penalty. By contrast, the court could induce the firm to experiment at q' by choosing the penalty appropriately. In this case, the firm will be penalized with positive probability - and so its expected payoff is strictly lower. Consequently, the continuation value Π depends upon the firm's beliefs about the court's future strategies.

For concreteness I will consider the case in which the firm believes that the court will always write a broad second period opinion to implement its desired outcome in the final period (i.e. the court will not induce experimentation in the final period). This approach has the benefit that such a strategy is always available (by Lemma 2) since the court can always implement $q = \lambda$ by choosing the penalty large enough. Whilst the alternative assumption - that the court writes narrow opinions and induces experimentation wherever possible - is probably more appealing, the analysis is significantly complicated by the fact that such a strategy is not always available.

Proposition 7. *The scope for first period experimentation is reduced when the firm behaves strategically.*

Proposition 7 shows that the set of output levels that the court can induce the firm to experiment at is strictly smaller when the firm behaves strategically. The intuition is straight forward. If the firm experiments at a high output level, then this output level will most likely be found to be inefficiently large, and the court will induce the firm to choose a much lower output in the future. By contrast, if the firm experiments at a lower output level, then there is a greater likelihood that this output level will be found to be acceptable and that the court will consequently induce the firm to choose a higher output in the future.

The effect of strategic behavior of the firm on the court's policy choice depends on whether the reduction in the scope for experimentation imposes a binding constraint upon the court or not. The court seeks to maximize the social welfare and this goal is independent of whether the firm behaves strategically or not. However, the court can only implement outcomes that are incentive compatible for the firm. As long as the optimal period 2 output level remains incentive compatible, the court will induce the firm to choose this output level. (Of course, with a strategic firm, the penalty that the court will use to induce this output will be lower.

But this does not affect the court's choice, since the court's preference over penalties is purely instrumental.) Hence, the simplification to myopic firms in the main analysis was important only insofar as it affected the set of output choices that the court can induce in the following period. The simplification does not affect the logic of the main results presented in the previous sections.

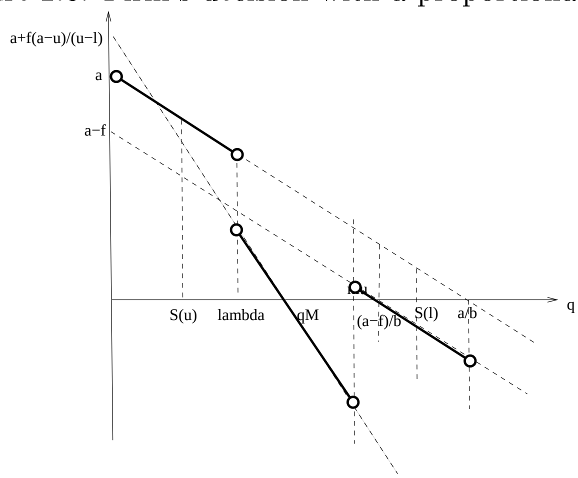
2.6.2 Functional Form

The analysis thus far has relied on specialized (and unrealistic) functional forms for the firm's profit function, the cost function (for the external party), the nature of the penalty and beliefs. In the most part, these assumptions were purely for convenience, to keep the analysis tractable. For example, the choice of a quadratic profit function for the firm generated linear first order conditions, which simplified the analysis. In principle, any continuous, strictly concave function that has a local maximum would suffice. Similarly, the choice of a linear cost function for the external party generated a constant marginal cost. Again, in principle, any continuous and increasing cost function would suffice.

The assumption that beliefs were drawn from a uniform distribution ensured that the beliefs in each period were drawn from the same family of distributions. It is a well known result that the uniform distribution is the only distribution with the property that a posterior generated by truncating the support of the prior, conforms to the prior distribution. Hence, the uniform beliefs assumption prevented the need to re-examine the firm's choice for different classes of posterior beliefs. However, the conformity of prior and posterior beliefs was never crucial to the analysis. In principle, any continuous prior distribution $f(\theta)$ with support (l_0, u_0) will suffice, and this implies a posterior distribution $\frac{f(\theta)}{F(u_t)-F(l_t)}$ for posterior beliefs with support (l_t, u_t).

The assumption of a flat rate penalty, however, is less benign. For example, it should be clear that with a proportional penalty schedule, $F = fq$, the firm will reduce its output for any

Figure 2.3: Firm's decision with a proportional fine



positive penalty (unlike in the above analysis where the firm continued to produce its most desired output, $\frac{\alpha}{\beta}$, for small positive fine.) Nevertheless, even with a proportional penalty schedule, the equilibrium retains many of the same features as the equilibrium in the simple model above. The diagram below shows the firm's marginal expected profit at different levels of output, for a given environment. (With a proportional fine, the marginal expected profit becomes steeper when $q \in (\lambda, \mu)$ because a marginal increase in output both increases both the probability that the firm will be penalized, and the size of the fine.) It should be clear from the diagram that the firm's supply function will exhibit the same discontinuous behavior as in the main model. For f small enough, the firm will produce $q_H = \frac{\alpha - f}{\beta}$ (which is obviously decreasing in f). At some threshold penalty level, the firm's output drops from q_H to $q_M = \frac{\alpha(u-l)+(\alpha-u)f}{\beta(u-l+2f)}$, which (analogous to q_M in the above model) is decreasing in f. Finally, at some higher threshold penalty level, the firm's output drops to $q = \lambda$, where it remains fixed - even if the penalty increases further.

The model was solved with a flat-penalty to keep the analysis tractable. Whilst the quantitative results will be different, since the firm's supply function retains the same qualitative properties, I assert that the main insights of the paper will continue to hold if the court adopts a proportional fine. Indeed, the insights should be robust to any continuous and weakly increasing penalty schedule.

2.7 Conclusion

Should judges always write narrow opinions? What is the value of writing broad opinions, and how does this affect agents' choices and long run efficiency under the common law? I addressed these questions by developing a model of law creation in which the court learns about the ideal legal rule over time, through the cases it hears, and in which agents' choices are responsive to the court's prior decisions. The model is developed in the context of a court that seeks to regulate the behavior of an agent whose actions harm others. Examples include a firm whose production generates pollutions, or law enforcement officials whose behavior intrudes upon the privacy rights of individuals. The model trades off the costs of having the court write broad opinions that potentially entrench inefficient rules, against the uncertainty cost of leaving the law unsettled.

The model generated several predictions. First, I showed that there is an asymmetry in the efficacy of different types of rules; permissive rules, that encourage the agent's behavior, are more effective at affecting the agent's choice than restrictive rules, that seek to limit the agent's actions. This result arises because of the agents' incentive to over-engage in the regulated activity. An important consequence is that the court will optimally write broad permissive opinions - since these are more efficacious at affecting the agent's behavior - and narrow restrictive opinions. To this extent, a judge who always writes narrow opinions will tend to create law that under-provides the regulated activity, and this is inefficient.

Second, I showed that the court will write opinions and set penalties that induce the firm to experiment with its output, in order to learn about the nature of the optimal rule. The model demonstrated the many limitations on court learning that arises when the court faces an agent who rationally responds to the court's prior rulings. Importantly, I showed that learning is possible only insofar as the court can make experimentation incentive compatible for the agent, and that, as a result, experimentation and learning eventually stop, before the court learns the truth. These results demonstrate the importance of modeling the agent's behavior when analyzing the court's decision making process. Moreover, I showed that the court will *not* induce efficient learning (wherein it's learning minimizes the expected future cost of uncertainty). This occurs because efficient learning doesn't guarantee that the court will be able to revise its more efficacious policy instruments in response to the new information that it receives. Rather, the court induces experimentation at a level where it learns 'less', but can more effectively utilize the information that it receives.

Third, I showed the institutional rules that constrain common law courts - the requirement that courts only resolve actual controversies, and the requirement that courts respect precedent - create incentives for the court to pre-emptively write broader permissive opinions than it ideally would, just in case the opportunity to revise that opinion does not arise soon enough in the future.

This paper makes an important contribution towards understanding the complexities of legal decision making, in an environment with imperfect information and when the agents being regulated are rational and risk averse. The model presented captured the tensions between the costs uncertainty on the one hand, and enabling learning on the other, in a tractable way, and provides a useful framework to consider other related problems in judicial politics.

2.8 Appendix

2.8.1 Full Statement of Proposition 2

For each $F \geq 0$, let $q_M(F) = \frac{\alpha}{\beta} - \frac{F}{u-l} = S\left(\frac{\beta}{u-l}F\right)$. Define $F_1 = 2l\frac{u-l}{\beta}$, $F_2(\lambda) = \frac{u-l}{\beta}\left(u - \sqrt{u^2 - T(\lambda)^2}\right)$, $F_3(\lambda) = \frac{1}{2\beta}T(\lambda)^2$, $F_4(\mu) = \frac{u-l}{2\beta}\frac{T(\mu)^2}{(T(\mu)-l)}$ and $F_5(\lambda,\mu) = \frac{u-l}{\beta}\frac{\mu-\lambda}{\mu-S(u)}T\left(\frac{\mu+\lambda}{2}\right)$. Further, let $F_\lambda = \frac{u-l}{\beta}T(\lambda)$ and $F_\mu = \frac{u-l}{\beta}T(\mu)$; and let $\hat{q}(\lambda) = S\left(u - \sqrt{u^2 - T(\lambda)^2}\right)$ and $\hat{\mu}(\lambda) = S\left[\frac{T(\lambda)^2 + T(\lambda)\sqrt{T(\lambda)^2 - 4l(u-l)}}{2(u-l)}\right]$.

Proposition (***Proposition 2***)**.** *If* $\lambda \leq S\left(2\sqrt{l(u-l)}\right)$ *and* $u \geq 2l$*, then:*

- *if* $\mu \geq S(2l)$*, then* $q^* = \begin{cases} \frac{\alpha}{\beta} & F < F_1 \\ q_M(F) & F_1 \leq F \leq F_2(\lambda) \\ \lambda & F > F_2(\lambda) \end{cases}$

- *if* $\hat{q}(\lambda) \leq \mu < S(2l)$*, then* $q^* = \begin{cases} \frac{\alpha}{\beta} & F < F_4(\mu) \\ \mu & F_4(\mu) \leq F < F_\mu \\ q_M(F) & F_\mu \leq F \leq F_2(\lambda) \\ \lambda & F > F_2(\lambda) \end{cases}$

- *if* $\hat{\mu}(\lambda) < \mu < \hat{q}(\lambda)$*, then* $q^* = \begin{cases} \frac{\alpha}{\beta} & F < F_4 \\ \mu & F_4(\mu) \leq F \leq F_5(\lambda, \mu) \\ \lambda & F > F_5(\lambda, \mu) \end{cases}$

- *if* $\mu \leq \hat{\mu}(\lambda)$*, then* $q^* = \begin{cases} \frac{\alpha}{\beta} & F < F_3(\lambda) \\ \lambda & F \geq F_3(\lambda) \end{cases}$

where $\hat{q}(\lambda) \geq \hat{\mu}(\lambda) \geq \lambda$*, with strict inequality whenever* $\lambda > S(u)$ *and equality when* $\lambda = S(u)$*.*

If $\lambda > S\left(2\sqrt{l(u-l)}\right)$ *or* $u < 2l$*, then* $q^* = \begin{cases} \frac{\alpha}{\beta} & F < F_3(\lambda) \\ \lambda & F \geq F_3(\lambda) \end{cases}$.

2.8.2 Proofs of Results

Proof of Proposition 1. Suppose the court holds the firm liable whenever $q > \frac{\alpha-\theta}{\beta}$. The marginal profit is $\pi'(q) = \alpha - \beta q$ whenever $q \neq \frac{\alpha-\theta}{\beta}$, and there is a jump-down discontinuity at $q = \frac{\alpha-\theta}{\beta}$. Hence there are two candidate optimal output choices for the firm: $q = \frac{\alpha}{\beta}$ (which solves $\pi'(q) = 0$) and $q = \frac{\alpha - \theta}{\beta}$. Computing the profit in each case:

$$\begin{aligned}
\pi\left(\frac{\alpha-\theta}{\beta}\right) &= \alpha\left(\frac{\alpha-\theta}{\beta}\right) - \frac{1}{2}\beta\left(\frac{\alpha-\theta}{\beta}\right)^2 = \frac{1}{2\beta}\left(\alpha^2 - \theta^2\right) \\
\pi\left(\frac{\alpha}{\beta}\right) &= \alpha\left(\frac{\alpha}{\beta}\right) - \frac{1}{2}\beta\left(\frac{\alpha}{\beta}\right)^2 - F = \frac{\alpha^2}{2\beta} - F
\end{aligned}$$

Then clearly, $q = \frac{\alpha-\theta}{\beta}$ is optimal whenever $F > \frac{\theta^2}{2\beta}$. □

Proof of Proposition 2 . The optimal output either occurs at a point of discontinuity (in the profit function) or at a level that causes $\pi'(q) = 0$. (The latter is guaranteed to be a local maximum since the marginal profit function is strictly decreasing wherever it is continuous.) Hence, there are 4 candidate solutions: (i) $q = \lambda$, (ii) $q = \mu$, (iii) $q = q_M(F) = \frac{\alpha}{\beta} - \frac{F}{u-l}$, and (iv) $q = \frac{\alpha}{\beta}$. Note further that $q_M(F) = S\left(\frac{\beta}{u-l}F\right)$. Consistency requires that q_M is only a solution if $\lambda < \frac{\alpha}{\beta} - \frac{F}{u-l} < \mu$ - or alternatively, that $F_\mu = \frac{\alpha-\beta\mu}{\beta}(u-l) < F < \frac{\alpha-\beta\lambda}{\beta}(u-l) = F_\lambda$.

The firm's utility at each of the candidate levels of output are: $\pi(\lambda) = \alpha\lambda - \frac{1}{2}\beta\lambda^2$, $\pi(\mu) = \alpha\mu - \frac{1}{2}\beta\mu^2 - \frac{u-T(\mu)}{u-l}F$, $\pi\left(\frac{\alpha}{\beta}\right) = \frac{\alpha^2}{2\beta} - F$, and

$$\begin{aligned}
\pi(q_M) &= \alpha S\left(\frac{\beta}{u-l}F\right) - \frac{1}{2}\beta\left[S\left(\frac{\beta}{u-l}F\right)\right]^2 - \frac{u - T\left(S\left(\frac{\beta}{u-l}F\right)\right)}{u-l}F \\
&= \frac{\alpha^2}{2\beta} + \frac{\beta}{2}\frac{F^2}{(u-l)^2} - \frac{u}{u-l}F
\end{aligned}$$

If $q_M(F) < \mu$, then $\pi(q_M(F)) > \pi(\mu)$. This follows since π is continuous over the interval $[q_M(F), \mu]$ and $\pi'(q) < 0$ over this interval. Hence, if $q_M(F)$ is feasible/consistent, then μ cannot be the optimizer.

First, ignore μ, and consider the firm's optimal choice between λ,$q_M(F)$ and $\frac{\alpha}{\beta}$ (assuming that $q_M(F)$ is feasible). (I refer to this as the baseline analysis.) $q_M(F)$ is preferred to $\frac{\alpha}{\beta}$ if:

$$\begin{aligned} \frac{\alpha^2}{2\beta}+\frac{\beta}{2}\frac{F^2}{(u-l)^2}-\frac{u}{u-l}F &\geq \frac{\alpha^2}{2\beta}-F \\ F &\geq \frac{2l(u-l)}{\beta}=F_1 \end{aligned}$$

Similarly $q_M(F)$is preferred to λ if:

$$\begin{aligned} \frac{\alpha^2}{2\beta}+\frac{\beta}{2}\frac{F^2}{(u-l)^2}-\frac{u}{u-l}F &\geq \alpha\lambda-\frac{1}{2}\beta\lambda^2 \\ F &\leq \frac{u-l}{\beta}\left(u-\sqrt{u^2-T(\lambda)^2}\right)=F_2(\lambda) \end{aligned}$$

There are two cases to consider. Suppose $F_1 \leq F_2$. Then $q^* = \frac{\alpha}{\beta}$ whenever $F < F_1$, $q^* = q_M(F)$ whenever $F \in [F_1, F_2]$, and $q^* = \lambda$ whenever $F > F_2(\lambda)$. (To see this, note that when $F < F_1$, $\frac{\alpha}{\beta}$ is preferred to $q_M(F)$ and $q_M(F)$ is preferred to λ - so by transitivity, $\frac{\alpha}{\beta}$ is the optimal output choice. A similar syllogism verifies the optimality of $q_M(F)$ and λ in the remaining regions.) Suppose instead that $F_1 > F_2$. Then obviously $q^* = \frac{\alpha}{\beta}$ for $F < F_2$ and $q^* = \lambda$ for $F > F_1$. But for $F \in [F_2, F_1]$, $q_M(F)$is dominated by both λ and $\frac{\alpha}{\beta}$. In this case, λ is chosen if:

$$\begin{aligned} \alpha\lambda-\frac{1}{2}\beta\lambda^2 &\geq \frac{\alpha^2}{2\beta}-F \\ F &\geq \frac{1}{2\beta}T(\lambda)^2=F_3(\lambda) \end{aligned}$$

Hence, if $F_1 > F_2$, then $q^* = \frac{\alpha}{\beta}$ whenever $F < F_3(\lambda)$ and $q^* = \lambda$ whenever $F \geq F_3(\lambda)$. (I show below that $F_2 < F_3 < F_1$ whenever $F_2 < F_1$.)

Now, since F_1 is constant and F_2 is decreasing in λ, there must exist some threshold $\hat{\lambda}$ such that $F_1 \leq F_2$ whenever $\lambda \leq \hat{\lambda}$. This requires:

$$\begin{aligned} \frac{2l(u-l)}{\beta} &\leq \frac{u-l}{\beta}\left(u-\sqrt{u^2-T(\lambda)^2}\right) \\ \sqrt{u^2-T(\lambda)^2} &\leq u-2l \\ \lambda &\leq S\left(2\sqrt{l(u-l)}\right)=\hat{\lambda} \end{aligned}$$

provided that $u \geq 2l$. Clearly if $u < 2l$, then the inequality can never be satisfied. It can be easily verified that $F_2 < F_3 < F_1$ whenever $F_2 < F_1$. Suppose $F_2 < F_1$ which implies $T(\lambda)^2 < 4l(u-l)$. Then: $F_3 = \frac{1}{2\beta}T(\lambda)^2 < \frac{2l(u-l)}{\beta} = F_1$. To see that $F_2 > F_3$, suppose not. Then $F_2 \geq F_3$ implies:

$$\begin{aligned} \frac{u-l}{\beta}\left[u-\sqrt{u^2-T(\lambda)^2}\right] &\geq \frac{T(\lambda)^2}{2\beta} \\ \frac{1}{2}\left[\sqrt{u^2-T(\lambda)^2}-(u-l)\right]^2 &\geq \frac{1}{2}l^2 \\ T(\lambda)^2 &\leq 0 \end{aligned}$$

which is a contradiction, since $T(\lambda) \geq l$.

Now, introduce μ into the analysis. Suppose $\lambda \leq \hat{\lambda}$ so that $F_1 \leq F_2$ and $q_M(F)$ is possibly chosen over some range of F. It has been shown that μ is never chosen if $\mu > q_M(F)$. I showed above that $q_M(F)$ is only chosen if $F_1 \leq F \leq F_2$, and so the largest value of $q_M(F)$ that is ever chosen is $q_M(F_1) = S(2l)$. Hence, if $\mu > S(2l)$, μ is never chosen, and the above analysis holds. Similarly, the smallest value of $q_M(F)$ that is ever chosen is $q_M(F_2) = S\left(u-\sqrt{u^2-T(\lambda)^2}\right) = \hat{q}(\lambda)$.

Suppose $\hat{q}(\lambda) \leq \mu < S(2l)$. Then μ is preferred to $\frac{\alpha}{\beta}$ if:

$$\begin{aligned} \alpha\mu-\frac{1}{2}\beta\mu^2-\frac{u-T(\mu)}{u-l}F &\geq \frac{\alpha^2}{2\beta}-F \\ F &\geq \frac{u-l}{2\beta}\frac{T(\mu)^2}{T(\mu)-l}=F_4(\mu) \end{aligned}$$

Note that, since $\mu < S(2l)$, $F_4(\mu) < F_\mu$. (To see this, note that $T(\mu) > 2l$ and so $F_4 = \frac{u-l}{2\beta}\frac{T(\mu)^2}{T(\mu)-l} < \frac{u-l}{2\beta}T(\mu)\frac{2l}{2l-l} = \frac{u-l}{\beta}T(\mu) = F_\mu$.) Furthermore $\hat{q}(\lambda) < \mu$ implies that $F_\mu < F_2(\lambda)$. (To see this, note that $\hat{q}(\lambda) = q_M(F_2) = S\left(\frac{\beta}{u-l}F_2\right)$. Then $S\left(\frac{\beta}{u-l}F_2\right) < \mu$ implies $F_2 > \frac{u-l}{\beta}T(\mu) = F_\mu$.) This implies that, for $F > F_\mu$, $q_M(F)$ becomes feasible again, and the baseline analysis holds. Hence $q^* = \frac{\alpha}{\beta}$ whenever $F < F_4(\mu)$, $q^* = \mu$ whenever $F \in [F_4(\mu), F_\mu)$, $q^* = q_M(F)$ whenever $F \in [F_\mu, F_2(\lambda)]$ and $q^* = \lambda$ whenever $F > F_2(\lambda)$.

Now suppose $\mu < \hat{q}(\lambda)$. This implies $F_\mu > F_2(\lambda)$ and so $q_M(F)$ is never chosen. As with the above analysis, the firm prefers μ to $\frac{\alpha}{\beta}$ if $F \geq F_4(\mu)$. Similarly, μ is preferred to λ if:

$$\begin{aligned} \alpha\mu - \frac{1}{2}\beta\mu^2 - \frac{u-T(\mu)}{u-l}F &\geq \alpha\lambda - \frac{1}{2}\beta\lambda^2 \\ F &\leq \frac{u-l}{\beta}\frac{T(\lambda)-T(\mu)}{u-T(\mu)}T\left(\frac{\mu+\lambda}{2}\right) = F_5(\lambda,\mu) \end{aligned}$$

Again, there are two cases to consider. Suppose $F_4 \leq F_5$. Then $q^* = \frac{\alpha}{\beta}$ whenever $F < F_4(\mu)$, $q^* = \mu$ whenever $F \in [F_4(\mu), F_5(\lambda,\mu)]$, and $q^* = \lambda$ whenever $F > F_5(\lambda,\mu)$. (To see this, note that when $F < F_4$, $\frac{\alpha}{\beta}$ is preferred to μ and μ is preferred to λ - so by transitivity, $\frac{\alpha}{\beta}$ is the optimal output choice. A similar syllogism verifies the optimality of μ and λ in the remaining regions.) Suppose instead that $F_4 > F_5$. Then obviously $q^* = \frac{\alpha}{\beta}$ when $F < F_5$ and $q^* = \lambda$ when $F > F_4$. But for $F \in [F_5, F_4]$, μ is dominated by both λ and $\frac{\alpha}{\beta}$. As above, λ is chosen if $F \geq F_3(\lambda)$. Hence, if $F_4 > F_5$, then $q^* = \frac{\alpha}{\beta}$ whenever $F < F_3(\lambda)$ and $q^* = \lambda$ whenever $F \geq F_3(\lambda)$. (I verify below that $F_5 < F_3 < F_4$ whenever $F_5 < F_4$.)

It remains to describe conditions under which $F_5 \geq F_4$. This requires:

$$\begin{aligned} \frac{u-l}{2\beta}\frac{T(\mu)^2}{T(\mu)-l} &\leq \frac{u-l}{2\beta}\frac{T(\lambda)^2 - T(\mu)^2}{u-T(\mu)} \\ (u-l)T(\mu)^2 - T(\lambda)^2 T(\mu) + lT(\lambda)^2 &\leq 0 \\ T(\mu) &\leq \frac{T(\lambda)^2 + T(\lambda)\sqrt{T(\lambda)^2 - 4l(u-l)}}{2(u-l)} \end{aligned}$$

which implies $\mu \geq \hat{\mu}(\lambda) = S\left[\frac{T(\lambda)^2 + T(\lambda)\sqrt{T(\lambda)^2 - 4l(u-l)}}{2(u-l)}\right]$. (Note that $\lambda \leq S\left(2\sqrt{l(u-l)}\right)$ ensures that the expression under the square root is well defined.) Hence $F_4 \leq F_5$ whenever $\mu \geq \hat{\mu}(\lambda)$. Note that $\lambda \leq \hat{\mu}(\lambda) \leq \hat{q}(\lambda)$ and that the inequalities are strict whenever $\lambda > S(u)$. To see this, note that:

$$
\begin{aligned}
T(\hat{\mu}(\lambda)) &= \frac{T(\lambda)^2 + T(\lambda)\sqrt{(T(\lambda) - 2l)^2 - 4l(u - T(\lambda))}}{2(u-l)} \\
&\leq \frac{T(\lambda)^2 + T(\lambda)(T(\lambda) - 2l)}{2(u-l)} \\
&= T(\lambda)\frac{T(\lambda) - l}{u - l}
\end{aligned}
$$

and that $\frac{T(\lambda)-l}{u-l} \leq 1$, with strict inequality whenever $\lambda > S(u)$.

Suppose $\mu < \hat{\mu}(\lambda)$ so that $F_5 < F_4$. I must verify that $F_5 < F_3 < F_4$. Since $F_5 < F_4$, $T(\lambda)^2 < \frac{(u-l)}{[T(\mu)-l]}T(\mu)^2$, which implies that $F_3(\lambda) < F_4(\lambda)$. Suppose $F_3 \leq F_5$. Then $\frac{T(\lambda)^2}{2\beta} \leq \frac{u-l}{2\beta}\frac{T(\lambda)^2 - T(\mu)^2}{u - T(\mu)}$, which implies $T(\lambda)^2 \geq \frac{(u-l)}{[T(\mu)-l]}T(\mu)^2$, which is a contradiction.

Finally, if $u < 2l$, then $F_\lambda \leq F_1$. The firm must choose between $q = \lambda$ and $q = \frac{\alpha}{\beta}$. By the above arguments, it will choose the former if $F \geq F_3(\lambda)$. □

Proof of Lemma 1. By Corollary 1, we know that experimentation is only possible with perfectly narrow opinions if $u \geq 2l$. (Of course, for broader rules, the scope for experimentation is even smaller.) The region of experimentation is $[S(u), S(2l)]$. Since $q^* = S\left(\frac{u+l}{2}\right) \in (\lambda, \mu)$, to implement the *ex ante* optimal output, the court must induce the firm to experiment. This requires $S\left(\frac{u+l}{2}\right) \in [S(u), S(2l)]$. Hence, $\frac{u+l}{2} \geq 2l$ and so $u \geq 3l$. □

Proof of Lemma 2 . To implement $q \in [S(u), \mu]$, the court can simply set $\lambda = q$ and choose F large enough, so that the firm is induced to choose $\lambda = q$. It is sufficient to choose

$F = \frac{1}{2}\frac{\alpha^2}{\beta}$ - which is the maximum profit the firm can make in the absence of the penalty. With such an F, the firm will never risk a positive probability of receiving the penalty, and hence chooses q_L. To implement $q = \frac{\alpha}{\beta}$, the court can choose any feasible $\lambda \in \left[\frac{\alpha-u}{\beta}, \mu\right]$ and choose F low enough. It is sufficient to choose $F = 0$. $\square$

Proof of Proposition 3. By proposition 2, the court can never induce the firm to choose an output level strictly between μ and $\frac{\alpha}{\beta}$, since if the firm knows it will be fined for sure, it may as well produce the output level that maximizes its pre-penalty profit. Hence if $q^* > \mu$, it is impossible to implement the efficient output. Noting that $\frac{\alpha}{\beta} = S(0)$, the second best solution is to choose $q = \mu$ so long as $S\left(\frac{u+l}{2}\right) - \mu < S(0) - S\left(\frac{u+l}{2}\right) = \frac{u+l}{2\beta}$. But this condition always holds, since $S\left(\frac{u+l}{2}\right) - \mu \leq S\left(\frac{u+l}{2}\right) - S(u+l) = \frac{u+l}{2\beta}$. By Lemma 2, the court can always entice the firm to choose $q = \mu$ by simply setting $\lambda = \mu$ and choosing $F \geq F_3(\lambda) = \frac{1}{2\beta}T(\mu)^2$. Alternatively, if $u \geq 2l$, the court can write a narrower opinion and use the penalty to target $q = \mu > \lambda$. By proposition 2, the firm will choose μ only if $\hat{\mu}(\lambda) < \mu < S(2l)$. Suppose $\lambda = S(u)$, then it immediately follows that $\mu > \hat{q}(\lambda) = \hat{\mu}(\lambda) = S(u)$ and so the narrow opinion will always suffice. In fact, for any $\lambda \leq S\left(\sqrt{u^2 - (u - T(\mu))^2}\right)$, $\mu \geq \hat{q}(\lambda)$, and if so, then choosing $F \in [F_4(\mu), F_\mu]$ will entice the firm to implement μ, where $F_4(\mu) = \frac{u-l}{2\beta}\frac{T(\mu)^2}{T(\mu)-l}$ and $F_\mu = \frac{u-l}{\beta}T(\mu)$. (I need to verify that $\lambda \leq S(2l(u-l))$, but this is implied by $\lambda \leq S\left(\sqrt{u^2 - (u - T(\mu))^2}\right)$ whenever $u \geq 2l$.) For $S\left(\sqrt{u^2 - (u - T(\mu))^2}\right) < \lambda < S\left(\sqrt{\frac{u-l}{T(\mu)-l}}T(\mu)\right)$, $\hat{\mu}(\lambda) < \mu < \hat{q}(\lambda)$, and so choosing $F \in [F_4(\mu), F_5(\lambda, \mu)]$ will entice the firm to implement μ, where $F_5 = \frac{u-l}{2\beta}\frac{T(\lambda)^2 - T(\mu)^2}{u - T(\mu)}$.

Suppose instead that $q^* \leq \mu$. Then it is possible to implement the efficient output. Again, by Lemma 2, the court can always entice the firm to choose q^* by simply setting $\lambda = S\left(\frac{u+l}{2}\right)$ and choosing F appropriately. Now, with $\lambda = S\left(\frac{u+l}{2}\right)$, $\lambda \leq S\left(2\sqrt{l(u-l)}\right)$ requires either $u \leq \left(7 - \sqrt{32}\right)l < 2l$ or $u \geq \left(7 + \sqrt{32}\right)l$. Hence, unless $u \geq \left(7 + \sqrt{32}\right)l$, it suffices to choose $F \geq F_3(\lambda) = \frac{1}{2\beta}\left(\frac{u+l}{2}\right)^2$. In any case, it is always sufficient to choose $F \geq \frac{\alpha^2}{2\beta}$. Alternatively, if $u \geq 3l$, the court can write a narrower opinion and use the penalty to target $q_M(F) = q^*$. Clearly, since $q_M(F) = S\left(\frac{\beta}{u-l}F\right) = S\left(\frac{u+l}{2}\right)$, the court must choose $F = \frac{u^2-l^2}{2\beta}$. (To see why beliefs are restricted to $u \geq 3l$, note that, by proposition 2, choosing q_M requires $q_M = S\left(\frac{u+l}{2}\right) \leq S(2l)$.) Choosing $\lambda \leq S\left(\sqrt{u^2 - (u - T(\mu))^2}\right)$ guarantees that $\hat{q}(\lambda) \leq \mu$. □

***Proof of Proposition** 4.* If $u < 3l$ or $\lambda > S\left(\sqrt{2}\frac{u+l}{2}\right)$, then the efficient allocation is not implementable. If $u < 2l$ or $\lambda > S\left(2\sqrt{l(u-l)}\right)$, then this result follows immediately from Proposition 2, since the firm will choose either λ or $\frac{\alpha}{\beta}$ (neither of which are efficient, except in the special case of $\lambda = S\left(\frac{u+l}{2}\right)$. Now suppose $u \geq 3l$ and $S\left(\sqrt{2}\frac{u+l}{2}\right) < \lambda < S\left(2\sqrt{l(u-l)}\right)$. Since $\lambda > S\left(\sqrt{2}\frac{u+l}{2}\right)$ and $u \geq 3l$, then $S\left(\frac{u+l}{2}\right) < \hat{q}(\lambda)$ and so the court cannot target the efficient output using q_M. Moreover, $S\left(\sqrt{2}\frac{u+l}{2}\right) < \lambda$ implies $S\left(\frac{u+l}{2}\right) < \hat{\mu}(\lambda)$ and so court cannot target implement efficiency using μ. (To see this last point, note that $\hat{\mu}(\lambda)$ is defined such that $F_4(\mu) = F_5(\lambda, \mu)$. Setting $\mu = S\left(\frac{u+l}{2}\right)$ implies $F_4 = \frac{1}{\beta}\left(\frac{u+l}{2}\right)^2$ and $F_5 = \frac{1}{\beta}\left[T(\lambda)^2 - \left(\frac{u+l}{2}\right)^2\right]$, and so $F_4 \leq F_5$ whenever $\lambda \leq S\left(\sqrt{2}\frac{u+l}{2}\right)$.)

Now, if $u < 2l$ or $\lambda > S\left(2\sqrt{l(u-l)}\right)$, then the firm will choose $q \in \left\{\lambda, \frac{\alpha}{\beta}\right\}$. Since $\left|\lambda - S\left(\frac{u+l}{2}\right)\right| < \left|\frac{\alpha}{\beta} - S\left(\frac{u+l}{2}\right)\right|$, implementing $q = \lambda$ is the second best outcome. The court can provide incentives for the firm to choose λ by setting $F \geq F_3(\lambda) = \frac{T(\lambda)^2}{2\beta}$. If $2l \leq u < 3l$, the second best policy is for the court to target $q = S(2l)$. By proposition 2, it can do this by setting $F = \frac{2l(u-l)}{\beta} = F_1$ and $\mu \geq S(2l)$. Finally, if $S\left(\sqrt{2}\frac{u+l}{2}\right) < \lambda < S\left(2\sqrt{l(u-l)}\right)$, then $\hat{\mu}(\lambda) > S\left(\frac{u+l}{2}\right)$. If $\left|\hat{\mu}(\lambda) - S\left(\frac{u+l}{2}\right)\right| < \left|\lambda - S\left(\frac{u+l}{2}\right)\right|$, then court will set $\mu = \hat{\mu}(\lambda)$ and choose the penalty $F = F_4(\hat{\mu}(\lambda)) = F_5(\lambda, \hat{\mu}(\lambda))$ that implements it. If $\left|\hat{\mu}(\lambda) - S\left(\frac{u+l}{2}\right)\right| < \left|\lambda - S\left(\frac{u+l}{2}\right)\right|$, the court will seek to implement $q = \lambda$ and it can do this by choosing F large enough.

Suppose $u \geq 3l$ and $\lambda \leq S\left(\sqrt{2}\frac{u+l}{2}\right)$. Then it is possible for the court to implement the efficient outcome. Then by proposition 2, the court can entice the firm to experiment and choose $q_M(F) = S\left(\frac{u+l}{2}\right)$ by setting $F = \frac{u^2-l^2}{2\beta}$ and $\mu \in \left(S\left(\frac{u+l}{2}\right), S(l)\right]$. (To see this, note that $u \geq 3l$ implies that $S\left(\frac{u+l}{2}\right) \leq S(2l)$ and so $F \geq F_1$. Since $\lambda \leq S\left(\sqrt{u^2 - \left(\frac{u-l}{2}\right)^2}\right)$, then $S\left(\frac{u+l}{2}\right) \geq S\left(u - \sqrt{u^2 - T(\lambda)^2}\right)$ and so $F \leq F_2$. Moreover, since $\mu > S\left(\frac{u+l}{2}\right)$, then $F > F_\mu$ and $\mu > \hat{q}(\lambda)$. Finally, $\lambda \leq S\left(\sqrt{u^2 - \left(\frac{u-l}{2}\right)^2}\right)$ and $u \geq 3l$ implies that $\lambda \leq S\left(2\sqrt{l(u-l)}\right)$. Hence all the conditions for experimentation, as given in proposition 2, are satisfied.) Alternatively, the court could set $\mu = S\left(\frac{u+l}{2}\right)$ and entice the firm to choose $q = \mu$. This requires $\hat{\mu}(\lambda) \leq S\left(\frac{u+l}{2}\right) \leq S(2l)$, which implies that $u \geq 3l$ and $\lambda \leq S\left(\sqrt{2}\frac{u+l}{2}\right)$. (If $\lambda \leq S\left(\sqrt{u^2 - \left(\frac{u-l}{2}\right)^2}\right)$ (i.e. if $S\left(\frac{u+l}{2}\right) \geq \hat{q}(\lambda)$), then any $F \in \left[\frac{1}{\beta}\left(\frac{u+l}{2}\right)^2, \frac{u^2-l^2}{2\beta}\right]$ will implement the efficient allocation. If $S\left(\sqrt{u^2 - \left(\frac{u-l}{2}\right)^2}\right) < \lambda \leq S\left(\sqrt{2}\frac{u+l}{2}\right)$, then any $F \in \left[\frac{1}{\beta}\left(\frac{u+l}{2}\right)^2, \frac{1}{\beta}\left(T(\lambda)^2 - \left(\frac{u+l}{2}\right)^2\right)\right]$ will implement the efficient allocation (again by Proposition 2.) It is easily verified that $\lambda \leq S\left(\sqrt{2}\frac{u+l}{2}\right)$ implies $\lambda \leq S\left(2\sqrt{l(u-l)}\right)$. □

Proof of Proposition 5. First, note that there is no strict incentive to write a broad opinion. Since the lower opinion starts narrow, $\lambda = \hat{\mu}(\lambda) = \hat{q}(\lambda) = S(u)$ - and so, by Lemma 3, any first period outcome that can be generated by a broad opinion can also be generated by a narrow opinion. Moreover, there is no second period benefit to writing a broad opinion, since the court will never choose μ_1 in the second period. (If the experimental level if found to be acceptable, the court will revise up λ to perfectly target the socially efficient outcome. If it is found unacceptable, then the court will be able to revise μ again anyway.)

