

Self-Signaling and Voting for Redistribution*

Juho Alasalmi[†]

July 2023

[Latest version here](#)

Abstract

There seems to be more to voting than policy. Nevertheless, models of voting tend to focus on aggregate policy preferences. I model voting for redistribution under uncertainty over future incomes with imperfect memory and anticipatory utility. Imperfect memory gives votes meanings as signals of future income prospects. Anticipatory utility motivates concern over the meanings of votes. Voting becomes self-signaling: Voting for low taxation is consistent with the desirable belief of high future consumption. Voting and policy preferences diverge, and voting does not aggregate policy preferences: fewer voters vote for redistribution than prefer redistribution. Higher income risk increases the value of self-signaling, which decreases the demand for redistribution. If voters do not perceive themselves as pivotal, a policy trap, where a decisive coalition votes against its best interest, may arise.

JEL classification: D31; D72; D83; D91

Keywords: self-signaling, anticipation, redistribution, voting

1 Introduction

Voters seem to sometimes vote against their self-interest, at least in the narrow sense of material utility. I show how voters wishing to expect high incomes choose apparently dominated actions in the context of voting for a redistributive policy. I model voting as a possibility for self-signaling or expressing desirable future outcomes: voting for low taxation identifies the voter with the party that supports low taxation and with its supporters who benefit from low taxation. Such identification is consistent with the desirable belief of high future income and consumption.

The question of why the poor do not expropriate the rich by voting for a high tax rate and large transfers is a subject of extensive literature. The focus of this literature has been on understanding why policy preferences in aggregate may not support redistribution. The question of why the poor do not *vote* for redistribution has received less attention. The

*I thank Toke Aidt, Luna Bellani, Benjamin Blumenthal, Frank Bohn, Allan Drazen, Florian Dorn, Holly Dykstra, Susanne Goldlücke, Arye Hillman, Michelle Jordan, Panu Poutvaara, Konstantinos Theocharopoulos, Janne Tukiainen, Arnaud Wolff and many other participants at workshops and conferences at International Institute of Public Finance Video Conference, 8/2020; TWI Microseminar, Kreuzlingen, 12/2020; Public Choice Society Meeting, Nashville, 3/2022; European Public Choice Society Meeting, Braga 4/2022; RUCAM Workshop on Analytical Modelling Approaches to Understanding Democracy, Nijmegen 8/2022; RGS Doctoral Conference in Economics, Bochum 3/2023; 7th Swiss Theory Day, St. Gallen 6/2023; 13th Conference on Economic Design, Girona 6/2023 for valuable comments and discussions.

[†]Department of Economics, University of Konstanz. Email: juho.alasalmi@uni-konstanz.de

distinction between policy preferences and voting is moot if all voters care about is the policy. However, if voting serves a dual purpose, on the one hand contributing to the popularity of the policy voted for but on the other hand expressing or affirming political identity, then voting and policy preferences may diverge.

People seem to be willing to take actions that correlate with but do not cause good outcomes. Such diagnostic actions have been discussed by Nozick (1969) and experimentally studied by Quattrone and Tversky (1986), Sloman et al. (2010), and Fernbach et al. (2014). Facing a choice where an underlying unknown state of the world is known to influence both the action and a future outcome, the action taken may be interpreted as revealing information about the state of the world and, hence, about the forthcoming outcome. Such behavior has also been called self-signaling (Bodner and Prelec, 2003; Bernheim and Thomadsen, 2005; Bénabou and Tirole, 2004, 2006b, 2009, 2011; Mijović-Prelec and Prelec, 2010; Grossman, 2015). If we remember our actions better than the motives of taking those actions, we may condition our self-inference and thus expectations about future outcomes on the actions we took. Our actions then influence our beliefs.

While imperfect recall makes actions informative, the willingness to take diagnostic actions may arise from anticipatory utility. Expecting good outcomes increases current anticipation. Utility from anticipation then motivates us to take actions that are consistent with the expectations of good outcomes. The role of anticipation of future utility as a part of currently experienced utility has been discussed by Bentham (1789) and Jevons (1905), experimentally studied by Löwenstein (1987) and formally modelled, for instance, by Caplin and Leahy (2001), Bernheim and Thomadsen (2005), Brunnermeier and Parker (2005), and Bénabou (2008, 2013).

Voting for redistribution is potentially a diagnostic action: First, future incomes are uncertain. Second, since voters expecting higher incomes tend to demand less redistribution and more goods, future incomes are associated both with our voting and our future consumption. Third, anticipating high consumption is more desirable than anticipating low consumption. Furthermore, there seems to be more to voting than reporting policy preferences: an extensive literature studies the direct or expressive benefits of voting (e.g. Tullock (1971), Hamlin and Jennings (2011), Brennan and Brooks (2013), Hamlin and Jennings (2018)).

I add two components, imperfect recall and anticipatory utility, to a model of voting for redistribution under uncertainty about future income. In the model, voters vote either for a low or high income tax rate. At the time of voting, voters know their current income and have expectations about how their incomes will evolve in the future. However, due to imperfect recall, the memory of voters is malleable, and after voting, voters form expectations about their future income partly based on their voting choice. This gives votes meanings as signals of future income prospects. After the election and before the realization of future incomes, voters anticipate their future consumption. Anticipation creates an incentive for self-signaling: expectations of higher consumption lead to higher anticipatory utility. In equilibrium, voting

for low taxation is associated with high income and voting for high taxation is associated with low income. Thus, recalling having voted for low taxation allows inference of more desirable beliefs than recalling having voted for high taxation. Expecting this at the time of voting, voters face a trade-off between the instrumental benefits of contributing to the popularity of their preferred policy and the self-signaling benefits derived from the meanings of votes. There is more to voting than policy, and voting and policy preferences diverge.

I show how self-signaling desirable mobility prospects limits the demand for redistribution: As voting and policy preferences diverge, voting does not aggregate policy preferences and there are fewer voters who vote for high taxes than prefer high taxes. I also study how the value of self-signaling depends on the distribution of mobility prospects. Deterioration of the prospects at the low end of the income distribution leads to incentives of “wanting to look away” to take refuge in focusing on the prospects of upward mobility. Such denial is facilitated by voting for the low tax rate and identifying with the voters who benefit from the low tax rate. On the other hand, as the prospects of being at the high end of the income distribution improve and the prizes with which the economy rewards the most successful individuals increase, voting to self-signal these high mobility prospects becomes more attractive. Combining these effects, I find that an increase in income risk decreases demand for redistribution.

Even if voting does not aggregate policy preferences and the policy preferred by the minority gains the majority of votes, voters do not need to be voting against their self-interest. Voters’ anticipation is a part of their utility. If a voter votes as if she was pivotal and votes for low tax, she reveals that anticipation of high income brings her higher utility than more extensive redistribution. Voters may, however, not perceive themselves as pivotal. Such voters underweight the consumption utility relative to the voters who vote as if they were pivotal. Thus, they may vote for the low tax rate even if, if decisive, they would vote for the high tax rate. If such voters form a decisive coalition, they face a collective action problem: they would be best off voting for the high tax rate. Nevertheless, it is individually rational to not face reality but rather to identify with the supporters of the low tax policy to enjoy anticipation. Thus, a policy trap, where a decisive coalition of voters votes against the high tax rate while collectively preferring to vote and induce the low tax rate, may arise.

I further show how such policy traps are more likely when voters value the meanings of their votes, do not perceive their votes as pivotal, and perceive the voters of different parties as facing very different income prospects. Thus, political campaigning that focuses on policy and the importance of voting in determining policy makes policy traps less likely. Political campaigning that focuses on the political identification of voters and frames voters of different parties in extreme ways makes policy traps more likely.

The next section positions this work in the literature and highlights contributions. Section 3 lays out the model. The voting rules and their responses to changes in model parameters are studied in Section 4. Section 5 derives the relevant measure of individual demand for redistribution that can be aggregated in Section 6. Section 7 studies welfare and Section 8

concludes. All proofs are in the appendix.

2 Literature

Formal modelling of voting for redistribution started with the contributions of Romer (1975) and Meltzer and Richard (1981).¹ In the simplest model, below-mean-income voters vote for high redistribution and above-mean-income voters vote for low redistribution. With a right-skewed income distribution, the decisive median voter is among the below-mean-income voters. The poor are thus in the majority and, hence, in a position to expropriate the rich. Later work has attempted to explain what restricts the demand for redistribution among the below-mean-income voters. Explanations most closely related to this paper are prospects of upward mobility, biased beliefs, and social identity.

In terms of the structure of the economy, polity, and policy preferences, this paper is related to the models of income mobility (Hirschman and Rothschild, 1973; Bénabou and Ok, 2001), where voters vote for a redistributive policy that will apply to their future incomes. Currently poor voters may vote against redistribution in the expectation of upward income mobility. Bénabou and Ok (2001) show how when future income is increasing and concave in today's income, a mean income voter expects to have above mean income in future by Jensen's inequality, and, thus, how the majority of voters with rational expectations caring only about their after-tax income may prefer low taxation. This restrictive class of income dynamics has led to a critique of this model as an explanation of the prospects of upward mobility hypothesis (Alesina and Giuliano, 2011), and it has been suggested that overoptimism may be more plausible an explanation (Alesina et al., 2001; Alesina and Glaeser, 2004; Alesina and Giuliano, 2011). Indeed, the perceived and actual prospects of upward mobility have been observed to diverge (Alesina et al., 2018). In my model, low demand for redistribution does not mechanically arise from the structure of mobility prospects. Rather, a belief in upward mobility is valued directly and low demand for redistribution arises from voters' choices to vote consistent with favorable mobility prospects.

In terms of enriching the utility of voters to contain their beliefs, this paper relates to models that have connected systematic biases in beliefs and voting (Piketty, 1995; Bénabou and Tirole, 2006a; Minozzi, 2013; Bénabou, 2013). Minozzi (2013) models how endogenous overoptimism over future income decreases the demand for redistribution. In his model, anticipation of future income motivates overly optimistic beliefs. However, optimism is restricted by its influence on voting behavior: too much optimism leads to economically suboptimal choices in the voting booth. Instead, I model voting as a tool for belief management conditioning voting on the hopes of the future: voting serves the general outlook voters have for their future rather than the reverse.

In terms of the interpretation of voting as identification, closely related are also models

¹Harms and Zink (2003), Borck (2007), Alesina and Giuliano (2011), Acemoglu et al. (2015), Gallice (2018), Bellani and Ursprung (2019), and Hillman (2019) review this literature.

of social identity. In Shayo (2009), voting is influenced by chosen group identity. Group identity manifests as the internalization of the group status, and status is modelled as the average payoff in the group. If the status of the poor deteriorate, shifting identification from the poor to the broader group of the nation may be attractive to the poor as the latter identification provides higher status. Such a shift in identification then decreases demand for redistribution. In Shayo (2009), following Akerlof and Kranton (2000), identity is modelled as preference-based: identity corresponds to preferences, identity is what is valued. Here, following Bénabou and Tirole (2011), identity is modelled as belief-based: identity is a belief about oneself.² Also, while Shayo (2009), as well as Penn (2008) and Lindqvist and Östling (2013), have focused on identification along voter attributes such as social class, nationality, ethnicity, and race, the model here can be seen as a model of partisan identification.

In all these models, voting is a consequent to preference or belief formation or identification, policy preferences coincide with voting preferences, and the role of voting is solely to aggregate honestly reported policy preferences. The act of voting itself is left implicit. This approach to understanding voting has sometimes been called instrumental: Voters vote for the policy that they expect to leave them best off. Another approach to understanding voting is expressive voting: voters, rather than being motivated by inducing their preferred policy outcome, are motivated by more direct benefits that flow from the act of casting a vote itself (Buchanan, 1954; Tullock, 1971; Brennan and Lomasky, 1993; Brennan and Hamlin, 1998; Hamlin and Jennings, 2018). An often suggested source of expressive benefits is confirming or managing identity: Tullock (1971) suggests that high-income voters' support for redistribution is due to a desire to obtain a feeling of being charitable, Glazer (1987) suggests voting is motivated by the possibility of conveying a certain image to oneself or to others, Schuessler (2000) defines the content of expressive benefits to be the identification with the other voters making the same voting choices. Hillman (2010) defines expressive benefits arising from the self-interested confirmation of one's identity. Hamlin and Jennings (2011) pinpoint the expressive benefits to flow from the vote's "*symbolic or representational aspect, [...] from its meaning*" (page 649). Following these ideas, I model voting as tool to manage beliefs and identity rather than as a result of belief formation or identification.

Following Bernheim and Thomadsen (2005) and Bénabou and Tirole (2006b, 2011), I model expectations or identity management by voting as self-signaling. Bernheim and Thomadsen (2005) model the appeal of diagnostic actions as a signaling game between temporal selves with self-signaling motivated by anticipatory utility. In their model, actions correlate (in equilibrium) with some random variables whose realizations agents know as they choose their actions but forget before later anticipating the outcomes of these actions. Knowing that the actions will, at a later date, be used to infer about the underlying random variable that also

²For the distinction between preference-based and belief-based models of identity, see Charness and Chen (2020). This distinction also parallels the literature on "esteem theory" (Brennan and Pettit, 1993; Cowen, 2002; Brennan and Pettit, 2004; Brennan and Brooks, 2007) which explains norm compliance as a desire to maintain social image or esteem rather than as a preference.

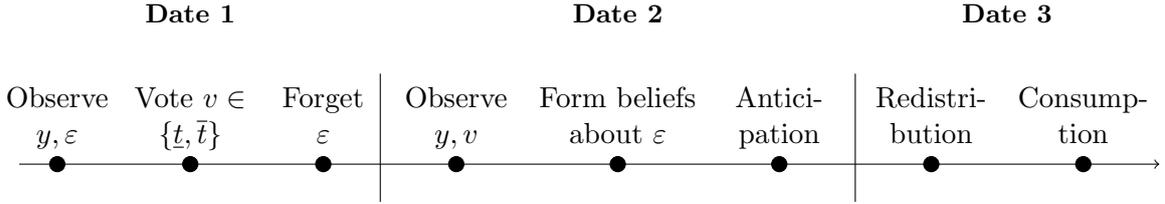


Figure 1: Timeline.

correlates with the outcomes, agents are motivated to choose actions that lead to desired inferences. In the honor-stigma model of Bénabou and Tirole (2006b), agents' choices are motivated by status, which directly depends on the observer's posterior belief of the agent's type. In Bénabou and Tirole (2011), identity is modelled as agent's posterior belief about her own type, and agents care about their identity directly or by anticipating the outcomes their believed type expects.

