The Politics of the Slippery Slope

Giri Parameswaran[a], Gabriel Sekeres[b] , and Haya Goldblatt[c]

[a]Haverford College, gparames@haverford.edu

[b]Stanford Institute for Economic Policy Research, gsekeres@stanford.edu

[c]Haverford College, haya.goldblatt001@gmail.com

September 14, 2023

Abstract

Slippery slope arguments — the idea that otherwise beneficial reforms should be rejected lest they beget further undesirable one — are ubiquitous in political discourse. We provide a learning-based policy-feedback mechanism to explain why slippery slope dynamics arise. Additionally, we provide conditions under which, in equilibrium, sophisticated agents will successfully manipulate policy to either induce or prevent a slippery slope dynamic.

Key Words: Slippery slope, Learning, Policy momentum.

1 Introduction

'Slippery slope' arguments are commonly invoked in political discourse. They express the idea that even though a policy may be desirable on its own merits, it should nevertheless be rejected because of the fear that its adoption will cause more extreme (and undesirable) policies to arise in the future.[1]

The public discourse surrounding the Affordable Care Act (ACA) provides a useful case study. Despite largely mirroring a proposal from the conservative Heritage Foundation, and notwithstanding its adoption by a Republican administration in Massachusetts, the ACA did not command the support of congressional Republicans, and was even met with suspicion by conservative Democrats. For example, during negotiations over the bill, Democratic Senator Ben Nelson expressed opposition to a proposed Medicare buy-in worrying that it would be a "forerunner of single payer" healthcare (Raju, 2009). His concern was not unfounded. After the ACA had passed, then Democratic Senate Majority Leader Harry Reid confirmed that his goal was "absolutely" to transition the ACA to a single payer system (Roy, 2013). The consideration by courts of the ACA's legality also raised slippery slope concerns. Justice Antonin Scalia famously worried that, absent a clear limiting principle, a government mandate to buy health insurance today would invite future governments to mandate the purchase of more mundane items such as broccoli.

Both examples have the feature that the immediate policy in question acts as a stepping stone that makes possible a more extreme policy, which would be politically infeasible to implement directly today, but which might become feasible as the public becomes accustomed to the moderate change. The slippery slope dynamic is generated by policy feedback: experience with a moderate policy may cause the public to re-evaluate their beliefs about the value of that reform, and potentially demand even more of it. This insight reflects Schattschneider's

[1] For examples of slippery slope arguments, see Dent (1999) and Kurtz (2003) on same-sex marriage, Nix (2012) and Somin (2012) on the Affordable Care Act mandate, and Volokh (2003) for an exhaustive primer.

(1935) aphorism that ‘a new policy creates a new politics’.

Implicit in the logic of policy feedback is that agents learn about the value of certain policies as they interact, and become acquainted with, those (or similar) policies. Experience with ACA programs, for example, has been shown to positively influence agents’ opinions about the ACA, as well as other governmental healthcare schemes, such as Medicare (see Campbell, 2020; Jacobs & Mettler, 2018; Lerman & McCabe, 2017).[2] There is strong empirical evidence (summarized below) that policy feedback occurs in many other contexts, as well. Importantly, the literature finds that the extent of this feedback depends on a range of factors, particularly the size, scope, and import of the policy, and the likelihood that agents experience or engage with it directly.

In this paper, we first explore a particular mechanism that explains why a slippery-slope dynamic — in which a moderate reform today begets a more extreme reform in the future — might arise. We also investigate the conditions under which (some) agents’ awareness of this policy feedback might create an incentive to strategically manipulate policy to either induce or prevent the dynamic from arising.

To answer these questions, we present a simple stylized model of public goods provision under majority rule. Agents are distinguished by their income; a majority have low income whilst the remainder have high income. Low income earners have a higher demand for the public good than high income earners, and this generates the baseline political disagreement between the groups.

Additionally, each agent may either be correctly informed about the value of the public good, or misinformed. We focus most attention on the case where misinformed agents *undervalue* the public good; thus expressing a lower demand than their informed counter-parts. This reflects the public’s typical skepticism towards unfamiliar projects and reforms. We discipline

[2] For specific channels through which this policy feedback arises, see: Jacobs and Mettler (2016), McCabe (2016), Pacheco and Maltby (2017), Hopkins and Parish (2019), Hobbs and Hopkins (2021), and Sances and Clinton (2021).

the model by assuming that a majority of agents are informed — so that our results are not purely driven by misinformed majorities. Importantly, though a majority are poor and a majority are informed, we assume that the informed poor are a minority.[3]

Misinformation has two effects: First, the median voter's preferred level of public goods provision will be below that of the informed poor. There will be policy 'skepticism' relative to the 'correct information' baseline. Second, the preferences of the misinformed poor and the informed rich will be more closely aligned, and these groups may potentially form a cohesive voting bloc, even though their intrinsic preferences (if correctly informed) diverge.

Agents may learn about the true value of the public good by acquaintance. We consider a very simple learning technology wherein agents learn the correct value of the public good whenever it is provided in a sufficiently large quantity to be consequential to their utility. This is consistent with empirical findings, noted above, that learning about policy is strongest when the policy is salient and visible to the agent.

Taken together, these features of our model imply several noteworthy results. First, if learning occurs in some period, it causes the ranks of the informed to grow, which increases future social demand for the public good, *ceteris paribus*. Learning shifts political power between the different groups. A moderate policy today combined with a skeptical public who can learn from acquaintance, induces a more extreme policy tomorrow. This is the slippery-slope dynamic at work. Policy momentum arises endogenously, as a consequence of learning by acquaintance.

Second, since the slippery slope dynamic hurts the informed rich (by moving policy farther from their ideal), they have an incentive to downwardly distort policy to prevent learning. To be successful, the informed rich must enlist the support of the misinformed poor, to build a majority coalition around this distorted policy. But this can only occur if the ideal

[3]If they were a majority, then they would constitute a decisive coalition in their own right, and there would be no interesting political economy analysis.

policy of the misinformed poor is *even lower* than that of the informed rich (absent strategic considerations). Thus, strategic manipulation of policy will only occur if misinformation creates a larger wedge in policy preferences between the informed and misinformed poor than is the inherent wedge between the (informed) rich and poor, ensuring that a natural alliance exists between the informed rich and misinformed poor against the informed poor.

Of course, distorting policy is costly to the informed rich, and so the incentive to behave strategically extends only as far as the benefits from preventing learning exceed the costs. This requires that the distortion needed to prevent learning is not too large.

The logic of strategic behavior requires that some agents are forward looking, and understand the policy dynamic that arises when there is learning by acquaintance. Our third result suggests a complement to this insight: the ability of some agents to strategically manipulate policy is limited by the degree of sophistication of other agents, and their awareness of being manipulated. The informed rich will be most able to strategically prevent the slippery slope dynamic when sufficiently many misinformed agents are myopic. By contrast, if the misinformed poor are sophisticated in sufficient numbers, then opportunities for strategic manipulation will disappear. Moreover, the competing incentives to manipulate and prevent manipulation will often result in policy incoherence, where no equilibrium policy exists.

In light of the motivating example, the public finance setting is a natural one to study the interplay between beliefs, intrinsic preferences (captured by income), and policy. However, the model's insights can be extended beyond this narrow setting. What is important is the interaction between the nature of misinformation and the learning technology, such that those who seek to strategically prevent learning can make common cause with the misinformed.

As a counter-point, we note (in Appendix A.2) that policy neither evolves endogenously nor is there policy distortion if the misinformed minority *overvalue* the public good, and are thus *optimistic* (rather than skeptical) of the policy reform. The reason is that the political incentives now pull in opposite directions; the informed rich still prefer a lower policy than the informed poor, but the misinformed poor will now prefer a higher policy than the informed rich. A natural coalition that supports strategic under-provision of the public good no longer exists.[4]

Somewhat more subtly, we show that in some instances, misinformation can cause policy to be over-provided. In particular, a pessimistic minority (who perceive the majority to be overly optimistic) may have an incentive to strategically over-provide the public good in order to induce learning that teaches the majority that they have been too optimistic. A similar dynamic can arise when the misinformed are a majority and optimistic. Here, learning serves a 'teach them a lesson by giving them what they want' flavor. Our analysis highlights conditions under which such a dynamic may arise.

Though perhaps less commonplace, examples of the lesson-teaching motive exist. Consider the response from law enforcement to the 'defund the police' movement. In many instances, police responded by withdrawing services, for example by patrolling neighborhoods less intensely. A widely reported case was the occupation of the Capital Hill area of Seattle in 2020 as part of broader anti-policy violence protests. The Seattle Police Department responded by abandoning the area, enabling the creation of the Capital Hill Autonomous Zone. The result, as examined by Piza and Connealy (2022), was an increase in crime in the zone and the surrounding area of Seattle. The CHAZ ended just three weeks after it began, and the police re-occupied the area. In over-providing the policy of 'less police', law enforcement were seemingly able to demonstrate that the benefits of such a policy were overstated. More generally, following short periods of reduced funding for police, many communities have restored or even increased that funding (see Fegley & Murtazashvili, 2023).

Though our leading example focused on efforts by conservatives to prevent healthcare reform,

[4]Our maintained assumption that a majority of agents are informed plays an important role here. If instead, a majority are optimistic, then there can be an equilibrium with policy reversal; a larger policy is enacted in the first period that begets learning, and a subsequent roll-back of policy in the second period. The 'war on terror' and Brexit are both policies that majorities initially supported but came to regret after implementation.

the logic of the slippery slope is not inherently partisan. The dynamic can arise whenever some voters are suspicious of a beneficial reform, including reforms typically associated with conservative parties. For example, slippery slope arguments are often invoked in debates over school choice, asserting that any increase in the number of charter schools would lead to a two-tiered system of education (Bailey, 2013).

The two key behavioral assumptions under-pinning our model — that there is policy feedback and that agents often undervalue reforms — are supported by considerable empirical evidence. First, there is strong evidence that policy feedback occurs, and that voters learn by acquaintance. In the case of Social Security, Campbell (2011) finds that the introduction of Social Security led to both increased knowledge of the program and increased support for it. Similarly, Cook, Jacobs, and Kim (2010) find that a program which sent information on social security to a random sample of beneficiaries increased confidence in the program among those who received the information. In a different context, Baccini and Leemann (2012) show that voters are more likely to be sensitive to climate issues after being exposed to a natural disaster. Similar effects exist regarding attitudes towards gay and lesbian people. Herek and Glunt (1993) and Herek and Capitanio (1996) show that interpersonal contact was the strongest predictor of positive attitudes towards homosexuals. And, of course, public policy affects the opportunities for learning by acquaintance to occur. Day and Schoenrade (1997, 2000) and Griffith and Hebl (2002) show that anti-discrimination policies cause individuals to be more open about their sexuality, thereby enabling known interpersonal contact between homosexuals and heterosexuals.

Second, the idea that voters are often mistaken about the value of reforms or public goods is also plainly evident. In a survey study, Koch and Mettler (2012) found that over 50% of respondents receiving some type of government benefit (such as the Home Mortgage Interest Deduction, the Earned Income Tax Credit, Pell Grants or Food Stamps) were unaware that those benefits were indeed provided by the government. This perceived absence of government in their lives suggests that agents will be more skeptical of the value of public spending than they would ideally, if they were correctly informed. Conversely, when government spending is seen to be wasteful or directed towards ends that do not directly improve the public welfare, voters tend to inflate the costs of such programs. U.S. spending on foreign aid provides a stark example. The median respondent in a 2010 World Public Opinion Poll of 848 Americans believed that the foreign aid budget accounted for 25% of the federal spending, whilst only 19% believed it was below 5%. In fact, it was less than 1% of total federal spending. By over-attributing the share of public spending on 'non-beneficial' projects, voters effectively undervalue public spending as an aggregate bundle.

Moreover, history is replete with examples of policies that voters were originally suspicious or skeptical about, but eventually came to appreciate. Social Security, which is now extremely popular amongst voters, was, at its inception, feared by many as a socialist scourge that would enslave Americans.[5]

This paper contributes to, and extends, several strands of the political economy literature. At its core, the inefficiency in this model arises from an endogenous time inconsistency in the decision makers' preferences, arising out of the changing identity of the pivotal voter. This feature is common to many models of inefficiencies in policy making, including Persson and Svensson (1989), Roberts (1989), Alesina and Tabellini (1990), Tabellini and Alesina (1990), Dewatripont and Roland (1992), Benabou (2000), and Battaglini and Coate (2008), amongst many others. However, in contrast to many of those models, and similar to Acemoglu and Robinson (2001), Benabou and Ok (2001), the shifting political power is not exogenous, but endogenous to the current pivotal agent's policy choice, in our model.

Policy momentum is another feature of this model that is present in Benabou and Ok (2001). In that paper, policy is sticky. This creates a fear in the current poor that a redistributive pol-

[5]Unsurprisingly, slippery slope concerns formed part of the objection to Social Security. During Congressional hearings, a senator from Oklahoma asked Secretary of Labor Frances Perkins, "Isn't this socialism?". When she answered no, he responded: "Isn't this a teeny-weeny bit of socialism?" Altman (2005)

icy that will benefit them in the short run will persist long enough to eventually expropriate their future wealth. Thus, policy inertia is hard-wired into their model. Our paper is more standard in that it allows the polity to change its policy in every period. Reform momentum arises as an equilibrium phenomenon rather than as a feature of the model technology. Although the mechanism that generates the behaviors are distinct, reform momentum in this model generates gradualism in policy making, similar to Dewatripont and Roland (1992). We explore this further in Section 5.

Besley and Coate (1998) present a model, similar to ours, with endogenously evolving policy. In the first period, the polity must both choose a (redistributive) tax policy and decide whether to undertake a public investment that changes the distribution of second period incomes. In equilibrium, the public investment may be rejected even when it weakly increases all incomes, if the change in the income distribution causes political power to shift. A similar dynamic exists in Bénabou, Ticchi, and Vindigni (2022), where a first period investment (in research) affects second period preferences for religious public goods.

Though both these papers involve policy feedback, we think that our approach, which highlights the role of beliefs and learning, is apt for the slippery slope context. To see why, notice that in only affecting beliefs, nothing fundamental about the economy has changed between the first and second periods; what was feasible before remains so after, and vice versa. Yet behavior changes. This is precisely the puzzling dynamic behind the slippery slope. By contrast, in Besley and Coate (1998), the first period policy changes the set of feasible second period policies. Additionally, the fact that the first period tax policy *itself* affects preferences over the second period tax policy makes the connection to policy momentum natural. This is less true in the above models, where some other dimension of first period policy affects second period preferences over tax policy. For example, in those models, though there is strategic behavior, there is never distortion of the first period tax policy. By contrast, in our model, the strategic behavior manifests precisely in distorting the first period tax policy.

A related literature examines the incentives for policymakers to distort (or even sabotage) policy. Groseclose and McCarty (2001) and Hirsch and Kastellec (2022) both present models in which policy distortion serves to harm the reputation of an incumbent whose ability is imperfectly observed by voters. Kang (2022) explores the opposite logic of a Congress that is overly deferential to the president, in the hope of showcasing subsequent presidential failure — analogous to the lesson-teaching motive for over-providing the public good in our model.

Finally, this paper extends upon a growing literature on learning in a political economy context. Fernandez and Rodrik (1991) consider a model in which asymmetric information about the identity of winners and losers from a reform may cause the reform to fail, even if the reform makes the average agent better off. Similar to this paper, they find an endogenous bias towards status quo policies. More recent work consider the incentives for agents to choose policies that affect the learning of others. Strulovici (2010) studies learning in bandit problems where decisions (about how to experiment) are made collectively by majority vote. Heidhues, Koszegi, and Strack (2018) develop a model of individual Bayesian learning with an overconfident prior. Hirsch (2014) studies the dynamic interaction between a principal and agent who share common preferences but different beliefs, where learning is possible from past choices. Baker and Mezzetti (2012), Fox and Vanberg (2014), and Parameswaran (2018) consider models of the judiciary in which learning occurs after courts observe the outcomes of agent choices. In the electoral setting, Dewan and Hortala-Vallve (2019) present a model voters learn about an incumbent's ability based on the success of past reforms. In each case, the learning motive distorts the incentives to efficiently provide the reform.

The rest of the paper is organized as follows. Section 2 introduces the characteristics of the model. Section 3 establishes basic analytical insights, and Section 4 analyzes the model in a dynamic equilibrium setting. Section 5 presents an extension in which policy and beliefs evolve gradually, and Section 6 concludes. All proofs, as well as several additional extensions, appear in the Appendix.

2 Model

We present a dynamic model with two periods, $t = 1, 2$. There is a unit mass of agents. Each agent may either be rich or poor. Poor agents have (exogenous) income $y_L > 0$ while rich agents have income $y_H > y_L$.[6] We assume that a majority of agents are poor, so that the median income earner has low income.

In each period, the government can provide a quantity $g \geq 0$ of a public good. The public good has unit cost normalized to 1, and is financed through a non-distortionary, proportional tax on income, $\tau \in [0, 1]$. The government's budget constraint is $g = \tau \overline{y}$.

An agent with income y_i has preferences over feasible policies (τ, g) given by:

$$u(\tau, g; y_i) = (1 - \tau)y_i + A \ln g$$

where A parameterizes the marginal benefit of public good spending. The log-linear functional form choice is purely to keep expressions simple; the basic insights will continue to hold for any concave preference.

Each agent may either be correctly informed (I) or misinformed (M) about the value of the public good. Informed agents know the true value of A (which we denote by A_I), while misinformed agents believe that it takes a different value A_M. In our main analysis, we assume that misinformed agents undervalue the public good ($A_M < A_I$), as this will be shown to be the most interesting case. In Appendix A.2 we consider the opposite case of misinformed agents over overvalue the good ($A_M > A_I$). Additionally, in Section 5, we consider an extension in which agent beliefs are drawn from a continuous distribution. To ensure that the first order conditions produce interior solutions, we assume that $A_I < y_L$.

So far, we have identified four types of agents; each agent having one of two possible incomes and one of two possible beliefs. For each type $i \in \{LI, HI, LM, HM\}$, let ϕ_i denote the

[6] In Appendix A.4 we explore an extension where agent incomes are drawn from a continuous distribution. All the key insights continue to hold in this richer setting.

proportion of type-i agents in the economy. No group constitutes a majority in its own right, so $\phi_i < \frac{1}{2}$ for all i. As above, we assume that a majority of agents are poor (i.e. $\phi_{LI} + \phi_{LM} > \frac{1}{2}$). To ensure that our results are not purely driven by the existence of a large number of misinformed voters, we assume that a majority of agents are informed (i.e. $\phi_{LI} + \phi_{HI} > \frac{1}{2}$). Finally, for technical convenience, we assume that the informed rich and the misinformed poor together constitute a majority (i.e. $\phi_{HI} + \phi_{LM} > \frac{1}{2}$). This latter assumption simplifies the analysis, though our insights will continue to hold even if the condition is violated. Taken together, these assumptions imply that any two of the three largest groups — informed poor, informed rich, and misinformed poor — will together constitute a majority.

We study a simple and stark model of learning. In each period, after a policy is implemented, each agent compares their actual utility against the utility they were expecting, given their belief about A. When these are sufficiently different, the agent realizes that their belief must have been incorrect, and updates their belief to the true value. Formally, an agent with belief A learns whenever:

$$|u(\tau, g; y_i, A) - u(\tau, g; y_i, A_I)| > \mu$$

where $\mu > 0$ parameterizes the agent's sensitivity to information.[7] Though stark, this learning technology operationalizes our story in the simplest possible way. In Appendix A.5, we show that our insights would continue to hold if agents were Bayesian and had non-degenerate priors — though at the cost of considerable complexity.