Next, consider the optimal level of experimentation. If $u \leq 2l$, then by Proposition 2, the firm will choose either $q = \lambda = S(u)$ or $q = \frac{\alpha}{\beta}$. Since $\left|S(u) - S\left(\frac{u+l}{2}\right)\right| < \left|\frac{\alpha}{\beta} - S\left(\frac{u+l}{2}\right)\right|$, the court will implement $q = \lambda$ by choosing a penalty $F \geq \frac{u^2}{2\beta}$ and writing any opinion $\mu \in [S(u), S(l)]$.

If $u > 2l$, the court can induce experimentation in the region $x \in [S(u), S(2l)]$. This output will be found acceptable in the second period with probability $\frac{T(x)-l}{u-l}$, in which case, the court can revise up its lower opinion λ_2 and induce the socially efficient output in the final period. The second period expected social loss from such a policy (relative to the full information optimum) is:

$$L_2^+(x) = = \frac{1}{2\beta}Var[\theta] + \frac{1}{2\beta}(T(q_2) - E_2[\theta])^2 = \frac{(T(x)-l)^2}{24\beta}$$

With probability $\frac{u-T(x)}{u-l}$, x is found to be unacceptable. The new beliefs are $(T(x), u)$ and the court will revise down its upper opinion μ_2. Since the lower opinion remains narrow, there is again no strict incentive for the court to write a broad opinion in the second period. Recall, by Proposition 4 that the court can only implement the efficient allocation if $u \geq 3T(x)$. (Note, this requires $u \geq 6l$ since $T(x) \geq 2l$.) If $u < 2T(x)$, the best the court can do is to implement $q = \lambda = S(u)$. If $2T(x) \leq u < 3T(x)$, the court will implement $q = S(2(T(x))) < S\left(\frac{T(x)+u}{2}\right)$. (Note again, this requires $u \geq 4l$, since $T(x) \geq 4l$.) Hence, the expected second period social loss is given by:

$$L_2^-(x) = \begin{cases} \frac{(u-T(x))^2}{6\beta} & x \leq S\left(\frac{u}{2}\right) \\ \frac{u^2-5uT(x)+7T(x)^2}{6\beta} & S\left(\frac{u}{2}\right) < x < S\left(\frac{u}{3}\right) \\ \frac{(u-T(x))^2}{24\beta} & x \geq S\left(\frac{u}{3}\right) \end{cases}$$

In addition, the expected first period loss from choosing x is: $L_1(x) = \frac{(u-l)^2}{24\beta} + \frac{\left(T(x)-\frac{u+l}{2}\right)^2}{2\beta}$. The court chooses x to minimize the sum of current and future expected losses:

$$\min_{x\in[S(u),S(2l)]} \mathcal{L} = L_1(x) + \frac{T(x)-l}{u-l}L_2^+(x) + \frac{u-T(x)}{u-l}L_2^-(x)$$

Suppose the optimal output x^* satisfies $x^* \leq S\left(\frac{u}{2}\right)$. $\mathcal{L}$ is strictly convex in this region, and so the Kuhn-Tucker conditions are sufficient for optimality. The first order condition is:

$$\frac{\partial \mathcal{L}}{\partial T(x)} = \frac{T(x)-\frac{u+l}{2}}{\beta} + \frac{(T(x)-l)^2}{8\beta(u-l)} - \frac{(u-T(x))^2}{2\beta(u-l)} = 0$$

which implies that $T(x) = \frac{8-2\sqrt{10}}{3}u + \frac{2\sqrt{10}-5}{3}l$. Note that since $T(x) \geq 2l$, this solution is only feasible if $u \geq \frac{11-2\sqrt{10}}{8-2\sqrt{10}}l \simeq 2.79l$. Hence, for $2l \leq u < 2.79$, $x^* = S(2l)$. For $u > \frac{11-2\sqrt{10}}{8-2\sqrt{10}}l$, the optimal net social loss is: $\mathcal{L}(x^*) = \frac{(u-l)^2}{24\beta}\left[\frac{520-160\sqrt{10}}{9}\right]$.

I now show that the optimal solution must be contained in the region $x \leq S\left(\frac{u}{2}\right)$. Suppose the optimal solution is in the region $x \geq S\left(\frac{u}{3}\right)$. Now, in this region, $\frac{\partial\mathcal{L}}{\partial T(x)} = \frac{\left(T(x)-\frac{u+l}{2}\right)}{\beta} + \frac{(T(x)-l)^2}{8\beta(u-l)} - \frac{(u-T(x))^2}{8\beta(u-l)} < 0$ since $T(x) \leq \frac{u}{3}$. Hence the optimum is at $T(x) = \frac{u}{3}$. But $\mathcal{L}\left(S\left(\frac{u}{3}\right)\right) = \frac{5(u-l)^2}{72\beta} + \frac{6ul-5l^2}{36\beta} > \mathcal{L}(x^*)$. Hence, the optimal solution cannot exist in region $x \geq S\left(\frac{u}{3}\right)$. A similar argument shows that it cannot be in the region $S\left(\frac{u}{2}\right) < x < S\left(\frac{u}{3}\right)$ either. □

Proof of Proposition 6. First, using the same argument as in the proof of Proposition 5, the court will optimally experiment in the region $y \leq S\left(\frac{u}{3}\right)$. However, unlike the previous case, there is now a strict incentive for the court to write a broad opinion λ. To see why, if, in the second period, the experimental level y^* is found to be unacceptably high, then by Proposition 4, the best the second period court can do is to implement λ. Writing a broad opinion can bring this second period choice closer to the efficient second period level, than writing a narrow opinion. (Of course, the opinion cannot be so broad as to prevent experimentation.)

The court's problem is:

$$\begin{aligned}\min_{y,\lambda}\mathcal{L} &= \frac{(u-l)^2}{24\beta} + \frac{\left(T(y)-\frac{u+l}{2}\right)^2}{2\beta} + \frac{(T(y)-l)^3}{24\beta(u-l)} + \frac{u-T(y)}{u-l}\left[\frac{(u-T(y))^2}{24\beta} + \frac{\left(T(\lambda)-\frac{u+T(y)}{2}\right)^2}{2\beta}\right] \\ \text{s.t } T(y) &\leq u-\sqrt{u^2-T(\lambda)^2}\end{aligned}$$

Letting ϕ denote the Lagrange multiple, the first order conditions are:

$$\frac{\partial L}{\partial T(y)} = \frac{T(y)-\frac{u+l}{2}}{\beta} + \frac{(T(y)-l)^2}{8\beta(u-l)} - \frac{(u-T(y))^2}{8\beta(u-l)} - \frac{\left(T(\lambda)-\frac{u+T(y)}{2}\right)^2}{2\beta(u-l)} - \frac{u-T(y)}{u-l}\frac{T(\lambda)-\frac{u+T(y)}{2}}{2\beta} - \phi = 0 \quad (2.2)$$

$$\frac{\partial L}{\partial T(\lambda)} = \frac{u-T(y)}{u-l}\frac{\left(T(\lambda)-\frac{u+T(y)}{2}\right)}{\beta} + \phi\frac{T(\lambda)}{\sqrt{u^2-T(\lambda)^2}} = 0 \quad (2.3)$$

Hence, if the constraint does not bind, (2.3) implies that$T(\lambda) = \frac{u+T(y)}{2}$, and so (2.2) implies that $T(y) = \frac{u+l}{2}$. But this solution does not satisfy the constraint. (It is easy to verify that $\frac{u+l}{2} \leq u - \sqrt{u^2-\left(\frac{3u+l}{4}\right)^2}$ only if $-\frac{5}{3}l \leq u \leq l$.) Hence, the constraint binds and so

$T(y) = u - \sqrt{u^2 - T(\lambda)^2}$. (Alternatively, $T(\lambda) = \sqrt{u^2 - (u - T(y))^2}$.)

The court's problem now becomes:

$$\max_y \mathcal{L} = \frac{(u-l)^2}{24\beta} + \frac{\left(T(y) - \frac{u+l}{2}\right)^2}{2\beta} + \frac{(T(y)-l)^3}{24\beta(u-l)} + \frac{u - T(y)}{u-l}\left[\frac{(u-T(y))^2}{24\beta} + \frac{\left(\sqrt{u^2-(u-T(y))^2} - \frac{u+T(y)}{2}\right)^2}{2\beta}\right]$$

and the first derivative is:

$$\frac{\partial \mathcal{L}}{\partial T(y)} = \frac{T(y) - \frac{u+l}{2}}{\beta} + \frac{(T(y)-l)^2}{8\beta(u-l)} - \frac{(u-T(y))^2}{8\beta(u-l)} - \frac{\left(T(\lambda) - \frac{u+T(y)}{2}\right)^2}{2\beta(u-l)} + \frac{u-T(y)}{u-l}\frac{T(\lambda) - \frac{u+T(y)}{2}}{\beta}\left[\frac{u-T(y)}{T(\lambda)} - \frac{1}{2}\right] \tag{2.4}$$

If $T(y) = u$, then $T(\lambda) = u$ and so: $\frac{\partial \mathcal{L}}{\partial T(y)} = \frac{5(u-l)}{8\beta} > 0$. Hence $T(y^*) < u$ and so $y^* > S(u)$. By contrast, $y^* < S\left(\frac{u+l}{2}\right)$, by the same argument as in the proof of Proposition 5. (Since $T(\lambda) > \frac{3u+l}{4}$, the loss when $xy = \frac{u+l}{2}$ is discovered to be too high is larger than the loss when it is found to be acceptable. But these each occur with equal probability.) Hence $S(u) < y^* < S\left(\frac{u+l}{2}\right)$. □

Proof of Proposition 7. Suppose the firm chooses q in the first period. With probability $\frac{T(q)-l}{u-l}$, this policy is found to be acceptable. The court will write a broad opinion at the *ex ante* optimal level $\lambda_2 = S\left(\frac{T(q)+l}{2}\right)$ and set a penalty large enough so that the firm chooses this in the final period. By contrast, with probability $\frac{u-T(q)}{u-l}$, the policy is found to be unacceptable, and the court will implement $q = \lambda_1$ in the following period. Noting that $\alpha q - \frac{1}{2}\beta q^2 = \frac{\alpha^2 - T(q)^2}{2\beta}$, the firm's continuation payoff is:

$$\begin{aligned}\Pi(q; \lambda_1, \mu_1, l_1, u_1) &= \frac{u_1 - T(q)}{u_1 - l_1}\frac{\alpha^2 - T(\lambda_1)^2}{2\beta} + \frac{T(q) - l_1}{u_1 - l_1}\left[\frac{\alpha^2 - \frac{T(q)+l_1}{2}}{2\beta}\right] \\ &= \frac{\alpha^2}{2\beta} - \frac{T(\lambda_1)^2}{2\beta} \cdot \frac{u - T(q)}{u-l} + \frac{T(q)^2 - l_1^2}{4\beta(u-l)}\end{aligned}$$

and

$$\frac{\partial \Pi}{\partial q} = -\frac{T(\lambda_1)^2}{2(u-l)} - \frac{T(q)}{2(u-l)} = -\frac{T(\lambda_1)^2 + T(q)}{2(u-l)} < 0$$

. If the firm experiments, then it will choose q to maximize its stream of current and future payoffs. Hence, it maximizes: $\alpha q - \frac{1}{2}\beta q^2 - \frac{u-T(q)}{u-l}F + \Pi(q)$. The first order condition is:

$$\alpha - \beta q - \frac{\beta F}{u-l} + \frac{\partial \Pi}{\partial q} = 0$$

Recall that the largest experimental level that the firm could be induced to choose was $q = S(2l)$, and it would choose this when $F = F_1 = 2l\frac{u-l}{\beta}$. This output had the property that $\alpha - \beta S(2l) - \frac{\beta F_1}{u-l} = 0$. But the clearly, $\alpha - \beta S(2l) - \frac{\beta F_1}{u-l} + \frac{\partial \Pi}{\partial q} < 0$. □

2.9 References

Aghion, P., P. Bolton, B. Jullien et al., "Learning through price experimentation by a monopolist facing unknown demand," *Mimeo*, 1988.

Baker, Scott and Claudio Mezzetti, "A theory of rational jurisprudence," *Journal of Political Economy*, 2012, *120* (3), 513–551.

Beim, Deborah, "Learning in the Judicial Hierarchy," *Mimeo*, 2012.

Brown, J.P., "Toward an economic theory of liability," *The Journal of Legal Studies*, 1973, *2* (2), 323–349.

Cameron, C.M., J.A. Segal, and D. Songer, "Strategic auditing in a political hierarchy: An informational model of the Supreme Court's certiorari decisions," *American Political Science Review*, 2000, pp. 101–116.

Clark, T. and J. Kastellec, "The Supreme Court and percolation in the lower courts: an optimal stopping model," *Journal of Politics*, 2010.

Cooter, R., L. Kornhauser, and D. Lane, "Liability Rules, Limited Information, and the Role of Precedent," *Bell Journal of Economics*, 1979, *10* (1), 366–373.

Fox, J. and G. Vanberg, "Law as a Discovery Procedure: An Informational Rationale for Broad Judicial Decisions," *Mimeo*, 2011.

Gennaioli, N. and A. Shleifer, "The evolution of common law," *Journal of Political Economy*, 2007, *115* (1), 43–68.

Kornhauser, L.A., "Modeling collegial courts I: Path-dependence," *International Review of Law and Economics*, 1992a, *12* (2), 169–185.

_ , "Modeling collegial courts. II. Legal doctrine," *JL Econ. & Org.*, 1992b, *8*, 441.

Mesquita, E.B. De and M. Stephenson, "Informative precedent and intrajudicial communication," *American Political Science Review*, 2002, *96* (04), 755–766.

Niblett, A., "Case-by-Case Adjudication and the Path of the Law," *Mimeo*, 2010.

Ponzetto, G.A.M. and P.A. Fernandez, "Case law versus statute law: an evolutionary comparison," *The Journal of Legal Studies*, 2008, *37* (2), 379–430.

Rothschild, M., "A two-armed bandit theory of market pricing," *Journal of Economic Theory*, 1974, *9* (2), 185–202.

Siegel, N.S., "A Theory in Search of a Court, and Itself: Judicial Minimalism at the Supreme Court Bar," *Michigan Law Review*, 2005, *103* (8), 1951–2019.

Smith, T., "Reckless Caution: The Perils of Judicial Minimalism," *NYU JL & Liberty*, 2010, *5*, 347–745.

3 Psychological Belief Distortions and Debt

3.1 Introduction

Debt in the United States has increased in every year since 2001. The public's response to this growing debt has varied from complacency (during the pre-recession years) to alarm (since the beginning of the Great Recession in 2008). The differing nature of the public's attitude towards this growing debt has the following curious feature: The public's complacency coincided with periods where standard models imply that debt should optimally fall, whilst their alarm coincided with periods where standard models imply that debt should optimally rise. The public's attitudes towards debt seem to be unusually sensitive to the state of the economy, and responsive in a way that is time inconsistent.

During the boom years between 2001 and 2007, government debt grew by 52%, far exceeding the 36% growth in GDP. (CBO, 2011), (CBO, 2012). During this period, there was little public concern about the size of the debt - epitomized by Vice President Dick Cheney's flippant remark that "deficit's don't matter". Since then, debt has grown even further. By the end of 2012, the Congressional Budget Office estimates that U.S. federal debt held by the public will reach 73% of GDP - about twice the fraction of debt held at the end of 2007. Of course, much of this debt increase can be attributed to lower tax revenues and higher government spending stemming from the severe recession of 2008. For example, in 2009 and 2010, federal tax revenue was about 14.6% of GDP - well below the 40 year average of 18.0%, and the high of 20.6% in 2000 (CBO, 2011).

This latter debt increase has been accompanied by an increased unease amongst the electorate about public debt. A survey of 1,008 registered voters in February 2009 (in the midst of the recession) found that 47% of respondents believed that the federal debt was serious threat to the U.S. economy, and that government needed to act immediately to address the government's deficit and debt problems (PGPF, 2009). Similarly, in a Gallup (2011) poll of

1,018 adults in May 2011, 47% of respondents were against raising the federal debt ceiling, whilst only 19% were in favor (with 34% not knowing enough to say). Responding to this sentiment, 261 members of the House of Representatives voted to amend the U.S. constitution to require the federal government to balance its budget, although the motion fell 23 votes short of the two-thirds majority required for a constitutional amendment.

What makes this anti-debt sentiment strange is that it comes at precisely the time when optimal behavior suggests that the government ought to be accumulating debt. One of the principle roles of debt is to decouple the timing of consumption from the time that the revenue that funds such consumption is generated. This allows agents to enjoy a steady stream of consumption, even if their income stream is volatile. The role for the government to engage in deficit spending during recessions to kick start the economy is a contested one. But even ignoring this contested role, the government's tax smoothing imperative implies that it ought to accumulate deficits during recessions, when revenues are lower and public goods (for example in the form of unemployment insurance, food vouchers etc.) spending is more valuable, and to retire this debt during booms, when the opposite is true. Moreover, this basic insight ought to be well understood by the voting public, even if only from their own experiences of borrowing or drawing down upon savings in tough times, and repaying loans or saving (for a rainy day) during good times.

In this paper, I propose an explanation for this 'anti-debt' phenomenon that draws upon a new insight about the political economy of debt - that voters may be systematically mistaken about the future trajectory of the economy, and this leads them to demand inefficient fiscal policy. For concreteness, consider an economy that can be in one of two states - G and B - indicating periods of booms and slumps or peace and war. I assert that voters systematically over-estimate the likelihood that the current state will persist. This implies that voters are unduly optimistic about the likely duration of booms and unduly pessimistic about the likely duration of recessions.

The idea that psychological factors play a role in business cycles is not a new one. In the

General Theory, Keynes (1936) wrote about the effects of "animal spirits" - of spontaneous optimism or pessimism - on agents' investment decisions. Pigou (1927) similarly noted that agents' 'errors of undue optimism or undue pessimism in their business forecasts were responsible (in part) for fluctuations in industrial activity'. In a 2002 speech, Alan Greenspan, the former Reserve Board chairman stated:

> "The often-repeated pattern in financial markets has been the periodic shift in risk attitudes, initiated by the state of the economy, among lenders and asset holders. History instructs us that, during recoveries and booms, risk discounts erode as the level of optimism lowers barriers to prudence."[7]

The current debate about the size of the U.S. debt relative to GDP and the fraction of GDP accounted for by government spending, provide more recent anecdotes. Much of the concern (as evidenced by comparisons to the historical average level of debt and spending) seems to be that the current levels of debt and spending are unsustainably high. Such a concern might be valid, but it surely relies upon an assumption that debt and spending will continue to remain high for a significant period of time.

Cases of mistaken future beliefs have been documented in the economics and psychology literature. Loewenstein et al. (2003) document that, in a variety of cases, individuals' dynamic behavior exhibits a *projection bias* - they exaggerate the extent to which their future preferences will reflect their current ones. They find that, although individuals correctly perceive the direction in which their tastes will change, they systematically under-estimate the magnitude of this change. In their paper, Loewenstein et al. (2003) give particular attention to the effect of projection bias in a model of life-time consumption, when agents under-estimate the effect of habit formation. Fuster et al. (2011) demonstrate that the projection bias is consistent with decision making by agents who perceive the economy as being governed by a parsimonious model, when in fact the true dynamics are 'hump-shaped'. The authors argue

[7] I wish to acknowledge that I first become aware of these comments upon reading Rotheli (2012).

that since “simple models robustly pick up the short-term momentum in fundamentals but often fail to capture the full extent of long-run mean reversion”, natural expectations will be characterized by a projection bias.

Recent empirical work has also corroborated the claim. Rotheli (2012) shows that banks appear to under-price default risk (i.e. to be over-optimistic) 3 to 5 years into a boom, and overprice risk (i.e. to be unduly pessimistic) during recessions. Evidence from the University of Michigan Consumer Sentiment Index (in which a representative sample of consumers are asked if they except business conditions in 12 months time to be better, the same, or worse) also corroborates the claim. The plurality response in all but one month between January 2006 and February 2012 was “stay same”, and in most months, this option was preferred by a large majority of respondents.

These psychological factors are important because they affect voters’ attitudes towards debt. Voters’ preferences over debt are induced by their preferences over the expected future stream of taxes and spending - and this depends upon their beliefs about the trajectory of the state of the economy. In this paper, I show that differences in beliefs about the future alone are sufficient to generate inefficient fiscal policy - even if both the voters and government are otherwise rational, and have time consistent preferences. To isolate the effect of beliefs about the future as the source of inefficiency, I assume that voters and political parties all share identical intra-period preferences - i.e. they value the immediate costs and benefits of taxation and public goods provision in the same way. However, since they potentially assess the path of future states differently, their continuation preferences - i.e. their attitudes towards debt - will differ.

Consider the world I describe - in which voters believe that the current state will persist for longer than is actually the case. I assume that in the bad state, either public goods become more valuable, or revenue from taxation falls (or both). Defined this way, the bad state can capture a variety of phenomena, including recession times (when public spending in the form of welfare benefits become valuable and taxation revenues fall) and periods of war and strife.

In such a world, the electorate will put political pressure on the government to over-tax and under-accumulate debt during good states (booms), and under-tax and under-accumulate savings during bad states (recessions). To see why, since voters expect the bad state to last for longer, they will expect the government to need to provide higher spending (or generate lower revenues) for longer. Financing this spending purely with debt will cause debt to rise to such a high level, that voters anticipate that taxes will eventually need to rise. To avoid this, the voters demand that the government accumulate less debt, and finance current spending with higher (than optimal) taxes. By contrast, since the voters are optimistic about the duration of the good state, they believe that the government will have a longer time to accumulate assets or to retire the debt incurred from previous bad states. Hence, they will choose lower than optimal taxation, and accumulate assets (or retire debt) insufficiently quickly. This will cause fiscal policy to have a pro-cyclic flavor, even if the voters would ideally choose counter-cyclic policies.

The intuition for this result can be made clear by considering the following stark example. Suppose the government begins with no outstanding debt or assets. Further, suppose voters believe that the current state will persist with probability one, whilst the government knows that it will actually switch in each period with probability one. (I assume that the government and political parties, having access to better data, know the true transition probabilities.) This example satisfies the assumption that voters expect the current state will persist for longer than the government expects it to. It also has the feature that there is no uncertainty in the model from the perspective of either voters or the government, and this simplifies the analysis. Since taxation has convex deadweight cost, all agents will seek to equalize the tax rate across periods. Moreover, since the voters do not expect the state to change, they will expect to choose the same policy (spending level) in each period. Hence, they will choose taxes and spending at levels that balance the budget in every period.[8] The voters will choose

[8] To see that the voters will demand a balanced budget, suppose the government ran a deficit in each period. Then eventually, the debt would reach a maximal level beyond which bond markets will not lend money to the government. At this stage, the government must raise taxes to finance its spending and service

a high tax rate in the bad state (to finance the more valuable public good from a smaller tax base) and a low tax rate in the good state - expecting whichever state they currently experience to persist.

By contrast, the government knows that the state will switch, and that for any two adjacent periods, it will desire higher spending in one and lower spending in the other. It will choose a constant, intermediate tax rate that raises enough revenue to finance spending over the cycle. This requires that it runs a deficit when in the bad state, and run a surplus whilst in the good state. Relative to the government's optimal policy, the voters demand a policy that over-taxes in low states and under-taxes in high states. Moreover, the voters' ideal policy blunts the tax-smoothing role for debt. These differences arise only from the differences in their beliefs about how the state will evolve. If the voters shared the government's beliefs, their ideal policy would correspond to the government's ideal.

The basic force that generates this distortion remains when uncertainty about future states is introduced. However, uncertainty has the additional effect of creating an incentive to precautionarily save. With convex deadweight costs, agents will choose a current tax rate that is higher than expected future taxes. This follows because the utility cost of higher taxes, if the bad state is realized in the future, is larger than the utility gain from lower taxes, if the good state is realized. To hedge against the risk of needing to raise taxes too much in the future, agents will acquire more assets (or issue less debt) in the current period to build a buffer of savings. This incentive exists whenever agents have non-degenerate beliefs about the trajectory of the state. Precautionary savings causes both types of agents to accumulate savings over time, until any future stream of government spending can be financed purely from interest earned on government assets. In this paper, I show that, regardless of the agents' beliefs (so long as they are non-degenerate) - the economy will converge to a long

the accumulated debt. But this predictable increase in taxes cannot be optimal, given the convex deadweight cost of taxation. Similarly, the government should not run surpluses and accumulate assets, since by doing so, it will eventually be able to finance all future spending from the interest on assets alone, and taxes will fall to zero. But again, such a predictable decrease in taxes cannot be optimal.

run steady state in which it has accumulated this maximal stock (the 'full endowment level') of assets, and at which taxation is zero. A fully endowed government is able to perfectly self-insure itself against shocks to the economy, since it can maintain a zero tax rate, regardless of the future sequence of states that are realized. Hence, although the voters' mistaken beliefs causes the government to make inefficient short-run fiscal policy choices, the economy will eventually converge to the same long run steady state as would be the case if the distortion were not present. Fiscal policy remains efficient in the long run. Of course, the path to the long-run equilibrium will differ from the trajectory of the planner's economy, and the nature of this difference will depend on the magnitude of the voters' long-run optimism or pessimism.[9] With sufficiently high long-run optimism, the voter-distorted economy will take longer than the planner economy to converge to the steady state. The converse is true for sufficiently high long-run pessimism. When voters' long run beliefs are characterized by neither significant optimism nor pessimism, then either outcome is possible.

The discussion thus far has been intentionally vague about the nature of political competition and the parties' preferences and incentives. Of course, these factors will affect the extent to which voters' mistaken beliefs distort fiscal policy from its ideal. In this paper, I allow for a broad range of institutional assumptions on the nature of political competition. Suppose there are two political parties with identical preferences. Following Black (1948) and Downs (1957), if the parties are solely office motivated, then, in equilibrium, political competition will drive both parties to choose the voters' ideal policy. In fact, as demonstrated by Wittman (1977) and Calvert (1985), this will remain true even if the parties are partly policy motivated, as long as there is no aggregate uncertainty about the voters' preferences. By contrast, if the parties are partly-policy-partly-office motivated and if there is aggregate uncertainty about the voters' ideal policy, then the political parties will choose a platform that trades off the loss in utility from choosing a platform away from its ideal, against the

[9] I define this explicitly in the following section. Naturally the level of long-run pessimism or optimism will depend on the magnitude of the level of short-run optimism in the good state relative to the pessimism in the bad state.

expected utility gain from being elected with higher probability. In this case, the parties will partially, but not completely internalize the voters' preferences. In this paper, I consider electoral competition of both types.

The latter type of electoral competition generates important modeling questions. Since the parties only partially internalize the voters' preferences, the policies that are chosen depend upon the continuation preferences of both the voters and parties. But these continuation preferences are formed by each type of agent based on their forecast of the future stream of taxes and public goods. Hence, the voters' continuation preferences will depend upon their beliefs about the parties preferences. In this model, this is equivalent to their higher order beliefs about the evolution of the state - in particular whether they are aware that the parties have different beliefs to their own. (This problem, does not arise when the parties completely pander to voters, since then the voters' expectations of future policy will depend solely on their own beliefs about the evolution of the state.)

There are two natural approaches to modeling voter forecasts of future policies. First, the voters may believe that the parties share their beliefs - and so simply assume that the government will choose the voters' ideal policy at every period in the future. Following O'Donoghue and Rabin (1999), I refer to voters of this type as 'naive', since they expect that the future decision maker will share their current preference. Second, the voters may be fully aware that the parties have different beliefs to their own, and forecast future policy accordingly - understanding that the parties will skew the policy away from the voters' ideal and towards the parties' ideal. I refer to voters of this type as 'sophisticates'. In this paper, I demonstrate the existence of equilibria under both approaches. This extends the methodological contribution of Battaglini and Coate[10] to dynamic games with strategic agents who have different continuation payoffs.

To simplify the analysis, I assume there is no learning in this model. Voters never update their beliefs about the transition dynamics, even after observing a long enough history of the state

[10]See for example, Battaglini and Coate (2008),Battaglini and Coate (2011) and Battaglini (2011)

of the economy. This assumption can be partly justified by noting that the voters' beliefs about the transition dynamics may be consistent with the long run empirical distribution of high and low states. (In fact, as I show in Section 2, in spite of having mistaken conditional beliefs, it is possible for the voters' unconditional beliefs to match every moment of the empirical distribution of states in the long run.) Hence, the behavioral assumption has its greatest bite in the short run. In Section 3.7, I consider a simple model of passive learning.

This model stands in contrast to much of the existing literature on the political economy of fiscal policy - which tend to focus on distortions driven by differences in (stage) preferences between the various agents. Persson and Svensson (1989), Alesina and Tabellini (1990) and Tabellini and Alesina (1990), consider models in which political power switches between two political parties who perceive the value of public goods differently. In the Persson and Svensson (1989) framework, if the low-valuation party is currently in power, it has an incentive to 'starve the beast' by under-taxing and ladening the future government with a large debt, to prevent the high-valuation party from providing a large amount of the public good, should it come to power. The inability of current governments to commit to a future path for fiscal policy causes them to choose inefficient policies. (In the Persson and Svensson (1989) framework, the high valuation party also has an incentive to over-accumulate debt, (or to under-accumulate assets) since it knows that the low valuation party will 'raid the kitty' as soon as it comes to power.) Persson and Svensson limit their analysis to a two-period model. Nevertheless, if the above logic is extended to the infinite horizon, debt in the economy will progressively increase until it reaches the government's debt ceiling - which may either be a statutory limit or the natural ceiling imposed by bond markets.[11]

A more recent series of papers, considers the dynamics of fiscal policy when political agents

[11] In their paper, Persson and Svensson consider a broad class of possible party preferences, and show that the 'starve the beast' result obtains for a subclass of preferences. It can be verified that the above result holds for the class of preferences considered in this paper. (Of course, the details need to be modified as appropriate - e.g. the high spending party is assumed to value public goods in the same way as public goods are valued in the bad state in my model, and the low spending party values public goods in the way that agents in my model value public goods in the good state, and there are no technology shocks.)

have an incentive to divert pork spending towards their constituencies. Battaglini and Coate (2008) consider policy made by an n-member legislature, where all agents share common preferences over taxes and public goods, but each agent has an incentive to allocate targeted transfers to their own district. In this model, the winning coalition in the legislature has an incentive to raise general taxes to provide directed transfers to their own constituents. Moreover, the government will never realize a large stock of assets, since the winning coalition always has an incentive to transfer these surplus assets to their own constituents. In a series of papers (Battaglini and Coate (2007), Battaglini and Coate (2011) and Battaglini (2011)) the authors apply this basic model to a number of different settings, including debt outcomes under different electoral systems, and debt over the business cycle. A similar literature (e.g. Acemoglu et al. (2008), and Yared (2010)) considers the long run dynamics in models where politicians can extract rents whilst in office.

In each of these models, the decision makers chose policies that were time-inconsistent, even though each agent had individually time-consistent preferences. The political process distorts choices in a way that makes the decision makers' debt choice inefficient. A similar result obtains in my model - although the source of the endogenous time-inconsistency is the voters' mistaken beliefs, rather than the fact of changing political power in a democracy. The fact that this mechanism is different has important policy implications and long run consequences. As noted, above, the distortion in the existing papers systematically resolves towards the government accumulating too much debt. Consequently, in a model analogous to Persson and Svensson (1989), public debt will increase in each period until it reaches the government's debt ceiling. Similarly, in a model analogous Battaglini and Coate (2008), debt will converge to a non-degenerate distribution, whose support is strictly above the full-endowment level reached in the planner model. By contrast, in this model, the long run debt level still converges to the planner's long run equilibrium, in which the government perfectly self-insures itself against risk, by accumulating a maximal stock of assets.

Hence, although all three models generate endogenous time inconsistent policies, the consequence of this time inconsistency is markedly different in this paper. The different nature of the distortion also suggests important policy implications that differ from the prescriptions in the existing literature. A typical policy prescription to prevent 'starve the beast'-type behavior is to enact limits on the amount of debt that a government can accumulate (or limit the sorts of conditions under which the government may accumulate debt - e.g. only during recessions, but not during booms), such as balanced budget amendments to the constitution, which have been enacted by 49 of the 50 states. However, it should be the clear that such a policy response will not solve the mistaken-voter-induced time-inconsistency problem. In fact, such a policy restriction may exacerbate distortions during recessions, when the government debt is inefficiently small. Solving the belief problem requires implementing mechanisms that amplify the size of government surpluses and deficits in booms and recessions, respectively. Such a mechanism can be implemented, for example, by making the tax code more progressive - thereby increasing the sensitivity of automatic stabilizers to productivity shocks.

The remainder of this paper is organized as follows. In section 2, I present the formal model. In section 3, I characterize the solution to the planner's problem, which in this framework is identical to the political equilibrium in the absence of belief distortions. In section 4, I analyze the effect of political competition on the nature of the parties' platform choices. In section 5, I characterize the political equilibrium assuming that parties are solely office motivated. In section 6 I consider the equilibrium under more general voter and party preferences. Section 7 contains an extension. All proofs are contained in the appendix.

3.2 Model

This model is based on the model presented in Battaglini and Coate (2011). I consider an infinite horizon dynamic model of two-party competition over fiscal policy. The model contains two types of agents - voters and two identical political parties competing for office.

3.2.1 Household/Voter

There are a continuum of identical and infinitely lived households. Each household has stage preferences represented by $c - \frac{1}{1+\varepsilon} l^{\frac{1+\varepsilon}{\varepsilon}}$, where c is consumption, l is time spent working, and $\varepsilon > 0$ is the elasticity of labor substitution. The voter receives a wage w for each hour worked. Households discount the future at rate $\delta \in (0,1)$. There is a competitive bond market, with interest rate $r = \frac{1}{\delta} - 1$. Each household contains one voter.

Since preferences are time separable and quasi-linear, the household's labor supply choice depends only upon the wage rate in that period. Moreover, the household is indifferent as to the timing of consumption. It is without loss of generality to assume that the household consumes his income in each period. Hence, the household's maximization problem in each period is:

$$\begin{aligned} \max_{c,l} \quad & c - \frac{1}{1+\varepsilon} l^{\frac{1+\varepsilon}{\varepsilon}} \\ \text{s.t. } c \quad \leq & \quad wl \end{aligned}$$

By the first order conditions, the labor supply is $l^* = [\varepsilon w]^{\varepsilon}$ and so the household's indirect utility function is given by:

$$u(w) = \frac{\varepsilon^{\varepsilon}}{1+\varepsilon} w^{1+\varepsilon}$$

3.2.2 The government

The government may provide a public good g, which can be produced at cost p. Public goods enter the household preferences separably and contribute utility $A \ln g$. (The choice of the log utility form is purely for tractability. Any increasing and strictly concave function will suffice.) In addition, the government may provide a uniform transfer T to all households.

The value of public goods and labor productivity (and hence wages) varies through time. There are two states of the economy $\theta \in \{G, B\}$. In state θ public goods have value A_θ and wages are w_θ, where $w_G \geq w_B$ and $A_G \leq A_B$, with at least one inequality holding strictly. One can think of the good state as representing times when productivity is high and there is lower need for government spending, such as during booms or peace-time. The bad state corresponds to periods of war or recessions. The state evolves according to a Markov process $\Pi^p = \{p^p_{\theta\theta'}\}$, where $p^p_{\theta\theta'} \in (0,1)$ is the probability that the state will be θ' in the following period, given that the current state is θ. The current state is common knowledge, and the parties are assumed to know the true transition probabilities.

Government spending may be financed by levying a proportional tax on labor τ or by issuing debt in the form of one-period bonds. Let $R_\theta(\tau) = \tau w_\theta l^* = \tau [\varepsilon (1-\tau)]^\varepsilon w_\theta^{1+\varepsilon}$ be the government revenue in state θ, given tax rate τ. Further, let b be the stock of outstanding debt that must be repaid in the current period, and let x be the stock of new debt issued. I assume that $x \in [\underline{x}, \bar{x}]$. The lower bound, $\underline{x} = -\frac{A_L}{r}$, is the amount of assets needed for the government to finance all current and future spending through interest earnings alone, for any sequence of realizations of states. I refer to this as the *full endowment level* of assets. Whilst this lower bound on debt is arbitrary, political parties never have a strict incentive to accumulate more assets - and so the restriction is non-binding. The upper bound $\overline{x} < \frac{\max_\tau R_L(\tau)}{r}$ is the maximum debt that the government may accumulate. $\frac{\max_\tau R_L(\tau)}{r}$ is the natural debt limit imposed by bond markets - it is the amount of debt whose interest the government can just afford to service if it taxes at the peak of the Laffer curve in the bad state. Naturally, creditors will never lend an amount in excess of the government's ability to service its debt. For technical reasons, I assume there is a statutory limit on the government's debt obligations below the natural limit, although I allow the statutory limit to be arbitrarily close to the natural limit.