These models are signaling games with continuous sender type and binary signal with the optimal sender strategy characterized as a threshold rule. There is thus partial pooling, and the type distribution determines the payoffs. A comparative static effect of interest is then to study how the sender's signaling incentives and the demand for signals depend on the distribution of sender types. Such distributional comparative static effects have recently been studied by Adriani and Sonderegger (2019). They show that an increase in the dispersion of the type distribution can either decrease or increase the demand for signals depending on the signaling strategy: While the effect on signaling strategy is unambiguous, the change in type distribution changes the demand for signals mechanically and the resulting composition effect may either dampen or reinforce the effect of a change in signaling strategy. I extend their results by considering a population of senders that, in addition to their private types, are heterogeneous in an observable parameter. This results in heterogeneity in signaling strategies such that there are both negative and positive demand effects in the population of senders. I show how the ambiguous effects in such a population can aggregate to an unambiguous aggregate demand effect.

3 The Model

Consider a model of voting for a redistributive policy to be applied in future under uncertainty about future income. Voting takes place at date 1. At date 2, voters anticipate their date 3 consumption. At date 3, future incomes realize, are redistributed, and consumed. The timeline is depicted in Figure 1.

3.1 Income Dynamics

At dates 1 and 2, a voter earns current income $y \in Y$ distributed G with a continuous and unimodal density. Each voter observes and knows their current income at all dates. At date

3, voter's income changes by ε such that at date 3, the voter earns $y + \varepsilon$. These income shocks are independent across voters and represent increases in idiosyncratic productive ability such as career advancement or (un)successful entrepreneurial activity or effects of events such as health problems or job loss. The process that generates the changes in incomes is represented by mean zero distribution F with support $[\underline{\varepsilon}, \bar{\varepsilon}]$ with $\underline{\varepsilon} \leq 0 \leq \bar{\varepsilon}$ and with continuously differentiable strictly log-concave density f . The zero prior expectations of income changes normalize aggregate income growth to zero. While the distribution F is known to voters, only the first moment of distribution G is known to voters.

More generally, G is the distribution of the deterministic part of date 3 income in the electorate: the transition function mapping current incomes to future incomes could more generally be defined as $\phi(y) + \varepsilon$, where ϕ is increasing. We would then define the distribution of current incomes y to satisfy $\phi(y) \sim G$. Concavity of ϕ reduces right-skewness of G . As the focus will be on weakly right-skewed³ income distributions G , assuming that ϕ is an identity function is without loss of generality as long as y is distributed right-skewed enough for $\phi(y)$ to be distributed weakly right-skewed. Also, as noted by Bénabou and Ok (2001), with concave ϕ , we have $\int \phi(y)dG(y) \leq \phi(\int ydG(y))$ and so a mean current income voter expects weakly above mean income in future. Thus, concavity of ϕ decreases the demand for redistribution, and, hence, the demand for redistribution derived for linear ϕ is the upper bound in the class of concave transition functions. Furthermore, with linear ϕ , none of the results here are driven by the mechanism studied by Bénabou and Ok (2001), which requires strict concavity of ϕ .

3.2 Information

Each voter knows her future income shock ε at date 1 before voting. With risk neutrality, the assumption of noiseless information at date 1 is without loss of generality as ε can be interpreted as the expected income change. However, a key assumption in the model is that voters have limited knowledge of their lifetime income when forming their future income and consumption expectations at date 2. This limited knowledge is modelled as imperfect information about ε at date 2. That is, voters' recall is imperfect, and after the election, they lose their date-1-information about their date-3-income changes ε .

This gain and loss of information captures two necessary components of a rational choice theory of self-signaling. First, a voter who knows her (expected) income change knows the true (expected) consequences of her vote. Such a voter thus knows the price of self-signaling and is able to trade material utility to utility from signaling.⁴ The observability of ε at date 1 can be interpreted as voter's estimation of her idiosyncratic mobility prospects based on, say, stories or observations of upward or downward mobility of relatable people or perceptions of

³For continuous unimodal distributions, the exceptions to the rule of right-skew being equivalent to mean exceeding mode and median are rather pathological (MacGillivray, 1981; Groeneveld, 1986). Thus, throughout, this rule is assumed to apply. More formally, right-skew is defined in this text as positive Pearson's first and second coefficients of skewness.

⁴This is also the crux in Caplan's (Caplan, 2001a,b) model of *rational irrationality* where a rational agent knows the price of her irrationality which depends on the true state of the world.

abilities or skills relative to others, existing networks and connections or economic conditions. The unbiasedness of such an estimation reflects the rational voter's best efforts to assess the consequences of her voting. However, an unbiased expectation of ε at date 1 is without loss of generality: Biased information about ε would simply amount to the voter being wrong about the cost of self-signaling and experiencing unexpected consumption utility at date 3.

Second, if the voter remembered her ε at date 2, then voting would not be informative of ε , and there could not be self-signaling. Forgetting ε should be interpreted as the malleability of beliefs and expectations: ε is voter's assessment of her future income expectations at one point in time, and such an assessment is not verifiable at later time. Rather, when forming beliefs at date 2, voting provides a stronger signal than recollections of past expectations of future incomes. The forgetting of ε may also be motivated since belief manipulation has benefits, and beliefs can only be manipulated if there is imperfect recall. Recall of ε could be modelled as probabilistic as well, but as will be shown in Section 3.4, forgetting ε with certainty is without loss of generality.

3.3 Polity and Policy Preferences

At date 1, two distinct policy platforms restricted to linear income tax rates with lump-sum transfers are proposed. One party proposes a low tax rate t ; the other a high tax rate \bar{t} with $t < \bar{t}$. Then all voters vote. The policy receiving the majority of votes is chosen, and the tie-breaking rule is to choose the high tax rate. The policy outcome is revealed immediately after the election, and the policy chosen will be in place at date 3.

Voters may not perceive themselves as pivotal: By voting for $v \in \{t, \bar{t}\}$ a voter induces a perceived probability distribution over election outcomes $((\bar{t}, q(v)), (t, 1 - q(v)))$ with $q(\bar{t}) > q(t)$ such that voting for high tax rate increases the perceived probability of the implementation of the high tax rate. The degenerate specification $q(\bar{t}) = 1, q(t) = 0$ nests the case where the voter perceives herself as decisive.

Since the current income distribution G is unknown to voters, the voters cannot correctly predict the policy outcome in equilibrium. If G was known to voters, they could, knowing G, F , and the equilibrium voting strategies, correctly predict the equilibrium policy almost always implying $q(\bar{t}) = q(t)$.⁵ Thus, the assumption of voters not knowing G merely ensures that the voters do not know the winner of the election.

At date 3, the income changes realize, income is redistributed and all disposable income is consumed. Saving is not allowed and when policy t is implemented, risk neutral voters derive utility linearly from the consumption of their after-tax income:⁶

$$u(t, y, \varepsilon) = (1 - t)(y + \varepsilon) + t\bar{y}, \quad (1)$$

⁵The unique exception occurs when exactly half of the electorate votes for the high tax rate, and the tie-breaking rule is to choose the high tax rate, such that we have $q(\bar{t}) = 1, q(t) = 0$ for voters voting for the high tax rate.

⁶Risk aversion would induce demand for redistribution in the form of insurance against downward mobility. To focus on the effects of self-signaling, risk neutrality is assumed.

where the mean income at date 3

$$\bar{y} := \int_{y \in Y} \left(y + \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \varepsilon dF(\varepsilon) \right) dG(y) = \int_{y \in Y} y dG(y) \quad (2)$$

is the tax base.⁷

I do not model the electoral competition but take the policy platforms as exogenous. This deviates from the applications of the median voter theorem where the offered platforms respond to voters' preferences and are driven toward the median voters' bliss points as a result of Downsian electoral competition (Black, 1948; Downs, 1957). First, in focusing on modelling voting, distinct policy platforms give voters a nontrivial choice. The announced policy platforms create the voting choice set $\{\underline{t}, \bar{t}\}$. Crucially, here the set of policy platforms is also the set of possible signals for the voters. Converging electoral competition would collapse the set of possible signals into a singleton making self-signaling trivially uninformative in equilibrium. Thus, in Downsian equilibrium, votes cannot have value as signals. Second, an election with distinct policy platforms is not an unrealistic assumption as the Downsian electoral convergence is rarely observed in reality. The two tax rates can more generally be interpreted to refer to the general party programmes or ideologies of the two parties with respect to the desired extent of redistribution. In the absence of credible commitment to platforms, these are the platforms voters expect the parties to implement if elected.

3.4 Voting Preferences

Casting vote v induces expected benefit over electoral outcomes

$$U(v, y, \varepsilon) = q(v)u(\bar{t}, y, \varepsilon) + (1 - q(v))u(\underline{t}, y, \varepsilon). \quad (3)$$

At date 2, voters experience a flow anticipatory utility when expecting their future consumption. The total utility of voter (y', ε') who votes for tax rate $v \in \{\underline{t}, \bar{t}\}$ is

$$V(v, y', \varepsilon') = sE[U(v, y', \varepsilon)|v, y'] + \delta U(v, y', \varepsilon'), \quad (4)$$

where $\delta > 0$ weighs consumption and $s \geq 0$ is a preference parameter measuring the value of anticipation. The flow consumption utilities of dates 1 and 2 are dropped as exogenous additive terms. The latter also means that we could, without loss of generality, allow the voter's income to change at the date 2 as well if this income change is not informative of the income change at date 3. This is natural if we interpret date 2 and the anticipatory utility to flow immediately after voting while the future date 3 is further away in time.

The income expectations at date 2 are conditional on the voting choice v and not on the date-3-income change ε . Voting works as a self-signaling device: a part of the motive of the

⁷With exogenous income, taxation does not have distortionary effects and wastage of taxation would simply shift the demand for redistribution downwards.

vote, the expected date-3-income change ε , is not recalled, and expectations are conditioned on the vote v instead. At date 1, voters fully know their types and, hence, their date-3 income. They, however, also know that their date-2 inference and, hence, anticipation is based on their voting choice but not on their current information about expected income change ε . That is, their anticipatory emotions at date 2 depend on their voting.

Voters' recall of ε could be probabilistic as well, but the boundary case of always forgetting ε is without loss of generality. To see this, suppose voters forget ε only with probability κ . That is, with probability $1 - \kappa$, date 1 information is recalled and date 2 anticipation conditions on the date 1 information about ε . The total utility of voter (y', ε') who votes for tax rate v is then

$$V(v, y', \varepsilon') = \kappa s E[U(v, y', \varepsilon) | v, y'] + [(1 - \kappa)s + \delta] U(v, y', \varepsilon'). \quad (5)$$

Thus, the parameters s and δ in (4) capture the possibility of probabilistic recall in the sense that a higher probability of forgetting is equivalent to an increase in s relative to δ .

4 Voting Behavior

Imperfect recall and self-signaling are modelled as a signaling game between the voter's temporal selves, and each voter's voting behavior emerges as an equilibrium of this signaling game. I characterize the voting rules that the equilibria in these within-voter signaling games imply and study how voting is affected by anticipation and the distribution of income shocks.

4.1 Voting Rule

The meanings of votes are determined by who votes for what and who votes for what is, in addition to economic interests, determined by the meanings of votes. Such interaction between voting behavior and meanings of votes is captured by Perfect Bayesian Equilibrium. Off-equilibrium-path beliefs are restricted by the divinity criterion D1 (Banks and Sobel, 1987; Cho and Kreps, 1987). Thus, off-equilibrium-path beliefs are assigned such that the deviator is believed to be the voter with the highest deviation payoff. Intuitively, for pooling low-income voters, a vote for the low tax rate is a very favorable signal of high upward mobility and, for pooling high-income voters, a vote for the high tax rate is a very menacing signal of downward mobility.

Let \mathcal{M}_+ denote the equilibrium meaning of a vote for the low tax rate and \mathcal{M}_- the equilibrium meaning of a vote for the high tax rate. Given these meanings of votes, the expected utility difference between voting for the low tax rate and the high tax rate $V(\underline{t}, y, \varepsilon) - V(\bar{t}, y, \varepsilon)$ is

$$s [(1 - \underline{\tau})\mathcal{M}_+ - (1 - \bar{\tau})\mathcal{M}_- - (\bar{\tau} - \underline{\tau})(\bar{y} - y)] - \delta(\bar{\tau} - \underline{\tau})(\bar{y} - y - \varepsilon), \quad (6)$$

where $\underline{\tau} := q(\underline{t})\bar{t} + (1 - q(\underline{t}))\underline{t}$ is the expected tax rate having voted for low tax rate \underline{t} and $\bar{\tau} := q(\bar{t})\bar{t} + (1 - q(\bar{t}))\underline{t}$ is the expected tax rate having voted for high tax rate \bar{t} .

For each income y , voting behavior is described by a mapping from the set of potential income changes $[\underline{\varepsilon}, \bar{\varepsilon}]$ to the set of possible votes $\{\underline{t}, \bar{t}\}$. The expected utility difference (6) is clearly strictly increasing in ε . Hence, we focus on finding monotonic mappings such that the binary choices can be characterized by thresholds $\varepsilon^*(y)$ such that expected income changes $\varepsilon < \varepsilon^*(y)$ map to \bar{t} and expected income changes $\varepsilon \geq \varepsilon^*(y)$ map to \underline{t} . These voting strategies determine the meanings of votes as $\mathcal{M}_+(\varepsilon') := E[\varepsilon | \varepsilon \geq \varepsilon']$ for $\varepsilon' \in [\underline{\varepsilon}, \bar{\varepsilon})$ and $\mathcal{M}_-(\varepsilon') := E[\varepsilon | \varepsilon < \varepsilon']$ for $\varepsilon' \in (\underline{\varepsilon}, \bar{\varepsilon}]$. To implement D1, define $\mathcal{M}_+(\bar{\varepsilon}) := \bar{\varepsilon}$ and $\mathcal{M}_-(\underline{\varepsilon}) := \underline{\varepsilon}$. These meanings arise as the voters, under uncertainty about their future incomes, form their expectations about future consumption and look back to their voting knowing that it is generally the voters with good prospects that vote for low taxation and vice versa.

The marginal type's expected utility difference between voting for the low tax rate and voting for the high tax rate is

$$\Psi(\varepsilon; y) := s[(1 - \underline{\tau})\mathcal{M}_+(\varepsilon) - (1 - \bar{\tau})\mathcal{M}_-(\varepsilon)] + \delta(\bar{\tau} - \underline{\tau})\varepsilon - (s + \delta)(\bar{\tau} - \underline{\tau})(\bar{y} - y). \quad (7)$$

In an interior equilibrium, the marginal type is indifferent, and for all current incomes y the equilibrium threshold $\varepsilon^*(y)$ is defined by $\Psi(\varepsilon^*(y); y) = 0$.

Assumption 1. $\delta(1 - \underline{\tau}) > (s + \delta)(1 - \bar{\tau})$.

Assume Assumption 1 throughout the text unless otherwise mentioned. Assumption 1 is sufficient to ensure that Ψ is increasing as shown by Lemma 2 and thus that for all y , $\varepsilon^*(y)$ is unique. Assumption 1 sets limits to how strong the signaling concerns are relative to economic concerns. Since $q(\bar{t}) > q(\underline{t})$ we have $\bar{\tau} > \underline{\tau}$ and there are positive values of s such that Assumption 1 holds.

