2013

Misinformed Voters and the Politics of the Slippery Slope

Giri Parameswaran
Haverford College, gparames@haverford.edu

Misinformed Voters and the Politics of the Slippery Slope*

Giri Parameswaran†

October 29, 2013

Abstract

Reform opponents often argue that beneficial reforms should be rejected, just in case implementation leads the polity down the slippery slope (of implementing additional reforms) that ends at an outcome that is worse than the status quo. What rationalizes this fear of policy overshooting its target? In the context of public goods provision, I explain the slippery slope sentiment as the consequence of manipulation by some informed voters, of the beliefs of misinformed voters who systematically undervalue the public good. Inefficiently under-providing the public good reduces the opportunities for the misinformed to learn the true value of the good, which suppresses aggregate demand for the good. This incentive to distort is larger when the income of the pivotal voter is further from the median income, and exists even when the number of misinformed are small. Using an inequality measure that is analogous to, but distinct from, Lorenz dominance, I show that slippery slope inefficiencies are more likely to arise when inequality increases.

*I would like to thank Richard Ball, James Fogarty, Edoardo Grillo, Tim Lambie-Hansen, Daniel Lewis, Daniel Stackman and Kai Steverson for their helpful comments and suggestions.

†Department of Economics, Haverford College, 370 Lancaster Rd, Haverford PA 19041. Phone: (610) 896-2905 Email: gparames@haverford.edu

1 Introduction

> "Once we let go of the exclusivity of a one man-one woman relationship with procreation linking the generations, then why stop there? If it is 'about love and commitment' then it is entirely logical to extend marriage to, say, two sisters bringing up children together. If it is merely 'about love and commitment' then there is nothing illogical about multiple relationships, such as two women and one man." — Lord Carey, former Archbishop of Canterbury

So argued Lord Carey (2013) against same-sex marriage, invoking the fear that embracing marriage equality would be the first step down the *slippery slope* to celebrating incestuous or polyamorous relationships. In an earlier debate, Justice Antonin Scalia and other conservatives argued against extending many civil rights and protections to gays and lesbians that were enjoyed by heterosexuals[1], on the basis that doing so would be the first step towards legalizing same-sex marriage — even though there was little popular support for marriage equality at the time. Both cases are examples of arguments against a proposed policy — not because the policy is itself objectionable, but because of the objectionable additional reforms that would likely follow due to the policy's implementation.

Anti-reform arguments of this sort are commonplace in political discourse. And, although slippery-slope arguments are by their nature 'anti-reform', they are not unique to conservatives. For example, liberals in societies with universal healthcare often argue against limited privatization initiatives as the first step along the slippery slope towards dismantling the welfare state. In Canada, for example, there is much opposition to the evolution of a two-tiered healthcare system in which people can choose to pay for private alternatives.

[1]For example, in *Lawrence v Texas,* 539 U.S. 558 (2003), Justice Scalia wrote in a minority opinion that gay men ought to not enjoy the right to engage in consensual intercourse in the privacy of their homes. The opinion argues at 600: "That review is readily satisfied here by the same rational basis...that certain forms of sexual behavior are 'immoral and unacceptable,'... This is the same justification that supports many other laws regulating sexual behavior that make a distinction based upon the identity of the partner-for example, laws against adultery, fornication, and adult incest, and laws refusing to recognize homosexual marriage."

Indeed, until a 2005 Supreme Court decision ruled such laws illegal, six of Canada's ten provinces had bans on individuals using private insurance to access services that were generally available through the public system.

Whilst the contexts differ, the arguments share a common feature. Reform opponents hold that society should reject otherwise beneficial or efficiency-enhancing policies or reforms, just in case those reforms push society down a *slippery slope* that ends in policies that are much worse than the status quo. The argument is puzzling in at least the following sense: it assumes that reforms generate momentum. Once a primary reform that enacts a desirable policy is implemented, successive reforms will inevitably follow, causing policy to over-shoot its target. But surely Congress (or the appropriate decision-making body) must approve those successive reforms as well? Presumably, Congress may reject further reforms if they are indeed undesirable or inefficient.

This paper explains the slippery slope sentiment in the context of a polity where some agents are misinformed about the value of proposed policies or reforms, but may come to learn of their value through acquaintance. As I will demonstrate, under majority rule, a partially informed polity will choose differently from a perfectly informed one — even if a large majority of the polity are correctly informed. This creates an incentive for agents who dislike the expected policy outcomes when the polity is perfectly informed (i.e. the outcome at the end of the slippery slope) to not implement the initial policy or reform, and thereby prevent learning. In effect, the identity of the future decision maker (pivotal voter) is endogenous to the current policy choice. The current decision maker has an incentive to inefficiently maintain the status quo in order to retain control of the agenda.

That voters learn by acquaintance is plainly evident. 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.[2] In a different context, recent work by Baccini and Leemann (2012) shows that voters are more likely to be sensitive to climate issues when voting after being exposed to a natural disaster. Social science research suggests a similar effect regarding attitudes towards gays and lesbians. Herek and Glunt (1993) and Herek et al. (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), Day and Schoenrade (2000), and Griffith and Hebl (2002), amongst others, demonstrate that individuals are more likely to be open about their sexuality (and thereby enable known interpersonal contact between homosexuals and heterosexuals) in environments where anti-discrimination laws and policies are present.[3]

The idea that voters are often mistaken about the value of reforms or public goods is also plainly evident. In a survey of 1021 individuals, Koch and Mettler (2012) found that over 50% of respondents receiving government benefits were unaware that those benefits were indeed provided by the government.[4] This

[2] A Republican congressman from New York claimed: "The lash of the dictator will be felt, and 25 million free American citizens will for the first time submit themselves to a fingerprint test." Another opponent worried that it would "establish a bureaucracy in the field of insurance in competition with private business" that would destroy private pensions. Unsurprisingly, slippery slope concerns formed part of the objection to Social Security. During hearings before the Senate Finance Committee, 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)

[3] Indeed, as this article is being written, Congress is debating a federal Employment Non-Discrimination Act (ENDA) to protect homosexual and transgendered people from discrimination in the workplace. That there is considerable opposition to the bill is perfectly consistent with this paper's thesis that politicians will reject desirable reforms — does anyone genuinely favor arbitrary discrimination in the workplace? — due to slippery slope concerns. Anti-discrimination protections that allow homosexuals to become more visible may result in co-workers favorably amending their attitudes towards homosexuals, and by consequence, supporting social policies that are more inclusive of homosexuals, such as marriage equality.

[4] In their study, Koch and Mettler first asked respondents if they had ever used a government social program or not. Only 43% responded affirmatively. Respondents were then read a list of 21 government programs, and then asked if their response would change. After hearing the list, 96% of respondents admitted to having benefited from government programs. The study relied purely upon self reporting. Amongst respondents who originally claimed to have not benefited from government programs, 60% later admitted to having claimed the Home Mortgage Interest Deduction, 47% had claimed the Earned Income Tax Credit, 44%

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 cost of such programs. U.S. spending on foreign aid provides a stark example. In a 2010 World Public Opinion Poll of 848 Americans, the median respondent believed that the foreign aid budget accounted for 25% of the federal spending, whilst only 19% believed it was below 5%. The median respondent believed that foreign aid should ideally comprise 10% of federal spending.[5] In fact, the foreign aid budget in 2010 was less than 1% of total federal spending.[6] By over-attributing the share of public spending on 'non-beneficial' projects, voters effectively undervalue public spending as an aggregate bundle. This is especially true if voters underestimate the positive externalities associated with foreign aid.

To give context to the analysis, I consider a stylized model that focuses on the provision of a public good that has an objective marginal benefit, but which some agents undervalue. I refer to the latter agents as misinformed. Agents in the economy are distinguished by their income, and the public good is financed by a proportional tax on income. These assumptions imply that informed agents with higher incomes will demand lower levels of public goods provision — even though they value the public good identically to informed agents with lower incomes — because they are liable to finance a greater share of the public good. The assumptions also imply that at each income level, misinformed agents demand less of the public good than their informed counterparts. Combining these results, each misinformed voter can be associated

accessed Social Security, 43% benefited from Pell Grants, 40% were on Medicare, 28% were on Medicaid, and 25% received Food Stamps.

[5]The mean responses were even larger - the average respondent believed that foreign aid comprised 27% of federal spending, but should only be 13%.

[6]A 2000 poll by the same group found that 61% of respondents believed the foreign aid budget was too large. (In that survey, the median estimate of the foreign aid budget was 10% of total federal spending.) When asked to imagine that the federal government actually spent 1% of its budget on foreign aid - which was actually the case - only 13% still claimed that this amount was too large.

(in principle) with a wealthier correctly informed voter with the same preferences over public policy. Misinformed agents express preferences as if they were informed and richer than they actually are. Observing the system from without, the partially informed polity appears to behave in the same way that a much richer fully informed polity would. In particular, since misinformed agents behave as if they are richer than they actually are, the income of the pivotal voter will be larger than the true median income.

The possibility of learning introduces an additional dynamic. To make things stark, suppose all misinformed voters perfectly learn the value of the public good if a positive quantity is provided. Following any period in which there is positive provision of the public good, all the misinformed voters will learn the truth, and demand a larger level of the public good in the future. This changes the identity of the pivotal voter. By choosing his ideal (positive) level of the public good, the pivotal voter in the partially-informed economy (whose income will be above the true median income) will cede political power to the true median income earner, causing future public goods provision to increase. If this increase is indeed large enough, the old pivotal voter may prefer to prevent learning by not providing the public good at all. By entrenching the inefficient status quo, the pivotal voter can prevent learning by the misinformed, and accordingly, retain control of the agenda. The decision maker is willing to tolerate short run inefficiencies to prevent sustained long run inefficiencies that, from her perspective, are much worse.

Several results are worth mentioning. First, as the examples in sections 4 and 5.1 make clear, the result can hold even when a significant majority of agents are perfectly informed. (Indeed, if a majority of agents were misinformed, the results would be trivial.) The key insight of this paper is that in the presence of political competition, a small (but significant) amount of misinformation can result in a complete breakdown in the provision of public goods — a result which is quite stark. The very existence of misinformed voters creates an incentive for a subset of the informed majority to exploit the misinformed, by distorting policy in such a way that causes the misinformed to remain as such. Moreover, the result is not strongly predicated on any interaction between income and informedness. In particular, it is not crucial to the analysis that the poor are more likely to be misinformed than the rich. Certainly, and in contrast to Frank (2007), I do not suggest that the poor are systematically duped by the rich to vote against their interests.

As a benchmark, I consider the case where an agent's informedness status is independent of her income. Relative to this benchmark, the inefficient outcome becomes even more likely[7] if, holding constant the total number of informed and misinformed agents, the misinformed are more likely to be drawn from the poor. However, I also demonstrate that inefficiencies can arise if only a small number of below-median income earners are misinformed, but a large number of slightly above-median income earners are misinformed. (Such a situation might be plausible with public goods such as social insurance, which the poor are much more likely to be acquainted with than the middle class.) This latter case makes clear that the culpability for slippery-slope inefficiencies need not lie with the misinformed poor — misinformation at other portions of the income distribution can also cause inefficiencies to arise.

Second, the model exhibits an *ends-against-the-middle* flavor (see Epple and Romano (1996)). As will become clear, political competition pits the informed poor against a coalition of the informed rich and all misinformed agents (including the misinformed poor). The stability of this coalition depends upon the misinformed poor not becoming informed. Indeed — although the misinformed all vote the same way, regardless of income — the misinformed poor tend to work against their interests much more than the misinformed rich. Since the informed rich have a strategic interest in under-providing the public good, the misinformed rich may still be acting in their best interests in demanding less of the public good — even if only unwittingly. By contrast, the misinformed poor really would regret their choice, after becoming informed.

Third, in the context of financing public goods, the prevalence of the slippery slope phenomenon is related to the amount of inequality in the polity. In a

[7] There is no uncertainty in this model. Throughout this paper, I use terms such as 'likelihood' to refer to the size of the set of parameters for which the outcome is efficient or not.

relatively equal society, the cost to the pivotal voter of ceding power is relatively small, since the new pivotal voter will demand only a slightly larger level of public good in the future. Whilst this is not optimal from the perspective of the current pivotal voter, it is preferable to the inefficient outcome where none of the public good is provided. By contrast, if society is very unequal, then the future median voter may be much poorer, and so demand a much larger level of public spending than the current pivotal voter is willing to tolerate. As such, inefficient under-provision of the public good does not arise from the learning mechanism alone — but rather through the interaction with inequality. In this paper, I define a measure of inequality that is analogous to, but distinct from, Lorenz dominance. I show that under this measure, the slippery-slope motivated incentive to under-provide public goods (to prevent learning) is increasing in the amount of inequality in the economy.

A notable feature of the model is that it endogenously explains both a status quo bias towards inefficient policies, as well as the incentive for gradualism in policy making. The status quo bias arises directly from the incentive to prevent learning. By contrast, the incentive for gradualism stems from the understanding that even small changes will — if they cause learning — result in political power shifting in a favorable way, which will make further reforms possible. Gradualism is not merely born from a pragmatic notion of 'taking what one can get', given political constraints. As the analysis in section 6 demonstrates, these choices are made with the expectation of a 'domino'-like effect as the identity of the pivotal voter changes.

This paper contributes to, and extends, several strands of the existing political economy literature. At its core, the inefficiency in this model arises from the 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 (especially fiscal policy), including Persson and Svensson (1989),Roberts (1989), Alesina and Tabellini (1990), Tabellini and Alesina (1990), Dewatripont and Roland (1992), Benabou (2000), Battaglini and Coate (2007), Battaglini and Coate (2008), amongst many others. However, in contrast to many of these models, and similar to Acemoglu and Robinson (2001), Benabou and Ok (2001), the shifting political power is not exogenous, but endogenous to the current agent's policy choice. Interestingly, and in strong contrast to this paper, the redistribution technology in Benabou and Ok (2001) causes the pivotal agent to be relatively richer in the future when redistribution is provided - causing a similar disincentive to favor redistributive policies, when there is policy stickiness. Indeed, policy momentum is another feature of this model that is present in Benabou and Ok (2001). (In that paper, it is the fear by the current poor that the redistributive policy that will make rich in the short run, will persist long enough to eventually expropriate their future wealth.) However, policy inertia is hard-wired into their model. This 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 results in an incentive for gradualism in policy making, similar to Dewatripont and Roland (1992).

