Policy Herding as a Challenge to Political
Persuasion*

Mauricio Fernández Duque

April 14, 2013

Abstract
Can rationally ignorant voters and informed politicians lead to persistently inefficient policies in a democracy? A long tradition of thought in political science and economics has answered negatively, due to the averaging out of randomly uninformed voters or persuasion attempts in a competitive setting. Rather, explanations for the persistence of inefficient policies has focused on interest group politics. In this paper, I present an argument for why voters may come to support inefficient policies. In particular, support may come from policy herding. The existence of a policy implies that choices were made to support it. If the policymaking process sometimes filters inefficient policies but rationally ignorant voters cannot distinguish when, they will incorrectly weigh the efficiency of an observed policy. Politicians without vested interests have incentives to persuade voters to support a more efficient policy, but policy herding biases against these persuasion efforts. On the other hand, politicians with vested interests have incentives to maintain the policy, further counterbalancing persuasion efforts. The longer a policy is in place, the further the herding effect is compounded. So even if vested politicians eventually leave the polity, the policy may persist due to voter support. I formalize this logic through an electoral model with incomplete information, and discuss the types of policies and institutions that make policy herding more likely.

1 Introduction
Can rationally ignorant voters and informed politicians lead to persistently inefficient policies in a democracy? There is a strong tradition in political science and economics that answers ‘no’ (Popkin [1991], Page and Shapiro [1992], Stigler [1992], Wittman [1995], Lupia and McCubbins [1998]). There are two flavors of argument as to why this is so. The first is the aggregation argument, that states that as long as voter’s mistakes over the best policy are random, it only takes a small proportion of informed voters to favor the best policies since others are averaged out. The second is the information argument, which states that

*This is a very preliminary draft. Please do not cite or distribute.
voters actually have more information that they are given credit, due in part to politician’s competitive incentives to keep them informed of the best policies. Explanations for inefficient policies have either focused on the role of special interests in making policy deviate from the voters’ preferences (e.g. Acemoglu and Robinson [2006], Coate and Morris [1999]) or have argued that voters’ preferences are systematically irrational (Caplan [2007]). What’s left unexplained by the former is that voters do seem to support policies that do not favor them, such as the widespread support of protectionist policies. What’s unsatisfying about the latter is that the irrationality of beliefs seems an ad-hoc explanation, which doesn’t address the incentives for politicians to persuade voters. In this article, I provide an explanation with rational voters and persuasive politicians.

The starting point for my argument is the idea of *policy herding*. The existence of a policy implies that choices were made to support it. If the policymaking process sometimes filters inefficient policies but rationally ignorant voters cannot distinguish when, they will incorrectly weigh the efficiency of an observed policy. Voters will consider an inefficient policy that makes it through this filter to be more efficient than it is. Thus, it is the observation of an action—the policy being implemented—without observing the conditions under which it was passed—that the filter was working—that makes voters ‘herd’ on the efficiency of the policy. The more persistent a policy, the more unobserved decisions that had to be made in order to keep it in place. If voters believe that these decisions served in some sense as filters, but are (rationally) ignorant about the conditions under which it passed, persistence will make individuals more confident about the appropriateness of the policy. Herding therefore leads persistence to breed persistence.

The effect of policy herding may be counteracted through several mechanisms. In this paper, I will focus mostly on the incentives for challengers to persuade voters to support an alternative policy. By persuasion, I mean the impact that a politician’s claims have on voter’s actions. A politician will be more persuasive for higher levels of trust and credibility, where the former is the probability that his incentives are aligned with the voter and the latter is that he has private information about what benefits the voter (Lupia and McCubbins [1998]). I focus on persuasion to address the argument that political competition leads to the most efficient policies being chosen and show why this may not hold. Other mechanisms that may counteract herding, and the conditions under which they fail to do so, have received wider discussion in the literature. I discuss them below as a way to support the main argument.

