Learning While Voting: Determinants of Collective Experimentation

Bruno Strulovici*

May 30, 2008

Abstract

This paper analyzes collective decision making when individual preferences evolve through learning. Votes are affected by their anticipated effect on future preferences. The analysis is conducted in a two-arm bandit model with a safe alternative and a risky alternative whose payoff distribution, or "type", varies across individuals and may be learned through experimentation. Society is shown to experiment less than any of its members would if he could dictate future decisions, and to be systematically biased against experimentation compared to the utilitarian optimum. Control sharing can even result in *negative value of experimentation*: society may shun a risky alternative even its expected payoff is higher than the safe one's. Commitment to a fixed alternative can only increase efficiency if *aggregate* uncertainty is small enough. Even when types are independent, a positive news shock for anyone raises everyone's incentive to experiment. Ex ante preference correlation or heterogeneity reduces these inefficiencies.

*Nuffield College and Economics Department, Oxford University. I am very grateful to Meg Meyer, Paul Milgrom, Andrea Patacconi, and Kevin Roberts for insightful conversations, as well as to Daron Acemoglu, Patrick Bolton, Alessandra Casella, Jeff Ely, Godfrey Keller, Andrea Prat, John Quah, Marzena Rostek, Marek Weretka, and seminar participants at Oxford University, Gerzensee, 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, and the Stanford GSB for numerous comments.

1 Introduction

The positive social choice literature usually assumes that individuals perfectly know their own preferences. In reality, preferences may evolve through learning. For example, a reform may have uncertain consequences, which can only be learned by experimenting with that reform. Similarly, decisions in committees are often analyzed under the assumption that payoff distributions are perfectly known to committee members at the time of their decisions, although this is rarely the case in practice. This difference is important if those decisions are made repeatedly, because one's vote at any given time must take into account the impact of elected decisions on everyone's future preferences, willingness to experiment, and votes.

This paper analyzes how experimentation - the fact that an alternative with comparatively lower immediate expected payoff is chosen to learn more about its value - is affected by the nature and amount of individual control over collective decisions. The main analysis is conducted in a two-arm bandit model in which a "safe" alternative yields a constant, homogeneous payoff to all and a "risky" alternative yields payoffs according to some distribution, or *type*, which varies across individuals. At any time, society elects one of the two actions according to some voting rule. Individuals may learn their type through experimentation with the risky alternative.[1] The paper first analyzes a setting in which each individual's type is either "good" or "bad" and gets revealed according to a particular process of news arrival that is simple enough to provide many results and their intuition. The question is then reformulated as an abstract stochastic control problem allowing to extend several results to arbitrary news arrival processes. Several news arrival processes are considered in turn, with an increasing degree of complexity and nuance in the results.

A key feature of the model is the feedback effect occurring between individual preferences and collective decisions. Not only does preference uncertainty affect society's choices, but the reverse is also true. This social phenomenon has been described in the context of conservatism by Kuran (1988):

> (...) a complete model for the study of conservatism would have a circular dynamic structure, with individuals' choices driven by their beliefs and preferences; society's choices generated by its

[1] Focusing on two actions gets rid of Condorcet cycles and ensures robustness of the equilibrium concept used in the analysis. Section 13 studies an example with three actions where learning is correlated across actions.

> members' choices; and, completing the circle, these members' beliefs and preferences influenced by society's choices. It would thus incorporate three interactive processes: that by which individuals' seek and integrate information to form their beliefs and preferences regarding the alternatives they face; that by which society combines these choices to select policies, institutions, and technologies; and finally, that by which collective outcomes mold individuals' beliefs and preferences. (...) the aim of theoretical analysis on the subject should be to elucidate these three processes with an eye toward deriving propositions as to when, how, and to what extent individuals and collectivities adapt to changes in environmental factors.

The first and most general result of this paper is that sharing control always reduces the value of experimentation compared to an equivalent single decision maker setting. In some cases, the *value of experimentation can be negative*, meaning that society may reject a new alternative with a higher expected payoff than the current alternative. This latter result may be surprising.[2] Many economic fields use the concept of a positive "option value," which captures a decision maker's ability to react to news.[3] In all these cases, the ability to make a decision upon learning some payoff-relevant information amounts to a positive "option value." When decisions are made collectively this ability should intuitively persist, if weaker, as long as each individual has *some* control over collective decisions. Contrary to this intuition, the paper shows that, even when voter types are independently distributed and voters have identical voting weight, the value of experimentation can be negative. This phenomenon can only occur in truly dynamic settings, as it is caused by *adverse duration* of experimentation. This can be explained as follows:[4] good news for one voter increases other voters' risk of being imposed the risky action in the future, as that voter is more likely to support it. Bad news, on the other hand, reduces that risk, making experimentation relatively more appealing to other voters. Experimentation duration may thus create some endogenous adversity of the equilibrium voting policy against *all* voters. This phenomenon can occur with as few as three individuals, and is more likely to occur when learning is slow, i.e. when the time cost of experimentation is high.

Control sharing effects may be decomposed into a *loser trap* risk: society imposes the risky action to a bad-type individual, and a symmetric *winner frustration* risk: society imposes the safe action to a good-type individual. The relative strength of these risks helps understand the

[2]The result does not rely on commitment ability or on asymmetric information. It is purely due to control sharing effects, as shown here and in Section 11.

[3]In finance, the exercise a (European) stock option depends on the observation of stock price at maturity, giving the option its strictly positive value. Similarly, real options were introduced by Myers (1977) to model discretionary investment of firms in response to growth opportunities. In corporate finance, Leland (1994) and a growing literature model bankruptcy events as an endogenous decision of shareholders prompting them to postpone default even when interest payments exceed incoming cash flows.

[4]A clearer intuition, requiring more buildup, is presented in Section 11.

effect on experimentation of news shocks, group size, and the quorum required in supermajority or q-rules, among other things. For example, only the loser trap may arise when learning is infinitely fast, so that experimentation entails no time cost. Given the possibility of immediate revelation of everyone's type, society may still reject the risky action, provided that the risk of loser trap is high enough. As group size goes to infinity, voters behave *myopically* if types are independent. Intuitively, individual control over future decisions is infinitely diluted, which annihilates individual value of experimentation. This intuition is only partially correct however, as experimentation is not monotonically decreasing with respect to group size: the addition of new voters reduces the risk of winner frustration, an effect that may locally dominate.

In the benchmark setting, if an individual's type is good, the risky alternative pays him some lump-sums at exponentially distributed times. If his type is bad, the risky action pays him nothing. Therefore, an individual knows for sure that his type is good as soon as he receives a lump-sum (he is then a "sure winner"), but remains "unsure" about his type otherwise. Experimenting with the risky action results in a better assessment of individual valuations for that action, hence a better knowledge of individual rankings of alternatives.

In this setting, society experiments *too little* compared to the utilitarian policy.[5,6] This result stems from two effects. A social planner, having full control over decisions, can always exploit whatever information is revealed through experimentation, whereas individuals are constrained by control sharing.[7] This makes, other things equal, experimentation more attractive for him than it does for any individual subject to majority voting. Moreover, whenever a majority of unsure voters imposes the safe action, it ignores the utility of sure winners, unlike a utilitarian social planner. If one reverses the benchmark setting, as in Section 10, so that negative lump-sums reveal sure losers and the safe action entails a small cost for unsure voters, the direction of the second effect is reversed. This reversal generates two experimentation regimes, one with too little experimentation (society simply rejects the risky alternative), the other with too much experimentation (the majority ignores sure losers' welfare), compared to the utilitarian optimum.

Comparison with the utilitarian optimum can be reinterpreted in terms of commitment. If individuals are ex ante identical - placed behind a veil of ignorance - and able to commit at the outset

[5]The comparison is based on simple majority voting, though the argument can be generalized.

[6]Throughout the paper, experimenting "less" or "too little" or "too short" means that the domain of beliefs under which society picks the risky action is smaller or too small compared to some reference such as the utilitarian criterion.

[7]In that respect, the main analysis assumes that individual payoffs are publicly observed. Section 13 considers the case of privately observed payoffs and shows that the voting equilibrium derived in the benchmark setting with publicly observed payoffs is truthful.

to some (anonymous) policy, they choose the utilitarian policy. Indeed, their expected utilities are identical and proportional to utilitarian welfare. Therefore, collective experimentation lasts longer if individuals can initially commit to a state-dependent policy. If, however, the only form of commitment available is with respect to a *fixed action* rather to a state-dependent policy, there is a trade-off between value of commitment and value of experimentation. Imposing an action ex ante introduces rigidity which prevents society from exploiting incoming information. Such commitment can only increase efficiency if aggregate uncertainty is low enough. Section 6 suggests further interpretations of these results in terms of decision frequency and delegation.

In light of the above results, it is tempting to shift to a normative analysis and ask whether some voting rules are systematically more efficient than others. 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. The unanimity rule thus exacerbates winner frustration.[8] Whatever control is gained from being able to veto the risky action is balanced by a control loss for enforcing that same action. Examples indeed show that with the unanimity rule, experimentation may last longer or shorter than under the majority rule.

Another key issue concerns the impact of news arrival on incentives and welfare. With independent types, incoming news for an individual brings no information externality on other voters. However, it creates a payoff externality as it affects that individual's voting decisions. What is the direction of this payoff externality? In the benchmark setting, *good news for any individual is good news for all, and prompts society to experiment more.* This result may seem counterintuitive: the occurrence of a new winner brings unsure voters closer to the brink, where the risky action is imposed on them forever. However, it also makes unsure voters more likely to enjoy the risky action if they turn out to be winners, an effect that must dominate whenever society experiments in the first place and experimentation is contingent on the arrival of such positive news shocks.

The severity of control sharing effects diminishes if types are positively correlated. Correlation also allows individuals to learn from one another's payoffs and thus reduces the time cost of experimentation. This observation is particularly relevant for large societies composed of groups with high intra-group correlation.

