# ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

*An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics*

http://www.econometricsociety.org/

*Econometrica*, Vol. 78, No. 3 (May, 2010), 933–971


## LEARNING WHILE VOTING: DETERMINANTS OF COLLECTIVE EXPERIMENTATION


BRUNO STRULOVICI

*Northwestern University, Evanston, IL 60208-2600, U.S.A.*

*Econometrica*, Vol. 78, No. 3 (May, 2010), 933–971

# LEARNING WHILE VOTING: DETERMINANTS OF COLLECTIVE EXPERIMENTATION

BY BRUNO STRULOVICI[1]

This paper combines dynamic social choice and strategic experimentation to study the following question: How does a society, a committee, or, more generally, a group of individuals with potentially heterogeneous preferences, experiment with new opportunities? Each voter recognizes that, during experimentation, other voters also learn about their preferences. As a result, pivotal voters today are biased against experimentation because it reduces their likelihood of remaining pivotal. This phenomenon reduces equilibrium experimentation below the socially efficient level, and may even result in a negative option value of experimentation. However, one can restore efficiency by designing a voting rule that depends deterministically on time. Another main result is that even when payoffs of a reform are independently distributed across the population, good news about any individual's payoff increases other individuals' incentives to experiment with that reform, due to a positive voting externality.

KEYWORDS: Voting, experimentation, reform, social choice, learning, bandit.

## 1. INTRODUCTION

EVERY REFORM HAS CONSEQUENCES which cannot be fully known until it has been implemented. For example, the diverse effects of trade liberalization on a country's industrial sectors (e.g., which sectors will gain or lose from liberalization and when) cannot be easily or fully anticipated.[2] Similarly, although economic liberalization in the form of more business-friendly laws and fiscal policy can be expected to create opportunities for entrepreneurship and individual success, its specific beneficiaries are also unknown a priori. Moreover, those individuals who are *not* among its beneficiaries, but rather "losers" in its wake, may experience such hardships/penalties as reduced income redistribution or job security. Other contemporary examples of reforms whose benefits and costs are both uncertain and heterogeneous across the population are reforms in health care, national security policies, or environmental regulations.

[1]I am grateful to Meg Meyer, Paul Milgrom, Andrea Patacconi, and Kevin Roberts for insightful conversations, as well as to the editor and three anonymous referees, Daron Acemoglu, Patrick Bolton, Alessandra Casella, Eddie Dekel, Jeff Ely, Bard Harstad, Godfrey Keller, Marco Ottaviani, Andrea Prat, John Quah, Phil Reny, Marzena Rostek, Nicolas Vieille, Marek Weretka, and seminar participants at Oxford University, Gerzensee's ESSET, Cambridge University, the London School of Economics, the Kellogg School of Management, Princeton University, Northwestern University, the California Institute of Technology, the University of Pennsylvania, Columbia University, Stanford Graduate School of Business, University of Arizona, Penn State University, the University of Toronto, University of Texas at Austin, CESifo's Conference on Strategic Information Acquisition and Transmission, the University of Washington at St. Louis, the Winter Meeting of the Econometric Society 2009, the joint seminar of the École Polytechnique and HEC, the University of Tokyo and CIRJE, New York University, the University of Chicago, and UCLA for numerous comments.

[2]See Baldwin (1985), Bhagwati (1988), Fernandez and Rodrik (1990), and Rodrik (1993).

 DOI: 10.3982/ECTA8011

This paper studies incentives for collective experimentation when individual interests may be in conflict, are revealed gradually and at times which are random, and may also vary across individuals. It addresses the following questions: How do these incentives evolve as preferences and heterogeneity get revealed? How do they vary with group size? How do they compare to the utilitarian optimum? How do they vary with the voting rule chosen? Under which collective decision mechanisms do they result in efficient experimentation? How are they affected by the particular learning process?

The analysis is conducted in a two-armed bandit model in which a *safe* alternative yields a constant, homogeneous payoff to everyone, while a *risky* alternative yields payoffs whose unknown distribution, or type, may vary across individuals. At each instant, society elects one of the two alternatives according to some voting rule. Individuals learn their type only through experimentation with the risky alternative.[3] In the *benchmark setting* the risky action is, for each individual, either good or bad, and these types are independently distributed across individuals. Moreover, any news shock fully reveals to its recipient that the risky action is good for him, that is, he is a *sure winner*. By contrast, *unsure voters* are those individuals who have not yet received any positive news about their type and who become increasingly more pessimistic as experimentation goes on. The benchmark setting focuses on simple majority voting, with other voting rules considered in later sections. Payoffs are initially assumed to be publicly observed,[4] but Section 6 considers the case of privately observed payoffs. In the benchmark setting, learning occurs at the individual level only (Section 7 discusses the case of correlated types). The fact that an individual becomes successful, however, changes the unsure voters' expected payoffs, since it makes it more likely that the reform will not be overturned.

The first result is that incentives for experimentation are always weaker when power is shared, compared to the case of a single decision maker or to a dictatorship. Two kinds of risk shape incentives for collective experimentation, in addition to the well known trade-off between exploration and exploitation arising in individual experimentation. The *loser trap* occurs when reform winners have enough power to make the reform irreversible, in effect trapping reform losers into this new course of action. In contrast, *winner frustration* occurs when reform losers (more precisely, unsure voters with a low enough belief) are powerful enough to effect a return to the status quo, frustrating reform winners and whoever else still wishes to continue experimentation. Costly reforms and projects may thus be abandoned too early if they do not garner enough support, even when they turn out to be ex post efficient. These risks, which are specific to decisions taken collectively and where interests may turn out to be heterogeneous, reduce incentives for experimentation.

[3]Focusing on two actions gets rid of Condorcet cycles and ensures the robustness of the equilibrium concept used in the analysis. An example with three actions is studied in Section 7.

[4]Voters care only about the number of sure winners at any time, not about their identity.

To illustrate, consider a community of $N$ individuals with equal voting rights. Every month, these individuals must decide between a centralized production, where tasks are fixed and earnings are divided equally, and a decentralized one, where each individual chooses his task and keeps his earnings. There are two types in this community: individuals with talent and those without, where "talent" refers to an individual's ability to find a successful task. If the community tries decentralization, individuals gradually find out whether or not they are talented. As time elapses, two things can happen: A majority of talented people may form, in which case decentralization is imposed forever. Alternatively, if few talents are revealed under decentralization, voters who remain unsure can impose reversal to a centralized production.[5] In the first case, untalented people are trapped in a community that essentially abandons them economically. In the second case, talented people are frustrated by the collective constraint. If these risks are severe enough ex ante, the community may prefer not to experiment at all with decentralization, even if it is efficient to do so.

Loser trap and winner frustration have a systematic impact on welfare: *experimentation incentives are always too weak, compared to the utilitarian optimum.* This result stems from two effects. First, the utilitarian policy, which is the optimum of a single-decision-maker problem (faced by a utilitarian social planner), is not subject to the control-sharing effects described earlier. The value of the information acquired through experimentation is thus maximal, which makes experimentation more valuable from an efficiency perspective. Second, unsure voters only care about their own utility and thus sometimes impose the safe action even when including sure winners' utility would make the risky action more efficient.

This social inefficiency remains for any fixed voting rule. For example, if the risky action requires unanimity, the risk of loser trap disappears. However, this very fact also makes experimentation less attractive: winners are less likely to enjoy the risky action in the long run, for this would require that all society members turn out to be winners. Unanimity rule thus exacerbates winner frustration. Similarly, if the safe action requires unanimity, the risk of winner frustration disappears, but the risk of loser trap becomes maximal.

However, efficiency can be restored by a voting rule that depends deterministically on time. To implement the efficient policy, the number of votes required for the risky action increases deterministically over time, according to a schedule agreed upon at the outset. Intuitively, the more time elapses, the more numerous sure winners should be, if the reform is efficient. Therefore, one way to make sure that the reform continues only if it is efficient is to gradually raise the voting threshold required for it to be implemented. As the paper shows, the threshold can be set so as to exactly implement the utilitarian policy.

[5]For the sake of this example, we suppose that a centralized production is better for individuals who are sure of being untalented, and yields deterministic and homogeneous payoffs.

Another dynamic aspect of experimentation concerns the impact of incoming news on experimentation incentives. To return to the example given earlier, How do other voters react whenever someone discovers a talent? The answer is that, in equilibrium, *good news for anyone increases others' incentives to experiment*. Intuitively, individuals vote for experimentation because they hope to be winners and, hence, to enjoy the reform in the longer run. The appearance of a new winner makes it more likely that others too will be able to enjoy the reform and thus makes it more valuable to experiment.

As group size gets arbitrarily large, voters behave myopically, as if there were no value in experimentation. Indeed, individual control over future decisions becomes infinitely diluted, so one's ability to react to individual news vanishes in an arbitrarily large group. For small groups, however, *equilibrium incentives for experimentation do not monotonically decrease with respect to group size*. This is because the addition of new voters reduces the risk of winner frustration, a benefit that may locally dominate a higher risk of loser trap.

Several key results are extended to general experimentation environments, beyond the benchmark setting. Most importantly, the main result on weak incentives for experimentation holds even when individuals (i) never fully learn their types, (ii) receive both positive and negative news, and/or (iii) have correlated types. The analysis is based on a collective version of the Gittins index. However, it is not true any more that experimentation is inefficiently short. In particular, Section 5.1 shows that with negative shocks, experimentation may be inefficiently short or long depending on initial conditions. The paper also introduces a nonadversity condition on the collective decision process under which the value of experimentation is always positive. A collective decision rule is nonadverse to a given individual if, at any time, it is more likely to select the risky action if that individual is a winner than if he is a loser.

Surprisingly, however, even fair-looking decision rules, such as the simple majority rule, can violate the nonadversity condition. The value of experimentation may even be negative, in that society may reject a reform with a higher expected payoff than that of the status quo.[6] This means that the common intuition of a positive "option value," which captures a decision maker's ability to react to news (e.g., financial options, real options, options of waiting in endogenous bankruptcy models), may be inaccurate in settings with multiple decision makers. In contrast, the value of experimentation is always nonnegative when voters use the unanimity rule.

The paper contributes to a developing literature on experimentation with multiple agents, in which conservatism may arise as a consequence of strategic information acquisition. Bolton and Harris (1999), Li (2001), Décamps and Mariotti (2004), and Keller, Rady, and Cripps (2005) analyzed an informational free-riding problem in settings where agents can experiment individually with some risky action to learn about its common value. By contrast, the

[6]The result does not rely on commitment ability or asymmetric information, but is due solely to control-sharing effects, as shown here and in Section 5.2.

present paper considers a reverse setting, in which a single action taken collectively is made at any time, but the value of the action varies across individuals.[7] In these papers, experimentation is inefficiently low due to positive information spillovers that are not internalized by agents. In contrast, the control-sharing effects in the present paper reduce experimentation due to the negative payoff externalities that voters impose on one another, and which are decomposed and quantified in Equations (4) and (5). The analysis of the benchmark setting owes conceptual and technical clarity to the use of exponential bandits, building on Keller, Rady, and Cripps (2005).[8]

The paper is related to Fernandez and Rodrik (1991), who identified an asymmetry between ex ante and ex post support for reforms, which is due to uncertainty about winners' identity. Voters know that if the reform is implemented once, it will surely be implemented afterward. However, they do not know whether they are winners or losers under the reform, and hence bear the risk of loser trap. Their setting is similar to the case of immediate type revelation and an infinite population in the benchmark setting presented here. In the present paper, individuals learn at different points in time, and the efficient policy is unknown a priori. The evolution of informed agents introduces some of the interesting strategic issues that were absent from the model of Fernandez and Rodrik (1991).

The paper is organized as follows. Section 2 below analyzes the benchmark setting under the simple majority rule. Section 3 considers the utilitarian optimum and compares it to the majority-voting equilibrium. Section 4 takes a broader design approach to voting procedures, showing which rules can restore efficiency. Section 5 extends the analysis to more general type and learning structures, where types are never fully revealed, news can be positive and/or negative, and types may be correlated. Section 6 considers the case of publicly observed payoffs, showing that the majority voting equilibrium of Section 2 is truthful. Section 7 discusses several assumptions of the model, and Section 8 concludes.

[7]In this way, the paper is also related to a burgeoning literature analyzing collective search in various settings, where a group must choose, at any time, between accepting some outstanding proposal or trying a new proposal with independent and identically distributed (i.i.d.) characteristics. Compte and Jehiel (2008) showed, in particular, that more stringent majority requirements select more efficient proposals but take more time to do so. Albrecht, Anderson, and Vroman (2007) found that committees are more permissive than a single decision maker facing an otherwise identical search problem. Messner and Polborn (2008) discussed correlation across the two periods of their setting. In contrast to those papers, the present work focuses on social and individual learning and experimentation when voter types for a given action are lasting and permanently influence collective decisions. Callander (2009) also considered experimentation in a political setting. His focus is on the experimentation pattern of a single decision maker—the median voter—facing a continuum of correlated policies. Although the median voter is myopic and nonstrategic, the nature of uncertainty in that model produces interesting experimentation patterns.

[8]Exponential bandits have also been used in economics by Malueg and Tsutsui (1997), Bergemann and Hege (1998, 2001), and Décamps and Mariotti (2004).