Suppose an incumbent politician implemented an inefficient policy and consider a situation in which the incumbent and a challenger have the opportunity to persuade voters whether or not to support this policy, as in an election. The challenger politician that does not have a vested interest in the inefficient policy has incentives to persuade voters to support an alternative policy for at least two reasons. One is that they may gain from a more efficient policy, perhaps indirectly through increasing social welfare. The other is electoral: by diminishing support for the incumbent’s policy, trust in the incumbent may diminish. Support for the incumbent’s policy, in contrast, would increase his trust. The
incumbent therefore has incentives to persuade voters to support his own policy. In this scenario, there are two reasons that give the policy in place an advantage over the alternative policy. The first is the herding rationale outlined above. The second is that the electoral motivation turns persuasion attempts into cheap talk. Were there only politicians with non-vested interests trying to persuade voters, the electoral incentives would diminish, making candidates’ appeal to an alternative policy more persuasive.

Policy herding and vested politicians interact to allow for policy persistence. When a policy has recently been implemented, the effects of herding are weakest, but the policymakers directly involved in its implementation are also most likely to be in power. When campaigning, the cheap talk mechanism diminishes attempts to move away from the policy. At later campaigns, even if the politicians that are campaigning no longer have vested interests in maintaining the policy, the herding mechanism has been compounded. Indeed, campaigns serve as filters of inefficient policies, so rationally uninformed voters will put weight on the campaign as having filtered inefficient policies out. Therefore, politicians will require higher levels of trust than at earlier elections in order to be able to change the implemented policies.

In order to formalize this logic, I present an electoral model with asymmetric information. There are three stages in the model. The first is a policymaking stage, where an incumbent politician chooses whether or not to implement a policy. A politician may try to implement an efficient policy which benefits voters or an inefficient one which does not. It may be inefficient due to special interests, because the incumbent has wrong ideas about what benefits voters or because he wants to implement a policy for selfish reasons. A policy is inefficient if there is an alternative policy that yields higher outcomes for the voter, and it is efficient if the alternative policy yields lower outcomes. The policy needs to pass a policymaking procedure, which serves as a filter of inefficient policies.

In the second stage, the incumbent is up for re-election. A representative voter will choose the candidate he believes to be most trustworthy given the campaign and past actions. Politicians are of two types: they either have incentives aligned perfectly with the voters or imperfectly. Since the former types always choose the most efficient policy, the focus on the model is on the strategy of the latter types. Trustworthiness is the belief that the politician’s incentives are aligned. Given that the incumbent chose the policy to be implemented, the appropriateness of this policy is a particularly good signal of the incumbent’s trustworthiness. The electoral results therefore depend heavily on the appropriateness of the incumbent’s policy. The incumbent and a challenger must decide whether to campaign in favor of the policy or in favor of the alternative. We assume that they both have perfect information about the appropriateness of the policy. A representative voter who is rationally ignorant did not observe whether the filter worked or not, but rather has beliefs over the probability that the filter worked and observes that the policy was passed. Further, the voter has beliefs over the trustworthiness of the candidates, and knows they have perfect information regarding the state of the world.

In the third stage, a second round of elections are held and a new representa-
tive voter makes the electoral decision. This voter is uncertain both of whether
the policymaking filter worked in case of the policy and whether the first elec-
tions worked as a filter. Formally, the voter is uncertain of the first election
challenger’s trustworthiness. In these elections, neither of the candidates have
vested interests, in the sense of neither being involved (or specifically benefit)
from the campaign being implemented. The electoral results therefore depend
less heavily on their campaigns regarding the appropriateness of the policy. We
look for the perfect Bayesian equilibrium with restrictions on out-of-equilibrium
beliefs.