Several papers investigate the relationship between inequality and the demand for public goods or redistribution. Standard models (e.g. Romer (1975), Meltzer and Richard (1981) and Moene and Wallerstein (2001), amongst others) show that redistribution and higher public goods provision are more likely when inequality is high and the income profile is right skewed. As inequality worsens, the relatively poorer median voter has a greater incentive to expropriate the rich. Benabou (2000) presents a model in which redistribution generates aggregate social gains. In this model, the standard effect (increasing inequality begetting a greater impetus for redistribution) is present. However, a separate effect exists, whereby increasing inequality decreases the incentives for redistribution, since the social gains will not be shared as equitably. This results in a 'U-shaped' effect where increasing inequality first reduces the likelihood of redistribution, and then increases it. In my model, an increase in inequality uniformly decreases the likelihood of public goods provision, which puts it in stark contrast to the existing literature.

Finally, this paper extends upon a growing literature concerning 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, although using a different mechanism, that paper finds an endogenous bias towards status quo policies. A more recent set of papers consider the incentives for agents to choose policies that affect the learning of others. Strulovici (2010) studies learning in bandit problems when decisions (about how to experiment) are made collectively by majority vote. Baker and Mezzetti (2012), Fox and Vanberg (2011) and Parameswaran (2013), consider models of the judiciary in which learning occurs after courts observe the outcomes of agent choices. For example, in Parameswaran (2013), and Fox and Vanberg (2011), agents have an incentive to skew their choices (or to make choices that appear sub-optimal when dynamic considerations are ignored) to prevent learning by the court, and consequently affect the way that legal rules evolve.

The remainder of this paper is structured as follows: In section 2 I introduce the formal model. Section 3 characterizes the optimal stage game policies in the absence of any dynamic considerations, and formalizes the notion of effective income. Section 4 characterizes the Markovian equilibrium in a game with the simple learning technology described above. Section 5 analyzes the comparative static effects of varying inequality and informedness, on equilibrium public goods provision. Section 6 provides several extensions, including a demonstration that the results in section 4 are robust to general learning technologies. All proofs are contained in the Appendix.

2 Baseline Model

There are a continuum of voters with mass 1. Each voter $i \in [0,1]$ has income y_i drawn from a distribution $F(y)$ with support $Y \subset \Re^+$ (possibly finite). The mean and median incomes are $\overline{y}(F)$ (assumed finite) and $y_m(F)$, respectively.

There is a public good g which gives utility Ag^{α} with $A > 0$ and $0 < \alpha < 1$.[8] The public good costs p. The government can finance spending on public goods by levying a uniform linear tax on income τ.[9] To focus on the implications of mistakes and learning, I abstract from the distortionary effects of taxation by considering a pure endowment economy.

Voters can either be informed (I) or misinformed (M). An informed voter knows the true value of A, whilst misinformed voters believe $A = 0$. Let $\gamma(y, F)$ denote the conditional probability that an agent with income y is informed. Since there are a continuum of voters, this is also represents the proportion of informed voters at each income level. Denote by $\overline{\gamma}(F) = \int_Y \gamma(y)\, dF(y)$, the number of informed voters in the polity. In everything that follows, I assume that $\overline{\gamma} > \frac{1}{2}$.[10]

A voter's type $\theta = (y, t)$ indicates her income level and state of informedness. Let $\Theta = Y \times \{I, M\}$ be the set of possible voter types. The distribution of types is given by:

$$\begin{aligned}
\Pr[Y \le y,\, t = I] &= \int_0^y \gamma(x)\, dF(x) \\
\Pr[Y \le y,\, t = M] &= F(y) - \int_0^y \gamma(x)\, dF(x)
\end{aligned}$$

There is a technology through which misinformed agents become informed. Let $Q(\cdot,\cdot)$ denote this technology, where $\gamma' = Q(g, \gamma)$ denotes the future profile of informedness as a function of the current profile and the current level of

[8]The functional form choice is simply for tractability. The results generalize to any increasing, concave function that is bounded from below.

[9]Again, proportional taxation is purely for simplicity. More generally, let $T(y, g)$ be a feasible tax schedule, representing the amount of taxes paid by an agent with income y, if the government provides g units of the public good. Feasibility implies $\int_Y T(y, g)\, dF(y) = pg$. The results generalize to the case where $\frac{\partial^2 T}{\partial g \partial y} > 0$ — i.e. the marginal rate of taxation is increasing in the level of public goods, at every level of income. (Since the marginal cost of public goods provision is $\frac{\partial T}{\partial g}$, this condition implies that the marginal cost is rising with income.) The condition admits certain classes of regressive income tax schedules.

[10]The case of $\overline{\gamma} < \frac{1}{2}$ is trivial. The misinformed are a majority and have identical preferences that are maximized when $g = 0$. Hence, the public good is never provided.

public goods provision.[11] In the baseline model, I consider a special case of this technology in which all agents learn the true value of the public good whenever a positive quantity is provided. I.e.

$$Q(\gamma, g) = \begin{cases} \gamma & g = 0 \\ \mathbf{1} & g > 0 \end{cases}$$

where $\mathbf{1} \in \Gamma$ is the informedness profile in which $\gamma(y) = 1$ for all y. In section 6, I consider a more general class of learning technologies and show that the baseline results are robust to these alternative technologies.

The income distribution is common knowledge. I assume that the informedness profile and learning technology are known by the informed agents — i.e. the informed are aware that some agents are misinformed. By contrast, I assume that the misinformed are ignorant of their misinformedness and of the possibility that learning might take place. Since the dynamics of the model arise only through learning, this implies that the misinformed will support the policy that maximizes their stage utility.

There are two political parties who are purely office motivated.[12] In every period, each party announces a feasible fiscal policy (τ, g) that it is committed to implement if it is elected. All agents vote, and the party receiving a majority of the vote is elected. All agents discount the future at the common rate $\delta \in [0, 1)$.

3 The Stage Game

This section is in two parts. In the first part, I characterize the optimal stage-policy for each type of agent. As will become apparent, the game is dynamic

[11] To make stark the mechanism at play in this model, I assume that learning is only possible through acquaintance with the public good. Of course, other channels exist in the real world — although they would presumably not provide an incentive to skew the provision of public goods, which is the focus of this paper's inquiry.

[12] The results are robust to endowing parties with partial policy motivation, as long as some office motivation remains.

only insofar as current policy choices may change the identity of future decision makers. Hence, the optimal policy in any sub-game where the preferences of future decision makers is no longer subject to change will correspond to the optimal stage-game policy. In particular, this will be the case when all agents are informed ($\gamma = \mathbf{1}$) so that no further learning is possible.

3.1 Ideal Policies

Consider type $\theta = (y, t)$ agent. Her ideal stage policy is given by:

$$\max_{g \geq 0} \left(1 - \frac{pg}{\overline{y}}\right) y_i + A_t g^{\alpha}$$

where $pg = \tau \overline{y}$ represents the government's budget constraint. It is easily verified that agents' preferences are single-peaked in g. The optimal amount of public spending for a type-(y, t) agent is:

$$g(y, t) = \left(\frac{A_t \alpha}{p} \frac{\overline{y}}{y}\right)^{\frac{1}{1-\alpha}}$$

Obviously, for every y, $g(y, M) = 0$, and for $y' > y$, $0 < g(y', I) < g(y, I)$. Hence, misinformed agents will never provide the public good (and, hence, will always choose a zero tax rate). Informed agents will choose positive taxation and spending, although lower income earners prefer more of each than higher income earners.

The indirect-utility of a type-(y', t') agent, when a type-(y, t) agent's optimal policy is implemented, is:

$$u_{(y',t')}(g(y, t)) \quad = \quad y' + \left(A_{t'} - \alpha A_t \frac{y'}{y}\right) \left(\frac{\alpha A_t}{p} \frac{\overline{y}}{y}\right)^{\frac{\alpha}{1-\alpha}}$$

In particular, the pivotal agent herself receives utility:

$$u_{(y,t)}(g(y, t)) = y + (1 - \alpha) A_t \left(\frac{\alpha A_t}{p} \frac{\overline{y}}{y}\right)^{\frac{\alpha}{1-\alpha}}$$

Although they may differ about the ideal level of provision, all informed voters agree that this ideal level is strictly positive. Along the Pareto frontier, a positive quantity of the public good will always be provided. In this sense, providing none of the public good very clearly entails an inefficient under-provision.

3.2 Effective Incomes and the Effective Median

In the previous subsection, I showed that stage preferences are single peaked and that the ideal policies of informed agents are monotonic in their incomes. Then, if all agents were informed, the median income earner would be pivotal. However, this monotonicity no longer holds when some agents are misinformed, as all misinformed agents, regardless of income, demand none of the public good. More generally, with misinformed voters, the type-space of voters is multidimensional — voters are distinguished by both their income and their informedness status. Indeed, with multidimensionality of this sort, the notion of a 'median' voter is, in general, not well defined.

However, as I show below, the effect of misinformedness is for agents to express preferences that are identical to the preferences that would be asserted by a richer, correctly informed agent. To see this, consider an agent with income y and who values the public good at $A_t \geq 0$. (To show the generality of this approach, I allow the agent to be misinformed in any way, so I do not yet require that $A_M = 0$.) For any level of public goods provision, g, the agent's utility is:

$$u_{(y,t)}(g) \quad = \quad \left(1 - \frac{pg}{\overline{y}}\right) y + A_t g^{\alpha} = \frac{A_t}{A} \left[\left(1 - \frac{pg}{\overline{y}}\right) \frac{A}{A_t} y + A g^{\alpha}\right]$$

which represents the same preference as a correctly informed agent with income $x(y,t) = \frac{A}{A_t} y$. I refer to $x(y,t)$ as the agent's 'effective income'. It is the income level at which a correctly informed agent would evaluate policies in the same way as the agent in question. In the simple model, where agents are either correctly informed, or believe the public good is worthless, effective income is given by:

$$x(y,t) = \begin{cases} y & t = I \\ \infty & t = M \end{cases}$$

Naturally, the effective income of a correctly informed agent is simply their actual income, whilst the effective income of a misinformed agent is infinite.[13]

The notion of effective income reduces the type-space from two-dimensions (encoding income y and the valuation of public goods A_t) into a single-dimension. Effective income is a summary statistic for agent preferences. Since, by assumption, misinformed agents systematically undervalue public goods, the effective income profile of the polity appears richer than the true income profile.

Following the analysis in the previous subsection, agents' ideal stage-policies are monotonic in their effective income.[14] Hence, when voters are ordered according to their effective income, standard median voter results (Black (1948), Downs (1957) etc.) apply.

Let $x_m(\gamma, F)$ denote the median effective income, given income profile F and informedness profile γ. The effective median income is[15]:

$$x_m(\gamma, F) = \begin{cases} \inf_{z \in Y} \int_0^z \gamma(y)\, dF(y) \geq \frac{1}{2} & \overline{\gamma} > \frac{1}{2} \\ \infty & \overline{\gamma} < \frac{1}{2} \end{cases}$$

Since misinformed agents upwardly skew the effective income distribution, the effective median income will be (weakly) larger than the true median income. Indeed, for any $\gamma \in \Gamma$, $x_m(\gamma, F) \geq y_m(F)$. Misinformedness causes the effective median voter to be richer than the true median and — since the ideal level of public goods provision is monotonically decreasing in effective income — causes the polity to choose a lower level of public goods provision than

[13]Infinite effective income is an immediate consequence of the assumption that $A_M = 0$. In section 6, I discuss consider approaches to relaxing this assumption.

[14]Indeed, $g(y,t) = \left(\frac{A_t \alpha}{p} \frac{\overline{y}}{y}\right)^{\frac{1}{1-\alpha}} = \left(\frac{A\alpha}{p} \frac{\overline{y}}{x(y,t)}\right)^{\frac{1}{1-\alpha}} = g(x(y,t), I)$ - which is monotonically decreasing in x.

[15]If F is continuous, then the top expression simplifies to $\int_0^{x_m(\gamma,F)} \gamma(y)\, dF(y) = \frac{1}{2}$.

would be implied by the income profile alone. *En masse*, the electorate appears to be more 'fiscally conservative' than one would infer from the income distribution alone.[16]

4 Model with Simple Learning Technology

In this section, I characterize the dynamic equilibrium with the simple learning technology. I solve for a Markov perfect equilibrium. Let $V_\theta(\gamma)$ be the value function for a type θ voter, given an informedness profile γ. If $\gamma = \mathbf{1}$, then the true median is pivotal in the current period and all future periods. (This follows since the current policy does not affect the continuation game, and so agents' lifetime preferences over policy are simply given by their stage preferences — and these are single-peaked.) Then, the optimal policy is simply the one that maximizes the median agent's stage preferences, and so:[17]

$$\begin{aligned} g^*(\mathbf{1}) &= \arg\max_g \left\{ \left(1 - \frac{pg}{\overline{y}}\right) y_m + A g^\alpha + \delta V_{(y_m,1)}(\mathbf{1}) \right\} \\ &= g(y_m, I) = \left(\frac{A\alpha}{p}\frac{\overline{y}}{y_m}\right)^{\frac{1}{1-\alpha}} \end{aligned}$$

Since the state never changes, for any type (y,t), the value function is given by:

$$V_{(y,t)}(\mathbf{1}) = \frac{1}{1-\delta}\left[y + \left(A_t - \alpha A \frac{y}{y_m}\right)\left(\frac{\alpha A}{p}\frac{\overline{y}}{y_m}\right)^{\frac{\alpha}{1-\alpha}} \right]$$

[16]Note again that this result would remain true even if I relaxed the assumption that the misinformed believe that the public good is worthless (i.e. $A_M = 0$). Indeed, along as $A_M < A$, the results that: (i) the effective median was richer than the true median (making the electorate appear richer than it truly is) and (ii) the polity consequently chooses a lower level of public goods provision, continue to hold.