Finally, agents may either be sophisticated or myopic. A sophisticated agent understands that learning by acquaintance in the first period affects the polity's second period beliefs (and thus policy preferences). When evaluating policies in the first period, sophisticated agents take this dynamic effect into account. Myopic agents, by contrast, ignore this dynamic effect,

[7] In Appendix A.3, we explore an extension in which agents have heterogeneous sensitivities to information. We show that the key insights of our baseline model continue to hold.

and so evaluate policies purely based on their stage game payoff. An alternative interpretation of sophisticated and myopic types is that all agents understand the learning dynamics, but that sophisticated agents are future oriented (putting weight $\beta > 0$ on future utility), whilst myopic agents are purely present oriented (i.e. with discount factor 0). Sophistication and myopia only have their bite in the first period. Since the game ends after the second period, neither type entertains dynamic policy considerations in the second period.

A note about sophistication is in order. Though the model identifies some agents as informed and others as misinformed, all agents will naturally perceive themselves as being correctly informed. Thus, a sophisticated agent with belief A will also believe that, whenever there is learning, other agents will come to share *their* belief.

There are two political parties that are purely office motivated. In each period, each party announces a feasible fiscal policy (τ, g) that it is committed to implement if elected. Voters cast their ballots and the party receiving a majority of the vote is elected. A feasible policy (τ, g) is a majority winner if it is preferred to any other feasible policy (τ', g') by a majority of agents. The median voter theorem predicts that competition between the parities will lead them both to propose majority winning policies. Thus, we associate equilibrium with the majority winning policy whenever it exists.

3 Preliminaries

In this section, we establish several useful insights that will prove valuable in the subsequent dynamic analysis. In particular, we characterize equilibrium behavior in a static version of the game, and (separately) establish the conditions under which there will be learning.

3.1 Stage Game Benchmark

We begin by studying the equilibrium in a single period game in which there are no dynamic considerations. Since the game is static, the agents evaluate policies by their associated stage game utilities.

Consider an agent with income y and whose belief about the value of the public good is A. We make the standard assumption that all agents understand the government's budget constraint; there is no fiscal illusion. Recall that the budget constraint is $g = \tau\overline{y}$; the quantity of public goods provided is in direct proportion to the tax rate. A type-(y, A) agent's indirect utility over tax policies is given by:

$$v(\tau; y, A) = u(\tau, \tau\overline{y}; y, A) = (1-\tau)y + A\ln(\tau\overline{y})$$

It is easily verified that v is strictly concave — and therefore single peaked — in τ for each (y, A). By the first order condition, a type-(y, A)'s most preferred policy is:

$$\tau^*(y, A) = \frac{A}{y} = \frac{A_I}{y \cdot \frac{A_I}{A}} = \tau^*(x(y, A), A_I) \quad (1)$$

where $x(y, A) = y \cdot \frac{A_I}{A}$. The first equality gives a direct expression for $\tau(y, A)$ as a function of y and A. All else equal, agents who believe that public goods are more valuable will demand a higher tax rate to fund more public goods; and richer agents will demand a lower tax rate and fewer public goods than poorer agents. For notational convenience, we denote a type i's ideal stage game policy by τ_i, where $i \in \{LI, HI, LM, HM\}$.

The final equality in (1) reveals that the most preferred tax rate of a type-(y, A) agent coincides with the most preferred tax rate of a type-$(x(y, A), A_I)$ agent; i.e. an informed agent having income $x(y, A)$. We refer to $x(y, A)$ as the agent's 'effective income'. It is the income for which their most preferred policy would be truly optimal if they had correct beliefs. Naturally, the effective income of an informed agent is simply their income. However, since misinformed agents undervalue the public good ($A_M < A_I$), their effective income will be larger than their true income (i.e. $x(y, A_M) > y$). A misinformed agent who undervalues public goods expresses identical preferences to an agent with higher income who correctly values public goods.[8] Let x_i denote the effective income of a type-i agent.

Now recall that, fixing beliefs, agents' most preferred tax rates are decreasing in incomes. Hence, if the agents are ordered by their effective incomes, their most preferred policies will be monotone in that ordering. Then, since preferences are single peaked, the median voter theorem applies. The equilibrium tax rate will be the most preferred tax rate of the agent with the median effective income.

Which type has the median effective income? Since $y_H > y_L$ and $A_I > A_M$, it follows that the informed poor have the lowest effective income, and the misinformed rich have the highest effective income. The ordering of the remaining types' effective incomes depends on the size of the belief disagreement relative to the size of income disparities. Suppose that the divergence in beliefs is small relative to the disparity in incomes (formally, that $\frac{A_I}{A_M} < \frac{y_H}{y_L}$). Then, $x_{LI} < x_{LM} < x_{HI} < x_{HM}$, which implies that $\tau_{HM} < \tau_{HI} < \tau_{LM} < \tau_{LI}$. Because the belief distortions are small, the ideal policies of the informed and misinformed poor will be closer together than the ideal policies of the informed poor and informed rich. Low income earners, as a group, will have a larger demand for public goods than high income earners. If so, since low income earners collectively form a majority, but the informed poor are a minority, it must be that the misinformed poor are pivotal.

By contrast, if beliefs are farther apart, so that $\frac{A_I}{A_M} > \frac{y_H}{y_L}$, then $x_{LI} < x_{HI} < x_{LM} < x_{HM}$, which implies that $\tau_{HM} < \tau_{LM} < \tau_{HI} < \tau_{LI}$. The distortion in beliefs is sufficiently large that the ideal policy of the misinformed poor is further from that of the informed poor than is the ideal policy of the informed rich. Informed agents, as a group, have larger demand for public goods than misinformed agents. Then, since informed agents collectively form a majority, but the informed poor are a minority, it must be that the informed rich are pivotal.

[8] Indeed, it is not just that the agents share the same ideal policies. Their preferences coincide. To see this, note that $v(\tau; y, A) = \frac{A}{A_I}\left[(1-\tau) \cdot y\frac{A_I}{A} + A_I \ln(\tau\overline{y})\right] = \frac{A}{A_I} v(\tau; x(y, A), A_I)$.

In general, the median effective income will be $x_{med} = \min\{x_{LM}, x_{HI}\}$. The equilibrium policy will be $\tau_{med} = \frac{A_I}{x_{med}} = \max\{\tau_{LM}, \tau_{HI}\}$, and $g_{med} = \tau_{med}\overline{y}$.

3.2 Learning

There will be learning if, given feasible policy $(\tau, \tau\overline{y})$ that is implemented, an agent's anticipated utility (given their belief of A) differs sufficiently from their realized utility (given the true A). This will be true if:

$$|[(1-\tau)y + A_I \ln(\tau\overline{y})] - [(1-\tau)y + A_M \ln(\tau\overline{y})]| > \mu$$

$$\tau > \frac{1}{\overline{y}} \exp\left\{\frac{\mu}{|A_I - A_M|}\right\} = \tau^\dagger$$

i.e. if the implemented policy is sufficiently large. This captures the ideas previously discussed, that agents typically learn from policy only if the policy is sufficiently salient. $\tau^\dagger$ denotes either the highest policy that does not result in learning, or the lowest policy that induces learning.[9] Notice that $\tau^\dagger$ is increasing in μ and decreasing in $|A_I - A_M|$. Intuitively, the less sensitive agents are to information (i.e. the higher is μ), the more salient the policy must be to induce learning. By contrast, the larger is the disparity in beliefs, the less salient the policy needs to be to convince the agent. Since there is a one-to-one relationship between μ and $\tau^\dagger$ (holding $|A_I - A_M|$ fixed), it suffices to present results in terms of $\tau^\dagger$ rather than μ. For simplicity, we will assume that $\tau^\dagger < \tau_{LI}$, which rules out the possibility of no learning even when the highest stage-game consistent policy is chosen.

Note importantly that an agent's propensity to learn is independent of their income. Thus, in any period, there will either be learning by misinformed agents of both income types, or neither income type. Moreover, given our two-types assumption regarding informedness, after learning occurs, all agents will be correctly informed. (In section 5, we explore an extension with a continuum of beliefs in which learning is gradual.)

[9] For technical reasons related to maximizing on an open set, we need to allow either possibility at the threshold. We could avoid the ambiguity by discretizing the policy space.

3.3 Dynamics with Myopic Agents

We end this section by briefly noting the benchmark dynamics that would arise in a world with only myopic agents. Since myopic agents ignore the effect of learning on future outcomes, they will simply express stage game preferences in each period. If so, the stage game equilibrium outcome will prevail, period by period. The first period policy will reflect the (original) median effective income earner's ideal policy, whom we showed would belong to either the informed rich or the the misinformed poor — whichever group had the lower effective income (i.e. $\tau_{med} = \max\{\tau_{LM}, \tau_{HI}\}$.) If $\tau_{med} < \tau^{\dagger}$, then there will be no learning; the second period environment will be identical to the first, and the static outcome will repeat.

By contrast, if $\tau_{med} > \tau^{\dagger}$, then all agents become informed, and then since the poor constitute a majority, the informed poor will be pivotal in the second period. Learning will cause political power to shift between the groups. Moreover, since the informed poor had the lowest effective income, and thus the highest ideal policy, taxes and public goods provision will be higher in the second period than the first. This is is the slippery slope at work. A smaller equilibrium policy today begets a larger equilibrium policy tomorrow. There is endogenous policy momentum.

The point here was to demonstrate how learning could generate a slippery slope. Because we assumed that agents were myopic, they did not attempt to manipulate policy to either prevent or ensure a slide down the slippery slope. We take up that concern in section 4.

4 Dynamics

In the previous section, we characterized the stage game preferences of agents, given their income and beliefs, and showed that these were single-peaked in the policy variable τ. Since the second period of our model is effectively a stage game, this characterization also reflects the agents' preferences in the second period. Additionally, since myopic agents are purely present-oriented, these also reflect the preferences of myopic agents in the first period.

The first period preferences of sophisticated agents, however, may differ insofar as those agents understand that learning in the first period may affect second period outcomes. Absent learning, sophisticated agents understand that the equilibrium second period policy will reflect the ideal policy of the (original) median effective income earner, as characterized in the previous section, so that $\tau_2 = \tau_{med}$. By contrast, if there is learning, sophisticated agents understand that all agents will have the same second period beliefs and that the poor will be pivotal. Moreover, since all agents believe that *they* have correct beliefs, a sophisticated agent with belief A will assume that all other agents will arrive at that same belief.

A sophisticated agent's assessment of their lifetime utility given a first period policy τ is:

$$V(\tau; y, A) = \begin{cases} (1-\tau)y + A\ln(\tau\overline{y}) + \beta\left[(1-\tau_{med})y + A\ln(\tau_{med}\overline{y})\right] & \text{if } \tau < \tau^{\dagger} \\ (1-\tau)y + A\ln(\tau\overline{y}) + \beta\left[(1-\tau^{*}(y_L, A))y + A\ln(\tau^{*}(y_L, A)\overline{y})\right] & \text{if } \tau > \tau^{\dagger} \end{cases}$$

The first part of the agent's lifetime utility (compromising the first two terms) corresponds to first period utility. It is continuous and concave (and thus single-peaked) in the first period policy τ, and achieves a maximum at the stage game optimum $\tau^{*}(y, A)$. The second part corresponds to second period utility, and is affected by the first period policy τ only insofar as τ determines whether there is learning or not (i.e. whether τ is above or below $\tau^{\dagger}$). Thus, the second part is piece-wise constant in τ. If learning harms the agent, then there will be a discontinuous jump down in the agent's utility at $\tau = \tau^{\dagger}$. The opposite is true if learning benefits the agent. The size of this utility jump is given by $\Delta(y, A) = A[\ln(\tau(y_L, A)) - \ln(\tau_{med})] - (\tau(y_L, A) - \tau_{med})y$, which depends on both the agent's income and their beliefs, but not on the specific policy chosen in period 1.

Figure 1 illustrates the lifetime utility of a sophisticated informed rich agent as a function of the first period policy τ. Learning harms the informed rich, because it results in a second

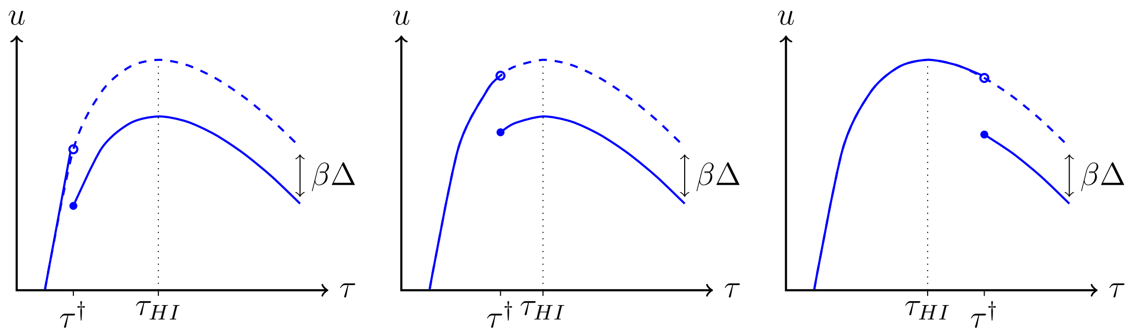

Figure 1: Lifetime utility for a sophisticated informed rich agent as a function of the first period policy τ. Each panel depicts utility for a different value of $\tau^\dagger$ — the threshold policy above which learning occurs; utility drops discontinuously at this threshold.

period policy that is farther from their ideal. Thus, learning reduces the agent's lifetime utility by $\beta\Delta$. The rightmost panel illustrates a situation where $\tau^\dagger > \tau_{HI}$; learning requires a first period policy above the agent's ideal stage game policy τ_{HI}. In this instance, the downshift in utility occurs in a region of the policy space where utility is already decreasing; single-peakedness is preserved. By contrast, in the other two panels, $\tau^\dagger < \tau_{HI}$, and so learning causes utility to decrease in a region where it is otherwise increasing, causing single-peakedness to be violated. Notice that, in the left-most panel, where a large distortion is required to prevent learning, the agent's most preferred policy remains unchanged; it is the stage game ideal. By contrast, her most preferred policy in the middle panel is $\tau^\dagger$. In this instance, the benefit of preventing learning exceeds the cost of distorting the policy away from the stage game optimum.

If learning was instead beneficial (e.g. for the sophisticated informed poor), then utility would discontinuously jump up at $\tau^\dagger$. This will cause single-peakedness to be violated whenever $\tau^\dagger$ is to the right of the agent's stage game ideal (because utility jumps up in a region where it is otherwise falling). Moreover, if the distortion needed to entice learning is small, then the agent's most preferred policy will be $\tau^\dagger$, rather than her stage game ideal.

With these insights in mind, we are ready to begin characterizing the dynamic equilibrium of the game. Since the benefit of learning $\Delta(A, y)$ depends on which agent is pivotal absent learning (i.e. the informed rich or the misinformed poor), our analysis will be in two parts.

4.1 Divergence in Beliefs is Relatively Small

Suppose the divergence in beliefs between the informed and misinformed is relatively small (formally, that $\frac{A_I}{A_M} < \frac{y_H}{y_L}$). We previously showed that, in this scenario, $\tau_{HI} < \tau_{LM} < \tau_{LI}$, so that income disparities create a larger wedge in ideal policies than belief differences do.

By our assumptions, if there were no learning, all agents will expect the second period policy to be τ_{LM}. Moreover, even if there was learning, the sophisticated misinformed agents would expect the same second period policy, because they would anticipate all other agents learning that $A = A_M$. Learning is inconsequential to sophisticated misinformed agents. Thus, all misinformed agents will express stage game preferences in the first period, whether they are sophisticated or not. In particular, the most preferred first period policy of the misinformed poor will be their stage game ideal, τ_{LM}.

By contrast, sophisticated informed agents will potentially face dynamic incentives. They each understand that if there is learning, the second period policy will shift from τ_{LM} to τ_{LI} — which the informed rich perceive as worse, and the informed poor perceive as better. Hence, the informed rich may wish to strategically prevent learning that would otherwise happen under the myopic benchmark (i.e. if $\tau^\dagger < \tau_{LM}$), and the informed poor may wish to strategically induce learning when it would otherwise not (i.e. if $\tau^\dagger > \tau_{LM}$). However, neither group can build a majority coalition around such a strategic choice. For example, none of the poor agents (whether informed or not, or sophisticated or not) would join the informed rich in supporting a policy below τ_{LM}. Similarly, none of the other groups would join the informed poor in supporting a policy above τ_{LM}. This implies the following result:

Proposition 1. *Suppose the divergence in beliefs is small (i.e. $\frac{A_I}{A_M} < \frac{y_H}{y_L}$). The equilibrium first period policy is $\tau_1^* = \tau_{LM}$ (regardless of the value of $\tau^\dagger$).*

When the divergence in beliefs is small, behavior in the dynamic game coincides with the benchmark equilibrium with myopic players. The policy chosen in each period will simply be the most preferred policy of the median effective income earner in that period. In particular, the misinformed poor will implement their ideal stage game policy τ_{LM} in the first period. This will be true even if some measure of agents are sophisticated and have long run concerns.

If $\tau_{LM} < \tau^{\dagger}$, then there will not be any learning, and the first period outcome will repeat in the second period. If $\tau_{LM} > \tau^{\dagger}$, then there will be learning, and the informed poor will become pivotal in the second period. The second period policy will be $\tau_{LI} > \tau_{LM}$. There will be endogenous policy momentum.

When the divergence in beliefs is small, there may be policy momentum insofar as learning causes political power to shift from the misinformed poor to the informed poor. However, awareness of this dynamic by sophisticated agents cannot sustain policy manipulation to prevent policy from going down the slippery slope.

4.2 Divergence in Beliefs is Relatively Large

Next, suppose that the divergence in beliefs between the informed and misinformed is relatively large (formally, that $\frac{A_I}{A_M} > \frac{y_H}{y_L}$). Then, $\tau_{LM} < \tau_{HI} < \tau_{LI}$; differences in beliefs create a larger wedge in ideal policies than income disparities do. If there is no learning, then the informed rich will be pivotal in the second period. By contrast, if there is learning, a poor agent will become pivotal. Unlike the previous case, in this setting, learning will be salient to all sophisticated agents, since it will be understood to shift political power from rich agents to poor agents. As before, learning hurts the informed rich since they perceive it as shifting political power away from them to the informed poor. For the same reason, the informed poor perceive learning as being favorable to them. Additionally, all misinformed agents perceive learning as favorable, since they perceive it shifting political power from the informed rich to the misinformed poor, whose ideal policy is closer to their own.

It turns out the dynamic incentives created by learning affect the strategies of the informed sophisticated agents quite differently from the misinformed sophisticated agents. Accordingly, we conduct the analysis in two parts. First, we suppose that all informed agents are sophisticated and all misinformed agents are myopic. Second, we suppose that all agents — informed and misinformed — are sophisticated. A comparison of these cases will shed light on the role that sophistication plays in sustaining strategic behavior.[10]

4.2.1 Only Informed Agents are Sophisticated

Suppose that all informed agents are sophisticated and that all misinformed agents are myopic. In this case, the strategic incentives are similar to those in Section 4.1. The misinformed will express their stage game preferences. The informed rich may seek to strategically prevent learning that would otherwise happen (if $\tau^{\dagger} < \tau_{HI}$), and the informed poor may seek to strategically induce learning that would otherwise not (if $\tau^{\dagger} > \tau_{HI}$).