The following presents the preliminary results of the model. We focus on
the case where an inefficient policy is passed, so the incumbent politician has
unaligned incentives. The voter’s posteriors over the efficiency of the policy and
the trustworthiness of the incumbent are higher the higher the belief that the
filter works. This first instance of herding gives the incumbent an advantage,
who pools on a campaign in favor of his policy in order to avoid signaling that he
is not trustworthy. We derive conditions under which the challenger cheap talks
in the sense of campaigning against the incumbent’s policy independently of the
state of the world. These conditions imply that this cheap talk will be more
the higher the voter’s belief that the filter functioned properly (due to policy
herding), the lower the trust in the challenger (who cannot persuade voters to
support the alternative policy) and the lower the probability the voter assigns
to the incumbent having had a chance to implement an inefficient policy. Since
both players are pooling on the state of the world, the campaigns are not as
informative about the state of the world as it could be. This statement makes
two claims. The first is that the campaigns are still informative, which is driven
by the trust of the candidates. The second is that it would be more informative
were the strategy to separate on the state of the world. Indeed, we will argue
that there is a range of values in which, if both candidates had the same levels
of trust as an incumbent and a challenger but neither had vested interests, an
equilibrium can be sustained where both separated the state of the world with
their campaign strategies. This result is driven by how we operationalize vested
interests. When non-vested candidates campaign, they are free to choose to
campaign for or against the specific policy because they have not committed to
any policy. On the other hand, they pay a smaller electoral cost if they are not
to be trusted on the basis of the campaign because it is a weaker signal of their
type. The benefits of separating on the efficient policy can therefore outweigh
the electoral costs in situations where they would not for the challenger facing
an incumbent. Finally, in the third stage of the model, the voter now infers
the efficiency of the policy from having passed both the policymaking stage and
the election. Having passed both of these filters increases the voter’s confidence
in the efficiency of the policy. Non-vested candidates will therefore have more
trouble getting support for the alternative policy than they would have in the
first election, in the sense that they require higher trustworthiness to be willing
to campaign on the alternative policy in equilibrium.

Policy herding and the challenge it poses to persuasion can help explain why
support for inefficient policies develop in rational voters and why it can persist

4
in a democracy. It also provides a novel framework with which to think about the conditions under which support for an inefficient policy is more or less likely to survive. A standard way of thinking about how individuals learn about the state of the world, and which they may use to judge the efficiency of a policy, is the observed outcome. Observing an outcome that is clearly different from the one they would have expected had the policy been appropriate would be an indication that the policy, and therefore the politician who proposed it, are not to be trusted. However, there are numerous reasons why observed outcomes do not rule out the policy herding logic. For one thing, the consequences may not always be observed by the general public. This may be because the effect of the policy are long term, which allows for policy herding to compound and increases the incumbent’s incentives to pander (Canes-Wrone et al [2001]). Another reason outcomes may be hard to observe is that there may be plausible confounds, particularly when rationally ignorant voters are using aggregate outcomes such as the state of the economy to judge politician’s performance. Indeed, Cole, Healy and Worker (2012) find that rain shocks increase support for a politician in an agricultural setting, presumably because the impact on the economy is partly attributed to the incumbent. Politicians that implement an inefficient policy also have incentives to make the outcomes of the policy hard to distinguish from a bad draw on an efficient policy (Coate and Morris [1995]), or to simply act as if it were not a bad draw (Majumdar and Mukand [2004]). Outcomes may therefore be hard to use to distinguish good from bad policies. Further, a higher degree of policy herding implies that the outcomes have to be more distinct in order to convince voters that they are in the wrong state of the world.

Institutions also affect whether inefficient policies will persist due to herding. They may do so through their effect on the trustworthiness of the candidates or on the amount of information that is available. Belief in the effectiveness of the policymaking filter and trust in the incumbent in principle reflect that there are fewer inefficient policies that get implemented. However, inefficient policies that are implemented persist longer since voters are more confident that they are efficient. For example, policies that require congressional approval as opposed to executive fiat may be considered more filtered, and harder to change in future periods. A complementary reason congressional approval may exacerbate policy herding is that congress may represent voter’s biases and pose a challenge to the executive’s persuasion attempts.