Let

$$B_\theta(\tau, g, x; b) = R_\theta(\tau) - pg + x - (1+r)b$$

be the government's pre-transfer surplus. The government's budget constraint is given by:

$B_\theta(\tau, g, x; b) \geq 0$. If $B_\theta(\tau, g, x; b) > 0$, then the government is providing positive transfers to the voters.

Given the above discussion, the household's stage utility in state θ is given by:

$$u\left((1-\tau)\,w_\theta\right) + A_\theta \ln g + B_\theta(\tau, g, x; b)$$

where τ is a linear tax on labor income and $T \geq 0$ is a uniform transfer from the government to households. For notational convenience, I write $u_\theta(\tau)$ to mean $u\left((1-\tau)\,w_\theta\right)$ and write $u'_\theta(\tau)$to mean $\frac{\partial u((1-\tau)w_\theta)}{\partial \tau}$.

3.2.3 Voters' Beliefs

Recall that the state evolves according to a Markov process with transition probabilities $\Pi^p = \{p^p_{\theta\theta'}\}$. The main behavioral assumption in this paper is that voters misperceive this process. Instead, they assume that the state evolves according to a different Markov process $\Pi^v \neq \Pi^p$ with $p^v_{GG} \geq p^p_{GG}$ and $p^v_{BB} \geq p^p_{BB}$. Hence, in each state, the voter believes the current state will be more likely to persist (or will persist for longer) than is actually the case. I say that voter are (short-run) optimists if they expect the good state to arise in the immediate future with higher probability - i.e. if $p^v_{\theta G} > p^p_{\theta G}$. By contrast, voters are (short-run) pessimists if they expect the bad state to arise in the immediate future with higher probability. Hence voters are optimistic during booms and pessimistic during recessions. In addition, I assume that $p^i_{GG} > p^i_{BG}$ for each i, and so all agents agree that the next period state is more likely to be G if the current state is also G.

The behavioral assumption is strong in that it claims not only that the voters misperceive the trajectory of the economy, but that they do so in a systematic and state-dependent way. The assumption about the nature and direction of optimism and pessimism is not crucial to generating time-inconsistent fiscal policy or to the properties of the long run equilibrium. It will, however, affect the nature of the distortion - whether the government will over- or under-accumulate debt - and the path to the long-run equilibrium. At various points in the analysis, I describe how the results would differ if the nature of belief distortions were different - for example, if voters were uniformly optimistic.

It is worth noting that the behavioral assumption only requires that the voters' short-run (state-contingent) beliefs are incorrect. However, the voters need not be mistaken about the unconditional probability of each state arising in the long run. Let π_θ^p denote the true probability that the state is θ under the invariant distribution. This is also the empirical long run fraction of periods in which the state is θ. It is easily shown that Π^p has a unique invariant distribution, and that $\pi_G^p = \frac{p_{BG}^p}{p_{BG}^p + p_{GB}^p}$. Similarly, there is a unique invariant distribution of the distorted process, and $\pi_G^v = \frac{p_{BG}^v}{p_{BG}^v + p_{GB}^v}$. I say that the voters are long run optimists if they expect the good state to arise more frequently in the long run - i.e. if $\pi_G^v > \pi_G^v$. They are long run pessimists if the converse is true. The behavioral assumption makes no prediction about the nature of long-run optimism or pessimism. Although the voters expect the high state to persist for longer, once the low state arises, they also expect it to persist for longer. Hence, their forecast of the long-run time spent in the high state may be higher or lower than is actually the case. In particular, if $\frac{p_{BG}^v}{p_{GB}^v} = \frac{p_{BG}^p}{p_{GB}^p}$, then the voters will have correct unconditional long run beliefs, even though their conditional short run beliefs are mistaken. In fact, since in this case, the voters' unconditional beliefs are correct, their beliefs will match each of the moments of long run history. Hence, the model does not require voters to be grossly ignorant. Their beliefs may be perfectly consistent with the observed long run history of the economy.

3.2.4 The political process

There are two political parties with identical preferences. Parties are office motivated and receive utility s if elected. In certain sections of this paper, I also assume that parties may be partly-policy motivated, and seek to implement policies that maximize the voter's utility. (Since voters and parties have different beliefs about the evolution of the state, the parties' perception of the ideal policy may differ from the voters' perception - even when the parties seek to maximize voter utility.) Let γ be the weight on the office motivation component of the parties' utility, and let $1-\gamma$ be the weight on the policy motivation component. In the analysis below, I consider two cases: $\gamma = 1$ and $\gamma \in (0,1)$. (If $\gamma = 0$, then electoral competition has no effect on the platforms that the parties offer.)

I also consider two variant cases regarding the parties' beliefs about the voters' preferences. In the first case, the parties are assumed to perfectly know the voters' policy preferences, and that fiscal policy is the only relevant policy dimension that informs the voters' choices. If voters are indifferent between the parties, then they randomly choose between them with equal probability. In the second case, I assume that, in addition to preferences over policy, voters share a common non-policy preference for one party over the other. (This preference may arise from non-policy attributes of the candidates - such as their charisma or looks - or from other unmodeled policies in the parties' platforms.) This is analogous to the assumptions on voter preferences in Lindbeck and Weibull (1987), Persson and Tabellini (1999) and Battaglini (2011). Let η be the bias of the voters towards party B. This bias is unobservable to the political parties, who treat it as a random variable. Let η have continuously differentiable distribution F, and let the density f be symmetric about 0. I.e. $F(0) = \frac{1}{2}$, $F(-x) = 1 - F(x)$ and $f(-x) = f(x)$. I assume that F is time-independent[12] and that F has the appropriate curvature properties to ensure that the parties' objective functions are concave. (Lindbeck and Weibull (1987) provide sufficient conditions for this to be true.) In this latter case, there is aggregate uncertainty about which party will win the election. Even if both parties offer the same policy, the voters may choose one party over the other, due to the effect of non-policy preferences. However, I assume that the voters are *ex ante* unbiased - and so neither party has an inherent advantage in the electoral game.

Let $v(q)$ be the voter's expected utility if policy $q = (\tau, g, x)$ is implemented. Then, the

[12] In Battaglini (2011), the distribution functions may vary through time - in each period, the distribution is drawn from a meta-distribution over distribution functions, which is time-invariant. I discuss the consequences of this assumption, and my simplification, at the end of Section 4.

voter chooses party A if:

$$v(q_A) > v(q_B) + \eta$$

and so the probability that party A is elected is:

$$\begin{aligned} p^A(q_A, q_B) &= \Pr[v(q_A) > v(q_B) + \eta] \\ &= F(v(q_A) - v(q_B)) \end{aligned}$$

Naturally, the probability that party B is elected is $p^B = 1 - p^A$.

I focus on symmetric Markov perfect equilibria, where parties condition their strategies only upon pay-off relevant variables. In the the context of this model, the relevant state variables are the current debt stock b and the state $\theta \in \{G, B\}$.

3.3 Planner's Problem

To establish a benchmark, I first consider the policies that a benevolent social planner would choose, and analyze the long run characteristics of taxes, spending and debt in the planner's economy. (The results in the section are analogous to the results for the planner economy in Battaglini and Coate (2011).) Let $V_\theta^o(b)$ be the planner's value function when the existing debt level is b and the state of the economy is θ. Let $V^o = (V_G^o, V_B^o)$.

Proposition 8. *The planner's problem has a unique Markovian solution, and is characterized by a pair of continuous, differentiable, strictly decreasing and concave value functions.*

The planner's problem is defined recursively by:

$$\begin{aligned} V_\theta^o(b) &= \max_{\tau,g,x} u(\tau) + A_\theta \ln g + B(\tau, g, x; b) + \delta E[V_{\theta'}^o(x)|\theta] \\ \text{s.t. } B_\theta(\tau, g, x; b) &\geq 0 \\ x &\in [\underline{x}, \overline{x}] \end{aligned}$$

Since V_θ^o is concave, the first order conditions are sufficient for a global maximum. The first order conditions are:

$$-\frac{u_\theta'(\tau)}{R'(\tau)} = 1+\lambda \tag{3.1}$$

$$\frac{A_\theta}{gp} = 1+\lambda \tag{3.2}$$

$$-\delta E\left[\frac{\partial V_{\theta'}^o(x)}{\partial x}\right] \leq 1+\lambda \tag{3.3}$$

where λ is the Lagrange multiplier associated with the budget constraint, and (3.13) holds with strict equality whenever $x < \overline{x}$. In addition, the envelope theorem gives:

$$\frac{dV_\theta^o(b)}{db} = -(1+r)(1+\lambda) \tag{3.4}$$

and so the Euler Equation is:

$$\lambda_\theta^o(b) \geq E\left[\lambda_{\theta'}^o(x_\theta(b))\,|\theta\right] \tag{3.5}$$

since $\delta(1+r)=1$. Note that (3.5) holds with equality whenever $x_\theta(b) < \overline{x}$.

Let $D(\tau) = -\frac{u_\theta'(\tau)}{R_\theta'(\tau)}$ be the deadweight (or shadow) cost of taxation (or alternatively the marginal cost of public funds). $D(\tau)$ is the compensating variation associated with the increase in taxation required to raise revenue by one dollar. It can be easily shown that $D(\tau) = \frac{1-\tau}{1-(1+\varepsilon)\tau}$, which implies both $D'(\tau) > 0$ and $D(0) = 1$. (Note that D does not depend on θ - the deadweight cost of taxation is independent of the worker's productivity.) Two features of the optimal planner policies are immediately apparent. First, the planner never simultaneously taxes and makes transfers. (To see this, note that the planner provides transfers only if the budget constraint doesn't bind - i.e. $\lambda = 0$ - and by (3.1), this implies that $\tau = 0$.) Since taxation is distortionary, it is never efficient to levy taxes merely to return the revenue to households in the form of transfers.

Second, the level of government spending is exactly pinned down by the tax rate, and given by the modified Samuelson rule. Equations (3.1) and (3.12) imply that the optimal level of public goods provision is:

$$g = \frac{1}{D(\tau)} \frac{A_\theta}{p}$$

$\frac{A_\theta}{p}$ is the Samuelson level of the public good - it is the level at which the marginal benefit of the public good is equal to the marginal cost (p). The expression $\frac{1}{D(\tau)} \frac{A_\theta}{p}$ is the *modified* Samuelson level - which takes into account the distortionary costs taxation, rather than just production costs. Since government spending is financed through taxation, the effective marginal cost is not p, but $D(\tau)\, p$ - the utility cost of raising \$$p$ at the margin.

The planner's policy can be in one of two possible regimes - one in which the planner provides transfers and one in which it does not. The optimal policies are characterized in the following Lemma:

Lemma 5. *Let* $b_B^o = \underline{x}$ *and* $b_G^o = \underline{x} + \frac{A_L - A_H}{1+r} > \underline{x}$*. Then, the optimal policies in each state* $\theta \in \{G, B\}$ *are:*

$$\begin{aligned}
\tau_\theta^o(b) &= \begin{cases} 0 & b < b_\theta^o \\ \hat{\tau}_\theta(b) & b \geq b_\theta^o \end{cases} \\
g_\theta^o(b) &= \frac{1}{D(\tau_\theta^o(b))} \frac{A_\theta}{p} \\
x_\theta^o(b) &= \begin{cases} \underline{x} & b < b_\theta^o \\ \hat{x}_\theta(b) & b \geq b_\theta^o \end{cases}
\end{aligned}$$

where $\hat{\tau}_\theta(b)$ *and* $\hat{x}_\theta(b)$ *solve the following system of equations:*

$$\begin{aligned}
R(\tau) - \frac{A_\theta}{D(\tau)} + x &= (1+r)\, b \\
-\delta E_\theta \left[\frac{dV_{\theta'}(x)}{dx} \right] &= D(\tau)
\end{aligned}$$

When the economy is in the transfer regime, the planner sets zero taxes, provides the Samuelson level of the public good, and accumulates the full endowment level of assets $\underline{x}$. Lemma 5 shows that the economy can only be in the transfer regime if the state is G and if the existing debt stock is low enough. Whenever the existing debt stock is large enough (i.e. for any $b > \underline{x}$, if the state is B), then the planner will choose a positive tax rate and provide a level of public goods below the Samuelson level. There are never positive transfers in the bad state.

The policy functions satisfy the properties in the following lemmata:

Lemma 6. *The policy functions are continuous in b. For any $b' > b \geq b^o_\theta$, and for each $\theta \in \{G, B\}$, $\tau^o_\theta(b') > \tau^o_\theta(b)$, $g^o_\theta(b') < g^o_\theta(b')$ and $x^o_\theta(b') \geq x^o_\theta(b)$ (with $x^o_\theta(b') = x^o_\theta(b)$ only if $\theta = B$ and $b = \overline{x}$).*

Lemma 6 states that the planner's policy functions are monotonic in the current debt level. As the debt stock increases, the planner will raise taxes, reduce government spending and increase future debt (if possible).

Lemma 7. *For any $b > \underline{x}$, $\tau^o_B(b) > \tau^o_G(b)$ and $x^o_B(b) \geq b > x^o_G(b)$ (with $x^o_B(b) = b$ only if $b = \overline{x}$).*

Lemma 7 compares the planner's policies across different states of the economy. It shows that, in the bad state, the planner will implement higher taxes and increase debt, whilst in the good state, the planner will implement lower taxes and retire debt. The intuition is straight forward. Since the planner has an incentive to (roughly) smooth taxes over time, and since government spending is more valuable in the bad state, the planner will issue debt when in the bad state, and retire the debt when in the good state. However, since bad states are more likely to follow a current bad state, the planner expects a longer period of high spending following a bad state than a good state. Hence, it must increase taxes during the bad state, to accommodate for the increase in expected lifetime spending by the government.

The Euler Equation implies that, in equilibrium, the deadweight cost of taxation follows a martingale. The planner chooses the tax level in each period to smooth the deadweight cost of taxation over time. The government's debt choice is efficient when the Euler condition is a satisfied. If not, then the government is not efficiently smoothing the costs of taxation. Note, however, that the Euler condition does not imply that the tax rate should follow a martingale - i.e. it is not the case that the current tax rate should equal the expected future tax rate. Since the costs of taxation are increasing and concave, the cost in the bad state is more than proportionally higher than the cost in the good state. It follows that the current tax rate will be larger than the expected future tax rate. The government has an incentive to raise taxes slightly in the current period and accumulate a buffer of savings to partially insure itself against the need to set even higher taxes in the future, should the bad state arise. Hence, although the level of government debt will rise and fall as the economy fluctuates between low and high states, the government will precautionarily save, on average, in each period. In the long run, the government's asset holdings will grow until it reaches the full endowment level of assets. The following proposition characterizes the long run dynamics of the planner's economy:

Proposition 9. *In the long run, the planner's economy will settle at a steady state in which $\tau_\theta = 0$, $g_\theta = \frac{A_\theta}{p}$ and $x_\theta = \underline{x}$. The government will make positive transfers of $T = A_B - A_G$ whenever the state is G.*

Proposition 9 is analogous to proposition 1 in Battaglini and Coate (2011) and results in section III of Aiyagari et al. (2002). In the long run, the planner will accumulate enough assets to finance the government's entire future stream of spending entirely through interest earnings. (This is true for any realization of the stochastic process governing the transition between states.) The tax rate converges to zero, and the level of government spending settles at the Samuelson level in each state. Fiscal policy is efficient.

3.4 Electoral Competition

In this section, I analyze the effect of political competition on the strategic incentives for the political parties. As mentioned in the introduction, I consider several different assumptions on the parties' preferences and the nature of uncertainty surrounding the voters' preferences. Suppose the continuation preferences of the parties and voters (respectively) can be represented by the functions $W_\theta(x)$ and $V_\theta(x)$, respectively. (I show that such functions exist in subsequent sections.) Let $q = (\tau, g, x)$ be some policy, and let $\upsilon_\theta(q)$ and $\omega_\theta(q)$ be the voters' and parties' utility from policy q, respectively. Hence:

$$\begin{aligned} \upsilon_\theta(q) &= u_\theta(\tau) + A_\theta \ln g + B(\tau, g, x; b) + \delta V_\theta(x) \\ \omega_\theta(q) &= u_\theta(\tau) + A_\theta \ln g + B(\tau, g, x; b) + \delta W_\theta(x) \end{aligned}$$

Party i's total utility from offering platform q_i, given that its opponent offers q_{-i}, is:

$$p^i(q_A, q_B)\left[(1-\gamma)\,\omega_\theta(q_i) + \gamma s\right] + \left(1 - p^i(q_A, q_B)\right)(1-\gamma)\,\omega_\theta(q_{-i})$$

where $p^i(q_A, q_B) = G\left[\upsilon_\theta(q_i) - \upsilon_\theta(q_{-i})\right]$ and G is as defined in the proof of Proposition 10. Each party chooses q_i to maximize:

$$p^i(q_A, q_B)\left[(1-\gamma)\left(\omega_\theta(q_i) - \omega_\theta(q_{-i})\right) + \gamma s\right]$$

given their opponent's platform q_{-i}.

The following propositions characterize the nature of the parties' optimal platforms, given the variant assumptions on party preferences and knowledge of voter preferences:

Proposition 10. *Suppose the parties are solely office motivated* $(\gamma = 1)$ *or that there is no aggregate uncertainty. Then, in equilibrium, both parties will offer identical platforms, corresponding to the voters' ideal policy.*

Proposition 10 captures the essence of the median voter theorem (see Black (1948) and Downs (1957)) - that when parties are office-motivated, political competition drives the equilibrium policies towards the ideal policies of the median voter. (In this case, the median voter is the 'representative voter', since all voters are identical.) As Wittman (1977) and others note, the median voter result is robust even to environments in which there is uncertainty about voter preferences, or when the parties have policy preferences. The intuition for the latter case - when parties have are partially policy motivated, but there is no aggregate uncertainty - is worth noting. Since voters and parties have different ideal policies, the parties may not wish to offer the voters' ideal policy. Consider a party and suppose its opponent does not offer the voters' ideal policy. Then the party can always offer a policy that is marginally closer to the voters' ideal policy and win the election for sure. This causes the pandering party's utility to jump discontinuously along the office-motivation dimension, but to only fall marginally along the policy-motivation dimensions. Hence, a profitable deviation of this sort always exists. In the absence of aggregate uncertainty about voter preferences, electoral competition with (partial) office motivation, creates a 'Bertrand competition' like effect. The incentive to 'outbid' the opponent in order to win the election causes the equilibrium policy to converge to the voters' ideal - even when the parties' policy preferences differ from the voters'. The next proposition shows that policy preferences, along with aggregate uncertainty about the voters' preferences, is sufficient to force the equilibrium policy away from the voters' ideal.

Proposition 11. *Suppose that parties are partially policy-motivated* ($\gamma \in (0,1)$) *and that there is aggregate uncertainty about voter preferences. Then, both parties will offer identical platforms, and the equilibrium policy maximizes a weighted sum of the parties' and voters' utility. More precisely, each party maximizes:* $\phi \upsilon_\theta (q) + (1-\phi)\, \omega_\theta (q)$*, subject to the government's budget constraint, where* $\phi = \frac{\gamma s f(0)}{\gamma s f(0) + (1-\gamma) F(0)}$ *is the exogenous weight on the voters' utility.*

Proposition 11 shows that when parties are partially policy oriented and when there is aggregate uncertainty about voter preferences, then political competition drives parties to partially, but not completely internalize voter preferences. Proposition 11 draws upon insights from Calvert (1985) and Lindbeck and Weibull (1987), that uncertainty about voter preferences can cause the median voter to no longer be decisive. (In Lindbeck and Weibull (1987), the 'mean' voter, rather than the median is decisive, which has a parallel to this model, where the equilibrium policy is a weighted mean of the parties' and voters' preferences.) The intuition for why the parties only partially internalize the voters' preferences is straightforward. Unlike the no-aggregate uncertainty case, a deviation by one party to a platform that the voters marginally prefer only marginally increases the probability of being elected, and comes at the cost of marginally reducing the parties policy-payoff. In equilibrium, the parties optimally trade off the increase in the probability of winning against the loss of utility from offering a less-preferred policy.[13] The extent to which the parties internalize the voters' preferences depends on three factors: (i) the density of swing voters ($f(0)$), (ii) the size of office rents (s), and (iii) the weight that parties place on the office-motivation component of their preferences (γ). As each of these factors increase, the marginal gain from accommodating voters is larger, and so the parties will offer a platform closer to the voters' ideal. In particular, as $\gamma \to 1$ (so that parties are solely office motivated), the parties' equilibrium offer converges to the voters' ideal policy.

Electoral competition plays an important role in Battaglini (2011), and at this juncture it is worth understanding the similarities and differences between its role in the two models. In Battaglini (2011), two parties compete for votes in electoral regions which are comprised of one or more (sub)-districts. The government may offer directed transfers at the district level, and there is aggregate uncertainty about voter preferences in each district. Since the marginal utility of transfers is the same in each district, the parties will provide directed transfers only to the district with the most swing voters. Battaglini notes that if the distributions of

[13] I noted, above, that, in the absense of aggregate uncertainty, political competition has a Bertrand competition-like effect. When there is aggregate uncertainty and policy motivation, the relevant analog is to Bertrand competition with heterogeneous goods. In such models, firms choose a price which is a weighted average of the monopoly price and their marginal cost, where the weight is endogenously determined as a function of the substitutability of the goods.

aggregate biases varies across time, then the identity of the district receiving transfers will also change. Anticipating that they will not receive transfers in the future, the voters in the current decisive district demand larger than efficient transfers funded by higher (than optimal) taxes and higher (than optimal) debt. The changing identity of the decisive district generates the inefficiency in fiscal policy.[14] Battaglini further notes that this inefficiency will not arise if either the distributions of the aggregate biases do not change, or if there is only one district in each region - since in either case, the identity of the decisive district will not change through time. If the identity of the decisive district changes, then fiscal policy will be inefficient in both the short and long run, and in particular, the government's accumulation of assets will be bounded away from the full endowment level. By contrast, if the identity of the decisive district does not change over time, then fiscal policy will be efficient in both the short and long run. In this paper, there is only one district, and so the inefficiency arising from directed transfers cannot arise. In addition, the distribution from which the aggregate biases are drawn does not change over time. This assumption is purely made for simplicity and does not bias the results of the model (with belief distortions) towards efficiency. (Since there is only one district, it will be the decisive district, regardless of the nature of the biases.) If time varying biases were introduced, then the parameter ϕ would change through time, and consequently the level of responsiveness of the parties to the voters. This has an intuitive interpretation. If parties expect future electorates to be more responsive to policy (for example, if the next election coincides with a presidential contest), then they know that they will have to pander more to voters in the future - and so will propose more efficient policies in the current period. In the analysis that follows, I show that psychological distortions will cause the parties to choose inefficient short run fiscal policies -

[14] In this sense, the frictions in Battaglini (2011) are very similar to those in Battaglini and Coate's 2008 baseline model. That model does not contain electoral competition, but rather focuses on bargaining over fiscal policy in a legislature with one representative from each of N districts. The inefficiency arises because, in each period, the proposer is randomly selected - and so the current proposer will be unlikely to be the proposer in the future. This creates an incentive for the current proposer to divert more (than optimal) pork towards her district whilst she retains proposal power, by choosing higer (than optimal) taxes and debt. Indeed, one can think of the electoral game in Battaglini (2011) as endogenizing the proposal recognition rule.

but that nevertheless - these policies will converge to the long run efficient equilibrium. This result stands in contrast to Battaglini (2011) in that it shows that the nature of the short run inefficiencies do not create a systematic barrier towards achieving long run efficiency.

On a broader level, the role that electoral competition plays varies between the models. In Battaglini (2011), the parties are purely office motivated, and so in equilibrium they will enact the policies most preferred by some district in the electorate. Electoral competition determines which district's preferences are decisive. In the model presented in this paper, parties may have policy preferences, and if so, electoral competition determines the extent to which the parties choose policies according to their own preferences, and the extent to which they pander to the voters.

Propositions 10 and 11 show that the parties' decision in the game (in which they must best respond to their opponent's strategy) can be reformulated as a decision problem, in which they each maximize a weighted sum of their utility and the voters' utility. In the next section, I characterize the equilibrium policies under both types of electoral competition.

3.5 Political Equilibrium

In this section, I characterize the equilibrium policies and implications for efficiency in the political game where political parties completely pander to voters. In section 6, I show that these results continue to hold in environments in which there is incompletes - and so the results are robust to different specifications of the parties utilities and information about voter preferences.

Suppose electoral competition satisfies the assumptions of Proposition 10, and so is characterized by complete pandering. In equilibrium, the parties offer the voters' ideal policies. Since the voters have time consistent preferences, these coincide with the policies that would be chosen by a planner who had the same beliefs as the voters. The voters' continuation preferences reflect their belief that their ideal policies will be chosen in the continuation game.

This result is made formal in the following Proposition. Let $V_{\theta}^{f}(b)$ be the voters' value function when the existing debt level is b, the state of the world is θ, and competition leads parties to fully pander to voters.

Proposition 12. *Suppose political competition is characterized by complete pandering. There is a unique Markovian equilibrium characterized by a pair of continuous, differentiable, strictly decreasing and concave value functions that satisfy:*

$$\begin{aligned} V_{\theta}^{f}(b) &= \max_{\tau,g,x} u_{\theta}(\tau) + A_{\theta}\ln g + B_{\theta}(\tau,g,x;b) + \delta E_{\theta}^{v}\left[V_{\theta'}^{f}(x)\right] \\ s.t. &\quad B_{\theta}(\tau,g,x;b) \geq 0 \text{ and } x \in \{\underline{x},\overline{x}\} \end{aligned}$$

3.5.1 The Short Run

The equilibrium policies have the same qualitative properties as the socially efficient policies. The equilibrium policy functions are monotonic, and taxes and debt are higher in the bad state relative to the good state. The government will accumulate debt in the bad state (i.e. $x_B(b) > b$) and retire debt in the good state ($x_G(b) < b$). Moreover, the government never simultaneously taxes and makes transfers, and so it only makes positive transfers after accumulating the full endowment level of assets. However, the government's policies are not efficient, since they are chosen assuming an incorrect specification of the dynamics of the economy. Since the policies have same qualitative features, belief distortions affect only the magnitudes, but not the direction, of government policy responses to economic shocks. As will become evident in the next subsection, this will have important implications for the long run efficiency of fiscal policy.

Although the equilibrium policies are qualitatively similar to the policies chosen by a social planner, they are not dynamically efficient, since the policies are determined according to the wrong beliefs. The equilibrium policies satisfy $E^{v}\left[D\left(\tau_{\theta'}^{f}\left(x_{\theta}^{f}(b)\right)\right)|\theta\right] \leq D\left(\tau_{\theta}^{f}(b)\right)$, where the inequality is strict only if $x^f_\theta(b) = \overline{x}$. This implies:

$$E^p_\theta\left[D\left(\tau^f_{\theta'}\left(x^f_\theta(b)\right)\right)\right] + (p^v_{\theta B} - p^p_{\theta B})\left[D\left(\tau^f_B(x_\theta(b))\right) - D\left(\tau^f_G(x_\theta(b))\right)\right] \leq D\left(\tau^f_\theta(b)\right) \quad (3.6)$$

Equation (3.6) demonstrates that the belief distortion creates a wedge between the current cost of taxation and the (true) expected future cost. The size of this wedge is increasing in the size of the belief distortion, and the difference in the costs of taxation in the two states of the economy. Moreover, the direction that the wedge pushes current taxes from expected future taxes differs between the two states, as outlined in the following Lemma:

Lemma 8. *In the political equilibrium, whenever $b > b^*_\theta$, the government will over-accumulate debt (or under-accumulate savings) in the good state and under-accumulate debt in the bad state .*

Lemma 8 shows that fiscal policy is no longer efficient when political competition skews the policies that the government implements. In the good state, both parties choose platforms with a tax rate whose deadweight cost is below the expected future cost of taxation. The tax rate in the good state is inefficiently low and debt is inefficiently high. Given the expected continuation policies, the government could do better by choosing a higher tax rate in the current period and accumulating savings to fund future spending without needing to raise taxes by as much. By contrast, in the bad state, each party chooses a platform with an inefficiently high level of taxation. Again, the government could do better by choosing a lower tax rate and accumulating more debt, since even with a lower tax rate, it will be able to service this debt when the next good state arrives.

The intuition for why voters demand inefficient policy is straightforward. In the good state, the voters want the government to accumulate an asset buffer, to pay for higher government spending in the bad state. However, since the voters are over-confident about the duration of the good state, they will accumulate fewer assets in each period - since they expect to have longer to accumulate the buffer. By contrast, since they are pessimistic about the duration of bad state, when the bad state arrives, voters expect that the government will need to provide higher spending for longer. Anticipating a much larger increase in the present value of government spending, the voters will demand higher (than optimal) taxes, instead of issuing debt.

Fiscal policy appears to be chosen by an agent with time-inconsistent preferences. However, similar to the other papers cited in the introduction, time inconsistency is not hardwired into agents' preferences. All of the agents in this model have perfectly time consistent preferences. The distortion in beliefs causes time inconsistency to arise as an endogenous feature of the model, even though at every stage, the government actors who chose policies have correct beliefs.

Although taxes are inefficiently low in the good state when voters are optimistic, it does not follow that the equilibrium tax rate is below the optimal tax chosen by the planner. Similarly, although taxes are inefficiently high in the bad state when voters are pessimistic, it need not be the case that the equilibrium tax rate is higher than the optimal taxes chosen by the planner. Dynamic efficiency of fiscal policy is not assessed by comparison of the level of taxes and debt to the level under the planner policies. Rather, efficiency is measured by the fit of current policies to the continuation policies, given true beliefs about the trajectory of the state of the economy. The inefficiency is revealed in more volatile taxes and less volatile debt in the distorted economy than in the planner's economy.[15]

The following Lemma makes this point clear:

Lemma 9. *There exist planner and voter beliefs under which equilibrium taxes are strictly lower (higher) than planner taxes in both states, and for all debt level.*

I first depart from the above framework slightly, and consider a world in which voters are either (weakly) optimistic in both states, or (weakly) pessimistic in both states. Recall, the

[15]This suggests a method to empirically test the predictions of this model, by comparing the observed volatility in taxes and debt, against the ideal levels predicted by a calibrated model.

voters exhibit short-run optimism in state θ if they assign a higher probability to the good state arising in the immediate future than the parties do (i.e. $p^v_{\theta G} > p^p_{\theta G}$).

The intuition for this Lemma can be most easily seen by considering the following simple case: Suppose the voter is optimistic in the good state, but has correct beliefs in the bad state. In the good state, the parties will naturally choose an inefficiently low tax rate, by the same logic as above. Since they have an incentive to smooth taxes in the other state (where they have correct beliefs), voters must also choose lower taxes than the planner in that state - even though there is no belief distortion. Optimism in one state of the economy causes the equilibrium tax rate to be strictly lower in *both* states.

Moreover, since this result is strict, it is robust to introducing small amounts of pessimism in the other state. Hence, even in the state where voters are pessimistic (and hence demand taxes that are inefficiently high), it is possible for equilibrium taxes to be lower than the planner's optimal tax rate. To understand why, note that when the degree of optimism in one state is much larger than the degree of pessimism in the other, then voters will be long-run optimists. (Although they expect the good state to arise less frequently following a bad state, they expect the good state to arise more frequently in the long run.) In the state where voters are pessimistic, there are two countervailing forces at play. On the one hand, they desire higher taxes to optimally save for the bad state which they expect to arise with greater probability in the immediate future. On the other hand, they desire lower taxes since they expect the good state to arise with greater probability farther into the future. When the degree of short run pessimism is small relative to the degree of long run optimism, then the second effect (which causes voters to demand lower taxes) dominates the first.

Conjecture 1. *If voters have correct long-run beliefs ($\pi^p_G = \pi^v_G$), then $\tau_G(b) < \tau^o_G(b)$ and $x_G(b) > x^o_G(b)$ whenever $b > b^*_G$, and $\tau_B(b) > \tau^o_B(b)$ and $x_B(b) < x^o_B(b)$ whenever $x_B(b) < \overline{x}$.*

Intuitively, if there is no long run optimism or pessimism (i.e. if the voters' unconditional beliefs are correct), then the second effect will not arise. The voters will seek to smooth taxes over the long run in the same way as the planner does, and so no long-run distortion arises. However, since voters exhibit short-run optimism in the good state, they will demand lower taxes (than the planner would ideally choose) in the good state. Similarly, since they exhibit short-run pessimism in the bad state, the voters will demand higher taxes (than the planner would ideally choose) in the bad state. Moreover, if the conjecture is true, then the qualitative results should be robust to moderate levels of long-run optimism or pessimism - since the equilibrium is continuous in the probabilities.

3.5.2 Long Run

It was noted, above, that the equilibrium policies inherit all of the qualitative properties of the planner properties. In particular, unless it is already fully endowed, the government will always accumulate assets or retire debt in the good state. Although psychological distortions to beliefs affect the size of this asset accumulation, they do not create a structural barrier that prevents asset accumulation. This feature of the equilibrium policies distinguishes this model from many of the other models discussed in this paper.

To see this, I compare the policies chosen by the actors in this model when the government is already fully endowed, against the policies chosen in the 'starve the beast' model and in typical variants of the Battaglini and Coate models. In this model, by Lemma 5, once it is fully endowed (i.e. $b = \underline{x}$), the government will ensure that this level of assets is maintained, and will only provide transfers in the good state of the economy, when surplus funds are available. This is true in spite of the dynamic inefficiency in the government's decision making. By contrast, in any of the Battaglini and Coate models where there is short-run dynamic inefficiency[16], the decision maker will 'raid the kitty' to provide directed transfers to

[16] These include Battaglini and Coate (2008), Battaglini and Coate (2011) and the generic case of the model in Battaglini (2011). In the latter model, two parties contest elections in districts that are sub-divided into one or more regions. Dynamic inefficiency exists whenever there is more than one region in a district.

its preferred district.[17] The same result is true in a 'starve the beast' model a'la Persson and Svensson (1989). Suppose the low spending party values the public good according to A_L and the high spending party values the public according to $A_H > A_L$, and let full-endowment refer to the level of assets that allows the high spending party to finance its desired level of the public good in every future period from interest earnings alone. From the low-spending party's perspective, the marginal benefit of public goods spending at the high-spending party's desired (Samuelson) level is strictly lower than the marginal benefit of providing transfers. The low-spending party has an incentive to provide an amount of transfers large enough (by drawing down on the government's endowment) that the high-spending party will be forced to levy positive taxes when it comes to power in the future. With higher taxes, the high-spending party must reduce the amount of public goods it provides. The low-spending party chooses its policy in such a way that the expected future marginal benefit of public goods provision is equal to the current marginal benefit of transfers.

In both of these models, the dynamic inefficiency creates an incentive for the decision maker to over-accumulate debt whenever the stock of assets is large enough. This creates a structural barrier that prevents the government from fully insuring itself against shocks and achieving the planner's long run optimum. By contrast, there is no such incentive that forces the economy away from the full endowment level in the model with belief-distortions. If the economy ever achieves this asset level, then it will remain at that level forever.

Will the belief-distorted economy accumulate the full endowment level of assets? Since the government over-accumulates debt in the high state, it under-accumulates debt in the low state, the long run dynamics of the debt stock are not immediately obvious. Proposition 13 describes the long run properties of the political equilibrium.

Proposition 13. *In the long run, the distorted economy will settle at the full endowment steady state in which* $\tau_\theta = 0$, $g_\theta = \frac{A_\theta}{p}$ *and* $x_\theta = \underline{x}$. *The government will make positive*

[17] See Proposition 3 in Battaglini and Coate (2008), Proposition 7 in Battaglini and Coate (2011) and Proposition 3 in Battaglini (2011).

transfers of $T = A_B - A_G$ whenever the state of the economy is G.

The long run characteristics of the political equilibrium coincide with the long run steady state in the planner's problem. Although the government's policies are inefficient in the short run, the precautionary savings motive still causes it to accumulate a buffer of savings in each period on average. In the long run, these savings will accumulate until the government is fully endowed. In spite of the short-run dynamic inefficiency, the government is able to fully insure itself against economic shocks, and hence its policy is efficient in the long run. This again stands in contrast to the other models discussed above, where the structural barrier to accumulating assets prevents the government from fully insuring itself.

Of course, the path to the long run steady state in the political equilibrium will likely not coincide with the path followed by the planner's economy. To understand the differences between the planner and equilibrium dynamics, I again first consider the case where equilibrium taxes are higher than planner taxes in both states of the economy. (Recall, by Lemma 9, this will be true if the voters' pessimism in the bad state overwhelms their optimisim in the good state - i.e. if there is considerable long run pessimism.) Intuitively, if equilibrium taxes are higher than planner taxes in each state, then the government will retire debt or accumulate savings more rapidly than a planner would - and so the distorted economy will converge to the full-endowment steady state faster than the planner economy.

This insight is formalized in the following Lemma: Let $\theta^{\infty} = (\theta_0, \theta_1, ...) \in \Theta$ be a countable sequence of shocks. Let $\zeta(b; \theta^{\infty})$ and $\zeta^o(b, \theta^{\infty})$ be the number of periods before the distorted and planner economies, respectively, first achieves the full endowment level of assets, given the current debt level and a trajectory for economic shocks.