[8]More generally, loser trap becomes weaker and winner frustration stronger as the number of voters required to implement the risky action increases.

An abstract formulation of the model as a general stochastic control problem makes it possible to prove some of the above results without specifying particular dynamics or direction for news arrival. The analysis introduces a collective version of the Gittins index, which provides a simple mathematic language to prove that collective experimentation is *always* shorter than what any individual would choose with dictatorial power. Provided that, for any individual, the risky action is more likely to be implemented if that individual benefits from it than otherwise, there is always *some* experimentation in equilibrium, i.e. a domain of beliefs where the risky action's expected payoff is lower than the safe action's for a decision number of individuals who nevertheless prefer to impose the risky action. Requiring unanimity for the risky action guarantees such non-adversity.

The issues described so far are concerned with how control sharing affects learning and experimentation. The paper briefly considers a dual question: how does the possibility that individual rankings of social alternatives evolve through learning, potentially resulting in majority shifts, affect collective decisions? In an example with multiple risky actions, Section 13 shows that even a slight risk of preference modification through learning can result in equilibrium breakdowns.

Section 2 discusses the related literature. Section 3 presents a simple example and potential applications. Section 4 analyzes the benchmark setting with simple majority voting and independent types. Section 5 compares voting-based experimentation to the utilitarian optimum. Section 6 considers various forms of commitment. Section 7 compares efficiency of various voting rules. Section 8 analyzes the effect of voter heterogeneity and correlation on experimentation. Section 9 formulates the model as an abstract stochastic control problem and extends several results to general news arrival processes. Sections 10 and 11 respectively analyze the cases of negative and mixed news shocks and show how the results are modified in these settings. Section 12 considers the case of privately observed payoffs. Section 13 discusses and relaxes some assumptions of the model, and in particular allows for multiple risky actions. Section 14 concludes.

2 Related Literature

The present analysis contributes to the literature on collective conservatism. In contrast to earlier literature, it does not rely on arguments such as exogenous transaction costs or sunk costs or bounded rationality as surveyed by Kuran (1988). The form of conservatism studied here is closely related to Fernandez and Rodrik (1991), who were the first to explain status quo bias as a consequence of uncertainty about winner identity. In their setting, however, there is no aggregate uncertainty (the utilitarian optimum is known from the beginning) and no experimentation, since whatever individuals learn about their type in the first period has no impact on the collective decision of the second period.[9] Explicitly considering such possibilities raises numerous issues tackled in the present paper.

The paper also contributes to a developing literature on games and experimentation, in which conservatism may arise as a consequence of strategic information acquisition.[10] Those papers identify an informational free-riding problem in settings where agents can experiment individually with some risky action to learn about its *common* value. The present paper considers a reverse setting, in which a single collective action is made at any time, but the value of the action may vary across individuals. The analysis of the benchmark setting owes conceptual and technical clarity to the use of exponential bandits, building on Keller Rady and Cripps (2005).[11]

The paper a burgeoning literature analyzing collective search in various settings (Albrecht, Andersen, Vroman (2007), Compte and Jehiel (2008), and Messner and Polborn (2008)). In these models a group must choose, at any time, between accepting some outstanding proposal or trying a new proposal with iid characteristics (Messner and Polborn also discuss correlation across the two periods of their setting). Compte and Jehiel show in particular that more stringent majority requirements select more efficient proposals but take more time to do so. Albrecht et al. find that committees are more permissive than a single decision maker facing an otherwise identical search problem. In those papers, committees never return to past proposals. In contrast, 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.

[9]The reference against which conservatism is established is also different. Their main result is that society can reject some risky alternative which would gain the *majority* if types were revealed. However, society does choose the *utilitarian* optimum.

[10]See Bolton and Harris (1999), Décamps and Mariotti (2004) and Keller, Rady, and Cripps (2005), and Li (2001).

[11]Exponential bandits were introduced by Presman and Sonin (1990) and used in economics by Malueg and Tsutsui (1997), Bergemann and Hege (1998, 2001), Décamps and Mariotti (2004) and Keller, Rady, and Cripps (2005).

3 Example and Applications

3.1 A Simple Example

Three friends, Ann, Bob, and Chris, go to a restaurant once every week-end. Each week-end, they choose their restaurant using the simple majority rule. A new restaurant has just opened. Should the friends try it? Do they try it? Suppose the alternative is a restaurant that gives utility 1 to all. For each voter, the new restaurant can be either bad (yielding 0 utility) or good (yielding utility $u > 1$). Suppose that preferences, or "types" are independently distributed across friends (e.g. Ann is no more or less like likely to appreciate the new restaurant if Bob likes it, etc.), with both types having an *ex ante* probability of 1/2.

Immediate Full Type Revelation Suppose that, if they try this new restaurant, all voters immediately learn their type. With probability 1/8, Ann and Bob will like it but Chris won't. In this case, Chris is trapped into always returning to that restaurant, as Ann and Bob have the majority. This situation will be referred to as the "loser trap". Also with probability 1/8, Chris is the only one who turns out to like the restaurant, but is blocked from exploiting this discovery for future dinners by Ann and Bob. This symmetric situation will be referred to as "winner frustration". Overall, there is a probability 1/4 that Chris loses control over the decision process, compared to the situation in which he could choose the restaurant by himself in the future. Depending on u and on how time is discounted, these control-loss effects may be such that Chris and, by symmetry, all voters prefer not to try the new restaurant even though each of them would have preferred to try it if he had full control over future decisions.

Gradual Type Revelation Now suppose that a voter likes the new restaurant if and only if he finds a dish there that he really likes. In this case several visits to the restaurant may be needed to find out one's type. This can lead to situations in which friends experiment with that restaurant until either a majority of them likes it, or a majority of them judges unlikely that they will find anything like there. With this assumption, suppose that, in their first try, only Chris discovers that he likes the new restaurant. What effect does it have on Ann and Bob? Does this incite them to try it more or, on the contrary, prompts them to block new experimentation? Good news for Chris reduces the risk of winner frustration for Ann and Bob, but increases the probability of the loser trap. It turns out that good news for Chris *always* makes Ann and Bob more willing to experiment, as shown in Section 4.

Social Efficiency What would a social planner, wishing to maximize the sum of utilities of the three friends, choose to do? Suppose that u is very close to 1, so that "winners" (those who like the new restaurant) appreciate it only slightly more than the incumbent. Then, the only case in which a social planner would impose the new restaurant in the long run is if all friends turn out to be winners. If there is a loser (i.e. someone who dislikes the new restaurant), the very small utility gain achieved by winners does not compensate the disutility experienced by the loser. However, this policy is incompatible with majority voting, which would result in the two winners imposing the new restaurant despite the much larger magnitude of the loss incurred by the third, losing voter. This difference may result in all friends voting against the new restaurant when their preferences are still unknown, due to the loser-trap effect, while a social planner will try it to see whether all friends like it. In fact, Section 5 shows that, with positive news shocks, majority-based experimentation is always shorter than the utilitarian optimum.

3.2 Applications

The effects described in this paper can arise whenever decisions are made collectively and repeatedly. Although the examination of any particular application is beyond the scope of this paper, the reader may keep in mind the following contexts when thinking about the phenomena analyzed in the following sections.

Reforms with Unknown Winners and Losers Even when they benefit a significant fraction of the population, reforms usually harm some individuals or groups. Whenever the identity of these losers is ex ante unknown, the "loser trap effect" becomes a source of conservatism, as this paper shows. Conservatism resulting from the interaction between preference uncertainty and collective decisions has been studied, both empirically and theoretically, in the context of trade liberalization.[12] Fernandez and Rodrik (1991), in particular, are motivated by the behavior of industry groups who lobby against trade reforms ex ante, but a majority of which benefits from these reforms once they are implemented. They explain this paradox by showing that reforms having a negative expected value ex ante may turn out to be beneficial to the majority once implemented. This will be the case, for example, if a reform generates a few "losers," whom it severely hits, while providing small benefits to a majority of "winners." In such a scenario, when the identity of winners and losers is a priori unknown, the reform is initially opposed by all, but eventually gains the support of a majority as the identity of winners and losers gets revealed.

[12]See Baldwin (1985) and Bhagwati (1988).

Ex Ante Public Goods Section 9 considers a general specification of preference uncertainty that encompasses a setting, symmetric to the benchmark setting, in which negative lump-sums occur if the risky action is bad, while the safe action quo entails a small cost compared to the payoff of the risky action if it is good. This setting is potentially useful to think about applications where coordination is valuable to avoid negative shocks whose effects may vary across countries in a way that is a priori unknown. One may think for example of international coordination of security policy or greenhouse gas emissions.

In such settings, the paper predicts two opposite regimes, depending on voters' initial beliefs. In the first regime, society experiments until, possibly, losers become numerous enough to impose the safe action. In the second regime, the risk of entering the first regime is so large ex ante that voters reject experimentation altogether. Both regimes are socially inefficient. The first regime pushes experimentation too far, as unsure voters ignore the disutility of losers.[13] In the second regime, voters pay a safety premium to avoid a phenomenon whose consequences remain largely unknown. Social efficiency would require a collective long-term commitment to an experimentation policy that depends on the observed consequences of the risky action. The paper also emphasizes the distinction between commitment to an observation-dependent policy and commitment to a fixed action. Commitment to a fixed action may not be sufficient to increase efficiency, as it adds even more rigidity to the decision process.

Coordination Breakdown Winner frustration has some resemblance to the holdup problem studied in contract theory, and can similarly lead to coordination breakdowns. At an abstract level, an "involuntary" holdup may occur ex post if preferences concerning a joint initiative turn out to be different. In that case, the disappointed party can cancel the initiative at the expense of the party who turned out to benefit from it. Contrary to a standard holdup problem, winner frustration results from revelation of preference gaps and can thus only occur through experimentation. Winner frustration is thus better suited to capture contexts where learning is significant, such as joint research projects between firms, laboratories, or co-authors.