A main result of the model is that the trustworthiness of the challenger also decreases the amount of herding. Institutions that increase the trustworthiness of the challenger, such as competitive primaries, therefore decrease the persistence of inefficient policies. Intermediate ideological strength also increases the claim of candidate’s credibility: it must be not too weak so that campaigns are considered vote-getting cheap talk nor too strong so that their claims are uninformative about the state of the world (Cukierman and Tommasi [1998]). Finally, the availability of information about the state of the world decreases the amount of policy herding. This suggests that inefficient policies will persist less with a more competitive media, education and research institutions.
In addition to the papers I’ve mentioned, this paper is related to several strands in the literature. There is a literature discussing how rational voters may prefer inefficient policies, which has focused on a war of attrition (Alesina and Drazen [1989]) or risk aversion (Fernandez and Rodrik [1991]) to redistributive outcomes. Neither of these papers highlight the policy herding mechanism or the challenges to persuasion. There is also a long literature on the persistence of policies due to policies justifying equilibrium beliefs, mostly in the context of the welfare state (e.g. Piketty [1995], Benabou and Tirole [2006]). In these arguments, beliefs are sustained by the change in behavior brought about by a certain policy, whereas in ours beliefs are sustained by the observation of a policy and the loss of information about how the policy was implemented. Finally, Acemoglu and Jackson (2012) recently looked at the impact of leadership on social change, which is similar to our focus on candidate persuasion and policy change. Their focus was on norms understood as a coordination equilibrium, and a leader did not use persuasion but rather visibility to affect change.

The rest of the paper is organized as follows. Section II presents the setup of the model. Section III presents the results. We leave future sections for later drafts.

2 The Setup

There are three periods, each with different players so I do not use time subscripts.

2.1 Period One: Policymaking

In the first stage, only the incumbent politician (I) makes a decision. He must decide whether or not to implement a policy \( P \in \{0, 1\} \). There are three states of the world: \( \theta \in \{H, h, L\} \) and the incumbent is of two types, \( \psi_I \in \{U, A\} \). The state of the world, the incumbent type and the policy determine the payoffs to the politician and to the voters. A type \( \theta = A \) politician has incentives that are (A)ligned with the voter, whereas a type \( \theta = U \) politician has (U)naligned incentives. We define an incumbent’s trustworthiness as the voters’ belief that he is of type \( A \) in state of the world \( \theta = H \) (\( \theta = L \)), voters and incumbents get a higher payoff from state \( P = 1 \) (\( P = 0 \)) than from state \( P = 0 \) (\( P = 1 \)). The contentious state is \( \theta = h \). Whereas voters and type \( A \) incumbents get a higher payoff from state \( P = 0 \), type \( U \) incumbents get a higher payoff from state \( P = 1 \).

Formally, the voter’s utility from the policy being implemented is given by
\[
 u_v(P = 0, \theta = L) > u_v(P = 1, \theta \neq L) = u_v(P = 1, \theta = H) > u_v(P = 0, \theta \neq H).
\]

Notice that because of the payoff structure, voters cannot distinguish through observing outcomes between choosing the wrong policy when \( \theta = L \) and choosing the right policy when \( \theta = H \). This argument is most convenient for herding, because voters cannot learn about the true state of the world through observing outcomes. Although it has been argued that politicians have incentives to make inefficient policies hard to distinguish from efficient ones (Coate and Morris
[1995]), it is not essential for policies to persist and can be relaxed.

The $\theta = h$ state is the state of the world where the $U$ type incumbent would want to implement an inefficient policy. We assume his payoff is such that in this case he implements the policy regardless of dynamic considerations. The setup is general enough to allow for different interpretations for why the policy is inefficient. It may be inefficient due to special interests (which we don’t model explicitly), because the incumbent has wrong ideas about what benefits voters or because he wants to implement a policy for other selfish reasons.

In order for the policy to pass, it must pass a policymaking procedure. We model this procedure in a reduced form. We simply assume that the filter ($\Phi$) was either (F)unctioning or it was (B)roken. If it was functioning, $P = 1$ would not have passed under $\theta = h$. If it was broken, it would have passed. The politician has complete information about $\Phi$, the voters only know the distribution of $\Phi$.

In what follows, the only case that interests us is the one where the politician was a $U$ type, the state of the world $h$ and the filter $B$. That is, an inefficient policy $P = 1$ was passed. We will focus on this case throughout.

### 2.2 Period Two: Incumbent and Challenger Election

In the second period the incumbent is campaigning against a challenger ($C$). The incumbent and the challenger must choose whether to campaign for or against policy $P$. In particular, the incumbent and challenger choose $\pi_I \in \{0, 1\}$ and $\pi_C \in \{0, 1\}$ respectively. We assume that campaign promises are binding, in the sense that what a candidate campaigns must be executed once in office. A more complicated model would require a credible signal that this was the case.

