TAXES AND TURNOUT

Felix Bierbrauer
Aleh Tsyvinski
Nicolas D. Werquin

ABSTRACT

We develop a model of political competition with endogenous turnout and endogenous platforms. Parties face a trade-off between maximizing their base and getting their supporters out to vote. We study the implications of this framework for non-linear income taxation. In equilibrium, both parties propose the same tax policy. This equilibrium policy is a weighted combination of two terms, one reflecting the parties' payoff from mobilizing their own supporters, one reflecting the payoff from demobilizing the supporters of the other party. The key determinant of the equilibrium policy is the distribution of the voters' party attachments rather than their propensity to swing vote. Our analysis also provides a novel explanation for why even left-leaning parties may not propose high taxes on the rich.

Felix Bierbrauer
CMR - Center for Macroeconomic Research
University of Cologne
Albert-Magnus Platz
50923 Cologne
Germany
bierbrauer@wiso.uni-koeln.de

Aleh Tsyvinski
Department of Economics
Yale University
Box 208268
New Haven, CT 06520-8268
and NBER
a.tsyvinski@yale.edu

Nicolas D. Werquin
Toulouse School of Economics
21 allee de Brienne
31000 Toulouse
nwerquin@gmail.com

Introduction

To win in an election, parties need to get their supporters out to vote. This paper develops a model of political competition with two political parties that seek to maximize the probability of winning an election. The electorate includes rule-utilitarian voters who participate in an election only if the benefits of a victory of their preferred party outweigh the costs of voting. That is, participation is endogenous. Parties propose policies taking into account these implications for turnout. That is, policies are endogenous. We show that parties face a trade-off between maximizing their base, defined as the mass of citizens who are better off if the party wins, and getting their supporters out to vote. This in turn translates into the central determinant of the choice of policies, namely, a trade-off between mobilizing one's own supporters and demobilizing those of the adversary.

Our formal analysis combines the probabilistic voting model due to Coughlin and Nitzan (1981) and Lindbeck and Weibull (1987) with the models of participation in elections due to Harsanyi (1980), Coate and Conlin (2004) and Feddersen and Sandroni (2006). In probabilistic voting models it is assumed that parties maximize their vote shares and voters trade off idiosyncratic party preferences with the utility they derive from the platforms that the competing parties propose.[1] Appropriate assumptions on the distribution of party preferences then ensure that equilibrium policies are well-defined and the focus is on the characterization of the policies that emerge in a probabilistic voting equilibrium. It is, however, assumed that everyone who is eligible to vote participates in the election. This assumption is relaxed by the literature that treats turnout as an endogenous political equilibrium outcome. We follow the literature that models participation in an election as the outcome of a rule-utilitarian calculation. The supporters, say, of party 1 formulate a rule for participating in an election that trades off what is at stake for the supporters of party 1 against their costs of voting. As has been shown by Coate and Conlin (2004) and Feddersen and Sandroni (2006), this approach yields comparative statics properties of turnout that are consistent with stylized facts about participation in elections, e.g., participation is larger the closer is the race between the two alternatives, or the more polarized preferences are. The literature on endogenous turnout views the alternatives in an election as exogenously given, however. Choices are between a competitor and

---

[1]The probabilistic voting model is a special case of what has been coined a random participation model by Rochet and Stole (2002).

an incumbent, or between a reform and a status quo so that there is no analysis of a party's problem to refine the alternatives that are proposed so as to achieve a higher turnout of its supporters and thereby a higher probability of winning. With our approach, both the platform choices of parties and the turnout decisions of voters are equilibrium outcomes.

It is a frequent observation that individuals with high incomes are more likely to participate in elections than individuals with low incomes. We therefore study the implications of our approach for a policy domain that allows for a differential treatment of individuals with different incomes. A natural policy space is therefore to consider non-linear income tax schedules, so that parties can propose tax schedules that are, e.g., progressive, flat or regressive; they can propose high marginal tax rates on the rich or earnings subsidies for the poor, etc. Note that the set of non-linear tax policies is a multi-dimensional policy space. With such a policy space, the existence of a Condorcet winner is not to be expected. This creates a complication for any political economy analysis of non-linear tax systems. Our main contribution is to provide a characterization of the non-linear income tax system that emerges in a political equilibrium with endogenous turnout. Another contribution is to provide sufficient conditions for the existence of a pure strategy equilibrium.

Equilibria with endogenous turnout are driven by strategic considerations that have no analogue in models with exogenous turnout. With exogenous turnout, parties "only" need to check whether they can deviate from a hypothetical equilibrium policy to an alternative that makes a larger set of voters better off – of course, taking account of the incentive and resource constraints that define the set of admissible tax policies. With endogenous turnout, a larger set of supporters no longer translates automatically into a larger probability of winning. An increase of the set of supporters may lead to a decrease in the probability of winning. For instance, if party 1's proposal becomes more similar to the proposal of party 2, this may increase the set of individuals who like party 1's proposal better while at the same time reducing their incentive to vote. Why bother to vote when the proposals are similar anyway, if the agent has no strong idiosyncratic preference for either party?

Our main result is to show that the equilibrium policy has a particularly simple form. The equilibrium is symmetric, i.e., both parties propose the same policy in equilibrium. The equilibrium policy is determined by a weighted sum of two terms, one reflecting a party's benefit from mobilizing its own supporters and one reflecting the benefit from demobilizing the supporters of the other party. The weights are gov-

erned by the distribution of the voters' party attachments. The intuitive explanation is as follows. A party that seeks to win an election has to strike a balance between providing benefits to its loyal supporters to get them out to vote and providing benefits to the most loyal supporters of the rival to make sure that they have no reason to vote. Moreover, if a party is unlikely to win, it focuses on its own supporters. If the party is the favorite, by contrast, it puts most weight on minimizing turnout in favor of the opponent. For this reason, a party that is left-leaning in the sense of having more loyal supporters with below average income chooses to avoid high taxes on the rich whenever it is the likely winner of an election. It chooses to propose significant taxes on the rich only if it cannot expect to win. The same logic can explain why a party that is supported to a larger extent by voters with above average income and expects to win chooses, e.g., not to propose to replace a progressive income tax system by a flat tax. It does so only when its chances of winning are low.