Similar to the previous section, the informed poor will not be able to build a majority coalition that successfully distorts policy; no other group desires policies above τ_{HI}. However, now, the informed rich may be able to successfully distort policy in a way that prevents learning. The reason is that, since $\tau_{LM} < \tau_{HI}$, the misinformed poor (who in conjunction with the informed rich constitute a majority) will support moves to push policy below τ_{HI}.

The willingness of the informed rich to actually distort policy depends on a comparison of the first period loss from the distortion against the second period gain from preventing learning. Naturally, the larger the required distortion, the less valuable it is to strategically prevent learning. Let $\underline{\tau}_{HI} < \tau_{HI}$ be the lowest first period policy (i.e. the most distorted policy)

[10]We note that a third alternative exists, where all informed agents are myopic and all misinformed agents are sophisticated. It will turn out that the equilibrium analysis in the second and third cases are quite similar. Accordingly, since this third case is seemingly the least empirically plausible, we leave the analysis to Appendix A.1.

that prevents learning, that is acceptable to the sophisticated informed rich. (Formally, $\underline{\tau}_{HI}$ is characterized by equation (3) in the Proof of Proposition 2.A in Appendix B.)

Before stating the main result, we note a possible complication. The assumptions of the model do not guarantee that $\underline{\tau}_{HI} \geq \tau_{LM}$. If this condition is not satisfied (i.e. if the informed rich are willing to distort policy below the ideal policy of the misinformed poor) and if $\tau^{\dagger} \in (\underline{\tau}_{HI}, \tau_{LM})$, then the ordering of agents by their most preferred policy will differ between the static and dynamic games, affecting the identity of the pivotal voter. We begin by studying the case where this complication does not arise, and then subsequently address the effect of the complication.

Proposition 2.A. *Suppose the divergence in beliefs is large (i.e. $\frac{A_I}{A_M} > \frac{y_H}{y_L}$), and that only informed agents are sophisticated. If $\underline{\tau}_{HI} \geq \tau_{LM}$, then the equilibrium first period policy is given by:*

$$\tau_1^*(\tau^{\dagger}) = \begin{cases} \tau_{HI} & \text{if } \tau^{\dagger} \leq \underline{\tau}_{HI} \\ \tau^{\dagger} & \text{if } \underline{\tau}_{HI} < \tau^{\dagger} < \tau_{HI} \\ \tau_{HI} & \text{if } \tau^{\dagger} \geq \tau_{HI} \end{cases}$$

The content of Proposition 2.A is summarized in Figure 2a. Implicitly, we assume that there is no learning when $\tau = \tau^{\dagger}$. If $\tau^{\dagger} \geq \tau_{HI}$, then there is no need for the informed rich to distort policy — implementing their stage game ideal policy is consistent with no learning, thus enabling them to retain power in the second period without any first period sacrifice. When $\tau^{\dagger} < \underline{\tau}_{HI}$, then the cost of distorting policy to prevent learning is so high that the informed rich 'throw in the towel', implementing their ideal policy today, and accepting that, by doing so, they will cede second period political power to the informed poor. Strategic manipulation occurs when $\tau^{\dagger} \in (\underline{\tau}_{HI}, \tau_{HI})$. In this case, the informed rich strategically under-provide the public good in order to prevent learning, enabling them to retain political power.

These insights can be restated in terms of the effectiveness of the policy feedback channel.

When the feedback channel is weak (i.e. μ is high), then $\tau^{\dagger}$ will be large, and the likelihood of a slippery slope dynamic arising will be small. By contrast, when policy feedback is strong (i.e μ is small), then $\tau^{\dagger}$ will be small, and a slippery slope dynamic will be at play. When policy feedback is extremely strong, the distortion needed to prevent learning will be so high as to make strategic manipulation unattractive. By contrast, when there is moderate feedback, the informed poor strategically distort policy to prevent learning.

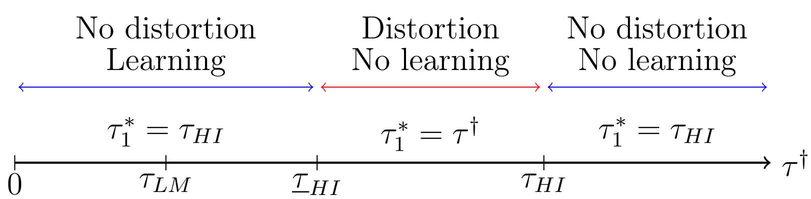

Figure 2a: Equilibrium when the divergence in beliefs is large, the informed (only) are sophisticated, and $\tau_{LM} \leq \underline{\tau}_{HI}$.

A comparison of Propositions 1 and 2.A — strategic behavior arises in the latter, but not the former — is instructive. Strategic manipulation will only succeed if there is a coalition that will support it. When the ideal policy of the misinformed poor lies below that of the informed rich, a natural coalition exists that supports downwardly distorting policy. By contrast, when the ideal policy of the misinformed poor lies above that of the informed poor, then no such common incentive exists. Now, the ideal policy of the misinformed poor will be (relatively) low when the misinformed have strongly pessimistic beliefs. Thus, strategic manipulation is most likely when misinformation highly skews beliefs in the polity.

Let us now return to the complication that arises when $\underline{\tau}_{HI} < \tau^{\dagger} < \tau_{LM}$. In this scenario, the informed rich have an ideal policy ($\tau^{\dagger}$) below that of the misinformed poor (τ_{LM}), and this policy prevents learning; however, limiting attention to policies that induce learning, the informed rich have an ideal policy (τ_{HI}) above that of the misinformed poor. This preference order reversal causes policy to become unstable; there will be no majority winning policy. Instead, a Condorcet cycle will exist.[11]

[11] In our baseline analysis, we associate the existence of Condorcet cycles with policy instability. In

Let $\underline{\tau}_{LM} < \tau_{HI}$ denote the policy below τ_{HI} that gives the misinformed poor the same utility as τ_{HI}.[12] We formally characterize $\underline{\tau}_{LM}$ in equation (4) in the Proof of Proposition 2.B, in Appendix B. In fact, since the misinformed poor are myopic, $\underline{\tau}_{LM} < \tau_{LM}$.

Proposition 2.B. *Suppose the divergence in beliefs is large (i.e. $\frac{A_I}{A_M} > \frac{y_H}{y_L}$), and that only informed agents are sophisticated. If $\underline{\tau}_{HI} < \tau_{LM}$, then the equilibrium first period policy is:*

$$\tau_1^*(\tau^\dagger) = \begin{cases} \tau_{HI} & \text{if } \tau^\dagger \leq \max\{\underline{\tau}_{LM}, \underline{\tau}_{HI}\} \\ \text{No Majority Winner} & \text{if } \max\{\underline{\tau}_{LM}, \underline{\tau}_{HI}\} < \tau^\dagger < \tau_{LM} \\ \tau^\dagger & \text{if } \tau_{LM} < \tau^\dagger < \tau_{HI} \\ \tau_{HI} & \text{if } \tau^\dagger \geq \tau_{HI} \end{cases}$$

Proposition 2.B is summarized in Figure 2b, and is broadly similar to Proposition 2.A, except in that there is policy inconsistency over a region of the parameter space. To see why, note that, in this region, $\tau^\dagger$ cannot be a majority winner; the informed and misinformed poor (who together constitute a majority) would replace it with τ_{LM} (which implies learning). But τ_{LM} cannot be a majority winner, since the informed rich and informed poor (who together constitute a majority) would replace it with τ_{HI}. But a coalition of the informed rich and the misinformed poor would, in turn, replace this with $\tau^\dagger$ (thus preventing learning), provided that $\tau^\dagger$ is not too far below τ_{LM}. There is a Condorcet cycle.

We end this subsection by noting that, though we assumed that the informed poor were sophisticated, this assumption was not important to the result, and Propositions 2.A and 2.B would continue to hold if some or all of the informed poor were myopic. Similarly, it wasn't essential that *all* the informed rich are sophisticated. All that is required is that the measure of sophisticated rich agents is large enough that they, along with the misinformed (rich and poor) jointly constitute a majority.

Appendix A.6 we explore more elaborate voting mechanisms that can fill these 'Condorcet Holes', and predict a focal policy.

[12] The notation intentionally highlights the analogy between $\underline{\tau}_{HI}$ and $\underline{\tau}_{LM}$. For $i \in \{HI, LM\}$, $\underline{\tau}_i$ denotes the policy below τ_{HI} that, absent learning, gives a type i agent the same utility as τ_{HI} does with learning.

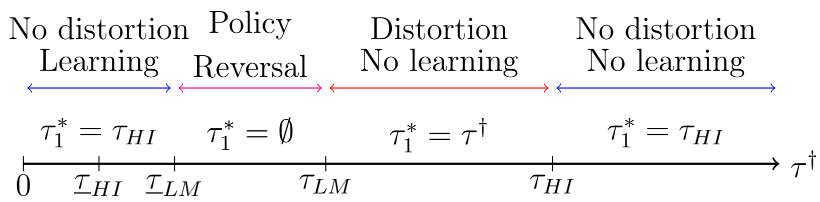

Figure 2b: Equilibrium when the divergence in beliefs is large, the informed (only) are sophisticated, and $\tau_{LM} > \underline{\tau}_{HI}$.

4.2.2 All Agents are Sophisticated

Now, suppose that misinformed agents are sophisticated as well. The strategic incentives for the misinformed agents are quite different to the informed rich. Both types of misinformed agents will be willing to distort policy *upwards* to generate learning, anticipating that political power will shift from the informed rich (whom they consider to be optimistically misinformed) towards the misinformed poor (whom they consider to be correctly informed).

What are the opportunities for strategic manipulation in this environment? When $\tau^\dagger \in (\tau_{LM}, \tau_{HI})$, it may be the the informed rich will seek to prevent learning by locating policy just below $\tau^\dagger$, while the misinformed rich seek to induce learning by locating policy just above $\tau^\dagger$. Obviously, it cannot be that both groups successfully manipulate policy. Indeed, the informed rich will never prevail, since the informed and misinformed poor agents (who together constitute a majority) will always prefer a policy slightly above $\tau^\dagger$ that induces learning, to one slightly below it that does not. Since the informed rich cannot downwardly distort policy, we should never expect an equilibrium policy below τ_{HI}.

This does not guarantee that the equilibrium policy will be τ_{HI}. If $\tau^\dagger \leq \tau_{HI}$, the static equilibrium policy will generate learning, and so there is no additional incentive for the misinformed to strategically distort policy. By contrast, if $\tau^\dagger > \tau_{HI}$, then the misinformed may wish to upwardly distort policy, to induce learning that would otherwise not happen. Naturally, their willingness to do so depends on a trade-off between the first period cost of distorting against the second period gain from having political power shift in their favor.

Let $\overline{\tau}_{LM}$ denote the highest policy (i.e. the most distorted policy) acceptable to the misinformed poor that induces learning, assuming the myopic benchmark policy τ_{HI} does not. We formally characterize $\overline{\tau}_{LM}$ in equation (5) in the Proof of Proposition 5, in Appendix B. As before, let $\underline{\tau}_i$ denote the policy below τ_{HI} that, absent learning, gives a type i agent the same utility as τ_{HI} would with learning.[13] These expressions are formally defined by equations (3) and (4) in Appendix B. Similar to Section 4.2.1, the equilibrium characterization will depend on the relative locations of $\underline{\tau}_{HI}$ and $\underline{\tau}_{LM}$.

Proposition 3. *Suppose the divergence in beliefs is large (i.e. $\frac{A_I}{A_M} > \frac{y_H}{y_L}$), and that all agents are sophisticated.*

1. *If $\underline{\tau}_{HI} \geq \underline{\tau}_{LM}$, then there exists $\tilde{\tau}$, with $\tau_{HI} \leq \tilde{\tau} < \overline{\tau}_{LM}$ such that the equilibrium first period policy is given by:*

$$\tau_1^*(\tau^\dagger) = \begin{cases} \tau_{HI} & \text{if } \tau^\dagger \leq \tau_{HI} \\ \tau^\dagger & \text{if } \tau_{HI} < \tau^\dagger < \tilde{\tau} \\ \text{No Majority Winner} & \text{if } \tilde{\tau} < \tau^\dagger \leq \overline{\tau}_{LM} \\ \tau_{HI} & \text{if } \tau^\dagger > \overline{\tau}_{LM} \end{cases}$$

2. *If $\underline{\tau}_{HI} < \underline{\tau}_{LM}$, then the equilibrium first period policy is given by:*

$$\tau_1^*(\tau^\dagger) = \begin{cases} \tau_{HI} & \text{if } \tau^\dagger \leq \underline{\tau}_{HI} \\ \text{No Majority Winner} & \text{if } \underline{\tau}_{HI} < \tau^\dagger < \overline{\tau}_{LM} \\ \tau_{HI} & \text{if } \tau^\dagger \geq \overline{\tau}_{LM} \end{cases}$$

To make sense of Proposition 3, there are two sets of deviations to consider. If $\tau^\dagger < \tau_{HI}$, then there will be learning under the static benchmark, and so the informed rich will seek to strategically under-provide the public good. If $\tau^\dagger > \tau_{HI}$, then learning will not occur under the static benchmark, and so the poor will seek to strategically over-provide the public good.

[13] We can show that $\underline{\tau}_{LM} > \tau_{LM}$, and that the misinformed poor will prefer a policy $\tau^\dagger$ that prevents learning to policy τ_{HI} that induces learning whenever $\tau^\dagger \in [\tau_{LM}, \underline{\tau}_{LM})$.

Let us take each motive in turn. Suppose $\tau^{\dagger} < \tau_{HI}$. By construction, the informed rich would prefer to replace τ_{HI} with a policy $\tau' < \tau^{\dagger}$ provided that $\tau' \geq \underline{\tau}_{HI}$. To be successful, they need the support of the misinformed poor. But since the misinformed poor benefit from learning, they will only support such a policy if it brings the first period policy much closer to their stage game ideal (i.e. if $\tau' < \underline{\tau}_{LM}$). If $\underline{\tau}_{LM} \leq \underline{\tau}_{HI}$ (as in part (1) of Proposition 3) then there is no policy τ' that satisfies both groups simultaneously, and so there is no possibility of successful downward distortion. If so, the equilibrium policy will be τ_{HI} whenever $\tau^{\dagger} < \tau_{HI}$.

By contrast, if $\underline{\tau}_{LM} > \underline{\tau}_{HI}$ (as in part (2) of Proposition 3), then any policy $\tau' \in (\underline{\tau}_{HI}, \underline{\tau}_{LM})$ that prevented learning would defeat τ_{HI} in a pair-wise majority contest. If so τ_{HI} cannot be a majority winner. But nor can any such policy τ' (assumed to be below $\tau^{\dagger}$) since a majority coalition of the informed and misinformed poor will replace it with a policy slightly above $\tau^{\dagger}$, and a coalition of the informed (rich and poor) would in turn replace *that* policy with τ_{HI}. A Condorcet cycle exists; there is policy inconsistency.

Next, suppose $\tau^{\dagger} > \tau_{HI}$. The misinformed poor would prefer to replace τ_{HI} with (a policy slightly above) $\tau^{\dagger}$ provided that $\tau^{\dagger} < \overline{\tau}_{LM}$, and in this endeavour, they will have the support of the informed poor. This policy is equilibrium consistent provided that it is immune to a counter-proposal by the informed rich that prevents learning and is closer to the stage-game ideal of the misinformed poor. Such a counter proposal will always exist when $\underline{\tau}_{LM} > \underline{\tau}_{HI}$ (for the reasons discussed above). Furthermore, even when $\underline{\tau}_{LM} \leq \underline{\tau}_{HI}$ a successful counter-proposal may exist if $\tau^{\dagger}$ is sufficiently large, so that the distortion required to induce learning is high. The threshold $\tilde{\tau}$ is the largest distortion in policy that is immune to a counter-proposal. Whenever $\tau^{\dagger} > \tilde{\tau}$, a Condorcet cycle will arise and there will be policy instability.

A comparison of Proposition 3 against Propositions 2.A and 2.B reveals two key insights. First, when the misinformed are sophisticated, strategic manipulation produces the opposite dynamic to the typical slippery slope behavior. Rather than downwardly distort policy to prevent other agents from learning that a policy is more desirable, here, the agents upwardly distort policy to ensure that other agents learn that the policy is less desirable. This strategic behavior has a 'learning from mistakes' flavor to it — with the understanding that the mistake must be large enough to ensure that the perceived optimists learn their lesson.

Second, as more agents are made sophisticated, the possibility of coherent policy making breaks down, as strategic incentives cause disparate coalitions to pull policy in different directions. Furthermore, as we argued in previous sections, our results are robust to allowing some agents from each group to be myopic. The insights in Proposition 3 did not rely on all agents being sophisticated *per se* — just that enough of them were.

Finally, recall that, to simplify the analysis, we assumed that the misinformed rich were never pivotal coalition partners. This made it unnecessary to consider policy deviations by a coalition of the informed poor and misinformed rich, say. Propositions 2.A, 2.B, and 3, will all continue to hold even if we relaxed this assumption, though the analysis would become more complicated.

4.3 Discussion

Given the preceding analysis, several insights become apparent. First, the slippery slope dynamic arises due to a specific interaction between the nature of misinformation and learning. It requires that: (i) misinformation causes the median agent to demand less of the public good than they would if perfectly informed; and that (ii) there is learning by acquaintance, so that beliefs evolve with the provision of the public good. The aggregated social preferences are time inconsistent, and reflect a shifting of political power between different groups of agents as learning occurs. Together, these features imply that, over time, the median agent's demand for the public good increases, creating an endogenous policy momentum whereby moderate policies today beget more extreme ones tomorrow.

We emphasize that the slippery slope dynamic is a political economy phenomenon — it is not enough that some voters be misinformed; what matters is how misinformation affects the aggregated social preference over outcomes. This insight will become particularly apparent in Appendix A.2, where misinformation causes agents to overvalue (rather than undervalue) the public good. There, we show that, although the demand by some agents may be higher, misinformation will not distort the preferences of the median agent. Hence, despite some agents being misinformed, and despite the possibility of learning by those agents, there will be no natural force causing policy to endogenously evolve.

Second, the possibility of a slippery slope dynamic arising may create incentives for particular groups of (sophisticated) voters to strategically manipulate policy — either to prevent or hasten the shift in political power between voters. However, we showed that in a political economy setting, their ability to successfully do so is constrained by several factors, including the ordering of agents' stage-game ideal policies. Indeed, for there to be scope for strategic manipulation, the divergence in beliefs between the informed and misinformed needed to be relatively large. This ensured that the wedge in stage-game ideal policies arising from misinformation (between agents having the same income) was larger than the wedge arising from income differences (between agents having the same beliefs) —i.e. $\tau_{LM} < \tau_{HI} < \tau_{LI}$. This arrangement of ideal policies created the possibility that the group seeking to strategically manipulate policy could make common cause with other groups to build a majority coalition around the distorted policy. If this condition were not met, then the groups would seek to pull policy in opposite directions, preventing the emergence of a coherent majority coalition that could shift policy away from the stage-game baseline.