A campaign $\pi_I = P$ ($\pi_C = P$) means that the incumbent (challenger) supports the policy. There are two types of candidates: $\phi_j \in \{U, A\}$ for $j \in \{I, C\}$ (the incumbent inherits his type from the first period). A candidate of type $A$ only receives utility from implementing the efficient policy. A candidate of type $U$ benefits from the efficient policy and from winning office. The utility of a candidate $j \in \{U, A\}$ is given by $u_j = u_v(P, \theta) + \alpha 1_{j = U} 1_{v = j}$, where $u_v$ is the utility the voter gets from policy $P$ in state $\theta$, $\alpha > 0$ is a parameter weighing the importance of the electoral motivation, and $1_X$ is an indicator function which takes on the value of 1 if $X$ holds and zero otherwise.

There is a representative voter who must decide which candidate to vote for, $v \in \{I, C\}$. The voter’s utility depends on the efficiency of the policy being implemented and the type of politician in power: $U_v(u_v(P, \theta), e_v(\psi_I, \psi_C, v))$. Regarding the second argument, the voter maximizes his utility by voting probabilistically for the most trustworthy candidate, which he infers using the candidates’ actions and his knowledge of the state of the world. In particular, beliefs over the appropriateness of a policy, whether the filter was functioning and the campaigns will shape the voter’s perceptions of the candidates’ trustworthiness. Note that we’re assuming that the voter does not know the strategy with which the candidate implemented a policy, so does not update based on that information. We assume that the incumbent’s trustworthiness depends
more heavily on knowledge of the appropriateness of the policy than the challenger’s. The intuition for this is that the choice of policy in a situation of power is a better signal of an individual’s type than the mere knowledge of its appropriateness. We capture this intuition simply by introducing it into the voter’s utility function. In particular, the voter does better off by voting for the incumbent if the incumbent is an A type, and worse off if the incumbent is a U type: \[
\text{arg max}_{v \in \{I, C\}} e_v(\psi_I, \psi_C, v) \text{ is } I \text{ if } (\psi_I, \psi_C) \in \{(A, A), (A, U)\} \text{ and } C \text{ if } (\psi_I, \psi_C) \in \{(U, A), (U, U)\},
\]
where \(u_v\) is the voter’s utility function.

The voter votes for the candidate with the probability that the candidate is more trustworthy. Probabilistic voting is a way to make the candidate’s decisions more continuous, and we allude to the standard models that justify these assumptions.

2.3 Period Three: Election Between Non-Vested Candidates

In the third period there are two non-vested candidates campaigning, call them candidates 1 and 2. These candidates’ trustworthiness is given by \(\psi_i\) for \(i \in \{1, 2\}\). Their utility function is as above, and have perfect information about \(\theta\). Again, they must choose whether to campaign for or against the policy \(P\), by setting \(\pi_i \in \{0, 1\}\) for \(i \in \{1, 2\}\). There is a representative voter that chooses who to vote for, and does so on the basis of who he finds more trustworthy. In contrast to the first election, knowledge of the state of the world is an equal signal of the trustworthiness of either candidate. We capture this through the voter’s utility function: \[
\text{arg max}_{v \in \{1, 2\}} e_v(\psi_1, \psi_2, v) \text{ is } 1 \text{ if } (\psi_1, \psi_2) = (A, U), 2 \text{ if } (\psi_1, \psi_2) = (U, A) \text{ and } \{1, 2\} \text{ if } (\psi_1, \psi_2) \in \{(A, A), (U, U)\}. \]
If the voter is indifferent between the voters, he randomizes his vote.

The voters have incomplete of the policymaking stage and of the first election. Incomplete information about the first election might be due to information being lost due to memory of the voter or because of a new generation of voters. Although a less rational story, loss of information due to imperfect memory is also consistent with incentives to be ignorant about public matters. We capture this incomplete information by assuming that the voters have uncertainty about the trustworthiness of the challenger. That is, voters hold beliefs \(\Psi_C \in \{\psi^H_C, \psi_C\}\) where \(\psi^H_C\) is a random variable with that is equal to A more often than \(\psi_C\): \(\Pr(\psi^H_C = A) > \Pr(\psi_C)\). Further, the voters are completely uninformed about the conditional strategy used by the politicians: they only observe the actual campaigns.