[17]Without confusion, I use $g^*(\gamma)$ to denote the level of public goods chosen in equilibrium when the informedness profile is γ, and $g(y,t)$ to denote the level of public goods that maximizes a (y,t)-type agent's stage preferences.

Now, take any $\gamma \neq \mathbf{1}$. An informed agent's utility from having g units of the public good provided is:

$$v(g;\gamma) = \begin{cases} \left(1 - \frac{pg}{\bar{y}}\right) y + A_t g^{\alpha} + \delta V_{(y,I)}(\mathbf{1}) & \text{if } g > 0 \\ y + \delta V_{(y,I)}(\gamma) & \text{if } g = 0 \end{cases} \tag{1}$$

It follows that if $g^*(\gamma) = 0$, then $V_{(y,I)}(\gamma) = \frac{1}{1-\delta} y$.

Unlike in the stage game, in the dynamic setting where learning is possible, preferences need no longer be single-peaked over the entire policy space. As (1) makes clear, continuation payoffs are constant for any $g > 0$, and so over this region, lifetime preferences are single-peaked, since stage preferences are single-peaked. However, continuation payoffs are discontinuous at $g = 0$. Moreover, the continuation payoff from choosing $g = 0$ may either be higher or lower than the continuation payoff from $g > 0$, depending on the agent's income. For an agent with income $y > \frac{1}{\alpha} y_m$, it can be verified that: $\frac{1}{1-\delta} y > V_{(y,I)}(\mathbf{1})$. For such agents, the median income earner's ideal policy is so far from their own ideal, that receiving none of the public good forever is preferable to receiving the median income earner's ideal level forever. For these agents, although lifetime utility falls as g decreases from their ideal level to 0, it jumps up discontinuously at $g = 0$. By contrast, for agents with $y < \frac{1}{\alpha} y_m$, lifetime utility is falling as g decreases to 0 and drops down discontinuously at $g = 0$, which preserves single-peakedness. Although preferences are no longer single-peaked for all agents in the dynamic setting, the following Lemma demonstrates that the effective median voter is still pivotal:

Lemma 1. *Let $x_m(\gamma, F)$ be the effective median voter given informedness profile γ. The following are true:*

1. Suppose $u_{(x_m(\gamma,F),I)}(g = 0) > u_{(x_m(\gamma,F),I)}(g(x_m(\gamma, F), I))$. Then for all $x > x_m(\gamma, F)$, $u_{(x,I)}(g = 0) > u_{(x,I)}(g(x_m(\gamma, F), I))$;

2. Suppose $u_{(x_m(\gamma,F),I)}(g = 0) < u_{(x_m(\gamma,F),I)}(g(x_m(\gamma, F), I))$. Then for all $x < x_m(\gamma, F)$, $u_{(x,I)}(g = 0) < u_{(x,I)}(g(x_m(\gamma, F), I))$.

Lemma 1 shows that, if the effective median income earner prefers to not provide the public good, then so will all agents whose effective income is larger (i.e. agents whose income is actually higher, or misinformed agents). By contrast, if the effective median prefers to provide her ideal stage-game level of the public good, then so will all informed agents with lower income. Hence, the governing coalition will either consist of the 'informed poor' or the 'rich' and misinformed (where 'rich' and 'poor' indicate incomes relative to the effective median).

The effective median faces a trade-off between choosing her ideal policy today and giving up future political power on the one hand, and retaining political power by not providing the public good at all. Whilst retaining political power is costly in that it requires an under provision of the public good, it may outweigh the detriment of ceding political power and facing a much larger over-provision of the public good in the future. This will be true if the difference in income levels (and hence, difference in ideal levels of public goods provision) between the true median and effective median is large enough.

Proposition 1. *There exists a unique* $\kappa(\alpha, \delta) > \frac{1}{\alpha}$ *s.t. the equilibrium level of public goods provision is:*

$$g^*(\gamma) = \begin{cases} g^*(x_m(\gamma), I) & \frac{x_m(\gamma)}{y_m} < \kappa \\ 0 & \frac{x_m(\gamma)}{y_m} > \kappa \end{cases}$$

Moreover, $\kappa(\alpha, \delta)$ *satisfies:* $(1-\delta)(1-\alpha)+\delta(1-\alpha\kappa)(\kappa)^{\frac{\alpha}{1-\alpha}} = 0$, *and* $\frac{\partial \kappa}{\partial \alpha} < 0$ *and* $\frac{\partial \kappa}{\partial \delta} < 0$.

Proposition 1 demonstrates that the equilibrium level of public goods provision can be in one of two regimes. If the effective median income is not too much larger than the true median income, then the public good will be provided in positive quantities in every period. Since the true median will demand a level of public goods provision only slightly higher than the effective median's ideal — the effective median voter finds it preferable to implement her ideal policy today and tolerate a slightly larger provision of public goods in every future period, rather than retain political power by never providing the public good. By contrast, if the true and effective median incomes differ by a large amount, then public goods provision breaks down altogether. The effective median would rather not provide the public good at all, than receive her desired amount today, and be subject to a sustained over-provision in the future. The current decision maker chooses to implement an inefficient policy today for fear that implementing the stage-game efficient policy will lead her down the slippery slope to policies that are even worse (from her perspective).

Given the above discussion, Proposition 1 effectively divides the parameter-space into two regions — an 'Efficient Provision' region and a 'Slippery-Slope Inefficiency' region. In the 'Efficient Provision' region, positive quantities of the public good are provided. These policies are efficient in the sense that they correspond to the ideal stage policies of some voter, and hence lie on the Pareto Frontier. By contrast, in the 'Slippery Slope' region, the public good is not provided at all, which is clearly inefficient, since every informed agent would ideally choose a positive level of public goods provision. It is worth noting that efficient provision need not be efficient in the sense of Samuelson (1954). The efficient (Samuelson) level of the public good is $g^* = \left(\frac{\alpha A}{p}\right)^{\frac{1}{1-\alpha}}$ — the level at which the aggregate marginal benefit $\alpha A g^{\alpha-1}$ and the marginal cost p coincide. In standard models, where the median income earner is pivotal and the income distribution is right-skewed, the equilibrium level of public goods provision, $g\left(y_m, I\right) = \left(\frac{\alpha A}{p}\frac{\overline{y}}{y_m}\right)^{\frac{1}{1-\alpha}}$, will exceed this level. Since the median voter is liable for less than the true marginal cost of providing the public good, the public good will be over-provided in the sense of Samuelson. In this model, since the effective income of the pivotal voter is larger than the median income, the equilibrium (short-run) level of public spending may be closer to the Samuelson level. Indeed, these will coincide if the effective median income coincides with the average income in the economy. Of course, if the effective median income is larger still, there will be an under-provision of the public good relative to the Samuelson level. Moreover, this effect only exists in the short run. In the long run, as learning takes place, and the pivotal voter converges to the median income earner, public spending will converge to the higher-than-efficient level. This suggests a limited basis for a small amount of misinformation to be desirable, in order to better align the costs faced by the pivotal voter to aggregate social costs.

I end this section by considering a simple example in which informedness is constant and independent of income. This is a 'judgment free' baseline of sorts — it does not appeal to any informedness differential between agents at different levels of the income spectrum. In particular, I do not assume that the poor are more or less likely to be informed than the rich. (In section 5.2, I analyze how the equilibrium level of public goods provision may vary with different informedness profiles.) I assume that income is log-normally distributed, since this has been shown to be a reasonable approximation of the income profiles in many countries. I calibrate the income distribution to the U.S. economy, using the Gini Coefficient as a measure of inequality.

Example 1. Suppose $\ln Y \sim N(\ln y_m, \sigma^2)$ and $\gamma(y) = \gamma > \frac{1}{2}$ is constant. The effective median income is given by $\int_0^{x_m} \gamma dF(y) = \frac{1}{2}$, where $F(y) = \Phi\left(\frac{1}{\sigma}\ln\left(\frac{y}{y_m}\right)\right)$ is the distribution function. This implies that $x_m = F^{-1}\left(\frac{1}{2\gamma}\right) = y_m e^{\sigma\Phi^{-1}\left(\frac{1}{2\gamma}\right)} \geq y_m$. By Proposition 1, the public good will not be provided if $\frac{x_m}{y_m} > \kappa$. This implies $e^{\sigma\Phi^{-1}\left(\frac{1}{2\gamma}\right)} > \kappa$ which will be true if $\gamma < \frac{1}{2\Phi\left[\frac{1}{\sigma}\ln\kappa\right]} = \tilde{\gamma}$. Hence, there is a threshold level of informedness below which the public good will not be provided.

To get a sense of how much misinformation is required, suppose $\sigma = 0.886$ and $\delta = 0.95$. The former is calibrated to the Gini coefficient measure of inequality in the U.S. in 2010[18] using the property that the Gini coefficient for log-normal income profiles is $2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1$. In the table below, I report the threshold value κ, for several different parametric values of α (which parametrizes the value of public spending), given the assumed value of δ. As predicted above, the threshold value κ is decreasing in α. The table below also reports the associated threshold informedness levels.

[18] According to the U.S. Census Bureau, the Gini Coefficient for the United States was 0.469 in 2010. United States Census Bureau (2012)

α	0.25	0.50	0.75
κ	4.0987	2.0260	1.3406
$\tilde{\gamma}$	0.5295	0.6351	0.7941

Example 1 demonstrates that misinformation about the value of public spending that is spread broadly throughout the population, can be sufficient to generate a complete breakdown in its provision. The result does not hinge on the relationship between income and inequality. Importantly, it is not crucial to the analysis that the poor are less informed than the rich. Moreover, a relatively small amount of misinformation may suffice to stifle provision of the public good. In the above example, when $\alpha = 0.75$, public goods provision ceases entirely if the effective median income is more than 34% above the true median income. This implies that a misinformedness rate of 21% is sufficient to induce slippery slope inefficiencies, even if the remaining 79% of voters are correctly informed and would ideally demand positive provision of the public good. In example 3, in the following section, I compare this result to the outcome in cases when information is not income independent.

5 Comparative Statics

In this section, I examine the comparative statics of changes in the income and informedness profiles on the equilibrium level of public goods provision.

5.1 Inequality and the Slippery Slope

The main result from the previous section is that inefficiencies arise when the effective and true median incomes are sufficiently disparate. Of course, as was shown in section 3, this difference in incomes is endogenous to the model, and is as a function of both the true income distribution and the informedness profile. In the introduction to this paper, I noted that many previous papers predict relationships between inequality and public goods provision — that these are positively related according to standard models, and 'U-shaped' in Benabou (2000). In this section, I demonstrate that for an appropriately defined measure of inequality, an increase in inequality will be associated with a larger separation between effective and true median incomes — and, as such, with a greater likelihood of a slippery-slope motivated under-provision of the public good.

Since our question concerns the effect of inequality on public goods provision, the comparative static analysis ideally considers the effect of a change in the income distribution, keeping the informedness profile fixed. However, in section 2, the informedness profile was defined conditionally upon the income distribution. Hence, as the income distribution changes, we must take care to appropriately modify the informedness profile as well. (To see why, suppose that γ was kept unchanged, which is equivalent to assuming that it is independent of the income distribution. Then a change in the income profile will generically change the number of agents associated with each income level, and this will affect the aggregate informedness level.) To do so, I assume that the informedness profile is *income-rank independent*. Let $Y \subset \Re_+$ be a convex set, and let $\mathcal{F}_Y$ denote the set of continuous distribution functions on Y.

Definition 1. An informedness profile $\gamma(\cdot|\cdot)$ is income-rank independent if for any $F_1, F_2 \in \mathcal{F}_Y$, and for any quantile $p \in (0,1)$, $\gamma\left(F_1^{-1}(p)\,|F_1\right) = \gamma\left(F_2^{-1}(p)\,|F_2\right)$.

Income-rank independence essentially says that the likelihood that an agent is informed depends only upon his rank in the income distribution (as measured by the quantile into which his income falls), and not the income level itself. This property ensures that the informedness of agents remains unchanged, even as their incomes are perturbed.

Proposition 2. *Let $F_1, F_2 \in \mathcal{F}_Y$ and let the informedness profile $\gamma(\cdot|\cdot)$ be income-rank independent. Then $\frac{x_m(\gamma,F_1)}{y_m(F_1)} \leq \frac{x_m(\gamma,F_2)}{y_m(F_2)}$ if and only if $F_1(x_m(\gamma, F_1)) \geq F_2\left(x_m(\gamma, F_1)\frac{y_m(F_2)}{y_m(F_1)}\right)$.*

Proposition 2 provides a sufficient condition on the distribution functions for the deviation between effective and true median incomes to diverge. To get the intuition, consider the case where $y_m(F_1) = y_m(F_2)$, so that the median incomes coincide under both income profiles. With income-rank independence, the median effective income must occupy the same income quantile regardless of the actual income distribution. It is sufficient to check whether the effective median income under the first income profile $x_m(\gamma, F_1)$ is associated with a higher or lower quantile under the second income profile. In the latter case, the new effective median income $x_m(\gamma, F_2)$ must be larger, since income level $x_m(\gamma, F_1)$ occupies a quantile that is too low under income profile F_2. When the medians differ, it suffices to simply scale income under profile F_2 up or down in a proportional manner such that the medians do coincide. Combined with Proposition 1, Proposition 2 addresses the relationship between different income profiles and the likelihood of slippery-slope inefficiencies.

Of course, Proposition 2 does not directly address the relationship between slippery slope incentives and inequality. To do so, I introduce a notion of inequality that is analogous to, but distinct from, Lorenz dominance. I say that Y_1 is more equal than Y_2 if Y_2 is a *median-normalized spread* of Y_1.

Definition 2. Let Y_1 and Y_2 be random variables, and for each $i \in \{1, 2\}$, let $F_i(y)$ and $y_m(F_i)$ be the associated distribution function and median, respectively. Y_2 is a *median-normalized spread* of Y_1, if $F_2(zy_m(F_2)) \geq F_1(zy_m(F_1))$ whenever $z < 1$ and $F_2(zy_m(F_2)) \leq F_1(zy_m(F_1))$ whenever $z > 1$.