4 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 discounted

[13] Allowing for transfers across voters would alleviate such inefficiency, although the transfers would have to be from losers to unsure voters, which may be undesirable or unethical in practice.

at rate $r > 0$. There is an odd number $N \geq 1$ of individuals who continually decide at 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 or bad) are independently distributed across the group (the case of correlated types is considered in Section 8). 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 λ. The arrival of lump-sums is independent across individuals. The magnitude of these lump sums[14] 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 be 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 13). In particular, the arrival of the first lump-sum to a given individual i makes him publicly a "sure winner". At any time t, therefore, the group is divided into k "sure winners" for whom R is good with probability one, 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 obeying the dynamic equation $dp/dt = -\lambda p(1 - p)$ if no lump-sum is observed, with p_j jumping to 1 if some voter j receives a lump sum.[15] Type independence implies that an unsure voter only learns from his payoff stream but not from others'.

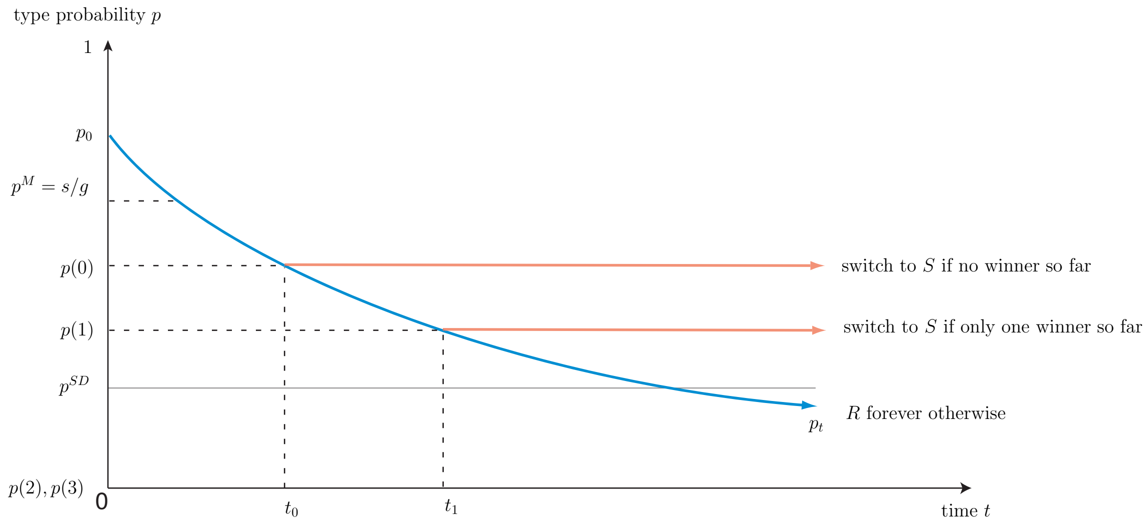
When $N = 1$, the setting reduces to the optimization problem of a single decision maker. The optimal experimentation strategy is Markov with respect to the current belief p, determined by a cut-off p^{SD} such that R is played if and only if $p \geq p^{SD}$. This cut-off is determined by an indifference condition (derived more generally in the proof of Theorem 1)

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

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

[14] 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 be good or bad. See Section 9 for a general specification of payoff distributions and beliefs.

[15] The simplest way to derive this dynamic equation is to observe that p_t is a martingale and jumps to 1 with probability rate $p\lambda$.



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 3 Voters.

For a group using the simple majority rule, the formal analysis to follow shows that collective decisions are determined by nonincreasing cut-offs $\{p(k)\}_{0 \leq k \leq N}$ 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 stops if no winner has been observed by then or continues until another threshold $p(1) < p(0)$ is reached, etc. These dynamics can are qualitatively represented by Figure 1 for the case of three voters.

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 $d\pi^i_R(\tau) = h dN^i_\tau$ or 0 depending on whether R is good or bad for i and $\{N^i\}_{1 \leq i \leq N}$ is a family of independent Poisson processes with intensity λ.

At any time t, within each category of voters (sure winners or unsure voters), individuals have the same value function since their payoffs are identically distributed. Let $w^{k,C}$ and $u^{k,C}$ respectively denote the value functions of sure winners and unsure voters where superscripts indicate the current number k of sure winners and the rule C that is followed. Letting $k_N = (N-1)/2$, winners have the majority if and only if $k > k_N$.

DEFINITION 1 *C is a* Majority Voting Equilibrium *(MVE) if for all t, it satisfies the following conditions:*

- *if $k_t \leq k_N$, C solves*

$$u_t^{k_t,C} = \sup_\theta E_t \left[\int_t^\sigma e^{-r(\tau-t)} d\pi^u_{\theta_\tau}(\tau) \right. \\ \left. + e^{-r(\sigma-t)} \left(\frac{1}{N-k_t} w_\sigma^{k_t+1,C} + \frac{(N-k_t-1)}{N-k_t} u_\sigma^{k_t+1,C} \right) \right], \quad (2)$$

- *if $k_t > k_N$, C solves*

$$w_t^{k_t,C} = \sup_\theta E_t \left[\int_t^\sigma e^{-r(\tau-t)} d\pi^w_{\theta_\tau}(\tau) + e^{-r(\sigma-t)} w_\sigma^{k_t+1,C} \right], \quad (3)$$

where σ is the first (possibly infinite) time at which a new winner is observed, and θ is any policy.

This definition means that at any time, the subgroup with the majority follows the policy that is optimal for itself, until a change occurs in the composition of the subgroups. When unsure voters have the majority, the conditional probability that any given unsure voter be that new winner is simply $1/(N-k_t)$, since there are $N-k_t$ unsure voters with identical payoff distributions. This explains the last term in (2). This definition extends to a non-Markov setting the standard notion of majority voting equilibrium for dynamic Markov policies (see for example Roberts (1989)). In particular, if one imposed at the outset that the collective decision rule only depend on the state (k,p), the above equations then reduce to the following Hamilton-Jacobi-Bellman (HJB) equations (to be explained in detail shortly):

- If $k \leq k_N$, $C = R$ if and only if

$$pg + \lambda p[w^C(k+1,p) - u^C(k,p)] \\ + \lambda p(N-k-1)[u^C(k+1,p) - u^C(k,p)] - \lambda p(1-p)\frac{\partial u^C(k,p)}{\partial p} > s \quad (4)$$

- If $k > k_N$, $C = R$ if and only if

$$g + \lambda p(N-k)[w^C(k+1,p) - w^C(k,p)] - \lambda p(1-p)\frac{\partial w^C(k,p)}{\partial p} > s \tag{5}$$

The equilibrium concept also corresponds, in a dynamic setting, to the elimination of weakly dominated strategies: it is the outcome obtained if each individual, at each time, votes for the action that maximizes his value function, given the policy, as if he were pivotal.[16]

Theorem 1 shows that there exists a unique[17] majority voting equilibrium, that this equilibrium has the Markov property, and that it is characterized by cut-offs. Equilibrium uniqueness is noteworthy and comes from a backward induction argument on the number of winners. Here is some intuition for the proof, assuming for now the Markov property. At any time t the state of the group can be summarized by k_t and p_t. Each voter category (sure winners or unsure voters) consists of individuals with perfectly aligned interests. If sure winners have the majority, they optimally impose R. If unsure voters have the majority and are very pessimistic about their type (p near zero), they impose S. Since an unsure voter can become a winner but the reverse is false, majority can only shift from unsure voters to winners. Starting with a majority of unsure voters, decisions are dictated by unsure voters' interest until they (possibly) lose the majority. The main question is therefore to determine unsure voters' preferences. These preferences are assessed by the following Hamilton-Jacobi-Bellman (HJB) equation:

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

The first part of the maximand corresponds to action R, the second 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 λ 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. Independence of the Poisson processes governing individual payoffs implies that only one lump-sum can be received during any infinitesimal period of time,

[16] The concept rules out trivial Nash equilibria, such as equilibria in which all individuals vote for the same dominated action. It also gets rid of some subtleties specific to continuous games identified by Simon and Stinchcombe (1989).

[17] As usual for the continuous-time stochastic control literature, uniqueness of the optimal policy is understood up to a subset of times of measure 0 on which actions can take any possible values without affecting value functions.

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 decided to stop experimentation if p becomes too low. They do so when the R part of (6) equals s. At such level p, the smooth pasting condition implies that the derivative term vanishes since the value function is constant, equal to s/r, below that level (see for example Dixit, (1993). This determines the equilibrium policy's cut-offs as state by Theorem 1, whose proof is in the appendix.

Theorem 1 (Existence and Uniqueness) *There exists a unique MVE. This equilibrium is characterized by cut-offs $p(k)$, $0 \leq k \leq N$, such that R is chosen in state (k, p) if and only if $p > p(k)$.*

The next result, Theorem 2, states that these cut-offs are decreasing in k: the larger the number of winners, and 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 that they lose majority and be imposed R forever increases? The intuition is as follows. Suppose that p is below the myopic cut-off p_M but above $p(k)$ 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, the only reason why they choose to experiment is that they hope to become 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 being 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.[18] Theorem 2 also states that $p(k) > p^{SD}$ for all $k \leq k_N$: 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 winner, he will be able to enjoy the risky action, and 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 be imposed the risky action forever if experimentation lasts long enough to reveal a majority of winners. This twofold control loss prompts unsure voters to experiment less than anyone of them would if he could dictate decisions in the future.

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

Theorem 2 (Cut-Offs Relations) *Equilibrium cut-offs satisfy the following relations:*

- $p^M > p(0)$.
- $p(k) > p(k+1)$ *for* $k \leq k_N$.
- $p(k_N) \geq p^{SD}$ *with strict inequality if* $N > 1$.
- $p(k) = 0$ *for* $k > k_N$.

Proof. In the appendix.