Our findings contrast with those of the probabilistic voting models. In these models, also known as *swing voter* models, participation is exogenous. The parties thus do not face a trade-off between maximizing their base and getting their supporters out to vote. All parties care about is to maximize their base, as a larger base translates mechanically into a larger number of votes. Policy proposals are therefore more favorable to voters whose idiosyncratic party preferences are less pronounced. There is no point to cater towards citizens who inelastically vote either for party 1 or for party 2. With endogenous turnout, by contrast, parties also need to pay attention to the incentives of their more loyal supporters to come to the ballot. That is, we show that the key determinant of policies is not the strength of the response of the swing voters but how the party attachments of voters vary with their incomes.

Our findings also contrast with those of the ethical voters models. With endogenous payoffs, we find that political equilibrium policies are symmetric, i.e., both parties propose the same tax policy. This is not compatible with the assumption in the literature on endogenous turnout that there is a choice between two well-defined and distinct political alternatives. It shows that endogenous turnout does not by itself imply that parties have an incentive to differentiate from each other. It is true that such differentiation would increase overall turnout. Parties, however, do not seek to maximize turnout, but the probability of winning the election. This objective makes both parties compromise between mobilizing their own supporters and demobilizing the other party's supporters – with the implication that their proposals are identical in equilibrium.

Our paper also contributes to the literature on optimal welfare-maximizing taxation (Mirrlees (1971), Diamond (1998) and Saez (2001)). Specifically, we address the fundamental question whether welfare-maximizing tax systems have a chance in the political process, or, put differently, whether the political process gives rise to political failures, i.e., to outcomes with undesirable welfare properties. A planner who maximizes a concave social welfare function assigns more weight to the well-being of the poor than to the well-being of the rich. All else equal, marginal tax rates should therefore increase with income. For parties who seek to win an election, the well-being of voters has only instrumental value. We derive conditions on primitives – i.e., the distribution of party preferences, attitudes towards participation in elections and productive abilities – under which parties assign more weight to the well-being of the rich than to the well-being of the poor, with the consequence that political equilibrium tax schedules are incompatible with the maximization of a social welfare function.

Our analysis of political failures provides support for the generalized social welfare weights of Saez and Stantcheva (2016). They argue that tax polices are as if a generalized social welfare function was maximized that may also reflect non-welfarist value judgments or political economy forces. We provide an explicit characterization of the generalized social welfare weights that correspond to the political equilibrium policies in our setup and thereby develop a political economy microfoundation for this idea. This provides in particular a rationale for the empirical finding of the high implicit social welfare weights on the rich in Lockwood and Weinzierl (2016).

Throughout we make use of functional derivatives for an analysis of political equilibrium policies. Piketty (1997) and Saez (2001) introduced the perturbation method as an intuitive approach for the characterization of optimal tax systems. A rigorous formal analysis of this idea requires the use of functional derivatives, see Golosov, Tsyvinski and Werquin (2014). We show that these techniques are also useful for an analysis of strategic interaction. First, a game in which two parties maximize the probability of winning is a zero sum game. Consequently, a pure strategy equilibrium exists if and only if the two parties' proposals are a saddle point of the function that gives party 1's probability of winning. We use functional derivatives to determine the conditions under which this function has a saddle point and the existence of pure strategy equilibria can be ensured. Second, we use functional derivatives to analyze whether the tax policies proposed by different parties are strategic substitutes or complements. Best response problems – one party choosing a tax policy to maximize its

probability of winning, taking as given the other party's proposal – are akin to the planning problems studied in the literature on optimal, welfare-maximizing taxation only if the best response of one party is independent of the other party's proposal, i.e., only if there is a dominant strategy equilibrium.

The literature on probabilistic voting has focused on the special case of idiosyncratic party preferences that are drawn from a uniform distribution. With a uniform distribution, political equilibrium policies are also the maximizers of a utilitarian social welfare function. We show that a uniform distribution, moreover, gives rise to a dominant strategy equilibrium, but only under the assumption that turnout is exogenous. With endogenous turnout, by contrast, the two parties' best response problems are strategically intertwined. An analysis of political competition with endogenous turnout is therefore also technically more demanding than the analysis of a probabilistic voting model with exogenous turnout. We characterize these best responses in closed for an arbitrary policy of the opponent.

Related literature. Our analysis draws on the literature on the paradox of voting, i.e., on the question why people vote even if the chance of affecting the outcome is negligible and going to the ballot box somewhat inconvenient. Downs (1957) and Riker and Ordeshook (1968) are the classical references. Riker and Ordeshook classify theories that give rise to the prediction that rational agents do not participate in large elections as "positive, but not descriptive." Various approaches have since been explored to come up with a model that explains both participation and abstentions in elections, see Feddersen (2004) for a survey. We draw on one strand of this literature due to Harsanyi (1980), Coate and Conlin (2004) and Feddersen and Sandroni (2006). In this approach, voters participate in elections because they feel a civic duty to vote. More specifically, the civic duty is to follow a participation rule that requires an individual to participate if her voting costs are below a cutoff and to abstain otherwise. The cutoff is determined by rule-utilitarian calculus that weighs the social benefits and the social costs of voting. The framework of Coate and Conlin (2004) and Feddersen and Sandroni (2006) not only has plausible comparative statics properties but also studies models with a continuum of voters. We are interested in redistributive income taxation and employ the workhorse of this literature due to Mirrlees (1971). The Mirrleesian model also has a continuum of agents. This makes it convenient to take Coate and Conlin (2004) and Feddersen and Sandroni (2006) as the starting point of our analysis.

