

Searching for Good Policies*

Steven Callander[†]

September 10, 2008

Abstract

I study a model of dynamic policy making in which citizens do not have complete knowledge of how policies are mapped into outcomes. They learn about the mapping through repeated elections as policies are implemented and outcomes observed. I characterize for this environment the policy trajectory with impatient voters. I find that through experimentation good policies are frequently found. However, I show that this is not always the case and demonstrate how policy making can get stuck at unappealing outcomes, revealing a novel informational failure of policy making. The model also provides insight into the size, direction, and sequencing of optimal policy experiments. Finally, I consider how the structure of political competition affects experimentation and learning.

*I thank Martin Osborne, Abhinay Muthoo, Alan Wiseman, and several seminar audiences for helpful comments, and Guillermo Diaz for research assistance.

[†]Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208; scal@kellogg.northwestern.edu.

1 Introduction

In October 1979 the United States embarked on a policy making experiment. For three years the Federal Reserve targeted growth in the money supply in its attempts to influence economic conditions, ushering in the era of monetarism. Although the experiment produced mixed results – and was abandoned in October 1982 – it revealed valuable information about how the economy works, information that was used to shape the replacement policy. In fact, Benjamin Friedman (1984) argues that the experiment was such a valuable learning experience precisely because it was so radical.

The monetarism experiment is not unique within policy making. Throughout history experimentation and learning have been central to policy choice. From laissez-faire to the New Deal, from tradeable pollution permits to school vouchers, the search for good policies has been guided by trial-and-error.

The objective of this paper is to study the challenges posed by policy making in a dynamic and uncertain world. In particular, the paper seeks to provide answers to such questions as: When do policy makers experiment with policy and when do they settle for what is known? When do they make radical changes to policy and when are changes incremental? Where in the policy space do policy makers search for good policies? How much is learned from the policy experience and how does this affect the trajectory of policy through time? And, in particular, does the search process identify good policies?

To answer these questions I develop a model of repeated two-candidate elections in a single dimensional policy space. The key ingredient of the model is that citizens – voters and candidates – have imperfect information about how policies are transformed into outcomes. Thus, finding the policy that delivers the desired outcome is not straightforward. Aiding the policy making process is the ability to learn from experience. In each period the winner of the election implements his campaign promise and the outcome is observed (and experienced). If the policy choice is *experimental* (not previously tried) its outcome reveals whether the policy is itself good and also provides information about the likely outcomes of other policy alternatives. Citizens update their beliefs accordingly and use the information to *predict* the outcomes of other policies, guiding their future choices. I characterize for this environment with impatient voters the optimal choice of policy in each period and describe the policy trajectory.

A novelty of the model is the specification of the policy process. I suppose that policies are mapped into outcomes by the realized path of a Brownian motion, where citizens know the parameters of the motion (the drift and variance) but not the path. Although used in a non-standard manner – with policies acting as the independent variable – the Brownian motion captures many realistic properties of policy making and does so in a highly tractable form.¹ This structure endows citizens with the ability to

¹The usefulness of the Brownian motion in modeling the policy process suggests other stochastic processes may also be applicable, an issue I take up in the discussion section.

order policies along the standard liberal-conservative continuum according to *expected* outcomes but not according to *realized* outcomes. Thus, citizens know which policies are *more likely* to deliver liberal (or conservative) outcomes but do not know which policies *do* deliver liberal outcomes. Moreover, the non-monotonicity of a Brownian path captures a key difficulty – and risk – of policy making: that outcomes may move in the opposite direction to that intended from a change in policy. This possibility formalizes Merton’s (1936) famous notion of “unanticipated consequences of purposive social action,” known popularly as the Law of Unintended Consequences. The Brownian motion also possesses attractive learning properties that I elaborate on in more detail momentarily.

The equilibrium trajectory of policy choices in this environment is path-dependent, varying in the outcomes realized as citizens progressively learn about the policy process. Yet basic patterns emerge, providing insight into the motivating questions listed above. I find that the search process is often effective at finding good policies, with experimentation continuing until an outcome close to the median voter’s ideal is obtained. More interestingly, however, I find that citizens often choose to stop experimenting *before* a good policy is found and I identify the possibility for policy making to get *stuck* at less appealing outcomes.

In fact, I prove that policy making can get stuck at policies that deliver *any* outcome, including outcomes arbitrarily distant from those preferred by voters. Policy making that is stuck represents a novel informational failure of policy making, showing how informational traps can emerge and lock-in in unappealing outcomes. Getting stuck requires at least three policy alternatives (and thus it is obscured in a binary policy model) and is induced by particular sequences of observed outcomes and the information they reveal. Notably, this logic differs from existing explanations of political failure, arising even in the absence of special interests, asymmetric information, or agency problems.

I derive analytically the conditions necessary for an outcome to be deemed good enough or to get stuck, providing a precise answer to the question of when citizens experiment and when they settle for what is known. Of particular interest is that the boundary for what is considered a good-enough outcome is not constant through time. Rather, it progressively tightens as more is learned about the policy process, reflecting the non-stationarity and history-dependence of policy making. Consequently, an outcome that earlier may have been deemed good enough can later be discarded and experimentation continued.

The model also provides insight into the question of how citizens experiment when they choose to do so. The richness of the Brownian structure pays dividends at this point as I am able to draw conclusions about the optimal size and direction of policy changes. Policy experimentation proceeds through two distinct phases – a *monotonic phase* and then a *triangulating phase* – that are followed almost surely by a *stable phase*, in which the same policy is retained thereafter.

In the monotonic phase the policy choice moves in a constant (monotonic) direction across the policy space as policy makers search for a region of the space where a good policy lies. In contrast, the direction of experimenting in the triangulating phase varies from period to period, reversing course repeatedly. At such a mature point in the issue’s life-cycle, the candidates oscillate between previously tried policies supported by either side of the political divide, attempting (*a la* Bill Clinton) to triangulate between the two sides in the hope of finding an appealing middle ground.

The distinction between experimental phases also generates a pattern in the size of policy changes. Early in an issue’s life-cycle – in the monotonic phase – optimal policy experiments are large, often bold, and unbounded. In contrast, later in the triangulating phase change is more incremental and bounded by previous choices. This pattern speaks to the long-running debate on the optimal size of policy (or organizational) change. The most famous theory on this question is *incrementalism* (Lindblom 1959) that argues change should be made only through incremental steps (see Bendor 1995 for an overview and critique). The results found here demonstrate the limitations of an incremental strategy and show how the optimal size of policy experiments varies across phases.²

A key feature of the model is that voters are impatient, choosing policy each period to maximize their immediate payoff (discounting the future entirely).³ In modeling elections this assumption is not entirely inappropriate. It is consistent with the myopia and inattention mass electorates are well-known for (see Bartels (2008) for an argument along these lines). Moreover, it provides a reasonable approximation to a rational world in which policies take considerable time to implement and produce results. That said, intertemporal incentives are surely relevant to policy making in some contexts. Although the results derived here are not knife-edged,⁴ extending the analysis to the case of patient voters is nevertheless of obvious interest (but also considerable difficulty).⁵

After presenting the main results, I turn to the issue of political design. A question raised by behavior in the baseline model is how the structure of political competition affects experimentation and policy outcomes. I explore this question by extending the model of political competition in two directions: including candidate uncertainty over voter preferences and allowing voters to abstain. Normatively, the extensions provide a

²The pattern found is also in contrast to other popular views, such as that bold change is always necessary or that small changes are needed early to “feel out” an issue before bold change can be adopted.

³An indirect benefit of this assumption is that it clearly distinguishes my results from other models of repeated elections in which patience and reputation drive policy dynamics (e.g., Duggan 2000).

⁴It is straightforward to establish that for generic histories, the policy choice when patience is positive but sufficiently low is in a neighborhood of the choice characterized here.

⁵In this way my focus differs from the experimentation literature in that learning is *passive* rather than *active*. I adopt the approach (see, for example, Piketty (1995)) of simplifying preferences to gain tractability in a general decision environment. The literature on active experimentation, in contrast, restricts attention to much simpler environments (e.g., independent bandits) or focuses on limit results (Aghion et al 1991).

first step to understanding how a social planner – who cares about future generations – may use the design of political systems to encourage (or discourage) experimentation and overcome the short-termism inherent in the policy choice of each generation. The extensions are also of positive interest, showing how behavior varies in features of real political competition.

1.1 Related Literature

This paper is related to several distinct literatures inside and outside of political economy. One broad connection is to the large literature on experimentation and learning. Formally, modeling uncertainty as a Brownian path corresponds to a bandit problem with a continuum of correlated, deterministic arms. To the best of my knowledge, no existing model considers such a structure.^{6,7}

Within political economy, most dynamic models assume the policy mapping to be common knowledge.⁸ An exception – and work closest in spirit to mine – is that on boundedly-rational policy making, derived largely from Lindblom’s (1959) theory of *muddling through* (incrementalism is the most famous element of this theory). Bendor (1995) formalizes and refines Lindblom, developing a boundedly-rational model in which a binary choice is made each period between a status quo and an exogenously supplied alternative. Among other things, he uses the model to subject incrementalism to formal analysis and finds it wanting. I obtain a similar conclusion in a rational model in which a continuum of policy alternatives are available each period (what Lindblom calls *comprehensiveness*), and show how optimal search behavior is affected by where an issue is in its life-cycle.

Kollman, Miller, and Page (2000) propose a related model of “structural search” that shares with the current paper the feature that outcomes from different policies are correlated (although the models are otherwise different).⁹ In their model policy makers search around the policy space for better policies according to some predetermined algorithm.¹⁰ I differ in that I solve for the fully rational policy choice. In addition

⁶To be sure, I am able to manage this general structure by simplifying voter preferences. See footnote 5 and the associated discussion on this point.

⁷The richness of uncertainty also distinguishes the Brownian framework from robustness issues in macroeconomics surrounding model uncertainty. That literature limits attention to recursive structures with finite (and typically few) unknown parameters (Brainard (1967) is the seminal reference, Hansen and Sargent (2007) provide a thorough treatment).

⁸A literature on repeated elections originating with Barro (1973) and Ferejohn (1986) deals with adverse selection and moral hazard issues in selecting candidates, neither of which is present here (Duggan (2000) extends these models to the choice of a policy rather than effort). A more recent literature studies repeated legislative bargaining in various environments (Baron 1996 is an early reference).

⁹Relaxing the structural aspect of the problem, Kollman, Miller, and Page also study a model of rational but “random” search in which policies are independent and sampled from a fixed distribution.

¹⁰Following Page (1996), Kollman et al define the difficulty of the policy problem as the number of nonlinearities in the policy process. As the Brownian path has infinite nonlinearities this definition is

to providing a contrast to boundedly rational results, my findings complement them by showing when particular search algorithms are optimal, thereby suggesting when a triangulating candidate or a candidate of bold change may do well.

Piketty (1995) develops a dynamic, rational model with impatient actors to explore the relative roles of luck and skill in economic outcomes. Although sharing a focus on learning with my model, his model is otherwise distinct: he assumes uncertainty is over only a few parameters, that learning is confounded by the coarseness of outcomes (binary), and that agents make individual economic decisions as well as vote. Moreover, Piketty’s focus is also different: he is interested in how heterogenous beliefs (about social mobility) persist through time, whereas in my model all players have symmetric (but incomplete) information.

Uncertainty over the policy process has found more frequent application in one-shot models of politics. In binary policy models, Fernandez and Rodrik (1991) observe that policy making may be inefficient if the identities of a policy’s beneficiaries are ex-ante unknown (see also Mitchell and Moro 2006), and recently Majumdar and Mukand (2004) show that incumbents’ electoral incentives can induce inefficient experimentation due to signaling effects. Neither of these sources of inefficiency is present in my model.¹¹

Policy uncertainty is also at the heart of expertise in policy making. Gilligan and Krehbiel (1987) explore how policy uncertainty affects the delegation of authority in legislatures, and Roemer (1994) develops a model of elections in which expert candidates communicate with voters (see also Schultz 1996). Tailored to one-shot games, however, the non-expert’s uncertainty in these settings is limited to a single piece of information, which in a dynamic environment induces full revelation of the policy mapping after the first period. In contrast, the richness of the Brownian motion implies that learning is always incomplete (in countable time).¹²

2 Model

In each period $t = 1, 2, 3, \dots$, a majority rule election is held between two candidates X and Y . The candidates compete by committing to *policies* $x_t, y_t \in \mathbb{R}$ that they implement if elected. Policies are transformed into *outcomes* by a *policy process*, ψ . Formally, a policy process is a function that maps from the policy space to the outcome space (also single dimensional) such that: $\psi : \mathbb{R} \rightarrow \mathbb{R}$.

The electorate consists of an odd number of voters. The voters care about outcomes

not applicable here. I instead view difficulty as the ratio of variance to drift of the generating Brownian process (and refer to this as issue *complexity*).

¹¹The framework can also be applied to *policy reform* more generally (Rodrik 1996 surveys the literature); the focus here on democratic elections is but one application.

¹²In a companion paper, Callander (2008), I apply the Brownian motion structure to a one-shot interplay between an expert and a non-expert.