In fact, it is possible to prove a stronger[19] result than cut-off monotonicity: when a new winner the *value function* of both winners and unsure voters jumps upwards, 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 cut-off monotonicity of Theorem 2. More surprising is the fact that the occurrence of a new winner results in an upward jump of unsure voters' value function, unless this new winner is the decisive voter that gives the majority to winners. The intuition here is that new winners reduce the risk of winner frustration which predominates as long as unsure voters keep control of the decision process.

Theorem 3 (Value Function Monotonicity) *The following holds:*

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

Proof. See the appendix.

[19] This result is use to analyze the case of privately observed payoffs, see Theorem 12.

When learning is extremely fast, so that types are immediately revealed whenever R is tried, a single-decision maker is always willing to experiment until he learns his type (almost) perfectly.[20] However, this result does not extend to the case of collective experimentation: 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, this will happen if the total number N of individuals is large enough and the initial probability p is low enough: experimentation cut-offs stay bounded away from 0 as learning intensity λ goes to infinity, provided that N large enough. The proof is a direct consequence of (16) in the appendix.

Corollary 1 (Immediate Type Revelation) *If* $N > 2g/s - 1$,

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

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

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

Corollary 1 suggests that the total number N of individuals plays an important role 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 S's. To state the result, let $p(k, N)$ denote the experimentation cut-off 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 (15) in the appendix. For the second part, (15) also implies that $p(k_N, N) \to s/g = p^M$ as N goes to infinity. To conclude the proof, observe that from Theorem 2, $p(k_N, N) \leq p(k, N) \leq p^M$ for fixed k and all $N \geq 2k + 1$. Taking the limit as N goes to infinity proves the result. ■

Figure 2 shows the numerical computation of cut-off policies for different values of N and of the number $\kappa = k_N + 1 - k$ of switches required for winners to gain the majority. In general, cut-offs $p(k, N)$ are *not* monotonic with respect to group size, as can be proved by numerical

[20] Mathematically, this result comes from the single-decision maker cut-off equation (1): as the intensity λ goes to infinity, μ goes to 0 and so does the cut-off p^{SD}.

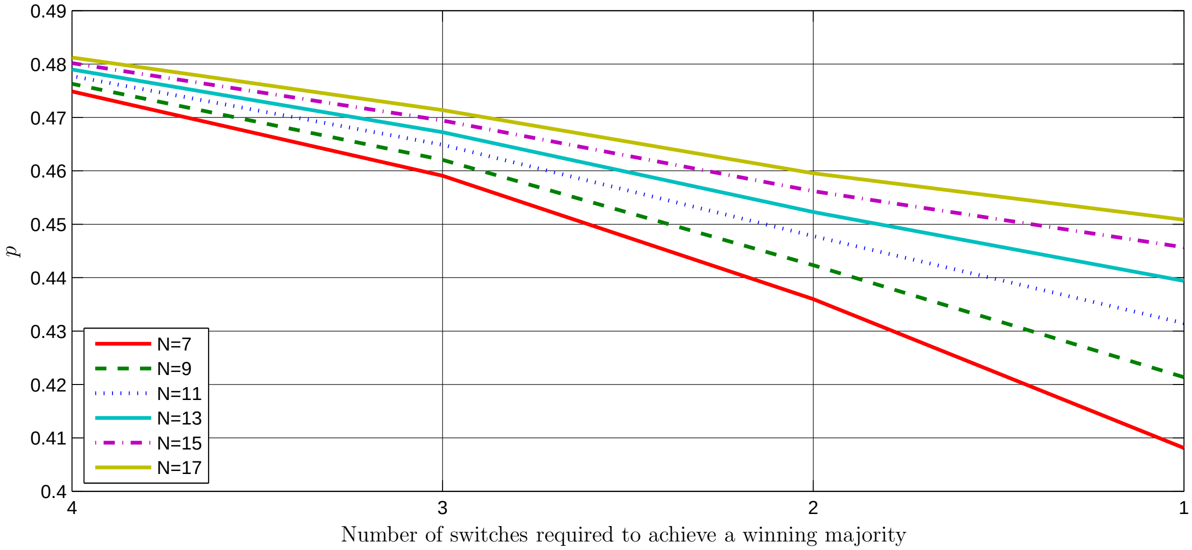


Figure 2: Policy Cut-Offs. $r = 1$, $\lambda = 1$, $s = 1$, $g = 2$.

counter-example. Such violations may seem counter-intuitive: as N increases, individual power gets more diluted, so shouldn't this reduce the value of experimentation? However, keeping k fixed, adding unsure voters increases the expected number of winners, hence the expected duration of experimentation, keeping cut-offs fixed. The addition of voters thus reduces the risk of winner frustration, which may increase the attractiveness of experimentation.

Proposition 1 is also related to 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 cut-offs are therefore defined as before, replacing k by the number of winners within the oligarchy, and 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 cut-off as a single decision maker.[21]

5 Utilitarian Policy

This section characterizes the optimal experimentation policy of a utilitarian social planner and shows that it lasts longer than majority-based experimentation.

[21]This however does not hold when types are correlated, as in Section 8.

Theorem 4 *Under the utilitarian criterion, the optimal policy is determined by cut-offs $q(k)$ such that $C(k,p) = R$ if and only if $p \geq q(k)$. These cut-offs are non-increasing in k, with $q(k) = 0$ if*

$$k \geq \bar{k} = \frac{s}{g}N.$$

Proof. See the appendix.

The next result shows that majority experimentation is inefficiently short compared to the utilitarian optimum. As the proof below illustrates, the result relies on the two effects described in the introduction: i) the social planner values information more than any individual voter, and ii) the social planner takes into account winners utility while unsure voters don't whenever they impose the safe action under the majority rule.

Theorem 5 (Majoritarian vs. Utilitarian Rules) *$q(k) \leq p(k)$ for all $k \leq k_N$.*

Proof. The utilitarian cut-off $q(k)$ solves

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

where W is the utilitarian value function, while the majority-voting cut-off, $p(k)$, solves

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

where $\bar{w}$ and $\bar{u}$ are the value functions obtained under the majority rule. Optimality of the utilitarian policy implies that for all k, p, $\frac{W(k,p)}{N} \geq (k/N)\bar{w}(k,p) + (1-k/N)\bar{u}(k,p)$. Since $\bar{w} > \bar{u}$, this also implies that $\frac{W(k+1,p)}{N} > 1/(N-k)\bar{w}(k+1,p) + (1-1/(N-k))\bar{u}(k+1,p)$, and subsequently that the left-hand side of (7) is higher than that of (8) for given p. Therefore, the root of the first equation must be lower than that of the second. ■

6 Commitment, Decision Frequencies, and Delegation

This section successively considers commitment to a state-dependent policy and commitment to a fixed action. While the first type of commitment always increases efficiency, the second prevents society from exploiting incoming information, resulting in a trade-off between value of commitment and value of experimentation. These results can be reinterpreted in terms of decision frequency and delegation.

If voters are initially homogeneous and can commit to an anonymous[22] policy at the outset, they share a common objective function. Since expected payoffs are identical, the optimal policy also maximizes the sum of these expected payoffs, i.e. utilitarian welfare. This shows the following result.

Theorem 6 (Commitment) *If voters can commit to an anonymous policy at time 0, they choose the cut-off policy determined by cut-offs* $\{q(k)\}_{0 \leq k \leq N}$.

Theorem 6 shows that social efficiency can be partially restored if voters can to some extent commit to a state dependent policy. However, such ability should not be confused with commitment 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. Commitment to an action may be harmful because too rigid. For example, consider the extreme case in which voters must initially commit to an action for the entire time horizon. In such circumstance, the risky action is chosen if and only if its expected payoff is above the myopic cut-off. 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 extent that the resulting policy is more inefficient.

The analysis suggests that, in the context of experimentation, commitment to a fixed action generates a trade-off between value of experimentation and value of commitment. Are there cases where committing to a fixed action is beneficial? The next example answers affirmatively, provided that aggregate uncertainty is small enough and initial beliefs are optimistic enough.

With a very large (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 is almost surely a fraction p of winners. Put differently, the law of large numbers implies that there is no learning at the aggregate level. 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

[22] 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.

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 cut-off $p^M = s/g$. This situation will occur almost surely if one starts with $p = p^M + \varepsilon$ for small ε for small enough.[23] This shows that commitment to the risky action is strictly more efficient than no commitment.

As another interpretation, suppose that voters can temporarily transfer the decision process to a delegate who makes decisions based on a mixture of electoral and welfare concerns. Such delegation can improve efficiency to the extent that welfare enters the delegate's objective, as he is able to adapt to incoming information. However, delegating action to a representative who prefers one action over another is equivalent to commitment to a fixed-action and subject to the same rigidity.

7 Supermajority Rules

This section analyzes how the choice of a voting rule affects experimentation, and shows that there is no systematic advantage of one voting rule above another.[24] 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 one of the two risks entirely vanishing when unanimity is required. Depending on the parameters of the model, which determines 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.[25]

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 will always prompt to choose immediate type revelation. However, once types are revealed, unanimity requires that R is only implemented if all voters are winners,

[23]Form Proposition 5 of 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 $\varepsilon < (g - s)/2g$.

[24]The analysis focuses on so-called "q-rules" or supermajority rules. If voting weights were allowed to vary with time according to news arrival, this would amount to a form of state-dependent commitment, and could clearly improve efficiency.

[25]For any supermajority rule one may, proceeding as in Section 4, prove the existence of a unique voting equilibrium characterized by monotonic cut-offs contained in $[p^{SD}, p^M]$. The analysis of this section, based on immediate type revelation, does not require this proof.

which typically is inefficiently too restrictive. Indeed, the social optimum is to get immediate type revelation and then choose the risky action if and only $kg > sN$. For $\nu \in \{1, \dots N\}$, define the ν voting rule as the rule requiring ν votes for the risky action. Letting $\nu^U = (sN)/g$, a ν 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 ν for $\nu \geq \nu^*$, because on this range ν is high enough to prompt experimentation and the probability of implementing the risky action if it is socially efficient ex post is decreasing in ν, while the probability of implementing the risky action if it is inefficient is zero. As is easily checked, ν^* can take any value between 1 and N ($\bar{\nu}$ decreases from N to 1 as p increases from 0 to 1).