The notion of a median-normalized spread is intimately related to that of a median-preserving spread[19]. Indeed, Y_2 is a median-normalized spread of Y_1 if and only if $\frac{Y_2}{y_m(F_2)}$ is a median-preserving spread of $\frac{Y_1}{y_m(F_1)}$. Two properties of median-normalized spreads recommend its use as an income inequality ranking. First, if Y_1 and Y_2 share a common median, then median-normalized spreads and median-preserving spreads are equivalent concepts. There is a

[19] Formally F_2 is a *median preserving spread* of F_1, if F_1 and F_2 share a common median, y_m (i.e. $F_1(y_m) = \frac{1}{2} = F_2(y_m)$), and if $F_2(y) \geq F_1(y)$ whenever $y < y_m$ and $F_2(y) \leq F_1(y)$ whenever $y > y_m$. See Mendelson (1987) and Malamud and Trojani (2009).

clear sense in which the notion of a median-preserving spread provides a (partial) inequality ranking of income distributions. Although in both cases the median agent receives the same income, more agents under F_2 receive incomes that are distant from the median than under F_1. Income under F_2 is less concentrated around the median income.

Second, normalizing by the median income ensures that the inequality ranking is invariant to arbitrary scaling. A failure of scale invariance would, for example, cause the measure of inequality to potentially vary depending on which unit of currency is used to measure income. Naturally, scale invariance is a property shared by many other measures of inequality, including Lorenz functions[20], and their associated measures, such as the Gini coefficient. Importantly for the analysis, scale invariance allows for a comparison of income distributions across economies with different median incomes.

Finally, there is a noteworthy analogy between median-normalized spreads and Lorenz dominance as measures of inequality. As Atkinson (1970), Thistle (1989) and the proof of Lemma 3 indicate, Y_1 is more equal that Y_2 in the sense of Lorenz dominance if and only if $\frac{1}{E[Y_2]}Y_2$ is a mean-preserving spread of $\frac{1}{E[Y_1]}Y_1$. Lorenz dominance requires that, after normalizing the income distributions to ensure that the means are equal, one distribution is a mean-preserving spread of the other. By contrast, the measure of inequality used in this paper requires that, after normalizing the income distributions to ensure that medians are equal, one distribution is a median-preserving spread of the other. The similarity between the approaches is obvious.

Lemma 2. *Let $F_1, F_2 \in \mathcal{F}_Y$ and let $\gamma(\cdot|\cdot)$ be any income-rank independent informedness profile. Suppose F_2 is more unequal than F_1 in the sense of median-normalized spreads. Then $\frac{x_m(\gamma,F_1)}{y_m} \leq \frac{x_m(\gamma,F_2)}{y_m}$.*

Lemma 2 demonstrates the relationship between inequality and the dispersion between the effective and true median incomes. Suppose income profile Y_2 is

[20]Given an income distribution F, the Lorenz function $L(p) = \frac{1}{\bar{y}}\int_0^{F^{-1}(p)} ydF(y)$ measures the fraction of total income held by the lowest p percent of income earners. Income distribution F_1 Lorenz dominates F_2 if $L_1(p) \geq L_2(p)$ for every $p \in [0,1]$.

more unequal than Y_1 in the sense of median-normalized spreads. Then the distance between the true and effective median incomes will be larger under F_2 than F_1. Combined with Proposition 1, this implies that the likelihood that the equilibrium will be in the slippery slope regime, and that public goods provision will break down altogether, is larger in the more unequal society. This is a stark result. In contrast to standard results, an increase in inequality may be associated with a significant decrease in public goods provision. As inequality worsens, the incomes of the pivotal voter and the median income earner diverge, and consequently, the current pivotal voter will be more wary of ceding political power.

Note that the sense in which public goods provision is less likely is that the economy is more likely to be in the slippery slope regime. However, if the equilibrium remains in the efficient provision regime after inequality worsens, then the level of public goods provision may actually go up. Although the decision to provide public goods or not depends upon the ratio of the effective and true median incomes, conditional upon positive provision, the quantity provided depends on the ratio between the effective median and *average* incomes. By construction, the medians of median-normalized spreads coincide (after normalization). Nevertheless, if average income is much larger under profile F_2, then the pivotal voter may indeed demand a larger quantity of the public good. Intuitively, this will be the case if $\frac{\bar{y}(F_2)}{\bar{y}(F_1)} \geq \frac{x_m(\gamma,F_2)}{x_m(\gamma,F_1)}$.

Of course, median-normalized spreads is just one approach to measuring income inequality. I have already argued that there is a strong analogy between median-normalized spreads and Lorenz dominance (which is related to mean-preserving spreads). Under certain conditions, it can be shown that median-normalized spreads imply Lorenz dominance.

Lemma 3. *Let $F_1, F_2 \in \mathcal{F}_Y$ and suppose F_2 is a median-normalized spread of F_1. Furthermore, suppose $\frac{\bar{y}(F_2)}{y_m(F_2)} \geq \frac{\bar{y}(F_1)}{y_m(F_1)}$ and suppose there is a unique z at which the functions $F_1(\bar{y}(F_1)z)$ and $F_2(\bar{y}(F_2)z)$ cross. Then F_1 Lorenz dominates F_2.*

Lemma 3 shows that, if two additional conditions are met, an increase in inequality in the sense of median-normalized spreads implies an increase in inequality in the sense of Lorenz dominance. The first condition requires that higher inequality causes the median and mean incomes to diverge, which implies that increasing inequality skews income towards the rich. The second condition is a single crossing property on distribution functions of the mean-normalized income profiles. In general, median-normalized spreads are, by themselves, not sufficient to guarantee that these two conditions hold, and hence are not generically sufficient for Lorenz dominance.[21] However, the following corollary provides a sufficient condition for a median-normalizing spread to be sufficient for Lorenz dominance.

Corollary 1. *Let $F_1, F_2 \in \mathcal{F}_Y$ and suppose F_2 is a median-normalized spread of F_1 satisfying $\frac{\bar{y}(F_2)}{y_m(F_2)} = \frac{\bar{y}(F_1)}{y_m(F_1)}$. Then F_1 Lorenz dominates F_2.*

Although the additional sufficient conditions make the results in Lemma 3 and Corollary 1 quite specialized, they are satisfied whenever a median-normalized spread exists within several classes of commonly used distributions. For example, suppose Y_1 and Y_2 are log-normal income profiles, and that Y_2 is a median-normalized spread of Y_1. (This will be true if $\sigma_2 \geq \sigma_1$, regardless of the median incomes in each profile.) Then both additional conditions in Lemma 3 are immediately satisfied, and so Y_1 Lorenz dominates Y_2. The same result holds for income profiles that are both drawn from the Pareto, Weibull and Uniform families of distributions - all of which are commonly used in modeling the distribution of income in society.[22] In fact, for income profiles

[21]To see that median-normalized spreads do not imply Lorenz dominance, consider the following discrete income distributions: $Y_1 \in \left\{\frac{1}{2}, 1, \frac{3}{2}\right\}$ each with probability $\frac{1}{3}$, and $Y_2 \in \left\{\frac{2}{5}, 1, \frac{8}{5}\right\}$ with probabilities $\frac{1}{3}, \frac{1}{5}$ and $\frac{7}{15}$, respectively. Clearly Y_2 is a median-normalized spread of Y_1 — indeed, it is a median-preserving spread. Furthermore $\bar{y}(F_2) = \frac{91}{75} > 1 = \bar{y}(F_1)$ and so the first sufficient condition is satisfied. However, it is easily verified that $F_1(\bar{y}(F_1) \cdot z)$ and $F_2(\bar{y}(F_2) \cdot z)$ do not satisfy single-crossing. Indeed, $F_2(\bar{y}(F_2) \cdot z) \geq F_1(\bar{y}(F_1) \cdot z)$ for $z \leq 1$ and for $z \geq \frac{120}{91}$. Consistent with Lemma 3, we can show that Lorenz functions are given by $L_1 = \begin{cases} \frac{1}{2}p & p < \frac{1}{3} \\ p - \frac{1}{6} & \frac{1}{3} \leq p < \frac{2}{3} \\ \frac{3}{2}p - \frac{1}{2} & p \geq \frac{2}{3} \end{cases}$ and $L_2 = \begin{cases} \frac{370}{999}p & p < \frac{1}{3} \\ \frac{25}{27}p - \frac{185}{999} & \frac{1}{3} \leq p < \frac{8}{15} \\ \frac{40}{27}p - \frac{13}{27} & p \geq \frac{8}{15} \end{cases}$ and these functions do not respect Lorenz dominance, since they intersect at $p = \frac{17}{26}$.

[22]I demonstrate that this is the case for the log-normal distribution. The other cases

drawn from these four classes of distributions, a median-normalized spread is necessary and sufficient for Lorenz dominance — the two notions of inequality are equivalent.

Of course, these results do not generally hold true for arbitrary pairs of distribution functions. As noted above, and in the example in footnote 21, median-normalized spreads do not generically imply Lorenz dominance. Conversely, as Example 2, below, demonstrates, Lorenz dominance does not imply that a median-normalized spread exists. In fact, as the example shows, it is possible for public goods provision to be more likely (and slippery slope concerns to be less likely) in an economy that is more unequal in the sense of Lorenz dominance.

Example 2. Let $Y = [0, \alpha\overline{y}]$ where $\alpha \in (1,2]$ and consider income distributions: $F_1 = \left(\frac{y}{\beta\overline{y}}\right)^{\frac{1}{\beta-1}}$ for $y \in [0, \beta\overline{y}]$ with $\beta \leq \alpha$, and F_2 piece-wise defined with $F_2(y) = 1 - \frac{1}{\alpha}$ for $y \in [0, \alpha\overline{y})$ and $F_2(\alpha\overline{y}) = 1$. (Under F_2, the lowest $1 - \frac{1}{\alpha} \leq \frac{1}{2}$ fraction of the population earn nothing, whilst the remainder all earn $\alpha\bar{y}$. The case of $\alpha = \beta = 2$ is especially straight-forward — income is uniformly distributed under F_1, whereas, under F_2, half the population receives nothing and the other half receives $2\overline{y}$.) Under both profiles the mean income is $\bar{y}$. The Lorenz functions are: $L_1(p) = p^{\beta}$ and $L_2(p) = 0$ for $p < \frac{\alpha-1}{\alpha}$ and $L_2(p) = \alpha p + 1 - \alpha$ for $p \geq \frac{\alpha-1}{\alpha}$. It is easy to verify that F_1 Lorenz dominates F_2. However, it is easily verified that F_2 is not a median-normalized spread of F_1. Consistent with Lemma 2, the distance between true and effective median incomes is smaller under F_2 even though F_2 is more unequal in the sense of

are demonstrated analogously. Let Y_1 and Y_2 be log-normally distributed — i.e. $\ln Y_i \sim N\left(\ln m_i, \sigma_i^2\right)$. Then $F_2(m_2 z) = \Phi\left(\frac{\ln z}{\sigma_2}\right)$ and $F_1(m_1 z) = \Phi\left(\frac{\ln z}{\sigma_1}\right)$. It is easily verified that F_2 can only be a median-normalized spread of F_1 if $\sigma_2 \geq \sigma_1$. For each i, $\frac{\bar{y}(F_i)}{y_m(F_i)} = e^{\frac{1}{2}\sigma^2}$. Hence, if F_2 is a median-normalized spread of F_1 (i.e. if $\sigma_2 > \sigma_1$), then $\frac{\bar{y}(F_2)}{y_m(F_2)} > \frac{\bar{y}(F_1)}{y_m(F_1)}$ and so the first condition is automatically satisfied. Furthermore, for each i, $F_i(\bar{y}(F_i) \cdot z) = \Phi\left(\frac{\ln z + \frac{1}{2}\sigma^2}{\sigma}\right)$. It is easily demonstrated that if F_2 is a median-normalized spread of F_1, then $F_2(\bar{y}(F_2) \cdot z) > F_1(\bar{y}(F_1) \cdot z)$ only if $z < exp\left\{\frac{1}{2}\sigma_1\sigma_2\right\}$, which verifies that single-crossing is satisfied. Finally, the Lorenz function associated with a log-normal distribution is $L(p) = \Phi\left(\Phi^{-1}(p) - \sigma\right)$. Clearly $L_1(p) \geq L_2(p)$ for any (and every) p iff $\sigma_1 \leq \sigma_2$. Hence, Y_1 Lorenz dominates Y_2 if and only if Y_2 is a median-normalized spread of Y_1.

Lorenz dominance. Indeed, this result is true even if the informedness profiles were different in the two economies. To see why, note that under F_2, the median income is also the maximum income, and so the effective median and true median share the same income $\alpha\bar{y}$. Hence, for any informedness profiles γ_1 and γ_2, $\frac{x_m(\gamma_1,F_1)}{y_m(F_1)} \geq 1 = \frac{x_m(\gamma_2,F_2)}{y_m(F_2)}$. The less equal society (with income profile F_2) will always choose the maximum amount of public goods feasible in any political equilibrium, whereas the public good may not be provided at all in the more equal society.

5.2 Nature of Informedness Profile

The previous subsection examined the effect of a changing income profile on the likelihood of public goods provision. In this subsection, I briefly consider the effect of different informedness profiles on the provision of public goods.

Remark 1. Let $\gamma_1, \gamma_2 \in \Gamma$ and suppose $\gamma_2 \geq \gamma_1$ (i.e. $\gamma_2(y) \geq \gamma_2(y)$ for every $y \in Y$). Then $\frac{x_m(\gamma_2,F)}{y_m(F)} \leq \frac{x_m(\gamma_1,F)}{y_m(F)}$.

Remark 1 makes the fairly obvious point that a change in the income profile that (weakly) increases the likelihood that every income earner is informed, will cause the the effective median agent to have lower income. As the polity becomes more informed across the board, the proportion of below median income earners who are informed increases, and so one does not need to travel as far up the income distribution to find the effective median.