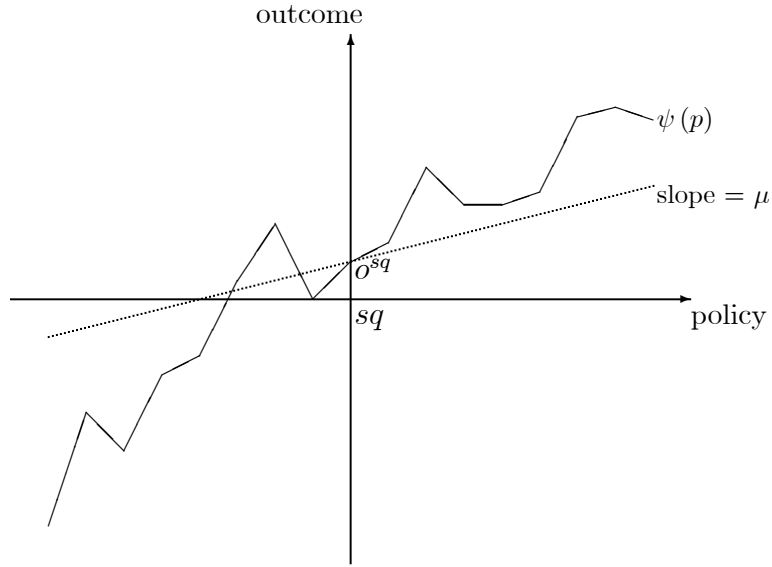


Figure 1: A Realized Path of Brownian Motion ψ : $\mu > 0$, σ^2 , and $\psi(sq) = o^{sq}$

(and indirectly about policies). Voter i 's ideal outcome is o_i and voters are ordered such that $o_i < o_j$ for $i < j$. Denote the median voter by m and set $o_m = 0$. Voters are impatient and discount the future entirely. The per period utility of voter i for policy p given outcome $\psi(p)$ is:

$$u_i(p) = -(o_i - \psi(p))^2.$$

Candidates have the same utility function over outcomes as do voters (although they need not be impatient). The ideal outcomes for candidates X and Y are $-d$ and $d > 0$. Candidates are also motivated by rents from office (ego or otherwise) that deliver a fixed benefit of $\kappa > 0$.

The true policy process ψ is determined randomly by Nature prior to period 1. To focus on learning, the same policy process is in effect for all periods. I model ψ as the realized path of a Brownian motion of drift μ and variance σ^2 .¹³ Voters and candidates know the parameters of the motion but not the realized path.¹⁴

Citizens begin with knowledge of one point in the mapping: the status quo policy and outcome, (sq, o^{sq}) . They observe a new point in the mapping each time an experimental policy is implemented, such that at election t they know up to t distinct points in the mapping. Outcomes are observed perfectly by all citizens (information is incomplete but symmetric). Figure 1 depicts one possible realization of the Brownian motion passing through the status quo point.

¹³Although Brownian motions are normally associated with movement through time, time plays no role in the realization of the path. Instead policy serves as the independent variable.

¹⁴In Section 3.4 I relax the assumption that voters know the parameters of the motion.

For all policies other than those previously tried the outcome is uncertain. For policies on the flanks, beliefs depend only on the nearest end-point (as the Brownian motion possesses the Markov property). Denoting the right-most known policy by r , beliefs for policies $p > r$ are distributed normally with:

$$\text{Expected Outcome:} \quad E\psi(p) = \psi(r) + \mu(p - r), \quad (1)$$

$$\text{Variance:} \quad \text{var}(\psi(p)) = |p - r| \sigma^2. \quad (2)$$

The drift parameter μ measures the expected rate of change and the variance the “noisiness” of the policy process. As beliefs are anchored by only one point I say they are *open-ended*. Beliefs on the left flank are defined analogously, replacing r and $\psi(r)$ with l and $\psi(l)$. At the first election beliefs are open-ended on either side of sq , as reflected in the constant drift line in Figure 1.

Policies between two known points in the mapping are said to lay on a *Brownian bridge*, with beliefs determined by the value at both ends of the bridge. For policies $p \in [q_1, q_2]$ on the bridge between q_1 and q_2 , beliefs are distributed normally with moments:

$$\text{Expected Outcome} \quad : \quad E\psi(p) = o^{q_1} + \frac{p - q_1}{q_2 - q_1} (o^{q_2} - o^{q_1}), \quad (3)$$

$$\text{Variance} \quad : \quad \text{var}(\psi(p)) = \frac{(p - q_1)(q_2 - p)}{q_2 - q_1} \sigma^2. \quad (4)$$

Expected outcomes on the bridge are given by the straight line between the two ends and are independent of the drift μ . The variance is concave over the domain, reaching its peak halfway between the ends of the bridge and equaling zero (obviously) at the ends. (Note that $\frac{d\text{var}(\psi(p))}{dp} = \pm\sigma^2$ at the ends of the bridge.)

The Brownian motion captures several key features of policy making in practice. The citizens are able to order policies according to *expected* outcomes but not according to *realized* outcomes. Thus, they know which policies are *more likely* to move outcomes in a certain direction but do not know which policies *do* move policy in that direction. The policy path is also *partially invertible* in that observing points in the mapping provides some but not all information about the outcomes of other policies and for so-far untried policies outcomes remain uncertain. Further, the accuracy of beliefs increases in the distance an untried policy is from a policy for which the outcome is known. This *proportionally invertibility* captures the intuition of Lindblom (1959) that greater uncertainty is incurred the more policy is moved from what is known.

Most intuitively, the non-monotonicity of Brownian paths captures the basic risk of policy making: that in trying to make outcomes better, changes to policy may actually make things worse. An experimental policy could produce an outcome that overshoots its intended target or moves the outcome in an unintended direction. Both of these possibilities can lead to worse outcomes, reflecting Merton’s (1936) famous Law of Un-

intended Consequences.

The volatility of policy making – and how much citizens learn from observing outcomes – is determined by the parameters of the motion and the ratio $\frac{\sigma^2}{|\mu|}$ provides a simple measure of the *complexity* of the underlying issue. The larger the ratio the less precise are citizens’ beliefs upon learning a point in the mapping and the more complex the issue.

The Brownian motion is also tractable. The simplicity of Equations 1-4 combines easily with quadratic utility as the mean and variance are sufficient to determine expected utility. For any policy p , expected utility reduces to:

$$Eu_i(p) = - [E\psi(p) - o_i]^2 - var(\psi(p)), \quad (5)$$

and this mean-variance representation of utility plays a central role in the analytic results to follow.

I restrict attention to equilibria in which voters use weakly dominant strategies, requiring that they vote for the candidate that maximizes their utility given the history at the time of the election. I do not allow abstention (although I relax this in Section 4.2). To break ties, I suppose an indifferent voter supports the candidate with greatest variance and mixes equally otherwise, although any non-degenerate tie-breaking rule can be used without meaningfully changing the results.

An equilibrium consists of strategies for both voters and candidates. Although voters are sophisticated and their preferences drive candidate behavior, their voting behavior is straightforward given the policy choices of the candidates. As such, and as is standard in models of electoral competition, I report only the strategies of candidates in describing equilibrium. Denote the equilibrium strategies of the candidates at time t by x_t^* and y_t^* .

3 Results

I begin with two preliminary results. The first shows that the power of the median voter is unencumbered by policy uncertainty and experimentation. Consequently, a repeated election analogue of Black’s (1958) median voter theorem holds and policy convergence obtains.

Lemma 1 *The (essentially¹⁵) unique equilibrium is convergent: for a given policy history at election t , platforms satisfy $x_t^* = y_t^* = \arg \max_{p_t \in \mathbb{R}} [Eu_m(p_t)]$.*

Although intuitive, Lemma 1 is not obvious. As will become apparent, voter preferences over policies are generally not single-peaked despite their outcome preferences

¹⁵Uniqueness in Lemma 1 is qualified as for some non-generic realizations of ψ divergence may occur in equilibrium. For example, if through experimentation citizens were to learn that $\psi(x) = -\psi(sq)$ for some x then it may be that the candidates diverge in equilibrium with one candidate offering policy x and the other sq . I hereafter I ignore these non-generic possibilities.

being so. The practical effect of Lemma 1 is that policy choice is reduced to a single person decision problem (with ideal outcome zero), reflecting the classic benchmark case of political competition. Later in the paper I enrich the model of political competition such that this property doesn't hold and show how policy experimentation is affected.

Policy convergence in Lemma 1 is *within* period and not *across* periods, raising the question of whether policy choices also converge through time. The next result establishes a sufficient condition for policy stability, showing that the willingness to experiment is a use-it-or-lose-it proposition. If in any period a previously implemented policy is reused citizens learn nothing new about the policy process. Consequently, the same choice is again optimal and experimentation stops. Formally, I say that policy z is *stable* on the equilibrium path if for some t' , $x_t^* = y_t^* = z$ for all $t \geq t'$. Throughout this section I focus on the utility of the median voter and refer only to the strategy of candidate X under the understanding that Y 's platforms is the same.

Lemma 2 *If $x_{t'}^* = x_t^*$ for some $t' > t$ then policy x_t^* is stable.*

I now turn to the equilibrium strategy, which I present constructively beginning with the opening election.

3.1 The First Election

The first election presents the citizenry with a basic trade-off: accept the known but imperfect status quo, or move policy in the hope of achieving a better outcome at the risk of making things worse. Put another way, should citizens trust the devil-they-know or the devil-they-don't-know? Proposition 1 provides the answer, showing when policy choice is conservative and when risk is undertaken, and for experimentation characterizes exactly the size and direction of the policy movement. Define $\alpha = \frac{\sigma^2}{2|\mu|}$ as half the complexity of the underlying issue.

Proposition 1 *The equilibrium strategy at $t = 1$ is:*

- (i) *Stable at $x_1^* = sq$ if $o^{sq} \in [-\alpha, \alpha]$.*
- (ii) *Experimental if $o^{sq} \notin [-\alpha, \alpha]$, where:*

$$x_1^* - sq = \begin{cases} \frac{\alpha - o^{sq}}{\mu} > 0 & \text{if } o^{sq} > \alpha, \\ \frac{-\alpha - o^{sq}}{\mu} < 0 & \text{if } o^{sq} < -\alpha. \end{cases}$$

If the outcome from the status quo is good enough it is immediately stable and no experimentation occurs in equilibrium. The stable outcome may diverge from the median voter's ideal outcome, although this divergence may be thought reasonable in the sense that only outcomes close to the median's ideal point can be stable.

The more interesting case is when the status quo is not good enough and the optimal response is to change policy and experiment. Experimentation leads to two questions:

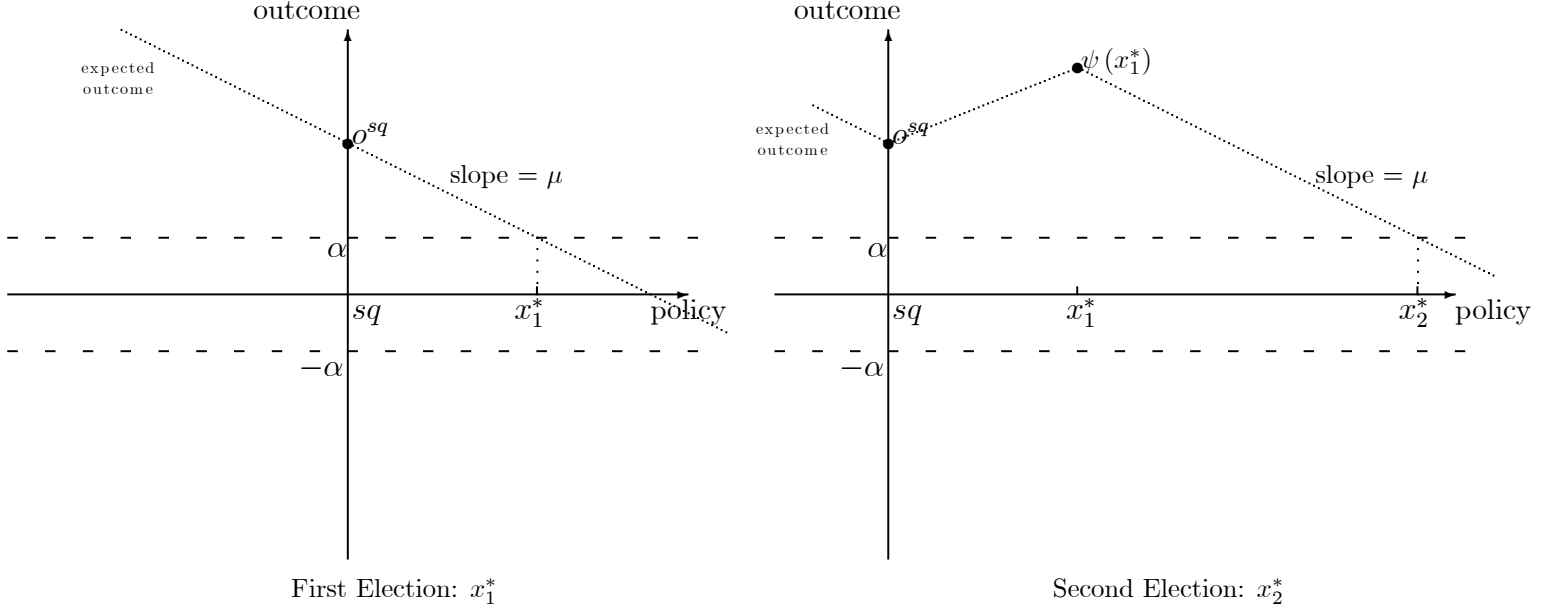


Figure 2: Monotonic Phase ($o^{sq} > \alpha, \mu < 0$)

In which direction does policy move and how far does it move? With uncertainty open-ended on either side of sq , the direction of movement depends on μ and the value of o^{sq} . The equilibrium choice involves a trade-off between greater variance and a more centrist expected outcome. Rearranging the solution for x_1^* reveals that the expected outcome is set to equal $\pm\alpha$, regardless of the location or outcome of the status quo policy (i.e., $+\alpha$ if $o^{sq} > \alpha$ and $-\alpha$ if $o^{sq} < -\alpha$). The left-hand panel of Figure 2 depicts the situation when $o^{sq} > 0$ and $\mu < 0$.