## 2. BENCHMARK SETTING

The benchmark setting embeds the exponential bandit model analyzed by Keller, Rady, and Cripps (2005) into a setting with majority voting. Time $t \in [0, \infty)$ is continuous and payoffs are discounted at rate $r > 0$. There is an odd number $N \geq 1$ of individuals who continually decide according to the simple majority rule which of two actions to choose. The first action $S$ is "safe" and yields a flow $s$ per unit of time to all individuals. The second action $R$ is "risky" and can be, for each player, either "good" or "bad." The types (good and bad) are independently distributed across the group. (The case of correlated types is discussed in Section 7.)

If $R$ is bad for some individual $i$, it always pays him 0. If $R$ is good for $i$, it pays him lump-sum payoffs at random times which correspond to the jumping times of a Poisson process with constant intensity $\lambda$. The arrival of lump sums is independent across individuals. The magnitude of these lump sums[9] equals $h$. If $R$ is good for $i$, the expected payoff per unit of time is therefore $g = \lambda h$. The assumption $0 < s < g$ rules out the uninteresting case in which either $R$ or $S$ is dominated for all beliefs.

Each individual starts with a probability $p_0$ that $R$ is good for him. This probability is the same for all and is common knowledge. Thereafter, all payoffs are publicly observed, so that everyone shares the same belief about any given individual's type (for privately observed payoffs, see Section 7). In particular, the arrival of the first lump sum to a given individual $i$ makes him publicly a sure winner. At any time $t$, the group is therefore divided into $k$ "sure winners," for whom $R$ is good with probability 1, and $N - k$ "unsure voters," who have the same probability $p$ of having a good type. Unsure voters' probability evolves according to Bayes' rule and obeys the dynamic equation $dp/dt = -\lambda p(1 - p)$ if no lump sum is observed, with $p_j$ jumping to 1 when some voter $j$ receives a lump sum.[10] Type independence implies that an unsure voter learns only from his payoff stream, but not from those of others.

When $N = 1$, the setting reduces to the optimization problem of a single decision maker. The optimal experimentation strategy is Markov with respect to the current belief $p$, determined by a cutoff $p^{SD}$ such that $R$ is played if and

[9]All results hold if these lump sums have random, independently distributed magnitudes with constant mean $h$. More generally, what matters to decision makers are the expected payoff rates of each action and the probability that the risky action is good or bad. See Section 5 for a general specification of payoff distributions and beliefs.

[10]One way to derive this dynamic equation is to observe that $p_t$ is a martingale and jumps to 1 with probability rate $p\lambda$; hence, $0 = E[dp_t|p_t] = \eta_t\,dt + \lambda p_t(1 - p_t)\,dt$, where $\eta_t$ is the rate of decrease of $p_t$ conditional on not observing any lump sum, that is, $\eta_t = dp_t/dt$ if no lump sum is observed, and where the factor $(1 - p_t)$ in the second term is the change in probability in case a lump sum is observed. This yields the equation for $dp/dt$. One may alternatively use Bayes' rule to directly show that $p_t = (p_0 e^{-\lambda t})/(p_0 e^{-\lambda t} + (1 - p_0))$, which yields the equation by differentiation.

only if $p \geq p^{SD}$. This cutoff is determined by the indifference condition[11]

$$p^{SD} = \frac{\mu s}{\mu g + (g - s)}, \tag{1}$$

where $\mu = r/\lambda$. Let $p^M = s/g$ denote the myopic cutoff, that is, the probability below which $R$ yields a lower expected flow payoff than $S$. The previous formula implies that $p^{SD} < p^M$. Indeed, experimentation really takes place only for all $p \in [p^{SD}, p^M]$, since the single decision maker then chooses the risky action, despite its lower payoff, so as to learn more about its true value for future decisions. Choosing $R$ in this range is optimal due to the option value of experimentation.

For a group using the simple majority rule, the formal analysis to follow in this section shows that collective decisions are determined by nonincreasing cutoffs $\{p(k)\}_{0\leq k\leq N}$ such that the risky action is played at time $t$ if and only if $p_t > p(k_t)$, where $k_t$ is the number of sure winners at that time. The dynamics of collective decisions can thus be described as follows. Starting with some (high enough) level $p_0$, $R$ is elected until the threshold $p(0)$ is reached, at which point experimentation either stops if no winner has been observed by then or continues until another threshold $p(1) < p(0)$ is reached, and so forth. These dynamics are qualitatively represented by Figure 1 for the case of three voters. Here and throughout, experimentation means choosing (or voting for) the risky action when one's type is unknown. (Thus, only unsure voters are experimenting.) The option value of experimentation is formally defined after the equilibrium concept is introduced.

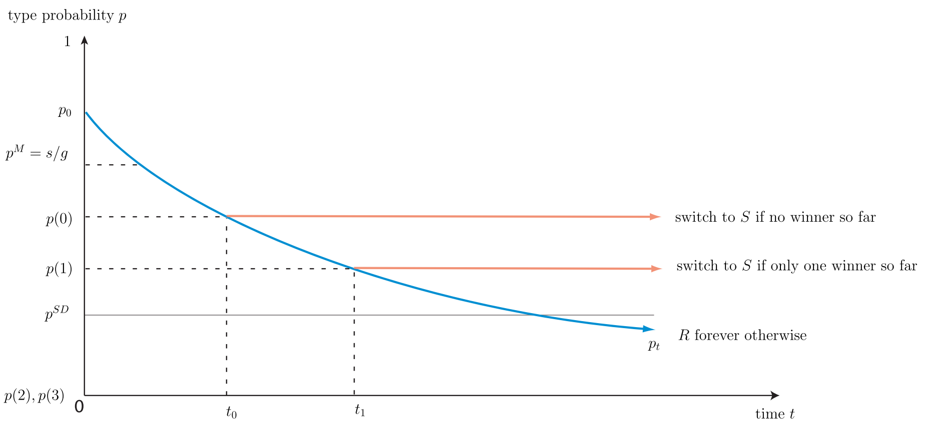
A collective decision rule, or policy, is a stochastic process $C = \{C_t\}_{t\geq 0}$ adapted to the filtration generated by the arrival of voters' lump sums and taking values in the action space $\{R, S\}$. Any collective decision rule determines a value function for each agent $i$:

$$V_t^{i,C} = E_t\left[\int_t^\infty e^{-r(\tau - t)}\, d\pi^i_{C_\tau}(\tau)\right],$$

where the payoff rate is $d\pi^i_S(\tau) = s\,d\tau$ and $d\pi^i_R(\tau) = h\,dZ^i_\tau$ or 0 depending on whether $R$ is good or bad for $i$, and $\{Z^i\}_{1\leq i\leq N}$ is a family of independent Poisson processes with intensity $\lambda$. At any given time, let $K$ denote the set of sure winners. The number $k$ of sure winners is thus the cardinal of $K$. A Markov strategy for voter $i$ is a function $d^i : (K, p) \mapsto \{R, S\}$.[12] For a

[11]Intuitively, if the decision maker experiments, his instantaneous payoff rate is $pg$, and with a probability rate of $\lambda p$, his value function jumps to $g/r$ from its current value. If he chooses $S$ instead, his immediate payoff rate is $s$. When the decision maker is indifferent, his current value is $s/r$, so the cutoff $p$ solves the indifference equation $pg + \lambda p(g/r - s/r) = s$, which is exactly (1). The result is derived formally and more generally in the proof of Theorem 1.

[12]We assume that, in the case of a jump at time $t$, the strategy depends only on the belief before the jump. Other assumptions would yield the same outcome, since they affect payoffs only over




$t_0$: experimentation end time if no winner is observed before reaching $p(0)$.

$t_1$: experimentation end time if only one winner is observed before reaching $p(1)$.

$p(2) = 0$: $R$ is elected forever if winners have the majority, no matter what $p_t$ for the remaining unsure voter.

$p^{SD} < p(1)$: a single decision maker always experiments more than a group with a majority of unsure voters.

FIGURE 1.—Dynamics of collective experimentation with three voters.

given profile $d = (d^1, \dots, d^N)$ of Markov strategies, let $C$ denote the resulting (Markov) collective decision rule, that is, such that $C(K, p) = R$ if and only if $|\{i : d^i(K, p) = R\}| > N/2$. $C$ is based on the simple majority rule. $V^{i,C}(K, p)$ denotes $i$'s value function under policy $C$ when the current state is $(K, p)$. Under any Markov rule, $C(K, p) = S$ implies that $S$ is played forever, since the state $(K, p)$ can evolve only when $R$ is played. Therefore, $V^{i,C}(K, p) = s/r$ for all $i$ whenever $C(K, p) = S$. This, among other things, rules out strategies of the grim-trigger type. To avoid trivial equilibria, the equilibrium concept used in this paper requires the elimination of weakly dominated strategies, iterated in the following sense.

DEFINITION 1: The profile $d$ is a Markov equilibrium in undominated strategies if for all $(K, p, i)$,

$$(2) \qquad d^i(K, p) = R \quad \Leftrightarrow \quad p^i g + \lambda p \sum_{j \notin K} \bigl(V^{i,C}(K \cup \{j\}, p) - V^{i,C}(K, p)\bigr) - \lambda p(1-p) \frac{\partial V^{i,C}}{\partial p}(K, p) > s,$$

a discrete time set, but do not affect information, since the probability that two jumps occur at exactly the same time is zero.

where $p^i = 1$ if $i \in K$ and $p^i = p$ if $i \notin K$.[13]

Thus, $i$ votes at each instant as if his vote were pivotal, realizing that voting for $S$ in any given state $(K, p)$ will result in a constant payoff flow $s$ forever. The left-hand side is $i$'s payoff from the risky action, including the impact of incoming lump sums and Bayesian updating on $i$'s immediate payoff and value function, as will be explained shortly. This equilibrium concept entails an iterated elimination of dominated strategies, where the iteration proceeds backward with respect to the state $(k, p)$. For example, the definition implies that if voter $i$, given other voters' strategies, prefers $S$ over $R$ at some state $(k', p')$, then $i$ will indeed vote for $S$ if that state is reached; it therefore also implies that, seen from any state $(k, p)$ with $k \leq k'$ and $p \geq p'$ from which the state $(k', p')$ may be reached, only strategies where $i$ chooses $S$ at $(k', p')$ should be considered. This concept is closely related to the elimination of conditionally dominated strategies as defined in Fudenberg and Tirole (1991), except that the present case corresponds to elimination of conditionally weakly dominated, rather than strictly dominated, strategies.[14]

The (option) value of experimentation of an unsure voter is the difference between his value function and the maximum payoff he could get if he had to decide today on one action played forever. Formally

$$X^{i,C}(K, p) = V^{i,C}(K, p) - \max\left\{\frac{s}{r}, \frac{p^i g}{r}\right\}. \tag{3}$$

This value is positive for a single decision maker, since choosing a fixed action forever is only one out of many policies over which the decision maker optimizes. In fact, this value is positive for a single decision maker for any news arrival process generated by the risky action. It is also positive under the majority rule in the present setting. However, Section 5.2 shows that when both positive and negative news shocks are allowed, $X^i_C$ can be negative, even under the simple majority rule.

Finally, we quantitatively define loser trap and winner frustration. Both risks depend on the probability of an action being imposed that is individually suboptimal, and on the magnitude of this prejudice. These probabilities depend on the particular collective policy being used. Therefore, loser trap and winner frustration depend not only on parameters $p_0$, $g$, $s$, $\lambda$, and $r$, but also on the voting rule (see also Section 4). Let $C$ denote any arbitrary policy and let $D$

[13] Since $p$ only decreases over time, here and throughout derivatives of value functions should be understood as left derivatives.

[14] The iterated elimination of dominated strategies, as described here, gets rid of the need to consider arbitrary, suboptimal strategies at future states. In the present setting, the concept yields the same solution as the majority-voting equilibrium of Roberts (1989).

denote the policy that $i$ chooses if he is a dictator. Then the expected loser trap under $C$ for individual $i$ is

$$L^i(C) = E\left[\int_0^\infty e^{-rt} 1_{(C_t=R\wedge D_t=S)} \cdot (d\pi_R^i(t) - s\,dt)\right], \tag{4}$$

where expectations, here and in the following equation, are taken with respect to player types and the arrival times of lump sums (and, therefore, independently of players' actions). Similarly, the expected winner frustration under $C$ for $i$ is

$$W^i(C) = E\left[\int_0^\infty e^{-rt} 1_{(C_t=S\wedge D_t=R)} \cdot (s\,dt - d\pi_R^i(t))\right]. \tag{5}$$

Thus, the expected loser trap is the expected relative loss that $i$ incurs from $R$ being imposed whenever he would have chosen $S$ had he had full control of the decision process. The difference between $i$'s value function under $C$ and $D$ is the sum of the two control-sharing effects.

Theorem 1 shows that there exists a unique Markov equilibrium in undominated strategies, and that this equilibrium is characterized by cutoffs. Equilibrium uniqueness comes from a backward induction argument on the number of winners. Here is some intuition for the proof. At any time $t$, the state of the group can be summarized by $k_t$ and $p_t$. Each of the two voter categories (i.e., sure winners or unsure voters) consists of individuals with currently perfectly aligned interests. If sure winners have the majority, they optimally impose $R$, since any policy involving $R$ is strictly better for them than having $S$ played forever. This determines the common value function of unsure voters when winners have the majority. Since an unsure voter can become a winner but a winner remains a winner forever, majority can only shift from unsure voters to winners. Proceeding by backward induction on the number of winners, one can show that unsure voters (or sure winners) always share a common voting strategy after the iterated elimination of weakly dominated ones.