A more interesting question considers the effect of a change in the informedness profile that reduces the likelihood of being informed at some income levels, and increases it at others. Recall, in section 4, I introduced Example 1 in which the informedness profile was constant, and hence independent of income. I referred to that case as the 'judgment-free' baseline as the poor were no less likely to be misinformed than the rich. In Example 3, below, I consider two variants of the baseline example, in which the informedness profile is no longer constant. In both cases, I keep the total number of informed agents in the economy unchanged. In the first case, I reduce the likelihood that below median income earners are informed and increase the likelihood that above median income earners are informed. In the second case, I increase the likelihood that below median income earners and very high income earners are informed, and reduce the likelihood that "upper-middle class voters" (those with slightly above median incomes) are informed.

Example 3. Suppose $\ln y \sim N(\ln y_m, \sigma^2)$ and $\gamma_1(y) = \begin{cases} 1 & y > y_m \\ 2\gamma - 1 & y \leq y_m \end{cases}$ so that $\overline{\gamma}_1 = \gamma$. Under informedness profile γ_1, only below median-income voters are misinformed. By contrast, let γ_2 be as defined below, which again satisfies $\overline{\gamma}_2 = \gamma$. Under profile γ_2, only 2% of below-median income earners are misinformed, however the next block of middle-income earners are completely misinformed, whilst the highest income earners are perfectly informed. (One might interpret this as a case where the public good involves social insurance, which the current poor are much more likely to be acquainted with than the middle class.)

$$\gamma_2(y) = \begin{cases} 0.98 & y < y_m \\ 0 & y \in [y_m, F^{-1}(1.49 - \gamma)] \\ 1 & y > F^{-1}(1.49 - \gamma) \end{cases}$$

It turns out that the effective median income is the same in both cases, and has income defined by $x_m(\gamma) = y_m e^{\sigma\Phi^{-1}\left(\frac{3}{2}-\gamma\right)}$. There will be slippery-slope inefficiencies if: $\kappa < e^{\sigma\Phi^{-1}\left(\frac{3}{2}-\gamma\right)}$ which will be true if $\gamma < \tilde{\gamma} = \frac{3}{2} - \Phi\left(\frac{1}{\sigma}\ln\kappa\right)$. Again calibrating to the U.S. economy, and using the sample values for α from Example 1, gives the following threshold values:

α	0.25	0.50	0.75
κ	4.0987	2.0260	1.3406
$\tilde{\gamma}$	0.5557	0.7127	0.8704

A comparison of the outcomes in Examples 1 and 3 highlights the important features of the role of the informedness profile in determining the effective median (and hence in the likelihood of public goods provision). Relative to the baseline, both cases in Example 3 involve a transfer of informedness from either poor or moderately-high income earners, to the super rich. Since the aggregate level of informedness was unchanged, this necessarily caused the median effective income to increase. (If more agents with incomes above the original median effective income are informed, and total informedness is unchanged, then fewer agents with incomes below the original median effective income can be informed. But this implies that the new effective median income must be larger.) Hence, any aggregate-information neutral change in the informedness profile that uniformly increases the informedness level of very high income earners is likely to cause the median-effective income to increase. *Ceteris paribus*, public goods provision is more likely to break down in economies where the poor are far less informed than the rich. A poorly informed lower-class exacerbates slippery-slope concerns, relative to the judgment-free baseline. This is illustrated in the first case above (with $\alpha = 0.75$), where the relative lack of information by the poor causes the public good to not be provided even if 87% of the population (and 74% of below median income earners) are correctly informed. By contrast, in the 'judgment-free' case, aggregate informedness needed to be below 80% for slippery slope concerns to take effect.

However, as Example 3 also makes clear, significant misinformation amongst the poor is not crucial to generating slippery slope concerns — culpability for the breakdown in public goods provision, need not always lie with the poor. Indeed, concentrating misinformation in the middle of the income distribution can equally exacerbate problems. The informedness profile interacts with the income profile in an obvious way to determine the median effective income. But for two informedness profiles that generate the same median effective income, the specific details of which agents were likely to be informed or not is inconsequential to the outcome. As Example 3 demonstrates, it is unimportant whether misinformation occurs amongst the poorest voters, or middle-class voters, as long as between these groups, sufficiently many agents are misinformed. Of course, if every below-median income earner were correctly informed, then the effective median income would be the true median income, and slippery slope concerns could never arise. Hence, misinformation amongst the poor is important to the analysis. However, it is not crucial that the poor are more likely to be misinformed than other members of the community.

6 Extensions

In this section, I consider two variants of the model that extend the results to more general settings. In the first case, I consider a general learning technology, and show that the results from the main section continue to hold. This extension demonstrates that focusing on the very special learning technology in the previous sections was without much loss of generality. In the second case, I discuss the implications of the assumption that misinformed agents do not value the public good at all, and suggest methods of relaxing this assumption that preserve the model's results.

6.1 General Learning Technologies and Gradualism

The previous section characterized equilibrium provision of public goods under a simple and stark learning technology. In this section, I show that the main results continue to hold when more general learning technologies are considered. In so doing, I show that the above results do not arise out of the special features of the learning technology — but rather from the very fact of misinformation and learning.

Consider a simple income distribution, with two income types $y_H > y_L$ and $F(y_L) = \phi > \frac{1}{2}$ poor agents.[23] The informedness profile is given by the vector

[23] The two-income-type assumption ensures that there is only one type of agent whose dynamic choices need to be modeled, since the choices of the uninformed and the informed poor are stationary. Extending to many income types requires the current dynamic decision maker to take into account the effect of his current choice on the future choices of other dynamically sophisticated decision makers with different preferences. This introduces all the usual time-inconsistency complications into the analysis. Note, however, that the limitation

$\gamma = (\gamma_L, \gamma_H) \in [0,1]^2$. Let $\Gamma = \left\{(y_L, y_H) \in \Delta | \, \phi\gamma_L + (1-\phi)\, \gamma_H \geq \frac{1}{2}\right\}$ be the set of informedness profiles where the effective median voter is informed. The effective median has income:

$$x_m(\gamma) = \begin{cases} y_L & \phi\gamma_L > \frac{1}{2} \\ y_H & \phi\gamma_L < \frac{1}{2} \end{cases}$$

Let $\Gamma^L = \left\{\gamma \in \Gamma | \gamma_L > \frac{1}{2\phi}\right\}$ be the informedness profiles under which the poor are pivotal, and let $\Gamma^H = \left\{\gamma \in \Gamma | \gamma_L \leq \frac{1}{2\phi}\right\}$ be the profiles under which the rich are pivotal. (For technical convenience, I assume that if $\gamma_L = \frac{1}{2\phi}$, so that exactly half of voters are informed poor, then the informed rich remain pivotal.) Note that, conditional upon the effective median being informed, the income type of the effective median depends only upon γ_L.

Consider a generalized learning technology $Q(g, \gamma)$ that is continuous, (strictly)[24] increasing in each argument, and that satisfies $Q(0, \gamma) = \gamma$. These assumptions imply natural relationships between public goods provision and learning. The assumption that $Q(0, \gamma) = \gamma$ maintains the working assumption that learning about the public good occurs only through acquaintance. Whilst this is obviously an over-simplification, it allows the analysis to abstract from other influencing factors, and focus on the effect of acquaintance-based learning. Monotonicity implies that more learning occurs when more of the public good is provided. Furthermore, a more informed polity will remain more informed after receiving the same amount of the public good as a less informed polity.

Let $\Theta = \{y_H, y_L\} \times \{I, M\}$ be the set of types, and let $(v_\theta)_{\theta \in \Theta}$ be a quadruple of functions, where $v_\theta : [0,1]^2 \to \Re$ denotes the continuation value of a type θ

to two income types does not limit the sense in which the analysis in this section is a generalization of the analysis in previous sections. Indeed, in those sections — although more income types were allowed — only two types of decision makers could exert political power with positive probability; those with the same income as the current pivotal voter, and those with the same income as the median income earner. Hence, the simplification to two income types preserves (rather than restricts) the nature of the transition dynamics.

[24] The monotonicity property is strict whenever $Q(g, \gamma) < \mathbf{1}$.

agent.

Proposition 3. *There exists a unique Markov Perfect Equilibrium, and is jointly characterized by:*

1. a set of bounded value functions $(v_\theta^*)_{\theta\in\Theta}$which satisfy:

$$\begin{aligned} v^*_{(y_i,I)}(\gamma) &= \left(1-\frac{pg^*(\gamma)}{\overline{y}}\right)y_i + A\left[g^*(\gamma)\right]^\alpha + \delta v^*_{(y_i,I)}\left(Q\left(g^*(\gamma),\gamma\right)\right) \\ v^*_{(y_i,M)}(\gamma) &= \frac{1}{1-\delta}\left(1-\frac{pg^*(\gamma)}{\overline{y}}\right)y_i \end{aligned}$$

2. a policy function:

$$g^*(\gamma) = \begin{cases} \left(\frac{\alpha A}{p}\frac{\overline{y}}{y_L}\right)^{\frac{1}{1-\alpha}} & \gamma\in\Gamma^L \\ \hat{g}(\gamma) & \gamma\in\Gamma^H \end{cases}$$

where

$$\hat{g}(\gamma) = \arg\max_{g\geq 0}\left(1-\frac{pg}{\overline{y}}\right)y_H + Ag^\alpha + \delta v^*_{(y_H,I)}\left(Q\left(g,\gamma\right)\right)$$

Proposition 3 demonstrates that there is a unique Markovian equilibrium of the generalized learning game, characterized by the above Bellman equations. As in the previous section, I assume that the misinformed are unaware of the dynamics arising out of learning. Accordingly they assume the game is stationary — that continuation play will resemble current policy choices. Since the poor can never lose political power (Γ^L is an absorbing state), they will choose their ideal policy whenever they are in power. When the rich are in power, they face a truly dynamic decision problem, and, hence, it is the value function of the rich that drives the results.

Let $\kappa(\alpha,\delta)$ be the threshold from Proposition 1 in the main section. An economy is a 5-tuple $e = (\alpha,\delta,y_H,y_L,\phi)$ which summarizes the relevant preference and income distribution parameters. The value function depends upon the parameters that constitute the economy, although this dependence is typically suppressed in the notation. Let $\mathcal{E}$ be the set of economies, $\mathcal{E}_{Eff} = \left\{ e \in \mathcal{E} | \frac{y_H}{y_L} < \kappa(\alpha, \delta) \right\}$ be the set of economies in which there is positive public goods provision under the simple learning technology from the previous section, and let $\mathcal{E}_{SS} = \left\{ e \in \mathcal{E} | \frac{y_H}{y_L} > \kappa(\alpha, \delta) \right\}$ be the set of economies in which there are slippery slope inefficiencies. Consider an economy e with initial informedness profile γ. Let $G_t(\gamma, e)$ be the equilibrium public goods provision at each time t.

As the following proposition shows, slippery slope concerns arise with the generalized learning technology whenever and only when they arise with the simple technology:

Proposition 4. *Public goods provision respects the following dynamics:*

1. If $e \in \mathcal{E}_{Eff}$, then $G_t(\gamma, e) > 0$ for all t and $\exists T(\gamma, e) \geq 0$ s.t. $G_t(\gamma, e) = g(y_L, I)$ whenever $t > T(\gamma, e)$.

2. If $e \in \mathcal{E}_{SS}$, then $G_t(\gamma, e) \longrightarrow 0$ as $t \to \infty$.

Proposition 4 is the analogue of Proposition 1 in the previous section. It shows that long run policy in an economy with a general learning technology is identical to the long run policy with the stark learning technology considered in the main section. This suggests a strong robustness of the results in sections 4 and 5.1. If the economy lies in the efficient region, then there will be strictly positive provision of the public good in the short run — although not necessarily at the ideal level of the rich, since the rich may still have an incentive to under-provide the public (relative even to their own ideal), in order to slow the process of learning and delay the time at which they completely cede power to the poor. Since this time will eventually arrive, in the long run, the median income earner's ideal policy will be eventually implemented. By contrast, if the economy lies in the slippery slope region, then public goods provision will very quickly disappear. Again, some public goods provision may occur in the short run; the rich may have some ‘wiggle room’ to provide a small amount of the public good without ceding power.

With additional assumptions, it can be shown the rich will never choose to provide their ideal level of the public good when in the slippery slope regime. Even if the rich could choose their ideal level for some periods without ceding political power, Proposition 4 ensures that they must eventually reduce the quantity of public goods provided by a significant amount. Intuitively, it cannot be inter-temporally optimal to expect such a dramatic decrease in public goods provision. The rich could do better by decreasing the original level of public goods provision, thereby slowing down the rate of learning, in order to sustain higher average public goods provision for a longer period of time. This intuition is formalized in the following Lemma: For notational convenience , let $\eta(g) = \left(1 - \frac{pg}{\bar{y}}\right) y_H + Ag^{\alpha}$ denote the stage utility of the rich when g units of the public good are provided, and let $Q^{-1}(y, \gamma)$ be the amount of the public good that is needed to shift the informedness profile from γ to y .

Lemma 4. *Suppose Q is differentiable and $\eta(Q^{-1}(y, \gamma))$ is concave. Then $G_t(\gamma, e) < g(y_H, I)$ whenever $e \in \mathcal{E}_{SS}$.*

The generalized model can explain both a status quo bias in equilibrium policy making that entrenches inefficiencies, as well as gradualism in policy making when efficiency enhancing reforms are embraced. The model predicts that the ‘reform-motivated-party’ (in this case, the party of the poor) will propose a sequence of policies that eventually result in their base’s ideal policy being implemented — and that they will embrace this gradual approach even if implementing their ideal policy is feasible in the short term. The virtue of the gradual approach is not merely the pragmatism of implementing the best politically feasible outcome. Rather, it sets in motion a ‘domino-like’ sequence of events that systematically improves the long-run welfare of the party’s base. This effect would be even more pronounced in a more general model with a generalized learning technology and multiple income types, since then, in every period, political power is shifting to agents whose preferences increasingly aligned with those of the poor.