2.4 Solution Concept

I am interested in the Perfect Bayesian equilibrium of the game. Since dynamic linkages are not important in this model, we will first solve the second period election and then the third period election. In each election, the voter will maximize his expected utility through his vote, where beliefs are consistent with the candidate’s equilibrium strategies. Out-of-equilibrium beliefs will be
restricted to be consistent with strict dominance: an out of equilibrium action will be taken by the type of candidate who does not have a strictly dominant strategy to take the alternative action. Given these beliefs, candidates will choose to campaign to maximize their utility.

3 Solving the Model

We begin with the following:

**Observation 1.** A higher belief in the trust in the policymaking process increases trust in an incumbent who passed an inefficient policy, and in the belief that the policy was efficient.

\[
\frac{\partial \Pr(\psi_I = A | P = 1)}{\partial \Pr(\Phi = F)} > 0, \quad \frac{\partial \Pr(\theta_I = H | P = 1)}{\partial \Pr(\Phi = F)} > 0
\]

This follows from directly from writing out the probabilities using the law of iterated expectations:

\[
\Pr(\theta = H | P = 1) = \Pr(\Phi = F) + \frac{\Pr(\theta = H)}{\Pr(\theta = H) + \Pr(\theta = H) \Pr(\psi_I = U)}
\]

\[
\Pr(\psi_I = A | P = 1) = \Pr(\Phi = F) \Pr(\psi_i = A) + \frac{\Pr(\theta = H) \Pr(\psi_I = A)}{\Pr(\theta = H) \Pr(\psi_I = A) + \Pr(\theta \neq H) \Pr(\psi_I = U)}
\]

and noting that in both cases the term multiplying \(\Pr(\Phi = F)\) is larger than the one multiplying \((1 - \Pr(\Phi = F))\), which is guaranteed because all the probabilities are between 0 and 1. □

The intuition for this is the herding rationale: when voters see a policy being implemented but are not sure of the conditions under which it passed, they will assign weights to the conditions under which it passed based on their priors. The more they believe the policymaking procedure blocks inefficient policies, the more likely it is that they’ll believe the policy that passed was efficient and the incumbent is trustworthy.

Now let us focus on the second period. We relabel these posteriors as the priors for this period. The first thing to establish is the following:

**Observation 2.**

- A type \(A\) incumbent has a dominant strategy to separate on his policy choice and campaign by setting \(P = \pi_I = 1\) when \(\theta = H\) and \(P = \pi_I = 0\) when \(\theta \neq H\).

- If \(\alpha > \alpha^* > 0\), the equilibrium strategy of a type \(U\) incumbent is to separate on his policy choice and campaign by setting \(P = \pi_I = 1\) when \(\theta \neq H\) and \(P = \pi_I = 0\) when \(\theta = L\).
A type A incumbent has a dominant strategy to separate on his campaign by setting $\pi_C = 1$ when $\theta = H$ and $\pi_I = 0$ when $\theta \neq H$.

Intuition.
That it is a dominant strategy for type A candidates to separate on policy choices follows from the observations that the only argument in their utility function they can affect at the second stage is the policy that will be implemented, and that by campaigning on a strategy different from the one in Observation 2 they decrease the probability that the efficient policy will be implemented.

That it is an equilibrium strategy for the $U$ candidates to separate on $\theta$ as stated in the proposition follows simply in the case of $\theta \neq h$, since maximizing the policy objective is aligned with increasing electoral incentives through pooling with the A type incumbent’s strategy. That it is also an equilibrium strategy to do so when $\theta = h$ if the electoral benefit is high enough follows from the fact that by setting $\pi_I \neq P$, the voter will set posterior beliefs $Pr(\psi_I = A | \pi_I \neq P) = 0$. With these posteriors, the incumbent would lose the election with certainty. The policy objective and the electoral objective of the incumbent are at odds when the policy implemented was inefficient, and if the latter is sufficiently important (\(\alpha\) high enough) he will choose $P = \pi_A$. □

Given Observation 2, for $\alpha > \alpha^*$ the only choice left to consider is the optimal strategy for the challenger candidate of type $U$. This candidate’s policy and electoral incentives are aligned when the incumbent proposed an inefficient policy: by campaigning for the alternative policy the candidate makes it more likely for the efficient policy to be implemented, and decreases the incumbent’s trustworthiness. So the question we’re interested in is the following: would the challenger campaign in favor of the policy the incumbent implemented if he knew it was the appropriate policy? If he did so, the challenger’s campaign would be most credible in the states of the world where the policy was inefficient.