Let $u(k, p)$ and $w(k, p)$ denote unsure voters' and sure winners' respective value function when the state is $(k, p)$. When there is a majority of unsure voters, decisions are dictated by their common interest unless and until they lose the majority. The goal is therefore to determine unsure voters' preferences. These preferences are determined by the following Hamilton–Jacobi–Bellman (HJB) equation, which is a simplified formulation of (2):

$$\begin{aligned} ru(k, p) = \max\Big\{ & pg + \lambda p[w(k+1, p) - u(k, p)] \\ & + \lambda p(N-k-1)[u(k+1, p) - u(k, p)] \\ & - \lambda p(1-p)\frac{\partial u}{\partial p}(k, p), s\Big\}. \end{aligned} \tag{6}$$

The first part of the maximand corresponds to action $R$, and the second corresponds to action $S$. The effect of $R$ on an unsure voter $i$ can be decomposed into four elements: (i) the expected payoff rate $pg$, (ii) the jump of the value function if $i$ receives a lump sum, which occurs at rate $\lambda$ with probability $p$—his value function jumps to $w$ and the number of winners increases by 1, (iii) the jump of $i$'s value function if another unsure voter receives a lump sum—$i$ is still an unsure voter, but the number of sure winners increases by 1, and (iv) the effect of Bayesian updating on the value function when no lump sum is observed. The independence of the Poisson processes governing individual payoffs implies that only one lump sum can be received during any infinitesimal period of time, so that no term involving two or more jumps appears in the HJB equation. In comparison, if $S$ is chosen, learning stops and $i$ simply receives payoff rate $s$.

Since unsure voters have identical value functions, they unanimously decide to stop experimentation if $p$ becomes too low, which occurs when the $R$ part of (6) equals $s$. At this level $p$, the smooth-pasting condition implies that the derivative term vanishes, since the value function is constant and equal to $s/r$ below that level (see, for example, Dixit (1993)). This determines the equilibrium policy's cutoffs as stated by Theorem 1, whose proof is in the Appendix. The theorem is proved for the simple majority rule, but the backward induction argument can also be applied to other voting rules.

Let $k_N = (N-1)/2$, where $k_N$ is the number of winners for which (i) sure winners are in the minority, but (ii) only one new winner is needed for the majority to change sides from unsure voters to sure winners.

THEOREM 1—Equilibrium Characterization: *There exists a unique Markov equilibrium in undominated strategies. This equilibrium is characterized by cutoffs $p(k)$ for $k \in \{0, \dots, N\}$, such that $R$ is chosen in state $(k, p)$ if and only if $p > p(k)$. Furthermore, for all $k \in \{0, \dots, k_N\}$, $p^M > p(k) > p^{SD}$,[15] $p(k)$ is decreasing in $k$ for $k \le k_N$, and $p(k) = 0$ for all $k > k_N$. The value functions $u$ and $w$ satisfy the following properties*:

- *$u(k, p)$ and $w(k, p)$ are nondecreasing in $p$.*
- *$w(k, p)$ is nondecreasing in $k$ for all $p$.*
- *$u(k+1, p) \ge u(k, p)$ for all $p$ and all $k < k_N$.*
- *$u(k_N+1, p) < u(k_N, p)$ for all $p$.*
- *$u(k, p) = pg/r$ and $w(k, p) = g/r$ for all $p$ and all $k > k_N$.*

Cutoffs are decreasing in $k$: the larger the number of winners, the more remaining unsure voters are willing to experiment. This result is perhaps surprising: why would unsure voters want to experiment more when the risk of losing their majority and having $R$ be imposed on them forever increases? The intuition is as follows. Suppose that $p$ is below the myopic cutoff $p^M$ but above

[15]The strict inequality $p(k) > p^{SD}$ holds only if $N > 1$.

$p(k)$ so that with $k$ current winners, unsure voters choose to experiment. By definition of $p^M$, unsure voters get a lower immediate expected payoff rate with $R$ than with $S$. Therefore, their only reason for experimenting is their hope of becoming winners. Now suppose by contradiction that $p(k+1) > p(k)$ and that $p$ lies in $(p_k, p_{k+1})$. Then, as soon as a new winner is observed, $k$ jumps to $k+1$, which implies that $S$ is imposed forever, since $p < p_{k+1}$. Therefore, the very reason why unsure voters wanted to experiment—namely, the hope of becoming sure winners—becomes moot: as soon as one of these unsure voters becomes a winner, he sees the safe action imposed on him forever, which prevents him from actually enjoying any benefit of being a winner.[16] Theorem 1 also states that $p(k) > p^{SD}$ for all $k \leq k_N$; that is, a single decision maker always experiments more than a group whose majority consists of unsure voters. The reason is the control-sharing effect mentioned in the Introduction: a single decision maker knows that if he turns out to be a winner, he will be able to enjoy the risky action, while if he turns out to be a loser, he can stop experimentation whenever he wants. In a group, even if a voter turns out to be a winner, he is not guaranteed that the risky action will be played forever, as a majority of unsure voters may block it. And if he turns out to be a loser, he may still have the risky action imposed on him forever if experimentation lasts long enough to reveal a majority of winners. This double risk of losing control prompts unsure voters to experiment less than any one of them would if he alone could dictate decisions in the future. In fact, a result stronger[17] than cutoff monotonicity obtains: when a new winner is revealed, the value function of both winners and unsure voters jumps upward, provided that $k < k_N$. For sure winners, this result is intuitive: a higher number of sure winners means a higher probability that a winning majority will be achieved. To be complete, this argument also requires that experimentation gets longer as the number of winners increases, which is guaranteed by cutoff monotonicity. More surprising is the fact that the revelation of a new winner results in an upward jump of the unsure voters' value function unless this new winner is the decisive voter who gives the majority to winners. The intuition here is that new winners reduce the risk of winner frustration, a risk that dominates as long as unsure voters keep control of the decision process. Another possible interpretation of this result is that the emergence of new winners increases the expected "pivotality" of unsure voters, as it reduces the imbalance between the majority and the minority. Finally, the utility of unsure voters jumps downward when winners gain the majority (i.e., $k$ jumps from $k_N$ to $k_N + 1$). This is true even if $p$ is large. This may seem surprising since, when $p$ is large, voters are happy to experiment and could appreciate a priori that the opportunity to experiment will not be

[16]That is, apart from receiving a lump sum at the time of the jump, but the possibility of that gain is already taken into account in the computation of the immediate expected payoff, which is still less than $s$ for $p < p^M$.

[17]This result is used to analyze the case of privately observed payoffs; see Theorem 10.

overturned. However, this opportunity would have been overturned only after $p$ became sufficiently low (below the myopic cutoff), and now that option is no longer available.

The simpler case where learning is extremely fast teaches something else. When types are immediately revealed as soon as $R$ is tried, a single decision maker is always willing to experiment.[18] However, this result does not extend to the case of collective experimentation, for even as the time cost of experimentation vanishes, the risk of loser trap remains. If that risk is severe enough, society may prefer to shun the opportunity of immediate type revelation and hence of making a perfectly informed decision (clearly what a utilitarian planner would choose!). Keeping other parameter values fixed, nonexperimentation will occur if the total number $N$ of individuals is large enough and the initial probability $p$ is low enough; experimentation cutoffs then stay bounded away from 0 as learning intensity $\lambda$ goes to infinity, provided that $N$ is large enough. The proof is a direct consequence of equation (17) in the Appendix.

COROLLARY 1—Immediate Type Revelation: *If* $N > 2g/s - 1$, *then*

$$\lim_{\lambda\to\infty} p(k_N) = \frac{(N+1)s/g - 2}{N-1} > 0.$$

*If* $N \leq 2g/s - 1$, *then*

$$\lim_{\lambda\to\infty} p(k_N) = 0.$$

Corollary 1 suggests that the total number $N$ of individuals has an important effect on experimentation. In fact, the next proposition states that with independent types, individuals behave myopically as group size becomes arbitrarily large, electing the risky action if and only if its expected payoff is higher than that of $S$. To state the result, let $p(k, N)$ denote the experimentation cutoff when there are $k$ winners and $N$ overall individuals.

PROPOSITION 1—Group Size: *$p(k_N, N)$ is nondecreasing in $N$. Moreover, for all $k$, $p(k, N) \to p^M$ as $N$ goes to infinity.*

PROOF: The first part of the proposition is an immediate consequence of (16) in the Appendix. For the second part, (16) also implies that $p(k_N, N) \to s/g = p^M$ as $N$ goes to infinity. Finally, Theorem 1 implies that $p(k_N, N) \leq p(k, N) \leq p^M$ for fixed $k$ and for all $N \geq 2k + 1$. Taking the limit as $N$ goes to infinity proves the result. *Q.E.D.*

[18]Mathematically, this result comes from the single-decision-maker cutoff equation (1): as the intensity $\lambda$ goes to infinity, $\mu$ goes to 0 and so does the cutoff $p^{SD}$.

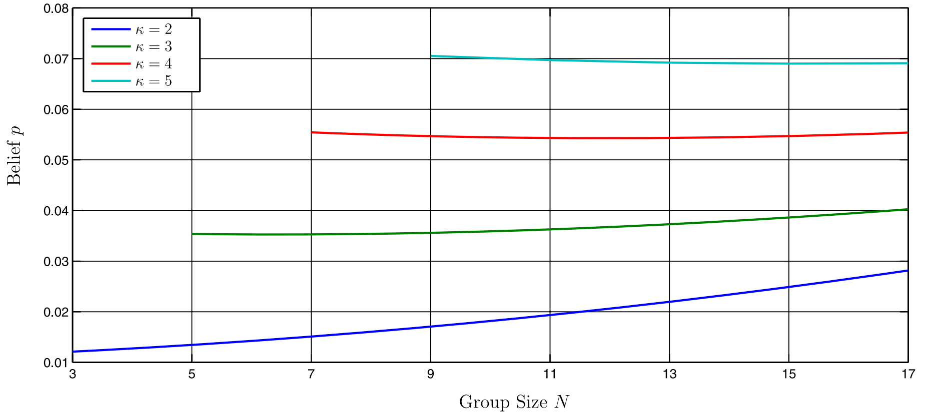



FIGURE 2.—Cutoffs as a function of group size $N$ and the switching number $\kappa$. Parameter values: $r = 1$, $\lambda = 10$, $s = 1$, $g = 10$. $N$ takes all odd values from 3 to 17. For $N = 1$, $p^{SD} = 0.01$.

In general, cutoffs $p(k, N)$ are not monotonic with respect to group size $N$, as can be proved by numerical counterexample. Such violations may seem counterintuitive: As $N$ increases, individual power gets more diluted. Should this not reduce the value of experimentation? However, adding unsure voters increases the expected number of winners, and thus the expected duration of experimentation, for given cutoffs. The addition of voters thus reduces the risk of winner frustration, which sometimes increases the attractiveness of experimentation.[19]

Figure 2 shows the numerical computation of cutoff policies for different values of $N$ and of the switching number $\kappa = k_N + 1 - k$ of switches required for winners to gain the majority. For $\kappa = 4$, the cutoff is not monotonic in $N$. For $\kappa = 5$, the cutoff is actually decreasing in $N$ when $N$ is in the range [9, 17].

## 3. UTILITARIAN POLICY

This section characterizes the optimal experimentation policy of a utilitarian social planner and shows that it lasts longer than majority-based experimentation. A social planner faces a single-decision-maker experimentation problem, the solution of which can be computed by backward induction on the number of observed winners and is characterized by monotonic cutoffs.

[19]The expected length of experimentation is also not generally monotonic in $N$. To see this, it is easy to build an example where the risky action is played forever when $N$ is arbitrarily large (see Theorem 6 and the discussion below it), whereas experimentation stops with positive probability when $N$ is small. Similarly, it is also easy to build an example where experimentation stops immediately when $N$ is large, but has positive duration when $N$ is small.

THEOREM 2: *Under the utilitarian criterion, the optimal policy is determined by cutoffs* $q(k)$ *such that* $C(k, p) = R$ *if and only if* $p \geq q(k)$. *These cutoffs are decreasing in* $k$ *for* $k < \bar{k}$ *and equal to zero for* $k \geq \bar{k}$, *where* $\bar{k} = \frac{s}{g}N$.

See the Appendix for the proof.

The next result shows that the equilibrium level of experimentation under the majoritarian rule is inefficiently short compared to the utilitarian optimum. This result is due to two concurring reasons. First, the social planner can exploit information to meet its objective better than individual voters, since he has full control over future decisions. That argument shows up in the proof below (see (9)). Second, the social planner takes into account winners' utility, while unsure voters do not. This implies that, under the majoritarian equilibrium, when unsure voters decide to stop, a social planner would take into account winners' utility which, other things equal, makes experimentation more attractive (see (10) in the proof below).

THEOREM 3—Majoritarian versus Utilitarian Rules: *Let* $\{p(k)\}_k$ *be the cutoff values associated with the majority rule. Then* $q(k) \leq p(k)$ *for all* $k \leq k_N$.