To generate the reverse inefficiency ranking, suppose that, in addition to immediate type revelation, p is 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 ν 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 ν. Since ν^U can take any value between 1 and N, this implies the following result.

Theorem 7 *For any voting rules $\nu \neq \tilde{\nu}$, there exist parameter values and an initial belief p such that the ν voting rule is strictly socially more efficient than the $\tilde{\nu}$ voting rule.*

Intuitively, 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 ν^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 cut-offs characterizing to the unanimity-voting policy.

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

Proof. Equation (15) in the appendix implies that

$$p(1) = \frac{\mu s}{\mu g + (g - s) - (s - pg)} \sim \frac{\mu s}{(\mu + 1)g} \tag{9}$$

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

$$pg + \lambda p[w(2, p) - s/r] + \lambda p[v^{SD}(p) - s/r] = s, \tag{10}$$

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 (10) imply that $\chi(1)$ must satisfy the inequality

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

or

$$p \geq \frac{\mu s}{\mu g + 2p(g-s)} \sim s/g \tag{11}$$

if $g \gg s$. Comparing (9) and (11) shows that $\chi(1) > p(1)$. ■

8 Correlation and Heterogeneity

This section considers the case of two voters, 1 and 2, who share a common belief about the initial joint distribution of their types, although this distribution may be asymmetric and exhibit type correlation across voters. Let θ^i denote Voter i's type, and let $p^{\vartheta_1\vartheta_2} = Pr[(\theta^1,\theta^2) = (\vartheta^1,\vartheta^2)$, where $\vartheta^i \in \{g,b\}$ describes the possible types (good or bad) of each voter. Also let $p^i = Prob[\theta^i = g]$ for $i \in \{1,2\}$, and $\alpha = p^{gg}/(p^1p^2)$. α is a measure of the correlation between voter types.[26] Let Δ denote the set of (p^2,α) that are achievable as elementary probabilities vary over the four-dimensional simplex. The following proposition is a simple exercise of Bayesian updating, whose proof is easy and omitted.

Proposition 2 (State Dynamics) *Beliefs are governed by the following dynamics equations. When no lump-sum is observed,*

- $\frac{dp^{gg}}{dt} = -\lambda p^{gg}(2 - p^1 - p^2)$
- $\frac{dp^{bb}}{dt} = \lambda p^{bb}(p^1 + p^2)$
- $\frac{dp^{gb}}{dt} = -\lambda p^{gb}(1 - p^1 - p^2)$, $\frac{dp^{bg}}{dt} = -\lambda p^{bg}(1 - p^1 - p^2)$
- $\frac{d\alpha}{dt} = -\lambda\alpha(1-\alpha)(p^1 + p^2)$

[26] The standard correlation measure and α have a one-to-one relationship for any given p^1 and p^2. If $\alpha = 1$, types are uncorrelated. In general, α takes values in $\mathbb{R}_+$, although not all values of $\mathbb{R}_+$ are achievable for given p^1, p^2. For example, $p^1 = 1$ implies that $\alpha = 1$, since in that case Voter 1's type is deterministic hence uncorrelated with Voter 2's type.

When Voter 1 receives a lump-sum,

- $p_+^{bb} = 0$, $p_+^{gb} = \frac{p^{gb}}{p^1}$, $p_+^{bg} = 0$, $p_+^{bb} = \frac{p^{bb}}{p^1}$
- $\alpha_+ = 1$, $p_+^1 = 1$, $p_+^2 = \alpha p^2$

where the subscript '+' denotes values immediately after the lump-sum is observed, and its absence denotes values immediately before the lump-sum. Symmetric formulas if instead Voter 2 receives a lump sum.

Suppose that R requires unanimity. Since voters may now be heterogeneous (i.e. $p^1 \neq p^2$), the equilibrium concept must be modified. In the spirit of Section 4, and given the unanimity rule, it is natural to assume that the voter who is the less likely of being a winner is in control: if that voter wants to play the risky action, so should the player with a higher expected type. This notion is also consistent with elimination of weakly dominated strategies, because the pivotal voter is always the voter who wishes to stop experimentation. For simplicity, let us therefore define a unanimity equilibrium (UE) as follows: at any time t, if $p^i \leq p^j$, then j votes for R whenever i does.

Theorem 8 *There exists a unique UE. This equilibrium determined by a cut-off function $\delta : \Delta \rightarrow [0,1]$ such that $C(p^1, p^2, \alpha) = R$ if and only if $p^1 > \delta(p^2, \alpha)$ whenever $p^1 \leq p^2$, with the reverse relation if $p^1 > p^2$.*

Proof. See the appendix.

Theorem 9 states that a voter's incentive to experiment increases both with the other voter's probability of being a winner and with voters' type correlation. Intuitively, if types are more positively correlated, the risk of winner frustration decreases. The risk of winner frustration is also lower for a given voter if the other voter is more likely to be a winner. In the extreme case in which, say, Voter 2 is a sure winner (i.e. $p^2 = 1$), Voter 1 has full control over collective decisions, and can behave in effect as a single-decision maker. In addition, positive correlation increases the speed of learning, which reduces the time cost of experimentation. In the extreme case of perfect type correlation, the setting is equivalent to one with a single decision maker with twice the initial learning intensity. Figure 3 shows numerical computations of the experimentation boundary for several values of the correlation measure α.

Theorem 9 *δ is decreasing in both components.*

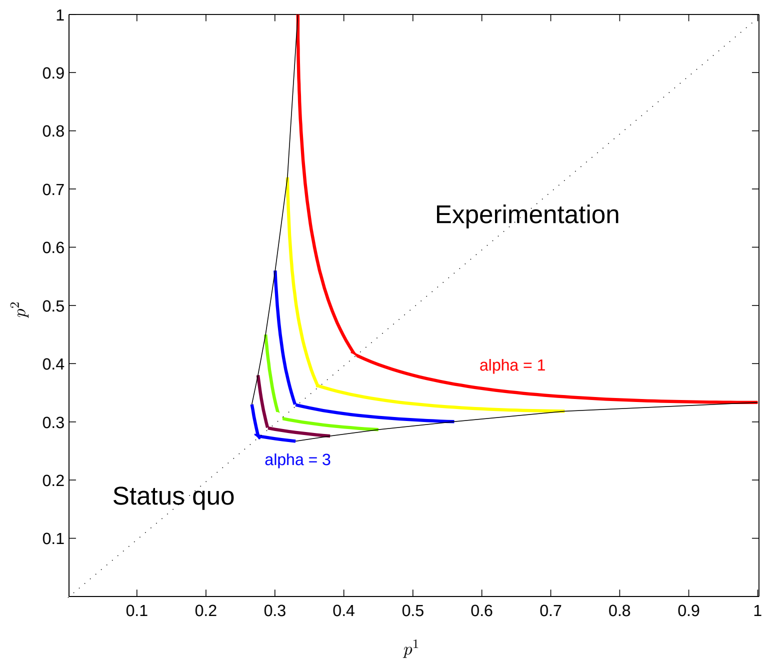


Figure 3: Experimentation Boundary δ as a function of α. $r = 1$, $\lambda = 1$, $s = 1$, $g = 2$.

Proof. The left-hand side of (22), is increasing in p, p^2 and α. Therefore, keeping α fixed, the root $\delta(p^2, \alpha)$ must be decreasing in p^2, and similarly keeping p^2 fixed, $\delta(p^2, \alpha)$ must be decreasing in α. ■

9 General News Arrival Processes

Several potential applications of this paper require a different specification of news arrival than the one studied in the benchmark setting. In order to determine which of the previous results generalize to other news arrival processes, this section reformulates the model of collective experimentation as an abstract stochastic control problem. It then reconsiders three results of the previous sections: i) collective experimentation is shorter than the single decision maker equivalent, ii) collective experimentation is shorter than the utilitarian optimum,[27] iii) there is always some experimentation, i.e. a set of voter beliefs where R's immediate payoffs is lower than S's but society still elects R. The results of this section are then exploited to analyze a

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

setting with negative news shocks and another with mixed news shocks.

Suppose that, for any given individual, the risky arm has a payoff distribution, or "type", θ lying in some finite set Θ. At any time, that individual's belief about his type is summarized by a probability distribution or "state" $\gamma \in \Gamma$, where $\Gamma = \Delta(\Theta)$ is the set of all probability distributions[28] over Θ. 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 γ, $G(\gamma)$ is the smallest value of s for which the single decision maker prefers the safe action over experimentation. Mathematically, $G(\gamma)$ solves

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

where σ is any policy, and the expectation is conditional on the current state γ and on the rate s of the safe action.[29]

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 γ. If types are independent, then $\gamma = \left(\gamma^1, \ldots, \gamma^N\right) \in \Gamma^N$. In general, γ may contain information about type correlation, as in Section 8). 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

$$\sup_{\sigma} E\left[\int_0^\infty e^{-rt} d\pi^i_{\sigma_t}(t)|\gamma, s\right] \geq E\left[\int_0^\infty e^{-rt} d\pi^i_{C_t}(t)|\gamma, s\right]. \qquad (12)$$

The inequality obtains because C is an element of the policy set over which the maximization is taken.[30] We may define a policy-dependent generalization of the Gittins index as

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

Inequality (12) implies that $G^i_D(\gamma) \geq G^i_C(\gamma)$ for all i, γ, and C, where $G^i_D(\gamma)$ is i's Gittins index if he has dictatorial power over all decisions.

The definition of voting equilibria is extended as follows. Let ν denote any integer in $\{1, \ldots, N\}$

Definition 2 (Voting Equilibrium) *C is a ν-voting equilibrium if for all belief γ, $C = S$ if and only if the number of voters i such that $G^i_C(\gamma) \leq s$ is greater than or equal to ν.*

[28]In the benchmark model, the type θ is either "good" or "bad" and the state γ is the probability p that the type is good.