Third, we highlighted the crucial role that sophistication played in generating and sustaining policy distortion. We demonstrated that the canonical case, of policy under-provision to prevent a slide down the slippery slope, could only arise when informed agents were more likely to be sophisticated than their misinformed counterparts. This made possible a stable majority coalition between the informed rich and the misinformed poor to keep policy low.

By contrast, when the misinformed agents were relatively more likely to be sophisticated, the opposite effect arose: the misinformed would upwardly distort policy to induce learning. Interestingly, the learning motive here had a ‘learning from mistakes’ flavor to it — the misinformed upwardly distorted policy by sufficiently much to teach their counterparts a lesson, by making it inescapably clear that the public good was not nearly so valuable.

Finally, we showed that sophistication amongst agents created the possibility of preference reversals, where the ordering of groups’ ideal policies were different in the region of policy-space where learning occurred, from the region where it did not. We showed that these preference reversals were associated with the existence of Condorcet cycles and incoherent policy making. Moreover, we showed that policy inconsistency became more likely as the number of sophisticated agents in the polity grew. This suggests that the role for actual strategic manipulation of policy motivated by a slippery slope dynamic is potentially quite limited. Though slippery slope arguments are common place as rhetorical devices, their translation to actual policy is necessarily more complicated.

5 Heterogeneous Beliefs and Gradualism

A stark feature of our baseline model was that only beliefs were possible: the correct belief A_I or the incorrect belief A_M. Together with the assumption of a common sensitivity parameter μ, this generated the all-or-nothing feature of learning. It also meant that agents would either distort policy just sufficiently to prevent (or induce) learning, or they would not distort at all. In this section, we do two things: First, we demonstrate how our model can be extended to allow agents to have a range of beliefs. Second, we demonstrate that in the extended model, learning will be more gradual, and that this may create incentives for distortion at intermediate levels — where, for example, agents do not prevent learning entirely, but slow the rate of learning, and thus the rate at which political power shifts.

There is a unit mass of agents playing an infinite horizon dynamic game, with time indexed by $t = 0, 1, 2, ...$ Let ϕ_i denote the fraction of agents with income y_i (with $i \in \{L, H\}$ and where $\phi_L > \frac{1}{2}$), $\gamma_i > \frac{1}{2}$ denote the fraction of income-i agents that are initially correctly informed, and suppose that $\phi_L \gamma_L < \frac{1}{2}$. We now allow misinformed agents to have beliefs that are drawn from a wider set. Each misinformed agent has a belief independently drawn from a continuous distribution $F_i(A)$, with support $[A_0, A_I]$.[14] We continue to assume that all agents share a common sensitivity to learning μ. The agents' preferences, the learning technology, and the government's budget constraint are unchanged from the baseline model.

We focus on the case where strategic voting to prevent a slippery slope dynamic is most likely: *i.e.* when the informed are sophisticated and the misinformed are myopic. In what follows, we provide a succinct analysis of this extension and highlight key results. A detailed analysis can be found in Appendix B.

Given a policy τ, the set of agents who update their beliefs are those whose beliefs are sufficiently far from the truth. As before, the propensity to learn is independent of income. Formally, any agent with belief $A < \mathcal{A}(\tau)$ will learn, where:

$$\mathcal{A}(\tau) = A_I - \frac{\mu}{\ln(\tau \overline{y})}$$

Thus, learning truncates the distribution of beliefs from below. Notice that $\mathcal{A}(\tau)$ is increasing in τ — a larger policy begets more learning and truncates the distribution at a higher point. Let A_t denote the lowest belief that an agent may have at the start of time t, given the sequence of policies (i.e. opportunities for learning) up to that time. We have $A_{t+1} = \max\{A_t, \mathcal{A}(\tau_t)\}$ for all $t = 0, 1, 2, \ldots$. At time t, all originally misinformed agents with beliefs $A \in [A_0, A_t]$ will have become correctly informed, while agents with belief $A \in (A_t, A_I)$ remain misinformed. Given this learning process, the mass of informed agents with income y_i at time t is $\phi_i[\gamma_i + (1 - \gamma_i)F_i(A_t)]$. Let $G(x \mid A_t) = \Pr[x(y, A) \leq x \mid A_t]$ denote

[14] For simplicity, we assume that the supports of the distribution are the same for rich and poor agents, but the distributions themselves may differ.

the (induced) distribution of effective incomes, given A_t.

Let $X(A_t)$ denote the median effective income at time t, which is the solution to $G(X(A_t)\,|\,A_t) = \frac{1}{2}$.[15] Define $\underline{A} = \inf\{A\,|\,X(A) < y_H\}$ and $\overline{A} = \sup\{A\,|\,X(A) > y_L\}$. When $A_t < \underline{A}$, the informed rich have the median effective income; when $\underline{A} < A_t < \overline{A}$, a misinformed poor agent (with effective income below that of the informed rich) has the median effective income; and when $A_t > \overline{A}$, the informed poor have the median effective income. We can show that $\frac{A_I}{\underline{A}} > \frac{y_H}{y_L}$, which is the familiar condition that the informed rich are pivotal when the the dispersion in beliefs is larger than the deviation in incomes.

We make 2 key assumptions, which are tantamount to joint restrictions on the behavior of the distribution function F_L and the sensitivity to learning parameter μ. These are: (i) $A_0 < \underline{A}$, and (ii) $\underline{A} < \mathcal{A}(\tau_{HI}) < \overline{A}$. The first assumption states that political power is initially held by the informed rich. The second assumption states that if they implement their ideal stage game policy τ_{HI}, the informed rich will cede political power, but not directly to the informed poor. These assumptions largely mirror assumptions in the baseline model.

Let $\underline{\tau}$ be the largest policy that is consistent with the informed rich retaining power. By construction, $\mathcal{A}(\underline{\tau}) = \underline{A}$, which implies that $\underline{\tau} < \tau_{HI}$.

Proposition 4. *There exists a threshold $\hat{\tau} < \tau_{HI}$ such that the period 0 policy is:*

$$\tau_0^* = \begin{cases} \underline{\tau} & \text{if } \underline{\tau} > \hat{\tau} \\ \in (\hat{\tau}, \tau_{HI}) & \text{if } \underline{\tau} < \hat{\tau} \end{cases}$$

Furthermore, if $\tau_0^* = \underline{\tau}$, then $\tau_t^* = \underline{\tau}$ for all t. Instead, if $\tau_0^* > \underline{\tau}$, then $\tau_t^* = \tau(X(A_t), A_I)$.

Proposition 4 has many similarities to Proposition 2.A. To retain political power, the informed rich must offer a policy no larger than $\underline{\tau}$. If the required distortion is small (i.e. if $\underline{\tau}$ is sufficiently close to τ_{HI}), then the informed rich will distort policy in this way to retain

[15] Because G is not everywhere continuous — it has point masses at y_L and y_H — this condition may not have a solution. This will occur if $G(x\,|\,A_t) > \frac{1}{2} > \lim_{y \uparrow x} G(y\,|\,A_t)$. If so, we take $X(A_T) = x$.

political power. Furthermore, given the stationary environment, they will implement this same policy in all future periods, thus never ceding political power.

By contrast, if the required distortion is large, then the informed rich will cede political power. Importantly, they will continue to distort policy below their stage game ideal (i.e. $\tau_0^* < \tau_{HI}$). This ensures that the effective income of the new median agent (a misinformed poor agent) is higher than would be the case without the distortion, and therefore that future policies are themselves lower than they would otherwise be. Of course, these subsequent policies may induce additional learning, such that the informed poor eventually become pivotal. But the distortion slows down this process, delaying the transfer of political power, and guaranteeing the informed rich more favorable policies along the transition path.

Proposition 4 shows that the key insights of the baseline analysis continue to hold in this richer environment. However, it also highlights the scope for more nuanced policy, and that transition dynamics along the equilibrium path may be characterized more by gradualism than an immediate change.

6 Conclusion

Slippery slope arguments are ubiquitous in political discourse. In this paper, we explored a political economy mechanism that rationalized the slippery slope concern. We first showed that misinformation (that creates policy skepticism) combined with learning by acquaintance, can create a dynamic in which a small reform today begets larger reforms in the future.

We then examined whether awareness of the slippery slope dynamic would result in strategic manipulation of policy to prevent learning — i.e. whether slippery slope concerns would actually cause agents to scuttle otherwise welfare enhancing reforms. Though some agents may always wish to manipulate policy, in a political economy equilibrium with majority rule, we show that policy can only be successfully manipulated if two conditions are satisfied. First, the degree of misinformation must be large relative to the baseline level of political disagreement in the polity. Second, informed agents must be more likely to be sophisticated (and thus understand the slippery slope dynamic) than misinformed agents. If these two conditions are satisfied, then a coalition of the sophisticated informed rich, along with misinformed agents can conspire to strategically manipulate policy. While we focus on a public goods provision setting with rich and poor agents, our insights would apply to any policy setting where the core features of policy skepticism and learning by acquaintance arise.

We also explored other possibilities. When the misinformed are relatively more likely to be sophisticated than the informed, we get the opposite effect — policy skeptics strategically over-provide the reform, to cause optimistic voters to learn that the reform is less worthwhile than they think. This behavior has a 'lesson-teaching' flavor, and generates the opposite dynamic — there is policy reversal rather than policy momentum. Additionally, we demonstrate that policy skepticism by the misinformed is crucial to the mechanism: when the misinformed are over-optimistic, policy does not evolve endogenously at all.

Though our model is simple and stylized, we believe that it captures important insights about the nature of decision making in a political economy setting. The robustness of our results to variant assumptions (see Appendix A) suggests that our insights will continue to hold in more complicated models. Amongst many issues worth investigating are the implications of relaxing various standard assumptions that our analysis takes granted, such as common knowledge assumptions. Certainly, there is scope for further theoretical development of the role of learning in a political economy setting, which we leave for future analysis.

References

Acemoglu, D., & Robinson, J. A. (2001). A theory of political transitions. *American Economic Review*, 938–963.

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

Altman, N. J. (2005). *The battle for social security.* Wiley.

Baccini, L., & Leemann, L. (2012). Information and voting-how voters update beliefs after natural disasters. In *Epsa 2013 annual general conference paper* (Vol. 435).

Bailey, M. (2013). Weingarten proclaims victory in new haven. *The New Haven Independent*. Retrieved 2022-10-20, from https://www.newhavenindependent.org/index.php/article/randi_weingarten_new_haven/

Baker, S., & Mezzetti, C. (2012). A theory of rational jurisprudence. *Journal of Political Economy*, *120*(3), 513–551.

Battaglini, M., & Coate, S. (2008). A dynamic theory of public spending, taxation, and debt. *The American Economic Review*, *98*(1), 201–236.

Benabou, R. (2000). Unequal societies: Income distribution and the social contract. *American Economic Review*, *90*(1), 96–129.

Benabou, R., & Ok, E. A. (2001). Social mobility and the demand for redistribution: the poum hypothesis. *The Quarterly Journal of Economics*, *116*(2), 447–487.

Bénabou, R., Ticchi, D., & Vindigni, A. (2022). Forbidden fruits: The political economy of science, religion, and growth. *The Review of Economic Studies*, *89*(4), 1785–1832.

Besley, T., & Coate, S. (1998). Sources of inefficiency in a representative democracy: A dynamic analysis. *The American Economic Review*, *88*(1), 139–156.

Campbell, A. L. (2011). *How policies make citizens: Senior political activism and the american welfare state.* Princeton: Princeton University Press.

Campbell, A. L. (2020). The affordable care act and mass policy feedbacks. *Journal of Health Politics, Policy and Law*, *45*(4), 567–580.

Cook, F. L., Jacobs, L. R., & Kim, D. (2010). Trusting what you know: Information, knowledge, and confidence in social security. *The Journal of Politics*, *72*(2), 397-412.

Day, N. E., & Schoenrade, P. (1997). Staying in the closet versus coming out: Relationships between communication about sexual orientation and work attitudes. *Personnel Psychology*, *50*(1), 147–163.

Day, N. E., & Schoenrade, P. (2000). The relationship among reported disclosure of sexual orientation, anti-discrimination policies, top management support and work attitudes of gay and lesbian employees. *Personnel review*, *29*(3), 346–363.

Dent, G. W. (1999). Defense of traditional marriage. *JL & Pol.*, *15*, 581.

Dewan, T., & Hortala-Vallve, R. (2019). Electoral competition, control and learning. *British Journal of Political Science*, *49*(3), 923–939.

Dewatripont, M., & Roland, G. (1992). Economic reform and dynamic political constraints. *The Review of Economic Studies*, *59*(4), 703–730.

Fegley, T., & Murtazashvili, I. (2023). From defunding to refunding police: institutions and the persistence of policing budgets. *Public Choice*, *196*, 123–140.

Fernandez, R., & Rodrik, D. (1991). Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *The American economic review*, 1146–1155.

Fox, J., & Vanberg, G. (2014). Narrow versus broad judicial decisions. *Journal of Theoretical Politics*, *26*(3), 355-383.

Griffith, K. H., & Hebl, M. R. (2002). The disclosure dilemma for gay men and lesbians:" coming out" at work. *Journal of Applied Psychology*, *87*(6), 1191.

Groseclose, T., & McCarty, N. (2001). The politics of blame: Bargaining before an audience. *American Journal of Political Science*, 100–119.

Heidhues, P., Koszegi, B., & Strack, P. (2018). Unrealistic expectations and misguided learning. *Econometrica*, *86*(4), 1159â-1214.

Herek, G. M., & Capitanio, J. P. (1996). "some of my best friends": Intergroup contact, concealable stigma, and heterosexuals' attitudes toward gay men and lesbians. *Personality and Social Psychology Bulletin*, *22*, 412–424.

Herek, G. M., & Glunt, E. K. (1993). Interpersonal contact and heterosexuals' attitudes toward gay men: Results from a national survey. *Journal of sex research*, *30*(3), 239–244.

Hirsch, A. V. (2014). Experimentation and persuasion in political organizations. *American Political Science Review*, *110*(1), 68–84.

Hirsch, A. V., & Kastellec, J. P. (2022). A theory of policy sabotage. *Journal of Theoretical Politics*, *34*(2), 191–218.

Hobbs, W. R., & Hopkins, D. J. (2021). Offsetting policy feedback effects: Evidence from the affordable care act. *The Journal of Politics*, *83*(4), 1800–1817.

Hopkins, D. J., & Parish, K. (2019). The medicaid expansion and attitudes toward the affordable care act: Testing for a policy feedback on mass opinion. *Public Opinion Quarterly*, *83*(1), 123–134.

Jacobs, L. R., & Mettler, S. (2016). Liking health reform but turned off by toxic politics. *Health Affairs*, *35*(5), 915–922.

Jacobs, L. R., & Mettler, S. (2018). When and how new policy creates new politics: Examining the feedback effects of the affordable care act on public opinion. *Perspectives on politics*, *16*(2), 345–363.

Kang, M. (2022). Let presidents fail: Congressional deference to presidents as gambling on failure. *Research Politics*, *9*(2), 1–6.

Koch, J., & Mettler, S. (2012). Who perceives government's role in their lives? *Working Paper*.

Kurtz, S. (2003). Beyond gay marriage. *The Weekly Standard*, *8*(45), 26–33.

Lerman, A. E., & McCabe, K. T. (2017). Personal experience and public opinion: a theory and test of conditional policy feedback. *The Journal of Politics*, *79*(2), 624–641.

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

McCabe, K. T. (2016). Attitude responsiveness and partisan bias: Direct experience with the affordable care act. *Political Behavior*, *38*(4), 861–882.

Nix, K. (2012). Comparative effectiveness research under obamacare: A slippery slope to health care rationing. *Heritage Foundation, Washington, DC*.

Pacheco, J., & Maltby, E. (2017). The role of public opinion–does it influence the diffusion of aca decisions? *Journal of Health Politics, Policy and Law*, *42*(2), 309–340.

Parameswaran, G. (2018). Endogenous cases and the evolution of the common law. *The RAND Journal of Economics*, *49*(4), 791–818.

Persson, T., & Svensson, L. (1989). Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences. *The Quarterly Journal of Economics*, *104*(2), 325.

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

Piza, E. L., & Connealy, N. T. (2022). The effect of the seattle police-free chop zoneon crime: A microsynthetic control evaluation. *Criminology and Public Policy*, *21*, 35–58.

Raju, M. (2009). Lieberman says no to medicare buy-in. *Politico*. Retrieved from `https://www.politico.com/story/2009/12/lieberman-says-no-to-medicare-buy-in-030553`

Roberts, K. (1989). *The theory of union behaviour: labour hoarding and endogenous hysteresis* (Tech. Rep.). Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Roy, A. (2013). Sen. harry reid: Obamacare 'absolutely' a step toward a single-payer system. *Forbes*. Retrieved from `https://www.forbes.com/sites/theapothecary/2013/08/10/sen-harry-reid-obamacare-abso`

Sances, M. W., & Clinton, J. D. (2021). Policy effects, partisanship, and elections: How medicaid expansion affected public opinion toward the affordable care act. *The Journal of Politics*, *83*(2), 498–514.

Schattschneider, E. E. (1935). *Politics, pressures and the tariff*. Prentice-Hall, Inc.

Somin, I. (2012). A mandate for mandates: is the individual health insurance case a slippery slope. *Law & Contemp. Probs.*, *75*, 75.

Strulovici, B. (2010). Learning while voting: Determinants of collective experimentation. *Econometrica*, *78*(3), 933–971.

Tabellini, G., & Alesina, A. (1990). Voting on the budget deficit. *The American Economic Review*, 37–49.

Tideman, T. N. (1987). Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, *4*, 185–206.

Volokh, E. (2003). The mechanisms of the slippery slope. *Harvard Law Review*, *116*(4), 1026–1137.

World Public Opinion Poll. (2010, 30 November). *American public opinion on foreign aid* (Tech. Rep.). World Public Opinion.Org. Retrieved from http://www.worldpublicopinion.org/pipa/pdf/nov10/ForeignAid$_Nov10_quaire.pdf$

Appendices

This Appendix is in two parts: Section A presents a series of variants and extensions to our baseline analysis, while Section B presents proofs of all claims in the paper. Within the Section A we present: (i) the case where only misinformed agents are sophisticated, completing the analysis in Section 4.2; (ii) a variant model in which the misinformed *overvalue* the public good; (iii) an extension in which agents are heterogeneous in their sensitivity to information; (iv) an extension in which agent incomes are drawn from a continuous distribution; (v) an extension in which agents have non-degenerate priors and update beliefs according to Bayes' Rule; and (vi) an extension with a voting rule that selects a focal policy whenever Condorcet cycles arise.

A Extensions and Variants

A.1 Equilibrium when Only Misinformed Agents are Sophisticated

In this subsection, we complete the analysis in Section 4.2 by considering the third case, where the misinformed are sophisticated and the informed are myopic. The equilibrium, here, is analogous to that in the first part of Proposition 3, since the condition $\underline{\tau}_{LM} \leq \underline{\tau}_{HI} = \tau_{HI}$ is trivially satisfied.

Similar to Section 4.2.2, the misinformed (both rich and poor) perceive themselves to benefit from learning, as it transfers political power from the informed rich to the misinformed poor (who demand less of the public good). However, unlike in that section, the informed rich do not perceive learning as detrimental to their interests, as they are myopic. Thus, when only the misinformed are sophisticated, no faction will actively seek to prevent learning.