PROOF: The utilitarian cutoff $q(k)$ solves

$$\left(\frac{k}{N}\right)g + \left(1 - \frac{k}{N}\right)pg + (N-k)\lambda p\left[\frac{W(k+1, p)}{N} - \frac{s}{r}\right] = s, \tag{7}$$

where $W$ is the utilitarian value function. The left-hand side is the sum of the per-capita immediate expected payoff given state $(k, N)$ and of the per-capita jump of welfare following the observation of a new winner, weighted by the probability rate of this event. The majority-voting cutoff, $p(k)$, solves

$$pg + (N-k)\lambda p\left[\frac{\bar{w}(k+1, p)}{N-k} + \frac{N-k-1}{N-k}\bar{u}(k+1, p) - \frac{s}{r}\right] = s, \tag{8}$$

where $\bar{w}$ and $\bar{u}$ are the value functions obtained under the majority rule. (The left-hand side is obtained from (6) simplified through value-matching and smooth-pasting conditions.) Optimality of the utilitarian policy implies that for all $k$, $p$,

$$\frac{W(k, p)}{N} \geq \frac{k}{N}\bar{w}(k, p) + \left(1 - \frac{k}{N}\right)\bar{u}(k, p). \tag{9}$$

Since $\bar{w} > \bar{u}$, this also implies that

$$\frac{W(k+1, p)}{N} > \frac{1}{N-k}\bar{w}(k+1, p) + \left(1 - \frac{1}{N-k}\right)\bar{u}(k+1, p) \tag{10}$$

and, hence, that the left-hand side of (7) is higher than that of (8), for each $p$. Therefore, the root of the first equation must be lower than that of the second. *Q.E.D.*

## 4. DESIGN OF VOTING PROCEDURES

This section considers which mechanism can improve efficiency compared to the simple majority rule.

### 4.1. *Fixed Quorum*

The first issue is to determine how changing the number of votes (hereafter, quorum) required for the risky action affects the length and efficiency of experimentation. The simpler case of a constant quorum is considered first.[20] In that case, Theorem 4 shows that there is no systematic advantage of one voting rule over another. As one moves across the entire spectrum of voting rules, from requiring unanimity for the safe action to requiring unanimity for the risky action, the risk of loser trap diminishes while the risk of winner frustration increases, with exactly one of the two risks entirely vanishing at the ends of the spectrum. Depending on the parameters of the model, which determine the magnitude of these risks, the optimal rule can be any rule in the spectrum. For simplicity, the analysis starts with the case of immediate type revelation, which is sufficient to show the lack of comparability of voting rules.

Suppose that learning is arbitrarily fast (i.e., $\lambda \to \infty$). In that case, there is no time cost of experimentation, hence no winner frustration. If one requires unanimity for the risky action, this also gets rid of loser trap, so it will always prompt society to choose immediate type revelation. However, once types are revealed, unanimity requires that $R$ is only implemented if all voters are winners, which typically is inefficiently too restrictive. Indeed, the utilitarian optimum is to get immediate type revelation and then choose the risky action if and only if $kg > sN$. For $\nu \in \{1, \dots, N\}$, define the $\nu$ voting rule as the rule requiring $\nu$ votes for the risky action. Letting $\nu^U = (sN)/g$, a $\nu$ rule with $\nu > \nu^U$ will never implement the risky action when it is socially inefficient to do so. Let $\bar{\nu}$ denote the smallest integer such that society is ready to experiment with the $\bar{\nu}$ voting rule and let $\nu^* = \max\{\bar{\nu}, \nu^U\}$. Then, social efficiency is decreasing in $\nu$ for $\nu \geq \nu^*$, because in this range $\nu$ is high enough to prompt experimentation and the probability of implementing the risky action if it is socially efficient ex post is decreasing in $\nu$, while the probability of implementing the risky action if it is inefficient is zero. As is easily checked, $\nu^*$ can take any value between 1 and $N$ ($\bar{\nu}$ decreases from $N$ to 1 as $p$ increases from 0 to 1).

[20]For any $q$ rule one may, proceeding as in Section 2, prove the existence of a unique equilibrium characterized by monotonic cutoffs contained in $[p^{SD}, p^M]$. The analysis of this section, based on immediate type revelation, does not require this proof.

To generate the reverse inefficiency ranking, suppose that, in addition to immediate type revelation, $p$ is arbitrarily close to 1. In that case, society always wishes to experiment, since the probability of loser trap is arbitrarily small. Social efficiency is increasing in $\nu$ for $\nu \leq \nu^U$: since $p$ is close to 1, initial experimentation takes place anyway, and ex post the probability of implementing the risky action if it is socially inefficient decreases in $\nu$. Since $\nu^U$ can take any value between 1 and $N$, this implies the following result.

THEOREM 4: *For any voting rules $\nu \neq \tilde{\nu}$, there exist parameter values and an initial belief $p$ such that the $\nu$ voting rule is strictly socially more efficient than the $\tilde{\nu}$ voting rule.*

It would seem that as the risk of loser trap becomes more salient compared to winner frustration, it becomes more efficient to have a more restrictive voting rule, that is, a higher $\nu$. However, this intuition may sometimes be wrong. For example, suppose that types are immediately revealed if $R$ is played. Owing to the risk of loser trap, there must exist a level $\nu^*$, depending on initial belief $p_0$, such that society experiments for all $\nu \geq \nu^*$ and sticks to $S$ for $\nu < \nu^*$. Now suppose that $s$ is decreased by a small amount $\varepsilon$ so that $\nu^*$ stays unchanged. This increase of $s$ can be interpreted as the risk of loser trap becoming marginally less salient than the risk of winner frustration. The change reduces voters' value function for $\nu < \nu^*$ by $\varepsilon/r$, since they still choose $S$, but it reduces their value function for $\nu > \nu^*$ by a lower amount, since under experimentation the discounted time spent playing $s$ is strictly less than $1/r$.[21] This shows that, at least in some cases, reducing the risk of loser trap relative to winner frustration does not make less restrictive rules more desirable.

Efficiency depends not only on voters' ex ante probability of falling in the loser trap but also on the magnitude of the loser trap (more generally, the relative values of $g$ and $s$ and 0). With slower learning, the risk and magnitude of winner frustration also influences voting rule efficiency in the opposite direction. The impact of magnitude, already implicit in the above analysis through $\nu^U$, is illustrated below for the comparison of the simple majority rule and the unanimity rule for $R$ (i.e., $\nu = N$). Let $\{\chi(k)\}_{0\leq k\leq N}$ denote the cutoffs characterizing to the unanimity-voting policy.

EXAMPLE 1: *Suppose that $N = 3$ and $s \ll g$. Then $\chi(1) > p(1)$.*

PROOF: Equation (16) in the Appendix implies that

$$(11) \qquad p(1) = \frac{\mu s}{\mu g + (g - s) - (s - pg)} \sim \frac{\mu s}{(\mu + 1)g}$$

[21]Cutoffs are also affected by this change, but this change is of second order by optimality.

if $g \gg s$. In particular, $p(1)$ is arbitrarily close to zero if $g \gg s$. With the unanimity rule and $k = 1$, unsure voters are indifferent when $p$ satisfies

$$pg + \lambda p[w(2, p) - s/r] + \lambda p[v^{SD}(p) - s/r] = s, \tag{12}$$

where $w(2, p)$ is the value of a sure winner under unanimity rule if there are two sure winners (and $N = 3$), and $v^{SD}(p)$ is the value function of a single decision maker. As can be easily checked, $v^{SD}(p) \leq pg/r + (1 - p)s/r$, while $w(2, p) \leq pg/r + (1 - p)s/r$. This and (12) imply that $\chi(1)$ must satisfy the inequality

$$pg + 2\lambda p^2(g/r - s/r) \geq s$$

or

$$p \geq \frac{\mu s}{\mu g + 2p(g - s)} \sim \frac{s}{g} \tag{13}$$

if $g \gg s$. Comparing (11) and (13) shows that $\chi(1) > p(1)$. *Q.E.D.*

### 4.2. *Time-Varying Quorum*

Suppose now that at each time $t$, $R$ is elected if and only if it gets $\nu_t$ of the votes. The next result shows that even if $\nu_t$ is deterministic, efficiency can be fully restored.

THEOREM 5—Deterministic Quorum: *There exists a quorum function $t \mapsto \nu_t$ such that the resulting unique Markov equilibrium in undominated strategies implements the utilitarian policy. Moreover, $\nu_t$ is increasing in time and is entirely determined by the initial belief $p_0$ and the utilitarian cutoffs $q(k)$.*

PROOF: Starting from a given belief $p_0$, let $t_k$ denote the time such that $p_{t_k} = q(k)$ for each $k \in \{0, \dots, N\}$. Since $p_t$ decreases over time, the sequence $t_k$ is increasing, and $t_N = \infty$. For $t \in [t_k, t_{k+1})$, let $\nu_t = k$. This entirely and uniquely determines the function $\nu_t$. By construction, if there are at least $k$ winners, they can impose $R$ whenever $t < t_k$, that is, whenever $p > q(k)$. Moreover, if there are exactly $k$ winners, they can only impose $R$ if $t < t_k$, that is, when $p > q(k)$. From Theorem 3,[22] unsure voters always want to impose $S$ when it is socially efficient to do so, which guarantees implementation of the efficient policy whenever $t \geq t_k$ and there are only $k$ sure winners or less. Proceeding by backward induction on $k$, as in the proof of Theorem 1, one

[22]The proof is actually slightly different and simpler in the present case. Unsure voters choose between continuing with the efficient policy versus imposing $S$ forever.

may therefore conclude that $t \mapsto \nu_t$ yields a unique Markov equilibrium in undominated strategies and that this equilibrium implements the utilitarian policy. *Q.E.D.*

The quorum is not unique. For example, near time zero, everyone wants to experiment (assuming $p_0$ is high enough), so any quorum initially yields the efficient policy. In general however, the quorum must be low enough to allow $R$ whenever unsure voters want to stop experimentation while it is efficient to pursue it, and must be high enough to prevent winners from imposing $R$ whenever $S$ is the socially efficient action. More generally, Theorem 5 suggests that in settings where news events amount to good news and no news is bad news, an efficient quorum should increase over time: as more time elapses, society should require a higher number of winners for $R$ to be elected. Although the precise times of quorum change rely on the exact map between $p_t$ and the utilitarian cutoffs, and on $p_0$, the insight that a gradual increase in cut-offs may improve efficiency over constant quorum does not depend on that map.

Theorem 5 has several corollaries and equivalent formulations. For example, increasing voting weights of unsure voters, relative to sure winners, offers an alternative way of implementing efficiency. Indeed, it is easy to determine a particular weighting scheme that mirrors the proof of Theorem 5, under which sure winners are in control of the collective decision process only if $R$ needs to be implemented. Naturally, history-dependent quorums, which contain deterministic ones as a particular case, can also be devised to implement the efficient outcome. Such quorums—as well as constant ones—amount to a form of commitment, as opposed to having the stronger side impose his choice at any given time. In a setting where winners achieve increasing political power (for example, their higher payoffs might give them higher lobbying power), the resulting variation in voting weights goes opposite to the one implementing efficiency. In that case, winners may impose the risky action, owing to their higher political power, when it is inefficient to do so. Anticipation of this potential outcome should result in even more conservatism ex ante. The next section considers efficiency under various forms of commitment assumptions.

### 4.3. *Commitment and Endogenous Quorum*

Theorem 5 shows that it is possible to implement the efficient outcome as long as one can commit to some time-varying quorum. If voters are initially homogeneous and can only commit to an anonymous policy at the outset, they share, initially, a common objective function.[23] Since expected payoffs are identical, the optimal policy also maximizes the sum of these expected payoffs, that

[23] Anonymity means that individuals cannot commit to a policy that favors or harms particular voters, such as imposing generous redistribution if some given individuals turn out to be poor and no redistribution if these same individuals turn out to be rich. This assumption is consistent with veil-of-ignorance arguments.

is, utilitarian welfare. Therefore, if given the possibility, voters would like to commit to the utilitarian policy, showing the following result.

THEOREM 6—Commitment: *If voters can commit to an anonymous policy at time* 0, *they choose the cutoff policy determined by cutoffs* $\{q(k)\}_{0\leq k\leq N}$.