[29]The results of this section are easily adapted to discrete-time settings. In fact, Theorem 10 does not assume anything about the time domain.

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

The following result shows that collective experimentation is shorter than dictatorial experimentation, in the following sense: if there are at least ν 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 ν-voting equilibrium.

THEOREM 10 *Suppose that C is a ν-voting equilibrium. Then, $C = S$ whenever $|\{i : G_D^i(\gamma) \leq s\}| \geq \nu$.*

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

When types are independent, $G_D^i(\gamma) = G(\gamma^i)$ where $G(\gamma^i)$ is the Gittins index of the single decision maker problem with (individual) belief γ^i. In that case, i's optimal policy is independent of other individuals' types. As a corollary of Theorem 10, 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, as shown by Section 8. 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.

In the benchmark setting, Theorem 2 showed that all cut-offs were below the myopic cut-off, meaning that there is always 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 γ^i over the type space, let $g(\gamma^i) = E[d\pi_R^i/dt|\gamma^i]$. $g(\gamma^i)$ is i' immediate expected payoff rate with action R given type distribution γ^i. For any individual type θ^i, let, slightly abusing notation, $g(\theta^i) = g(\delta_{\theta^i})$, where δ_{θ^i} is the Dirac distribution concentrated on type θ^i, denote i's true immediate expected payoff rate with action R when his actual type is θ^i. Say that i is a *winner* if $g(\theta^i) > s$, and a loser otherwise. i is a winner if R really is optimal for him given his true type. Θ 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.[31]

Nonnegative type correlation would, intuitively, seem to rule out adversity. However, some perverse effect may occur even in that case. For example, in a setting with both positive and negative news shocks, unsure voters may, upon observing a sure loser, want to push experimentation further as the risk of the loser trap is reduced, which adversely affects this loser. Such example is constructed in Section 11. Fortunately, adversity is ruled out by important cases, for example when R requires unanimity, or in setting with negative news shocks only.[32] Finally, non-adversity also holds in the mixed-shocks setting of Section 11 as long as learning is fast enough relative to voters patience.

THEOREM 11 *Suppose that C is a voting equilibrium for voting rule υ. Then, $G^i_C(\gamma) \geq g(\gamma^i)$ for all i for which C is non-adverse.*

Proof. See the appendix.

10 Negative News Shocks

Several applications mentioned in Section 3 require a setting with negative news shocks. To accommodate such applications, suppose 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

[31]In that case, however, majority would simply ignore i and proceed with experimentation. As a stronger case of adversity, suppose 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, 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 cut off is $p^M = 1/3$, so p is above the myopic cut-off and yet voters prefer to avoid R since $\pi < s$. Section 11 provides an example of endogenous adversity.

[32]Indeed, with negative news shocks, R is continually played, if at all, until losers gain the majority. See Section 10.

for unsure voters. It can be shown that the policy is also determined by cut-offs $\rho(k)$ 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 obtain the majority, or because the cut-off $\rho(k_t)$ jumps over p_t upon the observation of a new loser (cut-off variation is discussed below). Theorem 10 implies that, provided types are independent, $\rho(k) \geq \rho^{SD}$ for all $k < N/2$, where ρ^{SD} is the single-decision-maker cut-off.

The equilibrium policy resulting from the majority rule is non-adverse to any voter. Indeed, suppose that the risky action is elected by unsure voters at some time t. Then, the (possibly infinite) time at which the risky action is abandoned necessarily decreases with the number of sure losers since as time passes, the belief of unsure voters gets better and better. Therefore, for any unsure voter i, $Pr[\theta^i \text{ good}|C_t = R] \geq Pr[\theta^i \text{ bad}|C_t = R]$, which shows non-adversity of the equilibrium policy.[33] Theorem 11 then implies that all cut-offs lie below the myopic cut-off.

Section 5 showed that, in the benchmark setting, majority-based experimentation is inefficiently short compared to the utilitarian optimum. With negative news shocks, the reverse intuition holds to some extent: unsure voters, ignoring losers, may push experimentation further than a utilitarian social planner. Another effect arises, with an opposite direction. The value of experimentation is generally higher for the social planner than for any individual due to control sharing effects. For example, a social planner would always welcome immediate type revelation, whereas voters can shun it provided 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 before unsure voters under majority voting.

In view of Theorem 2, one may naturally wonder whether cut-offs are also monotonic in this negative news shock setting. The answer is negative. Counter-examples 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

[33] See the proof of Theorem 11 for the equivalence between this statement and non-adversity.

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.[34]

11 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 λ_g, and if it is bad, it pays negative lump sums according to the jumping times of a Poisson process with intensity λ_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[35], and that they stop experimentation at some cut-offs $p(k^W, k^L)$, provided they are pivotal. Theorem 10 implies that $p^{SD} \leq p(k^W, k^L)$, where p^{SD} is single decision maker setting cut-off. This inequality holds for all ν voting rules. If the risky action requires the unanimity rule, then Theorem 11 implies that $p(k^W, k^L) \leq p^M$, where p^M is the myopic cut-off: unanimity guarantees at least some experimentation.

Negative Value of Experimentation With other voting rules, non-adversity 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*, i.e. 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 them two. This increases their willingness to experiment.

[34] From a technical viewpoint, 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 cut-off. Instead, cut-offs are determined by direct comparison of value functions with and without starting experimentation.

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

If a winner is observed, the remaining two unsure voters are now on the brink: any winner among them will impose the risky action to the other. This risk reduces their willingness to experiment. Therefore, ex ante, the three voters know that if any 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 requires 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. Let $g > 0$ and $b < 0$ the expected payoff rates of the risky arm for sure winners and sure losers respectively. Let p^M, p^{SD}, p^L, p^W, p^3, respectively denote the myopic cut-off, the single decision maker cut-off, the two unsure voters' cut-off when the third voter is a loser, the two unsure voters' cut-off when the third voter is a winner, and the experimentation cut-off when all three voters are unsure. For the following parameter values $g = .1$, $b = -1$, $s = 0$, $r = 1$, $\lambda_b = .1$, $\lambda_g = .11$, cut-offs have the following values:

p^M	p^{SD}	p^L	p^W	p^3
.9091	.9001	.9016	.9083	.9095

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 than p^W,[36] meaning that if a voter is a loser, experimentation lasts much longer than if he is a winner.

As learning becomes faster (relative to patience), however, an argument similar in spirit to non-adversity implies that the value of experimentation is positive. Indeed, suppose that either λ_b or λ_g is arbitrarily large, so that R (almost) immediately reveals voter types. The expected value of the risky action for any unsure voter i, given that unsure voters are ex ante pivotal (i.e. $\max\{k^W, k^L\} < N/2$), is

$$\begin{aligned} V^i = Pr[k_+^W > N/2]\left(g/rPr[\theta^i \text{ good}|k_+^W > N/2] + b/rPr[\theta^i \text{ bad}|k_+^W > N/2]\right) \\ + Pr[k_+^L > N/2]\left(g/rPr[\theta^i \text{ good}|k_+^L > N/2] + b/rPr[\theta^i \text{ bad}|k_+^L > N/2]\right), \quad (13) \end{aligned}$$

where k_+^W (resp. k_+^L) denotes the number of winners after all types are revealed. Clearly, $Pr[\theta^i \text{ good}|k_+^W > N/2] > Pr[\theta^i \text{ good}] > Pr[\theta^i \text{ good}|k_+^L > N/2]$. This implies that

$$V^i > g/rPr[\theta^i \text{ good}] + b/rPr[\theta^i \text{ bad}] = pg/r + (1-p)b/r,$$

which is the myopic payoff. Therefore, Voter i is willing to experiment at least until p drops below the myopic cut-off p^M, defined by $p^M g/r + (1-p^M)b/r = s = 0$.

[36] Indeed, p^L is close to the single decision maker cut-off, while p^W is close to the myopic cut-off.

12 Privately Observed Payoffs

Previous sections assumed that all payoffs were publicly observable. What happens if payoffs are privately observed? Perhaps surprisingly, this need not affect the equilibrium policy in the context of Section 4. The intuition has two parts. Suppose voters can communicate their types, through their voting decisions or through cheap talk. Winners always benefit from revealing their type because this increases experimentation, by the cut-off monotonicity result of Theorem 2. Unsure voters cannot gain by manipulating the choice process because, conditional on being pivotal (i.e. $k \leq k_N$), they are already choosing their optimal action.

Voters can resort to several channels to communicate their type. Cheap talk is one possibility. Voters can simply communicate through their votes. Here is a natural protocol: each time a cut-off is reached at which unsure voters would want to stop given the number of winners that they last observed, 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 can thereafter decide whether to pursue experimentation to the next cut-off (everyone votes for the risky action), or impose the safe action (unsure voters vote for the safe action). With this protocol, voters know the current number of winners only when p reaches particular cut-offs, but that suffices to implement the policy of the public-information setting.[37]

The proof that the above protocol implements the public-information policy is sketched in the appendix.

Theorem 12 *The above protocol yields the same equilibrium as the experimentation policy based on publicly observed payoffs.*

13 Extensions and Discussion

Multiple Risky Actions

With multiple correlated risky actions, even a small probability of modification in the preference ranking of some voter can create a disagreement among members who could otherwise impose a better action for themselves.

[37] Other less natural protocols would be more informative, for example if a new winner reveals his type change by voting for S for an infinitesimal amount of time.

Consider three individuals voting at the majority rule. Suppose for now that there are two actions: S pays a constant rate $s = 1$ to all. R pays a certain rate of 2 to Voters 1 and 2, but has an uncertain payoff distribution for Voter 3: with probability p it gives off lump-sums to Voter 3 with expected payoff is $g = 0.1$, and gives 0 otherwise. There is a unique equilibrium: Voters 1 and 2 impose R forever, independently of p. For what follows, it is worth noting that, even if p were equal to 1, Voter 3 would prefer S.