6.2 Partial Undervaluation of Public Good

The assumption that the misinformed do not value the public good at all is an admittedly strong assumption. It proved useful in keeping the analysis tractable by ensuring that all misinformed agents demanded less of the public goods than all informed agents. Furthermore, it had the feature that the misinformed would be the natural ally of a rich decision maker who sought to prevent learning. For generic $A_M < A$, these need not be true. For example, if $A_M > 0$, then the misinformed would still demand a positive quantity of the public good, opening up the possibility of a coalition between the informed poor and the misinformed. (The informed poor have an incentive to form such a coalition, anticipating that political power will shift in their favor as learning occurs.)

Nevertheless, the assumption $A_M = 0$ is not crucial generically. Its necessity stemmed from the interaction of two other strong assumptions of the model — the strong monotonicity assumption that learning occurs whenever a positive level of public goods are provided, and the assumption that a positive level of public goods provision will be stage-game optimal whenever $A_t > 0$. It should be clear that either of these assumptions could be plausibly relaxed in a more general model. For example, one may plausibly assert that acquaintance-based learning is improbable (or negligible) if the quantity of public good provided is so small as to be essentially invisible to the public. Since voters are typically not monitoring government policy very thoroughly, the policy would need to be large enough in scale to attract the public's attention and enable them to appropriately interact with it.

Similarly, we may plausibly assert that agents who place only a small positive value on the benefit of the public good will demand zero public goods provision. This would be the case, for example, if the public good could only be provided in discrete increments. For example, if public goods provision was limited to integer quantities, then any agent with effective income $x(y_i, A_i) > \frac{A\overline{y}}{p}$ would ideally not have the public good provided at all. A similar result arises if the public good is divisible, but there is a fixed cost C was associated with its provision, for example the cost of financing the bureaucracy that oversees the provision the public good. It is easily shown that, under these conditions, an agent with effective income $x(y_i, A_i)$ will optimally demand that the public good not be provided if:

$$x(y_i, A_i) > (1-\alpha)\left(\frac{\alpha}{p}\right)^{\alpha}\frac{A}{C}\overline{y} = \overline{x}$$

Either approach allows one to relax the rather strong assumption that the misinformed completely undervalue the public, without changing the nature of the strategic interactions involved.

7 Conclusion

Slippery slope concerns are often used in political discourse to argue against beneficial or efficiency enhancing policies or reforms. Opponents argue that whilst the policy, taken in isolation, might be beneficial, its implementation will likely cause a sequence of further policies to be adopted, that results in a final outcome that is worse than the status quo. What rationalizes this fear of reform momentum that causes policy to overshoot its target? Why doesn't the polity simply reject the subsequent reforms, if they really are suboptimal?

This paper rationalizes slippery slope concerns as a consequence of decision making in a democracy, where some voters are originally misinformed about the value of a policy or reform, but may come to learn its value through acquaintance. For concreteness, I focused on the provision of a public good that has an objective marginal benefit, but whose value some voters under-estimate. Since they are liable to finance a greater share of the public good, richer voters demand less of the good than the poor. Moreover, at each income level, misinformed voters demand less of the good than a correctly informed voter would. Lower than optimal demand by the misinformed results in political power being held by agents whose incomes are larger than the true median income. Through acquaintance with the good, the misinformed come to learn of its true value. Hence, if the optimal level of the public good is provided, through time, political power will shift to agents with incomes closer to the true median, who prefer a larger provision of the public good. This creates an incentive for the current (relatively rich) decision maker to not provide the public good, to prevent learning and thereby retain political control of the agenda.

This paper's main result is that slippery slope inefficiencies are most likely to arise in polities where the income (or more generally preferences) of the current and future decision makers are sufficiently disparate. Importantly, I show that the incentive to distort policy to prevent learning may exist even if the number of misinformed agents is relatively small and large majorities of agents are informed. The relationship between income disparity the likelihood of slippery slope inefficiencies arising suggests a connection to the amount of income inequality in the polity. I introduce a measure of inequality that is distinct from, but analogous to, Lorenz dominance, and show that increasing inequality increases the likelihood that slippery slope inefficiencies will arise. This result is in contrast to standard models in which rising inequality causes the demand for public goods to rise. Hence, the analysis suggest a novel mechanism by which inequality can generate Pareto inefficient outcomes.

I also studied the relationship between the informedness of agents at different levels of the income profile and slippery slope inefficiencies. Taking the case of uniform informedness as a 'judgment-free' baseline, I demonstrated that slippery slope inefficiencies are more likely to arise when the poor are relatively less informed than the rich (holding the total number of informed agents constant). This invites the interpretation that the informed rich dupe the less informed poor into voting against their interests. However, I show that the same inefficiency can arise if almost all below-median income earners are correctly informed, but many 'middle-class' voters are misinformed. Although the misinformed poor contribute to the existence of slippery slope inefficiencies, culpability for these inefficiencies does not lie solely on their shoulders.

Although the analysis focused on the case of public goods provision, the mechanism can be applied more broadly. The motivating example of same-sex marriage can be mapped into the model by considering the (reasonably) objective benefits (including favorable tax treatment, legal and societal rights and recognitions, such as visitation rights at hospitals, etc.) of extending marriage rights to same-sex couples on the one hand, against the personal costs that individuals may suffer from changing the structure of institutions that they may be more or less invested in (perhaps for spiritual reasons). In this context, there is an incentive for conservative voters (for whom the personal ideological costs are large) to obscure the many legal and financial benefits that marriage affords same-sex couples, which, if made explicit, would cause more liberal voters to support same-sex marriage. In this context, the analogue of greater income inequality is greater political polarization, and it is clear to see that slippery slope inefficiencies will be much more likely to arise in more polarized societies.

This analysis focused on learning by acquaintance, and thereby abstracted from the many varied sources of information through which agents may learn about the value of policies and reforms. The purpose was to demonstrate that political actors have strong incentives to distort policy choices in order to prevent learning. In a broader context, one can think of additional resources, such as the media, that political actors can bring to bear to skew the learning of voters. Given the proliferation of media sources with the internet and cable television, the rise of a partisan media, and the increasing tendency of voters to consciously select their news sources, this suggests fruitful avenues for further research.

References

Acemoglu, Daron and James A Robinson, "A theory of political transitions," *American Economic Review*, 2001, pp. 938–963.

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

Altman, Nancy J., *The Battle for Social Security*, Wiley, 2005.

Atkinson, Anthony B, "On the measurement of inequality," *Journal of economic theory*, 1970, *2* (3), 244–263.

Baccini, Leonardo and Lucas Leemann, "Information and Voting-How Voters Update Beliefs After Natural Disasters," in "EPSA 2013 Annual General Conference Paper," Vol. 435 2012.

Baker, Scott and Claudio Mezzetti, "A theory of rational jurisprudence," *Journal of Political Economy*, 2012, *120* (3), 513–551.

Battaglini, M. and S. Coate, "Inefficiency in Legislative Policymaking: A Dynamic Analysis," *The American Economic Review*, 2007, *97* (1), 118–149.

_ and _ , "A Dynamic Theory of Public Spending, Taxation, and Debt," *The American Economic Review*, 2008, *98* (1), 201–236.

Benabou, Roland, "Unequal societies: Income distribution and the social contract," *American Economic Review*, 2000, pp. 96–129.

_ and Efe A Ok, "Social mobility and the demand for redistribution: the POUM hypothesis," *The Quarterly Journal of Economics*, 2001, *116* (2), 447–487.

Benveniste, L.M. and J.A. Scheinkman, "On the differentiability of the value function in dynamic models of economics," *Econometrica: Journal of the Econometric Society*, 1979, pp. 727–732.

Black, D., "On the rationale of group decision-making," *The Journal of Political Economy*, 1948, *56* (1), 23–34.

Carey, G.L., *The Meaning of Matrimony: Debating Same-Sex Marriage*, Civitas,

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

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

Dewatripont, Mathias and Gerard Roland, "Economic reform and dynamic political constraints," *The Review of Economic Studies*, 1992, *59* (4), 703–730.

Downs, A., "An economic theory of political action in a democracy," *The Journal of Political Economy*, 1957, *65* (2), 135–150.

Epple, Dennis and Richard E Romano, "Ends against the middle: Determining public service provision when there are private alternatives," *Journal of Public Economics*, 1996, *62* (3), 297–325.

Fernandez, Raquel and Dani Rodrik, "Resistance to reform: Status quo bias in the presence of individual-specific uncertainty," *The American economic review*, 1991, pp. 1146–1155.

Fox, J. and G. Vanberg, "Law as a Discovery Procedure: An Informational Rationale for Broad Judicial Decisions," *Mimeo*, 2011.

Frank, Thomas, *What's the matter with Kansas?: how conservatives won the heart of America*, Macmillan, 2007.

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

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

_ **, John P Capitanio et al.**, "" Some of my best friends": Intergroup contact, concealable stigma, and heterosexuals' attitudes toward gay men and lesbians," *Personality and Social Psychology Bulletin*, 1996, *22*, 412–424.

Koch, J. and S. Mettler, "Who Perceives Government's Role in Their Lives?," *Working Paper*, 2012.

Malamud, Semyon and Fabio Trojani, "Variance, Risk and Median Preserving Spreads," 2009.

Meltzer, Allan H and Scott F Richard, "A rational theory of the size of government," *The Journal of Political Economy*, 1981, pp. 914–927.

Mendelson, Haim, "Quantile-preserving spread," *Journal of Economic Theory*, 1987, *42* (2), 334–351.

Moene, Karle Ove and Michael Wallerstein, "Inequality, social insurance, and redistribution," *American Political Science Review*, 2001, *95* (04), 859–874.

Parameswaran, Giri, "Ruling Narrowly and Broadly: Learning, Experimentation and Law Creation," 2013.

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

Program on International Policy Attitudes, "Americans on Foreign Aid and World Hunger: A Study of U.S. Public Attitudes," Technical Report, World Public Opinion.Org 2 February 2001.

Roberts, Kevin, "The theory of union behaviour: labour hoarding and endogenous hysteresis," Technical Report, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE 1989.

Romer, Thomas, "Individual welfare, majority voting, and the properties of a linear income tax," *Journal of Public Economics*, 1975, *4* (2), 163–185.

Rothschild, Michael and Joseph E Stiglitz, "Increasing risk: I. A definition," *Journal of Economic theory*, 1970, *2* (3), 225–243.

Samuelson, Paul A, "The pure theory of public expenditure," *The review of economics and statistics*, 1954, *36* (4), 387–389.

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

Tabellini, G. and A. Alesina, "Voting on the budget deficit," *The American Economic Review*, 1990, pp. 37–49.

Thistle, Paul D, "Ranking distributions with generalized Lorenz curves," *Southern Economic Journal*, 1989, pp. 1–12.

United States Census Bureau, "Household Income Inequality Within U.S. Counties: 2006-2010," Technical Report 2012.

World Public Opinion Poll, "American Public Opinion on Foreign Aid," Technical Report, World Public Opinion.Org 30 November 2010.

8 Appendix

Proof of Lemma 1. (1) Take $x > x_m$. Suppose $u_{(x_m(\gamma),I)}(g=0) > u_{(x_m(\gamma),I)}(g(x_m(\gamma),I))$. Let $H(x;x_m(\gamma)) = (1-\delta)\left(1-\alpha\frac{x}{x_m(\gamma)}\right) + \delta\left(1-\alpha\frac{x}{y_m}\right)\left(\frac{x_m(\gamma)}{y_m}\right)^{\frac{\alpha}{1-\alpha}}$. It follows algebraically that $u_{(x,I)}(g=0) > u_{(x,I)}(g^*(x_m(\gamma),I))$ iff $H(x;x_m(\gamma_1)) < 0$. By assumption $H(x_m(\gamma_1);x_m(\gamma_1)) < 0$. Moreover:

$$\frac{\partial H}{\partial x} = -\frac{(1-\delta)\alpha}{x_m(\gamma)} - \frac{\delta\alpha}{y_m}\left(\frac{x_m(\gamma)}{y_m}\right)^{\frac{\alpha}{1-\alpha}} < 0$$

Hence, $H(x;x_m(\gamma_1)) < 0$ for all $x > x_m(\gamma)$.

(2) Proved analogously $\square$

Proof of Proposition 1 . Since $\overline{\gamma} > \frac{1}{2}$, we know that $x_m(\gamma) < \infty$. We know that a type (y,I) proposer prefers her ideal public good level if $H(y,y) > 0$ and no public goods otherwise. Let $h(k) = H(ky_m, ky_m)$. Since $h(k) = (1-\delta)(1-\alpha)+\delta(1-\alpha k)(k)^{\frac{\alpha}{1-\alpha}}$, it is straight-forward to show that, $h\left(\frac{1}{\alpha}\right) = (1-\delta)(1-\alpha) > 0$ and $\lim_{k\to\infty} h(x) < 0$. Hence, by the intermediate value theorem, there exists some $\kappa \in \left(\frac{1}{\alpha},\infty\right)$ s.t. $h(\kappa) = 0$. Moreover, since $\frac{\partial h}{\partial k} = \delta\frac{\alpha}{1-\alpha}(k)^{\frac{\alpha}{1-\alpha}-1}(1-k) < 0$ whenever $k > \frac{1}{\alpha}$, κ is unique, and $h(k) > 0$ iff $k < \kappa$. The comparative statics follow by the implicit function theorem:

$$\frac{\partial\kappa}{\partial\alpha} = -\frac{(1-\delta)+\delta(\kappa)^{\frac{1}{1-\alpha}} + \frac{\delta}{(1-\alpha)^2}(\alpha\kappa-1)(\kappa)^{\frac{\alpha}{1-\alpha}}\ln(\kappa)}{\delta\frac{\alpha}{1-\alpha}(\kappa)^{\frac{\alpha}{1-\alpha}-1}(\kappa-1)} < 0$$

and

$$\frac{\partial\kappa}{\partial\delta} = \frac{-(1-\alpha)+(1-\alpha\kappa)(\kappa)^{\frac{\alpha}{1-\alpha}}}{\delta\frac{\alpha}{1-\alpha}\frac{1}{y_m}(\kappa)^{\frac{\alpha}{1-\alpha}-1}(\kappa-1)} < 0$$

$\square$

Proof of Proposition 2. For notational simplicity, let $x_i = x_m(\gamma, F_i)$ and let $y_i = y_m(F_i)$. Let $f_i(y) = F_i'(y)$ be the density of F_i. Consider the transformed random variable $\frac{Y_i}{y_i}$, and let $\hat{F}_i$ be its distribution function. Clearly $\hat{F}_i(z) = \Pr\left[\frac{Y_i}{y_i} \leq z\right] = F_i(y_i z)$. Then $\hat{f}_i(z) = y_i f_i(y_i z)$.