If voters can only commit to a fixed action, such as imposing a new rule for the next five years no matter how well that rule performs over that period, efficiency need not be restored. To give an extreme illustration, suppose that voters must commit to an action for the entire time horizon. In that case, the risky action is chosen if and only if its expected payoff is above the myopic cutoff. This extreme case of action commitment thus entirely annihilates the value of experimentation. Commitment to an action is formally equivalent to reducing the frequency of decision making. For example, voting every five years amounts to a succession of five-year commitments. The previous observation can therefore be reinterpreted as follows: if votes take place at a low enough time frequency, individual control over collective decisions is reduced to such an extent that the resulting policy may be more inefficient. However, provided that aggregate uncertainty is small enough and initial beliefs are optimistic enough, commitment to a fixed action can restore efficiency.[24]

One may wonder whether a hybrid form of commitment, where voters commit to a fixed action over some horizon but also get to dynamically modify the quorum required for the risky action, can restore efficiency. The following result, which extends to a much more general setting than the present one, answers negatively. Consider the following recursive procedure, which may be called the endogenous quorum procedure. At time 0, voters decide on two policy dimensions: an initial horizon of experimentation, $T_1$, and a quorum, $\nu_1$, used at $T_1$ for deciding whether or not to continue experimenting and, if so, on a new horizon of experimentation, $T_2$, and on the quorum $\nu_2$ used at $T_2$ to vote on $T_3$ and $\nu_3$, and so forth. The key of this procedure is that voters cannot commit to the quorum function at the outset. Rather, they can, at any given election, mitigate the law of the strongest by controlling the quorum used at

[24]With an infinite population, the law of large numbers allows one to compute the socially optimal action: starting with an individual probability $p$ that the action is good, the risky action is the social optimum if and only if $pg > s$, since there surely is a fraction $p$ of winners. Suppose that, initially, $pg > s$. From Theorem 6, individuals find it optimal to commit to the risky action over the infinite horizon. What happens without commitment? The second part of Proposition 1 implies that unsure voters, if they have the majority, impose the safe action as soon as $p_t$ hits the myopic cutoff $p^M = s/g$. This situation will occur surely if one starts with $p = p^M + \varepsilon$ for $\varepsilon$ small enough. Indeed, from Proposition 2 in the Appendix, the probability that an unsure voter with initial probability $p$ receives a lump sum before $p_t$ reaches $q < p$ equals $(p-q)/(1-q)$. This and the law of large numbers imply that when society starts at $p^M + \varepsilon$, the fraction of remaining unsure voters when $p^M$ is reached equals $1 - \varepsilon/(1 - p^M)$, which is greater than $1/2$ for $\varepsilon < (g - s)/2g$. This shows that commitment to the risky action is strictly more efficient than no commitment.

the next election. This procedure is a natural way to capture the idea that voting rules cannot be committed to ex ante. As soon as sure winners meet the quorum, they impose the risky action forever by setting the next experimentation period to infinity. If sure winners do not meet the quorum, the procedure grants unsure voters the right to choose the next horizon and quorum. To be complete, the procedure must specify an initial quorum, $\nu_0$, at time 0. Since voters are ex ante homogeneous, this initial quorum plays no role in the present setting. In principle there may exist multiple equilibria. However, the endogenous quorum procedure must be inefficient for any equilibrium, as the next result shows.

THEOREM 7—Endogenous Quorum: *There exist initial belief $p_0$ and group size $N$ such that the endogenous quorum procedure does not implement the utilitarian policy*.

PROOF: Consider any policy $C$ consistent with an endogenous quorum procedure: $C$ is a right continuous stochastic process taking values in $\{S, R\}$ characterized by an increasing sequence of (possibly infinite) random times $T_j$ such that $C$ is constant on any interval $[T_j, T_{j+1}]$, and by a quorum process $\nu_j$ such that at each $T_j$, unsure voters impose $S$ forever or set $\nu_{j+1}$ and $T_{j+1}$ optimally if there are at least $\nu_j$ of them; otherwise, sure winners set $T_{j+1} = \infty$. With positive probability there exist both sure winners and unsure voters at time 1. Suppose that sure winners do not meet the quorum $\nu_1$, so that unsure voters can choose the next quorum and experimentation horizon. If Theorem 7 were false, then these unsure voters, starting from their current belief $p_1$ and group size $N - k_1$, could choose $\nu_1$ and $T_1$ as part of a policy that maximizes their expected welfare, that is, ignoring the utility of current sure winners. Since it would be optimal for them to do so, they implement that policy, contradicting efficiency of $C$. *Q.E.D.*

## 5. GENERAL NEWS ARRIVAL PROCESSES

Section 2 assumed that individuals perfectly learned their types upon receiving some lump sum, that news events amounted to good news, and that types were independently distributed. Relaxing these assumptions, this section reconsiders in a very general setting whether (i) collective experimentation is shorter than the single-decision-maker equivalent, (ii) collective experimentation is shorter than the utilitarian optimum,[25] and (iii) there is always some experimentation, that is, a set of voter beliefs where $R$'s immediate payoffs is lower than $S$'s but society still elects $R$.

Suppose that, for any given individual, the risky arm has a payoff distribution or type $\theta$ lying in some finite set $\Theta$. At any time, that individual's belief about

[25]That result is considered in the negative-news setting to follow.

his type is summarized by a probability distribution or state $\gamma \in \Gamma$, where $\Gamma = \Delta(\Theta)$ is the set of all probability distributions[26] over $\Theta$. The safe arm still pays a constant rate $s$. For a single decision maker, the Gittins index of the risky arm is the map $G : \Gamma \to \mathbb{R}$ such that, given state $\gamma$, $G(\gamma)$ is the smallest value of $s$ for which the single decision maker prefers the safe action over experimentation. Mathematically, $G(\gamma)$ solves

$$G(\gamma) = \inf\left\{s : \frac{s}{r} = \sup_{\sigma} E\left[\int_0^{\infty} e^{-rt}\, d\pi_{\sigma_t}(t) \middle| \gamma, s\right]\right\},$$

where $\sigma$ is any policy, and the expectation is conditional on the current state $\gamma$ and on the rate $s$ of the safe action.[27]

Now consider the case of $N$ decision makers and let $\{\mathcal{F}_t\}_{t\geq 0}$ denote the filtration generated by all voters' payoffs. At any time, the state, known to all, is denoted $\gamma$. If types are independent, then $\gamma = (\gamma^1, \dots, \gamma^N) \in \Gamma^N$. In general, $\gamma$ may contain information about type correlation. A policy is a process adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and taking values in $\{S, R\}$.

For any rate $s$, policy $C$, and voter $i$, necessarily

$$\sup_{\sigma} E\left[\int_0^{\infty} e^{-rt}\, d\pi^i_{\sigma_t}(t) \middle| \gamma, s\right] \geq E\left[\int_0^{\infty} e^{-rt}\, d\pi^i_{C_t}(t) \middle| \gamma, s\right]. \tag{14}$$

The inequality obtains because $C$ is an element of the policy set over which the maximization is taken.[28] We may define a policy-dependent generalization of the Gittins index as

$$G^i_C(\gamma) = \inf\left\{s : \frac{s}{r} = E\left[\int_0^{\infty} e^{-rt}\, d\pi^i_{C_t}(t) \middle| \gamma, s\right]\right\}.$$

Inequality (14) implies that $G^i_D(\gamma) \geq G^i_C(\gamma)$ for all $i$, $\gamma$, and $C$, where $G^i_D(\gamma)$ is $i$'s Gittins index if he has dictatorial power over all decisions.

The definition of Markov equilibrium in undominated strategies is extended as follows. Let $\nu$ denote any integer in $\{1, \dots, N\}$.

DEFINITION 2—Voting Equilibrium: $C$ is a $\nu$ voting equilibrium if, for any belief $\gamma$,

$$C(\gamma) = S \quad \Leftrightarrow \quad \left|\{i : G^i_C(\gamma) \leq s\}\right| \geq \nu.$$

[26]In the benchmark model, the type $\theta$ is either "good" or "bad," and the state $\gamma$ is the probability $p$ that the type is good.

[27]The results of this section are easily adapted to discrete-time settings. In fact, Theorem 8 does not assume anything about the time domain.

[28]In general, $C$ depends on all voters' types and need not be anonymous.

This definition should be interpreted as follows. If a voter is pivotal, the Markov property implies that imposing $S$ at any time amounts to imposing $S$ forever, since the state $\gamma$ is frozen whenever $S$ is played. Therefore, $i$ votes for $S$ if and only if he prefers getting the constant payoff $s$ forever over pursuing policy $C$, a choice that is determined by $i$'s $C$ Gittins index at belief $\gamma$. The following result shows that collective experimentation is shorter than dictatorial experimentation in the following sense: if there are at least $\nu$ individuals who, taken individually, would prefer the safe action if given dictatorial power over future decisions, then society also picks the safe action in any $\nu$ voting equilibrium. This result is an extreme generalization of the fact that all equilibrium cutoffs in Section 2 were above the single-decision-maker cutoff.

THEOREM 8: *Suppose that $C$ is a $\nu$ voting equilibrium. Then $C = S$ whenever $|\{i : G^i_D(\gamma) \leq s\}| \geq \nu$.*

The proof is an immediate consequence of the general inequality $G^i_D(\gamma) \geq G^i_C(\gamma)$ for all $i$ and, $C$ and $\gamma$.

When types are independent, $G^i_D(\gamma) = G(\gamma^i)$, where $G(\gamma^i)$ is the Gittins index of the single-decision-maker problem with (individual) belief $\gamma^i$. In that case, $i$'s optimal policy is independent of other individuals' types. As a corollary of Theorem 8, therefore, collective experimentation is shorter than in an equivalent single-decision-maker setting. If types are positively correlated, however, collective experimentation can last longer than in a single-decision-maker setting, as positive type correlation increases learning speed and thus reduces the time cost of experimentation (see also Section 7). In contrast, collective experimentation is always shorter, even with positive correlation, than what any voter would like if he could dictate all decisions, because a dictator benefits from the same learning speed as society, unlike a single decision maker.

Theorem 1 also stated that all cutoffs were below the myopic cutoff, meaning that there always was some experimentation. How general is this result? Are there cases where society elects the safe action even when the risky action yields a higher payoff? To answer this question, the following definitions will be used. For any probability distribution $\gamma^i$ over the type space, let $g(\gamma^i) = E[d\pi^i_R/dt|\gamma^i]$. $g(\gamma^i)$ is $i$'s immediate expected payoff rate with action $R$ given type distribution $\gamma^i$. For any individual type $\theta^i$, slightly abusing notation, let $g(\theta^i) = g(\delta_{\theta^i})$, where $\delta_{\theta^i}$ is the Dirac distribution concentrated on type $\theta^i$, denote $i$'s true immediate expected payoff rate with action $R$ when his actual type is $\theta^i$. Say that $i$ is a winner if $g(\theta^i) > s$ and a loser otherwise. Hence, $i$ is a winner if $R$ is optimal for him given his true type. $\Theta$ can thus be partitioned into good (winner) types and bad (loser) types.

DEFINITION 3: A policy $C$ is adverse for voter $i$ if the set

$$\{t : \Pr[C_t = R|\theta^i \text{ good}] < \Pr[C_t = R|\theta^i \text{ bad}]\}$$

has positive Lebesgue measure.

Adversity means that $R$ is more likely to be chosen if $i$ is a loser, at least for some nonzero time set. Adversity can occur, for example, if a voter's type is perfectly negatively correlated with a majority of voters. The majority then blocks $R$ whenever that voter is a winner and imposes it when he is a loser.[29]

THEOREM 9: *Suppose that $C$ is a voting equilibrium for voting rule $v$. Then $G^i_C(\gamma) \geq g(\gamma^i)$ for all $i$ for which $C$ is nonadverse.*

See the Appendix for the proof.

It would seem a priori that in settings where types are independent or positively correlated, usual voting rules would be nonadverse. However, this intuition is incorrect, as explained in Section 5.2.

### 5.1. *Negative-News Shocks*

Several potential applications concern a setting, which is symmetric to the benchmark, where news events amount to catastrophes and no news is good news. One models such applications by assuming that the risky arm pays a positive constant rate if it is good and, in addition, pays some negative lump sums according to some Poisson process if it is bad. One may assume without loss of generality that the payoff rate of $S$ is zero, since all payoffs can be translated by the same constant without affecting voters' decision problem. The state variables are the number $k$ of sure losers and the probability $p$ that the arm is good for unsure voters. It may be shown that the policy is also determined by cutoffs $\rho(k)$ such that unsure voters impose the risky action if and only if $p \geq \rho(k)$ provided $k \leq k_N$, and losers impose $S$ when $k > k_N$. In this setting, $p_t$ increases over time since no news is good news for unsure voters. Therefore, the risky action can only be stopped, if used at all, when enough sure losers are observed, either because those losers obtain the majority or because the cutoff $\rho(k_t)$ jumps over $p_t$ upon the observation of a new loser (cutoff variation is discussed below). Theorem 8 implies that, provided that types are independent, $\rho(k) \geq \rho^{SD}$ for all $k < N/2$, where $\rho^{SD}$ is the single-decision-maker cutoff. One may prove that the equilibrium policy resulting from the majority rule is nonadverse to any voter.

[29]In that case, however, majority would simply ignore $i$ and proceed with experimentation. As a stronger case of adversity, suppose that 10 individuals face the following problem. Either they elect the safe action forever or they try $R$, in which case types are immediately revealed and a dictator is randomly, uniformly chosen, such that the dictator has an opposite type from all other voters (i.e., either $R$ is good for him and bad for all others or vice versa), with a 50% chance of being a winner. Ex ante, $R$ yields an individual expected value of $\pi = 1/10 * [pg + (1-p)s] + 9/10 * (1-p)s = pg/10 + (1-p)s$ (letting $r = 1$). On the other hand, a voter's probability of being a winner is $p/10 + (1-p)9/10 = 1/2$. Choosing $g = 3s$, the myopic cutoff is $p^M = 1/3$, so $p$ is above the myopic cutoff and yet voters prefer to avoid $R$ since $\pi < s$. Section 5.2 provides an example of endogenous adversity.

With negative-news shocks, it is not true any more that experimentation is inefficiently short. Unsure voters, ignoring losers, may push experimentation further than a utilitarian social planner. However, a social planner still has a higher value of experimentation than control-sharing voters.[30] For example, a social planner would always welcome immediate type revelation, whereas voters may prefer playing $S$ forever rather than learn their type, however fast, provided that the risk of loser trap is high enough. At the outset, a social planner may thus be more willing to experiment than individual voters. As the number of observed losers increases, the first effect starts to dominate, with the social planner stopping experimentation sooner than unsure voters under majority voting.