The size of a policy change is increasing in the unattractiveness of the status quo outcome and decreasing in the complexity of the underlying issue (as $\alpha = \frac{\sigma^2}{2|\mu|}$). This relationship formalizes the intuition that the citizenry is more willing to engage in risky policy making the more dissatisfied they are with the current state of affairs and the more they feel they understand an issue (lower complexity).

An interesting aspect of equilibrium behavior is that despite the candidates choosing the policy that maximizes median voter utility, it is not the median voter who most likes the policy that is ultimately offered. Rather, for an experimental policy with expected outcome α it is the voter with ideal outcome α that has the highest expected utility, with the expected utility for other voters arrayed symmetrically around α . To a naive observer, therefore, it may appear that candidates exhibit a partisan bias in their platforms where one doesn't exist.

3.2 The Second and Subsequent Elections

Following the first election the winning policy is implemented and the outcome realized. If policy at $t = 1$ was stable then learning ceases. If it was experimental voters now know two points in the policy process and update their beliefs and policy preferences accordingly. The structure of subsequent policy choices depends on whether the outcome from x_1^* is on the same or opposite side of 0 from o^{sq} . I begin with the *monotonic phase* in which $\psi(x_1^*)$ and o^{sq} (and all subsequent outcomes) are of the same sign. For ease of exposition, and without loss of generality, I describe equilibrium behavior for the case $\mu \leq 0$ and $o^{sq} \geq 0$.

Monotonicity

The case $\psi(x_1^*) \geq 0$ is depicted in the right-hand panel of Figure 2. It is easy to see that policies on the bridge between the known points are dominated by the end-point closer to 0 (policy sq in Figure 2) as this yields a more attractive expected outcome and with no risk. If experimentation is optimal, therefore, it must be on a flank and, as is the case at the first election, the right flank is preferable and the search for a good policy continues monotonically in the direction of the original movement.

Given the similarity of conditions in the second period to the first period, one may conjecture that the cut-points of Proposition 1 apply again at $t = 2$ and thereafter. Surprisingly, this conjecture is false. The first period calculations again apply *if* experimentation is optimal. What changes is that the willingness of citizens to experiment may diminish. In fact, policy stability can now be induced by sufficiently bad outcomes as well as outcomes that are sufficiently good.

The logic for candidate behavior at $t = 2$ extends to all subsequent periods in which policy has not stabilized and previous outcomes have been of the same sign. For this to be the case policy movements must have been monotonic and I refer to this as the *monotonic phase* of policy making.

Definition 1 *Policy making at election t is in the monotonic phase if $sq < x_1^* < x_2^* < \dots < x_{t-1}^*$ and $o^{sq}, \psi(x_1^*), \dots, \psi(x_{t-1}^*) \geq 0$.*

Define $\tau_t^* = \arg \min_{t' < t} [|\psi(x_{t'}^*)|, |o^{sq}|]$ as the most attractive outcome realized up to election t . Equilibrium behavior in the monotonic phase is described by the following.

Proposition 2 *In the monotonic phase at election $t \geq 2$, the equilibrium strategy is:*

(i) *Stable at:*

$$x_t^* = x_{t-1}^* \text{ if } \psi(x_{t-1}^*) \in [0, \alpha].$$

(ii) *Stable at:*

$$x_t^* = \tau_t^* \neq x_{t-1}^* \text{ if } \psi(x_{t-1}^*) > \delta_t = \frac{\alpha^2 + \psi(\tau_t^*)^2}{2\alpha}.$$

(iii) *Experimental with: $x_t^* > x_{t-1}^*$ if $\psi(x_{t-1}^*) \in (\alpha, \delta_t]$, where*

$$\psi(x_{t-1}^*) + \mu(x_t^* - x_{t-1}^*) = \alpha.$$

Good-enough stability (part i) occurs only at the most recent policy choice, with the boundary of α holding constant throughout the monotonic phase.

The second type of stability that emerges (part ii) is rather different and when it happens I say that policy making gets *stuck*. The stable policy in this case is τ_t^* (the previous most attractive policy) and not the most recent choice. For policy to backslide in this way to a previously chosen – and discarded – policy it is necessary for the outcome at x_{t-1}^* to have moved in the opposite direction to that anticipated. The importance of moving in the wrong direction is not that the policy will be chosen again, but that it reduces the expected utility of further experimentation. Consequently, although experimentation at time t is always preferable to a bad outcome of policy x_{t-1}^* , it may not be preferred to less unattractive policies that had been previously implemented. For a sufficiently bad outcome at $t-1$, therefore, policy making gets stuck. The critical value δ_t is that which equates the expected utility from experimenting to the sure outcome at τ_t^* .¹⁶

Policy making that is stuck represents a novel informational failure of policy making. Outcomes can persist that are significantly divergent from the median voter's ideal, showing how informational traps can appear and lock-in bad outcomes. This possibility arises, notably, without the presence of special interests, asymmetric information, or an agency problem. Getting stuck is also not due to the limitations of a hill-climbing algorithm. Voters evaluate all possible policies fully rationally and know that with probability one an ideal policy exists, yet looking for this policy involves a large change in policy and is costly. Consequently, policy gets stuck due to endogenous costs of searching and not an exogenously imposed limitation on policy choice.¹⁷

The possibility of getting stuck at an unattractive outcome raises the question of

¹⁶Thus in this situation the median voter's expected utility is necessarily not concave over policies.

¹⁷Getting stuck is exacerbated by voter impatience but does not depend on complete impatience. For any discounting of the future policy making gets stuck with positive probability, although the necessary condition is weakening in the level of patience.

whether there is a bound on where policy can stabilize. Corollary 1 shows that the bound is given only by the status quo outcome (since sq can always be chosen). As the status quo outcome can itself be set arbitrarily, it is possible to construct paths such that policy making gets stuck at any arbitrarily inefficient outcome.

Corollary 1 *In the monotonic phase policy making can get stuck at any outcome in $[-o^{sq}, o^{sq}]$ when $o^{sq} > \alpha$.*

The substantive difference between getting stuck and good-enough stability is best illustrated by the expected utility each delivers relative to that of the policy choices in preceding periods. When policy stabilizes because it is good enough, the expected utility is greater than in all previous periods, reflecting the achievement of a better (and good enough) outcome. In contrast, when policy gets stuck the utility level is lower than the expected utility in the immediately preceding periods, reflecting an unlucky set of realized outcomes.

Triangulation

The monotonic phase continues indefinitely until a policy proves stable (for any reason) or policy making over-shoots and an outcome is realized on the opposite side of the median voter's ideal point. In this case a Brownian bridge spans the median voter's ideal point and policy making transitions to the *triangulating phase*. The triangulating phase can begin as early as the second election and once it starts it cannot be reversed. Formally, the triangulating phase is defined as follows.

Definition 2 *Policy making at election $t \geq t^\Delta$ in the triangulating phase if it has not yet stabilized, $sq < x_1^* < \dots < x_{t^\Delta-1}^*$ and $o^{sq}, \psi(x_1^*), \dots, \psi(x_{t-2}^*) > 0 > \psi(x_{t^\Delta-1}^*)$.*

The beginning of the triangulating phase marks the end of monotonic search. Hereafter experimentation must be on a *spanning bridge* (across zero) as all other experimental policies are dominated by known points. The left panel of Figure 3 depicts the situation at election t^Δ . Hereafter all experimental policies will be in the interval between $x_{t^\Delta-2}^*$ and $x_{t^\Delta-1}^*$.

Behavior at the beginning of the triangulating phase mimics in its simplicity that at the beginning of the monotonic phase: Stabilize at the most recent policy choice if it is good enough, otherwise continue experimenting. Despite the similarity, the trigger for stability and the degree of experimentation differ. For the bridge $\widehat{w \cdot z}$ between generic policies w and z define:

$$\alpha(\widehat{w \cdot z}) = \frac{\sigma^2}{-2 \left(\frac{\psi(z) - \psi(w)}{z - w} \right)},$$

where the bracketed term in the denominator is the slope of the bridge between w and z . Formally, $\alpha(\widehat{w \cdot z})$ generalizes α by replacing μ with the slope of the bridge. Equilibrium behavior upon first entering the triangulating phase is as follows.

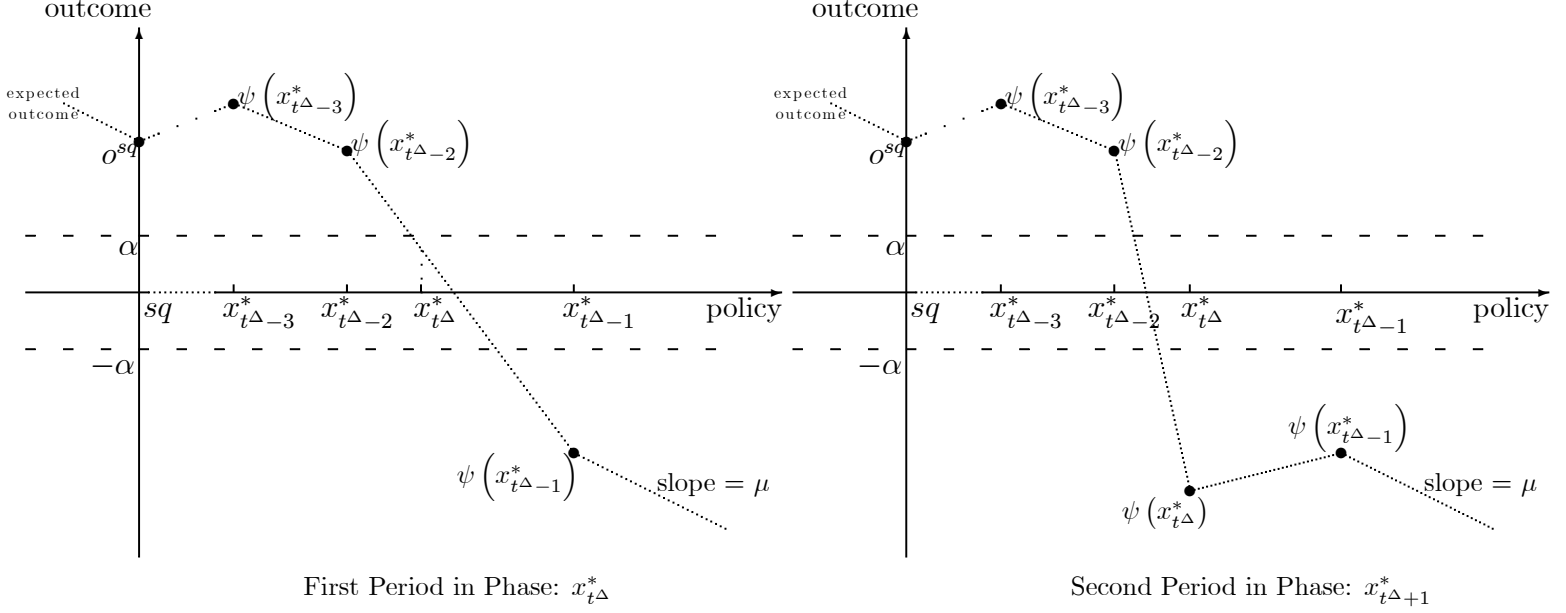


Figure 3: Triangulating Phase ($o^{sq} > \alpha, \mu < 0$).

Proposition 3 *At election t^Δ in the triangulating phase the equilibrium strategy is:*

- (i) *Stable at $x_t^* = x_{t^\Delta-1}^*$ if $|\psi(x_{t^\Delta-1}^*)| \leq \alpha \left(\widehat{x_{t^\Delta-2}^* \cdot x_{t^\Delta-1}^*} \right)$.*
- (ii) *Experimental otherwise, where $x_t^* \in (x_{t^\Delta-2}^*, x_{t^\Delta-1}^*)$ solves:*

$$E[\psi(x_{t^\Delta}^*)] = \alpha \left(\widehat{x_{t^\Delta-2}^* \cdot x_{t^\Delta-1}^*} \right) \left[1 - 2 \frac{x_{t^\Delta}^* - x_{t^\Delta-2}^*}{x_{t^\Delta-1}^* - x_{t^\Delta-2}^*} \right].$$

In addition to the direction of search, behavior here is distinguished from the monotonic phase in the nature of experimentation. Driving this difference is two factors. First, by construction the slope of spanning bridge is steeper than μ . And second, the variance of outcomes across the bridge is concave. Both factors make experimentation more lucrative than when uncertainty is open-ended.