However, he has electoral incentives to deviate. If he does, the informativeness of the campaign itself would be dampened by the cheap talk motivations of the $U$ types, and the campaign’s information value would depend on their trustworthiness.

To disentangle the voter’s motivations, we consider what the voter’s beliefs over types would be were the voter to observe a pooling $\pi_C = 0$ strategy versus observing $\pi_C = 1$ when the challenger was separating on $\theta$. In this latter case, the voter would find both candidates agreeing on the state of the world and would only be able to infer their type through his priors: $Pr(\psi_I, \psi_C | P = 1, \pi_I = 1, \pi_C = 1) = Pr(\psi_I)Pr(\psi_C)$. In contrast, when the challenger pools on $\pi_C = 0$, the voter’s beliefs over the candidate’s types depends on $\theta$. If $\theta = H (\theta \neq H)$, then it must be the case that the challenger (incumbent) is a $U$ type and the probability that the incumbent (challenger) is an $A$ type is given by the voter’s priors: $Pr(\psi_I, \psi_C | P = 1, \pi_I = 1, \pi_C = 1) = Pr(\psi_I) \left( Pr(\psi_I, \psi_C | P = 1, \pi_I = 1, \pi_C = 1) = Pr(\psi_C) \right)$.

It remains to be determined how beliefs over $\theta_I$ are formed. This brings us to the following:

Observation 3. When $P = \pi_I = 1$, voter’s posterior beliefs over $\theta_I$ are higher
for $\pi_C = 1$. These posterior beliefs are lower the higher the belief that the policymaking process was functioning.

We use Bayes’ rule to determine the posterior beliefs of $\theta_I = H$.

$$Pr(\theta_I = H|P = 1, \pi_I = 1, \pi_C = 0) = \frac{Pr(\psi_C = U)Pr(\theta_I = H)}{Pr(\theta_I = H)Pr(\psi_C = U) + Pr(\theta_I = h)Pr(\psi_I = U)} < 1$$

$$Pr(\theta_I = H|P = 1, \pi_I = 1, \pi_C = 1) = 1. \Box$$

We can now state our first main result.

**Proposition 4.** If $Pr(\theta_I = h)Pr(\Phi = B) > Pr(\theta_I = h)Pr(\psi_I = U)Pr(\Psi = B) + Pr(\theta_I = H)Pr(\psi_C = U)$ and $\alpha > \alpha^{**} > 0$, then the $U$ type candidate will choose to pool on $\pi_C = 0$ in equilibrium.

By the above discussion, we know that the probability the voter will vote for the challenger when he sets $\pi_C = 1$ is $Pr(\psi_I = U)Pr(\psi_C = A) + Pr(\psi_I = U)Pr(\psi_C = U)$. In contrast, when setting $\pi_C = 0$ the probability becomes $Pr(\theta_I \neq H|P = 1, \pi_I = 1, \pi_C = 0)$. Setting the second term greater than the first and taking the value of $Pr(\theta_I \neq H|P = 1, \pi_I = 1, \pi_C = 0)$ from the proof of Observation 3 yields the first condition. The second condition requires the payoff from the electoral motivation to be high enough to compensate for utility loss coming from the fact that the wrong policy will be chosen with a positive probability when $\theta = H$. $\Box$

**Corollary 1.** $Pr(\theta_I \neq 1|P = 1, \pi_I = 1, \pi_C = 0) = 1$ is easier to achieve the higher the voter’s belief that the filter functioned properly, the lower the trust in the challenger, the lower the probability assigned to a situation where incentives were unaligned and the higher assigned to a situation where they were aligned $Pr(\theta_I = H)$.