Proposition 1. *For all y , there exists a unique equilibrium characterized by a threshold $\varepsilon^*(y)$ such that voters with $\varepsilon \geq \varepsilon^*(y)$ vote for low tax rate \underline{t} and voters with $\varepsilon < \varepsilon^*(y)$ vote for high tax rate \bar{t} . For voters with income $y > y^H$, $\varepsilon^*(y) := \underline{\varepsilon}$, and they always vote for low taxation; for voters with income $y < y^L$, $\varepsilon^*(y) := \bar{\varepsilon}$, and they always vote for high taxation; for voters with income $y \in [y^L, y^H]$, $\varepsilon^*(y)$ is defined by $\Psi(\varepsilon^*(y); y) = 0$, where*

$$y^L := \bar{y} - \frac{\bar{\varepsilon}}{s + \delta} \left(\delta + s \frac{(1 - \underline{\tau})}{(\bar{\tau} - \underline{\tau})} \right) \quad \text{and} \quad y^H := \bar{y} + \frac{|\underline{\varepsilon}|}{s + \delta} \left(\delta - s \frac{(1 - \bar{\tau})}{(\bar{\tau} - \underline{\tau})} \right), \quad (8)$$

where $y^L < \bar{y}$ always and $\bar{y} < y^H$ if and only if Assumption 1 holds. The threshold income change $\varepsilon^*(y)$ decreases in y on $[y^L, y^H]$ and is constant otherwise.

The voting rule partitions the electorate into two groups: the group of voters with high future incomes and the group of voters with low future incomes. A voter's consumption expectations are based on her vote and on the consumption expectations of voters who vote

like her. If she votes for high taxation, a consistent belief with this action is that she is among those voters who will benefit from high taxation and high transfers, that is, among those voters who expect low incomes. If she votes for low taxation, a belief consistent with this behavior is that she will be among the voters who benefit from low taxation, that is, among those voters with high future incomes and consumption.

Clearly, the voters expecting above-mean incomes at date 1 vote for low taxation as it is in their interest to do so both with respect to consumption and anticipatory utility. The voters expecting below-mean incomes face a trade-off: voting for high redistribution will later on result in the inference of low income and consumption expectations which brings little utility in anticipation. On the other hand, high redistribution is in their economic interest. Voting for low redistribution allows the voters expecting below-mean incomes to infer at date 2 that their increasing incomes will allow high future consumption. However, by voting for low redistribution, these voters contribute to the popularity of a policy that is not in their economic interest. The threshold $\varepsilon^*(y)$ determines the lower bound for the expected income changes for which self-signaling concerns dominate the economic concerns. Whenever the expected income change is below $\varepsilon^*(y)$, the potential economic loss in the form of lower taxation outweighs signaling benefits.

Proposition 1 establishes a correlation between voting for redistribution and perceived mobility prospects: those who vote for low redistribution expect higher future incomes than those who vote for high redistribution. A correlation between self-reported demand for redistribution and mobility prospects has been documented, for instance, by Ravallion and Lokshin (2000), Corneo (2001), Corneo and Grüner (2002), and Alesina and La Ferrara (2005). These studies use a theory of instrumental voting to interpret the causality to run from mobility prospects to demand for redistribution. In contrast, Proposition 1 suggests that there are theoretical reasons to interpret the causality to run in the other direction as well: Voting may influence the perceived mobility prospects.

4.2 Anticipation

With anticipatory utility, the income shock threshold required to vote for low taxation is strictly lower than in the case of no anticipation:

Proposition 2. *Assume $\underline{\varepsilon} < 0 < \bar{\varepsilon}$. For $s = 0$, $\varepsilon^*(y) = \hat{\varepsilon}(y) := \bar{y} - y$, $y^L = \hat{y}^L := \bar{y} - \bar{\varepsilon}$ and $y^H = \hat{y}^H := \bar{y} - \underline{\varepsilon}$. For $s > 0$, $\varepsilon^*(y) \in [\underline{\varepsilon}, \hat{\varepsilon}(y))$, $y^H < \hat{y}^H$ and $y^L < \hat{y}^L$. An increase in s strictly decreases the threshold ε^* for $y \in [y^L, y^H]$ and has no effect otherwise, decreases y^H , and decreases y^L .*

Anticipatory utility creates the demand for self-signaling. Without anticipatory utility, voting preferences coincide with policy preferences. With anticipatory utility, self-signaling shifts the threshold voting rule downward, and some voters who prefer high redistribution vote for low redistribution. The value of self-signaling increases in the value of anticipation, and

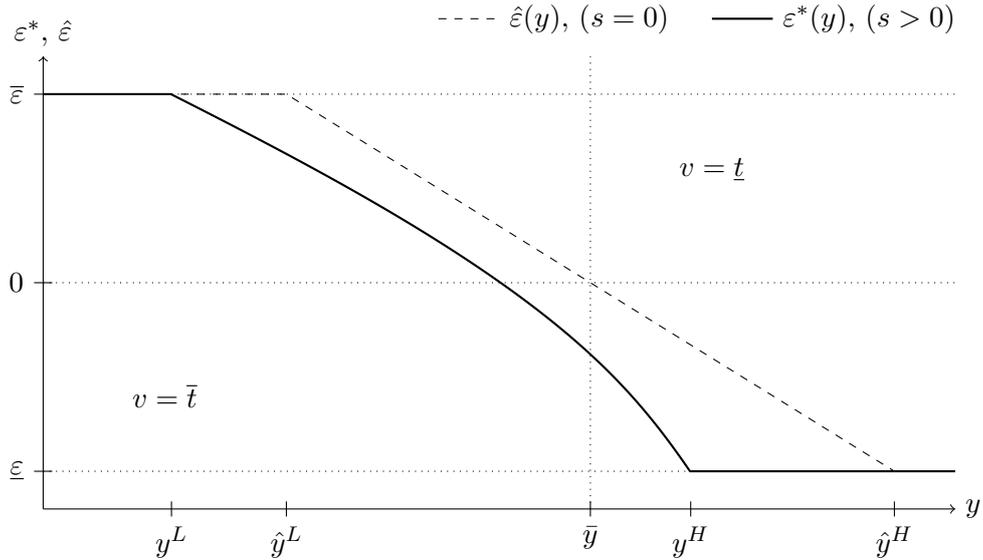


Figure 2: **Voting rule as a function of income and anticipation** *Notes:* The solid line depicts the voting rule ε^* for the case with anticipatory utility $s > 0$. The dashed line depicts the voting rule $\hat{\varepsilon}$ for the case without anticipatory utility $s = 0$. Figure drawn for normal F and G .

higher value of anticipation decreases the threshold income shock that determines the vote. Also, the more a voter values anticipation, the lower are the pooling thresholds. Figure 2 visualises: The dashed line depicts the voting strategy in the absence of anticipatory concerns $\hat{\varepsilon}$. The thick solid line depicts the voting strategy in the presence of anticipation ε^* . The voters $(y, \varepsilon) \in \{(y, \varepsilon) : \varepsilon^*(y) < \varepsilon < \hat{\varepsilon}(y)\}$ between the two lines are motivated by self-signaling concerns to change their vote or party affiliation from high tax rate to low tax rate. Thus, these voters choose apparently or economically dominated actions and thus seem to vote against their best interest.⁸

4.3 Income Risk

The economic environment that defines a voter's economic opportunities, mobility, and degree of risk is represented by the distribution of income shocks F . This distribution does not need to represent the actual possibilities for income mobility but may also be more loosely interpreted as measuring the perceived possibilities for income mobility. Thus, the measure of economic opportunities F can, for instance, be interpreted as being an outcome of learning (Piketty, 1995), individual histories (Alesina and Giuliano, 2011) or perceptions of economic fortunes of others (Hirschman and Rothschild, 1973). A comparative static effect of interest is that of an increase in income risk. Top incomes have increased (Atkinson et al., 2011; Piketty and Saez, 2014) whereas the risks at the low end of income distribution have increased (Kalleberg, 2003).

To study the changes in income risk, scale the date-3-income shock ε by factor $\sigma \in \mathbb{R}_{>0}$

⁸Imperfect recall is necessary for the voting rule to depend on anticipation. With perfect recall, the relative payoff from voting for the low tax rate rather than the high tax rate is $-(s + \delta)(\bar{\tau} - \underline{\tau})(\bar{y} - y - \varepsilon)$, producing the exact same voting rule $\hat{\varepsilon}$ as the case without anticipation.

such that the income change $\sigma\varepsilon$ is generated by

$$\sigma\varepsilon \sim F\left(\frac{\varepsilon}{\sigma}\right), \quad \sigma\varepsilon \in [\sigma\underline{\varepsilon}, \sigma\bar{\varepsilon}].$$

An increase in income risk is then defined as an increase in σ . Restricting the changes in the distribution to within the scale family does not, in any way, restrict the original distribution. Proposition 3 states the effect of an increase in risk on voting rule.⁹

Proposition 3. *Assume $\underline{\varepsilon} < 0 < \bar{\varepsilon}$. If and only if $s > 0$, an increase in income risk σ strictly decreases the threshold $\varepsilon^*(y)$ for $y \in [y^L, y^H]$ and has no effect otherwise. An increase in income risk decreases y^L , and increases y^H for all $s \geq 0$.*

The effect of income risk on the voting strategies for a discrete change in σ is depicted in Figure 3. Proposition 3 can be understood as two separate effects: a decrease in \mathcal{M}_- and an increase in \mathcal{M}_+ . First, as the left tail of future income distribution grows longer and fatter, the prospects of downward mobility become more menacing. Anxiety about these increased risks increases the value of expressing good prospects and makes self-signaling more attractive. The deteriorating prospects of downward mobility make the voters want to look away and focus more on the prospects of upward mobility. Identifying with the voters who prefer the low tax rate achieves this.

This mechanism provides an explanation for the *negative exposure effect* first empirically documented by Luttmer (2001). He finds that an increase in the welfare reciprocity rate in a survey respondent's area decreases her support for welfare spending. Sands (2017) randomizes passersby's exposure to poverty before asking for support for a tax for the wealthy and identifies a causal link between such exposure and decreased support for redistribution. Proposition 3 suggests that the perceived increased likelihood of low incomes motivates identification and behavior consistent with the belief of not needing to rely on welfare oneself. A stronger identification with the party that supports low welfare and lower demand for redistribution results.

The effect of increasing lower tail risk also proposes an explanation for the diverging policy preferences of the middle class and the poor. This has previously been explained as a growing social distance between the middle class and the poor (Lupu and Pontusson, 2011). This social affinity hypothesis suggests that the middle class is willing to support redistribution when they are socially close to the beneficiaries of redistribution. An increasing social distance between the middle class and the poor then decreases the support for redistribution in the middle class. The mechanism here does not rely on other-regarding preferences but is an

⁹The effect runs via the truncated expectations \mathcal{M}_- and \mathcal{M}_+ and either an increase in \mathcal{M}_+ or a decrease in \mathcal{M}_- is sufficient to decrease ε^* . Thus, increasing the range of the support of the date-3-income changes is not a necessary condition. Fattening of the tails, that is an increase in probabilities $F(\mathcal{M}_-)$ and $1 - F(\mathcal{M}_+)$ without changes in the support of F would yield similar result. Also note how, by Proposition 1, with no income mobility, without risk, that is, for $\underline{\varepsilon} = \bar{\varepsilon} = 0$, we have $\bar{y} = y^L = y^H$. That is, voters with above mean income always vote for low taxation, and voters with below mean income always vote for high taxation, reproducing the static benchmark model.

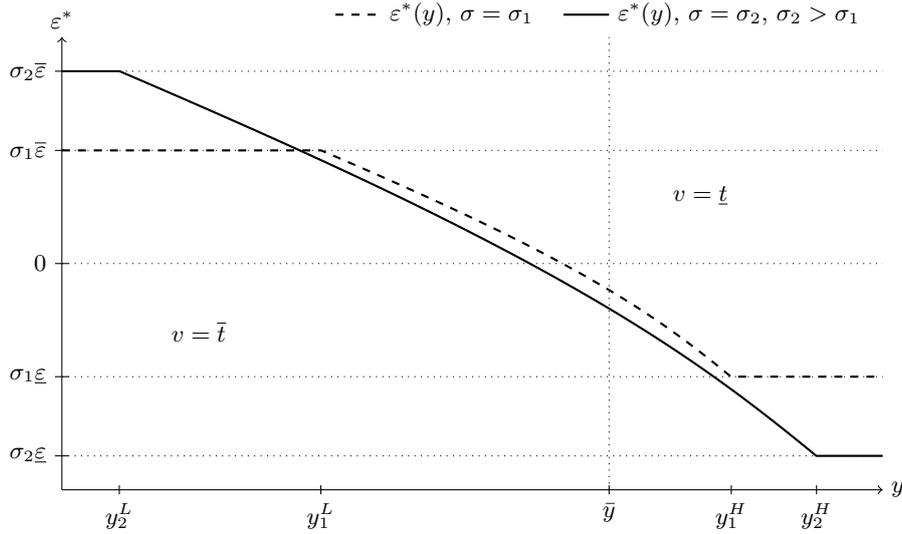


Figure 3: **Voting rule as a function of income and risk** *Notes:* The solid line depicts the voting rule ε^* for the degree of risk σ_2 . The dashed line depicts the voting rule ε^* for the degree of risk $\sigma_1 < \sigma_2$. Figure drawn for normal F and G .

affective reaction of “wanting to look away” as the economic outcomes of the poor deteriorate. By supporting low taxation, the middle class voter identifies herself with the more desirable economic opportunities of upward mobility and distances herself from the outcomes of the poor.

Second, as the right tail of future income distribution grows longer and fatter, the prospects of upward mobility become more improve, and the anticipation of these possibilities becomes more attractive making signaling good prospects more valuable. When the expected incomes in the group of voters who vote for the low tax rate increase, identification with this group becomes more attractive. Such a mechanism has gained evidence in laboratory: Mijović-Prelec and Prelec (2010) and Coutts (2019) find that the larger the prizes rewarded if an event occurs, the more optimistic subjects are about the occurrence of the event. In the experiment of Mijović-Prelec and Prelec (2010), such optimism was mediated by increased attempt to self-signal.

5 Demand for Redistribution

I now characterize the demand for redistribution of a voter with income y and income mobility distribution F . This measure of demand for redistribution can be interpreted either as the ex ante probability of a voter voting for high redistribution before the realization of income shock ε or, with an infinite number of voters, as the share of voters with income y that vote for high redistribution.