The steeper slope of the spanning bridge implies the boundary on stability (defined implicitly in part i) is strictly tighter than in the monotonic phase. Thus, the triangulating phase makes voters less willing to settle for any particular outcome. Combined with the concavity of variance across a bridge, the expected outcome of an experimental policy (defined implicitly in part ii) is more centrist than the stability boundary in part (i), also different from the monotonic phase.

An upshot of a greater preference for risk is that policy making cannot get stuck at election t^Δ . As earlier policies did not prove stable at election $t^\Delta - 1$, they cannot prove stable at t^Δ when the pay-off from experimentation is greater; thus only the most recent policy can prove stable and only because it is good enough.

If an experimental policy is chosen at t^Δ the triangulating phase continues. The realization of $\psi(x_{t^\Delta}^*)$ breaks into two the bridge between $x_{t^\Delta-2}^*$ and $x_{t^\Delta-1}^*$, only one of which is spanning, as depicted in the right side panel of Figure 3. As experimentation in the subsequent period must be on the spanning bridge the process repeats and a simple induction argument establishes that throughout the triangulating phase a unique spanning bridge exists.

Lemma 3 *In the triangulating phase one and only one Brownian bridge is spanning.*

The logic for behavior at election $t^\Delta+1$ (the second period of the triangulating phase) extends to all later periods and is described in Proposition 4. For election $t > t^\Delta$, denote the endpoints of the unique spanning bridge by x_l^* and x_r^* (omitting dependence on t for simplicity) where by construction one of the ends is x_{t-1}^* , the most recently chosen policy. Recall that policy τ_t^* delivers the most centrist outcome of those observed up until time t .

Proposition 4 *At election $t > t^\Delta$ in the triangulating phase the equilibrium strategy is:*

- (i) *Stable at $x_t^* = x_{t-1}^* \in \{x_l^*, x_r^*\}$ if $|\psi(x_{t-1}^*)| \leq \alpha \widehat{(x_l^* \cdot x_r^*)}$.*
- (ii) *Stable at $x_t^* = \tau_t^* \notin \{x_l^*, x_r^*\}$ if $|\psi(\tau_t^*)| < \frac{\sigma}{2} \sqrt{|x_r^* - x_l^*|}$ and $\psi(x_l^*) \approx -\psi(x_r^*)$.*
- (iii) *Experimental otherwise, where $x_t^* \in (x_l^*, x_r^*)$ solves:*

$$E[\psi(x_t^*)] = \alpha \widehat{(x_l^* \cdot x_r^*)} \left[1 - 2 \frac{x_t^* - x_l^*}{x_r^* - x_l^*} \right].$$

Proposition 4 differs from Proposition 3 in adding the possibility of getting stuck. The logic for getting stuck is similar to the monotonic phase but the trigger is different. In contrast to the monotonic phase, very bad outcomes make experimentation more attractive, ensuring it continues. Instead, policy making gets stuck only for moderately bad outcomes.

The logic of this result follows from the properties of experimentation on a bridge. A really bad outcome is more valuable than a moderately bad outcome as it makes the bridge steeper, meaning an expected outcome near zero can be obtained for smaller variance (as the corresponding policy is closer to the other end of the bridge). On the other hand, a moderately bad outcome is less valuable than a good outcome as expected outcomes near zero require a higher variance (as the corresponding policy is then further from the end of the bridge). In combination, this implies the utility of further experimentation is lower when the most recent outcome is moderately bad (when $\psi(x_l^*) + \psi(x_r^*)$ is small).

In fact, the expected utility of experimenting is minimized when $\psi(x_l^*) = -\psi(x_r^*)$, which delivers an expected utility of $-\frac{|x_r^* - x_l^*|}{4} \sigma^2$ that depends only on the width of the bridge (as the expected outcome of the optimal experimental policy is zero). This value

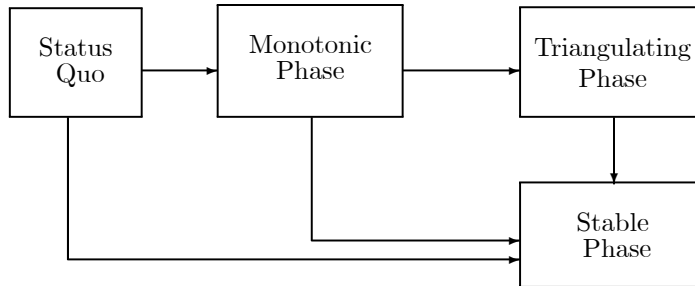


Figure 4: Timeline: Sequence of Phases.

provides a lower bound on experimenting and together with $\psi(\tau_t^*)$ a necessary condition for getting stuck.¹⁸

Good-enough stability is, of course, also possible. What voters deem good enough, however, is contracting throughout the triangulating phase. Driving this property is that the unique spanning bridge either gets steeper or policy stabilizes, thus if experimentation continues voters are less willing to settle. This property is described in Corollary 2.

Corollary 2 *The boundary on good-enough stability is strictly converging throughout the triangulating phase.*

Stability

The triangulating phase continues indefinitely until policy making stabilizes. Figure 4 depicts the sequence of policy making phases. An outstanding question, however, is whether policy making moves through all phases and ultimately stabilizes. This question is non-trivial as, by Corollary 2, the bound on stability approaches zero and if it converges too quickly may not be reached with positive probability. Nevertheless, Proposition 5 confirms that a stable policy emerges in equilibrium almost surely.

Proposition 5 *With probability one, a stable policy appears on the equilibrium path.*

This result confirms that along the equilibrium path learning eventually stops and policy settles down. However, as stability occurs in finite time, learning is incomplete and the convergence of outcomes to zero does not necessarily obtain. Almost surely, therefore, the policy that proves stable delivers an outcome divergent from the median voter's ideal.

¹⁸The precise bounds on getting stuck can be calculated but are not particularly illuminating and I do not consider them in detail. They derive from the solution to optimal experimentation in (iii), such that policy making gets stuck if and only if:

$$|\psi(\tau_t^*)| < |E[\psi(x_t^*)]|^2 + \frac{(x_t^* - x_l^*)(x_r^* - x_t^*)}{x_r^* - x_l^*} \cdot \sigma^2.$$

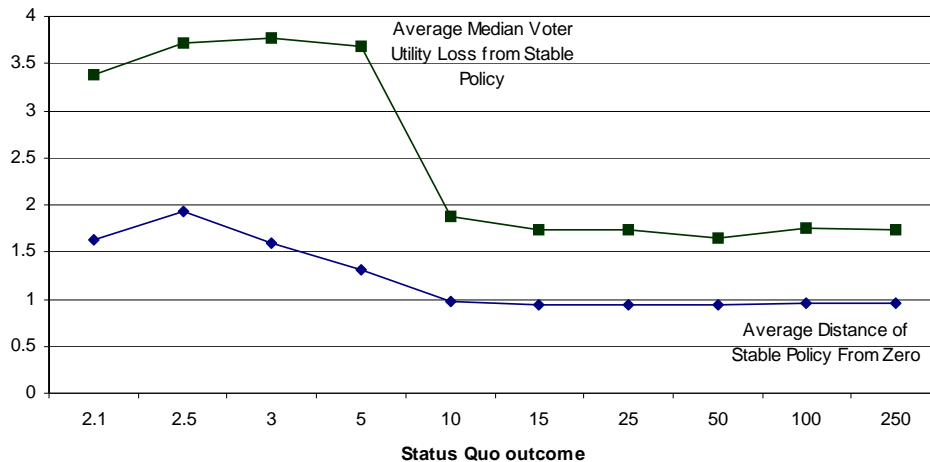


Figure 5: Properties of Stable Policy as a Function of Status Quo Outcome, o^{sq}

3.3 Simulations

Several questions of interest are not accessible analytically. In this section I offer simulations of the dynamic policy making process that provide insight to questions such as: How many periods pass before policy stabilizes? How close on average is the outcome of the stable policy to the preferences of the median voter? How often does policy making get stuck?¹⁹ The focus of the analysis is in varying the outcome of the status quo. For all simulations I fix $\mu = -1$ and $\sigma^2 = 4$, such that $\alpha = 2$.

I begin by looking at the average properties of the policies that ultimately prove stable. I depict in Figure 5 two related measures: the average utility loss for the median voter from the stable policy, and the distance of the policy from zero. The most striking feature is that the utility loss is not increasing in the unattractiveness of the status quo. In fact, the utility loss is non-monotonic in o^{sq} , at first increasing before slowly decreasing, with a precipitous drop between $o^{sq} = 5$ and $o^{sq} = 10$, and is relatively flat thereafter.

To a social planner who cares about all generations (adopting the interpretation that each generation chooses policy only once), this pattern implies that overall welfare is enhanced if the initial generation of voters faces a *worse* status quo policy.²⁰ Intuitively, two factors drive this result. First, if policy starts at an outcome just beyond the boundary α voters have a higher probability of observing an outcome inside $[-\alpha, \alpha]$ but near the boundaries and stopping there. Second, the relative attractiveness of the sq for lower o^{sq} increases the frequency with which policy making gets stuck at sq .

¹⁹The simulations were performed via a Matlab program that is available from the author. For each o^{sq} the values reported are for 10,000 iterations.

²⁰Contra to standard arguments (e.g., Drazen and Grilli 1993), therefore, society may be better off if reform does not *begin* until a “crisis” situation has been reached.

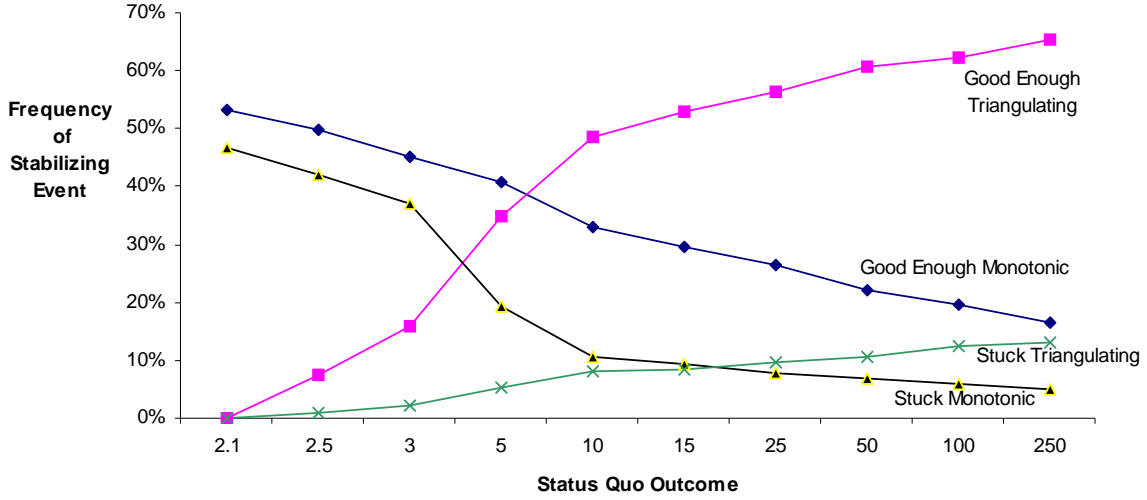


Figure 6: Frequency of Stabilizing Event as a Function of Status Quo outcome, o^{sq}

o^{sq}	2.1	2.5	3	5	10	15	25	50	100	250
Average	1.53	1.73	1.94	2.54	3.02	3.19	3.39	3.61	3.82	4.07
Min	1	1	1	1	1	1	1	1	1	1
Max	4	13	17	17	17	17	25	19	22	20
Stand. dev.	.56	.89	1.14	1.63	1.88	1.93	1.97	2.02	2.00	2.05
Ave. Mono	1.53	1.59	1.65	1.81	1.90	1.90	1.92	1.92	1.94	1.96
Ave. Triang.	0.0004	0.13	0.28	0.73	1.12	1.29	1.47	1.70	1.88	2.11

Table 1: Number of Policy Changes Before Stability

Figure 6 shows the relative frequency of events that trigger stability. For low o^{sq} policy making stabilizes directly from the monotonic phase and getting stuck is a real possibility. As o^{sq} increases the probability of entering the triangulating phase increases, and for large enough o^{sq} the probability of a good-enough outcome being achieved in the triangulating phase is the predominant trigger of stability. Even in these cases, however, the probability of getting stuck is non-trivial and the sum across the monotonic and triangulating phases always exceeds 17% and ranges up to 50%.