Let $I(x,F) = \int_0^x \gamma(y|F) f(y)\,dy$ be the number of informed people with income below x, given income profile F and informedness profile γ. Note that for any other income profile G, $G(y) = F(F^{-1}(G(y)))$ and so $g(y) = f(F^{-1}(G(y)))\frac{d}{dy}F^{-1}(G(y))$. Hence:

$$\begin{aligned} I(x,G) &= \int_0^x \gamma(y|G)\, g(y)\, dy \\ &= \int_0^x \gamma\left(F^{-1}(G(y))\,|F\right) f\left(F^{-1}(G(y))\right)\frac{d}{dy}F^{-1}(G(y))\,dy \\ &= \int_0^{F^{-1}(G(x))} \gamma(z|F)\, f(z)\, dz \\ &= I\left(F^{-1}(G(x)), F\right) \end{aligned}$$

where the second line follows from the income-rank independence of γ, and the third line involves a change of variable. Note also by construction that $I(x_m(\gamma,F),F) = \frac{1}{2}$.

Using the above property, note that: $I(x,F_i) = I\left(\hat{F}_i^{-1}(F_i(x)), \hat{F}_i\right) = I\left(\frac{x}{y_i}, \hat{F}_i\right)$ and so $I\left(\frac{x_i}{y_i}, \hat{F}_i\right) = \frac{1}{2}$. Using the property again, $I\left(\frac{x_i}{y_i}, \hat{F}_j\right) = I\left(\hat{F}_i^{-1}\left(\hat{F}_j\left(\frac{x_i}{x_j}\right)\right), \hat{F}_i\right)$. Finally, note that $I(\cdot,F)$ is strictly increasing in its first argument. Hence $\frac{x_1}{y_1} \leq \frac{x_2}{y_2}$ iff $\frac{1}{2} = I\left(\frac{x_1}{y_1}, \hat{F}_1\right) \geq I\left(\hat{F}_1^{-1}\left(\hat{F}_2\left(\frac{x_1}{y_1}\right)\right), \hat{F}_1\right) = I\left(\frac{x_1}{y_1}, \hat{F}_2\right)$. This will be true iff $\frac{x_1}{y_1} \geq \hat{F}_1^{-1}\left(\hat{F}_2\left(\frac{x_1}{y_1}\right)\right)$ or equivalently $\hat{F}_1\left(\frac{x_1}{y_1}\right) \geq \hat{F}_2\left(\frac{x_1}{y_1}\right)$. Finally, since $\hat{F}_i(z) = \bar{F}_i(zy_i)$, we have $F_1(x_1) \geq F_2\left(x_1\frac{y_2}{y_1}\right)$. $\square$

Proof of Lemma 2. Obviously $\frac{x_m(\gamma|F_i)}{y_m(F_1)} \geq 1$. Since F_2 is a median-normalized spread of F_1, then $F_2(zy_m(F_2)) \leq F_1(zy_m(F_1))$ for $z > 1$. Taking $z = \frac{x_m(\gamma,F_1)}{y_m(F_1)}$, we have $F_2\left(x_m(\gamma,F)\frac{y_m(F_2)}{y_m(F_1)}\right) \leq F_1(x_m(\gamma,F_1))$. A direct application of Proposition 2 completes the proof. $\square$

Proof of Lemma 3 . Let Y_1 and Y_2 be two income profiles. For each profile Y_i, let F_i and L_i be the associated distribution and Lorenz functions, and let

$\bar{y}(F_i)$ and $y_m(F_i)$ be the associated mean and median incomes. Furthermore, for each i, let $\bar{Y}_i = \frac{1}{\bar{y}(F_i)} Y_i$ and let $\bar{F}_i$ and $\bar{L}_i$ be the associated distribution and Lorenz functions. Clearly $\bar{F}_i(z) = F_i(\bar{y}(F_i) \cdot z)$, and $\bar{y}\left(\bar{F}_i\right) = 1$. Furthermore, by the scale-invariance property of Lorenz functions, $\bar{L}_i(p) = L_i(p)$.

First, using a variant of the proof in Thistle (1989), I show that Y_1 Lorenz dominates Y_2 only if $\bar{F}_2(y)$ second order stochastically dominates $\bar{F}_1(y)$. Recall, the Lorenz function is defined by: $L(p) = \frac{1}{E[Y]} \int_0^{F^{-1}(p)} y dF(y)$. Hence:

$$
\begin{aligned}
L_i(p) &= \bar{L}_i(p) \\
&= \int_0^{\bar{F}_i^{-1}(p)} y d\bar{F}_i(y) \\
&= p\bar{F}_i^{-1}(p) - \int_0^{\bar{F}_i^{-1}(p)} \bar{F}_i(y)\, dy
\end{aligned}
$$

where the third line uses integration by parts. Let $\bar{S}_i(x) = \int_0^x \bar{F}_i(y)\, dy$.

$$
\begin{aligned}
L_1(p) - L_2(p) &= p\left[\bar{F}_1^{-1}(p) - \bar{F}_2^{-1}(p)\right] - \left(\bar{S}_1\left(\bar{F}_1^{-1}(p)\right) - \bar{S}_2\left(\bar{F}_2^{-1}(p)\right)\right) \\
&= p\left[\bar{F}_1^{-1}(p) - \bar{F}_2^{-1}(p)\right] - \left(\bar{S}_1\left(\bar{F}_1^{-1}(p)\right) - \bar{S}_1\left(\bar{F}_2^{-1}(p)\right)\right) \\
&\quad + \left(\bar{S}_2\left(\bar{F}_2^{-1}(p)\right) - \bar{S}_1\left(\bar{F}_2^{-1}(p)\right)\right) \\
&= \int_{\bar{F}_1^{-1}(p)}^{\bar{F}_2^{-1}(p)} \left[\bar{F}_1(y) - p\right] dy + \left(\bar{S}_2\left(\bar{F}_2^{-1}(p)\right) - \bar{S}_1\left(\bar{F}_2^{-1}(p)\right)\right)
\end{aligned}
$$

Since $\bar{F}_2$ is a mean-preserving spread of $\bar{F}_1$, the second term (in parentheses) is non-negative, by Rothschild and Stiglitz (1970). So is the first term. Suppose $\bar{F}_1^{-1}(p) \leq \bar{F}_2^{-1}(p)$. Then $\bar{F}_1(y) - p \geq 0$ for $y \in \left[\bar{F}_1^{-1}(p), \bar{F}_2^{-1}(p)\right]$. Else, suppose $\bar{F}_1^{-1}(p) \geq \bar{F}_2^{-1}(p)$. Then $\bar{F}_1(y) - p \leq 0$ for $y \in \left[\bar{F}_2^{-1}(p), \bar{F}_1^{-1}(p)\right]$. In either case, the first term is positive. Hence $L_1(p) \geq L_2(p)$.

Second, using a variant of the proof in Malamud and Trojani (2009), I show that if F_2 is a median-normalized spread of F_1 and satisfies the single-crossing property, then, $\bar{F}_2$ second order stochastically dominates $\bar{F}_1$. It suffices to show that $\bar{S}_2(x) - \bar{S}_1(x) \geq 0$ for every x.

Let $\frac{\bar{y}(F_2)}{y_m(F_2)} \geq \frac{\bar{y}(F_1)}{y_m(F_1)}$. Then $\bar{S}_2(z) - \bar{S}_1(z) = \int_0^z \left(\bar{F}_2(y) - \bar{F}_1(y)\right) dy =$

$\int_0^z (F_2(\bar{y}(F_2)\cdot y) - F_1(\bar{y}(F_2)\cdot y))\,dy$. Since F_2 is a median-normalized spread of F_1, $F_2(y_m(F_2)\cdot z) \geq F_1(y_m(F_1)\cdot z)$ whenever $z \leq 1$. Hence $F_2(\bar{y}(F_2)\cdot y) \geq F_2\left(\frac{\bar{y}(F_1)}{y_m(F_1)} y \cdot y_m(F_2)\right) \geq F_1(\bar{y}(F_1)\cdot y)$ whenever $y \leq \frac{y_m(F_1)}{\bar{y}(F_1)}$. Hence $\bar{S}_2(z) - \bar{S}_1(z) \geq 0$ for $z \leq y_m(F_1)$.

By the single-crossing property, we know that there is a single z for which $F_2(\bar{y}(F_2)\cdot z)$ and $F_1(\bar{y}(F_1)\cdot z)$ cross. Hence if $F_2(\bar{y}(F_2)\cdot z) < F_1(\bar{y}(F_1)\cdot z)$ for some z, then $F_2(\bar{y}(F_2)\cdot y) < F_1(\bar{y}(F_1)\cdot y)$ for all $y > z$. This implies that $\bar{S}_2(z) - \bar{S}_1(z)$ is decreasing for z above the crossing point. Now, take $\lim_{z\to\infty} \bar{S}_2(z) - \bar{S}_1(z)$. Integrating by parts gives:

$$
\begin{aligned}
& \lim_{z\to\infty} y\left[F_2(\bar{y}(F_2)\cdot y) - F_1(\bar{y}(F_1)\cdot y)\right] - \int_0^\infty y\left(dF_2(\bar{y}(F_2)\cdot y) - dF_1(\bar{y}(F_1)\cdot y)\right) \\
= & -\frac{1}{\bar{y}(F_2)}\int_0^\infty w\,dF_2(w) + \frac{1}{\bar{y}(F_1)}\int_0^\infty w\,dF_1(w) \\
= & \ 0
\end{aligned}
$$

Hence $\bar{S}_2(z) - \bar{S}_1(z) \geq 0$ for all z, and so $\bar{F}_1$ second order stochastically dominates $\bar{F}_2$. This completes the proof. □

Proof of Corollary 1. Let Y_2 be a median-normalized spread of Y_1 and suppose $\frac{\bar{y}(F_2)}{y_m(F_2)} = \frac{\bar{y}(F_1)}{y_m(F_1)}$. Then $F_2(\bar{y}(F_2)z) = F_2\left(\frac{\bar{y}(F_1)}{y_m(F_1)} z \cdot y_m(F_2)\right)$. By median-normalized spreads, $F_2\left(\frac{\bar{y}(F_1)}{y_m(F_1)} z \cdot y_m(F_2)\right) \geq F_1\left(\frac{\bar{y}(F_1)}{y_m(F_1)} z \cdot y_m(F_1)\right)$ whenever $\frac{\bar{y}(F_1)}{y_m(F_1)} z < 1$, and $F_2\left(\frac{\bar{y}(F_1)}{y_m(F_1)} z \cdot y_m(F_2)\right) \leq F_1\left(\frac{\bar{y}(F_1)}{y_m(F_1)} z \cdot y_m(F_1)\right) = F_1(\bar{y}(F_1)\cdot z)$ whenever $\frac{\bar{y}(F_1)}{y_m(F_1)} z \geq 1$. Hence $F_2(\bar{y}(F_2)z) \geq F_1(\bar{y}(F_1)z)$ whenever $z < \frac{y_m(F_1)}{\bar{y}(F_1)}$ and $F_2(\bar{y}(F_2)z) \leq F_1(\bar{y}(F_1)z)$ whenever $z \geq \frac{y_m(F_1)}{\bar{y}(F_1)}$. Since $F_2(\bar{y}(F_2)\cdot z^*) = \frac{1}{2} = F(\bar{y}(F_1)\cdot z^*)$, there is a single-crossing at $z^* = \frac{y_m(F_1)}{\bar{y}(F_1)}$.Hence, by Lemma 3, median-normalization implies Lorenz dominance. □

Proof of Proposition 3. Suppose $\gamma \in \Gamma^L$. Then, by monotonicity, $Q(g,\gamma) \in \Gamma^L$ for any $g \geq 0$, and so the poor will be pivotal in every period, regardless of their choice of g. Since the dynamics of the game arise only out of potentially changing identity of the effective median (as learning occurs), the game strategically collapses to one in which the poor implement their ideal stage policy in every period. Hence, for any $\gamma \in \Gamma^L$, $g^*(\gamma) = \left(\frac{\alpha A}{p}\frac{\overline{y}}{y_L}\right)^{\frac{1}{1-\alpha}}$ and so:

$$v_\theta^*(\gamma) = \frac{1}{1-\delta}\left[y_t + \left(A_t - \alpha A\frac{y_t}{y_L}\right)\left(\frac{\alpha A}{p}\frac{\overline{y}}{y_L}\right)^{\frac{\alpha}{1-\alpha}}\right]$$

Now, take $\gamma \in \Gamma^H$. This implies that the pivotal voter is informed and has income y_H. (To see this, note that, by construction, a coalition between the informed rich and either the informed poor or the misinformed will command a majority. For any two feasible policies $(\tau, g) \leq (\tau', g')$ with $g' \leq g(y_L, I)$, the misinformed always prefer (τ, g) to (τ', g') and the informed poor always prefer (τ', g') to (τ, g). Hence, the preference of the informed rich over these policies will be decisive.)