Lemma 10. *Suppose equilibrium taxes are (weakly) higher than planner taxes in both states. Then, $\zeta(b; \theta^{\infty}) \leq \zeta^o(b, \theta^{\infty})$ for every $\theta^{\infty} \in \Omega$- i.e. the planner economy will not reach the long run steady state before the distorted economy. Conversely, if equilibrium taxes are (weakly) lower than planner taxes in both states, then $\zeta(b, \theta^{\infty}) \geq \zeta^o_{\theta}(b, \theta^{\infty})$ for every $\theta^{\infty} \in \Theta$.*

Lemma 10 shows that when voters' beliefs are characterized by sufficiently high long-run optimism (so that equilibrium taxes are weakly lower than planner taxes in each state), then the distorted economy takes longer to reach the long run steady state than the planner economy, for any sequence of economic shocks. Since the distorted economy has lower taxes in each period, agents receive higher stage utilities (than agents in the planner economy) at each period on the path to the long run steady - however this comes at the cost of a longer wait to receiving the sustained higher utility when the steady-state finally arrives. When voters' beliefs include both optimism and pessimism, so that equilibrium taxes are lower in the optimistic state, and higher in the pessimistic state, then the difference in the time to the long run steady state between the distorted and planner economies is not so clear cut.

Lemma 11. *Suppose equilibrium taxes are (weakly) higher than planner taxes in the bad state, but lower in the good state. Then:*

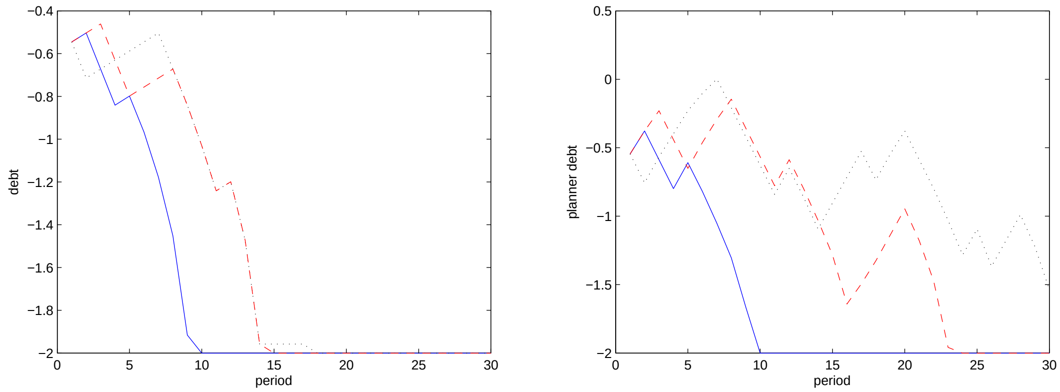
1. $\forall b \in (\underline{x}, \overline{x}]$ *there exists* $\theta^{\infty} \in \Theta$ *s.t.* $\zeta(b, \theta^{\infty}) \geq \zeta^{o}(b, \theta^{\infty})$

2. $\exists b \in (\underline{x}, \overline{x}]$ *and* $\theta^{\infty} \in \Theta$ *s.t.* $\zeta(b, \theta^{\infty}) < \zeta^{o}(b, \theta^{\infty})$

Lemma 11 demonstrates that there can be sequences of shocks to the economy in which the planner economy achieves the full endowment asset level before the distorted economy, and sequences of shocks in which the opposite is true. Let $E[\zeta(b)]$ and $E[\zeta^{o}(b)]$ be the expected arrival times at the long run steady state, of the distorted and planner economies, respectively. Simulations indicate that $E[\zeta(b)] < E[\zeta^{o}(b)]$ when there is long-run pessimism (i.e. $\pi_h^v > \pi_h^p$) and that the opposite is true when there is long-run optimism. (See Figures 1 and 2, below.) Assuming that expected arrival times are continuous in the transition probabilities, this is consistent with results in Lemma 10.

Figure 1 shows the path of debt in the distorted economy, and the corresponding path in the planner economy, for three separate sequence of economic shocks. The transition dynamics and the voters' beliefs about these are such that the voter exhibits long-run pessimism, since

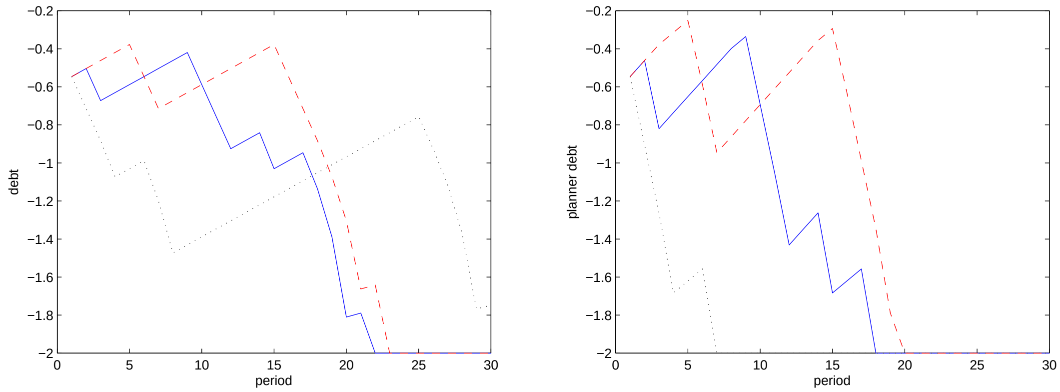
Figure 3.1: Simulations of path of debt in distorted and planner economies with long-run pessimism



The left hand panel shows the path of debt in the distorted economy in three separate simulations. The right hand panel shows the corresponding path of debt in the planner economy. The transition dynamics are given by: $p^p_{GG} = 0.8$, $p^p_{BG} = 0.4$, $p^v_{GG} = 0.9$ and $p^v_{GG} = 0.1$ which implies $\pi^p_G = \frac{2}{3}$ and $\pi^v_G = \frac{1}{2}$. The parameters used are: $A_H = 2$, $A_L = 1$, $w_H = 1.5$, $w_L = 1$, $p = 1$,$\delta = 0.5$, $\varepsilon = 1.5$ and $\phi = 0.5$.

they expect the good state to arise in the long run with probability $\frac{1}{2}$, whilst the good state will actually arise with probability $\frac{2}{3}$. As is evident from figure 1, the time for the planner economy to reach the long-run steady state is longer in each case. The intuition is as follows: Short-run optimism in the good state tends to cause equilibrium taxes to be below planner taxes in the good state, whilst short-run pessimism in the bad state has the opposite effect. However, as discussed in the previous section, long-run pessimism tends to cause equilibrium taxes to be larger than planner taxes in both states. The net effect is for equilibrium taxes to be slightly lower than planner taxes in the good state, but significantly larger than planner taxes in the bad state. The distortion is greatest in the bad state. Although the distorted economy over-accumulates debt in the good state, and under-accumulates in the bad state, the magnitude of the under-accumulation in the bad state overwhelms the size of the over-accumulation in the good state. The net effect is for the distorted economy to accumulate assets more rapidly, in expectation. As is evident from Figure 2, the converse appears to be true when there is long-run optimism.

Figure 3.2: Simulations of path of debt in distorted and planner economies with long-run optimism



The left hand panel shows the path of debt in the distorted economy in three separate simulations. The right hand panel shows the corresponding path of debt in the planner economy. The transition dynamics are given by: $p^p_{GG} = 0.6$, $p^p_{BG} = 0.2$, $p^v_{GG} = 0.9$ and $p^v_{GG} = 0.1$ which implies $\pi^p_G = \frac{1}{3}$ and $\pi^v_G = \frac{1}{2}$. The parameters used are the same as in Figure 1.

3.6 Incomplete Pandering

In Section 5, I considered the simple case where electoral competition is characterized by complete pandering, and in which the political parties behaved as social planners with the same beliefs about the transition dynamics as the voters. In this section, I consider the more general case where electoral competition is characterized by incomplete pandering. By Proposition 11, the equilibrium policies reflect a weighted sum of the voters' and parties' preferences. Since continuation preferences depend on agents beliefs about the continuation policies that will be chosen, the continuation preferences of both types of agents will depend upon their beliefs about the other type's preferences. In particular, the voters' continuation preferences will depend upon their higher order beliefs about the dynamics of the state of the economy. (Throughout the paper, it is assumed that the parties know the voters' (higher order) beliefs.)

There are two natural cases of higher order beliefs that serve as foci for analysis. In the first case, I assume that voters believe that the parties share their beliefs. In this case, the voters do not recognize the conflict between the parties and themselves. They naively assume that the continuation policies will coincide with their ideal policies. Hence the voters' continuation preferences coincide with their continuation preferences in the game with complete pandering. The parties will, of course, assess the future correctly, taking into account the policies that will actually be implemented (given their different preferences and beliefs). Following O'Donoghue and Rabin (1999), I refer to this as the *naive* case, since the voters fail to recognize that the continuation game will be different.

Let $W_\theta^n(b)$ and $V_\theta^n(b)$ be the parties' and voters' value functions (respectively), when the existing debt level is b, the state of the economy is θ, and political competition is characterized by incomplete pandering to naive voters. The naive equilibrium is characterized as follows:

$$\begin{aligned} V_\theta^n(b) &= \max_{\tau,g,x} u_\theta(\tau) + A_\theta \ln g + B_\theta(\tau,g,x;b) + \delta E_\theta^v\left[V_{\theta'}^n(x)\right] \\ \text{s.t.} &\quad B_\theta(\tau,g,x;b) \geq 0 \text{ and } x \in \{\underline{x}, \overline{x}\} \end{aligned}$$

and:

$$W_\theta^n(b) = u_\theta(\tau_\theta^n(b)) + A_\theta \ln g_\theta^n(b) + B_\theta(\tau_\theta^n(b), g_\theta^n(b), x_\theta^n(b); b) + \delta E_\theta^p\left[W_{\theta'}^n(x_\theta(b))\right]$$

where:

$$\begin{aligned} (\tau_\theta^n(b), g_\theta^n(b), x_\theta^n(b)) &= \arg\max_{\tau,g,x} u_\theta(\tau) + A_\theta \ln g + B_\theta(\tau,g,x;b) + \delta\left[\phi E_\theta^v\left[V_{\theta'}^n(x)\right] + (1-\phi) E_\theta^p\left[W_{\theta'}^n(x)\right]\right] \\ \text{s.t.} &\quad B_\theta(\tau,g,x;b) \geq 0 \text{ and } x \in \{\underline{x}, \overline{x}\} \end{aligned}$$

In the second focal case, I assume that voters know the parties' beliefs. Accordingly, the voters recognize the conflict between the parties and themselves, and understand that the continuation policies will maximize a convex combination of these different preferences. Voters with these higher order beliefs are *sophisticates*. Unlike the naive case, sophisticated voters are not mistaken about the continuation policies that will be enacted. However, their continuation preferences will still differ from the parties, since they assess the probabilities that each of these policies will be implemented differently.

Let $W_\theta^s(b)$ and $V_\theta^s(b)$ be the parties' and voters' value functions (respectively), when the existing debt level is b, the state of the economy is θ, and political competition is characterized by incomplete pandering to sophisticated voters. The sophisticated equilibrium is characterized as follows:

$$\begin{aligned}
V_\theta^s(b) &= u_\theta(\tau_\theta^s(b)) + A_\theta \ln g_\theta^s(b) + B_\theta(\tau_\theta^s(b), g_\theta^s(b), x_\theta^s(b); b) + \delta E_\theta^v [V_{\theta'}^s(x_\theta(b))] \\
W_\theta^s(b) &= u_\theta(\tau_\theta^s(b)) + A_\theta \ln g_\theta^s(b) + B_\theta(\tau_\theta^s(b), g_\theta^s(b), x_\theta^s(b); b) + \delta E_\theta^p [W_{\theta'}^s(x_\theta(b))]
\end{aligned}$$

where:

$$\begin{aligned}
(\tau_\theta^s(b), g_\theta^s(b), x_\theta^s(b)) &= \arg\max_{\tau,g,x} u_\theta(\tau) + A_\theta \ln g + B_\theta(\tau, g, x; b) + \delta \left[\phi E_\theta^v [V_{\theta'}^s(x)] + (1-\phi) E_\theta^p [W_{\theta'}^s(x)]\right] \\
\text{s.t.} \quad & B_\theta(\tau, g, x; b) \geq 0 \text{ and } x \in \{\underline{x}, \overline{x}\}
\end{aligned}$$

The difference between the naive and sophisticated equilibria can be seen in the recursive formulations of the value functions, above. In the naive case, the voters' continuation preferences are independently determined by their ideal policies and beliefs alone. The equilibrium policies and the parties' value functions are then determined, taking the voters' continuation preferences as given. By contrast, in the sophisticated case, the equilibrium policies and the parties' and voters' value functions are jointly determined.

In the previous section on complete pandering, it was unneccessary to distinguish between naive and sophisticated voters. In the case with complete pandering, it would still be possible to characterize voters as either naive or sophisticated. As above, naive voters assume that the parties share their preferences and will always offer their ideal policies in the future. By contrast, sophisticated voters understand that parties have different preferences. More importantly, sophisticated voters understand how electoral competition translates party and voter preferences into future policies. In particular, with complete pandering, sophisticated voters understand that electoral competition will rationally lead parties to offer the voters' ideal policies. Hence, the sophisticated voters' forecasts of the continuation policies will coincide with those of the naive voters. Given this coincidence, it was possible to remain agnostic about the voters' understanding of the game, in the previous section. By contrast, with incomplete pandering, these distinctions matter considerably.

I noted, above, that in the complete pandering case, the equilibrium policy functions inherited the same qualitative properties as the planner's policy functions. As the following Lemma demonstrates, the same remains true when electoral competition results in incomplete pandering.

Lemma 12. *In each of the games studied, the equilibrium policy functions have the same qualitative properties as the planner's policies outlined in Lemmata 5, 6 and 7.*

When there is incomplete pandering, the parties' policy choices are no longer simply the solution to a planner problem. Nevertheless, Lemma 12 shows that the optimal policies in this environment continue to exhibit the same qualitative properties as the policies that a planner would choose. Consequently, all of the main results in Section 5 (that depended upon the properties of the policy functions outline in Lemmata 5, 6 and 7) will continue to hold in the incomplete pandering case. This demonstrates that the paper's main results are robust to a variety of assumptions about the nature of party and voter preferences.

Finally, I consider how the policy functions differ between the cases of complete and incomplete pandering. With incomplete pandering, the parties seek to undo the effect of the voters' distorted preferences, by choosing policies closer to the ideal planner policy. When voters are naive, the parties will able to do this successfully, since the voters are unaware of the parties' motives. Of course, the parties' cannot completely correct for the voters' mistaken preferences, since the competition only enables the parties to only partially internalize their own preferences. Recall, τ_θ^o and τ_θ^f denote the parties' and voters' ideal taxes, respectively:

Lemma 13. *Suppose voters are naive. If* $\tau_\theta^o(b) < \tau_\theta^f(b)$*, then* $\tau_\theta^o(b) < \tau_\theta^n(b) < \tau_\theta^f(b)$ *and* $x_\theta^o(b) > x_\theta^n(b) > x_\theta^f(b)$*. Similarly, if* $\tau_\theta^o(b) > \tau_\theta^f(b)$*, then* $\tau_\theta^o(b) > \tau_\theta^n(b) > \tau(b)$ *and* $x_\theta^o(b) < x_\theta^n(b) < x_\theta^f(b)$*.*

Lemma 13 demonstrates that, with naive voters, the equilibrium policies are strictly between the ideal policies of the parties (i.e. the policies chosen by a planner with correct beliefs) and the ideal policies of the voters (i.e. the policies chosen by a planner with incorrect beliefs). Incomplete pandering results in policies that are less inefficient. Tax policies are still more volatile than is ideal, but the volatility is less pronounced than under the case where electoral competition results in complete pandering. This intuition need not remain true in the case of sophisticated voters. Since sophisticated voters understand that parties will seek to partially undo their ideal policies, they have an incentive to demand fiscal policies that are even more extreme than the naive voters would. The overall implications of this are not immediately clear.

3.7 Learning

The analysis in the above sections has relied on the assumption that the voter never learns about the true dynamics of the economy - even though the state of the economy is perfectly observable. In this section, I introduce a simple form of learning, to understand the validity of this assumption. Suppose the voter is Bayesian and has prior beliefs about the transition dynamics. For simplicity, suppose the voters' beliefs have support $\{P^v, P^p\}$ - so that the voter only assigns positive probability to the true process, and the specific distorted process considered in the model, above. Let $h^t \in H^t$ be a t-period history of the state of the economy, and let $\gamma_t(\theta_t)$ be the voters' belief that the process is governed by P^v at time t, when the state is θ_t . Voters' update their beliefs about the dynamic process in each period, after observing the realization of the state in that period. Suppose $\gamma_0 \in (0,1)$, so that the voters' initial beliefs are non-degenerate. Denote $\theta_t = \theta$ an $\theta_{t+1} = \theta'$, where $\theta, \theta' \in \{G, B\}$. Then, using Bayes' Rule:

$$\gamma_{t+1}(\theta') = \frac{\gamma_t(\theta)\, p^v_{\theta\theta'}}{\gamma_t(\theta)\, p^v_{\theta\theta'} + (1-\gamma_t(\theta))\, p^p_{\theta\theta'}} \tag{3.7}$$

Lemma 14. *Suppose $\gamma_t \in (0,1)$. Then $\gamma_{t+1}(\theta') > \gamma_t(\theta)$ whenever $\theta = \theta'$, and $\gamma_{t+1}(\theta') <$*

$\gamma_t(\theta)$ whenever $\theta \neq \theta'$. If $\gamma_t \in \{0,1\}$, then $\gamma_{t+1} = \gamma_t$, for any sequence of states.

Lemma 14 shows how voters update their beliefs about the Markov process, given the current and immediately prior states of the economy. Since the distorted process predicts greater stability in the current state, voters will interpret the lack of a change in the state between two adjacent periods as stronger evidence for the process being governed by the distorted process. By contrast, if the state does change, then voters will revise their beliefs in favor of the true process. Hence, voters will learn about the true process most rapidly when the state fluctuates frequently.

The following lemma characterizes the long run behavior of voters beliefs:

Lemma 15. *The voter learns the truth in the long run. Formally, $\lim_{t\to\infty} \gamma_t = 0$, for any $\gamma_0 < 1$.*

Lemma 15 states that, the voter will eventually assign probability 1 to the true Markov processP^p, after observing sufficiently many outcomes. As long as the voters' prior beliefs assign positive probability to the true process, then they will learn the truth in the long run, almost surely. Of course, Lemma 14 demonstrates that voters may rationally assign high probability to the distorted process of a long period of time, if the state of the economy does not fluctuate too quickly. Whilst learning occurs, it need not proceed rapidly, and the voters incorrect beliefs may be sustained over a long enough horizon for the baseline assumption of no-learning to be reasonable.

Moreover, the voters learn the truth in this model because learning is passive - voters observe the actual sequence of states, and this is independent of their policy choices. This feature may break down in a more expansive business cycle model where the transition probabilities depend upon the policies chosen in each period. For example, suppose the probability of transitioning to the good state were increasing in the size of the deficit, but the voters assume that the probability of transition is lower (than the parties do) for any given level of deficit spending. If so, the voters will demand lower deficit spending (i.e. higher taxes and lower debt) than is optimal, and this will decrease the likelihood that the economy will transition to the good state - which is consistent with the voters' belief. In a more expansive model, self-confirming equilibria of this sort may exist in which the voters' incorrect beliefs are sustained in the long run.

3.8 Conclusion

Fiscal policy choices in a democracy depend upon voters' forecasts of the trajectory of economic fundamentals. Anecdotal and empirical evidence suggest that voters are unduly pessimistic about the duration of recessions, and unduly optimistic about the duration of booms. These 'animal spirits', coupled with political competition that forces parties to accommodate voters' preferences, will result in the government choosing inefficient policies, by under-accumulating debt during recessions, and over-accumulating debt (or under-accumulating savings) during booms. This gives fiscal policy a pro-cyclic flavor. Taxes will be more volatile than they ought to be, and debt will not efficiently smooth the costs of taxation over time.

When the voters' beliefs are distorted, fiscal policy will be endogenously time inconsistent, even if all actors are rational and time consistent, and share common stage preferences. Unlike other political economy models of fiscal policy, the inefficiency in this model arises purely as a consequence of the distortion to beliefs, and manifests in agents having asymmetric continuation preferences. This distortion results in the government choosing inefficient short run policies, with higher than optimal taxation during recessions, and lower than optimal taxation during booms. The sense in which taxes are higher or lower is in comparison to continuation taxes - rather than planner taxes. I demonstrate that if the level of long-run pessimism is large enough (which will be true if the magnitude of optimism in the good state is small relative to the amount of pessimism in the bad state) then taxes will be uniformly higher in the distorted economy. Indeed, the inefficiency is revealed in the volatility of taxes and debt (which is higher and lower, respectively, in the distorted economy - rather than through a comparison of the levels of taxes and debt in the distorted and planner economies.

In spite of the short run inefficiencies, fiscal policy will become efficient in the long run. The distorted economy will converge to a long-run steady state in which the government is fully endowed, and therefore it is perfectly insured against future shocks. Taxes will converge to zero, and public goods provision will be at the (optimal) Samuelson level. Of course, the transition path to the long-run steady state is different in distorted-economy (relative to the planner's economy). The distorted economy will converge to the long-run steady state faster than the planner economy, almost surely, if voters are sufficiently pessimistic in the long run. The converse is true if voters are exhibit sufficiently high long-run optimism. If the degree of long-run optimism or pessimism is small enough, it is possible for either the distorted economy or the planner economy to first arrive at the long-run steady state. Simulations suggest that the expected time to the long-run steady state is longer for the distorted economy when voters exhibit long-run optimism, and vice versa.

The paper demonstrates that importance of identifying the particular mechanisms that cause inefficient fiscal policy in a democracy. In the other models discussed, the time-inconsistency generates a systematic barrier to the government accumulating assets and a bias towards choosing higher than efficient debt. By contrast, there is no such barrier in the model with distorted beliefs, and the bias in the government's debt choice does not act uniformly. (Recall, the government under-accumulates debt in the state in which it ought to run deficits.) As such policies aimed at curbing inefficiencies arising out of asymmetric preferences, such as balanced budget amendments and binding statutory debt ceilings, will be ineffective at addressing the inefficiencies arising out of belief-distortions. A more useful approach is to develop mechanisms that amplify the size of government budget surpluses and deficits. This suggests important avenues for future research into the behavioral respects of political economy and public finance.

3.9 Appendix

Proof of Proposition 8. I first prove existence. Let F denote the set of bounded, real valued functions $v(\cdot)$ defined over the set $[\underline{x}, \overline{x}]$. For each $\theta \in \{G, B\}$, define the operator $T_\theta : F \times F \to F$ as follows:

$$
\begin{aligned}
T_\theta [v_G, v_B](b) &= \max_{(\tau,g,x)} \{u_\theta(\tau) + A_\theta \ln g + B(\tau, g, x; b) + \delta [p_{\theta G}^p v_G(x) + (1 - p_{\theta G}^p) v_B(x)]\} \\
\text{s.t.} \quad & B_\theta(\tau, g, x; b) \geq 0 \qquad (3.8) \\
& x \in [\underline{x}, \overline{x}] \qquad (3.9)
\end{aligned}
$$

Let $v = (v_G, v_B)$ and $T[v](b) = (T_G[v_G, v_B](b), T_B[v_G, v_B](b))$. I confirm that T maps onto $F \times F$. (To see that $T_\theta[v]$ is bounded, it suffices to show that the policy functions τ_θ^v and g_θ^v are bounded, since the stage utility is continuous and the continuation utility is bounded. Clearly, $\tau \in \left[0, \frac{1}{1+\varepsilon}\right]$ in equilibrium, since $\tau = \frac{1}{1+\varepsilon}$ maximizes $R_\theta(\tau)$. It is also clear that the government never provides more than the Samuelson level of the public good, so $g_\theta^v(b) \leq \frac{A_\theta}{p}$. I need to show that there is some $\underline{g} > 0$, such that $g_\theta^v(b) \geq \underline{g}$ for all b. Let $\hat{g}_\theta(b) = \frac{R_\theta\left(\frac{1}{1+\varepsilon}\right) + \overline{x} - (1+r)b}{p}$. Since $\overline{x} < \frac{R_\theta\left(\frac{1}{1+\varepsilon}\right)}{r}$, then $\hat{g}_\theta(b) > 0$ for all b. Note that $(\tau, g, x) = \left(\frac{1}{1+\varepsilon}, \hat{g}_\theta, \overline{x}\right)$ is feasible. For each $v \in F \times F$, let $\Delta v = \max_\theta \{\max_b v_\theta(b) - \min_b v_\theta(b)\}$. Define $\underline{g}(v) > 0$ by $\ln \underline{g_\theta}(v) = \ln \hat{g}(\overline{x}) - \frac{\delta \Delta v + u_\theta(0) - u_\theta\left(\frac{1}{1+\varepsilon}\right)}{A_H}$. Clearly, $0 < \underline{g_\theta}(v) < \hat{g}(\overline{x}) \leq \hat{g}(b)$. Suppose the optimal policy is $(\tau_\theta^v, g_\theta^v, x_\theta^v)$ with $g_\theta^v < \underline{g_\theta}(v)$. Then:

$$
\begin{aligned}
T_\theta[v](b) &= u_\theta(\tau_\theta^v) + A_\theta \ln g_\theta^v + \delta E_\theta [v_{\theta'}(x_\theta^v)] \\
&= u_\theta\left(\frac{1}{1+\varepsilon}\right) + A_\theta \ln \hat{g}(b) + \delta E_\theta [v_{\theta'}(\overline{x})] + \left[u_\theta(\tau_\theta^v) - u\left(\frac{1}{1+\varepsilon}\right)\right] + A_\theta [\ln g_\theta^v - \ln \hat{g}(b)] \\
&\quad + \delta \{E[v_{\theta'}(x_\theta^v) | \theta] - E_\theta [v_{\theta'}(\overline{x})]\} \\
&\leq u_\theta\left(\frac{1}{1+\varepsilon}\right) + A_\theta \ln \hat{g}_\theta(b) + \delta E_\theta [v_{\theta'}(\overline{x})] + A_\theta [\ln g_\theta^v - \ln \hat{g}(b)] + \left[u_\theta(0) - u_\theta\left(\frac{1}{1+\varepsilon}\right)\right] + \delta \Delta v \\
&= u_\theta\left(\frac{1}{1+\varepsilon}\right) + A_\theta \ln \hat{g}_\theta(b) + \delta E_\theta [v_{\theta'}(\overline{x})] + A_H \left[\ln \hat{g}_\theta(\overline{x}) - \ln \underline{g_\theta(v)}\right] + A_\theta [\ln g_\theta^v - \ln \hat{g}(b)] \\
&\leq u_\theta\left(\frac{1}{1+\varepsilon}\right) + A_\theta \ln \hat{g}_\theta(b) + \delta E_\theta [v_{\theta'}(\overline{x})] + A_H [\ln \hat{g}_\theta(\overline{x}) - \ln \hat{g}_\theta(b)] + A_H \left[\ln g_\theta^v - \ln \underline{g_\theta}(v)\right] \\
&< u_\theta\left(\frac{1}{1+\varepsilon}\right) + A_\theta \ln \hat{g}_\theta(b) + \delta E_\theta [v_{\theta'}(\overline{x})]
\end{aligned}
$$

since $\hat{g}_\theta(\overline{x}) \leq \hat{g}_\theta(b)$ and $g_\theta^v < \underline{g_\theta}(v)$. But, since $\left(\frac{1}{1+\varepsilon}, \hat{g}_\theta(b), \hat{x}\right)$ is feasible, this implies that $(\tau_\theta^v, g_\theta^v, x_\theta^v)$ is not optimal, which is a contradiction. Hence $g_\theta^v \in \left[\underline{g_\theta}(v), \frac{A_\theta}{p}\right]$.

Then $T[v]$ is a contraction mapping. To prove this it suffices to show that Blackwell's conditions hold. Take any $v^A, v^B \in F \times F$ and suppose $v^A(b) \geq v^B(b)$ for all $b \in [\underline{x}, \overline{x}]$. Let

$$\left(\tau_\theta^i(b), g_\theta^i(b), x_\theta^i(b)\right) \in \arg\max_{(\tau,g,x)} \left\{u(\tau) + A_\theta \ln g + B_\theta(\tau, g, x; b) + \delta\left[p_{\theta G}^p v_G^i(x) + (1 - p_{\theta G}^p) v_B^i(x)\right]\right\}$$

and subject to (3.8) and (3.9), for each $\theta \in \{G, B\}$ and each $i \in \{A, B\}$. (For brevity, let $q_\theta^i(b) = (\tau_\theta^i(b), g_\theta^i(b), x_\theta^i(b)))$ Then:

$$\begin{aligned}
T_\theta\left[v^A\right](b) &= u_\theta\left(\tau_\theta^A\right) + A_\theta \ln g_\theta^A + B_\theta\left(q_\theta^A(b); b\right) + \delta\left[p_{\theta G}^p v_G^A\left(x_\theta^A\right) + (1 - p_{\theta G}^p) v_B^A\left(x_\theta^A\right)\right] \\
&\geq u_\theta\left(\tau_\theta^B\right) + A_\theta \ln g_\theta^B + B_\theta\left(q_\theta^B(b); b\right) + \delta\left[p_{\theta G}^p v_G^A\left(x_\theta^B\right) + (1 - p_{\theta G}^p) v_B^A\left(x_\theta^B\right)\right] \\
&\geq u_\theta\left(\tau_\theta^B\right) + A_\theta \ln g_\theta^B + B_\theta\left(q_\theta^B(b); b\right) + \delta\left[p_{\theta G}^p v_G^B\left(x_\theta^B\right) + (1 - p_{\theta G}^p) v_B^B\left(x_\theta^B\right)\right] \\
&= T_\theta\left[v^B\right](b)
\end{aligned}$$

where the second line follows from the optimality of q_θ^A (when continuation utilities are given by v^A) noting that q_θ^B was feasible, and the third line follows since $v^A \geq v^B$. Hence, T is monotonic. Next, for any $v_A \in F$, let $v_B = v_A + c$. Then:

$$\begin{aligned}
T_\theta\left[v^B\right](b) &= \max_{(\tau,g,x)} \left\{u_\theta(\tau) + A_\theta \ln g + B_\theta(\tau, g, x; b) + \delta\left[p_{\theta G}^p v_G^B(x) + \left(1 - p_{\theta G}^p\right) v_B^B(x)\right]\right\} \\
&= \max_{(\tau,g,x)} \left\{u_\theta(\tau) + A_\theta \ln g + B_\theta(\tau, g, x; b) + \delta\left[p_{\theta G}^p v_G^A(x) + \left(1 - p_{\theta G}^p\right) v_B^A(x)\right] + \delta c\right\} \\
&= T_\theta\left[v^A\right](b) + \delta c
\end{aligned}$$

which implies that T satisfies discounting. Hence, T is a contraction mapping. Hence, it contains a unique fixed point (V_G^o, V_B^o).

Next I show that the value functions are strictly decreasing, continuous and concave. Let $F_D \subset F$ be the set of bounded, strictly decreasing functions on $[\underline{x}, \overline{x}]$, and similarly let $F_C \subset F$ and $F^* \subset F_C$ be the set of continuous and concave (respectively) functions on $[\underline{x}, \overline{x}]$. I show that T_θ maps F into F_D for each $\theta \in \{G, B\}$. Take any $b, b' \in [\underline{x}, \overline{x}]$ with $b' > b$ and $\theta \in \{G, B\}$. Let (τ, g, x) and (τ', g', x') be the optimal policies for states b and b', respectively. Define $\hat{\tau}$ as follows: If $B_\theta(0, g', x'; b) > 0$, then $\hat{\tau} = 0$. Else, $\hat{\tau}$ solves:

$R_\theta(\hat{\tau}) - pg' + x' = (1+r)\,b$. Then $(\hat{\tau}, g', x')$ is feasible when the state is (b, θ). Note that $u_\theta(\hat{\tau}) \geq u_\theta(\tau')$ and $B_\theta(\hat{\tau}, g', x'; b) \geq B_\theta(\tau', g', x'; b') \geq 0$ and at least one of the inequalities is strict. (To see this, note that, by construction $\hat{\tau} \leq \tau'$ since $b < b'$. If $\hat{\tau} < \tau'$, then the claim holds straightforwardly. Suppose $\hat{\tau} = \tau'$. Then, it must be that $\hat{\tau} = \tau' = 0$. Then $B_\theta(\hat{\tau}, g', x'; b) = -pg' + x' - (1+r)\,b > -pg' + x' - (1+r)\,b' = B_\theta(\tau', g', x'; b')$.) Then:

$$\begin{aligned} T_\theta[v](b) &= u_\theta(\tau) + A_\theta \ln g + B_\theta(\tau, g, x; b) + \delta E_\theta[V_{\theta'}(x)] \\ &\geq u_\theta(\hat{\tau}) + A_\theta \ln g' + B_\theta(\hat{\tau}, g', x'; b) + \delta E_\theta[V_{\theta'}(x')] \\ &> u_\theta(\tau') + A_\theta \ln g' + B_\theta(\tau', g', x'; b') + \delta E_\theta[V_{\theta'}(x')] \\ &= T_\theta[v](b') \end{aligned}$$

Hence, $T_\theta[v](b) > T_\theta[v](b')$ and so T_θ maps into F_D. The equilibrium value functions must be strictly decreasing.

Next I show that T_θ maps onto F_C. Let $\{b_n\}$ be any monotonically decreasing sequence with $b_n \in [\underline{x}, \overline{x}]$ for each n and $b_n \to b$. Since $b_{n+1} \leq b_n$ for all n, then $T_\theta[v](b_{n+1}) \geq T_\theta[v](b_{n+1})$ for each n (since $T_\theta[v](b)$ is strictly decreasing in b). Then, $\{T_\theta[v](b_n)\}$ is a bounded monotone sequence, and so it converges. Let $T^+ = \lim_{b_n \downarrow b} T_\theta[v](b_n)$. Similarly, consider a monotonically increasing sequence $\{\beta_n\}$ and let $T^- = \lim_{\beta_n \uparrow b} T_\theta[v](\beta_n)$. Suppose $T^- \neq T^+$. Then $T^- > T^+$, since $T[v]$ is strictly decreasing. There is some $\zeta > 0$ s.t. $T^- - \zeta > T^+$

There exists some $N(\varepsilon) > 0$ s.t. $b + \varepsilon$ for all $n > N$ and $\beta_n > b - \varepsilon$. Fix some $n^* > N$. Then $b_n > \beta_n$ and $b_n - \beta_n < 2\varepsilon$. Let (τ, g, x) be the optimal policy at β_n, so $B(\tau, g, x; \beta_n) \geq 0$. Define $\hat{g}$ as follows: If $R_\theta(\tau) - pg + x > (1+r)\,b_n$, then $\hat{g} = g$; else $\hat{g} = \frac{R_\theta(\tau) + x - (1+r)b_n}{p}$. (Since $g > 0$, then $g' > 0$ for ε small enough.) Then $(\tau, \hat{g}, x)$ is a feasible policy at β_n. Then

$p\left(\hat{g}-g\right) \geq (1+r)\left(\beta_n - b_n\right)$

$$\begin{aligned}
T_\theta\left[v\right]\left[b_n\right] &\geq u_\theta\left(\tau\right) + A_\theta \ln \hat{g} + B_\theta\left(\tau,\hat{g},x;b_n\right) + \delta E_\theta\left[v_{\theta'}\left(x\right)\right] \\
&= T_\theta\left[v\right]\left[\beta_n\right] + A_\theta\left[\ln \hat{g} - \ln g\right] + B_\theta\left(\tau,\hat{g},x;b_n\right) - B_\theta\left(\tau,g,x;\beta_n\right) \\
&\geq T_\theta\left[v\right]\left[\beta_n\right] + \frac{A_\theta}{\hat{g}}\left(\hat{g}-g\right) - \left[p\left(\hat{g}-g\right) + (1+r)\left(b_n-\beta_n\right)\right] \\
&\geq T_\theta\left[v\right]\left[\beta_n\right] + \left[\frac{A_\theta}{\hat{g}} - p\right]\left(\hat{g}-g\right) + (1+r)\left(\beta_n - b_n\right) \\
&\geq T_\theta\left[v\right]\left[\beta_n\right] + \frac{A_\theta}{p\hat{g}}(1+r)\left(\beta_n-b_n\right) \\
&> T_\theta\left[v\right]\left[\beta_n\right] - 2\varepsilon\frac{A_\theta}{p\hat{g}}(1+r)
\end{aligned}$$

Now $\frac{A_\theta}{p\hat{g}} \geq 1$, since $\frac{A_\theta}{p g_\theta^s} = 1$ where g_θ^s is the Samuelson level of the public good, and $\hat{g} \leq g_\theta^s$, since $\hat{g} \leq g \leq g_\theta^s$. Hence, $T_\theta\left[v\right]\left(b_n\right) \geq T_\theta\left[v\right]\left[\beta_n\right] - 2\left(1+r\right)\varepsilon \max\left\{1, \frac{A_\theta}{p\underline{g}_\theta(v)}\right\}$. Take $\varepsilon < \frac{\zeta}{2(1+r)\max\left\{1,\frac{A_\theta}{p\underline{g}_\theta(v)}\right\}}$. Then $T^+ > T_\theta\left[v\right]\left(b_n\right) > T_\theta\left[v\right]\left[\beta_n\right] - \zeta > T^- - \zeta$, which is a contradiction. Hence $T^+ = T^-$, and so $T_\theta\left[v\right]$ is continuous at b. Since b was arbitrarily chosen, $T_\theta\left[v\right]$ is everywhere continuous, and so T maps into F_C. The equilibrium value functions must be continuous.