Table 1 reports the number of times the policy is changed before stabilizing. The notable feature is that stability is achieved on average relatively quickly, although the range is broad. Driving this speed is that voters know the drift and variance of the policy process, allowing them to make dramatic changes to policy with relative confidence that outcomes will move in the intended direction. Consistent with this interpretation is that most of the increase for larger values of o^{sq} is due to an increase in the length of the triangulating phase. I consider this issue further in the following section.

3.4 Robustness

Although voters in this environment face considerable difficulty in finding good policies, their task is simplified by knowledge of the underlying parameters of the policy process. In practice, they may lack even this much knowledge. I show here how parameter uncertainty can further constrain the willingness of voters to experiment with policy. For brevity I limit myself to the more interesting case of uncertainty over the drift parameter and present a partial characterization of behavior. I compare behavior from the benchmark model with known drift μ to the case where citizens' prior beliefs ascribe equal probability to two values $\mu_1 < \mu_2$, with $\frac{\mu_1 + \mu_2}{2} = \mu < 0$ (and $\alpha = \frac{\sigma^2}{2|\mu|}$).

I begin with the first election. It was shown previously that with no uncertainty over drift the status quo is stable if $o^{sq} \leq \alpha$ and otherwise the experimental policy chosen has an expected outcome of α . I show here that with uncertainty over drift the first of these properties holds whereas the second does not. Instead, citizens are less bold when they do experiment, producing an expected outcome that is strictly more divergent than α .

Corollary 3 *With drift uncertainty, the equilibrium strategy at $t = 1$ is:*

- (i) *Stable at $x_1^* = sq$ if $o^{sq} \in [-\alpha, \alpha]$.*
- (ii) *Experimental if $o^{sq} > \alpha$, where $E\psi(x_1^*) > \alpha$ and strictly increasing in o^{sq} .*

At the sq policy, the expected marginal gain of experimentation is $\frac{\mu_1 + \mu_2}{2}$, the same as when drift is known and this leads to the same stability cut-point. The equivalence breaks down, however, for positive levels of experimentation. The gain from the steeper possible drift value (μ_1) is tempered by the fact that if this is the true drift the expected outcome is already close to zero and the marginal gain is small. Although the reverse holds for the flatter drift μ_2 , the average of the two marginal gains leads to less experimentation.

In subsequent elections the effect of drift uncertainty is more subtle and substantial. Whereas with no drift uncertainty unfavorable outcomes are attributed to simple bad luck, they must now be interpreted for what they imply about drift. How extensively this inference problem affects behavior depends on whether uncertainty extends to the sign of the drift as well as the magnitude. I begin with $\mu_1 < \mu_2 < 0$ and behavior at the second election.

Corollary 4 *With drift uncertainty and $\mu_1 < \mu_2 < 0$, the equilibrium strategy at $t = 2$ is in part:*

- (i) *Stuck at $x_2^* = sq$ if $\psi(x_1^*) > \hat{\delta}_2$ where $\hat{\delta}_2 < \delta_2$, as defined in Proposition 2.*
- (ii) *Experimental if $\psi(x_1^*)$ is in a neighborhood of α .*

Drift uncertainty produces contrasting effects on second period behavior, depending on the outcome realized after the first election. A bad outcome at $t = 1$ leads voters

to assign more weight to μ_2 , rendering further experimentation less attractive and increasing voters' willingness to backslide to the *sq* policy. In contrast, a good outcome at $t = 1$ pushes more weight onto μ_1 , making further experimentation more attractive. Thus, policies that produce outcomes in a neighborhood of α – even outcomes closer to zero – are not stable.²¹

Behavior is not so straightforward when $\mu_1 < 0 < \mu_2$, although the relevance of this case to ideological policy making is unclear. It implies that citizens can order policies according to expectations but cannot identify which end of the policy spectrum delivers liberal outcomes and which end delivers conservative outcomes. Such extreme uncertainty may be relevant to some sorts of issues. When it does apply it gives rise to an additional justification for course reversal as following a sufficiently bad outcome policy choice may move to the other side of the status quo (violating the precepts of the monotonic phase).

Regardless of the impact of drift uncertainty throughout the monotonic phase, it has no impact once policy making enters the triangulating phase. At this point experimentation is on a bridge, where the true drift is irrelevant to beliefs (although variance uncertainty would still be relevant).

4 The Structure of Political Competition

Driving the results of Section 3 is a simple model of political competition. In this section I enrich the political environment in two ways. In both variations the power of the median citizen is relaxed such that the analysis does not reduce to a single person decision problem. These extensions contribute to a positive understanding of how realistic features of politics affect experimentation and the efficiency of search. More interestingly, they also offer a first step toward the broader normative question of how the design of a political system affects the level of experimentation and the efficiency of policy making.

4.1 Candidate Divergence and Long-Term Efficiency

In practice candidates are unsure about the preferences of voters. To incorporate this uncertainty I amend the model as follows. In addition to policy, voters evaluate candidates on a non-policy valence component. Specifically, voter i 's utility from policy p when offered by candidate $J \in \{X, Y\}$ is:

$$u_i^J(p) = -(o_i - \psi(p))^2 + \gamma_t^J.$$

²¹It does not follow immediately that the stability region at $t = 2$ is strictly tighter than $[0, \alpha]$ as it is possible that the domain of good-enough stability is not connected under drift uncertainty.

Let $\gamma_t = \gamma_t^X - \gamma_t^Y$ be the difference in valence evaluations, where γ_t is distributed symmetrically and with full support over $[-\lambda, \lambda]$. The valence evaluation γ_t is common to all voters and a new γ_t is drawn independently each period.²² Valence is not observed by candidates when choosing their policy positions. Unless otherwise specified, set $\lambda = \infty$ such that both candidates have a positive probability of winning for any pair of platforms. Hereafter candidates care only about policy outcomes ($\kappa = 0$) and are equally impatient as voters.²³

The effect of valence uncertainty is to “smooth out” candidate payoffs and Wittman (1983) and Calvert (1985) famously show in the traditional model that this induces the candidates to diverge in their policy offerings. In this section I show that within a learning environment policy divergence has contrasting effects: divergence is inefficient *within* period and strictly disliked by a majority of voters, whereas *across* periods it can improve efficiency by acting as a catalyst to experimentation. However, the increased experimentation across-periods does not hold always, for as will be shown in Example 1 below it is possible for divergent platforms to actually reduce a society’s willingness to experiment.

First, however, I present the surprising result that the logic of Wittman and Calvert depends on perfect knowledge of the policy process. With uncertainty over the policy process candidates may still converge in their platform choices.²⁴ Driving the difference here is the distinction between *ideal outcomes* and *ideal policies*, for while candidates differ in the former they may align on the latter when knowledge of the policy process is imperfect.

Lemma 4 *For a non-empty set of histories the candidates converge ($x_t^* = y_t^*$) in equilibrium. They do so if and only if they share a common ideal policy.*

The candidates can share an ideal policy only at a known point and the policy must also be the median voter’s most preferred. Political agreement is possible, therefore, even in this environment and when it occurs policy stabilizes and learning stops. In this case I say that stability is *convergent*, otherwise I say stability is *divergent*.

I turn now to the question of how divergence affects policy making within and across periods. Within period the effect is clear: disagreement among the candidates does not improve policy making. It is disliked by the median voter – whose favorite policy may no longer be offered. Less obviously, it is also disliked by at least all voters to one side of the median.

²²One can think of the parties nominating different candidates each period.

²³Equilibrium existence (in possibly mixed strategies) is assured by appropriately truncating the policy space and noting the utility of voters and candidates is continuous in policies (as the full support of γ_t implies the probability of winning for each candidate is continuous).

²⁴Wittman (1983) and Calvert (1985) show that convergence may obtain if candidates are motivated by winning office ($\kappa > 0$), whereas here candidates may converge even if $\kappa = 0$.

Lemma 5 *For any pair of platforms $x_t \neq y_t$ at election t , a strict majority of voters strictly prefer the convergent platforms $x_t^* = y_t^* = \arg \max_{p_t \in \mathbb{R}} [Eu_m(p_t)]$.*

To study the across-period impact of policy divergence I ask: For each history, is policy more likely to stabilize if candidates converge or diverge? The idea is that if stability occurs for strictly fewer histories here than in the baseline model then divergence can be said to encourage experimentation. In what follows I show that the answer to this question depends on the phase.²⁵

I begin with the monotonic phase and show that for this phase the possibility for candidate divergence strictly increases experimentation. Convergent stability is still possible here but it is more difficult to obtain than in the baseline model. For the candidates to share an ideal policy the policy's outcome must be within α of both candidates' ideal (which is impossible if $\alpha < d$). Getting stuck is also more difficult to obtain as, relative to the median voter and candidate Y , candidate X (with ideal outcome $-d$) is more willing to experiment than backtrack and get stuck. Denote by $z_t^* \in \{x_t^*, y_t^*\}$ the winning policy at election t , and retain the assumptions $\mu \leq 0$ and $o^{sq} \geq 0$.

Lemma 6 *At election t in the monotonic phase, convergent stability obtains only if:*

- (i) $\psi(z_{t-1}^*) \in [0, \alpha - d]$, or
- (ii) $\psi(z_{t-1}^*) > \hat{\delta}_t = \frac{\alpha^2 + \psi(\tau_t^*)^2 + d^2 + 2d(\psi(\tau_t^*) - \alpha)}{2\alpha} > \delta_t$.

Divergent stability is also possible in the monotonic phase although relatively rare. A necessary condition is that the median voter's ideal policy is a known point, a condition that is also necessary for stability in the baseline model (if this did not hold candidate X could deviate to the median's ideal policy and make himself better off). As histories can be found where stability obtains in the baseline model but not here (for example, a status quo policy with $o^{sq} \in (\alpha - d, \alpha)$), the following result obtains.

Proposition 6 *At election t in the monotonic phase, policy making stabilizes for strictly fewer histories than in the baseline model.*

In the triangulating phase the incidence of convergent stability is similarly reduced. Lemma 7 offers several conditions necessary for convergent stability based on the width of the spanning bridge and the distance of realized outcomes from zero. Define $\tau_t^{+*} = \arg \min_{x_{t'} < t} [\psi(x_{t'}^*), o^{sq}]$ and $\tau_t^{-*} = \arg \max_{x_{t'} < t} [\psi(x_{t'}^*)]$ as the most attractive outcomes realized up to election t on either side of zero (the previously defined τ_t^* is the policy among these two that is most attractive to the median voter).

²⁵Page and Zharinova (2006) explore a related question in a boundedly rational model of policy choice, showing how a pair of competing candidates who diverge in platforms can outperform a benevolent social planner.

Lemma 7 *In the triangulating phase, convergent stability obtains only if the requirements of Propositions 3 and 4 are satisfied, $d < \max \left[\frac{\psi(\tau_t^{+*})}{2}, \frac{|\psi(\tau_t^{-*})|}{2} \right]$, and $d^2 < \frac{\sigma^2}{4} (z_r^* - z_l^*)$, where z_l^* and z_r^* are the two ends of the spanning bridge.*

The triangulating phase makes it more difficult for candidates to share an ideal policy as they each have a known point that delivers an outcome on their side of zero. The conditions in the lemma require that at least one of the known points delivers a sufficiently divergent outcome and that the spanning bridge is not too narrow (as otherwise experimentation dominates). As the width of the spanning bridge only narrows through time, the width condition cannot be reversed once attained, ruling out convergent stability thereafter. This property is surprising as it contrasts with the baseline model in which convergent stability is eventually obtained almost surely.

Turning to the final case – divergent stability in the triangulating phase – reveals the possible perniciousness of platform divergence. Example 1 generates divergent stability where it otherwise would not obtain in the baseline model, demolishing the prospects for a general claim that divergence always increases experimentation.

Example 1 *Suppose at $t = 2$ the known points are sq and z_1^* , where $o^{sq} > \alpha$ and $\psi(z_1^*) \approx -o^{sq}$. Let γ_t be distributed uniformly over $[-\lambda, \lambda]$ for finite λ . If d is in a neighborhood of o^{sq} then, for λ sufficiently large, divergent stability obtains with $y_2^* = sq$ and $x_2^* = z_1^*$.*

Both candidates have ideal policies at known points whereas the median voter prefers experimenting further (as both known outcomes are more than α from 0). The candidates could deviate and increase their probability of victory and in the baseline model they would do so. With valence uncertainty and λ large, however, a deviator wins election only marginally more frequently and this does not compensate for the cost of committing to a less attractive policy. The logic of the example is general and similar examples can be generated for any history.

4.2 Abstention and Experimentation

If turning out to vote is voluntary the *median voter* of a population may differ from the *median citizen* and, more importantly, may vary from election to election. In this section I allow citizens the option to abstain and show how it reduces the willingness of a society to experiment. For simplicity, assume a continuum of citizens distributed according to the density function f , where f is symmetric around zero, has full support and is single-peaked, and return to the baseline model (with no valence term).