Now suppose that there is another risky action, X, such that i) X is good for Voters 2 and 3 if and only if R is good for Voter 3, ii) X surely pays -9 to Voter 1, and has expected payoffs $g_X^2 = 2.1$ and $g_X^3 = 1.1$ respectively to Voters 2 and 3 if it is good. Perfect correlation implies that any lump-sum observed by Voters 2 or 3 with action X or by Voter 3 with action R causes the common probability p to jump to 1.

To avoid ties, let us assume that S is imposed whenever no action receives a majority of votes. The following propositions show the impact on equilibrium of preference uncertainty. The proof of Proposition 3 is easy, as there is no learning, and omitted.

PROPOSITION 3 *If $p = 0$ or $p = 1$, there is a unique MVE: if $p = 0$ R is played forever, if $p = 1$, X is played forever.*

For the next result, suppose that $\mu = r/\lambda = 1$.

PROPOSITION 4 *If $p \in (0.1, 0.8)$, there is a unique MVE. In this MVE, S is played forever.*

Proof. Voter 1 is indifferent between R and S if p solves the equation

$$2 + \lambda p \left[-\frac{9}{r} - \frac{1}{r} \right] = 1,$$

which yields $\underline{p} = 0.1$. Thus if p is greater than $\underline{p}$, Voter 1 prefers the low payoff of S rather than risking that Voters 2 and 3 discover that X is good for them and imposing it forever, from Proposition 3. Voter 3's indifference equation relative to actions X and S is

$$p1.1 + \lambda p[1.1/r - 1/r] = 1,$$

which yields $\bar{p} = \mu/(0.1 + 1,1\mu) = 0.8$ if $\mu = 1$. Therefore, if $p \in (.1, .8)$, S is the preferred choice for 1 and 3, hence chosen forever. ■

An illustration: the Singaporean Restaurant To illustrate the content of these propositions, consider a variation of the restaurant example of Section 3. Three friends, Chris, Ian, and

Paul go the restaurant once every week-end, and choose each time their restaurant according to the majority rule. They start with the following preferences. Chris likes Chinese cuisine above anything else. Ian likes Indian cuisine but not Chinese one. Paul does not know Asian cuisine, and is thus uncertain about his preferences. There are three restaurants in town: i) a gourmet Chinese restaurant, clearly the best choice for those who like Chinese cuisine, ii) a Singaporean restaurant whose menu contains both Chinese and Indian dishes, and iii) a plain restaurant who has a well known common, relatively low value to all friends.

The following paradox may occur. When given the choice among the three restaurants, Ian and Paul impose the plain restaurant. However, if the gourmet Chinese restaurant closes down so that only two restaurants remain, Chris and Ian impose the Singaporean restaurant.

Why did Ian change his vote? In both cases, Ian prefers the Singaporean restaurant. However, if he agrees with Chris to go there, he runs the risk that Paul discovers that likes Chinese cuisine, resulting in Chris and Paul to impose the gourmet Chinese restaurant in the future. If this risk is high enough, Ian prefers to vote for the plain restaurant, which Paul also prefers if his expected value for Asian cuisines is low enough.

Factions and Heterogeneous Voting Weights. Given the results of this paper, it is natural to expect that voters with more decision weight will be more inclined to experiment longer. Although a detailed analysis of voting weights requires a separate paper, one already observe some important difference with the benchmark setting. Suppose that there are four voters with Voter 1 having two votes and made at the simple majority rule. Suppose that Voter 4 is the only sure winner so far. Then, 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 must then consider the risk that Voter 3 also becomes a winner resulting in a winning coalition that imposes R forever. Contrary to the benchmark setting, thus, experimentation can be interrupted by the occurrence of a new winner.

Why not a two-period model? Some results of this paper can be derived in a two period model. However, several key results of the paper, such as cut-off monotonicity, the impact of type correlation, and the possibility of a negative value of experimentation, rely on the impact of one's experimentation on other voters' future experimentation, so require at least three periods. In that case, an infinite horizon model may actually be simpler by virtue of time-homogeneity. Time-homogeneity also guarantees that cut-offs only depend on current beliefs about action value, not on proximity to the end of time. More generally, the paper analyzes reforms or projects where uncertainty is revealed gradually over time, so that one is interested in the impact of news arrival on decisions. Finally, some potential applications, such as joint R&D projects, can be seen as stopping games, where the time extension is an important part of the model. Some of the results in this paper can be interpreted as comparative statics about stopping.[38]

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. In a separate paper, I show how side-payments restore utilitarian efficiency in general experimentation settings with two actions (possibly both risky).

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 cut-offs.

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) show 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. What happens if unsure voters have heterogeneous type probabilities? As suggested by Section 8 for the case of two voters, heterogeneity may increase experimentation. 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 which are (almost) sure to be winners and 4 of which are (almost) sure to be losers. The remaining voter has (almsot) 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.

[38] Choosing $\lambda = \infty$ in effect reduces the model to two periods: before and after type revelation. That simplicity is used in Section 7.

14 Conclusion

In a dynamic setting, collective decisions tend to be too conservative compared to what any individual would choose if he had full control over future decisions. This control sharing effect can be decomposed into "loser trap" and "winner frustration" risks. Experimentation through voting exhibits a systematic conservative bias compared to the utilitarian optimum. This bias dominates when types are revealed fast enough, since a utilitarian planner facing lower time-cost of experimentation always wants to learn those types, whereas individuals may vote against uncovering the heterogeneity associated with unknown alternatives. That bias, however, coexists with another effect, whereby a pivotal group ignores the benefits or costs of the risky action incurred by remaining voters and, as a result, can impose more or less experimentation than is socially optimal. Control sharing effects may result in a collective experimentation policy that is *adverse* to *all* voters, even when types are independent. In such case, the value of experimentation is *negative*: society may reject a risky action whose expected payoff is higher than the status quo. Which voting rule performs better depends on the relative salience of loser trap vs. winner frustration: the higher the risk of loser trap and the more restrictive access to R should be. Commitment to a state-dependent policy restores full efficiency, but commitment to a fixed action may only increase efficiency if aggregate uncertainty is small enough. For large groups, control sharing effects can entirely annihilate the value of experimentation, causing individuals to vote myopically, but are reduced if actual preferences types are positively correlated. Section 9 shows how to reformulate the analysis as an abstract stochastic control problem, using a concept similar to the Gittins index, to allow for general news arrival processes. This formulation provides more insight and generality to several results of the paper. For some applications, such as the coordination of gas emission policies, it would be useful to modify the model of this paper to accommodate partial action irreversibility. The effects identified in this paper will also matter for such settings and can be summarized by the following rule of thumb: collective experimentation increases with the likelihood that *future* decisions will favor *current pivotal voters*.

15 Appendix

15.1 Proof of Theorem 1

Suppose that $k = N$, i.e. all voters are sure winners. Then, σ is necessarily infinite, so (3) reduces to

$$w_t^{N,C} = \sup_\theta E_t \left[\int_t^\infty e^{-r(\tau-t)} d\pi_{\theta_\tau}^w(\tau) \right].$$

The (essentially) unique solution is $C_\tau = R$ for all τ, since it provides winners at any time with the maximal possible expected payoff g. This gives them the constant value function $w_t^N = g/r$. The value function of unsure voters is also easily computed: if an unsure voter's type is good, which happens with probability p_t, he gets the same expected value as winners, g/r. Otherwise, he gets 0 forever. Therefore, $u_t^N = p_t g/r$. For $k = N-1$, (3) reduces to

$$w_t^{N-1,C} = \sup_\theta E_t \left[\int_t^\sigma e^{-r(\tau-t)} d\pi_{\theta_\tau}^w(\tau) + e^{-r(\sigma-t)} g/r \right],$$

where I use the fact that $w_t^N = g/r$. Again, the (essentially) unique solution is $C_\tau = R$ for all τ, value functions still equal $w_t^{N-1} = g/r$ and $u_t^{N-1} = p_t g/r$. By the same induction argument, $C_\tau = R$ and $w_t^k = g/r$ and $u_t^k = p_t g/r$ for all $k > k_N$. Now consider the case $k = k_N$, in which unsure voters have the majority, but only one new winner among them is needed for the majority to switch to sure winners. Then (2) reduces to

$$u_t^{k_N,C} = \sup_\theta E_t \left[\int_t^\sigma e^{-r(\tau-t)} d\pi_{\theta_\tau}^u(\tau) + e^{-r(\sigma-t)} \left(\frac{1}{N-k_N} \left(\frac{g}{r} + h \right) + \frac{(N-k_N-1)}{N-k_N} \frac{p_t g}{r} \right) \right], \tag{14}$$

using the relations $w^{k_N+1} = g/r$ and $u^{k_N+1} = p_t g/r$. The optimization problem (14) is formally identical to the optimization problem of a single decision maker, with known termination values. The solution of such problem is well known (see for example Fleming and Soner, 1993). The control is Markov in p, with any indifference threshold $p(k_N)$ determined by the smooth-pasting condition of the Hamilton-Jacobi-Bellman equation (6), which reduces to

$$pg + p\lambda(g/r - s/r) + p\lambda(N - k_N - 1)(pg/r - s/r) = s, \tag{15}$$

using the relation $u^{k_N+1}(p) = pg/r$. The left-hand side of (15) 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

$$p(k_N) = \frac{\mu s}{\mu g + (g-s) + (N-k_N-1)(p(k_N)g - s)}. \tag{16}$$

This shows that $C(p, k_N) = R$ if and only if[39] $p > p(k_N)$. 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 strategy entirely determines the value functions $w^C(k_N, p)$ and $u^C(k_N, p)$ of sure winners and unsure voters, which are in fact computable in closed-form by integration of their dynamic equation:

$$w(k_N, p) = \frac{g}{r} - \frac{g-s}{r}\left(\frac{1-p}{1-p(k_N)}\right)^{N-k_N}\left(\frac{\Omega(p)}{\Omega(p(k_N))}\right)^{\mu}, \tag{17}$$

and

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

for $p \geq p(k_N)$, where $\Omega(p) = (1-p)/p$. These functions are easily shown to be increasing in p, with $u^C(k_N, p) \geq pg/r$. Moreover, $u(k_N, p) = w^C(k_N, p) = s/r$ 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

$$pg + p\lambda(w(k_N, p) - s/r) + p\lambda(N - k_N - 2)(u(k_N, p) - s/r) = s. \tag{19}$$

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, $\frac{\partial u_C}{\partial p}(k_N - 1, p(k_N - 1)) = 0$. Therefore, (19) becomes $p(k_N - 1)g = s$, which contradicts the assumption that $p(k_N - 1) \leq p(k_N) < p^M$. Thus, necessarily, $p(k_N) < p(k_N - 1)$.

