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 by equation (4) in section 4.2.2.

Proposition 4. *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:*

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

The behavior of the equilibrium is analogous to that in Proposition 3, and is accordingly summarized in the left panel of Figure 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.

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.

In the analysis so far, we have 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 the 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 eagerness 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: 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 just 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. Then $F(y) = \gamma F_I(y) + (1 - \gamma)F_M(y)$ is the unconditional distribution of incomes.

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

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

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

Proposition 5 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.[1] 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.[2] 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

1. 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.

2. 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.[3] 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 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.[4] Additionally, we

3. 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.

4. 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

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.[5] 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 $A^{\dagger}(\tau)$). 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:

$$v(\tau; y_H, A_I) + \beta \left(\int_{\underline{A}_I}^{\max\{\underline{A}_I, A^{\dagger}(\tau)\}} v(\tau_{LI}; y_H, A_I) dF_I(A) + \int_{\min\{A^{\dagger}(\tau), \overline{A}_I\}}^{\overline{A}_I} v(\tau_{HI}; y_H, A_I) dF_I(A) \right) \quad (1)$$

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:

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

---

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.

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

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 (5), 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 analogous to the assumption in Proposition 2.A that $\underline{\tau}_{HI} \geq \tau_{LM}$.)

Proposition 6. *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 6 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.[6] 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.

6. A different approach would be to introduce probabilistic voting (see Lindbeck and Weibull (1987) and Persson and 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.

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 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[7]. 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 method corresponds to the most preferred policy (within the alternative set $[\tau^{\dagger}, \tau_{HI}]$) of the most populous group in the polity.[8] 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

7. 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 $[\tau^{\dagger}, \overline{\tau}_{HI}]$.

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

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 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:

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

It is easily verified that $\underline{\tau}_{HI}$ is the solution to:

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

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:

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

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. $\square$

Proof of Proposition 4. [9] 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:

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

9. The proof of Proposition 3 builds on the proof of Proposition 4, so we present this proof first.

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.

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:

$$\frac{\tau^\dagger - \tilde{\tau}_{HI}}{\tau_{HI}} + \ln\left(\frac{\tilde{\tau}_{HI}}{\tau^\dagger}\right) = \beta\left[1 - \frac{\tau_{LI}}{\tau_{HI}} + \ln\left(\frac{\tau_{LI}}{\tau_{HI}}\right)\right] \tag{2}$$

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:

$$\frac{\tau^\dagger - \tilde{\tau}_{LM}}{\tau_{LM}} + \ln\left(\frac{\tilde{\tau}_{LM}}{\tau^\dagger}\right) = \beta\left[\frac{\tau_{HI}}{\tau_{LM}} - 1 + \ln\left(\frac{\tau_{LM}}{\tau_{HI}}\right)\right] \tag{3}$$

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 4, 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 4 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 (2) 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 (2) 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 5. 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 6. 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 $\underline{\tau}(\mu) = \frac{1}{\overline{y}} \exp\{\frac{\mu}{\overline{A} - A_M}\}$ and $\overline{\tau}(\mu) = \frac{1}{\overline{y}} \exp\{\frac{\mu}{\underline{A} - A_M}\}$. 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 $\tau_{LM} \leq \underline{\tau}(\mu)$. □

References

Lindbeck, Assar, and Jörgen W Weibull. 1987. "Balanced-budget redistribution as the outcome of political competition." *Public choice:* 273–297.

Persson, Torsten, and Guido Tabellini. 1999. “The size and scope of government:: Comparative politics with rational politicians.” *European economic review* 43 (4-6): 699–735.

Tideman, T Nicolaus. 1987. “Independence of clones as a criterion for voting rules.” *Social Choice and Welfare* 4:185–206.