The voting of a voter with income y emerges as the equilibrium-path behavior of the within voter signaling game studied in Section 4. A voter with income y votes for high taxation if and only if an income expectation below the threshold $\varepsilon^*(y)$ realizes. Thus, the probability

that a voter with income y votes for high tax rate is $Pr[\varepsilon < \varepsilon^*(y)] = F(\varepsilon^*(y))$. Define the demand for redistribution h for a voter with income y as

$$h(y) := F(\varepsilon^*(y)). \quad (9)$$

For voters with $y \leq y^L$, $F(\varepsilon^*(y)) = F(\bar{\varepsilon}) = 1$, for voters with $y \geq y^H$, $F(\varepsilon^*(y)) = F(\underline{\varepsilon}) = 0$, and for voters with $y \in (y^L, y^H)$, $F(\varepsilon^*(y)) \in (0, 1)$. In the absence of income mobility, $\underline{\varepsilon} = \bar{\varepsilon} = 0$, we have $y^L = y^H = \bar{y}$ and the demand for redistribution is 1 for $y < \bar{y}$ and 0 for $y > \bar{y}$. Income mobility thus decreases the correlation between current income and demand for redistribution.

As F is increasing, all comparative static effects of demand for redistribution that keep F fixed inherit their signs from the corresponding effects on the voting rules. Thus, voter's demand for redistribution is decreasing in current income y by Proposition 1 and in the value of anticipation s by Proposition 2.

In contrast, in case of changes in the distribution of income shocks F , there are also composition effects. The vote share that the high tax rate receives depends, in addition to the distribution of current incomes, on the distribution of income prospects among the voters. As the distribution of date-3-income shocks changes, the realized income prospects in the electorate change as well shifting the demand for redistribution also in the absence of behavioral effects. Looking at Figure 3, when we scale F with σ_2 instead of $\sigma_1 < \sigma_2$, there are new types of voters with $\varepsilon \in [\sigma_1 \underline{\varepsilon}, \sigma_2 \underline{\varepsilon}]$ and $\varepsilon \in [\sigma_1 \bar{\varepsilon}, \sigma_2 \bar{\varepsilon}]$ whose contributions to the demand for redistribution we have to take into account.

To take this compositional change in the electorate's expected income dynamics into account, let

$$h(y, \sigma) := F\left(\frac{\varepsilon^*(y)}{\sigma}\right) \quad (10)$$

denote the demand for redistribution of a voter with income y given the scaling σ of income shocks. The total effect of a change in income risk on the demand for redistribution consists of a mechanical effect due to a change in the composition of future incomes and a behavioral effect of each voter type changing her voting behavior. This decomposition can be written as:

$$\frac{\partial}{\partial \sigma} h(y, \sigma) = \frac{d}{d\sigma} F\left(\frac{\varepsilon^*}{\sigma}\right) = \underbrace{f\left(\frac{\varepsilon^*}{\sigma}\right) \frac{1}{\sigma} \frac{\partial \varepsilon^*}{\partial \sigma}}_{\text{Self-signaling effect}} - \underbrace{f\left(\frac{\varepsilon^*}{\sigma}\right) \frac{1}{\sigma} \frac{\varepsilon^*}{\sigma}}_{\text{Composition effect}}. \quad (11)$$

Clearly, the self-signaling effect in (11) is always negative as $\partial \varepsilon^* / \partial \sigma < 0$ by Proposition 3. However, as the threshold ε^* may be positive or negative, the composition effect in (11) maybe be negative or positive. Specifically, as ε^* is large and positive for low-income voters such that (11) is unambiguously negative for low-income voters. However, for high-income voters ε^* is small and negative, and thus the sign of (11) is ambiguous.

Formally, evaluating, without loss of generality, (11) at $\sigma = 1$, and noting that, $\Psi_\sigma + \Psi'\varepsilon^* = \Psi(\varepsilon^*(y), \bar{y}) = \Psi_y(\bar{y} - y)$, where details of the first equality are shown in Appendix and the second inequality follows from the definition of $\varepsilon^*(y)$, we have

$$\begin{aligned} f(\varepsilon^*) \left(\frac{\partial \varepsilon^*}{\partial \sigma} - \varepsilon^* \right) &= f(\varepsilon^*) \left(-\frac{\Psi_\sigma + \Psi'\varepsilon^*}{\Psi'} \right) = -f(\varepsilon^*) \frac{\Psi(\varepsilon^*; \bar{y})}{\Psi'} \\ &= -f(\varepsilon^*(y)) \frac{\Psi_y}{\Psi'} (\bar{y} - y) = (\bar{y} - y) \frac{d}{dy} F(\varepsilon^*(y)). \end{aligned} \quad (12)$$

Thus, the effect is proportional to the income effect, and the sign of the effect changes at $y = \bar{y}$. Proposition 4 states the effect of an increase in income risk on the individual demand for redistribution.

Proposition 4. *The demand for redistribution $h(y)$ strictly decreases in income risk for all $y \in [y^L, \bar{y})$ and strictly increases in income risk for all $y \in (\bar{y}, y^H]$ and has no effect otherwise.¹⁰*

For voters with below-mean current income, the self-signaling effect dominates the composition effect, and their demand for redistribution decreases in income risk. For voters with above-mean current income, the new real prospects of downward mobility dominate the self-signaling effect, and their demand for redistribution increases in income risk.¹¹

It is noteworthy that an increase in income risk has an effect on the demand for redistribution also in the absence of anticipatory concerns. The composition effect alone changes the mobility prospects such that the demand changes by $-(\bar{y} - y)f(\bar{y} - y)$. Increased mobility makes it relatively more likely for below-mean income voters to expect above-mean incomes and above-mean income voters to expect below-mean incomes. Intuitively, keeping the threshold ε^* fixed, as σ approaches infinity, the demand for redistribution approaches $F(0)$ for all incomes. When income mobility is high enough, only future income matters when voting for redistribution. The nature of the composition effect is thus not to increase or decrease the demand for distribution but to reduce the correlation between current income and voting.

6 Aggregate Demand for Redistribution

Aggregate demand for redistribution is defined as the share of voters demanding high redistribution. Equivalently, this is the probability that a randomly drawn voter (y, ε) votes for high redistribution. Let $H(\tilde{\varepsilon}) := \int_{y \in Y} F(\tilde{\varepsilon}(y)) dG(y)$ be the fraction of voters voting for high

¹⁰Note that $\bar{y} < y^H$, i.e. $(\bar{y}, y^H]$ is non-empty if and only if Assumption 1 holds.

¹¹Proposition 4 is similar to Proposition 6 in Adriani and Sonderegger (2019). Here the cost of sending the favorable signal for the mean type $\varepsilon = 0$, $\Psi_y(\bar{y} - y)$, is endogenized in terms of current income. For voters with current income above mean, sending the favorable signal is intrinsically beneficial for the mean type $\varepsilon = 0$ and signaling demand effect is negative. For voters with current income below mean, sending the favorable signal is costly for the mean type $\varepsilon = 0$ and signaling demand effect is positive.

redistribution given voting rule $\hat{\varepsilon}$. The equilibrium aggregate demand for redistribution is

$$H := \int_{y \in Y} h(y) dG(y) = \int_{y \in Y} F(\varepsilon^*(y)) dG(y) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} G(y^*(\varepsilon)) dF(\varepsilon), \quad (13)$$

where $y^* = (\varepsilon^*)^{-1}$ is a threshold on current income such that given an income shock ε voters with current income $y < y^*(\varepsilon)$ vote for high tax rate and otherwise they vote for low tax rate.¹² Thus for each ε , $G(y^*(\varepsilon))$ is the fraction of voters voting for high redistribution. In the absence of income mobility, a voter's demand for redistribution $h(y)$ is either 0 or 1, and we have $y^*(\varepsilon) = \bar{y}$. The aggregate demand for redistribution is thus $G(\bar{y})$. Hence, the familiar results apply: if the income distribution is right-skewed, median income is less than mean income and $G(\bar{y}) > \frac{1}{2}$, that is, more than half of the voters vote for high tax rate.

Without anticipation, $s = 0$, and with income mobility, we have $y^*(\varepsilon) = \bar{y} - \varepsilon$ but the income threshold of voting for low taxation y^* still averages to mean income \bar{y} in the electorate. The effect of income mobility in the absence of anticipation on demand for redistribution is discussed below. With anticipation, $s > 0$, $y^*(\varepsilon) < \bar{y} - \varepsilon$, for all ε , that is, the income threshold for voting for low tax rate shifts downwards. This decreases the aggregate demand for redistribution. As the demand for redistribution decreases in the value of anticipation s for all income levels, aggregate demand for redistribution decreases in s .

Voting preferences coincide with the policy preferences only in the absence of anticipation $s = 0$. With anticipation, $s > 0$, a fraction $H(\hat{\varepsilon})$ of the electorate prefers the high tax rate but only a fraction $H(\varepsilon^*) < H(\hat{\varepsilon})$ of the electorate votes for the high tax rate. With self-signaling concerns, voting does not aggregate policy preferences. Votes favoring redistribution fall short of the aggregate policy preference for redistribution, and this limits the demand for redistribution. Moreover, if the coalition of voters $\{(y, \varepsilon) : \varepsilon^*(y) < \varepsilon < \hat{\varepsilon}(y)\}$ is decisive, that is, if $H(\varepsilon^*) < \frac{1}{2} < H(\hat{\varepsilon})$, then the policy preferred by minority gains the majority of votes.

6.1 Income Risk

The effect of income risk on the demand for redistribution depends on income: some voters decrease while others increase their demand for redistribution. Hence, the interesting question is whether the sign of the aggregate demand effect can be determined.

Note first that the effect of an increase in income risk on aggregate demand for redistribution can be written as (see Lemma 4)

$$\frac{d}{d\sigma} \int_{y \in Y} h(y, \sigma) dG(y) = \int_{y^L}^{y^H} \frac{d}{d\sigma} h(y, \sigma) dG(y). \quad (14)$$

The aggregate demand effect is thus the mean effect among the voters whose voting strategy is separating.

¹²More formally, since ε^* , by its definition, is not one-to-one at boundaries $\underline{\varepsilon}$ and $\bar{\varepsilon}$, define y^* as the inverse of ε^* on $\varepsilon \in (\underline{\varepsilon}, \bar{\varepsilon})$, $y^*(\bar{\varepsilon}) = y^L$, and $y^*(\underline{\varepsilon}) = y^H$.

Note second that a right-skewed current income distribution G , where right skew is defined as mean exceeding median, weighs the negative demand effects of below-mean-income voters heavier than symmetric distributions. Thus, an aggregate demand effect for a symmetric distribution gives us an upper bound of the aggregate demand effect in the class of weakly right-skewed income distributions that are formed from the original distribution by reallocating probability mass from above the mean to below the mean.

Note third that if the distribution of current incomes is right-skewed, an increase in income risk may change aggregate demand solely due to the composition effect. To see this, let $s = 0$ and consider some small amount of mobility $|\varepsilon| = \bar{\varepsilon} = \varepsilon' > 0$. Now, if G is unimodal and right-skewed such that its mode is strictly smaller than its mean, then there is a ε' -neighborhood around its mean such that on this neighborhood, the density of G is strictly decreasing. This means that G is strictly concave on ε' -neighborhood around \bar{y} and thus by Jensen's inequality

$$H = \int_{-\varepsilon'}^{\varepsilon'} G(\bar{y} - \varepsilon) dF(\varepsilon) < G \left(\int_{-\varepsilon'}^{\varepsilon'} (\bar{y} - \varepsilon) dF(\varepsilon) \right) = G(\bar{y}), \quad (15)$$

where $G(\bar{y})$, as discussed above, is the demand for redistribution in the absence of income mobility. Hence, income mobility alone may decrease the aggregate demand for redistribution when the distribution of current incomes is right-skewed.¹³ Intuitively, when there are more voters below mean income than above mean income, there are more voters to whom income mobility makes above mean future incomes possible than there are voters to whom income mobility makes below mean future incomes possible. If G is uniform, however, then G is affine and we have

$$\int_{-\varepsilon'}^{\varepsilon'} G(\bar{y} - \varepsilon) dF(\varepsilon) = G \left(\int_{-\varepsilon'}^{\varepsilon'} (\bar{y} - \varepsilon) dF(\varepsilon) \right) = G(\bar{y}), \quad (16)$$

and income risk has no effect on aggregate demand for redistribution without anticipatory concerns. Proposition 5 shows that with anticipatory concerns, aggregate demand for redistribution is nevertheless decreasing in income risk even when G is uniform.

Proposition 5. *Let G be uniform. Then if and only if $s > 0$ the aggregate demand for redistribution H decreases in income risk σ .*

If G weighs the negative demand effects relatively more than a uniform distribution with the same support, the aggregate demand for redistribution is decreasing in income risk for such G as well. Thus, a corollary to Proposition 5 is that the aggregate demand effect is negative for all G that are more right-skewed than uniform in the sense that, compared to uniform with the same support, they have weakly higher density for all $y < \bar{y}$ and weakly lower density for all $y > \bar{y}$. This is a large class of right-skewed distributions and, thus, the result

¹³Note that the aggregate composition effect does not need to be monotone in ε' : as noted above, $h \rightarrow F(0)$ and thus $H \rightarrow F(0)$ as $\varepsilon' \rightarrow \infty$. If $F(0) > G(\bar{y})$, then the aggregate demand for redistribution at the limit of infinite income mobility is larger than in the absence of income mobility.

of Proposition 5 is fairly general in the context of usually right-skewed income distributions.

Income mobility has no effect on the aggregate demand for redistribution in the absence of anticipation also if F and G are symmetric, that is, the equality in (16) holds if both F and G are symmetric (see Lemma 5). Intuitively, when F and G are symmetric and there are no anticipatory concerns, there is an equal measure of voters (\hat{y}^L, \bar{y}) on whose voting an increase in income risk has a negative effect as there are voters (\bar{y}, \hat{y}^H) on whose voting an increase in income risk has a positive effect. As s increases, however, \hat{y}^L and \hat{y}^H decrease, and thus, the interval of voters with negative demand effects increases while the interval of voters with positive demand effects decreases. As s increases, the interval (y^L, y^H) where voters separate in their voting strategies moves left. The aggregate demand effect of income risk (14), on the other hand, is the mean effect among the voters whose voting strategy is separating. Proposition 6 considers a case where F and G are symmetric.

Proposition 6. *Let F be uniform and G symmetric. Then if and only if $s > 0$ the aggregate demand for redistribution H decreases in income risk σ .*

Again, since right-skewed current income distributions G weigh the negative demand effects relatively more than symmetric income distributions, if aggregate demand for redistribution is decreasing in income risk for symmetric G then it is decreasing for right-skewed G as well.

Proposition 4 concludes that as income risk, that is, the dispersion of F , increases, the demand for redistribution may either increase or decrease depending on income: for high-income voters to whom sending the favorable signal of high expected future income is, on average, not costly, the demand for these favorable signals decreases; for low-income voters to whom sending the favorable signals is, on average, costly, the demand for these favorable signals increases. Alternatively, the ambiguity of the demand effects can be stated in terms of signaling strategies: for voters with $\varepsilon^* < \varepsilon^*(\bar{y})$ the demand for favorable signals decreases as income risk increases and for voters with $\varepsilon^* > \varepsilon^*(\bar{y})$ the demand for favorable signals increases as income risk increases.