In view of Theorem 1, one may wonder whether cutoffs are also monotonic in this negative-news setting. The answer is negative. Counterexamples can be observed numerically or constructed with analytical results omitted here. Such violations can be explained as follows. Essentially, the loser trap is more severe with negative-news shocks. In the benchmark setting, unsure voters can always impose the safe action when they have the majority, and the only shock that may occur in that case is to become a winner. With negative-news shocks, in contrast, any unsure voter can, upon receiving a negative lump sum, suddenly join the minority of sure losers and hence face the worst possible situation. Negative news is compounded by a sudden control loss. This explains why the "insurance" effect resulting from the apparition of a new loser can, paradoxically, encourage experimentation. Seen differently, in the negative-news setting, $p$ simply increases over time, which is enough to make experimentation more attractive. In contrast, in the positive-news setting, the apparition of news winners is necessary for experimentation to continue, for otherwise, $p$ decreases until it causes experimentation to stop.[31] Note however that although cutoffs need not be monotonic, it is always true that experimentation decreases with the number of sure losers. Indeed, experimentation can only stop when a new loser is observed, since otherwise unsure voters become more optimistic about their types and have no reason to switch to $S$.

[30]More precisely, the value of experimentation of an agent, for a given policy, is the difference between that agent's value function if that policy is followed and the value that the agent gets if the action that gives him the highest expected payoff is played forever (see (3)). This definition captures the potential gain achieved by reacting to incoming news. The social planner, being the sole decision maker of his welfare maximization problem, chooses the policy that maximizes his value of experimentation. By contrast, the equilibrium policy differs from the optimum of any given agent, and thus provides him with a lower value of experimentation.

[31]From a technical standpoint, another distinctive feature of the negative-news settings is that the smooth-pasting property does not hold any more. Indeed, as time elapses, $p$ moves away from its threshold $p(k)$, so the value function need not be smooth at that cutoff. Instead, cutoffs are determined by direct comparison of value functions with and without starting experimentation.

### 5.2. *Mixed Shocks and Negative Value of Experimentation*

Suppose that the benchmark setting is modified as follows: If $R$ is good, it pays positive lump sums according to the jumping times of some Poisson process with intensity $\lambda_g$; if it is bad, it pays negative lump sums according to the jumping times of a Poisson process with intensity $\lambda_b$. Without loss of generality, also suppose that the payoff rate of $S$ is zero. In this case, state variables consist of the number $k^W$ of observed winners, the number $k^L$ of observed losers, and unsure voters' probability $p$ that $R$ is good for them. Since the number of revealed winners and losers can only increase over time, a backward induction argument on $k^W$ and $k^L$ shows that there exists a unique majority-voting equilibrium policy. If $\lambda_g > \lambda_b$, then no news is bad news, since shocks are more likely to happen if $R$ is good than if it is bad. This implies that, under this assumption, unsure voters become more pessimistic over time[32] and that they stop experimentation at some cutoffs $p(k^W, k^L)$, provided they are pivotal. Theorem 8 implies that $p^{SD} \leq p(k^W, k^L)$, where $p^{SD}$ is the single-decision-maker setting cutoff. This inequality holds for all $\nu$ voting rules. If the risky action requires the unanimity rule, then Theorem 9 implies that $p(k^W, k^L) \leq p^M$, where $p^M$ is the myopic cutoff: unanimity guarantees at least some experimentation.

*Negative Value of Experimentation*

With other voting rules, nonadversity need not hold, due to the following perverse effect: if a loser is observed, this may prompt other voters to experiment more by reducing their risk of the loser trap. The value of experimentation can be negative, that is, voters may prefer to elect the safe action even if the risky action has a higher immediate expected payoff. Here is such an example. There are three unsure voters, voting at the simple majority rule. If a loser is observed, the remaining two unsure voters are "protected": it is as if $R$ required unanimity among the two. This increases their willingness to experiment. If a winner is observed, the remaining two unsure voters are now on the brink: any winner among them will impose the risky action on the other. This risk reduces their willingness to experiment. Therefore, ex ante, the three voters know that if any one of them turns out to be a winner, other voters will soon revert to the safe action, while if one of them receives a negative lump sum, others will experiment more. This endogenous adversity makes $R$ unattractive even if its expected payoff is higher than $S$'s. For the value of experimentation to be negative, it is required that (i) the magnitude of loser trap be severe and (ii) learning be slow, so that experimentation takes time and the adversity described above lasts long. An explicit example is given in the Appendix

This section highlights an important virtue of requiring unanimity for $R$: the unanimity rule guarantees a nonnegative value of experimentation, whereas other voting rules may yield a negative value of experimentation.

[32]Precisely, one may show that $dp/dt = -(\lambda_g - \lambda_b)p(1-p)$.

## 6. PRIVATELY OBSERVED PAYOFFS

This section shows that even when payoffs are privately observed, the equilibrium policy of Section 2 can be implemented. Suppose that individuals can only observe their own payoffs and, at each time, the aggregate number of votes for each alternative. Voters cannot condition their voting policy on the current number of winners, since when everyone votes for $R$, it is impossible to tell sure winners apart from unsure voters. However, voters do learn the number of sure winners when it matters, that is, when cutoffs are reached. Indeed, each time a cutoff is reached at which unsure voters would want to stop given the number of winners that were last revealed, unsure voters vote for the safe action and sure winners vote for the risky action. That way, everyone observes the current number of winners, and unsure voters then decide whether to pursue experimentation to the next relevant cutoff (everyone votes for the risky action) or to vote for the safe action if no new winner is revealed. With this protocol, voters know the current number of winners only when $p$ reaches particular cutoffs, but that suffices to implement the policy of the public-information setting.[33]

To understand why it is in everyone's interest to follow this scheme, the intuition is as follows. First, sure winners always benefit from revealing their type, because this increases the duration of experimentation by cutoff monotonicity (Theorem 1). Second, unsure voters cannot gain from manipulating the choice process, because, conditional on being pivotal (i.e., $k \leq k_N$), they are already choosing their optimal action. For example, if an unsure voter voted for $R$ at some cutoff where he is supposed to vote for $S$, prompting other voters to believe that there are more winners than there really are, he will manage to extend experimentation. However, this will increase his risk of loser trap, since other voters may become sure winners during that time. The benefits and costs of such an extension of experimentation are already incorporated into the cutoffs derived under public information, making deviations unprofitable. The proof that the above protocol implements the public-information policy is sketched in the Appendix.

THEOREM 10: *The above protocol yields the same equilibrium as the experimentation policy based on publicly observed payoffs*.

## 7. EXTENSIONS

### *Correlated Types*

Positive correlation across types reduces risks of loser trap and winner frustration, and thus increases experimentation, compared to the case of independent types. Moreover, this also increases the speed of learning, reducing the

[33]In small committees, cheap talk would be another natural way for voters to truthfully reveal their payoffs.

time cost of experimentation. Results are shown formally in Strulovici (2010), which studies correlation for the case of two voters and where unanimity is required for $R$. With positive type correlation, an advantage of collective decision making compared to individual experimentation is to get rid of the free-rider problem identified in papers such as Bolton and Harris (1999), Li (2001), and Keller, Rady, and Cripps (2005). With perfect type correlation, voting on joint decision fully restores efficient experimentation, where when types are independent, letting each individual experiment on his/her own is efficient. A natural conjecture here is that imposing a joint decision over individual ones (assuming both forms of experimentation are possible) becomes more efficient as types get more positively correlated.

If some voters have negatively correlated types, this may increase or reduce experimentation, depending on the particular correlation structure. For example, suppose that there are only two voters, that unanimity is required for the risky action, and that voters have perfectly negatively correlated types. Then, as soon as one voter receives a lump sum, the other voter knows that he is surely a loser and this imposes the safe action. This completely destroys the value of experimentation, and voters stop at the myopic cutoff. A similar argument holds if unanimity is required for the risky action. Consider now the mixed-correlation case in which, say, two voters have a perfectly negatively correlated type with a third voter, and decisions are made according to the majority rule. In that case, the first two voters have perfectly positively correlated types and so have full control over the decision process: if the third voter receives a lump sum, the first two voters know that they are losers and thus impose the status quo. If one of the first two voters gets a lump sum, these voters impose the risky action. Overall, negative correlation must be seen more broadly as part of a more general description of potential alliances which may be formed and may have a positive or negative impact depending on the context. If it reduces a powerful group's expected power, negative correlation is likely to reduce experimentation. If it creates a more powerful group, the minority wants to experiment less but has no say on decisions, and so only the majority's increased incentive to experiment matters.

### *Factions and Heterogeneous Voting Weights*

If some voters have a greater decision weight, they are less subject to control-sharing effects and wish to experiment longer. For example, consider a setting with four voters, where voter 1 (only) can cast two votes and decisions are made according to the simple majority rule. If, say, voter 4 is the only sure winner so far, voter 1 can impose experimentation by siding with voter 4. As long as no other winner is observed, voter 1 can push experimentation up to the single-decision-maker threshold. If, say, voter 2 becomes a winner, voter 1 becomes subject to the risk of loser trap, as further experimentation may reveal that voter 3 also is a winner, resulting in a decisive coalition of sure winners.

Contrary to the benchmark setting, thus, experimentation can be interrupted by the occurrence of a new winner.

*Why Not a Two-Period Model?*

Some features of collective experimentation such as cutoff monotonicity, the impact of news arrival and type correlation, and the possibility of a negative value of experimentation, rely on the impact of one's experimentation on other voters' future experimentation, and hence require at least three periods. Infinite horizon provides time homogeneity, guaranteeing that cutoffs only depend on beliefs and not on time. Some potential applications, such as joint research and development projects, can be seen as stopping games, where the time dimension is an important feature of the model. Some results in this paper can be interpreted as comparative statics pertaining to stopping times.[34]

*Risk Aversion*

The analysis above does not require risk neutrality: it is enough that voters have a von Neumann–Morgenstern utility function, where lump sums actually correspond to "lump utils" or certainty equivalents thereof if the magnitude of these lump utils is random.

*Side Payments*

Side payments may restore efficiency under some implementations. However, a free-rider problem occurs if multiple individuals are needed to influence some voters. For example, if sure winners must coordinate to influence some unsure voter, then obtaining this switch amounts to a public good provision problem for sure winners. Furthermore, transfers may also be impractical or unethical.

*Switching Costs*

With a safe and a risky action, switching costs are easily accommodated, because the equilibrium policy can only switch actions once, from the risky to the safe action. Adding a cost there simply reduces the value of experimentation ex ante and, once the risky action is started, modifies indifference cutoffs.

*Multiple Risky Actions*

Adding a second risky action to the benchmark setting may decrease experimentation. This will occur, for example, if the two risky actions are perfectly

[34]Setting $\lambda = \infty$ reduces the model to two periods: before and after type revelation.

correlated (for each voter) and the payoff of the new action exacerbates the loser trap.[35] In that case, unsure voters may be unwilling to experiment with $R$ for fear that winners impose the new action, causing an even lower payoff for losers. Depending on the payoff structure, even a small probability that the second risky action becomes relevant can have a dramatic impact on the equilibrium policy.

### *Two Risky Actions*

Using a safe and a risky action provides an ideal setting to analyze conservatism: conservatism means choosing the safe action when the risky action would be more efficient. With two risky actions, conservatism could still be interpreted as settling inefficiently too early on one of the two risky actions when it would be more efficient to continue learning about the other action's value. In this spirit, Albrecht, Anderson, and Vroman (2007) showed in their model of search by committees that collective search settles earlier (i.e., acceptance thresholds are lower) than in the equivalent single-decision-maker setting.

### *Voter Heterogeneity*

If voters start with different type probabilities, this heterogeneity may increase experimentation. Indeed, heterogeneity concentrates more power in the hands of those voters who are pivotal today, because they are more likely to be also pivotal in the future. To illustrate with an extreme case, suppose that there are 9 voters, 4 of whom are (almost) sure to be winners and 4 of whom are (almost) sure to be losers. The remaining voter has (almost) perfect control over collective decision today, but also in the future: he will be able to side with whichever group corresponds to his preferred action.

### *Power Concentration*

The impact of group size on experimentation, as described by Proposition 1 and the discussion that follows, can be reinterpreted as a result on power concentration. As a simple model of power concentration, define an oligarchy as a subset of $O$ (odd) voters such that, at any time, the collective decision is the action chosen by the majority of that subset. Experimentation cutoffs are therefore defined as before, replacing $k$ by the number of winners within the oligarchy and replacing the total number of voters by the cardinal of the oligarchy. With this interpretation, Proposition 1 conveys a sense in which experimentation lasts longer if power is concentrated into fewer hands. In particular, a dictator sets the same experimentation cutoff as a single decision maker.[36]

[35] Perfect correlation ensures that there are still only two types of voters at any time and, therefore, no Condorcet cycles.

[36] This assumes type independence. With positively correlated types, a dictator would learn from others and set a lower cutoff than the single-decision-maker cutoff.

## 8. CONCLUSION

This paper has introduced a framework to study collective decisions when individual preferences evolve through learning. In this framework, experimentation with new alternatives is influenced by the potential effect of learning on future preferences and votes. Control-sharing effects arise, which introduce a conservative bias compared to the case of a single-decision-maker setting or a utilitarian social planner. Equilibrium policy is influenced by group size, voting rule, voting frequency, voter heterogeneity and correlation, the relative strength of loser trap and winner frustration, the ability to commit to an observation-dependent policy or to a fixed action, the amount of aggregate uncertainty, and the particular process of news arrival, among other factors.