To show that $V_\theta\left(b\right)$ is concave, I show that $T_\theta : F^* \times F^* \to F^*$. Let $v \in F^* \times F^*$. The stage utility $u_\theta\left(\tau\right) + A_\theta \ln g + B\left(\tau, g, x; b\right)$ is concave whenever $\tau < \frac{1}{\varepsilon}$, which is always satisfied, since $\tau < \frac{1}{1+\varepsilon}$ in equilibrium. Then if v is concave, then the entire objective function. Moreover, $B_\theta\left(\tau, g, x; b\right)$ is concave everywhere, so $\left\{\left(\tau, g, x\right) | B_\theta\left(\tau, g, x; b\right) \geq 0\right\}$ is a convex set. Then, $T_\theta[v]$ is concave. (To verify this, let $b, b' \in \left[\underline{x}, \overline{x}\right]$ and let $b_\lambda = \lambda b + \left(1-\lambda\right) b'$ for $\lambda \in [0,1]$. Let $q = \left(\tau, g, x\right)$,$q' = \left(\tau', g', x'\right)$ and $q_\lambda = \left(\tau_\lambda, g_\lambda, x_\lambda\right)$ be the corresponding policy functions. Since the constraint set is convex, $\lambda q + \left(1-\lambda\right) q'$ is feasible for b_λ. For notational convenience, let

$\phi_\theta^v(q,b) = u_\theta(\tau) + A_\theta \ln g + B_\theta(\tau, g, x; b) + \delta E_\theta[v_{\theta'}(x)]$. Then:

$$\begin{aligned} T_\theta[v](b_\lambda) &\geq \phi_\theta^v(\lambda q + (1-\lambda) q') \\ &\geq \lambda \phi_\theta^v(q) + (1-\lambda)\phi_\theta^v(q') \\ &= \lambda T_\theta[v](b) + (1-\lambda) T_\theta[v](b') \end{aligned}$$

Hence $T_\theta : F^* \times F^* \to F^*$. Since $F^* \subset F$ and (V_H^o, V_L^o) is the unique fixed point of T on F, then $(V_H^o, V_L^o) \in F^*$, and so the equilibrium value functions are concave.

Finally, to show differentiability, suppose $b < b_\theta^o$. Then $V_\theta(b) = u_\theta(0) + A_\theta \ln\left(\frac{A_\theta}{p}\right) - A_\theta + \underline{x} - (1+r)b + \delta E_\theta[V_{\theta'}(\underline{x})]$ which is clearly differentiable, and $\frac{dV_\theta(b)}{db} = -(1+r)$. Suppose $b > b_\theta^o$. Let b' be in the neighborhood of b and let $\hat{g}_\theta(b') = \frac{R_\theta\left(\tau_\theta^o(b)\right) + x_\theta^o(b) - (1+r)b'}{p}$. Clearly $(\tau_\theta^o(b), \hat{g}(b'), x_\theta^o(b))$ is feasible for any b'. Let $\eta_\theta(b') = u_\theta(\tau_\theta^o(b)) + A_\theta \ln \hat{g}_\theta(b') + B(\tau_\theta^o(b), \hat{g}_\theta(b'), x_\theta^o(b)) + \delta E_\theta[V_{\theta'}(x_\theta^o(b))]$. Clearly $\eta_\theta(b')$ is differentiable in b' and $\eta_\theta(b') \leq V_\theta(b')$ (since the policy is feasible but not optimal for $b' \neq b$, and $\eta_\theta(b) = V_\theta(b)$. Then, by Benveniste and Scheinkman (1979), $V_\theta(b)$ is differentiable. Finally, note that $\lim_{b \to (b_G^o)^+} \frac{dV_G(b)}{db} = \lim_{b \to (b_G^o)^+} -(1+r)(1+\lambda_G(b)) = -(1+r) = \lim_{b \to (b_G^o)^-} \frac{dV_G(b)}{db}$. Hence V_G is differentiable at $b = b_G^o$ and $\frac{dV\left(b_G^o\right)}{db} = -(1+r)$. □

Proof of Lemma 5. Let (θ, b) denote the state, and suppose the economy is in the transfer regime. Then $B_\theta(\tau, g, x; b) > 0$ and $\lambda_\theta(b) = 0$. Hence by (3.1) and (3.12), $\tau^o = 0$ and $g_\theta^o = \frac{A_\theta}{p}$. By (3.13) that the future stock of debt, x, depends on the current stock, b, only through the dependence of λ on b. If $\lambda_\theta(b) = 0$, then $x_\theta(b) = x_\theta^o$ is a constant. This will be consistent if:

$$\begin{aligned} -A_\theta + x^o &> (1+r)b \\ b &< -\frac{A_\theta - x_\theta^o}{1+r} = b_\theta^o \end{aligned}$$

Since $\lambda_\theta(b) = 0$ and $\lambda \geq 0$, then (3.5) implies that $\lambda_{\theta'}(x_\theta) = 0$ for each $\theta' \in \{G, B\}$. Hence

$x^o_\theta \leq b^o_\theta$. Taking $\theta = B$, this implies:

$$\begin{aligned} x^o_B &\leq -\frac{A_B - x^o_B}{1+r} \\ x^o_B &\leq -\frac{A_B}{r} = \underline{x} \end{aligned}$$

Hence $x^o_B = \underline{x}$, and this implies that $b^o_B = \underline{x}$. Moreover, $\lambda_B(b) > 0$ for all $b > \underline{x}$, and so $x^o_G = \underline{x}$ and $b^o_G = \underline{x} + \frac{A_B - A_G}{1+r}$.

Now suppose the economy is in the no-transfer regime. Since the budget constraint binds in equilibrium (i.e. $\lambda(b,\theta) > 0$), then the optimal policies are the solution to the system:

$$\begin{aligned} g &= \frac{A_\theta}{p}\frac{1-(1+\varepsilon)\tau}{1-\tau} \\ -\delta E_\theta\left[\frac{\partial V_{\theta'}(x)}{\partial x}\right] &= \frac{1-\tau}{1-(1+\varepsilon)\tau} \\ R_\theta(\tau) - pg + x &= (1+r)\,b \end{aligned}$$

□

Proof of Lemma 6. The proof of this lemma, and Lemma 7 follow the same method as in Battaglini and Coate (2011). First I show that the policy functions are continuous. Since the objective function and constraint functions are continuous in b, then by Berge's Theorem, the policy correspondences must be upper-hemicontinuous. Moreover, the objective function is strictly concave (since the stage utilities are strictly concave), and so the maximizers must be unique. Hence the policy correspondences are indeed functions. Then, since the policy functions are upper-hemicontinuous, they are continuous.

Let $b' > b \geq b^o_\theta$ and suppose $\tau^o_\theta(b') \leq \tau^o_\theta(b)$. For notational convenience, let $\tau^o_\theta(b) = \tau$ and $\tau^o_\theta(b') = \tau'$, and similarly define g and g', and x and x'. Then since $g_\theta(b) = \frac{A_\theta}{pD\left(\tau^o_\theta(b)\right)}$, $g' \geq g$. Furthermore, $-\delta E_\theta\left[\frac{dV_{\theta'}(x')}{dx}\right] = D(\tau') \leq D(\tau) = -\delta E_\theta\left[\frac{dV_{\theta'}(x)}{dx}\right]$ and so $E_\theta\left[\frac{dV_{\theta'}(x')}{dx}\right] \geq E_\theta\left[\frac{dV_{\theta'}(x)}{dx}\right]$. Since $E[V_\theta(x)]$ is concave in x, this implies $x' \leq x$. Hence $B_\theta(\tau, g, x; b) =$

$R_\theta(\tau) - pg + x - (1+r)b > R_\theta(\tau') - pg' + x' - (1+r)b' = 0$. But since $b' > b_\theta^o$, it must be that $B_\theta(\tau, g, x; b) = 0$, which is a contradiction. Hence $\tau_\theta^o(b') > \tau_\theta^o(b)$. It immediately follows from the first order conditions that $g_\theta^o(b') < g_\theta^o(b)$. If $x_\theta^o(b') = \overline{x}$, then clearly $x_\theta^o(b') \geq x_\theta^o(b)$ and $x_\theta^o(b') = x_\theta^o(b)$ only if $x_\theta^o(b) = \overline{x}$. If $x_\theta^o(b') < \overline{x}$, then $-\delta E_\theta\left[\frac{dV_{\theta'}(x')}{dx}\right] = D(\tau') > D(\tau) \geq -\delta E_\theta\left[\frac{dV_{\theta'}(x)}{dx}\right]$ and so $E_\theta\left[\frac{dV_{\theta'}(x')}{dx}\right] < E_\theta\left[\frac{dV_{\theta'}(x)}{dx}\right]$. By the concavity of $E_\theta[V_{\theta'}]$, $x_\theta^o(b') > x_\theta(b)$. □

Proof of Lemma 7. Let $b > \underline{x}$, so that $\tau_B(b) > 0$. Suppose $\tau_G(b) \geq \tau_B(b)$. Then, since $g_\theta(b) = \frac{A_\theta}{pD(\tau_\theta(b))}$, $g_G(b) < g_B(b)$. Moreover, since $-\delta E_\theta\left[\frac{dV_{\theta'}(x_\theta(b))}{dx}\right] = D(\tau_\theta(b))$, then:

$$-\delta\left[p_{GG}\frac{dV_G(x_G(b))}{dx} + (1-p_{GG})\frac{dV_B(x_G(b))}{dx}\right] \geq -\delta\left[p_{BG}\frac{dV_G(x_B(b))}{dx} + (1-p_{BG})\frac{dV_B(x_B(b))}{dx}\right] \quad (3.10)$$

Suppose $\frac{dV_G(x)}{dx} > \frac{dV_B(x)}{dx}$ for all $x \in (\underline{x}, \overline{x})$. Then since $p_{GG} \geq p_{BG}$ (which follows from $p_{BB} \geq p_{GB}$):

$$p_{BG}\frac{dV_G(x_G(b))}{dx} + (1-p_{BG})\frac{dV_B(x_G(b))}{dx} < p_{GG}\frac{dV_G(x_G(b))}{dx} + (1-p_{GG})\frac{dV_B(x_G(b))}{dx} \quad (3.11)$$

Combining (3.10) and (3.11) gives $E_B\left[\frac{dV_{\theta'}(x_G(b))}{dx}\right] < E_B\left[\frac{dV_{\theta'}(x_B(b))}{dx}\right]$. Since $E[V_\theta]$ is concave, then $x_G(b) > x_B(b)$. These results and the fact that $R_G(\tau) > R_B(\tau)$ whenever $\tau > 0$ imply:

$$\begin{aligned} 0 &= R_B(\tau_B(b)) - pg_B(b) + x_B(b) - (1+r)b \\ &< R_G(\tau_G(b)) - pg_G(b) + x_G(b) - (1+r)b \end{aligned}$$

which is a contradiction, since $0 < \tau_B(b) \leq \tau_G(b)$ implies that the budget constraint must strictly bind in both states. Hence $\tau_G(b) < \tau_B(b)$ whenever $b > \underline{x}$.

To see that $x_G(b) < b \leq x_B(b)$ whenever $b > \underline{x}$, recall that (3.13) and (3.4) imply: $E\left[\frac{dV_{\theta'}(x_\theta(b))}{dx}|\theta\right] \geq \frac{dV_\theta(b)}{db}$ (with inequality only if $x_\theta(b) = \overline{x}$). Since $\frac{dV_B(x)}{dx} < \frac{dV_G(x)}{dx}$, then $\frac{dV_G(x_G(b))}{dx} > p_{GG}\frac{dV_G(x_G(b))}{dx} + (1-p_{GG})\frac{dV_B(x_G(b))}{dx} \geq \frac{dV_G(b)}{db}$. Then, since V_G is concave, it follows that $x_G(b) < b$. Now if $x_B(b) = \overline{x}$, then clearly $x_B(b) > b$ unless $b = \overline{x}$. Suppose $x_B(b) < \overline{x}$. Then $\frac{dV_B(x_B(b))}{dx} < p_{BG}\frac{dV_G(x_B(b))}{dx} + (1-p_{BG})\frac{dV_B(x_B(b))}{dx} = \frac{dV_B(b)}{db}$, and so by the concavity of

V_B, $x_B(b) > b$. Hence, $x_G(b) < b < x_B(b)$ whenever $b \in (\underline{x}, \overline{x})$ and $x_G(\overline{x}) < \overline{x} = x_B(\overline{x})$.

It remains to prove that indeed $\frac{dV_B(x)}{dx} < \frac{dV_G(x)}{dx}$ for all $x \in (\underline{x}, \overline{x})$. Recall from the proof of Proposition 8 that $T : F \times F \to F \times F$ is a contraction mapping, where F is the set of continuous, concave functions on $[\underline{x}, \overline{x}]$. Let $G \subset F \times F$ where $(v_G, v_B) \in G$ if these functions are everywhere differentiable and $\frac{dv_B(b)}{db} < \frac{dv_G(b)}{db}$ for all $b \in (\underline{x}, \overline{x})$. Suppose $(v_G, v_B) \in G$. Then:

$$T_\theta [v_H, v_L](b) = u_\theta (\tau^v_\theta (b)) + A_\theta \ln g^v_\theta (b) + B_\theta (\tau^v_\theta (b), g^v_\theta (b), x^v_\theta (b); b) + \delta E [v_{\theta'} (x^v_\theta (b)) | \theta]$$

By the above argument, $\tau^v_B(b) > \tau^v_G(b)$, since $\frac{dv_B(b)}{db} < \frac{dv_G(b)}{db}$ for all $b \in (\underline{x}, \overline{x})$. By the envelope theorem, $\frac{dT_\theta[v](b)}{db} = -(1+r)(1+\lambda^v_\theta(b)) = -(1+r) D(\tau^v_\theta(b))$. Hence $\frac{dT_B[v](b)}{db} < \frac{dT_G[v](b)}{db}$. Hence $T[v] \in G$ whenever $v \in G$ and this implies that the unique fixed point of T, $(V_G, V_B) \in G$. It follows that $\frac{dV_B(b)}{db} < \frac{dV_G(b)}{db}$ for all $b \in (\underline{x}, \overline{x})$. □

Proof of Proposition 9. The marginal cost of public funds follows a super-martingale since $E[\lambda_{\theta'}(x_\theta(b)) | \theta] \leq \lambda_\theta(b)$. Since $\lambda_\theta \geq 0$, then $\lambda \to c \geq 0$ by the Martingale Convergence Theorem. Suppose $c > 0$. Then $\lambda_G(x_\theta(b)) = \lambda_B(x_\theta(b)) = \lambda_\theta(b) = c$. But $\lambda_\theta > 0$ implies $\tau_\theta > 0$, since $1 + \lambda_\theta(b) = D(\tau_\theta(b))$. Moreover, by Lemma 7, $\tau_G(b) < \tau_B(b)$ whenever $\tau_B(b) > 0$ and this contradicts $\lambda_G(x_\theta(b)) = \lambda_B(x_\theta(b))$. Hence $c = 0$ and $\tau_B(b) = \tau_G(b) = 0$. This is only possible if $b = \underline{x}$. Hence $\tau \to 0$ and $x \to \underline{x}$. □

Proof of Proposition 10. Let q^* be the voters' ideal policy. (I establish in Proposition 12 that such a unique policy exists.) Let $G(x) = F(x)$ if there is aggregate uncertainty, and let:

$$G(x) = \begin{cases} 1 & x > 0 \\ \frac{1}{2} & x = 0 \\ 0 & x < 0 \end{cases}$$

if there is no aggregate uncertainty.

First, consider the case where the parties are solely office-motivated. Consider party i, and suppose the opponent offers q_{-i}. Then party i chooses q to maximize: $G(\omega_\theta(q) - \omega_\theta(q_{-i}))$ - i.e. the probability that it wins, given the opponent's strategy. This objective is maximized at $q = q^*$ for any opponent choice q_{-i}. Hence, it is a (weakly) dominant strategy for parties to choose q^*.

Next, consider the case where there is no aggregate uncertainty. Let q_A, q_B be the equilibrium policies and let $\Delta\upsilon = \upsilon_\theta(q_A) - \upsilon_\theta(q_B)$ and $\Delta\omega = \omega_\theta(q_A) - \omega_\theta(q_B)$. In equilibrium, $\Delta\upsilon = 0$. (To see why, suppose not. Without loss of generality, suppose let $\Delta\upsilon > 0$, so that A wins the election for sure. Then party B's payoff is $(1-\gamma)\omega_\theta(q_A)$. Party B can guarantee itself a strictly larger utility ($\left[(1-\gamma)\omega_\theta(q_A) + \frac{1}{2}\gamma s.\right]$) by imitating party A's platform. But if so, $\Delta\upsilon = 0$.)

Now suppose $q_B \neq q^*$, and let q_A be such that $\upsilon_\theta(q_A) = \upsilon_\theta(q_B)$. Let $\varepsilon > 0$. Since υ_θ and ω_θ are continuous in q and $\upsilon_\theta(q_B) < \upsilon_\theta(q^*)$ there exists a feasible q' s.t. $\upsilon_\theta(q') > \upsilon_\theta(q^*)$ and $\omega_\theta(q') > \omega_\theta(q_A) - \varepsilon$. If party A offers q', it's utility is

$$\begin{aligned}
(1-\gamma)\omega_\theta(q') + \gamma s &> (1-\gamma)[\omega_\theta(q_A) - \varepsilon] + \gamma s \\
&= (1-\gamma)\omega_\theta(q_A) + \frac{1}{2}\gamma s + \left[\frac{1}{2}\gamma s - (1-\gamma)\varepsilon\right] \\
&> (1-\gamma)\omega_\theta(q_A) + \frac{1}{2}\gamma s
\end{aligned}$$

provided that $\varepsilon < \frac{1}{2}\frac{\gamma}{(1-\gamma)}s$. Hence, there is a profitable deviation for party A. But this contradicts the assumption that q_A, q_B was an equilibrium. □

Proof of Proposition 11. Let q_A, q_B be the equilibrium policies. Let $\Delta\omega = \omega_\theta(q_A) - \omega_\theta(q_B)$ and let $\Delta\upsilon = \upsilon_\theta(q_A) - \upsilon_\theta(q_B)$. In equilibrium, it must be true that neither party could do strictly better by imitating their opponent. For party A, fixing party $B's$ policy

q_B, this implies that:

$$F(\Delta\upsilon)\cdot[\Delta\omega+s] \quad \geq \quad F(0)\,s \tag{3.12}$$

Analogously for party B, we have:

$$(1-F(\Delta\upsilon))(-\Delta\omega+s)\geq(1-F(0))\,s \tag{3.13}$$

Recall that $F(0)=\frac{1}{2}$ and $1-F(x)=F(-x)$. Rearranging (3.12) and (3.13), we have $F(\Delta\upsilon)\,\Delta\omega \geq \left(\frac{1}{2}-F(\Delta\upsilon)\right)s \geq F(-\Delta\upsilon)\,\Delta\omega$. This implies that $\left(\frac{1}{2}-F(\Delta\upsilon)\right)$ and $\Delta\omega$ must have the same sign - since $s>0$ and $F(\cdot)\geq 0$. Hence $\left(\frac{1}{2}-F(\Delta\upsilon)\right)\Delta\omega\geq 0$. Also, $(F(\Delta\upsilon)-F(-\Delta\upsilon))\,\Delta\omega\geq 0$ which implies that $(2F(\Delta\upsilon)-1)\,\Delta\omega\geq 0$. Together these imply either that $\Delta\omega=0$ or $F(\Delta\upsilon)=\frac{1}{2}$ (which in turn implies $\Delta\upsilon=0$). In fact, both of these must be true simultaneously. (To see this, note that if $\Delta\omega=0$, then $0\geq\left[\frac{1}{2}-F(\Delta\upsilon)\right]s\geq 0$ which implies $F(\Delta\upsilon)=\frac{1}{2}$. But since $f(0)>0$, this only happens when $\Delta\upsilon=0$. Conversely, if $\Delta\upsilon=0$, then $\frac{1}{2}\Delta\omega\geq 0\geq\frac{1}{2}\Delta\omega$, which is only possible if $\Delta\omega=0$.)

Since the parties' problems are symmetric, I consider the problem from the perspective of party A, without loss of generality. For a given state (b,θ), and given party B's policy q_B, party A's problem is:

$$\max\left\{F(\upsilon_\theta(q)-\upsilon_\theta(q_B))\left[(1-\gamma)(\omega_\theta(q)-\omega_\theta(q_B))+\gamma s\right]+(1-\gamma)\,\omega_\theta(q_B)\right\} \text{ s.t. } B_\theta(\tau,g,x;b)\geq 0$$

The first order conditions are:

$$[(1-\gamma)(\omega_\theta(q)-\omega_\theta(q_B))+\gamma s]\,f(\upsilon_\theta(q)-\upsilon_\theta(q_B))\frac{\partial\upsilon_\theta}{\partial q}+F(\upsilon_\theta(q)-\upsilon_\theta(q_B))(1-\gamma)\frac{\partial\omega_\theta}{\partial q}=-\mu\frac{\partial B}{\partial q}$$

By the above result, in equilibrium $\Delta\omega=0$ an $\Delta\upsilon=0$. Hence, the first order conditions reduce to:

$$\gamma s f(0)\frac{\partial\upsilon_\theta}{\partial q}+(1-\gamma)\,F(0)\frac{\partial\omega_\theta}{\partial q}=-\mu\frac{\partial B}{\partial q}$$

which can be re-written:

$$\phi \frac{\partial \upsilon_\theta}{\partial q} + (1-\phi) \frac{\partial \omega_\theta}{\partial q} = -\frac{\mu}{\gamma s f(0) + (1-\gamma) F(0)}$$

where $\lambda_t = \frac{\mu_t}{\gamma s f(0)+(1-\gamma)F(0)}$. But this is precisely the first order condition that results from maximizing the ϕ-weighted sum of voter and party utilities, subject to the budget constraint.

□

Proof of Proposition 12. Since the voters' preferences are decisive, the parties essentially behave as a planner would, if the planner's beliefs coincided with the voters' beliefs. Hence, existence is proved in the same way as Proposition 8. □

Proof of Lemma 8. The dynamics of the economy are governed by the Euler Equation (3.6). By Lemma 7, $\tau_B^f(b) > \tau_G^f(b)$ for all $b > \underline{x}$. Since $D(\cdot)$ is strictly increasing, then $D\left(\tau_B^f(x_\theta(b))\right) - D\left(\tau_G^f(x_\theta(b))\right) > 0$. Suppose $\theta = G$. Then $p_{GG}^v - p_{GG}^p > 0$. Furthermore, (3.6) holds with equality, since $x_G(b) < \overline{x}$ (by Lemma 7). Hence $E_G^p\left[D\left(\tau_{\theta'}^f\left(x_G^f(b)\right)\right)\right] > D\left(\tau_G^f(b)\right)$. Suppose instead that $\theta = B$. Then $p_{BG}^v - p_{BG}^p < 0$ and so $E_B^p\left[D\left(\tau_{\theta'}^f\left(x_B^f(b)\right)\right)\right] < D\left(\tau_B^f(b)\right)$ (and this true even if (3.6) holds as a strict inequality). □

Proof of Lemma 9. I prove the Lemma for the case where equilibrium taxes are uniformly lower. The proof for the case when equilibrium taxes are always larger is analogous.

I first show that the Lemma holds whenever the voter is weakly optimistic in both states. (In particular, it holds whenever the voter is optimistic during the good state and has correct beliefs during the bad state.) By continuity, the Lemma extends to cases when the voter is slightly pessimistic in the bad state.

Let $p_{\theta G}^v \geq p_{\theta G}^p$, with strict inequality for at least one $\theta \in \{G, B\}$. For each $\theta \in \{G, B\}$, I claim that $\frac{dV_\theta(b)}{db} > \frac{dV_\theta^o(b)}{db}$ for all $b > b_\theta^*$. Then $E_\theta\left[\frac{dV_{\theta'}(b)}{db}\right] > E_\theta\left[\frac{dV_{\theta'}^o(b)}{db}\right]$ for all $b > \underline{x}$ (since

$b_B^* = \underline{x}$). Then, by the first order conditions:

$$\begin{aligned} D\left(\tau_\theta\left(b\right)\right) &= -\delta E_\theta^v \left[\frac{dV_{\theta'}\left(x_\theta\left(b\right)\right)}{db}\right] \\ &< -\delta E_\theta^v \left[\frac{dV_{\theta'}^o\left(x_\theta\left(b\right)\right)}{db}\right] \\ &\leq -\delta E_\theta^p \left[\frac{dV_{\theta'}^o\left(x_\theta\left(b\right)\right)}{db}\right] \end{aligned}$$

where the final line follows from the voters' optimism and the fact that, by Lemma 7, $\frac{dV_G^o(b)}{db} > \frac{dV_B^o(b)}{db}$ for all b. Moreover, $D\left(\tau_\theta^o\left(b\right)\right) = -\delta E_\theta^p \left[\frac{dV_{\theta'}^o\left(x_\theta^o(b)\right)}{db}\right]$. Suppose $\tau_\theta^o\left(b\right) \leq \tau_\theta\left(b\right)$ which implies that $x_\theta^o\left(b\right) \geq x_\theta\left(b\right)$, by the budget constraint. Since $E_\theta^p\left[V_{\theta'}^o\left(b\right)\right]$ is concave, this implies $-\delta E_\theta^p\left[\frac{dV_{\theta'}^o(x_\theta(b))}{db}\right] \leq -\delta E_\theta^p\left[\frac{dV_{\theta'}^o\left(x_\theta^o(b)\right)}{db}\right]$. Hence $D\left(\tau_\theta\left(b\right)\right) < D\left(\tau_\theta^o\left(b\right)\right)$, which contradicts $\tau_\theta^o\left(b\right) \leq \tau_\theta\left(b\right)$. Hence $\tau_\theta^o\left(b\right) > \tau_\theta\left(b\right)$.

It remains to show that the claim indeed holds. Let $H = \left\{(v_G, v_B) \in F \times F \middle| \frac{dv_\theta(b)}{db} > \frac{dV_\theta^o(b)}{db}\, \forall b > b_\theta^*\right\}$. Recall from Proposition 12 that (V_G, V_B) is the unique fixed point of the operator $T_\theta\left[v\right]\left(b\right) = \max_{(\tau,g,x)}\left\{u_\theta\left(\tau\right) + A_\theta \ln g + B_\theta\left(\tau, g, x; b\right) + \delta E_\theta^v\left[v_{\theta'}\left(x\right)\right]\right\}$. It suffices to show that $T : H \to H$ and so $(V_G, V_B) \in H$. Since $E_\theta\left[\frac{dv_{\theta'}(x)}{dx}\right] > E_\theta\left[\frac{dV_{\theta'}^o(x)}{dx}\right]$, then by the above discussion, $\tau_\theta^v\left(b\right) < \tau_\theta^o\left(b\right)$ whenever $b > b_\theta^*$. By the envelope theorem, $\frac{dT_\theta[v](b)}{db} = -\left(1+r\right) D\left(\tau_\theta^v\left(b\right)\right)$ and $\frac{dV_\theta^o(b)}{db} = -\left(1+r\right) D\left(\tau_\theta^o\left(b\right)\right)$. Hence $\frac{dT_\theta[v](b)}{db} > \frac{dV_\theta^o(b)}{db}$ for each θ. Hence $T : H \to H$. □

Proof of Proposition 13. For any $b \in [\underline{x}, \overline{x}]$, define $x_G^n\left(b\right)$ recursively as follows: (i) $x_G^1\left(b\right) = x_G\left(b\right)$ and (ii) for any $n \geq 1$, $x_G^{n+1}\left(b\right) = x_G\left(x_G^n\left(b\right)\right)$. Hence, $x_G^n\left(b\right)$ is the debt level that is chosen if the economy remains in state L for n consecutive periods, beginning with debt level b. Consider the sequence $\{x_n\}$ where $x_n = x_G^n\left(b\right)$, for some arbitrary b. By Lemma 7, $x_G\left(b\right) \leq b$ for all b (with strict inequality whenever $b > \underline{x}$), and so $\{x_n\}$ is a monotonically decreasing sequence. Furthermore, since $x_n \geq \underline{x}$ for each n, then the sequence is bounded, and so it converges. Since $x_G\left(b\right)$ is continuous in b, it must be that $x_G\left(x\right) = x$. (To see why, suppose not. Then $x_G\left(x\right) < x$. Let $\varepsilon > 0$ be such that $x_G\left(x\right) + \varepsilon < x$. Since

x_G is continuous, then there exists $\delta > 0$ such that when $y < x + \delta$, then $x_G(y) < x_G(x) + \varepsilon$. Since $x_n \to x$, there exists $N(\delta)$ large enough s.t. $x_n < x + \delta$ for all $n > N(\delta)$. But then, for $n > N(\delta)$, $x_{n+1} = x_G(x_n) < x$, which contradicts $x_n \to x$.) But $x_G(x) = x$ only if $x = \underline{x}$. Hence $x_n \to \underline{x}$. In fact, this convergence happens in finitely many periods. Since $x_n \to \underline{x}$ monotonically, there exists $N \geq 1$ s.t. $x_n \leq b_G^*$ for all $n \geq N$. But then $x_G(x_n) = \underline{x}$, and so $x_n = \underline{x}$ for all $n \geq N + 1$.

For each $b \in [\underline{x}, \overline{x}]$, let $N(b) = \min\{n \geq 1 | \, x_G^n(b) = \underline{x}\}$. Let $S = [\underline{x}, \overline{x}] \times \{H, L\}$ denote the state space, where $s = (b, \theta)$ is an arbitrary state. Let Σ denote the Borel sets of S. Let Q be the transition function on (S, Σ) induced by the Markov process that governs the state θ and the equilibrium policies $\{x_\theta(b)\}$. Clearly

$$Q(s, s') = \begin{cases} p^p_{\theta\theta'} & \text{if } s = (b, \theta) \text{ and } s' = (x_\theta(b), \theta') \\ 0 & \text{otherwise} \end{cases}$$

for each $b \in [\underline{x}, \overline{x}]$. Let $S_T = (\underline{x}, \overline{x}] \times \{G, B\}$ and let $S_E = \{\underline{x}\} \times \{G, B\}$. I claim that S_T is the set of transient states and S_E is the unique ergodic set. It is easy to see that S_E is an ergodic set, since for each $\theta \in \{G, B\}$, $x_\theta(\underline{x}) = \underline{x}$, and hence $Q(s, \Sigma_E) = 1$ for each $s \in S_E$. Moreover, for each strict subset of S_E (either $(\underline{x}, G)$ or $(\underline{x}, B)$), there is a positive probability of transitioning away from that state (if θ switches). To see that S_T is transient, it suffices to show that, for any $s \in S_T$, there is a positive probability of transitioning to S_E in a finite number of steps. (If so, there is a positive probability of transitioning away from state s and never returning.) Take any $s = (b, \theta) \in S_T$. Then, the economy can transition to S_E in a minimum of $N(b)$ periods if $\theta = G$ and in a minimum of $N(x_H(b)) + 1$ periods if $\theta = B$. (By construction, $N(b)$ is the number of transitions needed to get from debt level b to $\underline{x}$, if the state is L in each period. Obviously, if the state is in state B, it must first transition to state G.) The probability of this sequence of transitions arising is $(p^p_{GG})^{N(b)} > 0$ if $\theta = G$ and $p^p_{BG}(p^p_{GG})^{N(x_B(b))} > 0$ if $\theta = B$. Hence, for any $s \in \Sigma_T$, there is a positive probability of transitioning away from s and never returning, and so every $s \in S_T$ is transient. Finally, note that since $S_E \cup S_T = S$, then S_E must be the unique ergodic set.

Since $x_G(b)$ is monotonic in b, $N(b) \leq N(b')$ whenever $b < b'$. Hence, for any $s \in S_T$, the probability of transitioning to S_E in $N(\overline{x}) + 1$ steps at least as large as $p^p_{BG}(p^p_{GG})^{N(\overline{x})}$. Let $\rho = p^p_{BG}(p^p_{GG})^{N(\overline{x})}$ and let $N' = N(\overline{x}) + 1$. Let $s \in S_T$ and let $Q^n(s, S_T)$ be the probability that the economy remains in S_T after at most n steps. Then $Q^{N'}(s, S_T) \leq \rho < 1$. Moreover, $Q^{nN'}(s, S_T) \leq \rho^n$. Hence $\lim_{n\to\infty} Q^{nN'}(s, S_T) = 0$. In the long run, the economy leaves S_T and enters the ergodic set (where it remains) with probability 1. □

Proof of Lemma 10. Consider the case with weakly higher taxes. The proof of the case with weakly lower taxes is analogous. Since $\tau_\theta(b) \geq \tau^o_\theta(b)$ for all (b, θ), then $x_\theta(b) \leq x^o_\theta(b)$. Moreover, by Lemma 6, $x_\theta(b) \leq x_\theta(b')$ whenever $b < b'$. For an initial debt level b, and a sequence of shocks $\theta^\infty = (\theta_1, \theta_2, ...)$, let $\chi_t(b; \theta^\infty)$ be the equilibrium debt level after the first t periods. This function is defined recursively by: $\chi_t(b; \theta^\infty) = x_{\theta_t}(\chi_{t-1}(b; \theta^\infty))$ and $\chi_1(b, \theta^\infty) = x_{\theta_1}(b)$. Let $\chi^o(b; \theta^t)$ - the planner's debt level after t periods - be defined analogously.

Let θ^t be any finite, arbitrary sequence of shocks. Then $\chi_t(b; \theta^\infty) \leq \chi^o_t(b; \theta^\infty)$. I prove this inductively. For $t = 1$, $\chi_1(b; \theta^\infty) = x_{\theta_1}(b) \leq x^o_{\theta_1}(b) = \chi^o_1(b; \theta^\infty)$, which verifies the claim in the base case. Suppose $\chi_{t-1}(b; \theta^\infty) \leq \chi^o_{t-1}(b; \theta^\infty)$. Then $\chi_t(b; \theta^\infty) = x_{\theta_t}(\chi_{t-1}(b, \theta^\infty)) \leq x_{\theta_t}\left(\chi^o_{t-1}(b, \theta^\infty)\right) \leq x^o_{\theta_t}\left(\chi^o_{t-1}(b, \theta^\infty)\right) = \chi^o_t(b, \theta^\infty)$, which verifies the inductive step.

Take any arbitrary $\theta^\infty \in \Theta$, let θ^t be its first t elements. Now $\zeta(b, \theta^\infty) = \min\{t | \chi_t(b; \theta^\infty) = \underline{x}\}$ and $\zeta^o(b, \theta^\infty) = \min\{t | \chi^o_t(b; \theta^\infty) = \underline{x}\}$. Since $\chi_t(b; \theta^\infty) > \underline{x}$, for all $t < \zeta(b, \theta^\infty)$ then $\chi^o_t(b; \theta^\infty) > \underline{x}$ as well. Hence $\zeta^o(b, \theta^\infty) > \zeta(b, \theta^\infty)$. Moreover, since θ^∞ was chosen arbitrarily, this is true for every $\theta^\infty \in \Theta$. □

Proof of Lemma 11. To see (1), let $\theta^G = (G, G, ...)$. Since $\tau_G(b) < \tau^o_G(b)$, then $x_G(b) > x^o_G(b)$. Clearly if $\chi_{t-1}(b, \theta^\infty) \geq \chi^o_{t-1}(b, \theta^\infty)$ and $\theta_t = G$, then $\chi_t(b, \theta^\infty) = x_G(\chi_{t-1}(b, \theta^\infty)) >$

$x_G^o(\chi_{t-1}(b,\theta^\infty)) \geq x_G^o(\chi_{t-1}^o(b,\theta^\infty))=\chi_t(b,\theta^\infty)$. Then, taking $\theta^\infty=\theta^G$ gives $\chi_t(b,\theta^G) > \chi_t^o(b,\theta^G)$ whenever $t<\zeta(b,\theta^G)$. Hence $\zeta^o(b,\theta^G) \leq \zeta(b,\theta^G)$ and this is true for any b. (More generally, it is true for any $\theta \in \Theta$ whose first $\zeta^o(b,\theta^G)$ terms are G.)

To see (2), take any $\theta^\infty \in \Theta$ s.t. $\theta_1 = B$ and $\theta_2 = G$. Recall x_B and x_B^o are continuous and onto $[\underline{x},\overline{x}]$. By assumption $b < x_B(b) < x_B^o(b)$ whenever $x_B(b) < \overline{x}$. Hence, there exists $b' \in (\underline{x}, b_B^*)$ s.t. $x_B(b') < b_B^* < x_B^o(b')$. But then $x_G(x_B(b')) = \underline{x}$ and $x_G^o(x_B^o(b')) > \underline{x}$ (since $x_G(b) = \underline{x}$ iff $b \leq b_\theta^*$). Hence there exists $b' \in [\underline{x},\overline{x}]$ s.t. $\zeta(b',\theta^\infty) < \zeta^o(b',\theta^\infty)$. □

Proof of Lemma 12. I need to show that Lemmata 6, 7 and 5 continue to hold when there is incomplete pandering. The proof of Lemma 6 depended only upon the concavity of the value functions. Since the value functions remain concave with incomplete pandering, the results carry through. Since the value functions are concave, the proof method for Lemma 7 will carry over, so long as we can prove that: $\frac{dV_G^i(b)}{db} > \frac{dV_B^i(b)}{db}$ and $\frac{dW_G^i(b)}{db} > \frac{dW_B^i(b)}{db}$ (where $i \in \{n,s\}$ indicates whether the model has naive or sophisticated voters).