Propositions 5 and 6, however, show that for a large class of distributions F and G , the aggregate demand effect in a population with heterogeneity in signaling strategy, that is, in threshold ε^* , is unambiguous. This finding adds to the results of Adriani and Sonderegger (2019) who study these so called distributional comparative statics in a belief-based model of esteem. They study how the demand for favorable signals responds to an increase in the dispersion of sender type distribution in a population of senders that is heterogeneous in their private type. However, often the senders are heterogeneous in an observable dimension such as direct signaling cost as well. Here such a dimension is the current income y . This makes the population of senders heterogeneous in their signaling strategies ε^* such that there are both negative and positive signaling demand effects in the population. Propositions 5 and 6 show how the self-signaling effects can dominate the composition effects such that the aggregate effect is unambiguous. Intuitively, the decomposition effects tend to cancel out, resulting in

no changes in the aggregate demand for redistribution, while the self-signaling effects tend to increase the weight on negative income risk effects and decrease the weight on positive income risk effects. This results in a decrease in the aggregate demand for redistribution.

Thus, when voters express their desires about their future incomes when voting, an increase in income risk can decrease the aggregate demand for redistribution. This result contributes to the literature that aims to understand why the positive relationship between income inequality and redistribution emerging from the simple models of redistribution (Meltzer and Richard, 1981) is often not observed in the data. As income inequality has increased (Atkinson et al., 2011; Piketty and Saez, 2014), the demand for redistribution has not followed or has decreased both across countries (Alesina and Glaeser, 2004; Lindert, 2004) and within countries (Georgiadis and Manning, 2012; Rehm et al., 2012; Korpi and Palme, 2003; Cavallé and Trump, 2015).

7 Policy Traps

The voters $(y, \varepsilon) \in \{(y, \varepsilon) : \varepsilon^*(y) < \varepsilon < \hat{\varepsilon}(y)\}$ vote for a policy that does not maximize their consumption. However, in assessing the welfare consequences of voters' self-signaling, we should take the utility from anticipation into account. Are the voters $(y, \varepsilon) \in \{(y, \varepsilon) : \varepsilon^*(y) < \varepsilon < \hat{\varepsilon}(y)\}$ better off with expectations of high income despite incurring losses in consumption? This depends on the policy outcome, which depends on G and F and is not determined here. However, it is clear that if the policy outcome is still high taxation, that is, if $H(\varepsilon^*) \geq \frac{1}{2}$, then these voters are better off as they do not incur the economic cost of their self-signaling. Thus, consider the case $H(\varepsilon^*) < \frac{1}{2}$ where the policy outcome is low taxation. Would these voters have been better off voting for the high tax rate if their voting for the high tax rate had induced the high-tax policy?

Proposition 7. *If and only if $q(\bar{t}) - q(\underline{t}) < 1$, there exists a threshold $\varepsilon^{**} \in (\varepsilon^*, \hat{\varepsilon})$ such that the coalition of voters $\{(y, \varepsilon) : \varepsilon^*(y) < \varepsilon < \varepsilon^{**}(y)\}$, if decisive, that is, if $H(\varepsilon^*) < \frac{1}{2} < H(\varepsilon^{**})$, would be better off voting for high tax rate but votes for low tax rate. For all y , $\varepsilon^{**}(y) - \varepsilon^*(y)$ is increasing in the value of anticipation s and income risk σ .*

If voters perceive their pivotality $q(\bar{t}) - q(\underline{t})$ to be low, there is a coalition of voters who vote for the low tax rate even if their anticipatory utility falls short of their economic loss that the low tax policy induces. If this coalition of voters is decisive, that is if $H(\varepsilon^*) < \frac{1}{2} < H(\varepsilon^{**})$, then these voters would be able to improve their payoffs with a collective action of voting for the high tax rate. However, they perceive the probability that their individual votes have an effect on policy to be small and thus face a collective action problem in a policy trap where it is individually rational to vote for the low tax rate. Thus, if voters do not perceive themselves as pivotal, a policy trap where a decisive coalition of voters votes against its best interest, may arise.

Suppose on contrary that the voters perceive themselves to be pivotal such that $\varepsilon^{**} = \varepsilon^*$ and that $H(\varepsilon^*) < \frac{1}{2} < H(\hat{\varepsilon})$, such that, the coalition of voters $\{(y, \varepsilon) : \varepsilon^*(y) < \varepsilon < \hat{\varepsilon}(y)\}$ is decisive. While this coalition of voters prefers the high tax rate to be implemented it votes for the low tax rate. However, these voters would not be better off if they voted for the low tax rate and thus are not voting against their interest. Hence, the role of pivotality is crucial, and there may exist a decisive coalition of voters voting against their best interest only if voters do not perceive their votes to be fully pivotal.¹⁴

Thus, when voters perceive themselves as pivotal, a policy preferred by the minority may win the election, but the decisive voters are making the best choices for themselves. Any policy intervention aiming to influence their voting may decrease their welfare. However, if the voters do not perceive themselves as fully pivotal, it is possible that the decisive voters are trapped in a collective action problem. The latter is also more likely to happen when the perceived pivotality $q(\bar{t}) - q(\underline{t})$ is lower, when the weight on the meanings of votes s is larger, and when income risk σ is perceived larger. Again, since a change in σ also changes the fraction of voters within interval $(\varepsilon^*, \varepsilon^{**})$, an increase in $\varepsilon^{**} - \varepsilon^*$ does not imply an increase in the fraction of voters within $(\varepsilon^*, \varepsilon^{**})$. However, changes in the perceptions about σ do not have composition effects.

Political campaigning that emphasizes the policy and the importance of voting in determining the policy makes policy traps less likely. Political campaigning that emphasizes the meanings of votes and political identity by making s larger and that frames the voters of different parties in extreme ways, increasing the perceived income risk σ makes policy traps more likely. A policy maker aiming to alleviate the possibility of policy traps may thus encourage political campaigning focused on policy issues and discourage political campaigning focusing on the identity of voters.

8 Conclusion

There is an extensive literature aiming to understand what limits the demand for redistribution in democracies. However, this literature's focus is on asking why aggregate policy preferences do not support redistribution. Instead, I focus on voting: Why does the electorate not vote for redistribution? I add two components, anticipatory utility and imperfect recall, into a model of voting for redistribution under uncertainty of future income. With these two components, votes gain meanings that the voters care about: Voting for a low tax rate identifies the voter with the voters who benefit from low taxes and is thus consistent with the desirable expectation of a high future income and consumption. At the same time, policy concerns in the form of consumption of after-tax income restrict self-signaling. This dual purpose of voting makes voting and policy preferences diverge such that voting does

¹⁴The potential for policy traps that arises when voters do not perceive themselves pivotal is central in the literature on expressive voting (see e.g., Tullock (1971), Glazer (1992), Brennan and Lomasky (1993), and Hillman (2010)).

not aggregate policy preferences. Not all voters who prefer high redistribution vote for high redistribution. Votes favoring redistribution fall short of the aggregate policy preference for redistribution. This limits the demand for redistribution.

The model of voting for redistribution as self-signaling provides a new view to understand how prospects of upward mobility, biased beliefs, and social identification interact with voting. In contrast to models of Minozzi (2013), Bénabou and Tirole (2006a) and Shayo (2009) where belief formation or identification occur prior to voting, viewing voting as self-signaling, belief formation or identification occurs *by* voting. To what degree our actions are biased by our beliefs and to what degree our beliefs are biased by our actions is unclear. This, however, has consequences for how voters react to information. Informing voters with biased beliefs may decrease their biases but they may still want to hold on to them or regenerate them by self-signaling with their voting.

This framework of voting as self-signaling can be interpreted as a hybrid model of instrumental and expressive voting. A characteristic of models of expressive voting has been their ad-hoc -style nonconsequentialist operationalisations of expressive benefits in the form of various exogenous utility flows that voting might bring. Many (e.g., Green and Shapiro (1994), Mueller et al. (2003) (p. 329) and Bellani and Ursprung (2019)) have argued that such models of expressive benefits do not constitute a proper theory and are rather rationalizations than explanations of behavior. On the one hand, such a model lacks predictive content. In particular, the exogeneity of expressive benefits robs the model of the possibilities of studying how value of expressive voting changes in changing political and economic environments. On the other hand, looking for the “deeper reasons” (Dowding, 2005) of expressive benefits brings the idea of expressive voting closer to the core of rational choice theory where benefits of actions flow from their consequences.

Self-signaling provides a way of endogenizing expressive benefits that arise from identification with parties, policies, coalitions, or fellow voters: Voting for certain policies is compatible with the desirable belief of being the kind of voter who benefits from these policies. The comparative statics results demonstrate how such endogenous expressive benefits react to changes in the economic environment and so mediate these changes to voting behavior. While in this paper, voters wish to express prospects of upward mobility, the framework of voting as self-signaling is more general. Different issues in political agenda may give rise to various expressive concerns. If certain policies are viewed, say, as more charitable or altruistic, more moral, more patriotic, or more nationalistic, than others, then voters may express these preferences, qualities, or values in their voting as well. This may lead to policy traps and inefficient voting outcomes. Such adverse outcomes may be more likely when political campaigning focuses on voters’ identities rather than policies and frames the supporters of different parties in extreme ways.

References

- Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2015). Democracy, redistribution, and inequality. In *Handbook of Income Distribution*, volume 2, pages 1885–1966. Elsevier.
- Adriani, F. and Sonderegger, S. (2019). A theory of esteem based peer pressure. *Games and Economic Behavior*, 115:314–335.
- Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.
- Alesina, A. and Giuliano, P. (2011). Preferences for redistribution. In *Handbook of Social Economics*, volume 1, pages 93–131. Elsevier.
- Alesina, A., Glaeser, E., and Sacerdote, B. (2001). Why doesn't the US have a European-style welfare system? *Brookings Paper on Economics Activity*, 2:187–278.
- Alesina, A. and Glaeser, E. L. (2004). *Fighting poverty in the US and Europe: A world of difference*. Oxford University Press, New York.
- Alesina, A. and La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics*, 89(5-6):897–931.
- Alesina, A., Stantcheva, S., and Teso, E. (2018). Intergenerational mobility and preferences for redistribution. *American Economic Review*, 108(2):521–54.
- Atkinson, A. B., Piketty, T., and Saez, E. (2011). Top incomes in the long run of history. *Journal of Economic Literature*, 49(1):3–71.
- Banks, J. S. and Sobel, J. (1987). Equilibrium selection in signaling games. *Econometrica*, 55(3):647–661.
- Bellani, L. and Ursprung, H. (2019). The political economy of redistribution policy. In Congleton, R. D., Grofman, B., and Voigt, S., editors, *The Oxford Handbook of Public Choice, Volume 2*, pages 520–541. Oxford University Press, New York.
- Bénabou, R. (2008). Ideology. *Journal of the European Economic Association*, 6(2-3):321–352.
- Bénabou, R. (2013). Groupthink: Collective delusions in organizations and markets. *Review of Economic Studies*, 80(2):429–462.
- Bénabou, R. and Ok, E. A. (2001). Social mobility and the demand for redistribution: the POUM hypothesis. *The Quarterly Journal of Economics*, 116(2):447–487.
- Bénabou, R. and Tirole, J. (2004). Willpower and personal rules. *Journal of Political Economy*, 112(4):848–886.
- Bénabou, R. and Tirole, J. (2006a). Belief in a just world and redistributive politics. *The Quarterly Journal of Economics*, 121(2):699–746.

- Bénabou, R. and Tirole, J. (2006b). Incentives and prosocial behavior. *American Economic Review*, 96(5):1652–1678.
- Bénabou, R. and Tirole, J. (2009). Over my dead body: Bargaining and the price of dignity. *American Economic Review*, 99(2):459–65.
- Bénabou, R. and Tirole, J. (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics*, 126(2):805–855.
- Bentham, J. (1789). *The collected works of Jeremy Bentham: An introduction to the principles of morals and legislation*. (1996). Clarendon Press.
- Bernheim, B. D. and Thomadsen, R. (2005). Memory and anticipation. *The Economic Journal*, 115(503):271–304.
- Black, D. (1948). On the rationale of group decision-making. *Journal of Political Economy*, 56(1):23–34.
- Bodner, R. and Prelec, D. (2003). Self-signaling and diagnostic utility in everyday decision making. In Brocas, I. and Carrillo, J. D., editors, *The Psychology of Economic Decisions. 1: Rationality and Well-being*, pages 520–541. Oxford University Press, New York.
- Borck, R. (2007). Voting, inequality and redistribution. *Journal of Economic Surveys*, 21(1):90–109.
- Brennan, G. and Brooks, M. (2007). Esteem-based contributions and optimality in public goods supply. *Public Choice*, 130:457–470.
- Brennan, G. and Brooks, M. (2013). Expressive voting. In *The Elgar Companion to Public Choice, Second Edition*. Edward Elgar Publishing.
- Brennan, G. and Hamlin, A. (1998). Expressive voting and electoral equilibrium. *Public Choice*, 95(1-2):149–175.
- Brennan, G. and Lomasky, L. (1993). *Democracy and decision: The pure theory of electoral preference*. Cambridge University Press, Cambridge.
- Brennan, G. and Pettit, P. (1993). Hands invisible and intangible. *Synthese*, 94:191–225.
- Brennan, G. and Pettit, P. (2004). *The economy of esteem: An essay on civil and political society*. OUP Oxford.
- Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, 95(4):1092–1118.
- Buchanan, J. M. (1954). Individual choice in voting and the market. *Journal of Political Economy*, 62(4):334–343.
- Caplan, B. (2001a). Rational ignorance versus rational irrationality. *Kyklos*, 54(1):3–26.

- Caplan, B. (2001b). Rational irrationality and the microfoundations of political failure. *Public Choice*, 107(3):311–331.
- Caplin, A. and Leahy, J. (2001). Psychological expected utility theory and anticipatory feelings. *The Quarterly Journal of Economics*, 116(1):55–79.
- Cavaillé, C. and Trump, K.-S. (2015). The two facets of social policy preferences. *The Journal of Politics*, 77(1):146–160.
- Charness, G. and Chen, Y. (2020). Social identity, group behavior, and teams. *Annual Review of Economics*, 12(1):691–713.
- Cho, I.-K. and Kreps, D. M. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221.
- Corneo, G. (2001). Inequality and the state: Comparing US and German preferences. *Annales d'Économie et de Statistique*, (63/64):283–296.
- Corneo, G. and Grüner, H. P. (2002). Individual preferences for political redistribution. *Journal of Public Economics*, 83(1):83–107.
- Coutts, A. (2019). Testing models of belief bias: An experiment. *Games and Economic Behavior*, 113:549–565.
- Cowen, T. (2002). The esteem theory of norms. *Public Choice*, 113(1-2):211–224.
- Dowding, K. (2005). Is it rational to vote? Five types of answer and a suggestion. *The British Journal of Politics and International Relations*, 7(3):442–459.
- Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2):135–150.
- Fernbach, P. M., Hagmayer, Y., and Sloman, S. A. (2014). Effort denial in self-deception. *Organizational Behavior and Human Decision Processes*, 123(1):1–8.
- Gallice, A. (2018). Social status, preferences for redistribution and optimal taxation: a survey. *Economics*, 12(1):2018–2055.
- Georgiadis, A. and Manning, A. (2012). Spend it like Beckham? Inequality and redistribution in the UK, 1983–2004. *Public Choice*, 151(3-4):537–563.
- Glazer, A. (1987). A new theory of voting: Why vote when millions of others do. *Theory and decision*, 22(3):257.
- Glazer, A. (1992). An expressive voting theory of strikes. *Economic Inquiry*, 30(4):733–741.
- Green, D. and Shapiro, I. (1994). *Pathologies of rational choice theory: A critique of applications in political science*. Yale University Press.
- Groeneveld, R. A. (1986). Skewness for the weibull family. *Statistica Neerlandica*, 40(3):135–140.