In addition to the points developed in Section 7, there are several other important extensions to consider. For example, the cost or benefit of experimentation relative to the safe action may be heterogeneous across voters. Voters may also have an outside option allowing them to leave the group. In political applications, there may be several subgroups with high intragroup correlation and low intergroup correlation, with different and possibly evolving voting weights. Finally, some risky decisions may be (at least partially) irreversible. Such features may be necessary to analyze realistic settings. For example, "experimenting" with gas emissions has long-lasting effects, implying irreversibility. A country's cost of reducing gas emissions much depends on its primary source of energy, which is a source of heterogeneity. It will be useful to investigate the effect of such features on the equilibrium policy. When fully observation-contingent commitments are not available, what forms of commitment can improve efficiency under such extensions?

## APPENDIX

### A.1. *Proof of Theorem 1*

Suppose first that $k > k_N = (N-1)/2$, that is, sure winners have the majority. We show that $C(K, p) = R$ for all $p$. If not, there exist $\bar{K}$ with $|\bar{K}| > N/2$ and $\bar{p}$ for which $C(\bar{K}, \bar{p}) = S$. In this situation, $S$ is played forever whenever $p$ reaches $\bar{p}$, resulting in a constant value function of $s/r$ for all voters. Suppose that sure winner $i$ is pivotal. Then voting for $R$ yields an immediate expected payoff of $g$ and a continuation value function that is weakly greater than $s/r$, since sure winners get a payoff rate of at least $s$ no matter which action is played. This strictly dominates $s/r$. So the only undominated strategies, starting from $(\bar{K}, \bar{p})$, must start with $i$ voting for $R$. Since this is true of all sure winners and since $|\bar{K}| > N/2$, necessarily $C(\bar{K}, \bar{p}) = R$. This means that in any Markov equilibrium in undominated strategies, $R$ is elected forever as soon as winners gain the majority. The value function of unsure voters is easily computed in that case: if an unsure voter's type is good, which happens with probability $p$, he gets the same expected value as winners, $g/r$. Otherwise, he gets 0 forever. Therefore, $u(k, p) = pg/r$ for $k > k_N$. Now consider the case $k = k_N$, in which unsure voters have the majority, but only one new winner suffices for sure winners to gain the majority. If $i$ is an unsure voter, (2) reduces to

$$pg + \lambda p\left(\frac{g}{r} - V^{i,C}(k, p)\right) + (N - k_N - 1)\lambda p\left(\frac{pg}{r} - V^{i,C}(k, p)\right) - \lambda p(1-p)\frac{\partial V^{i,C}}{\partial p}(K, \cdot) > s. \tag{15}$$

In any equilibrium, $C(K, p) = R$ if and only if condition (15) holds. This condition is formally identical to the HJB equation for the optimization problem of a single decision maker. The solution is characterized by some indifference threshold $p(k_N)$ determined by the smooth-pasting condition of the Hamilton–Jacobi–Bellman equation (6), which reduces to

$$pg + p\lambda(g/r - s/r) + p\lambda(N - k_N - 1)(pg/r - s/r) = s, \tag{16}$$

using the relation $u^{k_N+1}(p) = pg/r$. The left-hand side of (16) is increasing in $p$, equal to 0 if $p = 0$ and higher than $g > s$ if $p = 1$. Therefore, the equation has a unique root, which can be reexpressed as

$$p(k_N) = \frac{\mu s}{\mu g + (g - s) + (N - k_N - 1)(p(k_N)g - s)}. \tag{17}$$

This shows that $C(K, p) = R$ if and only if $p > p(k_N)$. If $p \leq p(k_N)$, $S$ is chosen by unsure voters. Since no more learning occurs, $p$ remains constant forever, hence $S$ is played forever. The above policy entirely determines the value functions $w(k, p)$ and $u(k, p)$ of sure winners and unsure voters, for $k = k_N$, which are in fact computable in closed form by integration of their dynamic equation (a similar derivation was done by Keller, Rady, and Cripps (2005)):

$$w(k_N, p) = \frac{g}{r} - \frac{g - s}{r}\left(\frac{1 - p}{1 - p(k_N)}\right)^{N - k_N}\left(\frac{\Omega(p)}{\Omega(p(k_N))}\right)^{\mu} \tag{18}$$

and

$$u(k_N, p) = \frac{pg}{r} + \frac{s - p(k_N)g}{r}\left(\frac{1 - p}{1 - p(k_N)}\right)^{N - k_N}\left(\frac{\Omega(p)}{\Omega(p(k_N))}\right)^{\mu} \tag{19}$$

for $p \geq p(k_N)$, where $\Omega(p) = (1 - p)/p$. These functions are easily shown to be increasing in $p$, with $u(k_N, p) \geq pg/r$. Moreover, $u(k_N, p) = w(k_N, p) = s/r$ for $p \leq p(k_N)$, since the status quo is imposed forever.

Now suppose that $k = k_N - 1$. Then any new winner results in the case $k = k_N$ just analyzed. Again, (2) is formally equivalent to the stochastic control problem of a single decision maker. Using again the smooth-pasting property in (6), which implies that the derivative of the value function vanishes, any indifference threshold $p(k_N - 1)$ must solve

$$pg + p\lambda(w(k_N, p) - s/r) + p\lambda(N - k_N - 2)(u(k_N, p) - s/r) = s. \tag{20}$$

Since the left-hand side is increasing in $p$, equal to 0 for $p = 0$, and above $s$ for $p = 1$, the equation has a unique root $p(k_N - 1)$. The choice rule thus defined entirely determines value functions $u(k_N - 1, \cdot)$ and $w(k_N - 1, \cdot)$.

To show that $p(k_N - 1) > p(k_N)$, suppose that the contrary holds. Then $u(k_N, p(k_N - 1)) = w(k_N, p(k_N - 1)) = u(k_N - 1, p(k_N - 1)) = s/r$ and, by the smooth-pasting property, $\frac{\partial u}{\partial p}(k_N - 1, p(k_N - 1)) = 0$. Therefore, (20) becomes $p(k_N - 1)g = s$, which contradicts the assumption that $p(k_N - 1) \leq p(k_N) < p^M$.

Let us now show that $u(k_N - 1, p)$ is nondecreasing in $p$. Suppose that $p_t = \tilde{p} > \bar{p}$ and that unsure voters behave as if $p_t$ were equal to $\bar{p}$, meaning that they will stop experimenting after the same amount of time $\sigma_S$, unless a new winner is observed before. Let $\sigma_W$ denote the (possibly infinite) time at which a new winner is observed. Until $\sigma = \min\{\sigma_S, \sigma_W\}$, unsure voters receive nothing since $R$ is played and no new winner is observed. The value function of this strategy is thus equal to

$$u(p_t) = E_t\left\{e^{-r(\sigma - t)}\left[q\left(\frac{1}{N - k_N + 1}(w(k_N, p_\sigma) + h) + \frac{N - k_N}{N - k_N + 1}u(k_N, p_\sigma)\right) + (1 - q)\frac{s}{r}\right]\right\},$$

where $q = \Pr[\sigma_W < \sigma_S | p_t]$. We saw that $u(k_N, \cdot)$ and $w(k_N, \cdot)$ were increasing in $p$. Moreover, these values are above $s/r$. Indeed, $s/r$ is the value achieved if voters chose the status quo, which is suboptimal by definition of $\sigma_S$ and given that $p(k_N) < p(k_N - 1)$. Also, $p_\sigma$ is increasing in $p_t$ given the Bayesian updating dynamics. Finally, $\sigma_W$ is decreasing in $p_t$, since a higher $p_t$ makes it more likely that a payoff will be observed.[37] This also implies that $q$ is increasing in $p_t$, by definition of $q$ and by the fact that $\sigma_S$ is independent of $p_t$, by construction. Combining the above implies that $u(\tilde{p}) > u(\bar{p})$. Since unsure voters optimize their value function with respect to $\sigma_S$, this yields $u(k_N - 1, \tilde{p}) \geq u(\tilde{p}) > u(\bar{p}) = u(k_N - 1, \bar{p})$, which proves monotonicity of $u(k_N - 1, \cdot)$. $w(k_N - 1, \cdot)$ is also increasing in $p_t$. Indeed, let $\sigma_1 < \sigma_2$ be the arrivals times of lump sum to the next two new winners. As is easily shown, these

[37]Conditional on $p_t$, $\sigma_W$ is the mixture of exponential variables with intensity $\lambda j$, $j \in \{0, \ldots, N - k_N + 1\}$, with mixture weights $\{\rho_j\}$ corresponding to the binomial distribution $B(N - k_N + 1, p_t)$. Monotonicity is in the sense of first-order stochastic dominance.

stopping times are decreasing in $p_t$ in the sense of first-order stochastic dominance. This, given the fixed experimentation thresholds $p(k_N)$ and $p(k_N-1)$, implies that the distribution of the (possibly infinite) stopping time $\sigma_S$ at which experimentation stops increases in $p_t$ in the sense of first-order stochastic dominance. Finally, since

$$w(k_{N-1}, p_t) = E_t\left[\frac{g}{r}\left(1-e^{-r(\sigma_S-t)}\right) + \frac{s}{r}e^{-r(\sigma_S-t)}\right],$$

this shows that $w(k_{N-1}, \cdot)$ is increasing in $p_t$. The remainder of the proof proceeds by backward induction on $k$, where the induction hypothesis is that (i) for all $k' > k$, $C(k', p) = R$ if and only if $p > p(k')$, where (ii) $p(k')$ is nonincreasing for $k' > k$, and (iii) the resulting value functions $u(k', \cdot)$ and $w(k', \cdot)$ are nondecreasing in $p$. The general induction step is then proved exactly as above.

We now show cutoff monotonicity. We have seen above that $p(k) = 0$ for $k > k_N$. The fact that $p(k_N) \geq p^{SD}$ with strict inequality if $N > 1$ comes from the comparison of (17) and (1). Monotonicity of $p(k)$ is part of the induction in the proof of Theorem 1. There remains to show that $p^M > p(0)$. The indifference condition for $p(0)$ is

$$\begin{aligned} p(0)g + p(0)\lambda\big(w(1, p(0)) - s/r\big) \\ + p(0)\lambda(N-1)\big(u(1, p(0)) - s/r\big) = s. \end{aligned} \tag{21}$$

Since $p(0) > p(1)$, unsure voters strictly prefer experimentation at $p = p(0)$ when $k = 1$. Therefore, $u(1, p(0)) > s/r$. Since winners always get a higher expected payoff than losers no matter what action is chosen, $w(1, p(0)) \geq u(1, p(0))$. Therefore, the second and third terms on the left-hand side of (21) are positive, which implies that $p(0)g < s$ or, equivalently, that $p(0) < p^M$.

Monotonicity of $u$ and $w$ with respect to $p$ was shown as part of the induction hypothesis of the above proof. If $k > k_N$, $R$ is elected forever since winners have the majority. This determines value functions for this case and yields the last claim. To show monotonicity in $k$ of $w$ for $k \leq k_N$, we proceed by induction. Clearly, $g/r = w(k_N + 1, p) \geq w(k_N, p)$. Suppose that $w(k, p) \leq w(k+1, p)$. We need to show that $w(k-1, p) \leq w(k, p)$. Let $\phi(p) = w(k+1, p) - w(k, p) \geq 0$ and $\psi(p) = w(k, p) - w(k-1, p)$. Since $p(k-1) \geq p(k)$, $\psi(p) \geq 0$ for $p \leq p(k-1)$. Recall the dynamic equation of $w$ for $p \geq p(k-1)$ and $\tilde{k} \geq k-1$:

$$\begin{aligned} -rw(\tilde{k}, p) + \lambda(N-\tilde{k})p(w(\tilde{k}+1, p) - w(\tilde{k}, p)) \\ - \lambda p(1-p)\frac{\partial w}{\partial p}(\tilde{k}, p) + g = 0. \end{aligned}$$

Taking the difference of the resulting equations for $\tilde{k} = k, k-1$ and rearranging terms yields

$$(r + \lambda p(N-k+1))\psi(p) = \lambda p(N-k)\phi(p) - \lambda p(1-p)\psi'(p).$$

Suppose $\phi$ is nonnegative by the induction hypothesis. Then the previous equation can be rewritten as $\psi'(p) \le \alpha(p)\psi(p)$ for function $\alpha$. A direct application of Gronwall's inequality along with $\psi(p(k-1)) \ge 0$ proves that $\psi$ is nonnegative, completing the induction step.