We know that the equilibrium policy in the static baseline is τ_{HI}. If $\tau^\dagger \leq \tau_{HI}$, then the static equilibrium will automatically generate learning, and so there will be no distortion in the dynamic game. By contrast, if $\tau^\dagger > \tau_{HI}$, then the misinformed may seek to upwardly distort policy to induce learning that would otherwise not happen. Naturally, their willingness to do so depends on a trade-off between the first period cost of distorting against the perceived second period benefit from having political power shift in their favor.

As before, let $\overline{\tau}_{LM}$ denote the highest policy (i.e. the most distorted policy) acceptable to the misinformed poor that induces learning, assuming that the static benchmark policy τ_{HI} does not. $\overline{\tau}_{LM}$ is formally characterized in equation (5) in the Proof of Proposition 5.

Proposition 5. *Suppose the divergence in beliefs is large (i.e. $\frac{A_I}{A_M} > \frac{y_H}{y_L}$), and that only misinformed agents are sophisticated. Then there exists $\tilde{\tau}$, with $\tau_{HI} \leq \tilde{\tau} < \overline{\tau}_{LM}$ such that the equilibrium first period policy is given by:*

$$\tau_1^*(\tau^\dagger) = \begin{cases} \tau_{HI} & \text{if } \tau^\dagger \leq \tau_{HI} \\ \tau^\dagger & \text{if } \tau_{HI} < \tau^\dagger < \tilde{\tau} \\ \text{No Majority Winner} & \text{if } \tilde{\tau} < \tau^\dagger \leq \overline{\tau}_{LM} \\ \tau_{HI} & \text{if } \tau^\dagger > \overline{\tau}_{LM} \end{cases}$$

Proposition 5 is summarized in Figure 3, below. The behavior of the equilibrium is analogous to that in Proposition 3. The possibility of strategic policy making arises when $\tau^\dagger \in (\tau_{HI}, \overline{\tau}_{LM})$; otherwise the static benchmark will obtain. When strategic incentives are at play, there are two possibilities. When the distortion necessary to induce learning is small (formally if $\tau^\dagger < \tilde{\tau}$), then $\tau^\dagger$ is a stable policy. By contrast, when this distortion is large, policy becomes unstable. To see why, note that a (majority) coalition of the misinformed and the informed poor will support replacing τ_{HI} with $\tau^\dagger$ whenever $\tau^\dagger \leq \overline{\tau}_{LM}$. If $\tau^\dagger$ is not too large (i.e. $\tau^\dagger < \tilde{\tau}$), this policy will be stable. By contrast, if $\tau^\dagger > \tilde{\tau}$ so that the required distortion is large, then a (majority) coalition of the misinformed poor and the informed rich will support replacing $\tau^\dagger$ with some $\tau' < \tau_{HI}$ which does not induce learning, but implies a much lower first period policy distortion. And a majority coalition of the informed would replace this policy with τ_{HI}. A Condorcet cycle exists.

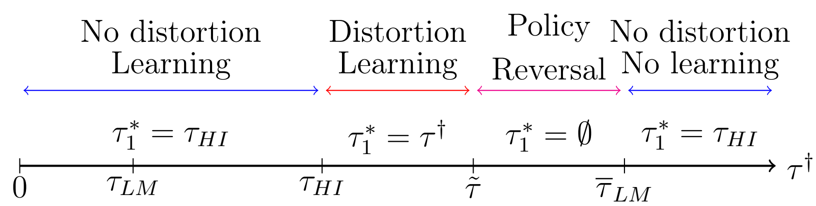

Figure 3: Equilibrium when the divergence in beliefs is large and the misinformed (only) are sophisticated.

We end this subsection by noting, as in the previous cases, that the insights here are robust to allowing some of the misinformed to be myopic or some of the informed to be sophisticated. All that is required is that the measure of sophisticated misinformed poor is sufficiently large that they, along with informed poor, jointly constitute a majority.

A.2 Agents Overvalue the Public Good

Crucial to our baseline analysis was the assumption that misinformed agents undervalued the public good. What if they instead overvalued it, so that $A_M > A_I$? With this change, there will neither be policy distortion, nor will policy evolve endogenously. And this will be true regardless of any assumptions we make about which agents are sophisticated.

To see why, notice that (depending on whether the divergence in beliefs is high or not, relative to the difference in incomes) there are two possible arrangements of effective incomes: (i) $x_{HI} < x_{LI} < x_{HM} < x_{LM}$, or (ii) $x_{HI} < x_{HM} < x_{LI} < x_{LM}$. Then, since the informed are a majority, and the poor are a majority, in both cases the median effective income earner must be an informed poor agent.

Two insights are worth noting. First, if all agents are myopic, then the first period policy will be τ_{LI} — the static ideal policy of the informed poor. Moreover, this policy will repeat in the second period, regardless of whether the policy begets learning or not. (The informed poor will be pivotal either way.) Though learning affects the beliefs of various agents in the polity, it does not affect the beliefs (or identity) of the median voter. Hence, there is no natural tendency for policy to evolve.

Second, sophistication will not change this basic dynamic. If the informed agents are sophisticated, they understand that learning has no dynamic effect on policy, so despite their sophistication, they express the same preferences as if they were myopic. By contrast, the misinformed agents, when sophisticated, may see an incentive to upwardly distort policy to beget learning and shift political power away from the informed poor (whom they perceive to be misinformed). But such policies will never be majority winners, since the informed agents (who together constitute a majority) will always prefer a policy closer to τ_{LI}. Thus, the informed poor's ideal policy will be the majority winner, no matter whether agents are sophisticated or myopic (or any mixture of the two).

What is going on? The underlying tension between the policy preferences of the rich and poor is unchanged — the rich want fewer public goods than the poor. However, when the misinformed overvalue the public good, they demand more of it — effectively behaving as if they were poorer than they are. The misinformed poor and the informed rich are no longer natural allies — in fact, they seek to pull policy in opposite directions. This secures (rather than undermines) the political power of the informed poor.

The analysis so far has retained the assumptions that informed agents constitute a majority, and that the poor have a higher demand for the public good than the rich. Modifying these assumptions can generate different dynamics. First, suppose that the optimistically misinformed constitute a majority, and for concreteness, suppose the informed are sophisticated and the misinformed are naive. If so, it may be that a misinformed poor agent is pivotal in the first period, and will seek to implement their (perceived) ideal stage policy. If that policy begets learning, then informed poor (who have a lower demand) will be pivotal in the second period. There will be policy reversal rather than policy momentum. Several recent examples, where a policy was implemented with majority support only to be subsequently rolled back or regretted, plausibly fit this scenario, amongst them, the 'war on terror' in the 2000s and Brexit.

Second, suppose that the informed are a majority, but that the poor have a lower demand for the policy than the rich. (When the policy in question relates to provision of public goods this might seem implausible, but in other contexts, the demand for reform may be switched. For example, the rich might have a greater incentive to support policies that expand the military industrial complex than the poor.) Again, suppose that the misinformed are optimistic. Then, the ordering of ideal policies is either $\tau_{LI} < \tau_{HI} < \tau_{LM} < \tau_{HM}$ or $\tau_{LI} < \tau_{LM} < \tau_{HI} < \tau_{HM}$, which is exactly the opposite ordering that arose in Section 4. In the stage game with optimistically misinformed agents, the informed rich are able to enjoy a higher policy than they would under the complete information benchmark. Thus, mirroring the logic from that section, equilibria may arise in which the informed rich strategically under-provide the policy, to prevent learning and a future roll-back of policy.

A.3 Heterogeneous Sensitivities to Information

Amongst the starker features of our baseline model was the assumption that all agents were equally sensitive to information, sharing a common learning parameter μ. In this section, we show that this assumption was entirely benign, and that all of our results will continue to hold even if we allowed agents to be heterogeneously sensitive to information.

We adapt our baseline in the following way, similar to the model in Section 5. Each agent is now characterized by a triple (y, A, μ), where $y \in \{y_L, y_H\}$ is the agent's income, $A \in \{A_I, A_M\}$ is the agent's belief (either correct or incorrect) with $A_M < A_I$, and $\mu > 0$ is the agent's sensitivity to information. Let ϕ_i denote the proportion of agents that that have income y_i where $i \in \{L, H\}$, and let γ_i denote the proportion of agents with income y_i that are initially correctly informed. As in the baseline model, we assume that $\phi_L > \frac{1}{2}$, $\gamma_i > \frac{1}{2}$ for each i, but that $\phi_L \gamma_L < \frac{1}{2}$. This ensures that a majority of agents are poor, and a majority are initially informed, but the informed poor are a minority. Additionally, suppose each agent's information sensitivity μ is a draw from a (possibly income contingent) continuous distribution $F_i(\mu)$ with positive support.

Given a first period policy τ, the set of misinformed agents who update their beliefs are those with $\mu < \mathcal{M}(\tau) = (A_I - A_M)\ln(\tau\overline{y})$. In the baseline, learning was all-or-nothing — either all misinformed agents learned, or none did. Now, generically, a fraction $F_i(\mathcal{M}(\tau))$ of misinformed agents with income y_i will learn, given a policy τ. Moreover, the larger is τ, the greater the measure of agents who will learn.

In the baseline model, we showed that a sophisticated group of agents may have an incentive to manipulate policy to either ensure or prevent learning, in order to affect which group had political power in the second period. That basic incentive continues to exist here. For concreteness, suppose that the informed rich are initially pivotal, and suppose that misinformed agents are myopic. Let $\tau^\dagger$ denote the threshold policy beyond which sufficiently many misinformed agents will learn, and the informed poor will become pivotal. $\tau^\dagger$ is defined implicitly by:

$$\phi_L[\gamma_L + (1 - \gamma_L)F_L(\mathcal{M}(\tau^\dagger))] = \frac{1}{2}$$

In the baseline analysis $\tau^\dagger$ was the threshold policy beyond which all misinformed agents would learn (which ensured that the informed poor would become pivotal). In the extension, $\tau^\dagger$ is now the threshold policy beyond which enough of the misinformed agents would have learned, to effectuate a transfer of political power to the informed poor. The all-or-nothing feature of our baseline analysis kept the analysis simple, but nothing turned on the assumption that all, rather than most, agents learned. With partial learning, what is important is that either sufficiently few agents learn (if the goal is to prevent a transfer of power) or that sufficiently many agents learn (if the goal to ensure the transfer). Having identified $\tau^\dagger$, all the results from the baseline analysis continue to hold exactly. For example, as in Proposition 2.A, the informed rich will trade-off the future benefit of retaining political power against the current loss from distorting policy, and thus only distort if $\tau^\dagger$ is close enough to τ_{HI}.

A.4 Continuous Distribution of Incomes

Another stark feature of our baseline model was the assumption that there were only two income types. In this section, we show that this assumption is again benign, and that our results will continue to hold even if incomes were drawn from a continuous distribution.

We adapt the model in Appendix A.3 in the following way. As before, each agent is either correctly informed or misinformed, having beliefs $A \in \{A_I, A_M\}$ with $A_M < A_I$. As in the baseline, all agents share a common sensitivity to information $\mu > 0$. In the extension, each agent's income y is an independent draw from a continuous (possibly belief-contingent) distribution $F_j(y)$ with support on (a subset of) the positive reals, where $j \in \{I, M\}$. (This allows for the income distributions to differ between the (initially) informed and misinformed.) Let $\gamma > \frac{1}{2}$ denote the proportion of agents who are (initially) informed.

Let y_{med} denote the median income earner; $F(y_{med}) = \frac{1}{2}$. Let $G(x)$ denote the distribution of effective incomes. Recall that the effective income of an informed agent is simply their true income, whilst a misinformed agent with income y has effective income $x = \frac{A_I}{A_M}y$. Then: $G(x) = \gamma F_I(x) + (1-\gamma)F_M\left(x\frac{A_M}{A_I}\right)$. Notice that $G(x) \leq F(x)$ for all x; G first order stochastically dominates F, so that the distribution of effective incomes is 'higher' than the distribution of true incomes. Let x_{med} denote the median effective income; $G(x_{med}) = \frac{1}{2}$. Stochastic dominance implies that $x_{med} > y_{med}$. As we will see, the political contest reduces to one between agents with effective incomes above x_{med} and below y_{med}, respectively. Thus x_{med} and y_{med} will take the roles of y_H and y_L from the baseline model.

For concreteness, we focus on the scenario analogous to Section 4.2.1, in which the informed are sophisticated and the misinformed are myopic; where the slippery slope dynamic was most likely to arise. As in the baseline model, learning is all-or-nothing; either all misinformed agents learn (regardless of income) or none do. Political power initially rests with agents with effective income x_{med}. However, if there is learning, a type (y_{med}, A_I) agent will become pivotal. This creates a strategic incentive for informed agents with higher incomes to manipulate policy, to prevent learning and political power from shifting.

Denote by $\tau_x = \tau^*(x_{med}, A_I)$ the ideal stage policy of agents with effective income x_{med}, and similarly define $\tau_y = \tau^*(y_{med}, A_I)$. These are analogous to τ_{HI} and τ_{LI} in the baseline model.

Proposition 6. *There exist thresholds $\tilde{\tau}_2 \leq \tilde{\tau}_1 < \tau_x$ such that:*

$$\tau_1^*(\tau^\dagger) = \begin{cases} \tau_x & \text{if } \tau^\dagger \leq \tilde{\tau}_2 \\ \textit{No Majority Winner} & \text{if } \tilde{\tau}_2 < \tau^\dagger \leq \tilde{\tau}_1 \\ \tau^\dagger & \text{if } \tilde{\tau}_1 < \tau^\dagger \leq \tau_x \\ \tau_x & \text{if } \tau^\dagger > \tau_x \end{cases}$$

If $\tilde{\tau}_2 = \tilde{\tau}_1$, then policy is never unstable.

Proposition 6 is identical to Propositions 2.A and 2.B. When $\tau^\dagger > \tau_x$, the pivotal agent can implement her stage-game ideal policy without fear of inducing learning. When $\tau^\dagger \leq \tilde{\tau}_2$, the distortion required to prevent learning is so large as to deter the pivotal sophisticated agent from behaving strategically; instead, she implements her ideal stage policy and cedes political power. Strategic voting occurs when $\tau^\dagger \in (\tilde{\tau}_2, \tau_x)$. If so, a coalition of the (effective) rich would prefer to prevent learning by implementing $\tau^\dagger$ than to implement the stage-game ideal τ_x and cede political power. Indeed, if $\tau^\dagger \in (\tilde{\tau}_1, \tau_x)$ so that the require distortion is small, this policy is the stable equilibrium policy. However, when the necessary distortion is larger, a different coalition will replace $\tau^\dagger$ with a more moderate policy $\tau' \in (\tau^\dagger, \tau_x)$ that induces learning. (This is analogous to the situation that arises in Proposition 2.B when $\tau^\dagger < \tau_{LM}$.) This will generate a Condorcet cycle, and unstable policy making.

A.5 Non-Degenerate Priors

In the baseline model, we considered a stark information environment in which the agents' beliefs were concentrated at a single point. In this section, we show that the key insights of the model will continue to hold in a more standard Bayesian setting, where agents' beliefs are represented by a non-degenerate prior distribution.

We modify the baseline setup (with two income groups, two information types, and common sensitivity to information) in the following way: Let $F_I(A)$ and $F_M(A)$ be continuous cumulative distribution functions that represent the (non-degenerate) beliefs of informed and misinformed agents, respectively, about the value of the public good A.[16] Let A_I denote the true value of A. We assume that the true belief A_I is in the support of both distributions. Further, suppose that $E_{F_I}[A] = A_I$ and $E_{F_M}[A] = A_M < A_I$, so that the informed have unbiased beliefs, whereas the misinformed have beliefs that are systematically downwardly biased.[17] This mirrors the setup in the baseline model.

Start with the stage game. Since A enters each agent's preferences linearly, expected stage utility is simply the utility associated with the agent's expected belief. We have:

$$E_{F_j}[v(\tau; y_i, A)] = (1-\tau)y_i + E_{F_j}[A]\ln(\tau\overline{y}) = v(\tau; y_i, A_j)$$

where $i \in \{H, L\}$ and $j \in \{I, M\}$. Thus, the stage preferences in a game with non-degenerate beliefs are identical to those in a game with degenerate beliefs (concentrated at the mean of the non-degenerate distribution). Moreover, since unsophisticated agents express stage

[16] These beliefs may themselves be interim beliefs, after the receipt of an (unmodeled) signal about A, where the informed types receive an unbiased signal and the misinformed types receive a pessimistic one.

[17] To be clear, as the modeler, we understand that the average belief of the informed agents coincides with the truth. But the informed agents themselves do not know this. They merely assign some probability to this being the case.

game preferences in the first period, and all agents express stage game preferences in the second period, those preferences remain unchanged from the baseline.

Now introduce learning. We adopt a learning dynamic that is close in spirit to that in the baseline. We retain our assumption that agents do not learn from acquaintance unless the policy generates salient differences from what the agent was expecting, given her prior. Thus, a type $j \in \{I, M\}$ agent will learn if:

$$|A_I - E_{F_j}[A]| \ln(\tau \overline{y}) > \mu$$

Moreover, since there is a one-to-one relationship between the true A and an agent's actual utility, we assume that, in observing her true utility, the agent perfectly learns the true value of A. Thus, when learning occurs, it is complete, and the agent's posterior places all weight on A_I. By contrast, if the policy does not generate salient utility differences, then the agent retains her prior belief.[18] With these assumptions, it is clear that the informed will never learn (though they will always have correct beliefs on average), whilst the misinformed will learn whenever the policy τ lies above the threshold $\tau^{\dagger}$ from the baseline model. Indeed, all of the section 3 results will continue to hold in this setting.

So far, modifying to a Bayesian framework has not affected our results. Differences arise when we consider the preferences and incentives for sophisticated agents in the first period. For concreteness, take the case most amenable to generating a slippery slope dynamic — where misinformed agents are myopic and informed agents are sophisticated. The fact that the informed agents do not know the true A precisely introduces two complications.

The first complication is that informed agents may have an incentive to distort policy to make their own learning more likely. This incentive is strongest for the informed poor, since learning both facilitates the transfer of political power in their direction, and enables future

[18] An alternative assumption would be that the agent learns that the true A lies within an interval around $E[A]$; formally that $A \in \left[E_{F_j}[A] - \frac{\mu}{\tau \overline{y}}, E_{F_j}[A] + \frac{\mu}{\tau \overline{y}}\right]$. But this alternative assumption violates the spirit of the story that agents are inattentive to policies that are roughly consistent with their prior beliefs.

policy to be more finely tailored to the true A. By contrast, for the informed rich, these incentives are in conflict. Retaining the spirit of the baseline model, we assume that learning poses a net harm to the informed rich. Notice that the benefit of a more finely tailored policy will be strongest when the true A is far from what the informed agents expect. It suffices, then, to restrict the beliefs of the informed to be sufficiently concentrated around A_I, so that the informed rich put probability zero on learning being beneficial.[19] Additionally, we suppose μ is sufficiently large that, for any policy that might be implemented in equilibrium, the informed agents will not foresee themselves directly updating their beliefs.[20] Thus, as in the baseline model, the role of learning is entirely focused on updating by the misinformed.

The second complication is that, within this framework, learning by the misinformed generates a role-reversal — the (initially) misinformed will perfectly learn A, whilst the (initially) informed, will continue to be uncertain about its true value. Moreover, although the (initially) informed do not directly learn themselves, since they are sophisticated, they will observe that the (initially) misinformed have learned through their changing behavior. How then should we assess the preferences of the (initially) informed? Should they continue to assess policies according to their prior? (That would require some sort of myopia on their part.) Or should they update their beliefs according to how they observe the (initially) misinformed change their behavior? We take the latter approach, assuming that whenever the informed agents notice that the misinformed have learned, that they update beliefs themselves.