- Grossman, Z. (2015). Self-signaling and social-signaling in giving. *Journal of Economic Behavior & Organization*, 117:26–39.
- Hamlin, A. and Jennings, C. (2011). Expressive political behaviour: Foundations, scope and implications. *British Journal of Political Science*, 41(3):645–670.
- Hamlin, A. and Jennings, C. (2018). Expressive voting. In Congleton, R. D., Grofman, B. N., and Voigt, S., editors, *The Oxford Handbook of Public Choice, volume 1*, pages 333–350. Oxford University Press.
- Harms, P. and Zink, S. (2003). Limits to redistribution in a democracy: a survey. *European Journal of Political Economy*, 19(4):651–668.
- Heckman, J. J. and Honoré, B. E. (1990). The empirical content of the Roy model. *Econometrica*, 58(5):1121–1149.
- Hillman, A. L. (2010). Expressive behavior in economics and politics. *European Journal of Political Economy*, 26(4):403–418.
- Hillman, A. L. (2019). *Public Finance and Public Policy: A Political Economy Perspective on the Responsibilities and Limitations of Government*. Cambridge University Press, 3 edition.
- Hirschman, A. O. and Rothschild, M. (1973). The changing tolerance for income inequality in the course of economic development: with a mathematical appendix. *The Quarterly Journal of Economics*, 87(4):544–566.
- Jevons, H. S. (1905). *Essays on economics*. Macmillan, London.
- Kalleberg, A. L. (2003). Flexible firms and labor market segmentation: Effects of workplace restructuring on jobs and workers. *Work and Occupations*, 30(2):154–175.
- Korpi, W. and Palme, J. (2003). New politics and class politics in the context of austerity and globalization: Welfare state regress in 18 countries, 1975–95. *American Political Science Review*, 97(3):425–446.
- Lindert, P. H. (2004). *Growing public: Volume 1, the story: Social spending and economic growth since the eighteenth century*. Cambridge University Press, Cambridge.
- Lindqvist, E. and Östling, R. (2013). Identity and redistribution. *Public Choice*, 155:469–491.
- Löwenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal*, 97(387):666–684.
- Lupu, N. and Pontusson, J. (2011). The structure of inequality and the politics of redistribution. *American Political Science Review*, 105(2):316–336.
- Luttmer, E. F. (2001). Group loyalty and the taste for redistribution. *Journal of Political Economy*, 109(3):500–528.
- MacGillivray, H. (1981). The mean, median, mode inequality and skewness for a class of densities. *Australian Journal of Statistics*, 23(2):247–250.

- Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, 89(5):914–927.
- Mijović-Prelec, D. and Prelec, D. (2010). Self-deception as self-signalling: a model and experimental evidence. *Philosophical Transactions of the Royal Society B*, 365:227–240.
- Minozzi, W. (2013). Endogenous beliefs in models of politics. *American Journal of Political Science*, 57(3):566–581.
- Mueller, D. C. et al. (2003). *Public Choice III*. Cambridge University Press, Cambridge.
- Nozick, R. (1969). Newcomb’s problem and two principles of choice. In *Essays in honor of Carl G. Hempel*, pages 114–146. Springer.
- Penn, E. M. (2008). Citizenship versus ethnicity: The role of institutions in shaping identity choice. *The Journal of Politics*, 70(4):956–973.
- Piketty, T. (1995). Social mobility and redistributive politics. *The Quarterly Journal of Economics*, 110(3):551–584.
- Piketty, T. and Saez, E. (2014). Inequality in the long run. *Science*, 344(6186):838–843.
- Quattrone, G. A. and Tversky, A. (1986). Self-deception and the voter’s illusion. In Elster, J., editor, *The Multiple Self*, pages 35–38. Cambridge University Press, Cambridge.
- Ravallion, M. and Lokshin, M. (2000). Who wants to redistribute?: The tunnel effect in 1990s Russia. *Journal of Public Economics*, 76(1):87–104.
- Rehm, P., Hacker, J. S., and Schlesinger, M. (2012). Insecure alliances: Risk, inequality, and support for the welfare state. *American Political Science Review*, 106(2):386–406.
- Romer, T. (1975). Individual welfare, majority voting, and the properties of a linear income tax. *Journal of Public Economics*, 4(2):163–185.
- Sands, M. L. (2017). Exposure to inequality affects support for redistribution. *Proceedings of the National Academy of Sciences*, 114(4):663–668.
- Schuessler, A. A. (2000). Expressive voting. *Rationality and Society*, 12(1):87–119.
- Shayo, M. (2009). A model of social identity with an application to political economy: Nation, class, and redistribution. *American Political Science Review*, 103(2):147–174.
- Sloman, S. A., Fernbach, P. M., and Haggmayer, Y. (2010). Self-deception requires vagueness. *Cognition*, 115(2):268–281.
- Tullock, G. (1971). The charity of the uncharitable. *Economic Inquiry*, 9(4):379.

A Proofs

Lemma 1 (Proposition 1 (16a), (16b), Heckman and Honoré (1990)). *If ε has density f with f log-concave, then for $\varepsilon \in (\underline{\varepsilon}, \bar{\varepsilon})$ $\mathcal{M}'_+(\varepsilon) \in [0, 1]$ and $\mathcal{M}'_-(\varepsilon) \in [0, 1]$. The unit intervals are open if f is strictly log-concave.*

Proof. See Heckman and Honoré (1990). □

Lemma 2. *Suppose Assumption 1 holds. Then $\Psi' := \frac{d}{d\varepsilon} \Psi(\varepsilon; y) > 0$.*

Proof. $\frac{d}{d\varepsilon} \Psi(\varepsilon; y) = s(1 - \underline{\tau})\mathcal{M}'_+(\varepsilon) - s(1 - \bar{\tau})\mathcal{M}'_-(\varepsilon) + \delta(\bar{\tau} - \underline{\tau}) > -s(1 - \bar{\tau}) + \delta(\bar{\tau} - \underline{\tau}) > 0 \iff \delta(1 - \underline{\tau}) > (s + \delta)(1 - \bar{\tau})$, where the first inequality follows from strict log-concavity of f and Lemma 1. □

Lemma 3. *Let ε have density f with f a nondegenerate zero mean log-concave density and $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$ and $\sigma \in \mathbb{R}_{>0}$. Then for $\varepsilon' \in (\underline{\varepsilon}, \bar{\varepsilon})$*

$$(i) \frac{\partial}{\partial \sigma} E[\sigma\varepsilon | \sigma\varepsilon > \varepsilon'] \geq 0, \quad \text{and} \quad (ii) \frac{\partial}{\partial \sigma} E[\sigma\varepsilon | \sigma\varepsilon < \varepsilon'] \leq 0.$$

The inequalities are strict if f is strictly log-concave.

Proof. Note as a corollary to Lemma 1 that for $\varepsilon' \in (\underline{\varepsilon}, \bar{\varepsilon})$

$$\mathcal{M}_+(\varepsilon') > E[\varepsilon] = 0 \quad \text{and} \quad \mathcal{M}_-(\varepsilon') < E[\varepsilon] = 0. \quad (17)$$

Part (i).

$$\begin{aligned} \frac{\partial}{\partial \sigma} E[\sigma\varepsilon | \sigma\varepsilon > \varepsilon'] &= \frac{\partial}{\partial \sigma} \left(\sigma E \left[\varepsilon | \varepsilon > \frac{\varepsilon'}{\sigma} \right] \right) \\ &= E \left[\varepsilon | \varepsilon > \frac{\varepsilon'}{\sigma} \right] + \sigma \frac{\partial}{\partial \sigma} E \left[\varepsilon | \varepsilon > \frac{\varepsilon'}{\sigma} \right] \\ &= E \left[\varepsilon | \varepsilon > \frac{\varepsilon'}{\sigma} \right] + \sigma \frac{\partial}{\partial \frac{\varepsilon'}{\sigma}} E \left[\varepsilon | \varepsilon > \frac{\varepsilon'}{\sigma} \right] \frac{\partial}{\partial \sigma} \frac{\varepsilon'}{\sigma} \\ &= E \left[\varepsilon | \varepsilon > \frac{\varepsilon'}{\sigma} \right] + \sigma \frac{\partial}{\partial \frac{\varepsilon'}{\sigma}} E \left[\varepsilon | \varepsilon > \frac{\varepsilon'}{\sigma} \right] \left(-\frac{\varepsilon'}{\sigma^2} \right) \\ &= E \left[\varepsilon | \varepsilon > \frac{\varepsilon'}{\sigma} \right] - \frac{\partial}{\partial \frac{\varepsilon'}{\sigma}} E \left[\varepsilon | \varepsilon > \frac{\varepsilon'}{\sigma} \right] \frac{\varepsilon'}{\sigma} \\ &= E[\varepsilon | \varepsilon > \tilde{\varepsilon}'] - \frac{\partial}{\partial \tilde{\varepsilon}'} E[\varepsilon | \varepsilon > \tilde{\varepsilon}'] \tilde{\varepsilon}' \quad \text{where } \tilde{\varepsilon}' = \frac{\varepsilon'}{\sigma} \\ &= \mathcal{M}_+(\tilde{\varepsilon}') - \tilde{\varepsilon}' \mathcal{M}'_+(\tilde{\varepsilon}'). \end{aligned} \quad (18)$$

First note that if $\tilde{\varepsilon}' \leq 0$, then by Lemma 1 and (17), (18) is strictly positive. Second, if $\tilde{\varepsilon}' > 0$ we have

$$\mathcal{M}_+(\tilde{\varepsilon}') - \tilde{\varepsilon}' \mathcal{M}'_+(\tilde{\varepsilon}') > \tilde{\varepsilon}' - \tilde{\varepsilon}' \mathcal{M}'_+(\tilde{\varepsilon}') = \tilde{\varepsilon}' [1 - \mathcal{M}'_+(\tilde{\varepsilon}')] \geq 0,$$

where the last inequality follows from Lemma 1 and is strict if f is strictly log-concave. The first inequality follows from the non-degeneracy of f .

Part (ii).

$$\begin{aligned}
\frac{\partial}{\partial \sigma} E[\sigma \varepsilon | \sigma \varepsilon < \varepsilon'] &= \frac{\partial}{\partial \sigma} \left(\sigma E \left[\varepsilon | \varepsilon < \frac{\varepsilon'}{\sigma} \right] \right) \\
&= E \left[\varepsilon | \varepsilon < \frac{\varepsilon'}{\sigma} \right] + \sigma \frac{\partial}{\partial \sigma} E \left[\varepsilon | \varepsilon < \frac{\varepsilon'}{\sigma} \right] \\
&= E \left[\varepsilon | \varepsilon < \frac{\varepsilon'}{\sigma} \right] + \sigma \frac{\partial}{\partial \frac{\varepsilon'}{\sigma}} E \left[\varepsilon | \varepsilon < \frac{\varepsilon'}{\sigma} \right] \frac{\partial}{\partial \sigma} \frac{\varepsilon'}{\sigma} \\
&= E \left[\varepsilon | \varepsilon < \frac{\varepsilon'}{\sigma} \right] + \sigma \frac{\partial}{\partial \frac{\varepsilon'}{\sigma}} E \left[\varepsilon | \varepsilon < \frac{\varepsilon'}{\sigma} \right] \left(-\frac{\varepsilon'}{\sigma^2} \right) \\
&= E \left[\varepsilon | \varepsilon < \frac{\varepsilon'}{\sigma} \right] - \frac{\partial}{\partial \frac{\varepsilon'}{\sigma}} E \left[\varepsilon | \varepsilon < \frac{\varepsilon'}{\sigma} \right] \frac{\varepsilon'}{\sigma} \\
&= E[\varepsilon | \varepsilon < \tilde{\varepsilon}'] - \frac{\partial}{\partial \tilde{\varepsilon}'} E[\varepsilon | \varepsilon < \tilde{\varepsilon}'] \tilde{\varepsilon}' \quad \text{where } \tilde{\varepsilon}' = \frac{\varepsilon'}{\sigma} \\
&= \mathcal{M}_-(\tilde{\varepsilon}') - \tilde{\varepsilon}' \mathcal{M}'_-(\tilde{\varepsilon}'). \tag{19}
\end{aligned}$$

First note that if $\tilde{\varepsilon}' \geq 0$, then by Lemma 1 and (17), (19) is strictly negative. Second, if $\tilde{\varepsilon}' < 0$ we have

$$\mathcal{M}_-(\tilde{\varepsilon}') - \tilde{\varepsilon}' \mathcal{M}'_-(\tilde{\varepsilon}') < \tilde{\varepsilon}' - \tilde{\varepsilon}' \mathcal{M}'_-(\tilde{\varepsilon}') = \tilde{\varepsilon}' [1 - \mathcal{M}'_-(\tilde{\varepsilon}')] \leq 0,$$

where the last inequality follows from Lemma 1 and is strict if f is strictly log-concave. The first inequality follows from the non-degeneracy of f . \square

Lemma 4.

$$\frac{d}{d\sigma} \int_{y \in Y} h(y, \sigma) dG(y) = \int_{y^L}^{y^H} \frac{d}{d\sigma} h(y, \sigma) dG(y).$$

Proof.