I suppose that citizens abstain from voting when they are sufficiently *alienated* from all candidates. That is, a citizen votes for her favorite candidate if that candidate is

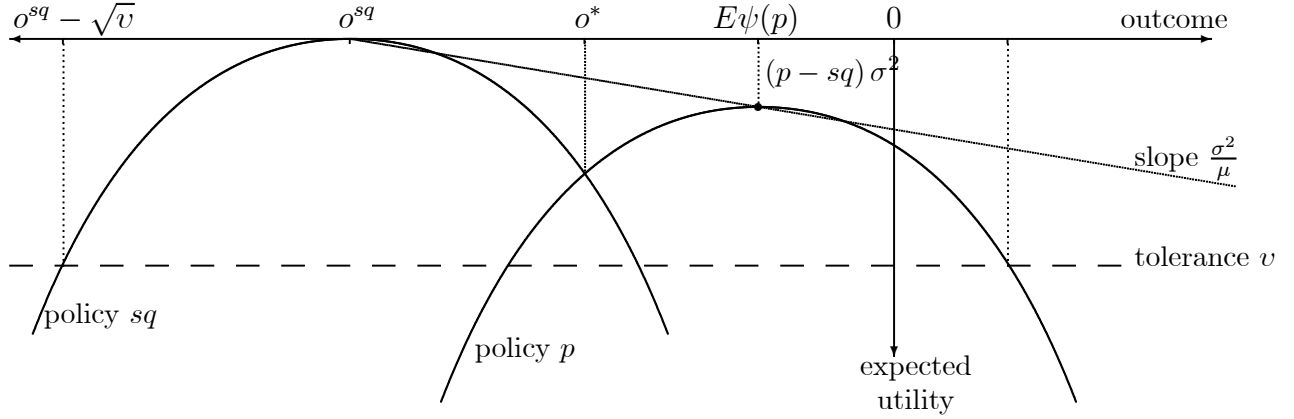


Figure 7: Expected utility for policies sq and p across citizens.

sufficiently attractive, otherwise she abstains due to alienation. Formally, citizens have a common *tolerance* level $v > 0$ such that they vote for their favorite candidate if the utility from that candidate exceeds $-v$. If no candidate meets this threshold a citizen abstains.²⁶

The threshold of tolerance implies that how much voters like a candidate, and not just who they like, matters to behavior. Proposition 7 shows that this distinction biases society toward known – and riskless – policies.

Proposition 7 *For any history at election t , there is a $v' > 0$ such that policy τ_t^* is stable if $v \leq v'$.*

The baseline model is equivalent to infinite tolerance ($v = \infty$). The proposition shows that as tolerance declines, more and more policies are stable relative to the full-turnout model. Driving this property is that the riskiness of policy experiments narrows their appeal and dampens turnout. This is most easily seen in Figure 7 that plots the expected utility across voters for a sq policy with outcome o^{sq} and an experimental alternative $p > sq$ with expected outcome $E\psi(p)$. The expected utility curves for the two policies have the same concavity but they differ in the maximum value. The lower utility from the experimental policy implies it draws support from a narrower section of the population. The sq policy gathers support for all voters \sqrt{v} to the left of o^{sq} , whereas policy p only gathers voters $\sqrt{v - (p - sq)\sigma^2}$ to the right of $E\psi(p)$ (and, as drawn, they split voters between o^{sq} and $E\psi(p)$). Nevertheless the experimental policy may still be majority preferred if voter density is greater around the center. Regardless, as voter tolerance declines the experimental policy sheds support faster than does the riskless status quo and policy stabilizes whenever tolerance is sufficiently low.

²⁶This is known as *expressive* voting. Other voting theories could be used here; I chose expressive voting as it is simple and popular. The main message of this section – that abstention limits experimentation – emerges regardless of the voting theory used.

An even stronger statement is possible for the first election. For tolerance less than α^2 no experimental policy can induce citizens to turnout and the status quo policy is stable *regardless* of the outcome it delivers.

Corollary 5 *For $v \leq \alpha^2$ the sq policy is stable at election $t = 1$ for any $o^{sq} \in \mathbb{R}$.*

Corollary 5 reflects the ability of known policies to gather in a basin of support. Regardless of tolerance, all voters within α of the status quo prefer the status quo to *any* policy experiment. Support for experimenting, then, comes only from those at least α away from the status quo, and these citizens require a level of tolerance bounded away from zero to be induced to turn out.

The results of this section show how abstention distorts political outcomes towards *safe* alternatives, reducing the willingness of society to experiment. Although a society may be clamoring for “change” – even with a majority of citizens even agreeing in which direction policy should move – the inability of those dissatisfied with the status quo to coalesce around a particular alternative means the *ability* to change is often less than the *desire*.

5 Discussion

Winning majorities. Obscured by the attention on the median voter is the nature of winning majorities and how they vary across phases. In the monotonic phase a broad consensus emerges in favor of experimenting: a majority of voters agree that experimenting is optimal and agree on the direction to move. The strongest advocates for change are the extremists who benefit the most from a shift in the expected outcome of policy.

In contrast, experimenting in the triangulating phase emerges as a compromise solution. The majority of voters generally want to stick to what is known but they cannot agree as to what is the best policy at which to stabilize. Extremists are most unhappy with the compromise solution and it may be only centrist voters who wish to experiment further.²⁷

Course Reversals. The directional pattern of policy choice that emerges in equilibrium illuminates a debate in applied policy making on the wisdom of *course reversals*. This debate turns on the issue of how a policy failure should be interpreted. One view, often attributed to the IMF and World Bank, is that a policy failure is due only to insufficient application. As Clark (20007) critically concludes:

“If the medicine fails to cure, then the only possible conclusion is that more is needed.”

²⁷The differences in coalition support across phases of policy making could be used to calibrate the theory to different points in the life-cycle of real policy issues.

The opposing view is that a policy failure reflects a mistaken choice of direction and should be immediately reversed. This view is satirized by Paul Krugman, who makes the case against knee-jerk course reversals by analogizing them as follows:²⁸

“A driver runs over a pedestrian; he looks back, realizes what he’s done. “I’m so sorry,” he says. “Let me fix the damage.” So he backs up, running over the pedestrian a second time.”

My results provide a theoretical underpinning to this informal debate. In particular, I find that the logic of course reversals is context-dependent, depending on the structure of the underlying issue and the complete history of policy choices rather than only the most recent outcome.

Stochastic Processes. The usefulness and flexibility of the Brownian motion in modeling the policy process suggests other stochastic processes may also be applicable. A natural generalization is to employ the Levy process and allow for discontinuities in the policy process. Discontinuities increase the volatility of policy making, rendering experimentation more risky. They also complicate the policy search as a policy with an outcome near zero need not exist. In the triangulating phase, therefore, citizens cannot be sure that a good policy lay somewhere between the two sides of the political divide, an uncertainty that may go some way to explaining the frequency and persistence of partisan divides in modern politics. Many other specifications of the policy process are possible, including mean-reverting processes, discrete processes, or even time-varying and non-Markovian processes.

Other Applications. That the baseline model reduces to a single person decision problem does offer one advantage: The result are directly applicable to other settings. For example, the direction and distance of change are also important in the search for a consumer product or a job (see Rogerson, Shimer, and Wright (2005) for a survey). To translate this model to labor market search, each policy may correspond to a particular job and outcomes represent the combination of pay and working conditions for that job. The job seeker has an ideal pay and working conditions combination, with outcomes above that level corresponding to a wage that is too low, and outcomes below that level to jobs requiring too many hours. The worker can order the jobs by expected outcome (as do voters with policies) but the realized job characteristics remain hidden until a job is tried.

Applied to this setting (with an ideal job outcome of zero), the results of previous sections describe the dynamics of impatient job search and explain why workers may sometimes get stuck in unattractive jobs. Moreover, interpreting each policy as a job suggests some natural extensions. For instance, employers may advertise by revealing (perhaps noisily) the characteristics of their job with the hope of enticing workers to

²⁸krugman.blogs.nytimes.com (March 3, 2008). Krugman attributes the anecdote to Jacob Frenkel.

join their firm. The logic of advertising in this sense is not straightforward, however, as if the advertisement creates a Brownian bridge it may induce a worker to experiment – and take a new job – but with a firm other than the one advertising.

6 Conclusion

The importance of experimentation and learning to policy making has long been acknowledged in practice. This paper is an attempt to address these issues theoretically through a novel model of policy making. The objective is to provide insight into when policy makers experiment with policy and the direction and size of policy experiments. The most interesting property to emerge is the possibility for policy making to get stuck. Getting stuck exposes the subtlety of learning in a complex environment with many policies, showing how undesirable political outcomes may persist even when candidates are fully responsive to the preferences of the median voter.

By necessity the model of politics developed here omits many features of real policy making. It would be interesting to explore how experimentation and policy making change in more detailed political environments or under more general policy processes. For instance, how experimentation is affected by the presence of expert lobbyists or bureaucrats, or the ability to observe other polities and learn from their experience, such as is possible in a federal system, are questions of obvious interest. The model introduced here should be sufficiently flexible to address these and other questions, although they must be left for another time.

7 Appendix

Proof of Lemma 1: Voter i 's utility for a policy p with expected value $E\psi(p)$ and variance σ_p^2 is given by Equation (5). Differentiating with respect to the ideal point:

$$\frac{du_i(p)}{do_i} = -2(o_i - E\psi(p)), \text{ and } \frac{d^2u_i(p)}{do_i^2} = -2.$$

As this holds for the policy offered by either candidate, single crossing holds and the median voter's preference is decisive.

As candidate utility functions are continuous in policy when the probability of victory is held constant, it follows by well-known arguments that in equilibrium both candidates locate at the median voter's most preferred policy (as for $\kappa > 0$ the probability of victory is discontinuous in policy). As with probability one the median voter has a unique most preferred policy, the essentially unique equilibrium is convergent. ■

Proof of Lemma 2: By Lemma 1 the strategies of candidates depend only on the

informational content of past play, a sufficient statistic for which is the set of unique winning policies. The lemma is then obvious by the stationarity of the problem. ■

Optimal Experimentation

Without loss of generality I assume hereafter that $\mu \leq 0$ and $\sigma^2 \geq 0$. At any time t' , define the right and left-most policies implemented by x_{\min} and x_{\max} . Define $\alpha(w|z, \psi(z))$ as the value of $\psi(w)$ that solves $|\psi(w)| = \alpha(\widehat{w \cdot z})$ for the Brownian bridge between policies w and z ; thus x_{t-1}^* is stable at election t in the triangulating phase if and only if $|\psi(x_{t-1}^*)| < \alpha(x_{t-1}^*|w^*, \psi(w^*))$, where w^* is the other end of the spanning bridge formed by x_{t-1}^* .

Open-Ended Uncertainty. Two properties are immediately clear: if $\psi(x_{\max}) \leq 0$ all experimental policies $z > x_{\max}$ are dominated by x_{\max} . Similarly all $w < x_{\min}$ are dominated by x_{\min} if $\psi(x_{\min}) \geq 0$. Consider then $\psi(x_{\max}) \geq 0$ and policies $z \geq x_{\max}$. Expected utility for the median voter is:

$$Eu_m(z) = -[\psi(x_{\max}) + \mu(z - x_{\max})]^2 - (z - x_{\max})\sigma^2.$$

Differentiating:

$$\begin{aligned} \frac{dEu_m(z)}{dz} &= -2\mu[\psi(x_{\max}) + \mu(z - x_{\max})] - \sigma^2, \\ \frac{d^2Eu_m(z)}{dz^2} &= -2\mu^2 < 0. \end{aligned}$$

The second derivative ensures a unique maximum; solving the first order condition for an internal solution gives:

$$\psi(x_{\max}) + \mu(z^* - x_{\max}) = \frac{\sigma^2}{-2\mu}, \quad (6)$$

and the expected outcome from policy z is set to the constant $\alpha = \frac{\sigma^2}{-2\mu}$. If $\psi(x_{\max}) \leq \alpha$ the corner solution is x_{\max} . ■

Brownian bridge. Behavior on non-spanning bridges is straightforward: the optimal policy is the end(s) of the bridge closest to zero. Consider then a spanning bridge $\widehat{x_l x_r}$ where $\psi(x_l) > 0 > \psi(x_r)$ ($\psi(x_r) > 0 > \psi(x_l)$ is analogous) and suppose $|x_l| \leq |x_r|$. The median voter's expected utility for $z \in [x_l, x_r]$ is:

$$Eu_m(z) = -\left[\psi(x_l) + \frac{(z - x_l)}{(x_r - x_l)}(\psi(x_r) - \psi(x_l))\right]^2 - \frac{(z - x_l)(x_r - z)}{x_r - x_l}\sigma^2.$$

Differentiating:

$$\begin{aligned}\frac{dEu_m(z)}{dz} &= -2\frac{\psi(x_r) - \psi(x_l)}{(x_r - x_l)} \left[\psi(x_l) + \frac{(z - x_l)}{(x_r - x_l)} (\psi(x_r) - \psi(x_l)) \right] - \frac{(x_r - z) - (z - x_l)}{x_r - x_l} \sigma^2, \\ \frac{d^2Eu_m(z)}{dz^2} &= -2 \left[\frac{\psi(x_r) - \psi(x_l)}{x_r - x_l} \right]^2 + \frac{2}{x_r - x_l} \sigma^2.\end{aligned}$$

As $u_m(x_l) \geq u_m(x_r)$ by construction and the second derivative is independent of z , $\frac{d^2Eu_m(z)}{dz^2} \geq 0$ implies x_l is the optimal policy (and x_r also if $|x_l| = |x_r|$). Straightforward algebra establishes the second derivative is positive if and only if:

$$\alpha(\widehat{x_l x_r}) = \frac{\sigma^2}{-2\frac{\psi(x_r) - \psi(x_l)}{x_r - x_l}} \geq \frac{\psi(x_l) - \psi(x_r)}{2},$$

where the right-hand-side is the average distance from 0 of the two ends of the bridge. As by definition $\psi(x_l) \leq \frac{\psi(x_l) - \psi(x_r)}{2}$, a positive second derivative requires $|\psi(x_l)| \leq \alpha(x_l|x_r, \psi(x_r))$ (with the inequality strict for $|x_l| < |x_r|$).