We use the Mirrleesian model of non-linear income taxation, except that we do not impose the assumption that the tax policy is chosen by a benevolent social planner. There is a literature that studies political economy forces in connection with non-linear taxation. There is, however, no predecessor that relates differences in participation in the political process among low, middle and high income classes to the non-linear tax policies that arise in a political equilibrium.[2]

We investigate the welfare properties of political equilibrium policies. We follow Saez and Stantcheva (2016) who propose to interpret political economy outcomes using the notion of generalized social welfare weights. A related approach is taken by Bierbrauer and Boyer (2016) who provide an analysis of whether analogues to the fundamental theorems of welfare economics hold in models of political competition. Lockwood and Weinzierl (2016) argue that it is difficult to reconcile the social welfare weights found empirically with those implied by a social welfare function that values redistribution from high to low incomes; taking abstention into account can explain this finding.[3]

The remainder is organized as follows. Section 1 introduces a general setup for an analysis of political competition that connects probabilistic voting with endogenous turnout. The implications of this framework for political equilibrium outcomes are spelled out in Section 2. We characterize political equilibrium tax policies in Section 3. This section contains the main results of the paper. Section 4 clarifies the significance of endogenous turnout by providing a comparison to a model with exogenous turnout. We discuss political failures and equilibrium turnout in Section 5. Section 6 concludes with a discussion of generalized social welfare weights. Unless stated otherwise, proofs are relegated to the Appendix.

---

[2]Acemoglu, Golosov and Tsyvinski (2008) and Acemoglu, Golosov and Tsyvinski (2010) embed dynamic problems of taxation into the political agency model due to Barro (1973) and Ferejohn (1986). Farhi et al. (2012) use a probabilistic voting model in which legislatures lack powers of commitment to study capital taxation. Bierbrauer and Boyer (2016) study Downsian competition when politicians not only propose income tax schedules but also engage in pork-barrel spending. Brett and Weymark (2017) and Roell (2012) embed non-linear income taxation into the citizen-candidate framework that has been developed by Besley and Coate (1997) and Osborne and Slivinski (1996).

[3]See also Weinzierl (2014).

1 Endogenous turnout: a general framework

The voters' preferences. There is competition between two political parties, $J = \{1, 2\}$. Parties choose policies from a set of feasible policies P. Party j's proposal is denoted by p^j. There is a continuum of voters of mass one. Voters differ in their preferences over policies. They evaluate policies according to a utility function $u : P \times \Omega \to \mathbb{R}$, where Ω is a set of possible preference types. Thus, $u(p, \omega)$ is the utility that a type ω-individual realizes under policy p. The cross-section distribution of types is taken to be common knowledge. It is represented by a cumulative distribution function F with density f.

Voters not only have preferences over policy outcomes but also idiosyncratic party preferences. We assume that type ω of voter i votes for party 1 if

$$u(p^1, \omega) \geq u(p^2, \omega) + \varepsilon_i \,,$$

where $\varepsilon_i \in \mathbb{R}$ is a random variable that we refer to as agent i's idiosyncratic preference for party 2. Conditional on ω, party preferences of different voters are independent and identically distributed. The probability that a type ω-individual votes for party 1 is given by $B\left(u\left(p^1, \omega\right) - u\left(p^2, \omega\right) \mid \omega\right)$, where $B(\cdot \mid \omega)$ is the cumulative distribution function that represents the distribution of party preferences among voters of type ω. We denote the corresponding density by $b(\cdot \mid \omega)$.

Special case. We frequently invoke the assumption that idiosyncratic party biases follow a uniform distribution.

Assumption 1. *For any one $\omega \in \Omega$, there are numbers $s^1(\omega) \in (0, 1)$ and $\alpha(\omega) > 0$ so that*

$$B(x \mid \omega) = \begin{cases} 0, & \text{if } x < -\frac{s^1(\omega)}{\alpha(\omega)} \,, \\ s^1(\omega) + \alpha(\omega)\, x, & \text{if } x \in \left[-\frac{s^1(\omega)}{\alpha(\omega)}, \frac{1-s^1(\omega)}{\alpha(\omega)}\right] \,, \\ 1, & \text{if } x > \frac{1-s^1(\omega)}{\alpha(\omega)} \,. \end{cases}$$

Whenever we invoke this assumption, we also assume that the support of the distribution is sufficiently wide so that $B(x \mid \omega) = s^1(\omega) + \alpha(\omega)\, x$, for all values of $x = u\left(p^1, \omega\right) - u\left(p^2, \omega\right)$ that we are concerned with.

Under Assumption 1, if both parties make the same proposal so that $p^1 = p^2$ and

hence $u(p^1, \omega) - u(p^2, \omega) = 0$, then

$$B(u(p^1, \omega) - u(p^2, \omega) \mid \omega) = s^1(\omega) ,$$

i.e., $s^1(\omega)$ is a measure of party 1's strength in the subset of type ω citizens. It gives the party's support among ω-types in a situation where the parties propose the same policies. Analogously, $s^2(\omega) := 1 - s^1(\omega)$ is a measure of party 2's strength. We refer to $s^1(\omega)$ and $s^2(\omega)$ as measuring the *attachment* of type ω-voters to parties 1 and 2, respectively. The parameter $\alpha(\omega)$, by contrast, measures the extra fraction of type ω-voters that swing to party 1 if it deviates from $p^1 = p^2$ and offers more utility to type ω-voters. As will become clear, the uniform distribution is useful in that it allows us to capture in a transparent way the difference between a policy that is targeted at swing voters – i.e., making more attractive offers to voter types ω with a comparatively high value of $\alpha(\omega)$ – and a policy that is targeted towards a party's most loyal voters – i.e., making more attractive offers to voter types ω with a comparatively high value of $s^j(\omega)$.

Votes. Given two policies p^1 and p^2, the mass of supporters of party 1, henceforth referred to as its base, is given by

$$\Gamma^1(p^1, p^2) = \mathbb{E}[B(u(p^1, \omega) - u(p^2, \omega) \mid \omega)]$$

and the base of party 2 equals $\Gamma^2(p^1, p^2) = 1 - \Gamma^1(p^1, p^2)$, where the expectation operator $\mathbb{E}$ always indicates the computation of a population average with respect to different types ω. The base of each party is split into three groups, a group that always votes, a group that always abstains, and a group of voters whose voting decision follows from a rule-utilitarian calculation. The previous literature often refers to this last group as *ethical* voters. We denote by $\tilde{q}^{jv}(\omega)$ the fraction of definite voters among the type ω supporters of party j, by $\tilde{q}^{ja}(\omega)$ the fraction of definite abstainers and by $\tilde{q}^{ju}(\omega)$ the fraction of rule-utilitarian or ethical supporters. Obviously, $\tilde{q}^{jv}(\omega) + \tilde{q}^{ja}(\omega) + \tilde{q}^{ju}(\omega) = 1$. We assume that these are random quantities both from the perspective of parties when choosing platforms and from the perspective of voters when choosing whether or not to vote. We write $\tilde{q}^j = \{\tilde{q}^{jv}(\omega), \tilde{q}^{ja}(\omega), \tilde{q}^{ju}(\omega)\}_{\omega \in \Omega}$ for the collection of random variables that refer to party j. We denote the expected value of the random variable $\tilde{q}^{ju}(\omega)$ by $\bar{q}^{ju}(\omega)$.