This follows from inspection of the first condition of Proposition 1 and noting that if the $U$ type challenger pools with the $A$ type challenger on separating on $\theta$, the voter will be fully informed of the state of the world by observing $\pi_C$. $\Box$

The comparative statics in the corollary map onto the discussion presented in the introduction regarding implications of policy herding, which we will not go into further detail for this draft.

We now establish that having an incumbent with vested interests makes it harder to move away from an inefficient policy. In order to do this, we look for the conditions under which it is an equilibrium for two non-vested candidates to be separating on $\theta$ with their campaigns.

**Claim 5.** There exists $\alpha^{***} > \alpha^{**}$ such that if $\alpha < \alpha^{***}$, $Pr(\Phi = B)$ high enough, $Pr(\psi_I = U) > 1/2$ and $Pr(\psi_I = A) > 1/2$, non-vested candidates can sustain an equilibrium in which they separate on $\theta$ with their campaigns in situations where a challenger would have pooled on $\pi_C$ against an incumbent.
Elements of Proof.

In order to establish the proof, we need to first show that the conditions for pooling on separating on $\theta$ are more relaxed for deviating at some state of the world. The rest of the proof, which I leave for a future draft, looks at the conditions under which different types of deviations are more profitable and from there derive conditions on $\alpha^{**}$. To establish the first part of the proof, we note that as with the incumbent and challenger, when both non-vested candidates are following the same equilibrium, the voter’s posterior beliefs over the joint distribution of types are equal to the joint prior distribution. Also as before, when the non-vested candidates are following different campaign strategies, belief over the joint distribution of types depends on what the voter thinks is the correct state of the world. To show some preliminary intuition, consider a deviation from the pooled strategy to one in which candidate 1 (say) deviates to pooling on $\pi_{C} = 0$. The probability that the voter will vote for candidate 1 under the original strategy and when $P = 1$ was:

$$\Pr(\psi_I = A)\Pr(\psi_C = U) + \frac{1}{2}[\Pr(\psi_C = A)\Pr(\psi_I = A) + \Pr(\psi_C = U)\Pr(\psi_I = U)]$$

Note that this probability takes into account that the voter randomizes his vote when he believes that the voters are both of the same type. This is where the assumption of non-vested voters is reflected. On the other hand, the probability that the voter will vote for candidate 2 under the deviation to pooling on $\pi_{C} = 0$ is:

$$\Pr(\theta_I \neq H)\Pr(\psi_C = A)\Pr(\Phi = B)$$

which is the same as the case for the incumbent and challenger we just considered. Note that the only difference in the electoral payoffs as compared to the incumbent and challenger case comes from the bracketed term under the original strategy. It is straightforward that if $\Pr(\psi_I = U) > 1/2$ and $\Pr(\psi_C = U) > 1/2$, this term is larger than the counterpart in the incumbent and challenger case, $\Pr(\psi_I = U)\Pr(\psi_C = U)$. Therefore, the conditions for pooling on separating on $\theta$ are laxer for the non-vested candidates case. To complete the proof, I have to specify the conditions under which this deviation is the most profitable one, or what the conditions for an alternative more profitable deviation are. □

The final main result is related to the third round of the model.

Claim 6. Non-vested candidates in the second round of the model would require higher levels of trustworthiness to separate on $\theta$ through their campaigns. Expressed alternatively, there is a larger set of levels of trustworthiness such that non-vested candidates would ‘pander’ in their campaigns by setting their equal to the policy chosen by the incumbent in stage 1.

Intuition.

The intuition from this result comes through the compounded effect of policy herding on the posterior over $\theta$. As with the uncertainty over the policymaking filter, uncertainty over the trustworthiness of the challenger gives positive weight to a state of the world where the challenger would have been able to persuade
voters with a campaign in favor of the alternative policy. This increases the voter’s belief that the state of the world is $\theta = H$, which from the Corollary we know leads to more stringent conditions for the candidates to separate on $\theta$.

4 Conclusion

I provided a rational explanation of why voters come to support inefficient policies. As suggested in the introduction, this explanation helps us think about the conditions under which an inefficient policy will persist in a democracy. Future drafts will include a further discussion of the conditions under which persistence will hold, as well as case studies to illustrate the argument.

References

[10] Cukierman and Tommasi, 1998. When does it take a Nixon to go to China?