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 σ_S, unless a new winner is observed σ_W. 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

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

[39] As before, this is up to action changes on a time subset of measure 0.

where $q = Prob[\sigma_W < \sigma_S | p_t]$. We saw that $u(k_N, \cdot)$ and $w(k_N, \cdot)$ are 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 σ_S and given that $p(k_N) < p(k_N - 1)$. Also, p_σ is increasing in p_t given the Bayesian updating dynamics. Finally, σ_W is decreasing in p_t, since a higher p_t makes it more likely that a payoff will be observed.[40] This also implies that q is increasing in p_t by definition of q and the fact that σ_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 σ_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 $\sigma_1 < \sigma_2$ the arrivals times of lump-sum to the next two new winners. As is easily shown, these 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 σ_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 remaining 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 non-increasing for $k' > k$, and iii) the resulting value functions $u(k', \cdot)$ and $w(k', \cdot)$ are non-decreasing in p. The general induction step is then proved exactly as above.

15.2 Proof of Theorem 2

Theorem 1 already shows 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 (16) 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

$$p(0)g + p(0)\lambda(w(1, p(0)) - s/r) + p(0)\lambda(N-1)(u(1, p(0)) - s/r) = s. \tag{20}$$

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

[40]Conditional on p_t, σ_W is the mixture of exponential variables with intensity λj, $j \in \{0, \ldots, N - k_N + 1\}$, 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.

on the left-hand side of (20) are positive, which implies that $p(0)g < s$, or equivalently that $p(0) < p^M$.

15.3 Proof of Theorem 3

Monotonicity of u and w with respect to p was shown as part of the induction hypothesis in the proof of Theorem 1. 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 $\tilde{k} \geq k - 1$:

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

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

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

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

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:

$$(r + \lambda p(N - k + 1))\psi(p) = \lambda p[\phi^w(p) + (N - k - 1)\phi(p)] - \lambda p(1 - p)\psi'(p). \tag{21}$$

We already know that ϕ^w is positive. Therefore, if ϕ were also nonnegative, the argument we just used for w would also show that ψ is nonnegative. In particular, if one can show that $u(k_N, p) \geq u(k_N - 1, p)$, a backward induction will prove the result for all $k \leq k_N$. Combining (17) and (18) implies that, for $k = k_N$,

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

Therefore, the left-hand side has the sign of $g - s - (N - k_N - 1)(s - p(k_N)g)$. From the cut-off formula (15), this latter term has the same sign as $s - p(k_N)g$, which is positive. Therefore, we can apply the first term in the right-hand side of 21 is nonnegative for $k = k_N$, which implies that ψ is nonnegative for $k = k_N$. This fills the missing step of the induction, concluding the proof that u is increasing in k for $k \leq k_N$.

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 (18).

15.4 Proof of Theorem 4

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 σ 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 remaining 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 non-increasing for $k' > k$, and iii) resulting value functions $w(k', \cdot)$, $u(k', \cdot)$, and $W(k', \cdot)$ are non-decreasing in p.

15.5 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 5 $p_S = (p - q)/(1 - q)$.

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 λ. Combining the previous formulas concludes the proof. ∎

15.6 Proof of Theorem 8

First suppose that $p^2 = 1$. Then voter 1 has full control over the collective decision. He therefore imposes his optimal policy, which is that of a single decision maker. This defines $\delta(1,1) = p^{SD}$. This also fully determines the value functions of both voters in that case. Let $p \mapsto w(p)$ denote the value function of voter 2, where p is voter 1's probability of being a winner is p, and v^{SD} is the value function of a single decision maker, which is also voter 1's value function in this case. More generally suppose that at time 0, $p_0^1 \leq p_0^2$. It follows from Proposition 2 that $p_t^1 \leq p_t^2$ for all t preceding the first arrival of a lump-sum. In particular, this implies that 1 has full control of the collective decision (under unanimity) over that period. Therefore, he chooses a policy θ that solves

$$u_t = \sup_\theta E\left[\int_t^\sigma e^{-r(\tau - t)} d\pi^1_{\theta_\tau}(\tau) + e^{-r(\sigma - t)}\left(qw(p^2_{\sigma+}) + (1-q)v(p^1_{\sigma+})\right)\right],$$

where, letting σ_i denote the (possibly infinite) time at which i receives his first lump sum, $\sigma = \min\{\sigma_1, \sigma_2\}$ and $q = Prob[\sigma_1 < \sigma_2]$. This is a standard control problem, whose solution is known to be Markov. Voter 1 is indifferent between R and S at probability level p, if p solves the equation

$$pg + \lambda p[w(\alpha p^2) - s/r] + \lambda p^2[v^{SD}(\alpha p) - s/r] = s. \tag{22}$$

The left-hand side is increasing in p, equal to 0 for $p = 0$ and greater than $g > s$ if $p = 1$. Therefore, it has a unique root $\delta(p_2, \alpha)$. This shows that $C(p^1, p^2, \alpha) = R$ if and only if $p^1 > \delta(p^2, \alpha)$. The case $p^1 > p^2$ obtains by symmetry.

15.7 Proof of Theorem 11

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 γ.

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

Therefore, if $E[d\pi_{C_t}^i(t)|C_t = R] > sdt$ 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, $E[d\pi_R^i(t)] = g(\gamma^i)dt > sdt$. Moreover, C's non-adversity 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 non-adverse 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.[41] Moreover,

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

Combining these results yields the inequality.

15.8 Proof of Theorem 12 (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 cut-off monotonicity established in Theorem 2. Under the protocol described in Section 12, unsure voters only observe the state k when particular cut-offs 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 3, $u(\bar{k}, p) \geq u(l, p)$ 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. If, upon reaching $p(l)$, it turns out that $k = l$, i.e. 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, equal to s/r.

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

References

Albrecht, J., Anderson, A., and Vroman, S. (2007) "Search by Committee," Working Paper.

Baldwin, R.E. (1985) *The Political Economy of U.S. Import Policy*, Cambridge, MA. MIT Press.

Bergemann, D., Hege, U. (1998) "Venture Capital Financing, Moral Hazard and Learning," *Journal of Banking and Finance*, Vol. 22, pp. 703–735.

Bergemann, D., Hege, U. (2001) "The Financing of Innovation: Learning and Stopping," *RAND Journal of Economics*, Vol. 36, No. 4, pp. 719–752.

Bhagwati, J. (1988) *Protectionism*, Cambridge, MA. MIT Press.

Bolton, P., Harris, C. (1999) "Strategic Experimentation," *Econometrica*, Vol. 67, No. 2, pp. 349–374.

Compte, O., Jehiel, P. (2008) "Bargaining and Majority Rules: A Collective Search Perspective," Working Paper.

Décamps, J.-P., Mariotti, T. (2004) "Investment Timing and Externalities," *Journal of Economic Theory*, Vol. 118, pp. 80–102.

Dixit, A. K. (1993) *The Art of Smooth Pasting,* in Fundamentals in Pure and Applied Economics, Vol. 55. Harwood Academic Publishers.

Fernandez, R., Rodrik, D. (1991) "Resistance to Reform: Status Quo Bias in the Presence of Individual- Specific Uncertainty," *American Economic Review*, Vol. 81, No. 5., pp. 1146–1155.

Fleming, W.H., Soner, H.M. (1993) *Controlled Markov Processes and Viscosity Solutions*, Springer, New York.

Keller, G., Rady, S., and Cripps, M. (2005) "Strategic Experimentation with Exponential Bandits," *Econometrica*, Vol. 73, No. 1, pp. 39–68.

Kuran, T. (1988) "The Tenacious Past: Theories of Personal and Collective Conservatism," *Journal of Economic Behavior and Organization*, Vol. 10, pp. 143–171.

Leland, H. (1994) "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *Journal of Finance*, Vol. 49, pp. 1213-1252.

Li, H. (2001) “A Theory of Conservatism,” *Journal of Political Economy*, Vol. 109, No. 3, pp. 617–636.

Malueg, D.A., Tsutsui, S.O. (1997) “Dynamic R&D Competition with Learning,” *RAND Journal of Economics*, Vol. 28, pp. 751-772.

Messner, M., Polborn, M. (2008) “The Option to Wait in Collective Decisions,” Working Paper.

Myers, S. (1977) “Determinants of Corporate Borrowing,” *Journal of Financial Economics*, Vol. 5, pp. 147–175.

Presman, E.L., Sonin, I.M. (1990) *Sequential Control with Incomplete Information*, Academic Press, New York.

Roberts, K. (1989) “The Theory of Union Behavior: Labour Hoarding and Endogenous Hysteresis,” Mimeo, London School of Economics.

Rodrik, D. (1993) “The Positive Economics of Policy Reform,” *American Economic Review*, Vol. 83, No. 2, pp. 356–361.

Simon, L., Stinchcombe, M. (1989) “Extensite Form Games in Continuous Time: Pure Strategies,” *Econometrica*, Vol. 57, pp. 1171–1214.