8

Denote by σ^j the fraction of ethical voters among party j who turn out to vote. The total number of votes for party 1 is then a random variable equal to

$$\tilde{V}^1(p^1, p^2, \sigma^1, \tilde{q}^1) = \mathbb{E}[(\tilde{q}^{1v}(\omega) + \sigma^1 \, \tilde{q}^{1u}(\omega))B(u(p^1, \omega) - u(p^2, \omega) \mid \omega)] \ .$$

Analogously, the total number of votes for party 2 equals

$$\tilde{V}^2(p^1, p^2, \sigma^2, \tilde{q}^2) = \mathbb{E}[(\tilde{q}^{2v}(\omega) + \sigma^2 \, \tilde{q}^{2u}(\omega))(1 - B(u(p^1, \omega) - u(p^2, \omega) \mid \omega))] \ .$$

Turnout. The rule-utilitarian supporters of any one party j adhere to a rule for participation in the election that maximizes the expected utility of the supporters of party j, taking the costs of voting into account. As a consequence, turnout depends on the parties' policy proposals.

In the Appendix, we show that this problem of the rule-utilitarian supporters of party 1 admits the following representation: Choose $\sigma^1 \in [0, 1]$ to maximize

$$\pi^1(p^1, p^2, \sigma^1, \sigma^2) \, W^1(p^1, p^2) - \kappa \, \sigma^1 \, \bar{\Gamma}^{1u}(p^1, p^2) \ , \tag{1}$$

where $\pi^1(p^1, p^2, \sigma^1, \sigma^2)$ is the probability that party 1 wins, $W^1(p^1, p^2)$ is the welfare gain realized by the supporters of party 1 if their party 1 wins relative to the utility they would get if party 2 won, and $\kappa \, \sigma^1 \, \bar{\Gamma}^{1u}(p^1, p^2)$ is a measure of expected voting costs. We now explain these expressions in more detail.

The welfare gain satisfies

$$W^1(p^1, p^2) = \mathbb{E}[G_W^1(u(p^1, \omega) - u(p^2, \omega) \mid \omega)] \ ,$$

where

$$G_W^1(x \mid \omega) := \int (x - \varepsilon_i) \, \mathbf{1}\{\varepsilon_i \leq x\} \, b(\varepsilon_i \mid \omega) \, d\varepsilon_i \ = \ B(x \mid \omega) \, x \ - \ \mathcal{E}^1(x \mid \omega) \ ,$$

with $\mathcal{E}^1(x \mid \omega) = \int_{-\infty}^{x} \varepsilon_i \, b(\varepsilon_i \mid \omega) \, d\varepsilon_i$, is the gain that type ω supporters of party 1 realize if party 1 wins and the utility difference $u(p^1, \omega) - u(p^2, \omega)$ is equal to x. This gain has two components: $B(u(p^1, \omega) - u(p^2, \omega) \mid \omega)(u(p^1, \omega) - u(p^2, \omega))$ is the welfare gain realized by the supporters of party 1 if party 1's platform is implemented instead of party 2's platform, and $\mathcal{E}^1(u(p^1, \omega) - u(p^2, \omega) \mid \omega)$ is the sum of the gains (or losses) that the supporters of party 1 realize because of their idiosyncratic party preference.

9

We denote per capita voting costs by κ and assume them to be the same for all rule-utilitarian voters. We interpret σ^1 as the rule-utilitarian voter's probability of voting so that expected voting costs per capita are equal to $k(\sigma^1) = \kappa\,\sigma^1$.[4] Total expected voting costs are then given by $\kappa\,\sigma^1\,\bar{\Gamma}^{1u}(p^1, p^2)$ where

$$\bar{\Gamma}^{1u}(p^1, p^2) = \mathbb{E}[\bar{q}^{1u}(\omega)\,B(u(p^1, \omega) - u(p^2, \omega) \mid \omega)]$$

is the expected value of the mass of rule-utilitarian supporters of party 1.[5]

The analogous problem for the ethical supporters of party 2 is to choose $\sigma^2 \in [0, 1]$ to maximize

$$(1 - \pi^1(p^1, p^2, \sigma^1, \sigma^2))\,W^2(p^1, p^2) - \kappa\,\sigma^2\,\bar{\Gamma}^{2u}(p^1, p^2)\,, \tag{2}$$

where $\bar{\Gamma}^{2u}(p^1, p^2) = E[\bar{q}^{2u}(\omega)\,(1 - B(u(p^1, \omega) - u(p^2, \omega) \mid \omega))]$ is the mass of rule-utilitarian supporters of party 2, $W^2(p^1, p^2) = E[G_W^2(u(p^2, \omega) - u(p^1, \omega) \mid \omega)]$ is the welfare gain realized by the supporters of party 2 if their party wins with, for any $x \in \mathbb{R}$,

$$G_W^2(x \mid \omega) := (1 - B(-x \mid \omega))\,x\; + \mathcal{E}^2(-x \mid \omega)\,,$$

where $\mathcal{E}^2(-x \mid \omega) = \int_{-x}^{\infty} \varepsilon_i\,b(\varepsilon_i \mid \omega)\,d\varepsilon_i$. Lemma 8 in the Appendix shows what the benefit functions W^1 and W^2 are equal to for the special case of uniformly distributed party biases.

Equilibrium turnout. Given p^1 and p^2, a pure strategy equilibrium of the turnout game can be characterized by a pair of participation thresholds $(\sigma^{1*}(p^1, p^2), \sigma^{2*}(p^1, p^2))$ that are mutually best responses.