Finally, to see that the results of Lemma 5 hold, I need to verify that $x_\theta^* \leq b_\theta^*$. Consider the sophisticated case, and suppose $x_\theta^* > b_\theta^*$ for some $\theta \in \{G,B\}$. By the first order conditions, $-\delta\left\{\phi E_\theta^v\left[\frac{dV_{\theta'}^i(x_\theta^*)}{db}\right] + (1-\phi)E_\theta^p\left[\frac{dW_{\theta'}^i(x_\theta^*)}{db}\right]\right\} = 1$. By the Envelope Theorem, $\phi\frac{dV_{\theta'}^i(b)}{db} + (1-\phi)\frac{dW_{\theta'}^i(b)}{db} = -(1+r)$ for any $b \leq b_\theta^*$ and $\phi\frac{dV_{\theta'}^i(b)}{db} + (1-\phi)\frac{dW_{\theta'}^i(b)}{db} < -(1+r)$ for any $b > b_\theta^*$. Hence $E_\theta\left[\phi\frac{dV_{\theta'}^i(b)}{db} + (1-\phi)\frac{dW_{\theta'}^i(b)}{db}\right] < -(1+r)$ whenever $b > \min\{b_G^*, b_B^*\}$, where the expectation is taken with respect to any non-degenerate beliefs. I claim that, for any b, there exists $\kappa(b) \in (0,1)$ s.t.

$$\begin{aligned}\phi E_\theta^v\left[\frac{dV_{\theta'}^i(b)}{db}\right] + (1-\phi)E_\theta^p\left[\frac{dW_{\theta'}^i(b)}{db}\right] &= \kappa E_\theta^v\left[\phi\frac{dV_{\theta'}^i(b)}{db} + (1-\phi)\frac{dW_{\theta'}^i(b)}{db}\right] \\ &\quad + (1-\kappa)E_\theta^p\left[\phi\frac{dV_{\theta'}^i(b)}{db} + (1-\phi)\frac{dW_{\theta'}^i(b)}{db}\right]\end{aligned}$$

If so, then since $x_\theta^* > b_\theta^*$, $\phi E_\theta^v\left[\frac{dV_{\theta'}^i(x_\theta^*)}{db}\right] + (1-\phi)E_\theta^p\left[\frac{dW_{\theta'}^i(x_\theta^*)}{db}\right] < -(1+r)$, which violates the first order condition. Hence $x_\theta^* \leq b_\theta^*$, and so by arguments in the proof of Lemma 5,

$x_B^* = x_G^* = \underline{x}$, $b_B^* = \underline{x}$ and $b_G^* = \underline{x} + \frac{(A_B - A_G)}{1+r}$.

It remains to show that my claim (above) is true. Note that:

$$\begin{aligned} \phi E_\theta^v \left[\frac{dV_{\theta'}^i(b)}{db}\right] + (1-\phi) E_\theta^p \left[\frac{dW_{\theta'}^i(b)}{db}\right] &= E_\theta^v \left[\phi \frac{dV_{\theta'}^i(b)}{db} + (1-\phi)\frac{dW_{\theta'}^i(b)}{db}\right] + \left(p_{\theta G}^p - p_{\theta G}^v\right)\left(\frac{dW_G^i(b)}{db} - \frac{dW_B^i(b)}{db}\right) \\ &= E_\theta^p \left[\phi \frac{dV_{\theta'}^i(b)}{db} + (1-\phi)\frac{dW_{\theta'}^i(b)}{db}\right] + \left(p_{\theta G}^v - p_{\theta G}^p\right)\left(\frac{dV_G^i(b)}{db} - \frac{dV_B^i(b)}{db}\right) \end{aligned}$$

We showed above that $\frac{dV_G^i(b)}{db} - \frac{dV_B^i(b)}{db} > 0$ and $\frac{dW_G^i(b)}{db} - \frac{dW_B^i(b)}{db} > 0$. For notational convenience, denote the LHS by α, the expectation on the first line of the RHS by β and the expectation on the second line on the RHS by γ. Clearly either $\beta < \alpha < \gamma$ or $\gamma < \alpha < \beta$. The claim follows by the intermediate value theorem. □

Proof of Lemma 14. Suppose $\theta = \theta'$. Then $p_{\theta\theta'}^v > p_{\theta\theta'}^p$ and so $\gamma_t(\theta) p_{\theta\theta'}^v + (1-\gamma_t(\theta)) p_{\theta\theta'}^p < p_{\theta\theta'}^v$. Then by (3.7)

$$\gamma_{t+1}(\theta') = \frac{\gamma_t(\theta) p_{\theta\theta'}^v}{\gamma_t(\theta) p_{\theta\theta'}^v + (1-\gamma_t(\theta)) p_{\theta\theta'}^p} > \frac{\gamma_t(\theta) p_{\theta\theta'}^v}{p_{\theta\theta'}^v} = \gamma_t(\theta)$$

If $\theta \neq \theta'$, then $p_{\theta\theta'}^v < p_{\theta\theta'}^p$ and so $\gamma_t(\theta) p_{\theta\theta'}^v + (1-\gamma_t(\theta)) p_{\theta\theta'}^p > p_{\theta\theta'}^v$, which implies the converse. The results for $\gamma \in \{0,1\}$ are obvious by inspection of (3.7). □

Proof of Lemma 15. I show that γ_t is a super-martingale. Then, by the Supermartingale Convergence Theorem, γ_t converges to its lower bound. Hence $\gamma_t \to 0$. The remainder of the proof establishes that γ_t is a super-martingale. Let $\theta_t = \theta$. Then, by (3.7):

$$\begin{aligned} E\left[\gamma_{t+1} | \gamma_t(\theta)\right] &= p_{\theta G}^p \gamma_{t+1}(H) + \left(1 - p_{\theta G}^p\right)\gamma_{t+1}(L) \\ &= p_{\theta G}^p \left[\frac{\gamma_t(\theta) p_{\theta G}^v}{\gamma_t(\theta) p_{\theta G}^v + (1-\gamma_t(\theta)) p_{\theta G}^p}\right] + \left(1 - p_{\theta G}^p\right)\left[\frac{\gamma_t(\theta)(1 - p_{\theta G}^v)}{\gamma_t(\theta)\left(1 - p_{\theta G}^v\right) + (1-\gamma_t(\theta))\left(1 - p_{\theta G}^p\right)}\right] \\ &= \gamma_t \left[\frac{p_{\theta G}^p}{1 + (1-\gamma_t(\theta))\frac{p_{\theta G}^p - p_{\theta G}^v}{p_{\theta G}^v}} + \frac{1 - p_{\theta G}^p}{1 - (1-\gamma_t(\theta))\frac{p_{\theta G}^p - p_{\theta G}^v}{1 - p_{\theta G}^v}}\right] \\ &= \gamma_t \frac{p_{\theta G}^p\left[1 - (1-\gamma_t(\theta))\frac{p_{\theta G}^p - p_{\theta G}^v}{1-p_{\theta G}^v}\right] + \left(1 - p_{\theta G}^p\right)\left[1 + (1-\gamma_t(\theta))\frac{p_{\theta G}^p - p_{\theta G}^v}{p_{\theta G}^v}\right]}{\left[1 + (1-\gamma_t(\theta))\frac{p_{\theta G}^p - p_{\theta G}^v}{p_{\theta G}^v}\right]\left[1 - (1-\gamma_t(\theta))\frac{p_{\theta G}^p - p_{\theta G}^v}{1-p_{\theta G}^v}\right]} \\ &= \gamma_t \frac{1 - A(\gamma_t(\theta), \theta)\left[p_{\theta G}^v + p_{\theta G}^p - 1\right]}{1 - A(\gamma_t, \theta)\left[(1+\gamma_t(\theta)) p_{\theta G}^v + (1-\gamma_t(\theta)) p_{\theta G}^p - 1\right]} \end{aligned}$$

where $A = (1-\gamma_t(\theta))\frac{p_{\theta G}^p - p_{\theta G}^v}{p_{\theta G}^v\left(1 - p_{\theta G}^v\right)}$. Let $\kappa = \frac{1 - A(\gamma_t(\theta),\theta)\left[p_{\theta G}^v + p_{\theta G}^p - 1\right]}{1 - A(\gamma_t,\theta)\left[(1+\gamma_t(\theta)) p_{\theta G}^v + (1-\gamma_t(\theta)) p_{\theta G}^p - 1\right]}$.

Suppose $\theta = G$. Then $p^p_{\theta G} - p^v_{\theta G} < 0$ and so $A < 0$. Moreover,

$$\begin{aligned} & (1 + \gamma_t(\theta)) p^v_{\theta G} + (1 - \gamma_t(\theta)) p^p_{\theta G} - 1 \\ = \; & p^v_{\theta G} + p^p_{\theta G} - 1 + \gamma_t(\theta) [p^v_{\theta G} - p^p_{\theta G}] \\ > \; & p^v_{\theta G} + p^p_{\theta G} - 1 \end{aligned}$$

Hence, $\kappa < 1$ and so $E[\gamma_{t+1} | \gamma_t(G)] < \gamma_t(G)$.

Suppose, instead, $\theta = B$. Then $p^p_{\theta G} - p^v_{\theta G} > 0$ and so $A > 0$. Moreover,

$$\begin{aligned} & (1 + \gamma_t(\theta)) p^v_{\theta G} + (1 - \gamma_t(\theta)) p^p_{\theta G} - 1 \\ = \; & p^v_{\theta G} + p^p_{\theta G} - 1 + \gamma_t(\theta) [p^v_{\theta G} - p^p_{\theta G}] \\ < \; & p^v_{\theta G} + p^p_{\theta G} - 1 \end{aligned}$$

Hence, $\kappa > 1$, and so $E[\gamma_{t+1} | \gamma_t(B)] < \gamma_t(B)$. Hence $E[\gamma_{t+1} | \gamma_t(\theta)] < \gamma_t(\theta)$ and so γ_t is a supermartingale. □

3.10 References

Acemoglu, D., M. Golosov, and A. Tsyvinski, "Political economy of mechanisms," *Econometrica*, 2008, *76* (3), 619–641.

Aiyagari, S.R., A. Marcet, T.J. Sargent, and J. Seppala, "Optimal Taxation without State-Contingent Debt," *Journal of Political Economy*, 2002, *110* (6).

Alesina, A. and G. Tabellini, "A positive theory of fiscal deficits and government debt," *The Review of Economic Studies*, 1990, *57* (3), 403.

Associates, Hart Research and Public Opinion Strategies, "Voters' Attitudes Toward the Budget Deficit & National Debt," February 2009.

Battaglini, M., "A Dynamic theory of electoral competition," 2011.

_ **and S. Coate**, "Inefficiency in Legislative Policymaking: A Dynamic Analysis," *The American Economic Review*, 2007, *97* (1), 118–149.

_ **and** _ , "A Dynamic Theory of Public Spending, Taxation, and Debt," *The American Economic Review*, 2008, *98* (1), 201–236.

_ **and** _ , "Fiscal Policy over the Real Business Cycle: A Positive Theory," 2011.

Benveniste, L.M. and J.A. Scheinkman, "On the differentiability of the value function in dynamic models of economics," *Econometrica: Journal of the Econometric Society*, 1979, pp. 727–732.

Black, D., "On the rationale of group decision-making," *The Journal of Political Economy*, 1948, *56* (1), 23–34.

Calvert, R.L., "Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence," *American Journal of Political Science*, 1985, pp. 69–95.

Downs, A., "An economic theory of political action in a democracy," *The Journal of Political Economy*, 1957, *65* (2), 135–150.

Fuster, A., B. Hebert, and D. Laibson, "Natural Expectations," *Macroeconomic Dynamics, and Asset Pricing, NBER Macroeconomics Annual, Forthcoming*, 2011.

Gallup, "Gallup Poll, May 5-8, 2011," May 2011.

Greenspan, A., "Cyclicality and banking regulation," in "Proceedings" number May Federal Reserve Bank of Chicago 2002, pp. 1–6.

Keynes, J.M., *The general theory of interest, employment and money*, London: Macmillan, 1936.

Lindbeck, A. and J.W. Weibull, "Balanced-budget redistribution as the outcome of political competition," *Public Choice*, 1987, *52* (3), 273–297.

Loewenstein, G., T. O'Donoghue, and M. Rabin, "Projection bias in predicting future utility," *The Quarterly Journal of Economics*, 2003, *118* (4), 1209–1248.

O'Donoghue, T. and M. Rabin, "Doing it now or later," *American Economic Review*, 1999, pp. 103–124.

of Michigan, University, "Consumer Sentiment Index."

Office, Congressional Budget, "CBO's 2011 Long-Term Budget Outlook," June 2011.

_ , "An Update to the Budget and Economin Outlook: Fiscal Years 2012 to 2022," August 2012.

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

_ **and L.E.O. Svensson**, "Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences," *The Quarterly Journal of Economics*, 1989, *104* (2), 325.

Pigou, A.C., *Industrial fluctuations*, Macmillan, 1927.

Rotheli, T.F., "Boundedly Rational Banks' Contribution to the Credit Cycle," *The Journal of Socio-Economics*, 2012.

Tabellini, G. and A. Alesina, "Voting on the budget deficit," *The American Economic Review*, 1990, pp. 37–49.

Wittman, D., "Candidates with policy preferences: A dynamic model," *Journal of Economic Theory*, 1977, *14* (1), 180–189.

Yared, P., "Politicians, taxes and debt," *Review of Economic Studies*, 2010, *77* (2), 806–840.

4 Less Representation is Better: How Bicameralism Can Benefit Large States

4.1 Introduction

Many of the world's democracies feature bicameral legislatures. The goal of bicameralism is to encourage desirable policy outcomes by requiring that proposed policies receive the support of concurrent majorities in multiple chambers. In almost all countries, representatives in the lower house are elected directly by the people and with equal weight given to each voter (Tsebelis and Money, 1997). However, the nature of representation in the upper house varies across different countries. In Italy and Japan and in all the bicameral state legislatures in the United States, the composition of the upper house tends to replicate the lower house - so that the chambers tend to be very similar. By contrast, in many countries, representation in the upper house is less congruent. In all federations, representation in the upper house is by geographic region rather than population size, and this often has the effect of over-representing regions with smaller populations (Tsebelis and Money, 1997). Nations such as Australia, Mexico, Russia, Switzerland and the United States, amongst others, have an upper house in which the constituent states are represented equally, regardless of population size (Lijphart, 1999). (Indeed, in Australia and the United States, the upper house is frequently referred to as the "States' House".) The European Council, which affords equal representation to member states, plays a similar role in the legislative branch of the European Union, and stands in contrast to the European Parliament, where countries are represented in proportion to population size. In other nations, such as Canada and Germany, states are not necessarily equally represented, however, the essential feature that smaller states are overrepresented, tends to remains true (Tsebelis and Money, 1997). Even in some unitary states - Burundi, South Africa and Spain amongst them - representatives in the upper house are drawn from broad administrative regions, not necessarily in proportion to population size (CIA, 2011).

A strong motivation for geographically-based representation stems from a tension between the interests of large and small regions. In the United States, the over-representation of small regions in the upper house was a compromise to the small states, who feared that their interests would be ignored in the popular chamber where the large states could typically command a majority between themselves. James Madison noted in the Federalist Papers:

> "The equality of representation in the Senate.... being evidently the result of compromise between the opposite pretensions of the large and the small States... " "A government founded on principles more consonant to the wishes of the larger States, is not likely to be obtained from the smaller States." - James Madison, Federalist no. 62

Indeed, the structure of the United States Senate was a compromise between the large states (led by Virginia) who favoured a bicameral legislature with representation in both chambers in proportion to population size, and the small states (led by Delaware and New Jersey) who favoured a unicameral legislature with equal representation for each state. A similar compromise, to entice smaller states, such as Western Australia, into the federation, informed the design of the Australian Senate. The logic seems straightforward enough - since the smaller states have relatively more representation in the upper house, they will be protected against usurpation of the policy agenda by the larger states.

In this paper, I argue that the above intiution is not as clear-cut as may first appear - and that over a range of scenarios that may plausibly prevail, large states may do even better in a bicameral legislature than they would if there were only a single chamber with representation in proportion to population size. I model the bicameral legislative decision making process using the framework developed by Baron and Ferejohn (1989). In this framework, the legislature must allocate a fixed surplus amongst various districts, each of which are represented by a legislator. A legislator is chosen to propose an allocation, and the proposal is implemented if a majority of legislators support the proposal. The proposer will build a majority coalition by allocating the surplus to the 'cheapest' legislators, since this maximises the residual that he keeps for himself. The cost of a legislator depends on the expected share of the surplus that a legislator would receive, if they were to reject a proposal and wait for the next proposal to arrive. This cost is determined in equilibrium, and is increasing in the probability that the legislator will be recognised as the proposer. A standard result in legislative bargaining is that the residual surplus is sufficiently large that it is optimal for the proposer to support his own proposal. This implies that the proposer needs to bribe one fewer legislator to support his proposal than the majority requirement - since he can always count on his own support.

I extend the Baron and Ferejohn framework to a model with two chambers - a lower house that represents districts and an upper house that represents states. Small states contain fewer districts than big states. Legislators in the lower house seek to maximise the allocation to their district alone (without reference to other districts in that state), whilst upper house legislators from a big state seek to maximise the total allocation to the districts in their state. This creates a complementarity between the payoffs of lower and upper house agents from the same state. In this extended framework, it remains true that the proposer - as the residual claimant - will support his own proposal. In addition, the complementarity in preferences ensures that the legislator(s) from the same state as the proposer, but in the other chamber, will also support the proposal. Lemma 17 states this claim more precisely. This reduces the number of legislators that need to be bribed in both chambers. Moreover, when the proposer is the upper house legislator from a big state, he needs to bribe many fewer legislators (than an upper house legislator from a small state) because complementarity guarantees that he gets the support of more legislators in the lower house for free. I refer to this as the *reduced requirement effect.* Since big state proposers do not need to buy as many coalition partners, the residual surplus that they keep for themselves is larger - and this tends to increase their *ex ante* expected share of the surplus. Furthermore, if the equilibrium were such that -

without this effect, the small state districts were cheaper - then adding the effect reduces the payoff to small states, since these are the coalition partners who are effectively being displaced. Proposition 17 makes precise this claim, that if the distribution of the surplus under unicameralism favours the big states, then the distribution under bicameralism will tend to favour big states even more. The key insight of the paper is that the correlation in preferences, along with the big states' numerical advantage in the lower house, reduces the number of agents that an upper house legislator from the big state needs to entice into his coalition - and this increases the expected payoff to big states.

To make this point stark, consider the following example: There are 4 small states each containing 1 district and 1 large state comprising 5 districts - so that there are 9 districts and 5 states in total. Each district has one representative in the lower house and each state has one representative in the upper house. A policy is adopted by the legislature if it has the support of a simple majority in each chamber (i.e. 5 lower house and 3 upper house agents). First consider a unicameral legislature and suppose that in equilibrium, the *ex ante* expected payoff to small state districts is smaller than the expected payoff to big state districts. This makes small districts cheaper to bribe and so they will be included in the coalition more frequently. Even when the proposer is from a big district - it is cheaper for her to invite the small state legislators into the coalition than to bribe the more expensive big districts. When the proposer is from a big state district, she will invite all four small state legislators into the coalition (since she will support the proposal herself). If the proposer is from a small state, then he will invite the remaining three small states and one large state district to join the coalition. Hence, small districts will be included in every coalition. They will receive their continuation value whenever they are not the proposer, and they will claim the residual surplus (which is strictly greater) whenever they are the proposer. (This is still consistent with the small state getting a smaller slice of the pie than a big district, if the probability of being the proposer and getting the residual is small enough.)

Now consider the bicameral legislature. If the proposer is from a small state, or is a lower house legislator from the big state, then the optimal coalition is unchanged. (The optimal coalition that generates a majority in the lower house also generates a majority in the upper house.) However, if the proposer is the upper house legislator from the big state, then a majority in the lower house automatically exists, since complementarity implies that all five big state legislators in the lower house will support the proposal. The proposer will simply invite two (out of the four) small state legislators to join the coalition, to satisfy the upper house majority constraint. Hence, in the bicameral framework, small state districts are sometimes excluded from the coalition, whereas they were never excluded in the unicameral legislature. Moreover, resources are more frequently retained by the big state districts, since they share these resources whenever the upper house legislator from the big state is the proposer. This has the effect of reducing the expected payoff to small states in the bicameral setting, relative to the expected payoff in the unicameral game. (For simplicity, I have assumed that the aggregate probability that the proposer is from a big state is the same in both unicameral and bicameral settings. In the formal analysis, I allow the proposal probabilities to change, and show that the result obtains as long as bicameralism does not increase too much the likelihood that the proposer is from a small state.)

The focus of this paper is on the distributional consequences of legislative institutions over a distributive policy space, where the legislators bargain over the division of a pie - as opposed to more general spatial models of policy. This restriction enables me to easily characterise the preferences of legislator - in this case legislators are assumed to prefer allocations in which a greater share of the pie is allocated to their constituency. In particular, the restriction allows for the preferences of upper and lower house legislators from the same state to be linked in a natural way. An upper house legislator prefers any policy that allocates more of the pie to at least one district within his state, *ceteris paribus*. As alluded to above, this complementarity between preferences of legislators from the same state is crucial to the paper's main insight.

The bargaining protocol is an extension of the seminal model of legislative bargaining by Baron and Ferejohn (1989) and its generalisation in Banks and Duggan (2000),(2006). Each period, the legislature convenes to consider a policy question, and a legislator is chosen to propose a policy. The legislature then simultaneously votes to either accept or reject the proposal. If a majority support the proposal, then it is implemented and the game ends. If not, the legislature reconvenes in the following period, and the above procedure is repeated. In these models, a legislator's equilibrium payoff (and consequently, her willingness to compromise) is a function of her recognition probability - the likelihood that she will be called upon to propose a policy - and the legislator's degree of patience. Kalandrakis (2006) shows that any division of the pie can be sustained as a stationary equilibrium, given an appropriate assignment of proposal power amongst the various legislators. In this paper, I show that if proposal rights are assigned such that big states are privileged under unicameralism, then they will likely be even further privileged under bicameralism. Furthermore, the equilibrium that I construct is unique, so it cannot be seen as merely an aberrant or exotic instance from amongst a set of possible equilibria. (Whilst the equilibrium of legislative bargaining models is typically not unique (see Banks and Duggan (2006)), Eraslan (2002) has shown the uniqueness of the stationary equilibrium payoffs when the policy space is distributional.) It is stressed that the result in this paper arises as a consequence of the composition of the equilibrium coalition - not from the recognition probabilities per se. Indeed, in contrast to the result in Kalandrakis (2006), the big states can benefit from bicameralism, even if their proposal power *falls* overall, when the second chamber is added.

Bicameralism, as a feature of legislative institutions, has generated significant interest in the recent literature. Tsebelis and Money (1997) and Cutrone and McCarty (2006) provide an extensive summary of the existing literature. The model of bicameralism used in this paper draws on the work of McCarty (2000), Ansolabehere et al. (2003) and in particular, Kalandrakis (2004). Kalandrakis presents a model in which there are two types of states - big and small - who send delegations to both houses of the legislature. The government surplus is allocated at the state level (rather than at the finer level of districts within states), and so the preferences of all legislators from a given state are perfectly aligned. By contrast, in this model, the preferences of lower house legislators from the same state are uncorrelated - each lower house legislator simply wants to maximise the share accruing to his district. However, the preferences of each lower house agent are correlated with the preferences of the upper house legislator from that state. This 'imperfect' correlation of preferences is crucial in generating the benefit of the upper house to the large states.

Ansolabehere et al. (2003) consider a model in which the distribution of the surplus is at the district level. In their model, legislators are perfect agents of the median voter in their district. The preferences of the upper house legislator from a given state are aligned with the preferences of the lower house legislator from the district within his state that receives the median allocation, given a proposed policy. This approach is slightly awkward in that the identity of the median district changes as different allocations are considered. Moreover, since the upper house legislator is concerned only with the payoff to a single district within his state, he will accept proposals for which the aggregate allocation to his state is smaller than the amount that his state could expect in the continuation game. Such behaviour is at odds with the standard equilibrium strategies in models of legislative bargaining. This problem does not arise in the model in this paper, because legislators are assumed to maximise the allocation to all the districts (voters) in their constituency - not just a subset of them.

The model is distinguished from other models of bicameralism, such as Hammond and Miller (1987), McCarty (2000) and Diermeier and Myerson (1999), that do not account for the complementarities between preferences of legislators from the same state. McCarty considers a bargaining model in the style of Baron and Ferejohn, with multiple veto players - however, his model does not explicitly introduce complementarity in legislator preferences. Diermeier & Myerson consider a model in which lobbies "buy" the votes of legislators, for example by making campaign donations. In this model, legislators' preferences are uncorrelated and depend only on the size of the bribe they individually receive. Hammond & Miller consider a spatial model of bicameralism. Their model predicts that the core may exist in a bicameral legislature (unlike in unicameral legislatures with a simple majority rule - where the core generically does not exist) - provided that upper and lower house agents have preferences that are sufficiently distinct. This result is curious in that, whilst it conforms to standard intuitions in the literature about the benefits of bicameralism in providing policy stability, it makes precisely that opposite assumption to this paper - that upper and lower house agents' preferences are sufficiently different.

This paper proceeds as follows. Subsection 2 outlines the model of bicameralism. Subsection 3 provides a benchmark for the analysis, by characterising the equilibrium of an analogous unicameral legislature and subsection 4 characterises the equilibrium of the bicameral model. Subsection 5 compares the equilibria in the unicameral and bicameral settings. Subsection 6 concludes.

4.2 Model

There is a polity that is divided into geographical regions - or states. There are two types of states - big and small. For simplicity, I assume that there are $s > 1$ small states, and $b \geq 1$ big state. States are indexed by $j \in \{1, ..., s+b\}$. The polity is also divided into electoral districts, each containing roughly the same number of voters. I assume that each small state contains just one electoral district (i.e. the state and district coincide), whilst each big state contains $k > 1$ districts. Electoral districts are indexed by$i \in \{1, ..., s+kb\}$. There is function $\rho : \{1, ..., s+kb\} \to \{1, ..., s+b\}$ that maps each electoral district into its respective state.

The government must determine the allocation of resources amongst the various districts in the polity. Examples of such policies include the allocation of funding for highways or for other local public goods. The size of the overall supply of government resources is normalised to 1. A policy is a vector $x = (x_1, ..., x_{s+bk})$, where x_i is the share of the pie that is received by the i^{th} district. It is assumed that the congressional district is the finest level at which legislators may direct resources. The set of feasible allocation is given by the $(s+bk-1)$-

dimensional simplex: $X = \left\{ x \in \Re^{s+bk} | x_i \geq 0 \, \forall i \text{ and } \sum_{i=1}^{s+bk} x_i \leq 1 \right\}$.

Government policy is determined by a legislature comprised of two chambers - an upper and lower house. In the upper house each state is represented by one legislator, whilst in the lower house each district is represented by a legislator. The procedure by which the legislature adopts policy is based on the framework of Baron and Ferejohn (1989). In a given period, a member of the legislature is chosen at random to propose a division of the 'pie'. Once the proposal is made, the legislators in both chamber simultaneously vote to either accept or reject the proposal. The proposal is accepted if $M_L > \frac{s+kb}{2}$ legislators in the lower house and $M_U > \frac{s+b}{2}$ legislators in the upper house vote to accept the proposal. If the proposal is accepted, then the allocation is implemented and the game ends. If the proposal is rejected, then the pie is not allocated. The legislature adjourns and reconvenes in the following period, when the above procedure is repeated. This process continues until a proposal is accepted.

The above decision procedure is consistent with many models of bicameralism including Ansolabehere et al. (2003), Banks and Duggan (2000), Kalandrakis (2004), McCarty (2000) and others. Nevertheless, the procedure simplifies the bargaining dynamic by assuming that offers can forever continued to be made until a resolution is achieved. The model asserts that the navette continues until the conflict is resolved. Tsebelis and Money (1997) note that the navette is the most common method of resolving these disputes, and as such, I argue that this is a reasonable modeling simplification. I acknowledge, however, that in some legislature, alternative dispute resolution mechanisms (such as a conference committee, or dissolution of the legislature) are used in the event that navette fails to resolve disputes within the first few rounds. Moreover, I assert that this simplification is benign, since the effect that this paper seeks to explain operates somewhat independently of the details of the bargaining protocol. (The bargaining protocol determines the continuation values along the equilibrium path, and so the details of the protocol typically matter only insofar as they affect these values. However, the effect described in this paper arises out of the differences in the composition of the optimal coalition between unicameralism and bicameralism, for *given* continuation values.)

Legislators are risk neutral and have stage game preferences represented by the size of the allocation to their district. Let u_i^L and u_j^U represent the von-Neumann Morgenstern indicies of the i^{th}lower house agent and j^{th}upper house agent, respectively. Then $u_i^L(x) = x_i \; \forall i \in \{1, ..., s + kb\}$, and $u_j^U(x) = \sum_{i \in \rho^{-1}(j)} x_i \, \forall j \in \{1, ..., s + b\}$. Since their constituencies overlap - there is a complementarity between the utilities of lower house agents from a given state and the upper house agent from that state. Indeed, the preferences of agents from the same small state perfectly coincide. In large states, the preferences of lower house agents are unrelated, however an allocation that improves the utility of at least one lower house agent also improves the utility of the upper house agent from that big state, *ceteris paribus*. This complementarity is crucial to the analysis that follows. All agents share a common discount factor $\delta \in [0, 1)$.

I will restrict attention to strategies and equilibria that are stationary and symmetric. This requires that, whilst agents might distinguish between districts or legislators from different sized states, they will not arbitrarily distinguish between districts or agents from states of the same size. Symmetry ensures that the outcomes generated in this paper are a consequence of differences in state size, rather than some other arbitrary or unmodelled factor. Since this paper focuses on the distributional consequences of bicameralism between large and small states, the symmetry assumption should be seen to be relatively benign.

Denote the set of legislator types by $T = \{S^L, S^U, B^L, B^U\}$, where S^L refers to a lower house legislator from a small state, and the other types are similarly defined. It should be clear that since the upper and lower house agents from small states have identical preferences (because their constituencies coincide), that they will choose identical strategies. However, the same is not true for upper and lower house agents from the large state, since the upper house agent cares about the allocation going to the entire state, whilst a lower house agent only cares about the allocation accruing to her own district.

Let p_t be the probability that a type-t legislator is recognized as the proposer. These probabilities must satisfy: $b\left(kp_{B^L}+p_{B^U}\right)+s\left(p_{S^L}+p_{S^U}\right)=1$.

Denote by $P = b\left(kp_{B^L}+p_{B^U}\right)$ the probability that the proposer is from a large state. Further, define α_B and α_S as the conditional probability that a legislator is from the lower house, given that she is from a big and small state, respectively:

$$\alpha_B = \frac{kp_{B^L}}{kp_{B^L}+p_{B^U}}$$

$$\alpha_S = \frac{p_{S^L}}{p_{S^L}+p_{S^U}}$$

The triple (P, α_B, α_S) fully characterises the recognition probability of each type. Indeed, it is easily verified that: $p_{B^L} = \frac{\alpha_B P}{bk}$, $p_{B^U} = \frac{(1-\alpha_B)P}{b}$, $p_{S^L} = \frac{\alpha_S(1-P)}{s}$ and $p_{S^U} = \frac{(1-\alpha_S)(1-P)}{s}$. It is stressed that this formulation places no restriction on the possible values that the recognition probabilities can take. In particular, recognition probabilities are not assumed to be independent across chambers. Finally, to simplify notation, I denote by P_L and P_U, the probabilities that the proposer is from a big state conditional upon the proposer being in the lower and upper house (respectively). These conditional probabilities are defined by:

$$\begin{aligned} P_L &= \frac{bkp_{B^L}}{bkp_{B^L}+sp_{S^L}} = \frac{\alpha_B P}{\alpha_B P + \alpha_S(1-P)} \\ P_U &= \frac{bp_{B^U}}{bp_{B^U}+sp_{S^U}} = \frac{(1-\alpha_B)P}{1-\left[\alpha_B P + \alpha_S(1-P)\right]} \end{aligned}$$

Let $\mu \in \Delta(X)$ be a probability mixture over the set of feasible offers. A stationary, symmetric strategy for a type-t legislator is a randomisation over proposals, μ_t, whenever they are the proposer and a decision rule $a_t : X \to \{0,1\}$ which indicates whether they will accept or reject a given proposal when they are not the proposer. Existence of an equilibrium may require that the agent randomises between several possible policies, when he is the proposer. A type-t agent's acceptance set A_t is the set of proposals that the agent will accept given the decision rule a_t (i.e. $A_t = \{x \in X | a_t(x) = 1\}$). It will often prove more convenient to consider an agent's acceptance set, rather than the decision rule itself. The assumption of stationarity requires that agents choose the same action in every structurally equivalent sub-game. This amounts to asserting that strategies are history independent. A stationary, symmetric, sub-game perfect equilibrium is a set of strategies $\left\{(\mu_t, a_t)_{t\in T}\right\}$, such that, for each type $t \in T$, there is no other strategy (μ_t', a_t') which gives the agent strictly higher utility, given the strategies of all other players.

I restrict attention to no-delay equilibria, since these are most efficient. Banks and Duggan (2000),(2006) show that no-delay equilibria exist in a general bargaining environment that embeds this model. Let v_B and v_S be the expected shares of the pie that are distributed to the big and small districts, respectively, in equilibrium. These are also the ex ante equilibrium payoffs to legislators in the lower house. By the complementarity in utilities, the ex ante equilibrium payoffs to upper house legislators are v_S and kv_B for small and big states, respectively. I denote the equilibrium payoff for a type-t agent by w_t. Since this paper is concerned with the relative shares of the pie received by big and small states, I will often consider the ratio $v = \frac{v_S}{v_B}$. (Note that v can take any positive real value, and that as v increases, that share of the pie going to small districts increases. In particular, when $v = 1$, the pie is distributed equally amongst districts, and when $v = k$, the pie is allocated equally amongst states.) Feasibility of pay-offs requires that $bkv_B + sv_S \leq 1$. Moreover, in a no-delay equilibrium $bkv_B + sv_S = 1$, since the entire pie is consumed immediately.

I restrict attention to equilibria in which the agents' decision-rule satisfies the weak dominance property. This requires that an agent accepts a policy only if she weakly prefers the outcome under that policy to her expected payoff if the proposal were rejected. This implies: $A_t = \{x \in X | u_t(x) \geq \delta w_t\}$, since if the proposal is rejected, stationarity implies that the agent will receive her expected payoff w_t in the following period. Weak dominance rules out perverse or implausible equilibria - for example, equilibria in which every agent accepts every proposal, independent of their preference. (Such strategies can be sustained in equilibrium, since no single legislator can change the outcome by changing their vote to accept.)

In a no-delay equilibrium, the optimal proposal maximses the proposer's share of the pie, subject to the proposal being accepted by a winning coalition of legislators. To entice legislator i to support a proposal, the proposal must allocate at least δw_i amongst the districts that affect $i's$ utility. The proposer retains whatever is left of the pie for his own district(s) after 'buying' the support of a winning coalition of legislators. Since the proposer seeks to maximise this residual, he will 'purchase' the cheapest coalition, by offering exactly δv_j ($j \in \{B, S\}$) to the districts with the least expensive legislators. Since the residual surplus is allocated amongst the proposer's district(s), other legislators who also care about the allocation to those districts may also support the proposal - even though they weren't actively recruited into the coalition. Lemma 17 states this intuition more precisely. In particular, the proposer will always support his own proposal in equilibrium. Hence, I distinguish between coalition members who are 'purchased' and those (including the proposer) who 'come for free'.

Let β_t and σ_t be the number of big and small state districts (respectively) that a type-t proposer purchases (i.e. that do not come for free). If there are more than β_t big state districts available for purchase, then symmetry requires that the proposer randomly choose amongst them with equal probability. The same is true for the small state districts. Then a type-t proposer allocates δv_B to β_t randomly chosen big state districts and δv_S to σ_t randomly chosen small state districts, whilst the remaining $1 - \beta_t \delta v_B - \sigma_t \delta v_S$ is distributed amongst the district(s) that he represents. To characterise the equilibrium, it suffices to find the optimal (σ_t, β_t) for each type-t. Of course, in equilibrium, the proposer may mix across different (σ, β)-coalition pairs. Let $\mu_t(\sigma, \beta)$ be the probability that a type-t proposer builds a (σ, β)-coalition. Define $\bar{\sigma}_t = \sum_{(\sigma,\beta)} \mu_t(\sigma, \beta) \sigma$ and $\bar{\beta}_t = \sum_{(\sigma,\beta)} \mu_t(\sigma, \beta) \beta$, which are the expected number of small and big state districts that are purchased by a type-t proposer.

4.3 Unicameralism

I being by characterising the equilibrium distribution in a unicameral legislature, where representation is in proportion to population size. This will provide a natural point of comparison for the outcome under bicameralism. To facilitate comparison, I assume that the composition of, and majority requirement in, the unicameral legislature are identical to those in the lower house of the bicameral legislature. Hence, there are bk legislators from big state districts and s legislators from small state districts, and a proposal needs the support of $M_L > \frac{s+bk}{2}$ legislators to be accepted. In contrast to the bicameral case - every legislator in the unicameral model is associated with a single district, and the constituencies of legislators do not overlap.