Consider $\frac{d^2Eu_m(z)}{dz^2} < 0$, noting this ensures a unique optimal policy. The end point x_l dominates experimenting iff $\frac{dEu_m(z)}{dz} \leq 0$ at $z = x_l$. By substituting $z = x_l$ into the first derivative and rearranging, this is true iff $\psi(x_l) \leq \alpha(x_l|x_r, \psi(x_r))$. For $\psi(x_l) > \alpha(x_l|x_r, \psi(x_r))$ a unique internal optimum exists and is found by rearranging $\frac{dEu_m(z)}{dz} = 0$, noting that the term in the square brackets is the expected value. ■

Thus, for any value of the second derivative the end point x_l is the optimal policy if and only if $\psi(x_l) \leq \alpha(x_l|x_r, \psi(x_r))$, otherwise experimenting on the bridge is preferred. I state here three properties of optimal experimentation on a bridge that later prove useful; suppose $z^* \in (x_l, x_r)$ and retain the assumption $|x_l| \leq |x_r|$:

Property i: $E[\psi(z^*)|\widehat{x_l x_r}] \geq 0$ and is closer to x_l than to x_r ,

Property ii: $E[\psi(z^*)|\widehat{x_l x_r}] < \alpha(x_l|x_r, \psi(x_r))$,

Property iii: The stable boundary in period $t + 1$, conditional on the spanning bridge being $\widehat{x_z^* x_r}$, satisfies: $E[\psi(z^*)|\widehat{x_l x_r}] < \alpha(z^*|x_r, \psi(x_r)) < \alpha(x_l|x_r, \psi(x_r))$.

Property (i) follows from the fact that for any \tilde{z} such that $E[\psi(\tilde{z})|\widehat{x_l x_r}] < 0$ there exists a corresponding \tilde{z}' such that $|E[\psi(\tilde{z})|\widehat{x_l x_r}]| = E[\psi(\tilde{z}')|\widehat{x_l x_r}]$ and with lower variance.

To establish Property (ii), substitute $E[\psi(z)|\widehat{x_l x_r}] = \alpha(x_l|x_r, \psi(x_r))$ into $\frac{dEu_m(z)}{dz}$.

Simplifying gives:

$$\begin{aligned}
\frac{dEu_m(z)}{dz}\Big|_z &= -2\frac{\psi(x_r) - \psi(x_l)}{(x_r - x_l)}\alpha(x_l|x_r, \psi(x_r)) - \frac{(x_r - z) - (z - x_l)}{x_r - x_l}\sigma^2 \\
&= -2\frac{\psi(x_r) - \psi(x_l)}{(x_r - x_l)}\frac{(x_r - x_l)}{\psi(x_r) - \alpha(x_l|x_r, \psi(x_r)) - 2} - \frac{(x_r - z) - (z - x_l)}{x_r - x_l}\sigma^2 \\
&= \sigma^2 \left[\frac{\psi(x_r) - \psi(x_l)}{\psi(x_r) - \alpha(x_l|x_r, \psi(x_r))} - \frac{(x_r - z) - (z - x_l)}{x_r - x_l} \right] \\
&> 0,
\end{aligned}$$

since $\psi(x_l) > \alpha(x_l|x_r, \psi(x_r))$ for experimentation to be optimal and the first term is greater than one, and from Property (i) $z < \frac{x_r + x_l}{2}$ such that the second term is less than one. The optimal z^* is further to the right than this point and the result follows from the negative slope of the bridge.

The first inequality of Property (iii) follows from the concavity of variance on a bridge: If $\psi(z^*) = E[\psi(z^*)|\widehat{x_l x_r}]$ then $\frac{dvar(\psi(z)|\widehat{z^* x_r})}{dz}\Big|_{z^*} = \sigma^2$ whereas on $\widehat{x_l x_r}$, the previous period's bridge, $\frac{dvar(\psi(z)|\widehat{x_l x_r})}{dz}\Big|_{z^*} < \sigma^2$ (and the bridges have the same slope, offering the same gain in expected value). In a similar vein, the second inequality follows because the $t + 1$ bridge is narrower and steeper if $\psi(z^*) = \alpha(x_l|x_r, \psi(x_r))$. Formally:

$$\alpha(x_l|x_r, \psi(x_r)) > \frac{\sigma^2}{-2\frac{\psi(x_r) - \alpha(x_l|x_r, \psi(x_r))}{x_r - z^*}}$$

since $x_r - z^* < x_r - x_l$, implying $\alpha(z^*|x_r, \psi(x_r)) < \alpha(x_l|x_r, \psi(x_r))$. ■

Proof of Proposition 1: $x_{\min} = x_{\max} = sq$ at $t = 1$. The result follows from the optimal response to uncertainty given in Equation 6. ■

Proof of Proposition 2: The requirement $\psi(x_1^*), \dots, \psi(x_{t-1}^*) \geq 0$ implies the optimal policy is in the set $\tau_t^* \cup (x_{t-1}^*, \infty)$. If $\tau_t^* = x_{t-1}^*$ the problem is equivalent to period 1: x_{t-1}^* is stable if $x_{t-1}^* \leq \alpha$ and is dominated by z^* from Equation 6 otherwise. So suppose $\tau_t^* \neq x_{t-1}^*$ and note that by induction this implies $\psi(x_{t-1}^*) > \alpha$. As x_{t-1}^* is dominated by both the optimal experimental policy z^* and τ_t^* , equilibrium behavior requires a comparison of utility. The utility from τ_t^* is straightforward. The expected utility from z^* is:

$$\begin{aligned}
Eu_m(z^*) &= -\left[\frac{\sigma^2}{-2\mu}\right]^2 - \left[\frac{\sigma^2}{-2\mu} - \psi(x_{t-1}^*)\right]\sigma^2 \\
&= \frac{1}{2}\frac{\sigma^2}{\mu} \left[\frac{1}{2}\frac{\sigma^2}{\mu} + 2\psi(x_{t-1}^*)\right],
\end{aligned}$$

which is strictly decreasing in $\psi(x_{t-1}^*)$. The result follows by setting $\frac{1}{2}\frac{\sigma^2}{\mu} \left[\frac{1}{2}\frac{\sigma^2}{\mu} + 2\delta_t\right] =$

$-[\psi(\tau_t^*)]^2$. ■

Proof of Corollary 1: As the Brownian motion has full support on \mathbb{R} , a path can be constructed for any $o \in [-o^{sq}, o^{sq}]$ such that $\psi(\tau_t^*) = o$ and $\psi(x_{t-1}) > \delta_t$ at period t . ■

Proof of Proposition 3: The optimal choice on the spanning bridge is given by properties *i-iii* above. I need then only show that this choice dominates all other policies. As by the definition of the triangulating phase $\psi(x_{t\Delta-1}^*) < 0 < E\psi(x_{t\Delta-1}^*)$ at $t-1$, the bridge is of steeper slope than μ . There exists, therefore, a $\hat{z} \in (x_{t\Delta-2}^*, x_{t\Delta-1}^*)$ such that $E[\psi(\hat{z}) | \widehat{x_{t\Delta-2}^* x_{t\Delta-1}^*}] = \alpha$ and $var(\psi(\hat{z}) | \widehat{x_{t\Delta-2}^* x_{t\Delta-1}^*}) < (\hat{z} - x_{t\Delta-2}^*) \sigma^2 < (x_{t\Delta-1}^* - x_{t\Delta-2}^*) \sigma^2$ by the properties of variance on a Brownian bridge. As $x_{t\Delta-1}^*$ dominated all $z \leq x_{t\Delta-2}^*$ at time $t-1$, policy \hat{z} dominates them also, implying the optimal policy is on the spanning bridge. ■

Proof of Lemma 3: I prove the result by induction. It is true by construction at t^Δ , the first period of the triangulating phase. Suppose it is true at some $t > t^\Delta$. If x_t^* is stable then uniqueness holds at $t+1$. If x_t^* is experimental it must be on the existing spanning bridge. If the unique spanning bridge is given by $\widehat{x_l^t x_r^t}$ with $\psi(x_l^t) > 0 > \psi(x_r^t)$, then $\psi(x_t^*) > 0$ implies $\widehat{x_l^t x_t^*}$ is spanning and $\widehat{x_t^* x_r^t}$ is not. The reverse holds for $\psi(x_t^*) < 0$, and uniqueness holds at $t+1$. The induction argument is complete. ■

Proof of Proposition 4: From Property (iii) above the threshold for stability of either end of a spanning bridge is decreasing through time. This implies that only the most recently formed end at x_{t-1}^* can prove stable, as claimed in parts (i) and (ii) of the proposition. The optimal choice on a spanning bridge (parts i and iii) are given by the properties above (the same as in Proposition 3).

Policy $\tau_t^* \notin \{x_l^*, x_r^*\}$ can prove stable if it dominates experimentation on the spanning bridge. For fixed x_l and x_r and assuming $|\psi(x_l)| \leq |\psi(x_r)|$, the expected utility of optimal experimentation is strictly decreasing in $|\psi(x_l)|$ and strictly increasing in $|\psi(x_r)|$. Utility is minimized at $\psi(x_l) = -\psi(x_r)$, in which case the optimal experimental policy is $z^* = \frac{x_l + x_r}{2}$, such that the expected outcome is 0 and:

$$\begin{aligned} Eu_m(z^*) &= -\frac{\left(\frac{x_e + x_l}{2} - x_e\right) \left(x_l - \frac{x_e + x_l}{2}\right)}{x_l - x_e} \sigma^2 \\ &= -\frac{\sigma^2}{4} (x_l - x_r). \end{aligned}$$

Thus, τ_t^* is stable iff $|\psi(\tau_t^*)| < \frac{\sigma}{2} \sqrt{|x_r - x_l|}$ and $\psi(x_l) + \psi(x_r)$ is in some neighborhood of zero. ■

Proof of Corollary 2: Property (iii) for experimentation on spanning bridges. ■

Proof of Proposition 5: In each period of the monotonic phase $\Pr[\psi(x_t^*) < \alpha] = \frac{1}{2}$ and the phase ends. Thus, the monotonic eventually ends with probability one. If it ends with a stable policy the result follows, so suppose the triangulating phase begins.

I complete the result via contradiction. Suppose the triangulating phase continues indefinitely with positive probability. This requires that the probability of stabilizing in each period goes to zero. Without loss of generality, suppose in each period the spanning bridge is such that $|\psi(x_r^t)| > \psi(x_t^t) > 0$ for $x_r^t > x_t^t$. As by property (iii) above the probability that the phase continues with $\psi(x_t^*) > 0$ is less than $\frac{1}{2}$, it must be that $\Pr[\psi(x_t^*) < 0] \rightarrow \frac{1}{2}$ as $t \rightarrow \infty$. By the symmetry of the normal distribution this implies $E[\psi(x_t^*) | \cdot] \rightarrow 0$ as $t \rightarrow \infty$ and the width of the spanning bridge approaches zero (the expected width in period $t + 1$ is less than $\frac{1}{2}$ of the length in period t). A narrowing bridge implies $var[\psi(x_t^*)] \rightarrow 0$ and by the law of large numbers the ends approach zero: $\psi(x_l^t), \psi(x_r^t) \rightarrow 0$. To determine whether experimentation on the bridge is optimal, it is sufficient to check $\frac{dEu_m(z)}{dz}$ at x_l^t . Substituting gives: $\frac{dEu_m(z)}{dz}|_{z=x_l} = -2\frac{\psi(x_r) - \psi(x_l)}{(x_r - x_l)}\psi(x_l) - \sigma^2$, which becomes negative unless $\frac{\psi(x_r) - \psi(x_l)}{(x_r - x_l)} \rightarrow -\infty$. For any given path, therefore, experimentation is suboptimal with probability one for some t , establishing the contradiction. ■

Proof of Corollary 3: The first order condition for experimentation becomes:

$$\begin{aligned} \frac{dEu_m(z)}{dz} &= -\mu_1[o^{sq} + \mu_1(z - sq)] - \mu_2[o^{sq} + \mu_2(z - sq)] - \sigma^2, \\ \frac{d^2Eu_m(z)}{dz^2} &= -\mu_1^2 - \mu_2^2 < 0. \end{aligned}$$

Without drift uncertainty, the optimal experimental policy z^* delivers expected outcome α . Substituting this into the first derivative and setting $\mu_1 = \mu - \nu$ and $\mu_2 = \mu + \nu$ gives:

$$\begin{aligned} \frac{dEu_m(z)}{dz}|_{z^*} &= -(\mu + \nu)[\alpha + (z - sq)\nu] - (\mu - \nu)[\alpha - (z - sq)\nu] - \sigma^2, \\ &= -2(z^* - sq)\nu^2 < 0. \end{aligned}$$

Thus, the optimal experimental policy under drift uncertainty is in (sq, z^*) when $z^* > sq$, and equal to z^* when $z^* = sq$. ■

Proof of Corollary 4: By stochastic dominance and $E\psi(x_1^*) < o^{sq}$, a realization $\psi(x_1^*) > o^{sq}$ induces more weight on μ_2 in posterior beliefs. Similarly, a realization $\psi(x_1^*) < E\psi(x_1^*)$ puts more weight on μ_1 . Both results then follow from straightforward algebra. ■

Proof of Lemma 4: I first prove that a policy p^* that is the most preferred by both candidates is also the median voter's most preferred. From the proof of Lemma 1 the

concavity of $u_i(p)$ in o_i is -2 for any policy p . Thus, if for some $q \neq p^*$, $u_m(q) > u_m(p^*)$ at least one candidate must also prefer q to p^* , a contradiction.