Best responses and equilibrium policies. We assume that parties seek to maximize their probability of winning. We will often take the policy proposal of party 2 as fixed and consider the best-response problem of party 1 so that, for given p^2, party

---

[4]Feddersen and Sandroni (2006) and Coate and Conlin (2004) assume, by contrast, that any one voter i has voting costs equal to $\kappa\,\sigma_i$, where σ_i is an idiosyncratic component that is i.i.d. across voters and uniformly distributed over the unit interval. For them, σ^1 is a cutoff so that all ethical supporters of party 1 with $\sigma_i \le \sigma^1$ are turning out to vote. In this case, $k(\sigma^1) = \kappa \int_0^{\sigma^1} \sigma_i\,di = \kappa\,\frac{1}{2}(\sigma^1)^2$. For our purposes, a linear cost function simplifies the exposition. In part B of the Appendix we clarify how the analysis would have to be modified in the case of a quadratic cost function.

[5]The costs of those who always vote do not appear here because they are an additive constant that does not affect the solution of the optimization problem.

1 chooses p^1 so as to maximize

$$\bar{\pi}^1(p^1, p^2) := \pi^1(p^1, p^2, \sigma^{1*}(p^1, p^2), \sigma^{2*}(p^1, p^2)) \,.$$

The characterization of best responses allows us to clarify the logic of the two parties' strategic interaction. For instance, it allows us to investigate whether the parties' policy proposals are strategic substitutes or complements. It can also serve as an input for an analysis of equilibrium policies. We distinguish between Nash equilibrium policies and the policies that arise in a subgame perfect equilibrium. With both solution concepts, the participation probabilities need to constitute an equilibrium of the participation subgame, i.e.,

$$\sigma^1 = \sigma^{1*}(p^1, p^2) \quad \text{and} \quad \sigma^2 = \sigma^{2*}(p^1, p^2) \,.$$

A pair of subgame perfect equilibrium policies (p^1, p^2) then satisfies

$$\bar{\pi}^1(p^1, p^2) \geq \bar{\pi}^1(\hat{p}^1, p^2) \,, \text{ for all } \hat{p}^1 \in P \quad \text{and} \quad \bar{\pi}^1(p^1, p^2) \leq \bar{\pi}^1(p^1, \hat{p}^2) \,, \text{ for all } \hat{p}^2 \in P \,.$$

A pair of Nash equilibrium policies, by contrast, is such that

$$\pi^1(p^1, p^2, \sigma^1, \sigma^2) \geq \pi^1(\hat{p}^1, p^2, \sigma^1, \sigma^2) \,, \text{ for all } \hat{p}^1 \in P \,,$$

and

$$\pi^1(p^1, p^2, \sigma^1, \sigma^2) \leq \pi^1(p^1, \hat{p}^2, \sigma^1, \sigma^2) \,, \quad \text{ for all } \hat{p}^2 \in P \,.$$

For (p^1, p^2) to be a Nash equilibrium, it is required that the parties correctly anticipate turnout and choose polices that maximize their winning probabilities, taking this turnout as exogenously given. In a subgame perfect equilibrium, by contrast, parties understand that alternative policy choices would also affect the equilibrium of the participation subgame and take this repercussion into account. We are mainly interested in subgame perfect equilibria that are interior, i.e., which are such that turnout responds at the margin to changes in proposed policies. The characterization of Nash equilibria is also of interest, however. First, a non-interior subgame perfect equilibrium forces parties to find best replies to turnout rates that are locally unaffected by changes in the platforms that the parties propose. Nash equilibrium policies describe what parties do if turnout does not respond to proposed policies. Second,

we show that Nash equilibrium policies are also similar to the equilibrium policies in a probabilistic voting model with exogenous turnout, a useful benchmark.

Remarks on equilibrium existence. The setup introduced so far is very flexible and nests prominent models of two-party competition. Models of Downsian competition and probabilistic voting models usually impose the assumption that all individuals always vote. Here, this amounts to assuming that, for all parties j and all voter types ω, $\tilde{q}^{jv}(\omega) = 1$ and hence, $\tilde{q}^{ja}(\omega) = \tilde{q}^{ju}(\omega) = 0$ with probability 1. These models differ, however, in their treatment of party preferences. They play no role in Downsian models so that for all i, $\varepsilon_i = 0$ with probability 1 and hence, for any ω,

$$
B(u(p^1, \omega) - u(p^2, \omega) \mid \omega) = \left\{
\begin{array}{ll}
0, & \text{if } u(p^1, \omega) < u(p^2, \omega) , \\
\frac{1}{2}, & \text{if } u(p^1, \omega) = u(p^2, \omega) , \\
1, & \text{if } u(p^1, \omega) > u(p^2, \omega) .
\end{array}
\right.
$$

As is well-known, with a multi-dimensional policy space, pure strategy equilibria typically fail to exist in Downsian models of competition. Our analysis below focuses on a particular multi-dimensional policy domain, the set of non-linear income tax policies. Since our framework nests the Downsian model as a special case, the existence of pure strategy equilibria cannot be taken for granted.

Appropriate assumptions on the distribution of party preferences ensure existence in probabilistic voting models with exogenous turnout. As will become clear below, Nash equilibria typically exist in our framework under the same conditions as in conventional probabilistic voting models.

The models on endogenous participation in large elections assume that for all voter types ω, $\tilde{q}^{jv}(\omega) = 0$ and hence, $\tilde{q}^{ju}(\omega) + \tilde{q}^{ja}(\omega) = 1$ with probability 1. An implication is that a party's probability of winning is equal to the probability of attracting more rule-utilitarian voters than the opponent. Below, we provide conditions on primitives that ensure the existence of a subgame-perfect equilibrium in this setting.

2   An equilibrium characterization

We provide an equilibrium characterization that allows to study specific policies in the context of our model. We later use these insights for an analysis of political competition over non-linear tax systems. As the first main result, we demonstrate in

this section that parties face a tradeoff between maximizing their turnout advantage and the size of their base.

From now on, we follow Coate and Conlin (2004) and Feddersen and Sandroni (2006) and assume that there are no voters who come to the ballot irrespectively of how high their voting costs are, i.e., every citizen is either of the rule-utilitarian or of the abstaining type.