With these assumptions, we are ready to characterize the optimal behavior of the informed rich. Define $A^{\dagger}(\tau) = A_M + \frac{\mu}{\ln(\tau\overline{y})}$. Given a policy τ, the misinformed will learn if $A > A^{\dagger}(\tau)$. Since the (initially) informed do not know A precisely, they will not know whether a given policy τ will induce learning or not (i.e. whether the true A will lie above or below

[19]Formally, let the support of F_I be $[\underline{A}, \overline{A}]$ (with $\underline{A}_I < A_I < \overline{A}_I$). We assume that $v\left(\frac{E_{F_I}[A]}{y_H}; y_H, E_{F_I}[A]\right) \geq v\left(\frac{A}{y_L}; y_H, A\right)$ for all $A \in [\underline{A}, \overline{A}]$. Hence, no matter what is learned *ex post*, the the informed rich always do better to retain political power and implement their *ex ante* ideal policy, than to cede political power, and have a more finely tailored policy implemented by the informed poor.

[20]It suffices that $\mu > \max\{\overline{A}_I - A_I, A_I - \underline{A}_I\} \ln\left(A_I \frac{\overline{y}}{y_L}\right)$.

$A^\dagger(\tau)$). From their perspective, learning by the (initially) misinformed will appear stochastic. Accordingly, the (initially) informed will choose policies assessing the likely second period outcomes that will follow.

The expected lifetime utility of the informed rich from a generic policy τ is:

$$v(\tau; y_H, A_I) + \beta \left(\int_{\underline{A}_I}^{\max\{\underline{A}_I, A^\dagger(\tau)\}} v(\tau_{LI}; y_H, A_I) dF_I(A) + \int_{\min\{A^\dagger(\tau), \overline{A}_I\}}^{\overline{A}_I} v(\tau_{HI}; y_H, A_I) dF_I(A) \right) \quad (2)$$

This expression is in three parts. The first part represents first period utility given a policy τ. The second and third terms (enclosed by parentheses) represent the expected second period utility, where the first period policy either does or does not induce learning.

The first order condition is:

$$-y_H + \frac{A_I}{\tau} - \beta \underbrace{\Delta(y_H, A^\dagger(\tau))}_{<0} \cdot \underbrace{\left(-\frac{\mu}{\tau[\ln(\tau\overline{y})]^2} \right)}_{<0} f_I(A^\dagger(\tau)) = 0$$

where $\Delta(y_H, A) = A[\ln(\tau(y_L, A)) - \ln(\tau(y_H, A_I))] - (\tau(y_L, A) - \tau(y_H, A_I)]y_H$ is the utility loss that the informed rich suffer after learning that the true value of public goods is A.

The first order condition is in two parts. The first two terms represent the net marginal benefit (to first period utility) of increasing τ. The third term represents the contribution to second period utility. It is itself the product of two terms. The first, $\beta\Delta(y_H, A^\dagger)$, is the discounted loss that the informed rich suffer from learning if the true A is $A^\dagger$. The second is the probability that a marginal increase in τ will induce learning. Notice that $\beta\Delta(y_H, A)$ was precisely the loss from learning that the informed rich suffered in the baseline model; the only difference here is that it is received probabilistically. It is straightforward to see that the second period component of marginal utility is weakly negative — hence, as in the baseline model, fear of the slippery slope will induce the informed rich to support policies below their stage game ideal, to strategically (though stochastically) prevent learning.

Let $\underline{\tau}(\mu)$ and $\overline{\tau}(\mu)$ be implicitly defined by the conditions: $A^\dagger(\underline{\tau}) = \overline{A}_I$ and $A^\dagger(\overline{\tau}) = \underline{A}_I$. $\underline{\tau}$ is the highest policy for which learning will definitely not occur, and $\overline{\tau}$ is the lowest policy for which it definitely will. If $\tau \in (\underline{\tau}, \overline{\tau})$, then informed rich will be uncertain whether learning will occur or not.

Let $\hat{\tau}(\mu)$ be the solution to the problem in (2), where τ is constrained to be in the interval $\tau \in [\underline{\tau}(\mu), \overline{\tau}(\mu)]$. We can show that $\hat{\tau} < \tau_{HI}$ whenever $A^{\dagger}(\tau_{HI}) < \overline{A}$. Define $\overline{\mu} = (\overline{A}_I - A_M)\ln(\tau_{HI}\overline{y})$. Finally, to avoid the problem of Condorcet cycles that arose in Proposition 2.B, we suppose that $\underline{\tau} \geq \tau_{LM}$. (This is the analogue to the assumption in Proposition 2.A that $\underline{\tau}_{HI} \geq \tau_{LM}$.)

Proposition 7. *Suppose that informed agents are sophisticated and that the (average) divergence in beliefs is large so that the the informed rich are initially pivotal. There exists a threshold $\underline{\mu} < (\underline{A}_I - A_M)\ln(\tau_{HI}\overline{y})$ such that the equilibrium first period policy satisfies:*

$$\tau_1^* = \begin{cases} \tau_{HI} & \text{if } \mu < \underline{\mu} \\ \hat{\tau}(\mu) & \text{if } \mu \in (\underline{\mu}, \overline{\mu}) \\ \tau_{HI} & \text{if } \mu > \overline{\mu} \end{cases}$$

Proposition 7 mirrors Proposition 2.A from Section 4.2.1. If at their ideal stage policy, the informed rich believe that learning will definitely not occur, then they will provide their ideal policy τ_{HI}. Analogous to Proposition 2.A, this will occur if either μ (or analogously $\tau^{\dagger}$ in Proposition 2.A) is sufficiently high. If at their ideal policy, learning is guaranteed, and the distortion needed to prevent (or reduce the chance of) learning is sufficiently large, then the informed rich will provide their stage-ideal policy and concede political power. For intermediate cases, the informed rich will downwardly skew policy, to reduce (though not necessarily eliminate) the probability of learning and transfer of political power. Finally, if we allowed $\underline{\tau} < \tau_{LM}$, then similar to Proposition 2.B, a Condorcet cycle may exist.

The purpose of this section was to demonstrate that our baseline results would broadly continue to hold under a more standard learning technology. Clearly, given the above analysis, our baseline results were not merely an artifact of our special learning technology – they capture a dynamic that remains present under alternative modelling assumptions. Naturally, there are opportunities to modify the learning technology in other ways. For example, rather than observing her true utility, we might instead posit that agents receive one of three signals: either utility is roughly what was expected, that it was significantly below what was expected, or that it was significantly above what was expected. This would make learning more coarse. And the complications we discussed above about what the sophisticated agents impute about what is learned by the misinformed will manifest even more strongly. Nevertheless, we expect that the underlying dynamic that we highlight will continue to be present, though perhaps muddied by these other considerations.

A.6 Plugging Condorcet Holes

In our baseline model, the preferences of sophisticated agents were not always single-peaked, and this occasionally resulted in the emergence of Condorcet cycles. In this section, we explore methods that identify a clear equilibrium policy even when a Condorcet winner fails to exist. We focus on 'Condorcet methods': those that select the majority winner whenever it exists.[21] Thus, these results will always agree with our baseline results whenever those predicted a stable equilibrium policy.

For concreteness, we consider the scenario in Proposition 2B, where the informed are sophisticated and the misinformed are not, and where $\tau^\dagger < \tau_{LM} < \tau_{HI}$. As before, assume that both the informed rich and misinformed poor strictly prefer $\tau^\dagger$ to τ_{HI}. If so, then by Proposition 2B, we know that no majority winner exists.

The methods we consider will require voters to provide a ranking over a set of alternatives, rather than merely vote for their most preferred policy — which is a departure from the

[21] A different approach would be to introduce probabilistic voting (see Lindbeck & Weibull, 1987; Persson & Tabellini, 1999). However, the equilibrium policy with probabilistic voting need not coincide with the majority winner when it exists. Moreover, given the discontinuities in the lifetime utilities of sophisticated agents, equilibria in pure strategies are not guaranteed to exist.

baseline model. (In particular, the behavior of political parties within the election mechanism becomes less clear.) In the analysis that follows, we assume that voters always sincerely rank the available outcomes. To simplify the analysis, we will assume that the misinformed rich are a negligible fraction of the voting population (though the methods do not rely on this assumption in any crucial way).

Condorcet methods are typically sensitive to the set of policies that voters are invited to rank. We focus on methods that satisfy the Generalized Condorcet criterion (or Smith criterion), which requires that the chosen policy lie within the top cycle[22]. Additionally, we focus on methods that are insensitive to the inclusion or exclusion of policies outside the top cycle (i.e. which satisfy the Independence of Smith-Dominated Alternatives property). In general, these methods will remain sensitive to the inclusion or exclusion of various policies within the top-cycle. For the purposes of this analysis, we will take the set of alternatives to be $\tau \in [\tau^{\dagger}, \tau_{HI}]$. The alternatives in this set form a strict subset of the top-cycle, however, our chosen set is reasonable in that it is the convex hull of the three 'focal' policies that generated the Condorcet cycle — $\tau^{\dagger}$, τ_{LM}, and τ_{HI}.

We consider two well studied methods. First is the 'ranked pairs' method (see Tideman, 1987), which works as follows: (i) for each pair of outcomes, calculate the 'strength of victory' (i.e. excess support) for the majority preferred outcome; (ii) sort the pairs by strength of victory; (iii) build the social ranking, starting with the pair with the largest strength of victory, and ignoring subsequent pairs that would introduce an intransitive cycle. In effect, the ranked pairs method generates a transitive social preference by removing from comparison the majority preferred pairs that have the lowest support and whose inclusion would induce a Condorcet cycle.

Applied to our model, the highest ranked outcome that emerges from the ranked pairs

[22]The top cycle is the smallest set of policies having the property that every policy within the set is majority preferred to every policy outside the set. Let $\overline{\tau}_{HI} > \tau_{HI}$ be the policy defined by $v(\overline{\tau}; y_H, A_I) = v(\tau^{\dagger}; y_H, A_I)$. The top cycle is $[\tau^{\dagger}, \overline{\tau}_{HI}]$.

method corresponds to the most preferred policy (within the alternative set $[\tau^\dagger, \tau_{HI}]$) of the most populous group in the polity.[23] If the informed rich are the largest group, the ranked pairs method selects the policy $\tau^\dagger$; if the informed poor are largest, then the policy τ_{HI} is selected; and if the misinformed poor are most sizeable, then the policy is τ_{LM}.

The second method that we study is Tideman's Alternative, which is essentially a method of instant run-off voting. Votes are distributed amongst the various alternatives according to each voter's highest preference; a plurality losing policy is eliminated and votes are redistributed by (the affected voters') next highest available preference. The procedure repeats until a single alternative remains.

In the context of our model, all but the three 'focal' policies are immediately eliminated (since they are not the most preferred policies of any voter). Then, the policy that survives the run-off procedure will be the second highest ranked policy (amongst these three) of the *least* popular group. (Intuitively, when that group is eliminated, its votes will be reallocated by their second preference, which is sufficient to give one of the two remaining outcomes a majority.) Hence, if the informed rich are smallest, Tideman's Alternative selects τ_{HI}; if the informed poor are smallest, the procedure selects τ_{LM}; and if the misinformed poor are smallest, it selects $\tau^\dagger$.

Several points are worth noting. First, though the methods contemplated a continuum of alternatives, the highest ranked policy selected by each method coincided with one of the three 'focal' policies that generated the Condorcet cycle. Second, for each method, the solution is sensitive to group size. Each of the focal policies can be rationalized for an appropriately chosen population profile of voters. Indeed, small changes in the relative sizes of the groups can produce dramatic changes in the selected policy. Third, even fixing the population profile, the variant methods are not guaranteed to select the same policy, and will often disagree. Thus, though these methods can serve as a refinement tool that selects

[23]Full details for how this result was computed are available by request from the authors.

a policy when a Condorcet cycle arises, the policy selected by any given method need not stand out as an obviously best policy.

B Proofs

Proof of Proposition 1. Recall that the ranking of stage game ideal policies is $\tau_{HM} < \tau_{HI} < \tau_{LM} < \tau_{LI}$, and that all myopic agents, as well as the sophisticated misinformed agents (since $\Delta(y, A_M) = 0$), have single-peaked preferences. The sophisticated informed agents have potentially non-single-peaked preferences. However, the preferences of the sophisticated informed rich are strictly decreasing whenever $\tau \geq \tau_{HI}$, while the preferences of the sophisticated informed poor are strictly increasing whenever $\tau \leq \tau_{LI}$

It suffices to show that τ_{LM} is a majority winner in the first period. Take any policy $\tau < \tau_{LM}$. Then all poor agents, whether informed or not, and whether sophisticated or not — a majority — strictly prefer τ_{LM} to τ, since they have increasing utility in this region. Similarly, all agents other than the informed poor — a majority — strictly prefer τ_{LM} to any $\tau > \tau_{LM}$, since utility is strictly decreasing in this region for those agents. Hence τ_{LM} is the majority winner. □

Proof of Propositions 2.A and 2.B. The ranking of stage game ideal policies is now: $\tau_{HM} < \tau_{LM} < \tau_{HI} < \tau_{LI}$. Let $\underline{\tau}_{HI} < \tau_{HI}$ be the lowest policy that the informed rich will accept that prevents learning. This satisfies:

$$v(\underline{\tau}_{HI}; y_H, A_I) + \beta v(\tau_{HI}; y_H, A_I) = v(\tau_{HI}; y_H, A_I) + \beta v(\tau_{LI}; y_H, A_I)$$

Simplifying, we have that $\underline{\tau}_{HI}$ is the solution to:

$$1 - \left(\frac{\underline{\tau}_{HI}}{\tau_{HI}}\right) + \ln\left(\frac{\underline{\tau}_{HI}}{\tau_{HI}}\right) = \beta\left[1 - \frac{\tau_{LI}}{\tau_{HI}} + \ln\left(\frac{\tau_{LI}}{\tau_{HI}}\right)\right] \tag{3}$$

Let t_{HI} denote the first period policy that maximizes the informed rich's lifetime utility. By construction, we know that:

$$t_{HI} = \begin{cases} \tau^\dagger & \text{if } \tau^\dagger \in [\underline{\tau}_{HI}, \tau_{HI}] \\ \tau_{HI} & \text{if } \tau^\dagger < \underline{\tau}_{HI} \text{ or } \tau^\dagger > \tau_{HI} \end{cases}$$

Recall that the informed poor have strictly increasing utility in the region $\tau \leq \tau_{LI}$, and that $t_{HI} \leq \tau_{HI} < \tau_{LI}$. It follows that, for any $\tau < t_{HI}$, t_{HI} is strictly preferred to τ by both the informed poor and the informed rich — who together constitute a majority.

Suppose $t_{HI} \geq \tau_{LM}$. (In particular, this will be true if $t_{HI} = \tau_{HI}$.) Recall that all misinformed agents have single-peaked preferences. Then, for any $\tau > t_{HI}$, t_{HI} is preferred to τ by the informed rich and all misinformed agents (since $\tau_{HM} < \tau_{LM} \leq t_{HI}$) — who together constitute a majority. If so, then t_{HI} is a majority winner.

Suppose instead that $t_{HI} < \tau_{LM}$. (This requires that $t_{HI} = \tau^\dagger$, which in turn requires that $\tau^\dagger \in (\underline{\tau}_{HI}, \tau_{LM})$.) Then t_{HI} cannot be majority preferred since τ_{LM} is preferred to it by both the informed and misinformed poor (who together constitute a majority). But τ_{LM} cannot be majority preferred since τ_{HI} is preferred to it by both the informed poor and the informed rich (who together constitute a majority). Can τ_{HI} be majority preferred? Clearly there is no policy $\tau > \tau_{HI}$ that is preferred to τ_{HI} by a majority. The informed rich prefer $\tau^\dagger$ to τ_{HI}. If the misinformed poor also prefer $\tau^\dagger$ to τ_{HI}, then τ_{HI} cannot be majority preferred. This requires that:

$$v(\tau^\dagger; y_L, A_M) \geq v(\tau_{HI}; y_L, A_M)$$

Let $\underline{\tau}_{LM}(\beta)$ be the lowest policy that the misinformed poor would accept in preference to τ_{HI}, if the former prevented learning and the latter did not. (It will prove useful to define $\underline{\tau}_{LM}$ for generic β, though of course, in this proposition, we take $\beta = 0$, so that the misinformed poor are indifferent to whether there is learning or not.) We have:

$$\frac{\tau_{HI}}{\tau_{LM}}\left(1-\frac{\underline{\tau}_{LM}(\beta)}{\tau_{HI}}\right)+\ln\left(\frac{\underline{\tau}_{LM}(\beta)}{\tau_{HI}}\right)=\beta\left[\frac{\tau_{HI}}{\tau_{LM}}-1+\ln\left(\frac{\tau_{LM}}{\tau_{HI}}\right)\right] \tag{4}$$

By assumption, the informed rich prefer $\tau^\dagger$ to τ_{HI} since $\tau^\dagger \in (\underline{\tau}_{HI}, \tau_{LM})$. If, in addition, $\tau^\dagger \geq \underline{\tau}_{LM}$, then the misinformed poor will prefer $\tau^\dagger$ to τ_{HI}. Hence, if $\tau^\dagger \in (\max\{\underline{\tau}_{HI}, \underline{\tau}_{LM}\}, \tau_{LM})$, then a majority prefer $\tau^\dagger$ to τ_{HI} and so there is no majority winner. (Instead, we have found a Condorcet cycle.) By contrast, if $\tau^\dagger < \max\{\underline{\tau}_{HI}, \underline{\tau}_{LM}\}$, then τ_{HI} is a majority winner. □

Proof of Proposition 5. [24] We first note that if τ is a majority winner, then $\tau \in [\tau_{HI}, \max\{\tau_{HI}, \tau^\dagger\}]$. To see this, note that τ_{HI} is preferred to any $\tau < \tau_{HI}$ by all informed agents — a majority. Similarly, the informed rich and the misinformed agents — who together constitute a majority — prefer $\max\{\tau_{HI}, \tau^\dagger\}$ to any larger policy τ. (For the informed rich, this is straightforward to see since their utility is decreasing beyond τ_{HI} and $\tau_{HI} \leq \max\{\tau_{HI}, \tau^\dagger\}$. Similarly, the utility of the misinformed poor is decreasing beyond $\max\{\tau_{LM}, \tau^\dagger\}$, and $\tau_{LM} < \tau_{HI}$.)

If $\tau^\dagger \leq \tau_{HI}$, it follows immediately that τ_{HI} is the majority winner. Next, suppose that $\tau^\dagger > \tau_{HI}$. Misinformed agents may be willing to support such a policy given that it induces learning and τ_{HI} does not. Let $\overline{\tau}_{LM} > \tau_{HI}$ be the highest policy that the misinformed poor will accept that induces learning, in preference to τ_{HI} which does not. This satisfies:

$$v(\overline{\tau}_{LM}; y_L, A_M) + \beta v(\tau_{LM}; y_L, A_M) = v(\tau_{HI}; y_L, A_M) + \beta v(\tau_{HI}; y_L, A_M)$$

Simplifying, we have that $\overline{\tau}_{LM}$ is the solution to:

$$\frac{\tau_{HI}}{\tau_{LM}}\left(1-\frac{\overline{\tau}_{LM}}{\tau_{HI}}\right)+\ln\left(\frac{\overline{\tau}_{LM}}{\tau_{HI}}\right)=\beta\left[1-\frac{\tau_{HI}}{\tau_{LM}}+\ln\left(\frac{\tau_{HI}}{\tau_{LM}}\right)\right] \tag{5}$$

If $\tau^\dagger > \overline{\tau}_{LM}$, then τ_{HI} is preferred to $\tau^\dagger$, and indeed to all $\tau > \tau_{HI}$, by both the informed rich and the misinformed poor. If so, τ_{HI} is the majority winner.