$$\begin{aligned}
& \frac{d}{d\sigma} \int_{y \in Y} h(y, \sigma) dG(y) \\
&= \frac{d}{d\sigma} \left[\int_{-\infty}^{y^L} h(y, \sigma) dG(y) + \int_{y^L}^{y^H} h(y, \sigma) dG(y) + \int_{y^H}^{\infty} h(y, \sigma) dG(y) \right] \\
&= \frac{d}{d\sigma} \left[\int_{-\infty}^{y^L} dG(y) + \int_{y^L}^{y^H} h(y, \sigma) dG(y) \right] \\
&= \frac{d}{d\sigma} G(y^L) + \frac{d}{d\sigma} \left[h(y^H, \sigma) G(y^H) - h(y^L, \sigma) G(y^L) - \int_{y^L}^{y^H} G(y) dh(y, \sigma) \right] \\
&= G(y^H) \frac{d}{d\sigma} h(y^H, \sigma) - G(y^L) \frac{d}{d\sigma} h(y^L, \sigma) - \frac{d}{d\sigma} \int_{y^L}^{y^H} G(y) \frac{\partial}{\partial y} h(y, \sigma) dy \\
&= G(y^H) \frac{d}{d\sigma} h(y^H, \sigma) - G(y^L) \frac{d}{d\sigma} h(y^L, \sigma) \\
&\quad - G(y^H) \frac{\partial}{\partial y} h(y^H, \sigma) \frac{dy^H}{d\sigma} + G(y^L) \frac{\partial}{\partial y} h(y^L, \sigma) \frac{dy^L}{d\sigma} - \int_{y^L}^{y^H} G(y) \frac{\partial^2}{\partial y \partial \sigma} h(y, \sigma) dy \\
&= G(y^H) \frac{\partial}{\partial \sigma} h(y^H, \sigma) - G(y^L) \frac{\partial}{\partial \sigma} h(y^L, \sigma) - \int_{y^L}^{y^H} G(y) d \frac{\partial}{\partial \sigma} h(y, \sigma) \\
&= \int_{y^L}^{y^H} \frac{\partial}{\partial \sigma} h(y, \sigma) dG(y),
\end{aligned}$$

where the second equality follows from (9), where $h(y^L, \sigma) = F(\varepsilon^*(y^L)) = F(\bar{\varepsilon}) = 1$ and $h(y^H, \sigma) = F(\varepsilon^*(y^H)) = F(\underline{\varepsilon}) = 0$ and where the derivatives of h with respect to the boundaries of integration y^L and y^H exist as defined in Propositions 1 and 2, and where $\frac{\partial}{\partial \sigma} h(y, \sigma)$ is continuous on $[y^L, y^H]$ in y and σ by Proposition 4 \square

Lemma 5. *Suppose F and G are symmetric. Then*

$$\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} G(\bar{y} - \varepsilon) dF(\varepsilon) = G(\bar{y}).$$

Proof. Since F is symmetric, we can let $\underline{\varepsilon} = -\varepsilon'$ and $\bar{\varepsilon} = \varepsilon'$. Then

$$\begin{aligned}
\int_{-\varepsilon'}^{\varepsilon'} G(\bar{y} - \varepsilon) dF(\varepsilon) &= \int_{-\varepsilon'}^0 G(\bar{y} - \varepsilon) dF(\varepsilon) + \int_0^{\varepsilon'} G(\bar{y} - \varepsilon) dF(\varepsilon) \\
&= \int_{-\varepsilon'}^0 G(\bar{y} - \varepsilon) dF(\varepsilon) - \int_0^{-\varepsilon'} G(\bar{y} + \varepsilon) f(-\varepsilon) d\varepsilon \\
&= \int_{-\varepsilon'}^0 G(\bar{y} - \varepsilon) dF(\varepsilon) + \int_{-\varepsilon'}^0 [1 - G(\bar{y} - \varepsilon)] f(\varepsilon) d\varepsilon \\
&= \int_{-\varepsilon'}^0 f(\varepsilon) d\varepsilon = \frac{1}{2} = G(\bar{y}),
\end{aligned}$$

where $f(-\varepsilon) = f(\varepsilon)$, $G(\bar{y} + \varepsilon) = 1 - G(\bar{y} - \varepsilon)$, and $G(\bar{y}) = \frac{1}{2}$ by symmetry. \square

Proof of Proposition 1. First, find the voters that always vote for low taxation. Let $\varepsilon^*(y) = \underline{\varepsilon}$.

Since these voters know they would have voted for low taxation for all ε , voting does not convey any information, and they expect an income change according to prior such that $\mathcal{M}_+(\underline{\varepsilon}) = 0$. Deviation to voting for the high tax rate is most profitable to the voter with the worst income prospects, and so the off-equilibrium-path belief is $\mathcal{M}_- = \underline{\varepsilon}$. Pooling requires that (7) is positive for all $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$ which, since (7) is increasing in ε and given the beliefs $\mathcal{M}_+(\underline{\varepsilon}) = 0$ and $\mathcal{M}_- = \underline{\varepsilon}$, is true if and only if $\Psi(\underline{\varepsilon}; y) \geq 0$. This is equivalent to

$$y \geq y^H := \bar{y} + \frac{|\underline{\varepsilon}|}{s + \delta} \left(\delta - s \frac{(1 - \bar{\tau})}{(\bar{\tau} - \underline{\tau})} \right). \quad (20)$$

Since Ψ is increasing, if $\Psi(\underline{\varepsilon}) \geq 0$ there are no $\varepsilon' \neq \underline{\varepsilon}$ such that $\Psi(\varepsilon') = 0$ and so this equilibrium is unique. Note that, $y^H > \bar{y}$ if and only if Assumption 1 holds.

Second, find the voters that always vote for high taxation. Let $\varepsilon^*(y) = \bar{\varepsilon}$. Since these voters know they would have voted for high taxation for all ε , voting does not convey any information, and they expect an income change according to prior such that $\mathcal{M}_-(\bar{\varepsilon}) = 0$. Deviation to voting for the low tax rate is most profitable to the voter with the best income prospects, and so the off-equilibrium-path belief is $\mathcal{M}_+ = \bar{\varepsilon}$. Pooling requires that (7) is negative for all $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$ which, since (7) is increasing in ε and given the beliefs $\mathcal{M}_-(\bar{\varepsilon}) = 0$ and $\mathcal{M}_+ = \bar{\varepsilon}$, is true if and only if $\Psi(\bar{\varepsilon}; y) \leq 0$. This is equivalent to

$$y \leq y^L := \bar{y} - \frac{\bar{\varepsilon}}{s + \delta} \left(\delta + s \frac{(1 - \underline{\tau})}{(\bar{\tau} - \underline{\tau})} \right). \quad (21)$$

Since Ψ is increasing, if $\Psi(\bar{\varepsilon}) \leq 0$ there are no $\varepsilon' \neq \bar{\varepsilon}$ such that $\Psi(\varepsilon') = 0$ and so this equilibrium is unique.

Third, find the voters whose vote depends on the income change they expect. Suppose now $y \in (y_L, y_H)$. Then $\Psi(\underline{\varepsilon}; y) < 0$ and $\Psi(\bar{\varepsilon}; y) > 0$ and as Ψ is continuous and strictly increasing, there exists unique threshold $\varepsilon^*(y) \in (\underline{\varepsilon}, \bar{\varepsilon})$ such that $\Psi(\varepsilon^*(y); y) = 0$ and such that for all $\varepsilon < \varepsilon^*(y)$ we have $\Psi(\varepsilon^*(y); y) < 0$ and thus a vote for the high tax rate is preferred, and for all $\varepsilon > \varepsilon^*(y)$ we have $\Psi(\varepsilon^*(y); y) > 0$ and thus a vote for the low tax rate is preferred.

To show that $\varepsilon^*(y)$ decreases in y , note that for $y < y^L$, $\varepsilon^*(y) = \bar{\varepsilon}$ and for $y > y^H$, $\varepsilon^*(y) = \underline{\varepsilon}$ and that $\varepsilon^*(y)$ approaches $\bar{\varepsilon}$ when y approaches y^L from the right and $\varepsilon^*(y)$ approaches $\underline{\varepsilon}$ when y approaches y^H from the left. Consider $y \in (y^L, y^H)$. Totally differentiate $\Psi(\varepsilon^*(y); y) = 0$ and rearrange to get

$$\frac{d\varepsilon^*(y)}{dy} = -\frac{\Psi_y}{\Psi'} < 0, \quad (22)$$

where $\Psi_y = (s + \delta)(\bar{\tau} - \underline{\tau}) > 0$ and, by Lemma 2, $\Psi' > 0$. Define the derivative at y^L as the right derivative at that point and the derivative at y^H as the left derivative at that point. Both these derivatives are negative as the derivative on (y^L, y^H) is negative. With these derivatives at the boundaries and by Implicit Function Theorem, $\varepsilon^*(y)$ is continuously differentiable with respect to y on $[y^L, y^H]$. \square

Proof of Proposition 2. With $s = 0$, $\Psi(\varepsilon; y) = -\delta(\bar{\tau} - \underline{\tau})(\bar{y} - y - \varepsilon)$, which is positive whenever $y + \varepsilon > \bar{y}$. Let $\hat{\varepsilon}(y) := \bar{y} - y$. Also, the income threshold for pooling on voting for the low tax

rate becomes $\hat{y}^H := y^H = \bar{y} + |\underline{\varepsilon}|$, that is, the expected future income is never below mean, and the income threshold for pooling on voting for the high tax rate becomes $\hat{y}^L := y^L = \bar{y} - \bar{\varepsilon}$, that is, the expected future income is never above the mean.

For $s > 0$, show that $\Psi(\varepsilon'; y) = 0$ cannot hold for any $\varepsilon' \geq \hat{\varepsilon}(y)$. Since by Lemma 2, Ψ is increasing it is enough to show that $\Psi(\hat{\varepsilon}(y); y) > 0$. We have

$$\begin{aligned}\Psi(\hat{\varepsilon}(y); y) &= s[(1 - \underline{\tau})\mathcal{M}_+(\hat{\varepsilon}(y)) - (1 - \bar{\tau})\mathcal{M}_-(\hat{\varepsilon}(y)) - (\bar{\tau} - \underline{\tau})(\bar{y} - y)] \\ &> s[(1 - \underline{\tau})\hat{\varepsilon}(y) - (1 - \bar{\tau})\hat{\varepsilon}(y) - (\bar{\tau} - \underline{\tau})(\bar{y} - y)] \\ &= s(\bar{\tau} - \underline{\tau})(\hat{\varepsilon}(y) - (\bar{y} - y)) = 0,\end{aligned}$$

where $s > 0$ and where the inequality follows from $\varepsilon' < \mathcal{M}_+(\varepsilon')$ and $\varepsilon' > \mathcal{M}_-(\varepsilon')$ for non-degenerate F . Hence, $\varepsilon^*(y) < \hat{\varepsilon}(y)$.

The income threshold required to always vote for the low tax rate decreases in s :

$$\frac{\partial y^H(s)}{\partial s} = -\frac{|\underline{\varepsilon}|\delta}{(s + \delta)^2} \left(1 + \frac{(1 - \bar{\tau})}{(\bar{\tau} - \underline{\tau})}\right) < 0.$$

The income threshold required to always vote for the high tax rate decreases in s :

$$\frac{\partial y^L(s)}{\partial s} = -\frac{\bar{\varepsilon}\delta}{(s + \delta)^2} \left(\frac{(1 - \underline{\tau})}{(\bar{\tau} - \underline{\tau})} - 1\right) < 0,$$

since

$$\frac{(1 - \underline{\tau})}{(\bar{\tau} - \underline{\tau})} > \frac{(\bar{\tau} - \underline{\tau})}{(\bar{\tau} - \underline{\tau})} = 1.$$

Consider an interior equilibrium $\varepsilon^*(y) \in (\underline{\varepsilon}, \bar{\varepsilon})$. Total differentiation of $\Psi(\varepsilon^*(y)) = 0$ with respect to s yields

$$\frac{\partial \varepsilon^*(y)}{\partial s} = -\frac{\Psi_s}{\Psi'} < 0, \tag{23}$$

where $\Psi' > 0$ by Lemma 2. Define the derivative at $\varepsilon^*(y^L) = \bar{\varepsilon}$ as the right derivative and at $\varepsilon^*(y^H) = \underline{\varepsilon}$ as the left derivative. It remains to show that $\Psi_s > 0$. The threshold $\varepsilon^*(y)$ satisfies $\Psi(\varepsilon^*(y); y) = 0$, that is,

$$\begin{aligned}s[(1 - \underline{\tau})\mathcal{M}_+(\varepsilon^*(y)) - (1 - \bar{\tau})\mathcal{M}_-(\varepsilon^*(y)) - (\bar{\tau} - \underline{\tau})(\bar{y} - y)] \\ = \delta(\bar{\tau} - \underline{\tau})(\bar{y} - y - \varepsilon^*(y)).\end{aligned} \tag{24}$$

Now suppose $\Psi_s \leq 0$. If $s > 0$, this implies that the left-hand side of (24) is weakly negative. That is,

$$\begin{aligned}0 &\geq (1 - \underline{\tau})\mathcal{M}_+(\varepsilon^*(y)) - (1 - \bar{\tau})\mathcal{M}_-(\varepsilon^*(y)) - (\bar{\tau} - \underline{\tau})(\bar{y} - y) \\ &> (1 - \underline{\tau})\varepsilon^*(y) - (1 - \bar{\tau})\varepsilon^*(y) - (\bar{\tau} - \underline{\tau})(\bar{y} - y) \\ &= (\bar{\tau} - \underline{\tau})(\varepsilon^*(y) - \bar{y} + y),\end{aligned}$$

where the second inequality follows from the fact that $\varepsilon' < \mathcal{M}_+(\varepsilon')$ and $\varepsilon' > \mathcal{M}_-(\varepsilon')$ for non-degenerate F . We thus have $(\bar{\tau} - \underline{\tau})(\varepsilon^*(y) - \bar{y} + y) < 0$ implying $\bar{y} - y - \varepsilon^*(y) > 0$ and thus the right hand side of (24) is strictly positive. Thus, weak negativity of the left-hand side of (24) implies strict positivity of the right-hand side of (24) violating the equilibrium condition $\Psi(\varepsilon^*(y); y) = 0$. Since $\Psi_s \leq 0$ gives us a contradiction, we must have $\Psi_s > 0$ whenever $\varepsilon^*(y) \in (\underline{\varepsilon}, \bar{\varepsilon})$. \square

Proof of Proposition 3. For $\varepsilon^*(y) \in (\underline{\varepsilon}, \bar{\varepsilon})$, total differentiation of $\Psi(\varepsilon^*(y)) = 0$ with respect to σ yields,

$$\frac{\partial \varepsilon^*(y)}{\partial \sigma} = -\frac{\Psi_\sigma}{\Psi'} < 0 \quad (25)$$

where $\Psi' > 0$ by Lemma 2 and

$$\Psi_\sigma = s[(1 - \underline{\tau})\frac{\partial}{\partial \sigma} E[\sigma\varepsilon | \sigma\varepsilon > \varepsilon^*(y)] - (1 - \bar{\tau})\frac{\partial}{\partial \sigma} E[\sigma\varepsilon | \sigma\varepsilon < \varepsilon^*(y)]] > 0, \quad (26)$$

where $\frac{\partial}{\partial \sigma} E[\sigma\varepsilon | \sigma\varepsilon > \varepsilon^*(y)] > 0$ and $\frac{\partial}{\partial \sigma} E[\sigma\varepsilon | \sigma\varepsilon < \varepsilon^*(y)] < 0$ follow from Lemma 3 and strict log-concavity of f . Define the derivative at $\varepsilon^*(y^L) = \bar{\varepsilon}$ as the right derivative and at $\varepsilon^*(y^H) = \underline{\varepsilon}$ as the left derivative. With these derivatives at the boundaries and by Implicit Function Theorem, $\varepsilon^*(y)$ is continuously differentiable with respect to σ on $[y^L, y^H]$. Clearly, for $s = 0$, $\Psi_\sigma = 0$. For the thresholds y^L and y^H we have

$$\frac{\partial y^H}{\partial \sigma} = \frac{|\underline{\varepsilon}|}{s + \delta} \left(\delta - s \frac{(1 - \bar{\tau})}{(\bar{\tau} - \underline{\tau})} \right) > 0$$

since if and only if Assumption 1 holds the term in the parenthesis is strictly positive and

$$\frac{\partial y^L}{\partial \sigma} = -\frac{\bar{\varepsilon}}{s + \delta} \left(\delta + s \frac{(1 - \underline{\tau})}{(\bar{\tau} - \underline{\tau})} \right) < 0.$$