Assumption 2. *For all* ω, $\tilde{q}^{1v}(\omega) = 0$ *and* $\tilde{q}^{2v}(\omega) = 0$ *with probability 1.*

Harsanyi (1980) also allowed for voters who feel a civic duty to vote but do not apply the rule-utilitarian logic. In the Appendix, we present a version of our model that includes such voters and gives rise to an equilibrium analysis that is equivalent to the one developed in the body of the text.

Lemma 1. *Suppose that Assumption 2 holds. Take the parties' proposals* p^1 *and* p^2 *as given. If the equilibrium of the turnout subgame* $(\sigma^{1*}(p^1, p^2), \sigma^{2*}(p^1, p^2))$ *satisfies the first order conditions of the optimization problems (1) and (2), then*

$$\chi^*(p^1, p^2) := \frac{\sigma^{1*}(p^1, p^2)}{\sigma^{2*}(p^1, p^2)} = \frac{W^1(p^1, p^2)/\bar{\Gamma}^{1u}(p^1, p^2)}{W^2(p^1, p^2)/\bar{\Gamma}^{2u}(p^1, p^2)} \ . \tag{3}$$

The left-hand side of this equation is a measure of relative turnout: The larger χ^*, the larger is the set of voters who turn out to vote for party 1 relative to the set of voters who turn out to vote for party 2. The right hand side compares two cost-benefit ratios; e.g., for the supporters of party 1, $W^1/\kappa \ \bar{\Gamma}^{1u}$ relates the benefit that comes with a victory of their preferred party to the expected cost of participating in the election. According to equation (3), relative turnout depends on the ratio of these cost-benefit-ratios.

We now impose additional structure that allows us to provide a clear exposition of the parties' objectives, while still taking into account the equilibrium in the turnout subgame.

Assumption 3.

i) A type ω *supporter of party 1 is as likely to be of the ethical type as a type* ω *supporter of party 2. Formally,* $\bar{q}^{u1}(\omega) = \bar{q}^{u2}(\omega) =: \bar{q}^u(\omega)$.

ii) There is uncertainty about the fraction of type ω *supporters of either party who are of the ethical type. Specifically, for party 1, there is a random variable* η^1 *with*

support bounded from above by 1 and mean 0 so that $\tilde{q}^{u1}(\omega) = \bar{q}^u(\omega)(1 - \eta^1)$, for all ω. For party 2, there is another random variable η^2 so that $\tilde{q}^{u2}(\omega) = \bar{q}^u(\omega)(1 - \eta^2)$, for all ω.

Part i) of Assumption 3 implies that

$$\bar{\Gamma}^{1u}(p^1, p^2) = E[\bar{q}^u(\omega)] - \bar{\Gamma}^{2u}(p^1, p^2) . \tag{4}$$

According to this equation, in expectation, the contributions of rule-utilitarian voters to the two parties' bases add up to a constant. Hence, a change in the proposed policies that increases, say, the base for party 1, translates, one for one, into a decrease of party 2's base. According to part ii) the shares of rule-utilitarian supporters in party j is a random variable and its deviations from its mean are driven by a random variable η^j. As a consequence, there is uncertainty whether party 1 or party 2 will be the winner of the election.[6]

Lemma 2. *Under Assumptions 2 and 3,*

$$\bar{\pi}^1(p^1, p^2) = \mathcal{P}\left(\chi^*(p^1, p^2) \frac{\bar{\Gamma}^{1u}(p^1, p^2)}{\bar{\Gamma}^{2u}(p^1, p^2)}\right) , \tag{5}$$

where $\mathcal{P}$ is the cdf of the random variable $\frac{1-\eta^2}{1-\eta^1}$.

According to this Lemma, the probability that party 1 wins the election is a non-decreasing function of $\chi^*(p^1, p^2) \frac{\bar{\Gamma}^{1u}(p^1,p^2)}{\bar{\Gamma}^{2u}(p^1,p^2)}$. Therefore, party 1's objective is to maximize this expression and that party 2's objective is to minimize it. Using equation (4) we can write this expression also as

$$\chi^*(p^1, p^2)\, M(\bar{\Gamma}^{1u}(p^1, p^2)) , \quad \text{where} \quad M(\bar{\Gamma}^{1u}(p^1, p^2)) = \frac{\bar{\Gamma}^{1u}(p^1, p^2)}{E[\bar{q}^u(\omega)] - \bar{\Gamma}^{1u}(p^1, p^2)} \tag{6}$$

is an increasing function of $\bar{\Gamma}^{1u}(p^1, p^2)$).

Equation (6) reveals that parties face a trade-off between maximizing their base and getting their supporters out to vote. If turnout were exogenous, party 1 would

---

[6]We are assuming here that voting types (i.e., whether a given individual is a rule-utilitarian voter or an always-abstaining) are drawn after policies have been proposed, i.e., after individuals have been sorted to the parties' bases (as a function of their payoff type ω and their preference type ε). This allows for the possibility that voting behaviors differ across parties. For instance, party 1 may have good luck in the sense of having many rule-utilitarian supporters while party 2 has bad luck, many abstaining supporters.

simply focus on maximizing its base, $\bar{\Gamma}^{1u}(p^1, p^2)$. If the base were exogenously given, party 1 would maximize its turnout advantage $\chi^*(p^1, p^2)$. With both endogenous turnout and an endogenous base, party 1 faces a trade-off between maximizing the number of supporters and maximizing turnout.