If $x_t = y_t = p^*$ a deviation by either candidate lowers the probability of winning for the deviator and leads to a strictly worse policy outcome (as λ has full support on \mathbb{R}). Alternatively, if $x_t \neq p^*$ a deviation to p^* increases X 's probability of winning and strictly improves the policy outcome. Suppose instead the candidates do not share an ideal outcome but converge. At least one of the candidates then can deviate to his favorite policy and strictly improve his utility as λ has full support and $\kappa = 0$.

Non-emptiness is established by example: for $\sigma^{sq} \in (0, \alpha - d)$ the candidates share sq as an ideal policy. ■

Proof of Lemma 5: By definition the median voter strictly prefers x_t^* to any lottery over less favorable policies. Using again that the concavity of $u_i(p)$ in o_i is -2 in the proof of Lemma 1, this implies all voters at least to one side of the median also prefer x_t^* as the utility over the lottery is a convex combination of x_t and y_t and also has concavity -2. ■

Proof of Lemma 6: Part (i) follows from Lemma 4 and Proposition 2. Candidate X prefers to back-slide to policy τ_t^* iff:

$$(\psi(\tau_t^*) + d)^2 < \alpha^2 + \frac{[\psi(z_{t-1}^*)]^2 - (-d + \alpha)}{-\mu} \cdot \sigma^2$$

Substituting $\frac{\sigma^2}{-\mu} = 2\alpha$ and rearranging gives the required condition for part (ii). ■

Proof of Proposition 6: Follows from Lemmas 4-6. ■

Proof of Lemma 7: From the proof of Lemma 4 convergent stability can only be at the median voter's ideal policy and the requirements of Propositions 4 and 3 must be satisfied. As the stable policy must be a known point, the median's preferred policy is either τ_t^{+*} or τ_t^{-*} .

Lemma 4 also implies the stable policy is the ideal for both candidates. For candidate X to prefer τ_t^{+*} to τ_t^{-*} it is necessary that $d < \frac{|\psi(\tau_t^{-*})|}{2}$, and similarly $d < \frac{\psi(\tau_t^{+*})}{2}$ is necessary for candidate Y to prefer τ_t^{-*} to τ_t^{+*} . Equilibrium requires at least one of these inequalities holds.

Without loss of generality, suppose $\psi(\tau_t^{+*}) > |\psi(\tau_t^{-*})|$ and τ_t^{-*} is stable. By the previous condition, $d < \frac{\psi(\tau_t^{+*})}{2}$ and a policy $q \in (x_l^*, x_r^*)$ exists with $E\psi(q) = d$ and variance weakly less than $\frac{z_r^* - z_l^*}{4} \sigma^2$. If $d^2 \geq \frac{\sigma^2}{4} (z_r^* - z_l^*)$ candidate Y strictly prefers experimenting to τ_t^{-*} ; a contradiction. ■

Proof of Example 1: Without loss of generality, fix $x_t^* = z_1^*$ and consider Y 's policy choice y (and assume $z_1^* > sq$). Y deviates only to policies that improve the median

voter's utility, so consider $y \in [sq, z_1^*]$. The probability candidate Y wins election is:

$$f(x_t^*, y) = \frac{1}{2} + \frac{[\psi(x_t^*)]^2 - E[\psi(y)]^2 - \text{var}[\psi(y)]}{2\lambda}.$$

And Y 's utility is:

$$u^Y(x_t^*, y) = -f(x_t^*, y) ([d - E[\psi(y)]]^2 + \text{var}[\psi(y)]) - (1 - f(x_t^*, y)) (d + |\psi(x_t^*)|)^2.$$

Differentiating:

$$\begin{aligned} \frac{du^Y(x_t^*, y)}{dy} &= -([d - E[\psi(y)]]^2 + \text{var}[\psi(y)]) \frac{df(x_t^*, y)}{dy} \\ &\quad - f(x_t^*, y) \left(-2[d - E[\psi(y)]] \frac{dE[\psi(y)]}{dy} + \frac{d\text{var}[\psi(y)]}{dy} \right) + \frac{df(x_t^*, y)}{dy} (d + |\psi(x_t^*)|)^2. \end{aligned}$$

On this domain $[d - E[\psi(p)]] > 0$, $\frac{dE[\psi(p)]}{dp} < 0$, $\frac{d\text{var}[\psi(p)]}{dp} \geq 0$, and $-2[d - E[\psi(y)]] \frac{dE[\psi(y)]}{dy} + \frac{d\text{var}[\psi(y)]}{dy} > 0$ is bounded away from zero. As for λ sufficiently large $\frac{df(x_t^*, y)}{dy}$ is arbitrarily small, the sign of $\frac{du^Y(x_t^*, y)}{dy}$ is negative and $y = sq$ is optimal. ■

Proof of Proposition 7: To be stable, a policy τ_t^* must win more votes when pitted against any policy p . Voter i 's expected utility over the two policies is:

$$u_i(\tau_t^*) = -(o_i - \psi(\tau_t^*))^2, \quad \text{and} \quad u_i(p) = -(o_i - E\psi(p))^2 - \sigma_p^2.$$

For $v > \alpha^2$ and assuming the intervals of support don't overlap, the measure of support for each alternative is:

$$V(\tau_t^*) = \int_{\psi(\tau_t^*) - \sqrt{v}}^{\psi(\tau_t^*) + \sqrt{v}} f(x) dx, \quad \text{and} \quad V(p) = \int_{E\psi(p) - \sqrt{v - \sigma_p^2}}^{E\psi(p) + \sqrt{v - \sigma_p^2}} f(x) dx.$$

Suppose the election is tied at tolerance level v' . As f is single peaked, this implies that p 's interval of support is closer to the median citizen at 0: formally, without loss of generality, if $\psi(\tau_t^*) + \sqrt{v} < 0$ and $E\psi(p) > 0$, then $f(E\psi(p) - \sqrt{v - \sigma_p^2}) > f(\psi(\tau_t^*) + \sqrt{v})$ and $f(E\psi(p) + \sqrt{v - \sigma_p^2}) > f(\psi(\tau_t^*) - \sqrt{v})$ (and that $V(\tau_t^*)$ doesn't span zero).

Differentiating vote shares with respect to voter tolerance:

$$\begin{aligned} \frac{dV(\tau_t^*)}{dv} &= \frac{1}{2\sqrt{v}} [f(\psi(\tau_t^*) + \sqrt{v}) + f(\psi(\tau_t^*) - \sqrt{v})], \\ \frac{dV(p)}{dv} &= \frac{1}{2\sqrt{v - \sigma_p^2}} \left[f\left(E\psi(p) + \sqrt{v - \sigma_p^2}\right) + f\left(E\psi(p) - \sqrt{v - \sigma_p^2}\right) \right], \end{aligned}$$

which gives $\frac{dV(p)}{dv} > \frac{dV(\tau_t^*)}{dv}$. As for tolerance $v < \sigma_p^2$ policy p receives no votes whereas τ_t^* does (and wins), there exists a v' such that policy τ_t^* wins over p for all $v \leq v'$. The analysis proceeds analogously if the intervals of support overlap (focusing on one flank). ■

Proof of Corollary 5: A citizen that prefers experimenting to the sq has an ideal policy α from her ideal outcome. As her utility for any policy then satisfies $Eu_i(p) < -\alpha^2$, she abstains for $v < \alpha^2$. Turnout for sq is positive in a neighborhood of α^{sq} as $v > 0$. ■

References

- [1] Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Jullien. 1991. "Optimal Learning by Experimentation." *Review of Economic Studies* 58: 621-654.
- [2] Baron, David. 1996. "A Dynamic Theory of Collective Goods Programs." *American Political Science Review* 90 (June): 316-330.
- [3] Barro, Robert J. 1973. "The Control of Politicians: An Economic Model." *Public Choice* 14: 19-42.
- [4] Bartels, Larry M. 2008. *Unequal Democracy: The Political Economy of the New Gilded Age* (Princeton University Press: Princeton, NJ).
- [5] Bendor, Jonathan. 1995. "A Model of Muddling Through." *American Political Science Review* 89 (December): 819-840.
- [6] Black, D. 1958. *The Theory of Committees and Elections* (Cambridge University Press: Cambridge, U.K.).
- [7] Brainard, William. 1967. "Uncertainty and the Effectiveness of Policy." *American Economic Review Papers and Proceedings* 57: 411-425.
- [8] Callander, Steven. 2008. "A Theory of Policy Expertise." *Quarterly Journal of Political Science*, forthcoming.
- [9] Calvert, Randall L. 1985. "Robustness of the Multidimensional Voting Model: Candidates' Motivations, Uncertainty and Convergence." *American Journal of Political Science* 29: 1056-1070.
- [10] Clark, Gregory. 2007. *A Farewell to Alms: A Brief Economic History of the World* (Princeton University Press: Princeton, NJ).

- [11] Drazen, Allan, and Vittorio Grilli. 1993. "The Benefit of Crises for Economic Reform." *American Economic Review* 83 (June): 598-607.
- [12] Duggan, J. 2000. "Repeated Elections With Asymmetric Information." *Economics and Politics* 12 (July): 109-135.
- [13] Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50: 5-26.
- [14] Fernandez, Raquel, and Dani Rodrik. 1991. "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty." *American Economic Review* 81(5): 1146-55.
- [15] Friedman, Benjamin M. 1984. "Lessons From the 1979-82 Monetary Policy Experiment." *American Economic Review Papers and Proceedings* 74: 382-387.
- [16] Gilligan, Thomas, and Keith Krehbiel. 1987. "Collective Decision Making and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures." *Journal of Law, Economics, and Organization* 3 (2): 287-335.
- [17] Hansen, Lars Peter, and Thomas J. Sargent. 2007. *Robustness* (Princeton University Press: Princeton, NJ).
- [18] Kollman, Kenneth, John Miller, and Scott Page. 2000. "Decentralization and the Search for Policy Solutions." *Journal of Law, Economics, and Organization* 16 (April): 102-128.
- [19] Lindblom, Charles. 1959. "The Science of 'Muddling Through'." *Public Administration Review* 19: 79-88.
- [20] Majumdar, Sumon, and Sharun W. Mukand. 2004. "Policy Gambles." *American Economic Review* 94 (September): 1207-22.
- [21] Merton, Robert K. 1936. "The Unanticipated Consequences of Purposive Social Action." *American Sociological Review* 1 (December): 894-904.
- [22] Mitchell, Matthew F., and Andrea Moro. 2006. "Persistent Distortionary Policies with Asymmetric Information." *American Economic Review* 96 (March): 387-393.
- [23] Page, Scott E. 1996. "Two Measures of Difficulty." *Economic Theory* 8: 321-346.
- [24] Page, Scott E., and Natalia Zharinova. 2006. "Can Electoral Competition Outperform a Social Planner?" Working paper.
- [25] Piketty, Thomas. 1995. "Social Mobility and Redistributive Politics." *Quarterly Journal of Economics* CX (August): 551-584.

- [26] Rodrik, Dani. 1996. "Understanding Economic Policy Reform." *Journal of Economic Literature* 34 (March): 9-41.
- [27] Roemer, John E. 1994. "The Strategic Role of Party Ideology When Voters are Uncertain About How the Economy Works." *American Political Science Review* 88 (June): 327-335.
- [28] Rogerson, Richard, Robert Shimer, and Randall Wright. 2005. "Search-Theoretic Models of the Labor Market: A Survey." *Journal of Economic Literature* XLIII (December): 959-988.
- [29] Schultz, Christian. 1996. "Polarization and Inefficient Policies." *Review of Economic Studies* 63 (April): 331-343.
- [30] Wittman, Donald. 1983. "Candidate Motivation: A Synthesis of Alternative Theories." *American Political Science Review* 77: 142-157.