\square

Proof of Proposition 4. First note that

$$\frac{\partial}{\partial \sigma} h(y, \sigma) = \frac{d}{d\sigma} F\left(\frac{\varepsilon^*(y)}{\sigma}\right) = f\left(\frac{\varepsilon^*(y)}{\sigma}\right) \frac{d}{d\sigma} \left(\frac{\varepsilon^*(y)}{\sigma}\right) = f\left(\frac{\varepsilon^*(y)}{\sigma}\right) \frac{1}{\sigma} \left(\frac{\partial \varepsilon^*(y)}{\partial \sigma} - \frac{\varepsilon^*(y)}{\sigma}\right).$$

Evaluating this without loss of generality at $\sigma = 1$ and noting that

$$\begin{aligned} \Psi_\sigma &= s[(1 - \underline{\tau})\frac{\partial}{\partial \sigma} E[\sigma\varepsilon | \sigma\varepsilon > \varepsilon^*] - (1 - \bar{\tau})\frac{\partial}{\partial \sigma} E[\sigma\varepsilon | \sigma\varepsilon < \varepsilon^*]] \\ &= s[(1 - \underline{\tau})[\mathcal{M}_+(\varepsilon^*) - \varepsilon^* \mathcal{M}'_+(\varepsilon^*)] - (1 - \bar{\tau})[\mathcal{M}_-(\varepsilon^*) - \varepsilon^* \mathcal{M}'_-(\varepsilon^*)]] \\ &= s[(1 - \underline{\tau})\mathcal{M}_+(\varepsilon^*) - (1 - \bar{\tau})\mathcal{M}_-(\varepsilon^*)] - s[(1 - \underline{\tau})\mathcal{M}'_+(\varepsilon^*) - (1 - \bar{\tau})\mathcal{M}'_-(\varepsilon^*)]\varepsilon^* \end{aligned}$$

and so

$$\Psi_\sigma + \Psi' \varepsilon^* = s[(1 - \underline{\tau})\mathcal{M}_+(\varepsilon^*) - (1 - \bar{\tau})\mathcal{M}_-(\varepsilon^*)] + \delta(\bar{\tau} - \underline{\tau})\varepsilon^* = \Psi(\varepsilon^*; \bar{y})$$

and thus we have

$$\frac{\partial \varepsilon^*}{\partial \sigma} - \varepsilon^* = -\frac{\Psi_\sigma}{\Psi'} - \frac{\Psi' \varepsilon^*}{\Psi'} = -\frac{\Psi_\sigma + \Psi' \varepsilon^*}{\Psi'} = -\frac{\Psi(\varepsilon^*(y); \bar{y})}{\Psi'} \quad (27)$$

and so

$$\frac{\partial}{\partial \sigma} h(y, \sigma) = f(\varepsilon^*(y)) \left(\frac{\partial \varepsilon^*(y)}{\partial \sigma} - \varepsilon^*(y) \right) = -f(\varepsilon^*(y)) \frac{\Psi(\varepsilon^*(y); \bar{y})}{\Psi'}. \quad (28)$$

Now, Ψ is increasing and ε^* is decreasing and so $\Psi(\varepsilon^*(y); \bar{y})$ is decreasing in y and by definition of ε^* , $\Psi(\varepsilon^*(\bar{y}); \bar{y}) = 0$. Assumption 1 ensures that $y^L < \bar{y} < y^H$ and so $\Psi(\varepsilon^*(y^H); \bar{y}) < 0 < \Psi(\varepsilon^*(y^L); \bar{y})$. For voters with $y \in [y^L, \bar{y})$ we have $\Psi(\varepsilon^*(y); \bar{y}) > 0$ and thus $\frac{\partial}{\partial \sigma} h(y, \sigma) < 0$; for voters with $y \in (\bar{y}, y^H]$ we have $\Psi(\varepsilon^*(y); \bar{y}) < 0$ and thus $\frac{\partial}{\partial \sigma} h(y, \sigma) > 0$; for voters with $y = \bar{y}$ we have $\Psi(\varepsilon^*(y); \bar{y}) = 0$ and thus $\frac{\partial}{\partial \sigma} h(y, \sigma) = 0$, where $\frac{\partial}{\partial \sigma} h(y, \sigma)$ is continuous on $[y^L, y^H]$ in y and σ as f is continuous and, by Proposition 3, ε^* is continuously differentiable on $[y^L, y^H]$. \square

Proof of Proposition 5. Let G be uniform. Then by Lemma 4, the effect of an increase in income risk on aggregate demand for redistribution is

$$\begin{aligned} \frac{d}{d\sigma} \int_{y \in Y} h(y, \sigma) dG(y) &\propto \int_{y^L}^{y^H} (\bar{y} - y) \frac{d}{dy} F(\varepsilon^*(y)) dy \\ &= - \int_{y^L}^{y^H} (\bar{y} - y) \frac{d}{dy} [1 - F(\varepsilon^*(y))] dy, \end{aligned}$$

where $1 - F(\varepsilon^*(y)) = Pr[y^*(\varepsilon) \leq y]$ is the distribution of y^* in the electorate. Since y^* , as its inverse ε^* , is decreasing in s for all ε and increase in s moves the distribution of y^* to the left in the sense of first order stochastic dominance. Thus,

$$\begin{aligned} - \int_{y^L}^{y^H} (\bar{y} - y) \frac{d}{dy} [1 - F(\varepsilon^*(y))] dy &\leq - \int_{y^L}^{y^H} (\bar{y} - y) \frac{d}{dy} [1 - F(\hat{\varepsilon}(y))] dy \\ &= - \int_{y^L}^{y^H} (\bar{y} - y) f(\bar{y} - y) dy, \end{aligned}$$

where the inequality follows from first order stochastic dominance: First note that $\frac{d}{dy} [1 - F(\varepsilon^*(y))]$ is a density on $[y_L, y_H]$: it is clearly non-negative as $[1 - F(\varepsilon^*(y))]$ is increasing in y . Noting that $\varepsilon^*(y^H) = \underline{\varepsilon}$ and $\varepsilon^*(y^L) = \bar{\varepsilon}$, it also integrates to 1: $\int_{y^L}^{y^H} \frac{d}{dy} [1 - F(\varepsilon^*(y))] dy = F(\bar{\varepsilon}) = 1$. Similarly, $\frac{d}{dy} [1 - F(\hat{\varepsilon}(y))]$ is a density on $[y_L, y_H]$. Second, by Proposition 2, $\varepsilon^*(y) \leq \hat{\varepsilon}(y)$ for $s \geq 0$ with equality at $s = 0$ and thus $1 - F(\varepsilon^*(y)) \geq 1 - F(\hat{\varepsilon}(y))$. Hence $\frac{d}{dy} [1 - F(\hat{\varepsilon}(y))]$ first order stochastically dominates $\frac{d}{dy} [1 - F(\varepsilon^*(y))]$. Next, substitution by $\varepsilon = \bar{y} - y$ such that $d\varepsilon = -dy$ gives

$$\begin{aligned} \int_{\bar{y}-y^L}^{\bar{y}-y^H} \varepsilon f(\varepsilon) d\varepsilon &= - \int_{\bar{y}-y^H}^{\bar{y}-y^L} \varepsilon f(\varepsilon) d\varepsilon \propto - \frac{\int_{\bar{y}-y^H}^{\bar{y}-y^L} \varepsilon f(\varepsilon) d\varepsilon}{F(\bar{y}-y^L) - F(\bar{y}-y^H)} \\ &= -E[\varepsilon | \bar{y} - y^H \leq \varepsilon \leq \bar{y} - y^L] \leq 0. \end{aligned} \quad (29)$$

To show that the weak inequality holds in (29), let $e(s) := E[\varepsilon|\bar{y} - y^H(s) \leq \varepsilon \leq \bar{y} - y^L(s)]$. For $s = 0$ we have

$$e(0) = E[\varepsilon|\underline{\varepsilon} \leq \varepsilon \leq \bar{\varepsilon}] = E[\varepsilon] = 0.$$

Also, $e(s)$ is increasing in s :

$$\frac{d}{ds}e(s) = \frac{de}{d(\bar{y} - y^L)} \frac{d}{ds}(\bar{y} - y^L) + \frac{de}{d(\bar{y} - y^H)} \frac{d}{ds}(\bar{y} - y^H) > 0,$$

since $\frac{dy^L}{ds} < 0$ and $\frac{dy^H}{ds} < 0$ by Proposition 2. Thus, for $s > 0$, $e(s) > 0$. Hence, the inequality in (29) is equality for $s = 0$ and strict inequality for $s > 0$ and so the aggregate demand for redistribution strictly decreases in income risk if and only if $s > 0$. \square

Proof of Proposition 6. Let F be uniform, $\varepsilon \sim U(-\bar{\varepsilon}, \bar{\varepsilon})$ and G symmetric. The threshold ε^* is now defined by

$$s \left[\frac{1}{2}(\bar{\tau} - \underline{\tau})\varepsilon^* + \frac{1}{2}\bar{\varepsilon}(2 - \underline{\tau} - \bar{\tau}) - (\bar{\tau} - \underline{\tau})(\bar{y} - y) \right] - \delta(\bar{\tau} - \underline{\tau})(\bar{y} - y - \varepsilon^*) = 0$$

and has a closed form solution

$$\varepsilon^*(y) = \frac{(s + \delta)(\bar{\tau} - \underline{\tau})(\bar{y} - y) - \frac{1}{2}s(2 - \underline{\tau} - \bar{\tau})\bar{\varepsilon}}{(\bar{\tau} - \underline{\tau})(\frac{1}{2}s + \delta)} = \frac{1}{\Psi'}(\Psi_y(\bar{y} - y) - \Psi_\sigma)$$

with $\Psi_y = (s + \delta)(\bar{\tau} - \underline{\tau}) > 0$ and $\Psi' = (\bar{\tau} - \underline{\tau})(\frac{1}{2}s + \delta) > 0$ independent of y . We thus have

$$\frac{\partial \varepsilon^*}{\partial \sigma} - \varepsilon^* = -\frac{\Psi_\sigma}{\Psi'} - \frac{1}{\Psi'}(\Psi_y(\bar{y} - y) - \Psi_\sigma) = -\frac{\Psi_y}{\Psi'}(\bar{y} - y)$$

and hence, the effect of an increase in income risk on aggregate demand is proportional to a strictly positive factor to

$$\begin{aligned} \int_{y^L}^{y^H} \left(\frac{\partial \varepsilon^*}{\partial \sigma} - \varepsilon^* \right) dG(y) &= -\frac{\Psi_y}{\Psi'} \int_{y^L}^{y^H} (\bar{y} - y) dG(y) \\ &= -\frac{\Psi_y}{\Psi'} \left(\bar{y}[G(y^H) - G(y^L)] - \int_{y^L}^{y^H} y dG(y) \right) \\ &\propto -\frac{\Psi_y}{\Psi'} \left(\bar{y} - \frac{\int_{y^L}^{y^H} y dG(y)}{G(y^H) - G(y^L)} \right) \\ &= -\frac{\Psi_y}{\Psi'} (\bar{y} - E[y|y^L \leq y \leq y^H]) \leq 0. \end{aligned} \quad (30)$$

To show that the weak inequality holds in (30), let $e(s) := E[y|y^L(s) \leq y \leq y^H(s)]$. For $s = 0$, for symmetric G , we have

$$e(0) = E[y|\bar{y} - \bar{\varepsilon} \leq y \leq \bar{y} + \bar{\varepsilon}] = \bar{y}.$$

Also, $e(s)$ is decreasing in s :

$$\frac{d}{ds}e(s) = \frac{de}{dy^L} \frac{dy^L}{ds} + \frac{de}{dy^H} \frac{dy^H}{ds} < 0.$$

since $\frac{dy^L}{ds} < 0$ and $\frac{dy^H}{ds} < 0$ by Proposition 2 and $\frac{de}{dy^L} > 0$ and $\frac{de}{dy^H} > 0$. Thus, for $s > 0$, $e(s) < \bar{y}$. Hence, the inequality in (30) is equality for $s = 0$ and strict inequality for $s > 0$ and so the aggregate demand for redistribution strictly decreases in income risk if and only if $s > 0$. \square

Proof of Proposition 7. If a voter votes for t and t is implemented the voter's utility flows sum to

$$\delta u(t, \varepsilon, y) + sE[U(t, \varepsilon, y)|t, y]. \quad (31)$$

Thus, the marginal voter ε^{**} who is indifferent between voting for low tax rate and high tax rate in terms of this experienced utility satisfies

$$s[(1 - \underline{\tau})\mathcal{M}_+(\varepsilon^{**}) - (1 - \bar{\tau})\mathcal{M}_-(\varepsilon^{**}) + (\bar{\tau} - \underline{\tau})(\bar{y} - y)] - \delta(\bar{t} - \underline{t})(\bar{y} - y - \varepsilon^{**}) = 0. \quad (32)$$

Since $(\bar{\tau} - \underline{\tau}) = (q(\bar{t}) - q(\underline{t}))(\bar{t} - \underline{t})$, we have $\varepsilon^* < \varepsilon^{**} < \hat{\varepsilon}$ whenever $q(\bar{t}) - q(\underline{t}) < 1$. Thus, for voters with $\varepsilon \in (\varepsilon^*, \varepsilon^{**})$ the left-hand side of (32) is negative and they would thus be better off if they voted for high tax rate and high tax rate was implemented. For $q(\bar{t}) - q(\underline{t}) = 1$ the left hand side of (32) is $\Psi(\varepsilon^{**})$ and we have $\varepsilon^{**} = \varepsilon^*$. Also since $(\bar{\tau} - \underline{\tau}) < (\bar{t} - \underline{t})$, ε^{**} is less responsive to changes in s and σ than ε^* . Thus, $\varepsilon^{**} - \varepsilon^*$ is increasing in s and σ . \square