I assume that the recognition probabilities in the unicameral system are identical to the probability that an agent is the proposer in the bicameral system, conditional upon that agent being in the lower house. As such, the recognition probabilities are given by:[18]:

$$\begin{aligned} p_B^u &= \frac{p_{B^L}}{bkp_{B^L} + sp_{S^L}} \\ p_S^u &= \frac{p_{S^L}}{bkp_{B^L} + sp_{S^L}} \end{aligned}$$

Moreover, the probability that the proposer is from a big state is given by $P_L = bkp_B^u$. (Note - since there is only one chamber, there are only two types of legislators in this setting - those from big state districts and those from small states.)

The remainder of this section is devoted to characterising the symmetric, stationary equilibrium of the unicameral game and replicates the equilibrium construction in Baron and Ferejohn (1989) or Eraslan (2002).

[18]I use a superscript u throughout to distinguish variables in the unicameral game

4.3.1 Optimal Coalition

Conjecture a feasible pair of expected equilibrium shares (v_S^u, v_B^u). To entice legislator i to support the proposal, the proposal must allocate at least δv_i^u to $i's$ district. Since the proposer seeks to maximise the residual that remains after 'buying' the support of a winning coalition, he will 'purchase' the cheapest coalition, by offering exactly δv_i^u ($i \in \{B, S\}$) to the districts with the least expensive legislators. If a winning coalition requires the assent of M_L legislators, the proposer needs only to buy the support of $M_L - 1$ other legislators, since he will always support his own proposal in equilibrium. To see why, note that even if the proposer builds the most expensive coalition by buying the support of every legislator other than himself, he still retains $1 - \sum_{j \neq i} \delta v_j^u = 1 - \delta \left(1 - v_i^u\right) > \delta v_i^u$. Since the proposer retains the residual surplus, the proposer is always strictly better off than he would be if he were simply a responder who was brought into the coalition.

Let β_t^u and σ_t^u be the number of big and small state districts (respectively) that a type-t proposer purchases (i.e. that do not come for free). The set of optimal coalitions satisfies the following property:

$$\begin{aligned} (\sigma_t^u, \beta_t^u) &\in \arg \min_{(\sigma,\beta)\in\mathbb{Z}_+^2} \sigma v_S^u + \beta v_B^u \\ \text{s.t. } \sigma + \beta &\geq M_L - 1 \\ \sigma &\leq s - \mathbf{1}_S\left[t\right] \\ \beta &\leq bk - \mathbf{1}_B\left[t\right] \end{aligned}$$

where $\mathbf{1}_S\left[t\right]$ ($\mathbf{1}_B\left[t\right]$) are indicator functions that takes value 1 if the proposer is from a small (big) state district, and 0 otherwise. The first constraint is the majority constraint, whist the last two constraints are the legislator supply constraints. The indicator functions account for the fact that the number of large and small state districts available for purchase depends upon whether the proposer is himself from a large or small state.

Let $\mu_t^u(\sigma,\beta)$ be the probability that a (σ,β) coalition pair is chosen by a type-t proposer, in equilibrium. The optimal proposal rule μ_t^u puts positive probability only on coalition pairs that are minimisers of the above problem. Since the objective and constraint functions in the cost minimisation problem are all linear, the set of minimisers will generally be a singleton, and if so, the proposal rule must put probability 1 on this coalition being proposed.

It should be obvious that if $v_S^u < v_B^u$ (i.e. $v^u < 1$), then the cost-minimising coalition will contain as many small state districts as needed, and will only include big state districts if all small districts are exhausted before the majority constraint is satisfied. If $v_S^u > v_B^u$ (i.e. $v^u > 1$), then the opposite is true, and the coalition is built by first including big state districts. The following proposition summarises the above discussion:

Lemma 16. *Given the equilibrium expected share ratio v^u, the optimal coalition for each type of proposer is given by:*

$$\left(\sigma_S^u,\beta_S^u\right) = \begin{cases}\left(\min\left\{M_L-1,s-1\right\},\max\left\{0,M_L-s\right\}\right) & v^u<1\\ \left(\max\left\{0,M_L-bk\right\},\min\left\{M_L-1,bk\right\}\right) & v^u>1\end{cases}$$

$$\left(\sigma_B^u,\beta_B^u\right) = \begin{cases}\left(\min\left\{M_L-1,s\right\},\max\left\{0,M_L-s\right\}\right) & v^u<1\\ \left(\max\left\{0,M_L-bk\right\},\min\left\{M_L-1,bk-1\right\}\right) & v^u>1\end{cases}$$

Furthermore, if $v^u=1$, then any mixture between the optimal coalitions for $v^u<1$ and $v^u>1$ is optimal.

The proof of the above proposition is straightforward and as such is omitted. The proposition can be viewed as a special case of Proposition 15 - which characterises the optimal coalition in the bicameral case - by setting $M_U=0$ and ignoring the cost problems of the upper house types. A proof of Proposition 15 is included in the Appendix.

Let $\bar{\sigma}^u(v^u)$ and $\bar{\beta}^u(v^u)$ denote the expected number of small and big districts that are included in the coalition for a given equilibrium share ratio v^u. $\bar{\sigma}^u$ and $\bar{\beta}^u$ are singleton-valued and piecewise constant on the intervals $v^u \in [0,1)$and $v^u \in (1,\infty)$, and convex-valued when $v^u = 1$. It is straightforward to show that $\bar{\sigma}^u$ and $\bar{\beta}^u$ are upper-hemicontinuous in v^u. Moreover, by inspection, $\bar{\sigma}^u$ is (weakly) decreasing and $\bar{\beta}^u$ is (weakly) increasing in v^u.

4.3.2 Equilibrium shares and payoffs

In the above section I characterised the optimal coalition given a pair of conjectured equilibrium shares. I now calculate the expected shares that are implied by these coalitions. Naturally, in equilibrium, the conjectured and implied expected shares must coincide. The expected share of the pie that is allocated to a big state district is characterised by:

$$\begin{aligned} v_B^u &= p_B^u \sum_{(\sigma,\beta)} \mu_B^u(\sigma,\beta)\left(1-\sigma\delta v_S^u - \beta_S \delta v_B^u\right) \\ &\quad + (bk-1)\, p_B^u \sum_{(\sigma,\beta)} \mu_B^u(\sigma,\beta)\frac{\beta}{bk-1}\delta v_B^u + s p_S^u \sum_{(\sigma,\beta)} \mu_S^u(\sigma,\beta)\frac{\beta}{bk}\delta v_B^u \\ v_B^u &= \frac{P_L}{bk} - \frac{P_L}{bk}\bar{\sigma}_B^u \delta v_S^u + (1-P_L)\frac{\bar{\beta}_S^u}{bk}\delta v_B^u \end{aligned} \tag{4.1}$$

The right hand side of this expression has three terms. The first, $\frac{P_L}{bk}$, is probability that the proposer is from a big state district. The second term is the expected amount that a big district proposer must offer to small districts to form a coalition, whilst the third term is the expected amount that a big district can expect to receive when the proposer is from a small district. (Note that $\frac{\bar{\beta}_S^u}{bk}$ is the probability that a big state district is included in the coalition when the proposer is from a small district.) The equilibrium share for a big state district then, is the amount that it would receive if the pie were allocated in proportion to its recognition probability, adjusted for its expected payouts to, and receipts from, small state districts.

The expected share of the pie allocated to small state districts is similarly characterised and interpreted:

$$v_S^u = \frac{1-P_L}{s} - \frac{1-P_L}{s}\bar{\beta}_S^u \delta v_B^u + P_L \frac{\bar{\sigma}_B^u}{s}\delta v_S^u \tag{4.2}$$

Solving (4.1) and (4.2) gives:

$$v^u = \frac{v_S^u}{v_B^u} = \frac{1-P_L}{P_L} \cdot \frac{bk - \delta\bar{\beta}_S^u}{s - \delta\bar{\sigma}_B^u} \tag{4.3}$$

which is analogous to (A.2) in Appendix A of Kalandrakis (2004) and Proposition 2 in McCarty (2000). Note that the equilibrium shares depend only upon $\bar{\beta}_S^u$ and $\bar{\sigma}_B^u$ - i.e. the number of districts of the *other* type that the proposer invites into the coalition.

An equilibrium is characterised by expected payoffs (v_S^u, v_B^u) and a probability assignment over coalitions $\{\mu_t^u\}_{t\in T}$ such that μ_t^u only puts positive weight on coalition pairs (σ, β) satisfying Lemma 16, given (v_S^u, v_B^u), and the payoff shares satisfy (4.3), given $\{\mu_t^u\}_{t\in T}$.

Define:

$$\phi^u(v^u) = \frac{bk - \delta\bar{\beta}_S^u}{s - \delta\bar{\sigma}_B^u} \tag{4.4}$$

Since the optimal coalitions $\{(\sigma_t^u, \beta_t^u)\}_{t\in T}$ are piece-wise constant on the intervals $[0,1)$and $(1,\infty)$, then so is $\phi^u(v^u)$. (For convenience, denote $\phi^u(v^u) = \phi_1^u$ when $v^u < 1$ and $\phi^u(v^u) = \phi_2^u$ when $v^u > 1$). Furthermore, ϕ^u is upper-hemicontinuous in v^u since it is a continuous function of $\bar{\sigma}_t^u$ and $\bar{\beta}_t^u$, which are upper-hemicontinuous. Moreover, ϕ^u is weakly decreasing in v^u, since $\bar{\sigma}_B^u$ is decreasing and $\bar{\beta}_S^u$ is increasing in v^u. Finally, since ϕ^u is upper-hemicontinuous, then for $v^u = 1$, ϕ^u is convex valued, and takes values in the interval $[\phi_2^u, \phi_1^u]$.

Given that the conjectured and implied shares must coincide, the equilibrium share ratio is characterised by $v^u \in \frac{1-P_L}{P_L}\phi^u(v^u)$ - i.e. v^u is a fixed point of the correspondence, $\frac{1-P_L}{P_L}\phi^u$. Since ϕ^u is upper-hemicontinuous, such a fixed point exists. Moreover, since ϕ^u is weakly decreasing, then $\frac{\phi^u(v^u)}{v^u}$ is strictly decreasing in v^u, and so the fixed point is unique for each $P_L \in [0,1]$. The following proposition provides a closed form expression for the expected equilibrium shares in the unicameral game:

Proposition 14. *The equilibrium share ratio can be expressed as follows:*

$$v^u = \begin{cases} \frac{1-P_L}{P_L}\phi_1^u & P_L \in \left(\frac{\phi_1^u}{1+\phi_1^u}, 1\right] \\ 1 & P_L \in \left[\frac{\phi_2^u}{1+\phi_2^u}, \frac{\phi_1^u}{1+\phi_1^u}\right] \\ \frac{1-P_L}{P_L}\phi_2^u & P_L \in \left[0, \frac{\phi_2^u}{1+\phi_2^u}\right) \end{cases}$$

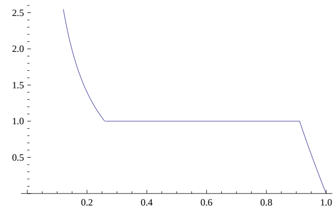
Two features of the equilibrium shares are worth noting. First, the expected payoff to a given district is continuous and weakly increasing in its recognition probability. Second, over a significant range of recognition probabilities, the pie is divided evenly amongst districts, in expectation, regardless of state size. To see why, note that the composition of the optimal coalition strongly favours the type that is cheaper. Hence, whilst an increase in the recognition probability of a given district puts an upward force on the equilibrium payoff to a given district (since it is more likely to capture the residual surplus), this effect is counterbalanced by that district being excluded from coalitions more frequently, when it is not the proposer.

As noted in the introduction, a significant motivation for bicameralism in federal states was the fear that a unicameral legislature would favour big state districts (in the sense that $v_B^u > v_S^u$). Given the above discussion, it should be clear that if legislators bargain optimally, such an equilibrium will only arise if the recognition probability strongly favours big states, or if legislators are highly impatient (i.e. δ is low).

The nature of the optimal coalition and the equilibrium shares can be seen in the following example:

Example 1. Consider the unicameral analogue of the example in the introduction where

Figure 4.1: Equilibrium share ratio under Unicameralism



there are 4 small state districts and 5 big state districts. The optimal coalitions are:

$$(\sigma_S^u, \beta_S^u) = \begin{cases} (3,1) & v^u < 1 \\ (0,4) & v^u > 1 \end{cases}$$

$$(\sigma_B^u, \beta_B^u) = \begin{cases} (4,0) & v^u < 1 \\ (0,4) & v^u > 1 \end{cases}$$

and the equilibrium share ratios are characterised by:

$$\hat{v} = \begin{cases} \frac{1-P_L}{P_L} \cdot \frac{5-\delta}{4(1-\delta)} & P_L \in \left(\frac{5-\delta}{9-5\delta}, 1\right] \\ 1 & P_L \in \left[\frac{5-4\delta}{9-4\delta}, \frac{5-\delta}{9-5\delta}\right] \\ \frac{1-P_L}{P_L} \cdot \frac{5-4\delta}{4} & P_L \in \left[0, \frac{5-4\delta}{9-4\delta}\right) \end{cases}$$

The following figure plots the equilibrium share ratio $\left(v^u = \frac{v_S^u}{v_B^u}\right)$ as the recognition probability of the big state, P_L varies, fixing $\delta = 0.9$:

The figure confirms that over a large range of recognition probabilities, the pie is equally divided between districts. Large states can only expropriate small states (i.e. $v^u < 1$) when their recognition probability is sufficiently high or the level of patience δ is sufficiently low.

4.4 Bicameralism

In this section, I characterise the equilibrium of the bicameral game, paying attention to the complementarities that exist between agents in the two chambers. As will become evident, the equilibrium characterisation shares many similarities with the unicameral case. The main difference lies in the composition of the optimal coalition - once proposals arising from the second chamber are introduced.

4.4.1 Optimal Coalition

The optimal coalition in the bicameral setting is built in the same way as in the unicameral case. To entice legislator i to support a proposal, the proposal must allocate at least δw_i amongst the districts that affect $i's$ utility. Allocating δv_S to a given small state buys the support of both the lower and upper house agents from that state, and allocating δv_B to a given big state district buys the support of the lower house agent from that district. To entice an upper house legislator from a big state into the coalition, the proposer must allocate $\delta k v_B$ amongst the k districts in that big state. It suffices to allocate δv_B to each of the k districts in the big state. (Note that there is no requirement that the allocation be equally divided amongst the k districts in that state. However, the equal distribution (weakly) dominates any other distribution, since it entices all k lower house legislators to join the coalition (in addition to the upper house agent) - whilst any other distribution will entice fewer lower house legislators to join. As such, I assume that the equal distribution is chosen in equilibrium.) Hence - as in the unicameral case - if the proposer wishes to entice a legislator into the coalition, he will allocate δv_i to each district that the legislator in question is concerned about.

In the unicameral case, as the residual claimant, the proposer always supported her own proposal. This remains true in the bicameral setting. Moreover, the complementarity in preferences implies that the legislators from the same state but opposite chamber to the proposer will also benefit from the residual claim. Lemma 17 formalises this insight:

Lemma 17. *In equilibrium, a proposal from a lower house agent will be accepted by that state's upper house legislator. Similarly, a proposal from an upper house agent will be accepted by each lower house agent from that state.*

Proof. The lemma is trivial for agents from small states, since they have identical preferences, and since the proposer always supports his own proposal. If the proposer is the upper house agent from a big state, then since he will support his own proposal, the big state must receive at least $\delta k v_B$ in total. But since the surplus is distributed equally amongst the k districts, each district receives at least δv_B - and so every lower house legislator from that big state will also support the proposal. If the proposer is a lower house agent from a big state, then the most he will allocate to districts outside of his state is $(b-1)\, k\delta v_B + s\delta v_S = \delta\,(1-kv_B)$, and so at least $1-\delta\,(1-kv_B) > \delta k v_B$ will be retained within the proposer's district. But this implies that the upper house legislator will support the proposal. $\square$

Lemma 17 shows that a proposer from a given chamber need not expend resources to get the support of the corresponding legislator(s) in the other chamber from his state. They join the coalition 'for free' - given the complementarity in preferences, the residual surplus that the proposer keeps for himself is sufficient to entice them into the coalition. This insight is important because it reduces the effective majority requirement in the lower house, when a B^U-type agent is the proposer. More specifically, whenever a non-B^U-type agent is the proposer, they must purchase the support of $M_L - 1$ lower house agents and $M_U - 1$ upper house agents. However, a B^U-type proposer needs only purchase the support of $M_L - k$ other lower house agents and $M_U - 1$ upper house agents. The key insight of this paper is this: Bicameralism allows big states to include $k-1$ fewer lower house agents in the coalition whenever the upper house agent is the proposer. If these $k-1$ agents would ordinarily have come from small states, then bicameralism results in small states being excluded more frequently from coalitions than would be true under unicameralism. This reduces the equilibrium payoff to small states and enhances the payoff to big state districts.

The set of optimal coalitions can be formulated as the solution to the following cost minimisation problem:

$$\begin{aligned}
(\sigma_t, \beta_t) &\in argmin_{(\sigma,\beta)\in\mathbb{Z}^2_+} v_S\sigma + v_B\beta \\
\text{s.t } \sigma + \beta &\geq M_L - 1 - (k-1)\mathbf{1}_{B^U}[t] \\
\sigma + \left\lfloor \frac{\beta}{k} \right\rfloor &\geq M_U - 1 \\
\beta &\leq bk - \mathbf{1}_{B^L}[t] - k\mathbf{1}_{B^U}[t] \\
\sigma &\leq s - \mathbf{1}_S[t]
\end{aligned}$$

where $\lfloor x \rfloor$ denotes the largest integer weakly less than x. I refer to the first two constraints as the lower and upper house majority constraints (respectively), whist the last two constraints are the legislator supply constraints. The indicator functions account for the fact that the effective majority requirements and the numbers of large and small districts available for purchase depends upon the proposer's type. As in the unicameral case, the optimal proposal rule μ_t may only put positive weight on minimisers to the above problem.

Let $l_t = 1 + (k-1)\,1_{B^U}[t]$, which is the number of lower house agents (including possibly the proposer herself) whose support the proposer gets for free. The following proposition characterises the composition of the optimal coalition in the bicameral setting:

Proposition 15. *Let* $\sigma'_t = \left\lfloor \frac{k(M_U-1)-(M_L-l_t)}{k-1} \right\rfloor$ *and* $\theta = mod_{k-1}(kM_U - M_L) + 1$. *Given the equilibrium expected share ratio* $v = \frac{v_S}{v_B}$, *the optimal coalition for a type-t proposer is given* *by:*

$$\sigma_t(v) = \begin{cases} \min\{\max\{M_U - 1, M_L - l_t\}, s - 1_S[t]\} & v < 1 \\ \min\{\max\{0, M_L - 1_S[t] - bk, M_U - 1_S[t] - b, \sigma_t' + 1\}, s\} & 1 < v < \theta \\ \min\{\max\{0, M_L - 1_S[t] - bk, M_U - 1_S[t] - b, \sigma_t'\}, s\} & \theta < v < k \\ \max\{0, M_L - 1_S[t] - bk, M_U - 1_S[t] - b\} & v > k \end{cases}$$

$$\beta_t(v) = \begin{cases} \max\{0, M_L - s - l_t 1_B[t], k(M_U - s - 1_B[t])\} & v < 1 \\ \max\{M_L - l_t - \sigma_t(v), k(M_U - 1 - \sigma_t(v))\} & v > 1 \end{cases}$$

Furthermore, unless $t = B^L$ *and* $\sigma_t(v) = M_U - b$, $\beta_t(v) = \min\{\max\{M_L - l_t, k(M_U - 1), bk - l_t 1_B[t]\}\}$ *whenever* $v > k$. If $v = \psi$, where $\psi \in \{1, \theta, k\}$, then any mixture of the optimal coalitions for $v = \psi - \frac{1}{2}$ and $v = \psi + \frac{1}{2}$ *is optimal.*

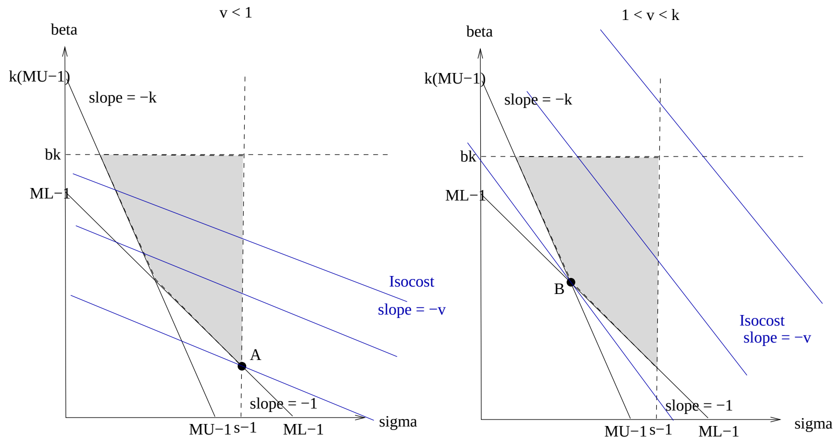
A proof - including a discussion about the variant expressions for β_t - can be found in the Appendix. The intuition for the above proposition is straightforward. If $v < 1$ (i.e. $v_S < v_B$), then in both chambers, small state districts are cheaper coalition partners than big state districts. Hence the optimal coalition is constructed by first purchasing small state legislators, and then 'topping-up' with big state districts, if necessary. The opposite is true if $v > k$.

When $1 < v < k$, big state legislators are cheaper in the lower house, but small state legislators are cheaper in the upper house. The construction of the optimal coalition uses the following insight: Suppose $M_L - l_t \geq k(M_U - 1)$, so that the number of big state districts needed to satisfy the lower house majority constraint is at least as large as the number of big state districts needed to satisfy the upper house constraint. I refer to this as the *lower house dominance property.* If so, then the optimal coalition will clearly include as many big state districts as possible - since this is the cheapest way to satisfy the lower house majority constraint, and doing so, satisfies the upper house constraint as well. However, if $M_L - l_t < k(M_U - 1)$, then this strategy is no longer optimal. To see why, note that after adding big state districts to the coalition to satisfy the lower house constraint, the proposer would still have to add some small state districts to satisfy the upper house constraint. But adding these small state districts creates a (large) excess majority in the lower house. Since the proposer must necessarily add small state districts to the coalition, he should anticipate this, and invite correspondingly fewer big state districts. In fact, the optimal coalition is built by adding small state districts one by one, until the dominance property holds for the residual majority requirements - and then adding as many big states as are needed or available. Indeed, $\max\{0, \sigma_t' + 1\}$ is precisely the number of small state districts that need to be added so that the dominance property holds on the residual majority requirement.

However there is a complication. If the above strategy is followed, then as big state districts are added to the coalition, the upper house majority constraint is satisfied before the lower house constraint. The last $k - \theta$ big state districts added to the coalition does not generate any complementary benefit in the upper house. The proposer may do better by adding a further θ big state districts (thereby getting the support of an additional upper house member) and reducing the number of small states in the coalition by 1 unit. Clearly, this alternative is desirable so long as $\theta v_B < v_s$ (i.e. $v > \theta$). Hence when $v \in (\theta, k)$, the optimal coalition contains 1 fewer small state and θ more big state districts than when $v \in (1, \theta)$.

The above complication in the equilibrium construction arises because of the need for solutions to take integer values. Nevertheless, the intuition behind Proposition 15 can be easily seen by considering a linearised version of the model, where non-integer solutions are permitted. The diagram below shows the optima of this cost minimisation problem for a type S proposer. The shaded area denotes the feasible set and is bounded from below by the majority constraints (diagonal lines) and from above by the legislator supply constraints. Since this is a linear program, the optima will generally be at the extreme points of the feasible set. If $v < 1$ (i.e. $v_S < v_B$), the lowest isocost line is achieved at point A and the

Figure 4.2: Optimal Coalition in the Linearized Model



optimal coalition exhausts all the small state districts. If $1 < v < k$, then the optimum is at point B. Note that σ and β will not both generally take integer values at this point. In the integer program, the optimal solution will be close to B, with σ being either rounded up or down depending on the relative cost of big and small districts. If v is relatively small ($v < \theta$), then σ is rounded up; it is rounded down otherwise. Hence, one can think of the complication discussed in the previous paragraph as simply being the result of an integer constraint.

The above diagram also illustrates the conditions under which excess majorities may arise in one chamber. An excess majority exists in a given chamber if a majority continues to exist after removing all the funds from any (arbitrarily chosen) state. (By this definition, an excess majority exists in the upper house if the coalition contains strictly more than M_U upper house agents. However, in the lower house, a coalition with slightly more than M_L members may not constitute an excess majority, if removing funding to a given state causes the coalition size to fall below M_L.) Suppose the majority requirements (M_L, M_U) are such that neither constraint strictly dominates the other. (I say that one majority constraint dominates the other, if the second constraint is always satisfied whenever the first is. Diagrammatically, this amounts to requiring the majority constraints to intersect within the interior of the rectangle enclosed by the axes and the legislator supply constraints. Point B in the above diagram satisfies this requirement.) If $v \in (1, k)$, then the optimal coalition is not in excess in either chamber. If $v < 1$, then there will generically be an excess majority in the upper house, and if $v > k$, there will be an excess majority in the lower house. To get an intuition for this result, we see that when $v \in (1, k)$, both the majority constraints are binding at the optimum, whilst when $v < 1$, only the lower house majority constraint binds. This result is similar in spirit to parts 4 and 5 of Proposition 1 in Kalandrakis (2004).

Two features of the optimal coalition are worth noting. First - as in the unicameral case - for any type of proposer, $\sigma_t(v)$ is (weakly) decreasing in v, whilst $\beta_t(v)$ is weakly increasing. This reflects the idea that proposers will optimally substitute away from particular coalition partners as they become more expensive. Let $\bar{\sigma}_t(v)$ and $\bar{\beta}_t(v)$ denote the expected number of small and big districts that are included in the coalition by a type-t proposer for a given equilibrium share ratio v. It is straightforward to show that $\bar{\sigma}_t$ and $\bar{\beta}_t$ are upper hemicontinuous in v. Moreover, the correspondences are singleton-valued and piecewise constant on the intervals $v \in [0, 1), v \in (1, \theta), v \in (\theta, k)$ and $v \in (k, \infty)$ and are convex valued when $v \in \{1, \theta, k\}$.

Second, $\sigma_{B^U} \leq \sigma_{B^L}$ - so that a B^U-type proposer will never build a coalition with more small states than a B^L-type proposer. This follows since the incentives for upper and lower house proposers from the big state differ only in that the upper house proposer does not need to purchase as many agents to satisfy the lower house constraint. Moreover, $\sigma_{B^U} < \sigma_{B^L}$ only when $v < 1$. To see why, again note that the advantage to the B^U-type proposer is that he doesn't need to purchase as many agents to satisfy the lower house majority constraint. When $v < 1$, both types of proposers will seek to build coalitions using small states. Since the lower house requirement is lower for B^U-type proposers, they will invite fewer small states into the coalition. Hence $\sigma_{B^U} < \sigma_{B^L}$. (Note - the previous discussion implicitly assumed that $M_L > M_U$. Of course, if $M_U > M_L$, then the upper house constraint is binding, and both types of proposers will choose identical coalitions.) I refer to this as the *reduced requirement* *effect.* However, if $v > 1$, both types of proposer would ideally satisfy the lower house constraint by adding big state districts to the coalition, and will only add small states to the coalition to satisfy the upper house constraint (ignoring legislator supply constraints for the moment). Since neither type has an advantage in the upper house, $\sigma_{B^L} = \sigma_{B^U}$ whenever $v \geq 1$.

4.4.2 Equilibrium shares and payoffs

In the above section I characterised the optimal coalition given a pair of conjectured equilibrium shares. I now calculate the expected shares that these (optimal) coalitions imply. Naturally, in equilibrium, the conjectured and implied expected shares must coincide. The expected share of the pie that is allocated to a big state district is characterised by:

$$v_B = \frac{P}{bk} - \frac{P}{bk}\left(\alpha_B \bar{\sigma}_{B^L} + (1-\alpha_B)\,\bar{\sigma}_{B^U}\right)\delta v_S + (1-P)\,\frac{\bar{\beta}_S}{bk}\delta v_B \tag{4.5}$$

This expression has an analogous interpretation to (4.1) in the unicameral case. The first term on the right hand side, $\frac{P}{bk}$, is the imputed probability that the proposer is concerned with the interests of a given big state district. The second term is the expected amount that a big district proposer must offer to small districts to form an optimal coalition (weighted by the conditional probabilities that such a proposer is from the lower and upper houses, respectively), whilst the third term is the expected amount the a big state district can expect to receive when a small district is the proposer. The equilibrium share for a big state district then, is its imputed recognition probability, adjusted for its expected payouts to, and receipts from, small state districts.

The expected share of the pie allocated to small state districts is similarly characterised and interpreted:

$$v_S = \frac{1-P}{s} - \frac{1-P}{s}\bar{\beta}_S \delta v_B + P\frac{\left(\alpha_B \bar{\sigma}_{B^L} + (1-\alpha_B)\,\bar{\sigma}_{B^U}\right)}{s}\delta v_S \tag{4.6}$$

Solving (4.5) and (4.6) gives:

$$\frac{v_S}{v_B} = \frac{1-P}{P} \frac{bk - \delta \bar{\beta}_S}{s - \delta \left(\alpha_B \bar{\sigma}_{B^L} + \left(1-\alpha_B\right) \bar{\sigma}_{B^U} \right)} \tag{4.7}$$

which is analogous to (4.3) in Section 4.3 above. Note again that the equilibrium shares depend only upon β_S, σ_{B^L} and σ_{B^U} - i.e. the number of districts of the *other* type that the proposer invites into the coalition.

An equilibrium is characterised by payoff shares (v_S, v_B) and a probability assignment over coalitions $\{\mu_t\}_{t\in T}$ such that μ_t only puts positive weight on coalition pairs (σ_t, β_t) satisfying Proposition 15, given (v_S, v_B); and the payoff shares satisfy (4.7), given $\{\mu_t\}_{t\in T}$.

Define

$$\phi\left(v\right) = \frac{bk - \delta \bar{\beta}_S \left(v\right)}{s - \delta \left(\alpha_B \bar{\sigma}_{B^L} \left(v\right) + \left(1-\alpha_B\right) \bar{\sigma}_{B^U} \left(v\right) \right)} \tag{4.8}$$

Since the optimal coalitions $\{(\sigma_t, \beta_t)\}_{t\in T}$ are piece-wise constant over the intervals $[0,1)$,$(1,\theta)\,,(\theta,k)$ and (k,∞), so is $\phi\left(v\right)$. (For convenience, denote $\phi\left(v\right) = \phi_1$ when $v<1$. Similarly let $\phi\left(v\right)$ be denoted by ϕ_2, ϕ_3 and ϕ_4 and over the remaining intervals, respectively.) $\phi\left(v\right)$ is upper-hemicontinuous in v since it is a continuous function of $\bar{\sigma}_{B^L}$, $\bar{\sigma}_{B^U}$ and $\bar{\beta}_S$ - all of which are upper-hemicontinuous. Moreover, ϕ is weakly decreasing in v, since $\bar{\sigma}_{B^L}$ and $\bar{\sigma}_{B^U}$ are weakly decreasing and $\bar{\beta}_S$ is weakly increasing in v. Finally, since ϕ is upper-hemicontinuous, then $\phi\left(1\right)$ is convex valued and takes values in the interval $[\phi_2, \phi_1]$. Similarly $\phi\left(\theta\right)$ and $\phi\left(k\right)$ are convex valued and take values in the intervals $[\phi_3, \phi_2]$ and $[\phi_4, \phi_3]$, respectively.

Given the above discussion, the equilibrium shares are characterised by $v \in \frac{1-P}{P}\phi\left(v\right)$ - i.e. v is a fixed point of the correspondence: $\frac{1-P}{P}\phi\left(v\right)$. Since $\phi\left(v\right)$ is upper-hemicontinuous and convex valued, such a fixed point exists. Furthermore, since $\frac{\phi(v)}{v}$ is strictly decreasing (in the sense that $v<v' \Rightarrow \min \frac{\phi(v)}{v} > \max \frac{\phi(v')}{v'}$), then the fixed point is unique for each $P \in [0,1]$. The following proposition characterises the equilibrium shares in the bicameral game, as a function of the recognition probabilities (P, α_B, α_S) - noting that $\phi_i = \phi_i\left(\alpha_B\right)$.

Proposition 16. *The equilibrium shares can be characterised (uniquely) as follows:*

$$v = \begin{cases} \frac{1-P}{P}\phi_1 & P \in \left(\frac{\phi_1}{1+\phi_1}, 1\right] \\ 1 & P \in \left[\frac{\phi_2}{1+\phi_2}, \frac{\phi_1}{1+\phi_1}\right] \\ \frac{1-P}{P}\phi_2 & P \in \left(\frac{\phi_2}{\theta+\phi_2}, \frac{\phi_2}{1+\phi_2}\right) \\ \theta & P \in \left[\frac{\phi_3}{\theta+\phi_3}, \frac{\phi_2}{\theta+\phi_2}\right] \\ \frac{1-P}{P}\phi_3 & P \in \left(\frac{\phi_3}{k+\phi_3}, \frac{\phi_3}{\theta+\phi_3}\right) \\ k & P \in \left[\frac{\phi_4}{k+\phi_4}, \frac{\phi_3}{k+\phi_3}\right] \\ \frac{1-P}{P}\phi_4 & P \in \left[0, \frac{\phi_4}{k+\phi_4}\right) \end{cases}$$

A proof is provided in the Appendix, however, the intuition is straightforward. Since $\phi(v)$ is decreasing in v, when $v < 1$, $\phi(v)$ is large. Since the equilibrium is a fixed point of the correspondence $\frac{1-P}{P}\phi$, then an equilibrium share ratio of v can only be sustained if $\frac{1-P}{P}$ is small enough - i.e. if P is sufficiently large. The proposition finds the values of P for which $\frac{1-P}{P}\phi_1 < 1$. The same logic applies for the other equilibrium shares. For $v = 1$, although any mixture of the optimal coalitions for $v = 1-\varepsilon$ and $v = 1+\varepsilon$ is optimal (from the perspective of the proposer), for a given P, there is a unique mixture that can sustained as a fixed point.

The equilibrium shares in the bicameral case share similar properties to the equilibrium in the unicameral case. The expected payoff to each district is continuous and weakly increasing in its recognition probability. Moreover, over certain ranges of recognition probabilities, the allocation is (locally) unresponsive to changes in the recognition probability. These 'focal' equilibrium shares include the cases where: (i) there is an equal allocation of resources amongst big and small state districts ($v = 1$), and (ii) there is an equal allocation of resources amongst big and small states ($v = k$) . The focal equilibria exist because at these foci, any pressure that would otherwise cause v to change (such as an increase in recognition probabilities) is counterbalanced by a change in the composition of the optimal coalition.

In particular, as v increases from $1-\varepsilon$ to $1+\varepsilon$, small states become more expensive than big state districts in the lower house, and are consequently excluded from coalitions more frequently. As v increases from $k-\varepsilon$ to $k+\varepsilon$, then small states also become more expensive in the upper house, and so are excluded from even more coalitions.

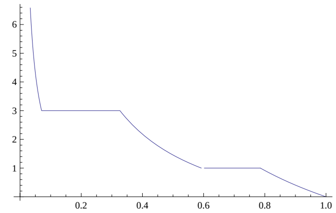
The nature of the optimal coalition and the equilibrium shares can be seen in the following example that was outlined in the introduction:

Example 2. Suppose there is one big state with $k=5$ and four small states, so that the big state has a bare majority in the lower house, but the small states control the upper house. The optimal coalitions are given by:

$$
\begin{aligned}
(\sigma_S, \beta_S) &= \begin{cases} (3,1) & v < 1 \\ (2,2) & 1 < v < 3 \\ (1,5) & v > 3 \end{cases} \\
(\sigma_{B^L}, \beta_{B^L}) &= \begin{cases} (4,0) & v < 1 \\ (2,2) & v > 1 \end{cases} \\
\left(\sigma_{B^U}, \beta_{B^U}\right) &= (2,0)
\end{aligned}
$$

It should be clear that $\sigma_B(v)$ is (weakly) decreasing in v, whilst $\beta_S(v)$ is increasing in v. Moreover $\sigma_{B^U} \leq \sigma_{B^L}$and this inequality is strict whenever $v < 1$.

Figure 4.3: Equilibrium share ratio under Bicameralism



The equilibrium share ratio are characterised by:

$$v = \begin{cases} \frac{1-P}{P} \cdot \frac{5-\delta}{4-2\delta(1+\alpha)} & P \in \left(\frac{5-\delta}{9-3\delta+2\delta\alpha}, 1\right] \\ 1 & P \in \left[\frac{5-2\delta}{9-4\delta}, \frac{5-\delta}{9-3\delta+2\delta\alpha}\right] \\ \frac{1-P}{P} \cdot \frac{5-2\delta}{4-2\delta} & P \in \left(\frac{5-2\delta}{17-8\delta}, \frac{5-2\delta}{9-4\delta}\right) \\ 3 & P \in \left[\frac{5(1-\delta)}{17-11\delta}, \frac{5-2\delta}{17-8\delta}\right] \\ \frac{1-P}{P} \cdot \frac{5(1-\delta)}{4-2} & P \in \left[0, \frac{5(1-\delta)}{17-11\delta}\right) \end{cases}$$

Suppose further that $\delta = 0.9$ and $\alpha_B = 0.6$ - i.e. big state proposals are more likely to originate in the lower house. Then the equilibrium share ratio is plotted as a function of the recognition probability P, in the figure below:

As before, the payoff to big state districts is weakly increasing in the recognition power of big states, and the allocation favours big state districts only when P is large enough. The figure confirms that over a range of recognition probabilities, the pie is equally divided between districts, and over a lower range of recognition probabilities, the pie is equally divided amongst states.