To illustrate this trade-off suppose that p^{1*} is a best response of party 1 to the policy p^2 proposed by party 2. Also suppose that the policy space P is such that local deviations from p^{1*} are well defined. Now consider a deviation p^1 from p^{1*} that takes the form $p^1 = p^{1*} + \tau^1 h^1$, where h^1 is a function that describes the direction of the deviation from p^{1*} and τ^1 is a scalar that measures the size of the deviation. The deviation induces a payoff for party 1 equal to

$$\chi^*(p^{1*} + \tau^1 h^1, p^2) \, M(\bar{\Gamma}^{1u}(p^{1*} + \tau^1 h^1, p^2)) \ .$$

Now, if p^{1*} is a best response, then it must be true that this expression is, for any feasible deviation h^1, maximized by choosing the parameter $\tau^1 = 0$. The corresponding first order condition is

$$\chi^*_{\tau^1}(p^{1*}, p^2)M(\bar{\Gamma}^{1u}(p^{1*}, p^2)) + \chi^*(p^{1*}, p^2)M'(\bar{\Gamma}^{1u}(p^{1*}, p^2))\bar{\Gamma}^{1u}_{\tau^1}(p^{1*}, p^2) = 0 \ , \quad (7)$$

where we let $\chi^*_{\tau^1}(p^{1*}, p^2) := \frac{d}{d\tau_1} \chi^*(p^{1*} + \tau^1 h^1, p^2)|_{\tau^1 = 0}$ and analogously for $\bar{\Gamma}^{1u}_{\tau^1}(p^{1*}, p^2)$.

To interpret this condition, suppose that P is the unit interval and that p^{1*} lies in the interior of it. (The logic of the argument applies for any policy space P and any direction h^1 for which the directional derivative is well-defined.) Models of linear taxation, for instance, typically use this policy space.[7] In these models, tax revenues are used for redistribution so that higher tax rates are associated with more redistributive or leftist policies. With P as the unit interval there is only one feasible direction of policy reform, so that h^1 is simply a constant. Now suppose that party 1's best response is such that its base could be increased by means of a more leftist policy as $\bar{\Gamma}^{1u}_{\tau^1}(p^{1*}, p^2) > 0$. Then, the first order condition (7) implies that $\chi^*_{\tau^1}(p^1, p^2) < 0$. Thus, a party whose base asks for more leftist policies sacrifices turnout (and increases the turnout of the other party) by moving to the left. The following Proposition summarizes.

Proposition 1. *If p^{1*} is a best response for party 1 that admits a characterization*

_____

[7]Prominent examples are Roberts (1977), Meltzer and Richard (1981) or Alesina and Angeletos (2005).

15

by the first order conditions in (7), then, for any feasible deviation h^1,

$$\chi_{\tau^1}^*(p^{1*}, p^2) < 0 \quad \text{if and only if} \quad \bar{\Gamma}_{\tau^1}^{1u}(p^{1*}, p^2) > 0 \, .$$

Proposition 1 shows that, with an interior equilibrium of the turnout subgame, parties run into a tradeoff that forces them to compromise the implications of their policy proposals for their turnout advantage with the implications of their policy proposals for the size of their base. What is good for the base is bad for turnout and vice versa. The Proposition highlights that this tradeoff is central to our analysis. It does not yet illuminate, however, how parties resolve that tradeoff. Via equation (3), $\chi^*(p^1, p^2)$ is an expression that also depends on $\bar{\Gamma}^{1u}(p^1, p^2)$ and, in addition, on $W^1(p^1, p^2)$ and $W^2(p^1, p^2)$. The following Proposition is more helpful in this respect. It shows in part i) that the analysis of subgame perfect equilibrium policies can be confined to the analysis of the saddle points of the function $\psi(p^1, p^2) := \frac{W^1(p^1, p^2)}{W^2(p^1, p^2)}$. We will also be interested in comparing subgame perfect equilibrium policies to Nash equilibrium policies that do not take the endogeneity of turnout into account. By part ii) of the Proposition, saddle points of the the function $\bar{\Gamma}^{1u}(p^1, p^2)$ correspond to Nash equilibrium policies.

Proposition 2.

i) Under Assumptions 2 and 3, if (p^1, p^2) is a pair of interior subgame perfect equilibrium policies, then it is a saddle point of the function $\psi(p^1, p^2) := \frac{W^1(p^1, p^2)}{W^2(p^1, p^2)}$.

ii) Under Assumptions 2 and 3, if (p^1, p^2) is a pair of Nash equilibrium policies, then it is a saddle point of the function $\bar{\Gamma}^{1u}(p^1, p^2)$.

According to part i) of the Proposition, if we want to study how party 1 resolves the tradeoff between turnout and base, we can as well assume that party 1 seeks to maximize $\psi(p^1, p^2) := \frac{W^1(p^1, p^2)}{W^2(p^1, p^2)}$, i.e., an expression where the base no longer appears. The reason is that the influence of the base on χ^* that appears in equation (3) exactly cancels with the expression $\frac{\bar{\Gamma}^{1u}(p^1, p^2)}{\bar{\Gamma}^{2u}(p^1, p^2)}$ that also appears on the right hand side of equation (5).[8]

---

[8]As we show in part C of the Appendix, this would be different with a quadratic cost function. Both with linear and with quadratic costs, however, subgame perfect equilibria differ from Nash equilibria because they require to take into account $\psi(p^1, p^2) := \frac{W^1(p^1, p^2)}{W^2(p^1, p^2)}$.

3 Subgame perfect equilibrium and non-linear income taxation

It is a stylized fact that individuals with high incomes are more likely to participate in an election than individuals with low incomes. Studying how the endogeneity of turnout affects equilibrium policies therefore is most interesting for polices that affect the rich and the poor differently. Our main focus is thus on the tax and transfer system – as opposed to, say, infrastructure investment or national defense (which we cover in Appendix E). Thereby we can also shed light on the following important question. Papers that provide calibrations of welfare-maximizing tax systems often suggest that top tax rates should be much higher than currently observed in modern democracies, see e.g. Diamond (1998), Saez (2001). Having higher taxes, say, on the top 10 percent and using the proceeds in a way that benefits the population at large should also be politically feasible. A vast majority would benefit from such a reform, see e.g. Bierbrauer and Boyer (2017). So why are the observed taxes on the rich lower than the theory predicts? We will show that a political economy analysis that takes the endogeneity of turnout into account can offer an explanation.