Let F be the set of bounded functions on Γ. Define the operator: $T : F \to F$ by

$$T[v](\gamma) = \begin{cases} \max_{g\geq 0}\left\{\left(1-\frac{pg}{\overline{y}}\right)y_H + Ag^\alpha + \delta v(Q(g,\gamma))\right\} & \gamma \in \Gamma^H \\ \frac{1}{1-\delta}\left[y_H + A\left(1-\alpha\frac{y}{y_L}\right)\left(\frac{\alpha A}{p}\frac{\overline{y}}{y_L}\right)^{\frac{\alpha}{1-\alpha}}\right] & \gamma \in \Gamma^L \end{cases}$$

Since v is bounded and $\left(1-\frac{pg}{\overline{y}}\right)y_H + Ag^\alpha$ has an upper bound (that is achieved when $g = \left(\frac{\alpha A}{p}\frac{\overline{y}}{y_H}\right)^{\frac{1}{1-\alpha}}$), it must be that $T[v]$ is bounded.

I show that $T[v]$ is a contraction mapping. It suffices to show that T satisfies Blackwell's conditions. Take $v, w \in F$ and suppose $v(\gamma) \geq w(\gamma)$ for all γ. For $\gamma \in \Gamma^L$, $T[v](\gamma) = T[w][\gamma]$. Suppose $\gamma \in \Gamma^H$ and let $g_v(\gamma)$ and $g_w(\gamma)$ be the optimal policy functions, given v and w, respectively. Then:

$$\begin{aligned} T[v](\gamma) &= \left(1-\frac{pg_v(\gamma)}{\overline{y}}\right)y_H + A(g_v(\gamma))^{\alpha} + \delta v(Q(g_v(\gamma),\gamma)) \\ &\geq \left(1-\frac{pg_w(\gamma)}{\overline{y}}\right)y_H + A(g_w(\gamma))^{\alpha} + \delta v(Q(g_w(\gamma),\gamma)) \\ &\geq \left(1-\frac{pg_w(\gamma)}{\overline{y}}\right)y_H + A(g_w(\gamma))^{\alpha} + \delta w(Q(g_w(\gamma),\gamma)) \\ &= T[w][\gamma] \end{aligned}$$

Hence $T[v](\gamma) \geq T[w][\gamma]$, which demonstrates monotonicity. Similarly, for $\gamma \in \Gamma^L$, $T[v+c](\gamma) = T[v](\gamma)$ and:

$$\begin{aligned} T[v+c](\gamma) &= \max_{g\geq 0}\left\{\left(1-\frac{pg}{\overline{y}}\right)y_H + Ag^{\alpha} + \delta(v(Q(g,\gamma))+c)\right\} \\ &= \max_{g\geq 0}\left\{\left(1-\frac{pg}{\overline{y}}\right)y_H + Ag^{\alpha} + \delta v(Q(g,\gamma))\right\} + \delta c \\ &= T[v](\gamma) + \delta c \end{aligned}$$

Hence $T[v+c](\gamma) \leq T[v](\gamma) + \delta c$, which verifies discounting. Hence, T is a contraction mapping and so it contains a unique fixed point $v^*(\gamma) \in F$.

In fact, this fixed point is the value function for a (y_H, I)-type agent. It follows that, when $\gamma \in \Gamma^H$, the policy is given by:

$$g^*(\gamma) = \arg\max_{g\geq 0}\left(1-\frac{pg}{\overline{y}}\right)y_H + Ag^{\alpha} + \delta v^*(Q(g,\gamma))$$

and the value functions for the remaining types (over the region $\gamma \in \Gamma^H$) are defined as in the statement of the proposition. $\square$

Proof of Proposition 4. For notational convenience, let $\eta(g) = \left(1-\frac{pg}{\overline{y}}\right)y_H + Ag^{\alpha}$. If the rich choose their desired stage policy $g_H^* = g(y_H, I)$ and immediately surrender political power to the poor, they will receive payoff $\bar{v} = \eta(g_H^*) + \frac{\delta}{1-\delta}\eta(g_L^*)$, where $g_L^* = g(y_H, I)$ is the ideal policy of the poor. Recall η is a concave function that is maximized at g_H^*, and that $g_L^* > g_H^*$.

By construction, if $e \in \mathcal{E}_{Eff}$, then $\bar{v} > \frac{1}{1-\delta}\eta(0)$, whilst the opposite is true if $e \in \mathcal{E}_{SS}$.

For each γ, define $\chi^t(\gamma, e)$ as the equilibrium informedness profile after t periods. I.e. $\chi^1(\gamma, e) = Q(g^*(\gamma), \gamma)$, and $\chi^k(\gamma, e) = Q\left(g^*\left(\chi^{k-1}(\gamma, e)\right), \chi^{k-1}(\gamma, e)\right)$ for each $k \geq 2$. Similarly, define $G^t(\gamma, e)$ as the equilibrium level of public goods provision at time t. Clearly $G^t(\gamma, e) = g^*(\chi^t(\gamma, e))$.

First I show that $e \in \mathcal{E}_{SS}$ implies that the rich should never concede to the poor. Suppose there is an equilibrium in which the rich concede power to the poor. Let $\hat{t}$ be the first period in which the poor are in power. Hence $\chi_L^{\hat{t}-1}(\gamma, e) \leq \frac{1}{2\phi} < \chi_L^{\hat{t}}(\gamma, e)$. At $t = \hat{t} - 1$, choosing $G^{\hat{t}-1}(\gamma, e)$ gives the rich utility $\eta\left(G^{\hat{t}-1}\right) + \frac{\delta}{1-\delta}\eta(g_L^*)$. But since $e \in \mathcal{E}_{SS}$, $\frac{1}{1-\delta}\eta(0) > \bar{v} \geq \eta\left(G^{\hat{t}}\right) + \frac{\delta}{1-\delta}\eta(g_L^*)$. This implies that there is a favorable deviation for the rich to offer $g = 0$ at every $t \geq \hat{t} - 1$, and so the rich should not concede power at $\hat{t}$. But this implies there is no first period when it is optimal for the rich to concede power.

Next, I show that $e \in \mathcal{E}_{SS}$ implies $G^t(\gamma, e) \to 0$ for every $\gamma \in \Gamma^H$. Suppose not. Then for some $\varepsilon > 0$, there exists a sub-sequence $\{G^{t_k}\}$ of $\{G^t\}$, such that $G^{t_k} \geq \varepsilon$ for each k. Define the sequence $\left\{\tilde{\chi}^k\right\}_{k=0}^{\infty}$, where $\tilde{\chi}^0(\gamma, e) = \gamma$ and for every $k \geq 1$, $\tilde{\chi}^k(\gamma, e) = Q\left(G^{t_k}, \tilde{\chi}^{k-1}(\gamma, e)\right)$. This is the sequence of informedness profiles that would arise by replacing $G^t = 0$ whenever $G^t < \varepsilon$. By monotonicity, $\chi^{t_k}(\gamma) \geq \tilde{\chi}^k(\gamma)$. Since Q is continuous in g, there exists $\rho(\varepsilon, \gamma) > 0$ s.t. $Q(g, \gamma) - Q(0, \gamma) > \rho(\varepsilon, \gamma)$ whenever $g > \varepsilon$. Moreover, ρ is continuous in γ (since Q is continuous in γ). Since Γ^H is compact, $\rho(\varepsilon, \gamma)$ achieves its lower bound on Γ^H. Let $\rho(\varepsilon) = \min_{\gamma \in \Gamma^H} \rho(\varepsilon, \gamma)$. Clearly $\rho(\varepsilon) > 0$. Hence, for each k, $\tilde{\chi}^k(\gamma, e) - \tilde{\chi}^{k-1}(\gamma, e) \geq Q\left(\varepsilon, \tilde{\chi}^{k-1}(\gamma, e)\right) - Q\left(0, \tilde{\chi}^{k-1}(\gamma, e)\right) > \rho(\varepsilon) > 0$ — i.e. $\tilde{\chi}^k(\gamma) > \tilde{\chi}^{k-1}(\gamma) + \rho(\varepsilon)$ for every $k \geq 1$. By induction, $\tilde{\chi}^k(\gamma) > k\rho(\varepsilon) + \gamma$. Then, for $k > \frac{\frac{1}{2\phi} - \gamma_L}{\rho(\varepsilon)} = K(\varepsilon)$, $\chi_{eL}^{t_k}(\gamma) \geq \tilde{\chi}_{eL}^k(\gamma) > \frac{1}{2\phi}$ — i.e. for $t \geq t_{K(\varepsilon)}$, $\chi_e^{t_k}(\gamma) \in \Gamma^L$; the rich will eventually surrender political power to the poor. But this cannot be.

Next, consider an economy $e \in \mathcal{E}_{Eff}$. This implies $\eta(0) < \eta(g_L^*) < \eta(g_H^*)$

and so there exists $\hat{g} \in (0, g_H^*)$ such that $\frac{1}{1-\delta}\eta(\hat{g}) = \eta(g_H^*) + \frac{\delta}{1-\delta}\eta(g_L^*) = \bar{v}$. The rich would rather concede power to the poor (and receive utility $\bar{v}$) than receive $\hat{g}$ (or fewer) public goods forever into the future. Using this fact, I show that $e \in \mathcal{E}_{Eff}$ implies the there exist $T(\gamma, \varepsilon) < \infty$ at which the rich concede power to the poor. Suppose not. Then, for every $t \geq 1$, $\chi^t(\gamma, e) \leq \frac{1}{2\phi}$. This requires that $G^t(\gamma, e) \to 0$. (If not, using the same argument as in the previous paragraph, we can find a sub-sequence that is bounded above zero that guarantees that power shifts within a finite number periods.) Convergence to zero implies that there is some $\hat{T}(\hat{g})$ s.t. $G^t(\gamma, e) < \hat{g}$ whenever $t > \hat{T}(\hat{g})$. Clearly, this policy path cannot be optimal from time $\hat{T}(\hat{g})$ onwards — the rich would be better off surrendering power to the poor at $t = \hat{T}(\hat{g})$ (or sooner). Hence, (by monotonicity) there must exist some $T(\gamma, e)$ s.t. $\chi^t(\gamma, e) > \frac{1}{2\phi}$ for every $t > T(\gamma, e)$. □

Proof of Lemma 4 . Suppose $e \in \mathcal{E}_{SS}$. Suppose the objective function in the agent's maximization problem is differentiable and concave, and so the first order conditions are sufficient for the maximum. By the first order conditions:

$$\frac{\eta'(Q^{-1}(y^*, \gamma))}{Q_g(Q^{-1}(y^*, \gamma))} + \delta v'(y^*) = 0$$

By the envelope theorem:

$$v'(\gamma) = -\eta'\left(Q^{-1}(y^*, \gamma)\right) \frac{Q_\gamma(Q^{-1}(y^*, \gamma), \gamma)}{Q_g(Q^{-1}(y^*, \gamma), \gamma)}$$

Hence, the Euler equation is:

$$\eta'(g^*(\gamma)) \frac{1}{Q_g(g^*, \gamma)} = \delta\eta'(g^*(y^*(\gamma))) \frac{Q_\gamma(g^*(y^*(\gamma)), y^*(\gamma))}{Q_g(g^*(y^*(\gamma)), y^*(\gamma))}$$

where $y^*(\gamma) = Q(g^*(\gamma), \gamma)$. Since $Q_g > 0$ and $Q_\gamma > 0$, this implies that if $\eta'(G_t(\gamma, e)) = 0$ for some t, then $\eta'(G_{t'}(\gamma, e)) = 0$ for all $t' > t$. If $G_t(\gamma, e) = g_H^*$ for some t, then it must remain at that level indefinitely. This is impossible, since $G_t \to 0$. Hence, $G_t(\gamma, e) < g_H^*$ for all t.

I am left to show that the objective function is indeed concave and differentiable. Since the rich never concede political power, the value function satisfies:

$$v^*(\gamma) = \max_{y\in[0,\bar{y}(\gamma)]} \left\{\eta\left(Q^{-1}(y,\gamma)\right) + \delta v^*(y)\right\}$$

where $\bar{y}(\gamma) = \lim_{g\to\infty}[Q(g,\gamma)]$.

I show that v^* is concave on Γ^H. Let F_C be the space of bounded concave functions. It suffices to show that $T : F_C \to F_C$. Let v be a concave function. Let $\gamma, \gamma \in \Gamma^H$ and let $\gamma_\lambda = \lambda\gamma + (1-\lambda)\gamma'$. Denote $y = Q(g^*(\gamma), \gamma)$, $y' = Q(g^*(\gamma'), \gamma')$ and $y_\lambda = Q(g^*(\gamma_\lambda), \gamma_\lambda)$. Then

$$\begin{aligned}
T[v](\gamma_\lambda) &= \eta\left(Q^{-1}(y_\lambda, \gamma_\lambda)\right) + \delta v(y_\lambda) \\
&\geq \eta\left(Q^{-1}(\lambda y + (1-\lambda)y', \lambda\gamma + (1-\lambda)\gamma')\right) + \delta v(\lambda y + (1-\lambda)y') \\
&\geq \lambda\eta\left(Q^{-1}(y,\gamma)\right) + (1-\lambda)\eta\left(Q^{-1}(y',\gamma')\right) + \delta[\lambda v(y) + (1-\lambda)v(y')] \\
&= \lambda T[v](\gamma) + (1-\lambda)T[v](\gamma')
\end{aligned}$$

where the second line uses the joint concavity of $\eta(Q^{-1}(y,\gamma))$. Hence the fixed point v^* is concave.

Next, I show that $v^*$is differentiable everywhere in the interior of Γ^H. Take any $\gamma^0 \in \left(1 - \frac{1}{2\phi}, \frac{1}{2\phi}\right)$ and let $y_0^* = Q(g^*(\gamma_0), \gamma_0)$. For γ in the neighborhood of γ_0, let $\psi(\gamma) = \eta(Q^{-1}(y_0^*, \gamma)) + \delta v^*(y_0^*)$. By the optimality of v^*, $v^*(\gamma) \geq \psi(\gamma)$ and by construction $\psi(\gamma_0) = v^*(\gamma_0)$. Moreover, $\psi(\gamma)$ is differentiable in γ, since η and Q are differentiable. Hence, by Theorem 1 in Benveniste and Scheinkman (1979), v^* is differentiable at γ_0. Since γ_0 was chosen arbitrarily, then v^* is differentiable everywhere in the interior of Γ^H. □