To show monotonicity of $u$ with respect to $k \le k_N$, fix some $k \le k_N$. The dynamic equation of $u$ for $p \ge p(k-1)$ and $\tilde{k} \ge k-1$ is

$$\begin{aligned}&-ru(\tilde{k}, p) + \lambda p(w(\tilde{k}+1, p) - u(\tilde{k}, p))\\&\quad + \lambda p(N-\tilde{k}-1)(u(\tilde{k}+1, p) - u(\tilde{k}, p))\\&\quad - \lambda p(1-p)\frac{\partial u}{\partial p}(\tilde{k}, p) + pg = 0.\end{aligned}$$

Let $\phi(p) = u(k+1, p) - u(k, p)$, $\phi^w(p) = w(k+1, p) - w(k, p)$, and $\psi(p) = u(k, p) - u(k-1, p)$. Taking the difference of the previous equation for $\tilde{k} = k, k-1$ and rearranging terms yields

$$\begin{aligned}(22)\qquad &(r + \lambda p(N-k+1))\psi(p)\\&\quad = \lambda p[\phi^w(p) + (N-k-1)\phi(p)] - \lambda p(1-p)\psi'(p).\end{aligned}$$

We already know that $\phi^w$ is positive. Therefore, if $\phi$ were also nonnegative, the argument we just used for $w$ would also show that $\psi$ is nonnegative. In particular, if one can show that $u(k_N, p) \ge u(k_N - 1, p)$, a backward induction will prove the result for all $k \le k_N$. Combining (18) and (19) implies that, for $k = k_N$,

$$\begin{aligned}\phi^w(p) + (N-k_N-1)\phi(p) &= \frac{g - s - (N-k_N-1)(s - p(k_N)g)}{r}\\&\quad \times \left(\frac{1-p}{1-p(k_N)}\right)^{N-k_N}\left(\frac{\Omega(p)}{\Omega(p(k_N))}\right)^{\mu}.\end{aligned}$$

Therefore, the left-hand side has the sign of $g - s - (N-k_N-1)(s - p(k_N)g)$. From the cutoff formula (16), this expression has the same sign as $s - p(k_N)g$, which is positive. Therefore, the first term on the right-hand side of (22) is nonnegative for $k = k_N$, which implies that $\psi$ is nonnegative for $k = k_N$. This fills the missing step of the induction, concluding the proof that $u$ is increasing in $k$ for $k \le k_N$.

To show the last statement, observe that $u(k_N + 1, p) = pg/r$ from Theorem 1, and that $u(k_N, p) > pg/r$ from (19). *Q.E.D.*

### A.2. *Proof of Theorem 2*

The proof is similar to that of Theorem 1, proceeding by backward induction on the number $k$ of winners. For $k \geq \bar{k}$, the utilitarian optimum is to choose $R$ forever even if $p = 0$, since sure winners' gains from $R$ outweigh the aggregate gain from $S$ even if all unsure voters get nothing from $R$. This fact can be expressed as $q(k) = 0$ for $k \geq \bar{k}$. The resulting welfare is $W(k, p) = k\frac{g}{r} + (N - k)\frac{pg}{r}$. Consider next $k = \bar{k} - 1$. Let $w^C(k, p)$ and $u^C(k, p)$ denote the value functions of sure winners and unsure voters if policy $C$ is used, given that $R$ is played forever if a new winner is observed, and let $W^C(k, p) = kw^C(k, p) + (N - k)u^C(k, p)$, denote utilitarian welfare under policy $C$. Then the utilitarian criterion $C$ must solve

$$W_t^{k_t,C} = \sup_\theta E_t\left[\int_t^\sigma e^{-r(\tau-t)}\sum_i d\pi^i_{\theta_\tau}(\tau) + e^{-r(\sigma-t)}W_\sigma^{k_t+1,C}\right],$$

where $\sigma$ is the first (possibly infinite) time at which a new winner is observed and where $W_\sigma^{k_t+1,C} = W(\bar{k}, p_\sigma)$, the welfare that was computed earlier for $k = \bar{k}$. This is a standard control problem, whose solution is Markov. The indifference boundary must satisfy the smooth-pasting condition

$$kg + (N - k)pg + (N - k)\lambda p\left[\frac{kg + (N - k)pg}{r} - \frac{Ns}{r}\right] = Ns,$$

which has a unique root $q(k)$, since the left-hand side is increasing in $p$, greater than $Ns$ if $p = 1$, and less than $Ns$ for $p = 0$, by definition of $\bar{k}$. Therefore, $C(k, p) = R$ if and only if $p \geq q(k)$. This entirely determines $w(k, \cdot)$, $u(k, \cdot)$, and $W(k, \cdot)$, which are easily shown to be increasing in $p$. The remainder of the proof proceeds by backward induction on $k$ as in Theorem 1, where the induction hypothesis is that (i) for all $k' > k$, $C(k', p) = R$ if and only if $p > q(k')$, where (ii) $q(k')$ is nonincreasing for $k' > k$, and (iii) resulting value functions $w(k', \cdot)$, $u(k', \cdot)$, and $W(k', \cdot)$ are nondecreasing in $p$. *Q.E.D.*

### A.3. *Probability of Receiving a Lump Sum Between $p$ and $q < p$*

Let $p_S$ denote the probability that an individual with initial probability $p$ of being a winner receives a lump sum by the time his belief has dropped to $q < p$.

PROPOSITION 2: $p_S = (p - q)/(1 - q)$.

PROOF: From the Bayesian updating equation, $p_t = (pe^{-\lambda t})/((1 - p) + pe^{-\lambda t})$. Therefore, $q$ is reached at a time $T$ such that $e^{-\lambda T} = \Omega(p)/\Omega(q)$, where $\Omega(p) = (1 - p)/p$. Conditional on the individual being a winner, the probability of getting a lump sum before time $T$ is simply $1 - e^{-\lambda T}$, since the arrival rate is an exponential random variable with parameter $\lambda$. Combining the previous formulas concludes the proof. *Q.E.D.*

### A.4. *Proof of Theorem 9*

For any safe rate $s$ and policy $C$, voter $i$'s expected payoff with policy $C$ is

$$V_C^i = E\left[\int_0^\infty e^{-rt}\, d\pi_{C_t}^i(t)\right] = \int_0^\infty e^{-rt} E[d\pi_{C_t}^i(t)], \tag{23}$$

where expectations are conditioned on $\gamma$:

$$E[d\pi_{C_t}^i(t)] = \Pr[C_t = S]s\,dt + \Pr[C_t = R]E[d\pi_{C_t}^i(t)|C_t = R].$$

Therefore, if $E[d\pi_{C_t}^i(t)|C_t = R] > s\,dt$ for all $t$, then $V_C^i > s/r$, implying that $G_C^i(\gamma) > s$. Suppose that $s < g(\gamma^i)$. Then, by definition of $g(\cdot)$ and by the fact that the probability of each type is a martingale, $E[d\pi_R^i(t)] = g(\gamma^i)\,dt > s\,dt$. Moreover, $C$'s nonadversity with respect to $i$ implies that $E[d\pi_{C_t}^i(t)|C_t = R] \geq E[d\pi_R^i(t)]$ as will be shown shortly. This inequality implies that $G_C^i(\gamma) > s$ for all $s < g(\gamma^i)$, which concludes the proof. To show the inequality, observe that by Bayes' rule, $C$ is nonadverse for $i$ if and only if $\Pr[\theta^i \text{ good}|C_t = R] \geq \Pr[\theta^i \text{ good}|C_t = S]$ for almost all $t$.[38] Moreover,

$$\begin{aligned} E[d\pi_{C_t}^i(t)|C_t = R] &= \Pr[\theta^i \text{ good}|C_t = R]E[d\pi_{C_t}^i(t)|C_t = R, \theta^i \text{ good}] \\ &\quad + \Pr[\theta^i \text{ bad}|C_t = R]E[d\pi_{C_t}^i(t)|C_t = R, \theta^i \text{ bad}]. \end{aligned} \tag{24}$$

Combining these results yields the inequality. *Q.E.D.*

### A.5. *Negative Value of Experimentation*

Let $g > 0$ and $b < 0$ be the expected payoff rates of the risky arm for sure winners and sure losers, respectively. Let $p^M$, $p^{SD}$, $p^L$, $p^W$, and $p^3$, respectively, denote the myopic cutoff, the single-decision-maker cutoff, the two unsure voters' cutoff when the third voter is a loser, the two unsure voters' cutoff when the third voter is a winner, and the experimentation cutoff when all three voters are unsure. For the parameter values $g = 0.1$, $b = -1$, $s = 0$, $r = 1$, $\lambda_b = 0.1$, and $\lambda_g = 0.11$, cutoffs have the values

| $p^M$ | $p^{SD}$ | $p^L$ | $p^W$ | $p^3$ |
|---|---|---|---|---|
| 0.9091 | 0.9001 | 0.9016 | 0.9083 | 0.9095 |

[38]Precisely, we have for all $t$, $\Pr[C_t = R|\text{good}] \geq \Pr[C_t = R|\text{bad}] \Leftrightarrow \Pr[C_t = R|\text{good}] \geq \Pr[C_t = R] \Leftrightarrow \Pr[\text{good}|C_t = R] \geq \Pr[\text{good}] \Leftrightarrow \Pr[\text{good}|C_t = R] \geq \Pr[\text{good}|C_t = B]$.

The most important result is that $p^3 > p^M$: voters stop experimentation at a probability level where $R$'s expected payoff is strictly above $S$'s. As explained above, $p^L$ is much lower[39] than $p^W$, meaning that if a voter is a loser, experimentation lasts much longer than if he is a winner. From (3), this implies that the value of experimentation is negative at $p^3$, since $V(p^3) = s/r < p^3 g/r$.

### A.6. *Proof of Theorem 10 (Sketch)*

For sure winners, voting $R$ forever is optimal as it maximizes their immediate payoff as well as the length of experimentation, due to the cutoff monotonicity established in Theorem 1. Under the protocol described in Section 6, unsure voters only observe the state $k$ when particular cutoffs are reached. Let $l$ denote the number of winners that was last revealed. For $p > p(l)$, unsure voters only know that the number $\tilde{k}$ of current winners is greater than or equal to $l$. Unsure voters are only pivotal if $\tilde{k} \leq k_N$. By Theorem 1, $u(\bar{k}, p) \geq u(l, p)$ for $l \leq \bar{k} \leq k_N$. Therefore, $E[u(\tilde{k}, p)|l \leq \tilde{k} \leq k_N] \geq u(l, p) > s/r$ for $p > p(l)$. Therefore, it is optimal for unsure voters to choose the risky action whenever indicated by the protocol, conditional on being pivotal. Upon reaching $p(l)$, if it turns out that $k = l$, that is, no new winner has been observed since the last release of public information, then it is optimal for unsure voters to stop: their value function is identical to the benchmark case, which is equal to $s/r$. *Q.E.D.*

## REFERENCES


ALBRECHT, J., A. ANDERSON, AND S. VROMAN (2007): "Search by Committee," Working Paper; *Journal of Economic Theory* (forthcoming). [937,962]

BALDWIN, R. E. (1985): *The Political Economy of U.S. Import Policy*. Cambridge, MA: MIT Press. [933]

BERGEMANN, D., AND U. HEGE (1998): "Venture Capital Financing, Moral Hazard and Learning," *Journal of Banking and Finance*, 22, 703–735. [937]

——— (2001): "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, 36, 719–752. [937]

BHAGWATI, J. (1988): *Protectionism*. Cambridge, MA: MIT Press. [933]

BOLTON, P., AND C. HARRIS (1999): "Strategic Experimentation," *Econometrica*, 67, 349–374. [936,960]

CALLANDER, S. (2009): "Searching for Good Policies," Working Paper, Northwestern University. [937]

COMPTE, O., AND P. JEHIEL (2008): "Bargaining and Majority Rules: A Collective Search Perspective," Working Paper, Paris School of Economics. [937]

DÉCAMPS, J.-P., AND T. MARIOTTI (2004): "Investment Timing and Externalities," *Journal of Economic Theory*, 118, 80–102. [936,937]

DIXIT, A. K. (1993): *The Art of Smooth Pasting*. Fundamentals in Pure and Applied Economics, Vol. 55. London: Routledge. [943]


[39]Indeed, $p^L$ is close to the single-decision-maker cutoff, while $p^W$ is close to the myopic cutoff.

FERNANDEZ, R., AND D. RODRIK (1990): "Why Is Trade Reform so Unpopular? On Status Quo Bias in Policy Reforms," Working Paper 3269, NBER. [933]

——— (1991): "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty," *American Economic Review*, 81, 1146–1155. [937]

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. Cambridge, MA: MIT Press. [941]

KELLER, G., S. RADY, AND M. CRIPPS (2005): "Strategic Experimentation With Exponential Bandits," *Econometrica*, 73, 39–68. [936-938,960,964]

LI, H. (2001): "A Theory of Conservatism," *Journal of Political Economy*, 109, 617–636. [936,960]

MALUEG, D. A., AND S. O. TSUTSUI (1997): "Dynamic R&D Competition With Learning," *RAND Journal of Economics*, 28, 751–772. [937]

MESSNER, M., AND M. POLBORN (2008): "The Option to Wait in Collective Decisions," Working Paper, Bocconi University and University of Illinois at Urbana–Champaign. [937]

ROBERTS, K. (1989): "The Theory of Union Behavior: Labour Hoarding and Endogenous Hysteresis," Mimeo, London School of Economics. [941]

RODRIK, D. (1993): "The Positive Economics of Policy Reform," *American Economic Review*, 83, 356–361. [933]

STRULOVICI, B. (2010): "Voting and Experimentation With Correlated Types," Note, Northwestern University. Available at http://faculty.wcas.northwestern.edu/~bhs675/VotExCor10Mar24.pdf. [960]

*Dept. of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208-2600, U.S.A.; b-strulovici@northwestern.edu.*

*Manuscript received July, 2008; final revision received December, 2009.*