[24] The proof of Proposition 3 builds on the proof of Proposition 5, so we present this proof first.

Suppose instead that $\tau^\dagger \in (\tau_{HI}, \overline{\tau}_{LM})$. Then both the informed and misinformed poor — a majority — will prefer $\tau^\dagger$ to τ_{HI}, so that τ_{HI} cannot be a majority winner. Is $\tau^\dagger$ a majority winner? We have previously shown that $\tau^\dagger$ is majority preferred to any $\tau > \tau^\dagger$. For a policy $\tau' < \tau^\dagger$ to be majority preferred to $\tau^\dagger$, it must be that τ' is strictly preferred to $\tau^\dagger$ by both the informed rich and the misinformed poor.

Let $\tilde{\tau}_{HI}(\tau^\dagger; \beta) < \tau_{HI}$ denote the lowest policy (not inducing learning) that the informed rich would accept in preference to $\tau^\dagger$. (As before, it will be helpful for future reference to allow $\tilde{\tau}_{HI}$ to be a function of β. But in this context, we know that $\beta = 0$ for the informed rich.) This satisfies:

$$v(\tilde{\tau}_{HI}; y_H, A_I) + \beta v(\tau_{HI}; y_H, A_I) = v(\tau^\dagger; y_H, A_I) + \beta v(\tau_{LI}; y_H, A_I)$$

Simplifying, we have that $\tilde{\tau}_{HI}$ is the solution to:

$$\frac{\tau^\dagger - \tilde{\tau}_{HI}}{\tau_{HI}} + \ln\left(\frac{\tilde{\tau}_{HI}}{\tau^\dagger}\right) = \beta\left[1 - \frac{\tau_{LI}}{\tau_{HI}} + \ln\left(\frac{\tau_{LI}}{\tau_{HI}}\right)\right] \tag{6}$$

It is straightforward to show, by the implicit function theorem, that $\frac{\partial \tilde{\tau}_{HI}}{\partial \tau^\dagger} < 0$. The further to the right of τ_{HI} is $\tau^\dagger$, the greater will be the range of policies to the left of τ_{HI} that the informed rich would be willing to accept instead. Additionally, $\tilde{\tau}_{HI} \uparrow \tau_{HI}$ as $\tau^\dagger \downarrow \tau_{HI}$.

Similarly, let $\tilde{\tau}_{LM}(\tau^\dagger; \beta) > \tau_{LM}$ denote the highest policy (not inducing learning) that the misinformed poor would accept in preference to $\tau^\dagger$. This satisfies:

$$v(\tilde{\tau}_{LM}; y_L, A_M) + \beta v(\tau_{HI}; y_L, A_M) = v(\tau^\dagger; y_L, A_M) + \beta v(\tau_{LM}; y_L, A_M)$$

Simplifying, we have that $\tilde{\tau}_{LM}$ is the solution to:

$$\frac{\tau^\dagger - \tilde{\tau}_{LM}}{\tau_{LM}} + \ln\left(\frac{\tilde{\tau}_{LM}}{\tau^\dagger}\right) = \beta\left[\frac{\tau_{HI}}{\tau_{LM}} - 1 + \ln\left(\frac{\tau_{LM}}{\tau_{HI}}\right)\right] \tag{7}$$

Again, by the implicit function theorem, $\frac{\partial \tilde{\tau}_{LM}}{\partial \tau^\dagger} > 0$. The further to the right of τ_{HI} is $\tau^\dagger$, the greater will be the range of policies to the right of τ_{LM} that the misinformed poor would be willing to accept instead. Additionally, $\tilde{\tau}_{LM} < \tau_{HI}$ when $\tau^\dagger = \tau_{HI}$.

If $\tilde{\tau}_{HI}(\tau^\dagger) < \tilde{\tau}_{LM}(\tau^\dagger)$, then any policy $\tau' \in (\tilde{\tau}_{HI}, \tilde{\tau}_{LM})$ is majority preferred to $\tau^\dagger$ (since it is preferred by the informed rich and the misinformed poor). Moreover, this generates a Condorcet cycle, since τ_{HI} is preferred to any such τ' by a majority (i.e. the informed agents), $\tau^\dagger$ is preferred to τ_{HI} by a majority (i.e. the poor agents), and τ' is preferred to $\tau^\dagger$ by a majority (i.e. the informed rich and the misinformed poor). By contrast, if the condition is not met, then $\tau^\dagger$ is a majority winner.

Finally, since $\tilde{\tau}_{LM}(\tau^\dagger) < \tilde{\tau}_{HI}(\tau^\dagger)$ when $\tau^\dagger = \tau_{HI}$, and since $\frac{\partial \tilde{\tau}_{HI}}{\partial \tau^\dagger} < 0$ and $\frac{\partial \tilde{\tau}_{LM}}{\partial \tau^\dagger} > 0$, then there exists some $\tilde{\tau}$ s.t. $\tilde{\tau}_{LM}(\tilde{\tau}) = \tilde{\tau}_{HI}(\tilde{\tau})$, and the condition $\tilde{\tau}_{HI}(\tau^\dagger) < \tilde{\tau}_{LM}(\tau^\dagger)$ is satisfied only if $\tau^\dagger > \tilde{\tau}$. □

Proof of Proposition 3. First, suppose that $\tau^\dagger \geq \tau_{HI}$. Then, the lifetime preferences of the (sophisticated) informed rich are single-peaked and achieve a maximum at τ_{HI} — which is qualitatively similar to what their preferences would be were they myopic. Then, by the same logic as in the proof of Proposition 5, if $\tau^\dagger > \overline{\tau}_{LM}$, the majority preferred policy will be τ_{HI}. If $\tau^\dagger \in [\tau_{HI}, \overline{\tau}_{LM}]$, then a majority winner will exist only if $\tilde{\tau}_{HI}(\tau^\dagger) \geq \tilde{\tau}_{LM}(\tau^\dagger)$, and if so, the majority winner will be $\tau^\dagger$.

The only potential difference from the proof of Proposition 5 is in the behavior of the functions $\tilde{\tau}_{HI}(\tau^\dagger)$ and $\tilde{\tau}_{LM}(\tau^\dagger)$, defined by equations (6) and (7) above. In particular, since the informed rich are now sophisticated, $\tilde{\tau}_{HI}$ should be calculated using $\beta > 0$. It remains the case that $\frac{\partial \tilde{\tau}_{HI}}{\partial \tau^\dagger} < 0$ and $\frac{\partial \tilde{\tau}_{LM}}{\partial \tau^\dagger} > 0$. However, now $\tilde{\tau}_{HI}(\tau_{HI}) = \underline{\tau}_{HI} < \tau_{HI}$ (where $\underline{\tau}_{HI}$ is defined by equation (3) in the proof of Proposition 2.A). Also, by construction, $\tilde{\tau}_{LM}(\tau_{HI}) = \underline{\tau}_{LM}$ (where $\underline{\tau}_{LM} > \tau_{LM}$ is defined by equation (3) in the proof of Proposition 2.B, though now with $\beta > 0$). Hence, if $\underline{\tau}_{HM} \leq \underline{\tau}_{LM}$, then $\tilde{\tau}_{HI}(\tau^\dagger) \leq \tilde{\tau}_{LM}(\tau^\dagger)$ for all $\tau^\dagger \geq \tau_{HI}$, and so there will be no majority winner. By contrast, if $\underline{\tau}_{HM} > \underline{\tau}_{LM}$, then there will exist $\tilde{\tau} > \tau_{HI}$ s.t. $\tilde{\tau}_{HI}(\tau^\dagger) > \tilde{\tau}_{LM}(\tau^\dagger)$ whenever $\tau^\dagger < \tilde{\tau}$. If so, $\tau^\dagger$ will be the majority winner, and if not, a Condorcet cycle will exist.

Next, suppose $\tau^\dagger < \tau_{HI}$. Now, the informed rich may have an incentive to strategically manipulate policy to prevent learning, whilst the misinformed may seek to do so to induce learning. Let us explicitly differentiate these. For some small $\varepsilon > 0$, let $\tau^\dagger_- = \tau^\dagger - \varepsilon$ denote the highest policy that prevents learning, and $\tau^\dagger_+ = \tau^\dagger + \varepsilon$ denote the lowest policy that induces it. The informed agents (who constitute a majority) will prefer $\tau^\dagger_-$ to any policy $\tau < \tau^\dagger_-$. The poor agents (who constitute a majority) must prefer $\tau^\dagger_+$ to $\tau^\dagger_-$, since they perceive learning as beneficial, and the informed agents will prefer τ_{HI} to any $\tau \in [\tau^\dagger_+, \tau_{HI})$. Hence, no policy $\tau < \tau_{HI}$ can be majority winning. Moreover, the informed rich and the misinformed agents (who constitute a majority) prefer τ_{HI} to any $\tau > \tau_{HI}$. The only candidate to be a majority winner is τ_{HI}.

We must check if there is a policy τ' that defeats τ_{HI} in a pair-wise contest. Given the above reasoning, if such a policy exists, it must be that $\tau' \leq \tau^\dagger_-$, and the coalition supporting it must include both the informed rich and the misinformed poor. Then, by construction, $\tau' \geq \underline{\tau}_{HI}$ (which guarantees that the informed rich prefer it to τ_{HI}) and that $\tau' \leq \underline{\tau}_{LM}$ (which is required for the misinformed poor prefer it to τ_{HI}). Hence, there will be no majority winner (and thus a Condorcet cycle will exist) if $\underline{\tau}_{HI} \leq \underline{\tau}_{LM}$. Else, τ_{HI} will be a majority winner. $\square$

Proof of Proposition 4. Begin by considering how learning affects beliefs. Given a policy τ, an agent with belief $A \leq A_I$ updates their belief if:

$$A < A_I - \frac{\mu}{\ln(\tau\overline{y})} = \mathcal{A}(\tau)$$

Notice that learning truncates the distribution of beliefs from below. Let A_t denote the lowest belief about A that some agent may profess at the start of time t, given learning to that point. It must be that $A_t = \max\{\mathcal{A}(\tau_{t-1}), A_{t-1}\}$. Then, given a sequence of policies $\{\tau_0, \ldots, \tau_{t-1}\}$, the induced sequence of minimal beliefs $\{A_0, \ldots, A_t\}$ is weakly increasing.

At the start of time t, the fraction of agents who have income y_i and are informed is $\phi_i[\gamma_i + (1-\gamma_i)F_i(A_t)]$ — of whom, a fraction $\phi_i\gamma_i$ were informed from the start, and a fraction $\phi_i(1-\gamma_i)F_i(A_t)$ became informed through acquaintance with policy. Recall, the effective income of a type (y, A) agent is $x(y, A) = y\frac{A_I}{A} \geq y$. Then, the distribution of effective incomes at time t is:

$$G(x, | A_t) = Pr[x(y, A) \leq x \,|\, A_t] = \sum_{i \in \{H,L\}} \mathbf{1}[x \geq y_i]\phi_i \left(\gamma_i + (1-\gamma_i)\left[F_i(A_t) + 1 - F_i\left(\frac{y_i}{x}A_I\right)\right]\right)$$

Suppose, at time t, the agent with the median effective income is informed and rich. This requires that $G(y_H \,|\, A_t) \geq \frac{1}{2}$ and $\lim_{x\uparrow y_H} G(x \,|\, A_t) < \frac{1}{2}$. The first condition is guaranteed to hold given the assumptions that $\gamma_L, \gamma_H > \frac{1}{2}$. The second condition requires that:

$$\phi_L\left(\gamma_L + (1-\gamma_L)\left[F_L(A_t) + 1 - F_L\left(\frac{y_L}{y_H}A_I\right)\right]\right) < \frac{1}{2}$$

$$A_t < F_L^{-1}\left(F_L\left(\frac{y_L}{y_H}A_I\right) - \frac{1-\frac{1}{2\phi_L}}{1-\gamma_L}\right) = \underline{A}$$

The assumptions that the poor are a majority ($\phi_L > \frac{1}{2}$), but that the informed poor are initially a minority (i.e. $\phi_L\gamma_L < \frac{1}{2}$), ensures that $\frac{1-\frac{1}{2\phi_L}}{1-\gamma_L} \in (0, 1)$. Then since F is strictly increasing, $\underline{A} < \frac{y_L}{y_H}A_I$.

Suppose, instead, that the agent with the median effective income is informed and poor. This requires that $G(y_L \,|\, A_t) \geq \frac{1}{2}$, and so:

$$\phi\left(\gamma_L + (1-\gamma_L)F_L(A_t)\right) \geq \frac{1}{2}$$

$$A_t \geq F_L^{-1}\left(\frac{\frac{1}{2\phi} - \gamma_L}{1-\gamma_L}\right) = \overline{A}$$

where $\frac{1}{2\phi_L} - \gamma_L > 0$ by assumption. Of course, if $A_t \geq \overline{A}$, then $A_{t'} \geq \overline{A}$ for all $t' > t$, since $\{A_t\}$ is a non-decreasing sequence.

If at time t, $A_t \in (\underline{A}_t, \overline{A}_t)$, so that neither the informed poor nor the informed rich are pivotal, then the median effective income must be below y_H but above y_L (i.e. the median effective income earner is from the misinformed poor). Denoting it by $X(A_t)$ we have:

$$\phi_L\left(\gamma_L + (1-\gamma_L)\left[F_L(A_t) + 1 - F_L\left(\frac{y_L}{X(A_t)}A_I\right)\right]\right) = \frac{1}{2}$$

$$X(A_t) = \frac{y_L A_I}{F_L^{-1}\left(F_L(A_t) + \frac{1-\frac{1}{2\phi_L}}{1-\gamma_L}\right)}$$

Moreover, since this agent is myopic, and using similar logic to Proposition 1, in this scenario, we know that the equilibrium policy at time t will simply be the ideal policy of the median effective income earner. Denote this:

$$\tau(A_t) = \frac{A_I}{X(A_t)} = \frac{F_L^{-1}\left(F_L(A_t) + \frac{1-\frac{1}{2\phi_L}}{1-\gamma_L}\right)}{y_L}$$

By assumption $A_0 < \underline{A}$, so the informed rich are initially pivotal. Let $\underline{\tau}$ be defined by $\mathcal{A}(\underline{\tau}) = \underline{A}$. $\underline{\tau}$ is the largest policy that the informed rich can offer and still retain political power. Also by assumption, $\underline{A} < \mathcal{A}(\tau_{HI})$, which implies that $\underline{\tau} < \tau_{HI}$ — the informed rich must downwardly distort policy to retain political power. (Obviously, they have no incentive to distort below $\underline{\tau}$.) If it is optimal to implement this policy, then they will retain political power, and by the stationarity in the problem, they will offer the same policy again in each future period. Their lifetime utility will be $\frac{1}{1-\beta}v(\underline{\tau}; y_H, A_I)$.

Suppose the informed rich implement a policy $\tau > \underline{\tau}$. Let $V_{HI}(A)$ denote the lifetime utility of the informed rich when the current lowest belief is $A > \underline{A}$. Given the above discussion, the current policy will be $\tau(A)$ and the evolution of beliefs will be governed by $A' = \max\{\mathcal{A}(\tau(A)), A\}$. Hence, V_{HI} is characterized by the Bellman Equation:

$$V_{HI}(A) = v(\tau(A); y_H, A_I) + \beta V_{HI}(\max\{\mathcal{A}(\tau(A)), A\})$$

It is easy to verify that the Bellman Operator satisfies Blackwell's sufficient conditions (since the current policy is chosen independently of the evolution of the state), and so a unique value function exists. Straightforwardly, since $\tau(A) = \tau_{LI}$ for all $A \geq \overline{A}$, then $V_{HI}(A) = \frac{1}{1-\beta}v(\tau_{LI}; y_H, A_I)$ for all $A \geq \overline{A}$ (which is constant in A).

Next, since $\tau(A)$ is increasing in A and $v(\tau; y_H, A_I)$ is strictly decreasing in τ for $\tau > \tau_{HI}$, it follows that the Bellman operator maps decreasing functions onto decreasing functions. Hence V_{HI} is weakly decreasing in A. Moreover, it is strictly decreasing whenever $A \in (\underline{A}, \overline{A})$. To see this, suppose $\underline{A} < A^1 < A^2 < \overline{A}$. Then, $\tau(A_1) < \tau(A_2)$, and so $v(\tau(A^1); y_H, A_I) > v(\tau(A^2); y_H, A_I)$. Furthermore, $\max\{\mathcal{A}(\tau(A^1)), A^1\} \leq \max\{\mathcal{A}(\tau(A^2)), A^2\}$, and so, since V is weakly decreasing:

$$V(A^1) = v(\tau(A^1); y_H, A_I) + \beta V(\max\{\mathcal{A}(\tau(A^1)), A^1\}) > v(\tau(A^2); y_H, A_I) + \beta V(\max\{\mathcal{A}(\tau(A^2)), A^2\}) = V(A^2)$$

The problem of the informed rich, if they decide to cede political power, is:

$$\max_{\tau \in [\underline{\tau}, \tau_{HI}]} v(\tau; y_H, A_I) + \beta V_{HI}(\mathcal{A}(\tau))$$

The first order condition implies that:

$$v'(\tau^*; y_H, A_I) + \beta V'_{HI}(\mathcal{A}(\tau^*)) \frac{\partial \mathcal{A}(\tau^*)}{\partial \tau} = 0$$

Since $\tau^* \leq \tau_{HI}$ and $\mathcal{A}(\tau_{HI}) < \overline{A}$ (by assumption), then $\mathcal{A}(\tau^*) \in (\underline{A}, \overline{A})$, and so $\frac{\partial \mathcal{A}(\tau^*)}{\partial \tau} > 0$ and $V'_{HI} < 0$. But then, the first order conditions imply that $v'(\tau^*; y_H, A_I) > 0$, and so $\tau^* < \tau_{HI}$. Even if the informed rich plan to cede power, they should still downwardly distort policy in the current period. Let $V^*_{HI} = v(\tau^*; y_H, A_I) + \beta V_{HI}(\mathcal{A}(\tau^*))$ denote the agent's lifetime utility if they cede power.

Finally, should the informed rich cede power or not? This involves a comparison of the utilities from the two approaches. They should retain power if:

$$\frac{1}{1-\beta} v(\underline{\tau}; y_H, A_I) \geq V^*_{HI} \tag{8}$$

Then, since $v(\underline{\tau}; y_H, A_I)$ is increasing in $\underline{\tau}$ and V^*_{HI} is constant in $\underline{\tau}$, the informed rich should retain power by choosing $\tau = \underline{\tau}$ provided that $\underline{\tau}$ is sufficiently large (i.e. $\underline{\tau} > \hat{\tau}$, $\hat{\tau}$ is the $\underline{\tau}$ that causes (8) to hold with equality). Else it should cede power by implementing $\tau^* \in (\hat{\tau}, \tau_{HI})$. □

Proof of Proposition 6. Recall, in the second period, the median effective agent is pivotal and will have their ideal policy implemented, in equilibrium. Thus, the second period policy will be $\tau_y = \frac{A_I}{y_{med}}$ if there is learning, and $\tau_x = \frac{A_I}{x_{med}}$ if there isn't. $\tau_x < \tau_y$ since $x_{med} > y_{med}$.