4.5 Comparison of Unicameralism and Bicameralism

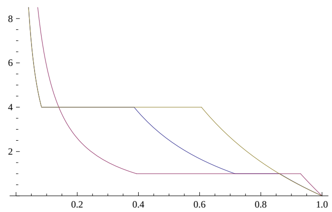
The previous sections characterised the equilibrium distribution of public funds under different institutional settings. In this section, I compare the outcomes across these institutions. In making this comparison, I implicitly assume that the lower house in a bicameral legislature has the same composition, recognition rule and decision rule as the unicameral legislature. Hence, M_L is the common majority requirement in both the lower house of the bicameral legislature and the unicameral legislature. Similarly, the recognition probabilities in the unicameral legislature are simply the recognition probabilities in the bicameral setting, conditional upon the proposer being in the lower house. These probabilities are calculated in Section 4.3.

To simplify the notation, I express all probabilities using the notation in the bicameral section. Hence the probability in the unicameral system of the proposer being from a big state (P_L) can be expressed as:

$$P_L = \frac{\alpha_B P}{\alpha_B P + \alpha_S (1 - P)} \tag{4.9}$$

At first glance this assumption appears to be quite restrictive - that the recognition probabilities in the bicameral legislature completely determine the recognition probabilities in the unicameral setting. One may plausibly assert that recognition rules evolve differently in different institutions. In fact, the above approach contains a degree of freedom, and so there is (almost) perfect freedom to move the recognition probabilities in the unicameral setting independently of the bicameral game. To see why, note that α_S has not featured in any part of the analysis, except in the determination of P_L. (Since the upper and lower house legislators from the same small state face identical cost minimisation problems when they are the proposer, the conditional probability that one is the proposer, rather than the other, becomes irrelevant.) Hence, the model allows for a wide range of possible recognition probabilities in the unicameral setting, simply by choosing α_S appropriately. Indeed, since $\alpha_S \in [0, 1]$, the model can accommodate any $P_L \in \left[\frac{\alpha_B P}{\alpha_B P + (1-P)}, 1\right]$. In particular, if $\alpha_S = \alpha_B$ (i.e. if being in the lower house confers no greater advantage in proposal power to the big state - relative to the small states - than being in the upper house), then $P_L = P$. Moreover - noting the irrelevance of α_S to any other part of the model - if the domain of α_S is expanded

Figure 4.4: Relative allocations under unicameralism and bicameralism under different majority requirements.



to $\alpha_S \in \Re_+$ (i.e. if it's interpretation as a conditional probability is abandoned), then any $P_L \in [0,1]$ can be sustained. Hence, since one can freely choose α_S, it is without loss of generality to use the same probability parameters in both institutional settings.
Define the following sets of majority requirement pairs:

$$A_1 = (M_L, M_U)|\, M_L < M_U < s + b\frac{s-(M_L-1)}{s-\delta(M_L-1)} \,\&\, M_L \leq s$$
$$A_2 = \left\{(M_L, M_U)\,|\, \frac{kM_U - M_L}{k-1} > s \,\&\, M_L > s, \text{ or } M_U \leq s \,\&\, M_U < M_L \leq s+k-1, \text{or } M_L \leq s \,\&\, M_U > s + b\frac{s-M_L}{s-\delta M_L}\right\}$$

These sets are illustrated in the diagram below. The sets A_1 and A_2 are depicted by the light and dark shaded areas (respectively).

Proposition 17 states the main result of this paper - which compares the allocation to small states under unicameralism and bicameralism. This depends upon the majority requirements and the relative recognition probabilities between the two chambers. The proposition shows that these factors may actually benefit big states.

Proposition 17. *Suppose the environment in the unicameral legislature favours the big state districts in equilibrium. (i.e. the parameters are such that $v^u < 1$). Then the distribution under a bicameral legislature favours the big district even more (i.e. $v < v^u$) if and only if $\alpha_S > \underline{\alpha}_S = \alpha_B \frac{\phi_1(\alpha_B)}{\phi_1^u}$. Furthermore $\underline{\alpha}_S < \alpha_B$ whenever $(M_L, M_U) \in A_2$ and $\underline{\alpha}_S > \alpha_B$ whenever $(M_L, M_U) \in A_1$.*

Proof. Let (P, α_B, α_S) be such that the equilibrium ratio satisfies $v^u < 1$. By (4.3) and (4.9), $v^u = \frac{\alpha_S}{\alpha_B}\frac{1-P}{P} \cdot \phi_1^u$. By similar argument, $v = \frac{1-P}{P}\phi(\alpha_B) \leq \frac{1-P}{P}\phi_1(\alpha_B)$, since ϕ is decreasing in v. Suppose $\alpha_S > \alpha_B \frac{\phi_1(\alpha_B)}{\phi_1^u}$. Then $v \leq \frac{1-P}{P}\phi_1(\alpha_B) < \frac{\alpha_S}{\alpha_B}\frac{1-P}{P}\phi_1^u = v^u < 1$. Hence $v < v^u$. Suppose instead that $v < v^u < 1$. Then by Propositions (16) and (14), $v = \frac{1-P}{P}\phi_1(\alpha_B) < \frac{\alpha_S}{\alpha_B}\frac{1-P}{P}\phi_1^u$, which implies that $\alpha_S > \alpha_B \frac{\phi_1(\alpha_B)}{\phi_1^u}$. This proves the first part of proposition. The proof of the second part appears in the Appendix. □

The proposition states that as long as small states are not recognised in the upper house with unduly high probability, then bicameralism (further) privileges big states whenever unicameralism does. This result is intuitive and follows as a straight-forward consequence of Propositions 14 and 16. It was shown in previous sections that the expected payoff to small states is increasing in their recognition probabilities. If the unicameral setting favours big state districts (in the sense that each big state district receives a larger share of the pie than each small state district), then it must be that the recognition probability of small states in the lower house is sufficiently low. By adding the second chamber, bicameralism may change the aggregate recognition probabilities of big and small states. The proposition states that bicameralism will continue to favour big state districts as long as the addition of the upper house does not skew the aggregate recognition probability too far in favour of small states.

Indeed, Proposition 17 provides a lower bound $\underline{\alpha}_S$ on the proposal power of small state legislators in the lower house (i.e. an upper bound on proposal power in the upper house) above which bicameralism privileges big states. If $\alpha_S > \underline{\alpha}_S$, then the upper house does not privilege small states enough (in terms of recognition power) to compensate for the bias in the unicameral legislature towards the big state districts. On the other hand, if $\alpha_S < \underline{\alpha}_S$, then small states are recognized in the upper house with large enough probability to swing the equilibrium allocation back in their favour.

The second part of the proposition characterises the lower bound $\underline{\alpha}_S$, and compares it to α_B - the likelihood that a big state proposer is recognised in the lower house. Recall, if $\alpha_B = \alpha_S$, then the likelihood that the recognized proposer is from a big state is independent across chambers - i.e. neither chamber confers any greater proposal advantage to big/small states than the other. If $\alpha_B > \alpha_S$, then legislators from big states are relatively more likely to be recognized in the lower house, whilst the upper house privileges small states in terms of proposal power relative to the lower house. If $\alpha_B < \alpha_S$, then the opposite is true. Hence, a comparison of $\underline{\alpha}_S$ and α_B determines the extent to which the upper house must privilege small state legislators as proposers before the equilibrium allocation under bicameralism favours small states. If $\underline{\alpha}_S < \alpha_B$, then the over-representation of small states in the upper chamber is not sufficient to increase their equilibrium share of the pie. Small states must also be recognized with a higher probability in the upper house, than the lower house. If $\underline{\alpha}_S > \alpha_B$, then the over-representation of small states is sufficient to increase their share, even if small states are less likely to be recognized in the upper house. Of course, the lower bound $\underline{\alpha}_S$ depends upon the majority requirements in both chambers and the composition of the optimal coalition.

Proposition 17 predicts that $\underline{\alpha}_S < \alpha_B$ whenever $(M_L, M_U) \in A_2$. This requires that either the upper house majority requirement is large enough, or that the majority requirement in the lower house is not too much larger than in the upper house. The first case (the upper section of A_2 in Figure 1) reflects the scenario where the upper house majority requirement is so large that, even after exhausting all the small states, big state districts must be drawn into the coalition. Moreover, to satisfy the upper house constraint, these districts must be added k at a time, and this provides a strong benefit to the big states.

The second case (the bottom section of A_2 in Figure 1) arises because of the *reduced requirement effect.* Since type-B^U proposers receive the support of k lower house agents for free, they do not need to purchase as many small state districts to satisfy the lower house majority constraint. As such, the residual surplus that accrues to the big state districts is larger, and the expected share accruing to small states is smaller, than would be the case under unicameralism. This effect will only arise if $\sigma_{B^U} < \sigma_{B^L}$ and this requires two conditions. First, the lower house majority requirement must be binding (i.e. $M_L > M_U$), or else the larger requirement in the upper house will pull more small states into the coalition than in the unicameral legislature. Second, the majority requirement in the lower house must not be so large as to exhaust the supply of small state legislatures. More precisely, $M_L < s + k$. If this were not true, then neither type of proposer can satisfy the lower house requirement by purchasing small state legislators alone, and so both types will exhaust the supply of small state legislators. But this implies $\sigma_{B^U} = \sigma_{B^L} = s$. Indeed, in this latter case, the payoff for large and small states is identical under unicameralism and bicameralism - and as such $\underline{\alpha}_S = \alpha_B$.

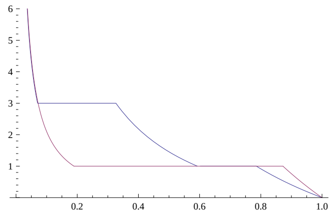
It is stressed that the two cases noted in the previous paragraphs arise for very different reasons. The first case arises when the majority requirement in the upper chamber is raised so high as to require the inclusion of many big state districts that had previously been excluded. This effect is reasonably intuitive and has been recognised in the literature by Kalandrakis (2004), amongst others. The second case is the novel feature that this paper seeks to highlight - that bicameralism may privilege big states, by reducing the number of small states that are needed to form a coalition, even when smalls states are the cheapest coalition partners.

For completeness, I consider the opposite case, when $\underline{\alpha}_S > \alpha_B$. In this case, the upper house may *disfavour* small state proposers, but still increase the equilibrium payoff to small states. Proposition 17 predicts that $\underline{\alpha}_S > \alpha_B$ whenever $(M_L, M_U) \in A_1$. This requires that two conditions are met. First, the lower house majority requirement must be small enough that it can be satisfied by a coalition of the small states alone. Second, the upper house majority requirement must be more demanding than the lower house majority requirement - although not too demanding. (Note - by 'more demanding', I mean that the total number of legislators - not just the proportion of legislators - required to satisfy the upper house majority requirement is larger than the number of legislators requires to satisfy the lower house majority requirement.) Under these conditions, bicameralism demands that more districts are included in the coalition (relative to unicameralism), and since the lower house constraint does not exhaust all the small state districts, more small states will be included in the coalition. This will tend to increase the expected payoff to small states. If the majority requirement in the upper house becomes too large, then after exhausting all the small state districts, big state districts must also be added to the coalition. If the number of big states added is large enough, this has the net effect of raising the expected share accruing to big state districts, relative to their share under unicameralism, where they are always excluded from the coalition when they are not the proposer. This explains why the majority requirement in the upper house cannot be too large.

Whilst the theoretical possibility may exist, it should be clear that the vast majority of bicameral legislatures are not characterised by majority requirements in the range of A_1. Given that the lower house typically contains many more members than the upper house, even if a large super-majority rule were employed in the upper house, it is still unlikely that the condition $M_U > M_L$ will obtain. (Two examples may serve to illustrate this point. In the United States, there are 435 legislators in the lower house and so even with a simple majority rule in the lower house, the 218 legislators required is larger than the 100 member Senate. Similarly, in Australia, a simple majority of the 150 member lower house is exactly equal to the size of the 76 member upper house - so even with a unanimity rule in the upper house, the majority requirement in that house cannot exceed the majority in the lower house.) Moreover, the requirement that $M_L < s$ is also unlikely to obtain. Even if the lower house adopts bills by simple majority - which is the most permissive majority rule available - this condition can only be met if the number of small states is larger than the total number of big state districts ($s > bk$). But in such a scenario, it is difficult to imagine there being any concern that big states will dominate a unicameral legislature.

Given the above discussion, and with reference to Figure 1, the following insights should be clear. First, increasing the majority requirement in the upper house tends to dampen the *reduced requirement effect*. Second, as the coalitions begin to exhaust the number of small states that are available, increasing the majority requirement in the lower house also tends to dampen the *reduced requirement effect*. Finally, to the extent that increasing the

Figure 4.5: Equilibrium share ratio - Comparison of Unicameralism and Bicameralism



number of big states in the polity tends to increase the majority requirement, adding more big states to the polity tends to dampen the *reduced requirement effect.* (Indeed, even if only a simple majority is required in the lower house, the *reduced requirement effect* disappears when $b > \frac{s-2}{k} + 2$, since this implies $M_L = \left\lfloor \frac{s+bk}{2} \right\rfloor > s + k - 1$).

It is worth noting that the *reduced requirement effect* is asymmetric. Bicameralism benefits big states when $v < 1$, however there is no corresponding benefit to small states when $v > 1$. The effect arises, not as a consequence of the equilibrium share ratio being larger or smaller than one, but because the complementarities in preferences reduces the majority requirement for type-B^U proposers. This causes B^U and B^L type proposers to choose different coalitions - and in particular, when $v < 1$, for type-B^U proposers to invite fewer small states into the coalition. Since their preferences are perfectly aligned, this wedge will never arise between the optimal coalitions for S^L and S^U type proposers - regardless of the value of v.

These results are demonstrated in the following example:

Example 3. Consider again the scenario outlined in Examples 2 and 1 above. Let s =4 and $k = 5$. Then $\phi_1(\alpha_B) = \frac{5-\delta}{4-2\delta(1+\alpha_B)}$ and $\hat{\phi}_1 = \frac{5-\delta}{4-4\delta}$. The figure below compares the equilibrium expected shares under unicameralism and bicameralism against the aggregate recognition probability of big states P , when $\delta = 0.9$, $\alpha_S = 0.4$ and $\alpha_B = 0.6$ (Note - these choice of parameters implies that big state legislators are more likely to be the proposer in the lower house and small state legislatures are more likely to be the proposer in the upper house).

By inspection of the diagram, $v^u > v$ whenever $v < 1$. More generally, if unicameralism favours the big state, then bicameralism is even more favourable, so long as:

$$\alpha_S > \alpha_B \frac{4 - 4\delta}{4 - 2\delta\left(1 + \alpha_B\right)}$$

If $\delta = 0.9$ and $\alpha_B = 0.6$, then bicameralism favours big states as long as $\alpha_S > 0.214$. Hence the benefit to big states exists even if the upper house confers a large proposal advantage on small states.

It is acknowledged that - in the above example, and as is evidenced in Figure 2 - the benefit to the big state only arises if the probability that the proposer is from a big state is sufficiently (and perhaps implausibly) large. However, this is less a criticism of the result presented in this paper, than a criticism of the general concern that small states will be relatively worse off under a unicameral legislature. As was noted in Section 4.3, the strong counterbalancing forces present in the construction of the optimal coalition tend to force the equilibrium shares towards equality. Small states can only be made worse off (relative to big state districts) if the recognition probability of big state districts is sufficiently (and perhaps implausibly) large. (Of course, in reality, other institutional features, such as the behaviour of political parties, may dampen the strength of this feature of coalition formation, such that small states lose out even when proposal rights are set a more 'reasonable' level.) The point that this paper seeks to make is that, if there are concerns about the welfare of small states under a unicameral setting, the solution may not be to introduce a second chamber, since this may possibly exacerbate the problem.

4.6 Conclusion

This paper examines the welfare implications of bicameral decision making over a distributive policy space. I present a standard legislative bargaining model of decision making in which the expected welfare of agents (in equilibrium) is determined by their recognition probabilities and the composition of the optimal minimum winning coalitions. As with the previous literature, the model predicts that the expected share of the pie accruing to small states is increasing in their proposal power. However, in contrast to the previous literature, I show that bicameralism can have the effect of skewing the composition of the optimal coalition in a way that favours big state districts - even if small states are *over-represented* in the upper house. The result is driven by the complementarity in preferences between upper and lower house legislators from big states - which arises since the upper house legislator from a big state will be responsive to the welfare of citizens spanning many lower house districts. This complementarity ensures that an upper house legislator from the big state receives the support of more lower house legislators for free, and so has to purchase fewer coalition partners than a lower house legislator from the same big state. Hence, bicameralism serves a 'coordinating' role amongst the various legislators from the same state, that allows legislators to channel more resources to their state when the proposer is amongst their number.

The main contribution of this paper is to show that the mere fact of requiring a concurrent majority in a second chamber that is disproportionately populated by small states is not sufficient to guarantee an increase in welfare for small states. Bicameralism may indeed improve the welfare of small states, but it does so by increasing the likelihood that small states can propose legislation, rather than biasing the composition of members in the second chamber.

4.7 Appendix

Proof of Proposition 15 . The legislator supply constraints imply that:

$$\max\left\{0, M_L - bk - \mathbf{1}_S[t], M_U - b - \mathbf{1}_S[t]\right\} \leq \sigma_t \leq \min\left\{s - \mathbf{1}_S[t]\right\}$$

$$\max\left\{0, M_L - s - l_t\mathbf{1}_B[t], k\left(M_U - s - \mathbf{1}_B[t]\right)\right\} \leq \beta_t \leq \min\left\{bk - l_t\mathbf{1}_B[t]\right\}$$

Consider the first expression. The lower bound is the minimum number of small states (in addition to the proposer, if she is from a small state) who must be enticed into the coalition to satisfy the majority constraints in both houses, given the supply of big state legislators. The upper bound is the maximum number of small states who can be enticed into the coalition (in addition to the proposer, if she is from a small state), given the supply of small state legislators. The second expression is analogous, for big state districts. I refer to these inequalities as the 'feasibility constraints'.

Define $\sigma_t' = \max\left\{\sigma \in \mathbb{Z} | \sigma < \frac{k(M_U-1)-(M_L-l_t)}{k-1}\right\}$ and for a given $\sigma \in \mathbb{Z}$, let $\beta_t(\sigma) = \max\{k(M_U - 1 - \sigma), M_L - l_t - \sigma\}$. For a given σ, β_t is the minimum number of big state districts that must be included in the coalition to ensure that the majority constraints are satisfied in both chambers (ignoring legislator supply constraints). (Note - at this stage I do not insist that σ_t' be feasible. Indeed - σ_t' may be negative.) Clearly:

$$\beta_t(\sigma) = \begin{cases} k(M_U - 1 - \sigma) & \sigma \leq \sigma_t' \\ M_L - l_t - \sigma & \sigma > \sigma_t' \end{cases}$$

Let $C_t(\sigma) = v_S\sigma + v_B\beta_t(\sigma)$ be the cost for a type-t proposer to build a (σ, β_t)-coalition. Let $\Delta C_t(\sigma) = C_t(\sigma+1) - C_t(\sigma)$. If $\Delta C_t(\sigma) < 0$, then there is strict incentive for the proposer to increase the number of small states in the coalition, and vice versa. From the above expression for $\beta_t(\sigma)$, it is easily verified that $\Delta C_t(\sigma)$ satisfies:

$$\Delta C_t(\sigma) = \begin{cases} v_S - kv_B & \sigma < \sigma_t' \\ v_S - [k(M_U - 1 - \sigma_t') - (M_L - l_t - \sigma_t')]\, v_B & \sigma = \sigma_t' \\ v_S - v_B & \sigma > \sigma_t' \end{cases}$$

Let $\psi_t = k\left(M_U - 1 - \sigma'_t\right) - \left(M_L - l_t - \sigma'_t\right)$. It can easily be shown that:

$$\begin{aligned}\psi_t &= mod_{k-1}\left[k\left(M_U - 1\right) - \left(M_L - l_t\right)\right] + 1 \\ &= mod_{k-1}\left[kM_U - M_L\right] + 1\end{aligned}$$

(To see this, note that $(k-1)\,\sigma'_t < k\left(M_U - 1\right) - \left(M_L - l_t\right) \leq (k-1)\left(\sigma'_t + 1\right)$, which implies that $0 < \psi_t \leq k - 1$. The second equality follows from the fact that $k - l_t \in \{0, k-1\}$.) Note that $0 < \psi < k$ implies that $\Delta^2 C_t(\sigma) \geq 0$, and so the marginal cost of adding one more small state to the coalition is weakly increasing.

Suppose $v < 1$ (i.e. $v_S < v_B$). Then $\Delta C_t(\sigma) < 0\ \forall\sigma$, and so the proposer can decrease the cost of a coalition by adding another small state, whenever it is feasible to do so. Adding the feasibility constraints, the optimal coalition will contain $\sigma_t = \min\left\{\max\left\{M_L - l_t, M_U - 1\right\}, s - \mathbf{1}_S[t]\right\}$ and $\beta_t = \max\left\{0, M_L - s - l_t\mathbf{1}_B[t], k\left(M_U - s - \mathbf{1}_B[t]\right)\right\}$.

Suppose $v > k$ (i.e. $v_S > v_B$). Then $\Delta C_t(\sigma) > 0\,\forall\sigma$, and so the proposer can decreasing the cost of a coalition by reducing the number of small states, whenever it is feasible to do so. Again adding the feasibility constraints, the optimal coalition will contain $\sigma_t = \max\left\{0, M_L - bk - \mathbf{1}_S[t], M_U - b - \mathbf{1}_S[t]\right\}$ and $\beta_t = \max\left\{M_L - l_t - \sigma_t, k\left(M_U - 1 - \sigma_t\right)\right\}$. Unless $0 \ngeq M_U - b > M_L - bk$ and $t = B^L$, this can be expressed more simply as $\beta^*_t = \min\left\{\max\left\{M_L - l_t, k\left(M_U - 1\right)\right\}, bk - l_t\mathbf{1}_B[t]\right\}$. (This is easily verified. If $\sigma_t = 0$, then there are sufficiently many big state districts to satisfy both majority constraints. Hence $\beta_t = \max\left\{M_L - l_t, k\left(M_U - 1\right)\right\} = \beta^*_t$. If $\sigma_t > 0$, then there are insufficiently many big state districts to satisfy both constraints. Hence $\beta^* = bk - l_t\mathbf{1}_B[t]$. Now, if the lower house majority constraint is binding, then the optimal coalition will use every available big state districts, and so $\beta_t = bk - l_t\mathbf{1}_B[t] = \beta^*_t$. However, if the upper house majority constraint (only) is binding, then the optimal coalition will purchase every available big state senator, and so $\beta_t = \left(b - \mathbf{1}_B[t]\right)k$. If $t \in \left\{S, B^U\right\}$, then $\beta_t = \beta^*_t$. By contrast, if $t = B^L$, then

$\beta_t = (b-1)k < bk - 1 = \beta_t^*$. When a lower house legislator from a big state is the proposer, he automatically gets the support of the upper house agent from his state - so the coalition needn't include the $k-1$ other districts in his state. This problem does not arise when $t = S$ (obviously) or when $t = B^U$,since in the latter case, all k districts from the proposer's state are automatically in the coalition.)

Suppose $1 < v < \psi$ (i.e. $v_B < v_S < \psi v_B$). Then $\Delta C(\sigma) > 0$ if $\sigma > \sigma_t'$ and $\Delta C(\sigma) < 0$ if $\sigma \leq \sigma_t$'. Hence, in the unconstrained problem, the optimal coalition contains $\sigma_t' + 1$ small states (since, when $\sigma = \sigma_t$, it is still profitable to add one more small state to the coalition). Adding the feasibility conditions gives $\sigma_t = \min\{\max\{0, M_L - bk - \mathbf{1}_S[t], M_u - b - \mathbf{1}_S[t], \sigma_t' + 1\}, s - \mathbf{1}_S[t]\}$ and $\beta_t = \beta_t(\sigma_t' + 1)$. Suppose $\psi < v < k$ (i.e. $\psi v_B < v_S < k v_B$). Then $\Delta C(\sigma) > 0$ if $\sigma \geq \sigma_t$ and $\Delta C(\sigma) < 0$ if $\sigma < \sigma_t$. Hence, in the unconstrained problem, the optimal coalition contains σ_t' small states. Adding the feasibility conditions gives $\sigma_t = \min\{\max\{0, M_L - bk - \mathbf{1}_S[t], M_u - b - \mathbf{1}_S[t], \sigma_t'\}, s - \mathbf{1}_S[t]\}$ and $\beta_t = \beta_t(\sigma_t')$.

Finally, if $v \in \{1, \psi, k\}$ then the optimal coalition is found by taking arbitrary mixtures of the adjacent coalitions. This follows since the correspondence $\left(\overline{\sigma}_t(v), \overline{\beta}_t(v)\right)$ is upper hemicontinuous. (To see this, let $\{v_n\} \to v$ and let $\{(\sigma_n, \beta_n)\} \to (\sigma, \beta)$ be a sequence s.t. $(\sigma_n, \beta_n)$is optimal for v_n. Suppose (σ, β) is not optimal for v. Then, $\exists(\sigma', \beta')$ feasible s.t. $v_S\sigma' + v_B\beta' < v_B\sigma + v_B\beta - 3\varepsilon$. But for n large enough, $v_S^n\sigma_n + v_B^n\beta_n > v_S^n\sigma + v_B^n\beta - \varepsilon > v_S\sigma + v_B\beta$-$2\varepsilon$. Moreover, for n large, $v_S\sigma' + v_B\beta'$> $v_S^n\sigma' + v_B^n\beta'$-$\varepsilon$. Then $v_S^n\sigma' + v_B^n\beta' < v_S^n\sigma_n + v_B^n\beta_n$, which contradicts the assumption that $(\sigma_n, \beta_n)$is optimal for v_n.) Take $v = 1$. Let (σ_1, β_1) be the optimal coalition whenever $v < 1$ and (σ_2, β_2) be the optimal coalition whenever $1 < v < \psi$. Then, by upper-hemicontinuity $\lim_{v \to 1^-}(\sigma_t(v), \beta_t(v)) = (\sigma_1, \beta_1) \in (\sigma_t(1), \beta_t(1))$. Similarly, $\lim_{v \to 1^+}(\sigma_t(v), \beta_t(v)) = (\sigma_2, \beta_2) \in (\sigma_t(1), \beta_t(1))$. Hence (σ_1, β_1) and (σ_2, β_2) are both optimal coalitions when $v = 1$, and so any mixture of these is also optimal. A similar argument holds for $v = \psi$ and $v = k$. □

Lemma 18. *Let $(\sigma, \beta) \neq (\sigma', \beta')$ both be optimal coalitions for $v = v_0$ and let μ be the* *probability that a (σ,β)-coalition is chosen. Let* $\phi(v_0,\mu)=\frac{bk-\delta\overline{\beta}}{s-\delta\overline{\sigma}}$*, where* $\overline{\sigma}=\mu\sigma+(1-\mu)\sigma'$ *and* $\overline{\beta}=\mu\beta+(1-\mu)\beta'$*, and let* $\phi=\phi(v_0,1)$ *and* $\phi'=\phi(v_0,0)$*. Then, for every* $\lambda\in[0,1]$*, there is a unique* $\mu\in[0,1]$ *s.t.* $\phi(v_0,\mu)=\lambda\phi+(1-\lambda)\phi'$*.*

Proof. Clearly, for $\lambda=1$ (respectively $\lambda=0$), $\mu=1$ (respectively $\mu=0$) satisfies the claim. Suppose $\lambda\in(0,1)$. $\phi(v_0,\mu)$ is continuous in μ, since it is the ratio of two non-zero continuous functions. Moreover, since $(\sigma,\beta)\neq(\sigma'\beta')$ and σ is decreasing whilst β is increasing, then $\phi\neq\phi'$. WLOG suppose $\phi>\phi'$. Let $\phi_\lambda=\lambda\phi+(1-\lambda)\phi'$ and note that $\phi>\phi_\lambda>\phi'$ and ϕ_λ is strictly increasing in λ. Then, by the intermediate value theorem, there is some $\mu(\lambda)\in(0,1)$s.t. $\phi(v_0,\mu)=\phi_\lambda$, for each $\lambda\in(0,1)$. Moreover, since ϕ_λ is strictly increasing, $\mu(\lambda)$ is unique. □

Proof of Proposition 16. Let $\Phi(P,v)=\frac{1-P}{P}\phi(v)$. It suffices to show that the conjectured v is a fixed point of $\Phi(P,v)$. The proof proceeds piecewise. First consider $v\in\Re\setminus\{1,\theta,k\}$, so that $\phi(v)$ is a singleton. To fix ideas, consider $v<1$. Then $\Phi(P,v)=\frac{1-P}{P}\phi_1$. Hence $v=\frac{1-P}{P}\phi_1$ is a fixed point of Φ as long as $\frac{1-P}{P}\phi_1<1$. But this implies that $P<\frac{\phi_1}{1+\phi_1}$. A similar argument is used for the remaining cases: $1<v<\theta$, $\theta<v<k$ and $v>k$.

Now, suppose $v\in\{1,\theta,k\}$, so that $\phi(v)$ is a convex correspondence. Again to fix ideas, consider $v=1$. It suffices to find some $\mu\in[0,1]$ s.t. $\frac{1-P}{P}\phi(1,\mu)=1$. This requires that $P=\frac{\phi(1,\mu)}{1+\phi(1,\mu)}$. By Lemma 18 there is unique $\mu\in[0,1]$ for each $\phi_\lambda\in[\phi_2,\phi_1]$ s.t. $\phi(1,\mu)=\phi_\lambda$. Since $\phi_2\leq\phi_\lambda\leq\phi_1$, then $\frac{\phi_2}{1+\phi_2}\leq\frac{\phi_\lambda}{1+\phi_\lambda}\leq\frac{\phi_1}{1+\phi_1}$. Hence, a fixed point of Φ exists whenever $P\in\left[\frac{\phi_2}{1+\phi_2},\frac{\phi_1}{1+\phi_2}\right]$. Moreover, since $\phi_\lambda$is strictly increasing in μ, this fixed point is unique. The proof for $v=\theta$ and $v=k$ is analogous.

Finally, I show that the expected equilibrium shares are unique. Suppose not. Then for some probability triple (P,α_B,α_S), there exist v,v' with $v\neq v'$ such that both are fixed points of Φ. WLOG suppose $v<v'$. Since $\sigma(v)$ is decreasing in v and $\beta(v)$ is increasing in

v, $\min \phi(v) \geq \max \phi(v')$. Then $\phi(v, \mu) \geq \phi(v', \mu')$ and so:

$$v = \frac{1-P}{P}\phi(v, \mu) \geq \frac{1-P}{P}\phi(v', \mu') = v'$$

which contradicts $v' > v$. $\square$

Proof of Proposition 17. I am left to prove the claims about value of $\underline{\alpha}_S$ relative to α_B. Let $\Delta = bk(\sigma_B - \hat{\sigma}_B) - s\left(\beta_S - \hat{\beta}_S\right) + \delta\left(\beta_S\hat{\sigma}_B - \sigma_B\hat{\beta}_S\right)$, where $\sigma_B = \alpha_B\sigma_{B^L} + (1-\alpha_B)\sigma_{B^U}$. Note that this expression simplifies to $\Delta = (bk - \delta\beta_S)(\sigma_B - \hat{\sigma}_B)$ if $\beta_S = \hat{\beta}_S$ and $\Delta = -(s - \delta\sigma_B)\left(\beta_S - \hat{\beta}_S\right)$ if $\sigma_B = \hat{\sigma}_B$. It is easily verified that $\Delta > 0$ iff $\phi_1(\alpha_B) > \hat{\phi}_1$ (which implies that $\frac{\phi_1(\alpha_B)}{\hat{\phi}_1} > 1$). To show that $\underline{\alpha}_S \lesseqgtr \alpha_B$ it suffices to show that $\Delta \lesseqgtr 0$.

Suppose $(M_L, M_U) \in A_1$. Then $M_L < M_U$ and $M_L \leq s$. If $M_U \leq s$, then $\sigma_B = M_U - 1 > M_L - 1 = \hat{\sigma}_B$ and $\beta_S = \hat{\beta}_S = 0$, and so $\Delta = bk(M_U - M_L) > 0$. If instead, $s < M_U < s + b\frac{s-(M_L-1)}{s-\delta(M_L-1)}$, then $\sigma_B = s > M_L - 1 = \hat{\sigma}_B$, and $\beta_S = k(M_U - s) > 0 = \hat{\beta}_S$. Then

$$\begin{aligned}
\Delta &= bk(s - (M_L - 1)) - sk(M_U - s) + \delta k(M_U - s)(M_L - 1) \\
&= k[b(s - (M_L - 1)) - (M_U - s)(s - \delta(M_L - 1))] \\
&> 0
\end{aligned}$$

where the last inequality is implied by $M_U < s + b\frac{s-(M_L-1)}{s-\delta(M_L-1)}$.

Suppose $(M_L, M_U) \in A_2$. There are three possibilities. (1) If $\frac{kM_U - M_L}{k-1} > s$ and $M_L > s$ (which implies that $M_U > s$), then $\sigma_B = \hat{\sigma}_B = s$ and $\beta_S = k(M_U - s) > M_L - s = \hat{\beta}_S$. Hence $\Delta = -(s - \delta\sigma_B)\left(\beta_S - \hat{\beta}_S\right) < 0$. (2) Suppose $M_U \leq s$ and $M_U < M_L \leq s + k - 1$. Since $M_U \leq s$, then $\beta_S = \hat{\beta}_S$, and since $M_U < M_L$, then $\sigma_{B^L} = \hat{\sigma}_B$. Hence $\Delta = (bk - \delta\beta_S)(1-\alpha)(\sigma_{B^U} - \sigma_B)$ and $\Delta < 0$ if $\sigma_{B^U} < \hat{\sigma}_B$. Since $M_L \leq s+k-1$ (i.e. $M_L - k < s$), then $\sigma_{B^U} \in \{M_L - k, M_U - 1\}$. Furthermore, by the assumptions on M_L and M_U, $M_L - k < \min\{M_L - 1, s\}$and $M_U - 1 < \min\{M_L - 1, s\}$. Hence $\sigma_{B^U} < \min\{M_L - 1, s\} = \hat{\sigma}_B$, and so $\Delta < 0$. (3) Suppose $M_L \leq s$ and $M_U > s + b\frac{s-(M_L-1)}{s-\delta(M_L-1)}$. Then $\beta_S = k(M_U - s) > 0 = \hat{\beta}_S$

and $\sigma_B = s > M_L - 1 = \hat{\sigma}_B$. Hence

$$\begin{aligned}\Delta &= bk\left(s-\left(M_L-1\right)\right)-sk\left(M_U-s\right)+\delta k\left(M_U-s\right)\left(M_L-1\right)\\ &= k\left[b\left(s-\left(M_L-1\right)\right)-\left(M_U-s\right)\left(s-\delta\left(M_L-1\right)\right)\right]\\ &< 0\end{aligned}$$

where the last inequality is implied by $M_U > s + b\frac{s-(M_L-1)}{s-\delta(M_L-1)}$. Hence $\Delta < 0$ whenever $(M_L, M_U) \in A_2$. □

4.8 References

Agency, Central Intelligence, "The CIA World Factbook," January 2011.

Ansolabehere, Stephen, Jr. Snyder James M., and Michael M. Ting, "Bargaining in Bicameral Legislatures: When and Why Does Malapportionment Matter?," *The American Political Science Review*, 2003, *97* (3), pp. 471–481.

Banks, J. and J. Duggan, "A general bargaining model of legislative policy-making," *Quarterly Journal of Political Science*, 2006, *1* (1), 49–85.

Banks, Jeffrey S. and John Duggan, "A Bargaining Model of Collective Choice," *The American Political Science Review*, 2000, *94* (1), pp. 73–88.

Baron, David P. and John A. Ferejohn, "Bargaining in Legislatures," *The American Political Science Review*, 1989, *83* (4), pp. 1181–1206.

Cutrone, M. and N. McCarty, "Does bicameralism matter?," *Oxford Handbook of Political Economy. Oxford: OUP*, 2006, pp. 180–195.

Diermeier, Daniel and Roger B. Myerson, "Bicameralism and Its Consequences for the Internal Organization of Legislatures," *The American Economic Review*, 1999, *89* (5), pp. 1182–1196.

Eraslan, H., "Uniqueness of stationary equilibrium payoffs in the Baron-Ferejohn model," *Journal of Economic Theory*, 2002, *103* (1), 11–30.

Hamilton, Alexander, James Madison, and John Jay, *The federalist papers*, Vol. 558, 1961.

Hammond, Thomas H. and Gary J. Miller, "The Core of the Constitution," *The American Political Science Review*, 1987, *81* (4), pp. 1155–1174.

Kalandrakis, T., "Proposal rights and political power," *American Journal of Political Science*, 2006, *50* (2), 441–448.

Kalandrakis, Tasos, "Bicameral Winning Coalitions and Equilibrium Federal Legislatures," *Legislative Studies Quarterly*, 2004, *29* (1), pp. 49–79.

Lijphart, A., *Patterns of democracy*, Yale University Press, 1999.

McCarty, N., "Proposal rights, veto rights, and political bargaining," *American Journal of Political Science*, 2000, *44* (3), 506–522.

Tsebelis, G. and J. Money, *Bicameralism*, Cambridge Univ Pr, 1997.