We begin by quantifying the benefit of learning for informed agents. We have:

$$\begin{aligned}\Delta(y) &= v(\tau_y; y, A_I) - v(\tau_x; y, A_I) \\ &= A_I \left[\left(1 - \frac{x_{med}}{y_{med}}\right) \frac{y}{x_{med}} + \ln\left(\frac{x_{med}}{y_{med}}\right) \right]\end{aligned}$$

Clearly, $\Delta(y)$ is continuous and strictly decreasing in y. Using the fact that $\ln(x) < x-1$ for any $x \neq 1$, we have $\Delta(y_{med}) > 0$ and $\Delta(x_{med}) < 0$. Then, by the intermediate value theorem, there exists a unique threshold $\tilde{y} \in (x_{med}, y_{med})$ s.t. $\Delta(\tilde{y}) = 0$ and $\Delta(y) > 0$ whenever $y < \tilde{y}$.

Suppose $\tau^\dagger > \tau_x$. Then preferences are single peaked for all sophisticated agents with $y \geq x_{med}$ (as well as for all naive agents). Hence all agents with effective income at least as large as x_{med} prefer τ_x to any $\tau' > \tau_x$. Furthermore, utility is strictly increasing in τ on the interval $(0, \tau_x)$ for all agents with effective income $y \leq x_{med}$ (a majority). Hence τ_x is majority preferred to any $\tau' \neq \tau_x$. Hence τ_x is a majority winner.

Suppose instead that $\tau^\dagger < \tau_x$. The proof here is more complicated. We proceed in 4 steps. First, we show that, if there is a majority winner, it must either be $\tau^\dagger$ or τ_x. Second, we provide conditions under which $\tau^\dagger$ is guaranteed to be the majority winner. Third, we provide conditions under which τ_x is guaranteed to be the majority winner. Finally, we explore cases where there is possibly no majority winner.

Step 1: Let $y(\tau^\dagger) = \frac{A_I}{\tau^\dagger}$ be the income for which $\tau^\dagger$ is stage optimal, and note that $y(\tau^\dagger) > x_{med}$. All agents with effective income $x < y(\tau^\dagger)$ prefer $\tau^\dagger$ to $\tau' < \tau^\dagger$, since neither choice induces learning and $\tau^\dagger$ is closer to their ideal. Since $y(\tau^\dagger) > x_{med}$ and $G(x_{med}) = \frac{1}{2}$, then $G(y(\tau^\dagger)) > \frac{1}{2}$, and so a $\tau^\dagger$ is majority preferred to every $\tau' < \tau^\dagger$.

Similarly, τ_x is preferred to $\tau' > \tau_x$ by all agents with effective income $x \geq x_{med}$ (a majority), and τ_x is preferred to $\tau' \in (\tau^\dagger, \tau_x)$ by all agents with effective income $x \leq x_{med}$ (a majority). Hence, if there is a majority winner, it must either be $\tau^\dagger$ or τ_x.

Step 2: Let us consider when $\tau^\dagger$ is majority preferred. Since $\Delta(y) < 0$ for sophisticated agents with $y > \tilde{y}$, these agents will be willing to distort policy from their stage ideal to prevent learning. Then $\tau^\dagger$ will be the globally optimal policy for a sophisticated agent if:

$$\phi(y; \tau^\dagger) = v(\tau^\dagger; y, A_I) - v\left(\frac{A_I}{y}, y, A_I\right) - \beta\Delta(y) \geq 0$$

It is easily verified that $\frac{\partial \phi}{\partial y} = \frac{A_I}{y} - \tau^\dagger - \beta\frac{\partial \Delta(y)}{\partial y}$. Now, $\Delta'(y) < 0$ and, by construction, $\frac{A_I}{y} - \tau^\dagger > 0$ whenever $y < y(\tau^\dagger)$. Hence $\frac{\partial \phi}{\partial y} > 0$ whenever $y < y(\tau^\dagger)$. Furthermore, $\phi(\tilde{y}) < 0$ (since $\Delta(\tilde{y}) = 0$), and $\phi(y(\tau^\dagger)) > 0$. Hence, by the intermediate value theorem, there exists $\hat{y}(\tau^\dagger) \in (\tilde{y}, y(\tau^\dagger))$ s.t. $\phi(\hat{y}(\tau^\dagger); \tau^\dagger) = 0$. Additionally, $\frac{\partial \phi}{\partial \tau^\dagger} = y(\tau^\dagger) - y > 0$ for $y < y(\tau^\dagger)$. Then, by the implicit function theorem, $\hat{y}(\tau^\dagger)$ is strictly decreasing in $\tau^\dagger$. [Intuitively, the smaller the distortion needed to prevent learning, the less rich an agent needs to make policy distortion optimal.]

Let $\underline{\tau}$ be defined implicitly by $\hat{y}(\underline{\tau}) = x_{med}$. Since $\hat{y}(\tau_x) < x_{med}$, and $\hat{y}$ is a strictly decreasing function, it must be that $\underline{\tau} < \tau_x$.

Now, $\tau^\dagger$ is preferred to any $\tau' > \tau^\dagger$ by each informed agent with $y > \hat{y}(\tau^\dagger)$ and by each misinformed agent with $y > \frac{A_M}{\tau^\dagger} = \frac{A_M}{A_I} y(\tau^\dagger)$. The measure of such agents is:

$$\rho(\tau^\dagger) = 1 - \gamma F_I(\hat{y}(\tau^\dagger)) - (1-\gamma) F_M\left(\frac{A_M}{A_I} y(\tau^\dagger)\right)$$

Since $y(\tau^\dagger)$ and $\hat{y}(\tau^\dagger)$ are both strictly decreasing in $\tau^\dagger$, $\rho(\tau^\dagger)$ is strictly increasing in $\tau^\dagger$. Now, if $\tau^\dagger = \tau_x$, then $y(\tau^\dagger) = x_{med}$ and $\hat{y}(\tau^\dagger) < x_{med}$. Then:

$$\begin{aligned}
\rho(\tau_x) &= 1 - \gamma F_I(\hat{y}(\tau_x)) - (1-\gamma) F_M\left(\frac{A_M}{A_I} x_{med}\right) \\
&> 1 - \gamma F_I(x_{med}) - (1-\gamma) F_M F_M\left(\frac{A_M}{A_I} x_{med}\right) \\
&= 1 - G(x_{med}) = \frac{1}{2}
\end{aligned}$$

By contrast, if $\tau^\dagger = \underline{\tau}$, then $y(\tau^\dagger) > x_{med}$ and $\hat{y}(\tau^\dagger) = x_{med}$. Then:

$$\begin{aligned}\rho(\tau_x) &= 1 - \gamma F_I(x_{med}) - (1-\gamma)F_M\left(\frac{A_M}{A_I}y(\underline{\tau})\right) \\ &< 1 - \gamma F_I(x_{med}) - (1-\gamma)F_M\left(\frac{A_M}{A_I}x_{med}\right) \\ &= 1 - G(x_{med}) = \frac{1}{2}\end{aligned}$$

Then, by the intermediate value theorem, there exists $\tilde{t}_1 \in (\underline{\tau}, \tau_x)$ s.t. $\rho(\tilde{t}_1) = \frac{1}{2}$. Moreover, whenever $\tau^\dagger > \tilde{t}_1$, $\rho(\tau^\dagger) > \frac{1}{2}$, and so $\tau^\dagger$ is majority preferred to any $\tau' > \tau^\dagger$. If so, $\tau^\dagger$ must be a majority winner.

Step 3: When is τ_x is majority preferred? Take any policy $\tau' \in (\tau^\dagger, \tau_x]$. Denote $y(\tau')$ as the income for which τ' is stage optimal. By construction, $x_{med} \leq y(\tau') < y(\tau^\dagger)$.

An informed agent will prefer τ' to $\tau^\dagger$ if:

$$\psi_I(y; \tau^\dagger, \tau') = v(\tau^\dagger; y, A_I) - v(\tau'; y, A_I) - \beta\Delta(y) \leq 0$$

Similarly, a misinformed agent will prefer τ' to $\tau^\dagger$ if:

$$\psi_M(y; \tau^\dagger, \tau') = v(\tau^\dagger; y, A_M) - v(\tau'; y, A_M) \leq 0$$

Notice that $\psi_I(y; \tau^\dagger, \tau') \geq \phi(y; \tau^\dagger)$, where the inequality is strict unless $y = y(\tau')$. (This follows because $v(\tau'; y, A_I) \leq v\left(\frac{A_I}{y}; y, A_I\right)$.) Now, $\frac{\partial \psi_I(y)}{\partial y} = (\tau' - \tau^\dagger) - \beta(\tau_x - \tau_y) > 0$. Furthermore, $\psi_I(\tilde{y}) < 0$ (since $\tilde{y} < y(\tau') < y(\tau^\dagger)$ and $\Delta(\tilde{y}) = 0$) and $\psi_I(\hat{y}(\tau^\dagger)) > 0$ (since $\psi_I(\hat{y}(\tau^\dagger)) > \phi(\hat{y}(\tau^\dagger)) = 0$). Hence, by the intermediate value theorem, there exists $y_I(\tau^\dagger, \tau') \in (\tilde{y}, \hat{y}(\tau^\dagger))$ s.t. $\psi_I(y_I(\tau^\dagger, \tau'); \tau^\dagger, \tau') = 0$, and $\psi(y; \tau^\dagger, \tau') < 0$ whenever $y < y_I(\tau^\dagger, \tau')$. Similarly, there exists $x_M(\tau^\dagger, \tau') \in (y(\tau'), y(\tau^\dagger))$ s.t. $\psi_M\left(\frac{A_M}{A_I}x_M(\tau^\dagger, tau'); \tau^\dagger, \tau'\right) = 0$ and $\psi_M(y; \tau^\dagger, \tau') < 0$ whenever $y < \frac{A_M}{A_I}x_M(\tau^\dagger, \tau')$.

The measure of agents who prefer τ' to $\tau^\dagger$ is:

$$r(\tau^\dagger, \tau') = \gamma F_I(y_I(\tau^\dagger, \tau')) + (1-\gamma)F_M\left(\frac{A_M}{A_I}x_M(\tau^\dagger, \tau')\right)$$

Now, since $y_I(\tau^\dagger, \tau')$ and $x_M(\tau^\dagger, \tau')$ are both strictly decreasing in $\tau^\dagger$ (for fixed τ'), $r(\tau^\dagger, \tau')$ is strictly decreasing in $\tau^\dagger$.

Now take $\tau' = \tau_x$. If $\tau^\dagger = \tilde{t}_1$, then:

$$\begin{aligned} r(\tilde{t}_1, \tau_x) &= \gamma F_I(y_I(\tilde{t}_1, \tau_x)) + (1-\gamma) F_M\left(\frac{A_M}{A_I} x_M(\tilde{t}_1, \tau_x)\right) \\ &< \gamma F_I(\hat{y}(\tilde{t}_1)) + (1-\gamma) F_M\left(\frac{A_M}{A_I} y(\tilde{t}_1)\right) \\ &= 1 - \rho(\tilde{t}_1) = \frac{1}{2} \end{aligned}$$

where we use the fact that $y_I(\tau^\dagger, \tau_x) < \hat{y}(\tau^\dagger)$ and $x_M(\tau^\dagger, \tau_x) < y(\tau^\dagger)$. By contrast, if $\tau^\dagger = \underline{\tau}$, then $y(\tau^\dagger) > x_{med}$ and $\hat{y}(\tau^\dagger) = x_{med}$. Then:

$$\begin{aligned} r(\tilde{\tau}, \tau_x) &= \gamma F_I(y_I(\underline{\tau}, \tau_x)) + (1-\gamma) F_M\left(\frac{A_M}{A_I} x_M(\underline{\tau}, \tau_x)\right) \\ &> \gamma F_I(x_{med}) + (1-\gamma) F_M\left(\frac{A_M}{A_I} x_{med}\right) \\ &= G(x_{med}) = \frac{1}{2} \end{aligned}$$

where we make use of the fact that $y_I(\underline{\tau}, \tau_x) = x_{med}$ (by construction) and that $y(\underline{\tau}) > x_{med}$. Then, by the intermediate value theorem, there exists $\tilde{\tau}_2 \in (\underline{\tau}, \tilde{t}_1)$ s.t. $r(\tilde{\tau}_2) = \frac{1}{2}$. Moreover, whenever $\tau^\dagger < \tilde{\tau}_2$, $r(\tau^\dagger, \tau_x) > \frac{1}{2}$, and so τ_x is majority preferred to any $\tau^\dagger$. Moreover, since this coalition of voters prefers $\tau^\dagger$ to any $\tau' < \tau^\dagger$, it must be that τ_x is majority preferred to any $\tau' < \tau^\dagger$. If so, τ_x must be a majority winner.

Step 4: What if $\tau^\dagger \in (\tilde{\tau}_2, \tilde{t}_1)$? Then, a majority prefer $\tau^\dagger$ to τ_x, so τ_x cannot be a majority winner. However, a majority of agents also have an ideal policy above $\tau^\dagger$ (since $\tilde{\tau}_1$ was constructed to have the property that exactly half of agents had ideal policy weakly below this when $\tau^\dagger = \tilde{t}_1$). However, this does not guarantee that there is a policy in $(\tau^\dagger, \tau_x)$ that is majority preferred to $\tau^\dagger$. If such a policy exists, then, a Condorcet cycle exists, and policy is unstable. If not, then $\tau^\dagger$ remains the equilibrium policy.

The measure of agents who prefer $\tau' > \tau^\dagger$ to $\tau^\dagger$ is $r(\tau^\dagger, \tau')$. Let $R(\tau^\dagger) = \sup_{\tau' \in (\tau', \tau_x]} r(\tau^\dagger, \tau')$.

When $\tau^\dagger = \tilde{t}_1$, we know that $r(\tilde{t}_1, \tau') < \frac{1}{2}$ for all relevant τ'. Hence $R(\tilde{t}_1) \leq \frac{1}{2}$. Since $r(\tau^\dagger, \tau')$ is strictly decreasing in $\tau^\dagger$ for each τ', then $R(\tau^\dagger)$ must be strictly decreasing in $\tau^\dagger$. Define $\tilde{\tau}_1 = \max\{\tau^\dagger \,|\, R(\tau^\dagger) \geq \frac{1}{2}\}$. Clearly $\tilde{\tau}_1 \leq \tilde{t}_1$. If $\tau^\dagger < \tilde{\tau}_1$, then a majority coalition exists that will replace $\tau^\dagger$ with some $\tau' > \tau^\dagger$. We know that $\tilde{\tau}_1 \geq \tilde{\tau}_2$, since a majority will replace $\tau^\dagger \leq \tilde{\tau}_2$ with τ_x. Hence, there will be policy instability when $\tau^\dagger \in (\tilde{\tau}_2, \tilde{\tau}_1)$. □

Proof of Proposition 7. The informed rich will implement the policy that maximizes their expected lifetime utility. First, suppose $A^\dagger(\tau_{HI}) \geq \overline{A}_I$. Then there will be no learning if the informed rich implement their stage optimal utility. Hence $\tau^* = \tau_{HI}$ is optimal. This requires that $\mu > (\overline{A}_I - A_M)\ln(\tau_{HI}\overline{y}) = \overline{\mu}$.

Next, suppose $A^\dagger(\tau_{HI}) \in (\underline{A}_I, \overline{A}_I)$. Then, the lifetime utility of the informed rich is decreasing at $\tau = \tau_{HI}$. The optimal policy must satisfy $\tau^* < \tau_{HI}$. Define $\underline{\tau}(\mu) = \frac{1}{y}\exp\{\frac{\mu}{\overline{A}-A_M}\}$ and $\overline{\tau}(\mu) = \frac{1}{y}\exp\{\frac{\mu}{\underline{A}-A_M}\}$. By construction $A^\dagger(\underline{\tau}(\mu)) = \overline{A}_I$ and $A^\dagger(\overline{\tau}(\mu)) = \underline{A}_I$. Then, since $\underline{A}_I < A^\dagger(\tau_{HI}) < \overline{A}_I$, it must be that $\underline{\tau}(\mu) < \tau_{HI} < \overline{\tau}(\mu)$. $\underline{\tau}(\mu)$ is the highest policy for which learning is guaranteed to not occur, and $\overline{\tau}(\mu)$ is the lowest policy for which learning is guaranteed to occur. The informed rich never have an incentive to propose $\tau < \underline{\tau}(\mu)$.

The optimal policy is the solution to:

$$\max_{\tau\in[\underline{\tau},\overline{\tau}]} v(\tau; y_H, A_I) + \beta\left(\int_{\underline{A}}^{A^\dagger(\tau)} v(\tau_{HI}; y_H, A_I)dF_I(A) + \int_{A^\dagger(\tau)}^{\overline{A}} v(\tau_{HI}; y_H, A_I)dF_I(A)\right)$$

Let $\hat{\tau}(\mu)$ be the solution to this problem. As noted, $\hat{\tau}(\mu) < \tau_{HI}$. If $\hat{\tau}\mu) \in (\underline{\tau}, \tau_{HI})$, then it will be characterized by the first order conditions. However, there may be a corner solution at $\underline{\tau}$.

Finally, suppose $A^\dagger(\tau_{HI}) \leq \underline{A}_I$, so that the ideal stage game policy guarantees that learning will occur. Let $\hat{\tau}(\mu) < \tau_{HI}$ denote the optimal solution to the problem:

$$\max_{\tau\in[\underline{\tau},\overline{\tau}]} v(\tau; y_H, A_I) + \beta\left(\int_{\underline{A}}^{A^\dagger(\tau)} v(\tau_{HI}; y_H, A_I)dF_I(A) + \int_{A^\dagger(\tau)}^{\overline{A}} v(\tau_{HI}; y_H, A_I)dF_I(A)\right)$$

The informed rich face a choice between implementing the distorted policy $\hat{\tau}(\mu)$ (which potentially avoids the transfer of power) or implementing the stage ideal policy τ_{HI} and conceding power to the informed poor. Let $\hat{V}(\mu)$ denote the expected lifetime utility of the informed rich under policy $\hat{\tau}(\mu)$. It is optimal to distort policy provided that:

$$\hat{V}(\mu) \geq v(\tau_{HI}; y_H, A_I) + E_{F_I}\left[v\left(\frac{A}{y_L}, y_H, A\right)\right]$$

Notice that the right hand side of this inequality is constant in μ. The left-hand side is increasing in μ since the necessary first period distortion becomes smaller as μ increases. Hence, there exists a threshold $\underline{\mu}$ s.t. the distortionary policy is preferred whenever $\mu > \underline{\mu}$. Moreover, since distortion is optimal when $A^{\dagger}(\tau_{HI}) = \underline{A}$, $\underline{\mu} < (\underline{A}_I - A_M)\ln(\tau_{HI}\overline{y})$.

Finally, we need to ensure that the ideal policy of the informed rich has the support of a majority coalition. It suffices to assume that the ideal stage game policy of the misinformed poor lies below the lowest policy that the informed rich would want to implement: i.e. $\tau_{LM} \leq \hat{\tau}(\mu)$. This condition is guaranteed to hold if $\tau_{LM} \leq \underline{\tau}(\mu)$. □