3.1 Non-linear income taxation

From now on we take the policy domain to consist of redistributive tax policies. Individuals value after-tax income or private goods consumption denoted by c and incur a disutility from the productive effort that is needed to generate pre-tax income y. The preferences of a type ω individual over (c, y)-pairs are taken to be quasi-linear in private goods consumption and represented by a utility function $c - v(y, \omega)$. The function v measures the cost of productive effort and satisfies $v_1 > 0$ and $v_{11} > 0$. Moreover, preferences have the Spence-Mirrlees single crossing property so that absolute and marginal effort costs are decreasing in the individual's type, $v_2 < 0$ and $v_{12} < 0$. This property implies that higher types will end up having higher (before and after tax) incomes than lower types. Thus, high types are "the rich" and low types are "the poor".

Party j proposes a function $T^j : \mathbb{R}_+ \to \mathbb{R}$ that assigns a tax payment to any level of pre-tax-income y. Voters then choose c and y so as to maximize utility subject to $c = y - T^j(y)$. Admissible tax functions are those that generate non-negative tax revenue. We use a mechanism design approach to obtain a characterization of

admissible tax systems, see part D of the Appendix for details. An admissible tax system T^j can be equivalently represented by a non-decreasing earnings function $y^j : \Omega \to \mathbb{R}_+$. Under such an earnings function, the utility realized by any one type-ω individual can be written as a sum of two terms, the minimal level of utility that is realized by the "poorest" type and the extra utility realized by higher types. More formally, the utility that a citizen of type ω realizes under earnings function y^j is given by

$$u^j(\omega) = \underline{u}^j - \int_{\underline{\omega}}^{\omega} v_2(y^j(z), z) \ dz \ , \tag{8}$$

where $\underline{u}^j$ is the poorest type's utility and $-\int_{\underline{\omega}}^{\omega} v_2(y^j(z), z)$ is the *information rent* realized by a higher type $\omega > \underline{\omega}$. This terminology reflects that private information on types is the impediment to first-best redistribution. The poorest type's utility equals the *virtual surplus* that is associated with an earnings function y^j. The virtual surplus is a surplus measure that takes into account the information rents that tax-payers realize and which reduces what is available for the lowest type. To arrive at the virtual surplus, the surplus of aggregate output over costs of effort $s(y^j) := \mathbb{E}\left[y^j(\omega) - v(y^j(\omega), \omega)\right]$ is reduced by the aggregate information rent $-\mathbb{E}[\int_{\underline{\omega}}^{\omega} v_2(y^j(z), z) \ dz] = -\mathbb{E}[\frac{1-F(\omega)}{f(\omega)} v_2(y^j(\omega), \omega)]$, where the equality follows from an integration by parts. Thus,

$$\underline{u}^j = s_v(y^j) = \mathbb{E}\left[y^j(\omega) - v(y^j(\omega), \omega) + \frac{1 - F(\omega)}{f(\omega)} v_2(y^j(\omega), \omega)\right] \ . \tag{9}$$

To summarize, for the application of non-linear income taxation, the policy space is the set of all non-decreasing earnings function. Any such function generates a payoff profile that is characterized by equations (8) and (9). We are particularly interested in the marginal tax rates that are associated with the tax systems that the parties propose. To get from a monotonic earnings function to the associated tax schedule we use the first order condition of the utility-maximization problem that individuals face in the presence of this tax system. If tax system T^j implements earnings function y^j, then

$$1 - T^{j\prime}(y^j(\omega)) = v_1(y^j(\omega), \omega) \ .$$

Hence, $1 - v_1(y^j(\omega), \omega)$ is the marginal tax rate that type ω agents face.

3.2 Subgame perfect equilibrium

We now turn to the tax functions that constitute a subgame perfect equilibrium. By Proposition 2, subgame perfect equilibrium policies are associated with a saddle point of the function $\psi(y^1, y^2) := \frac{W^1(y^1, y^2)}{W^2(y^1, y^2)}$. In the following we provide a characterization. We begin with an analysis of best responses. As will become clear, a best response of party 1 is a weighted average of the policy that maximizes the welfare gains that the supporters of party 1 realize if their party wins, $W^1(y^1, y^2)$, and the policy that minimizes the gains of the supporters of party 2 if their party wins, $W^2(y^1, y^2)$. As a preliminary step, we consider two auxiliary best response problems for party 1: first, the problem to maximize what is at the stake for the the supporters of party 1; and second, the problem to minimize what is at stake for the supporters of party 2.

3.2.1 Mobilizing the supporters of party 1, demobilizing the supporters of party 2

We analyze the problem to choose y^1 so as to maximize $W^1(y^1, y^2)$.[9] The following notation will prove helpful. Remember that $W^1(y^1, y^2) = \mathbb{E}[G_W^1(u^1(\omega) - u^2(\omega) \mid \omega)]$. We denote the derivative of the function $G_W^1(\cdot \mid \omega)$ by $g_W^1(\cdot \mid \omega)$ and write

$$
\begin{aligned}
\bar{g}_W^1(\omega \mid y^1, y^2) & := & \mathbb{E}[g_W^1(u^1(\omega') - u^2(\omega') \mid \omega') \mid \omega' \geq \omega] \\
& = & \int_\omega^{\bar{\omega}} g_W^1(u^1(\omega') - u^2(\omega') \mid \omega') \frac{f(\omega')}{1 - F(\omega)} d\omega'
\end{aligned}
$$

for the average value of $g_W^1(u^1(\omega') - u^2(\omega') \mid \omega')$ among individuals with a type ω' above some cutoff ω. To interpret these expressions, suppose that party 1 offers slightly more utility to type ω' individuals. Then $g_W^1(u^1(\omega') - u^2(\omega') \mid \omega')$ measures the extra gain that type ω'-supporters of party 1 realize in the event that party 1 wins rather party 2. Therefore, $\bar{g}_W^1(\omega \mid T^1, T^2)$ is the gain that party 1 can generate by offering all agents with types above ω slightly more utility. The gain that party 1 can generate by slightly raising everybody's utility is given by $\bar{g}_W^1(\underline{\omega} \mid y^1, y^2)$ and the

---

[9]We thereby provide a characterization of the solution to a relaxed problem , as opposed to the full problem to maximize $W^1(y^1, y^2)$ subject to the constraint that y^1 is a non-decreasing function. Obviously, if the solution to the relaxed problem is non-decreasing then it is also a solution to the full problem. Otherwise, the solution of the full problem will give rise to bunching. While it is well known how the analysis would have to be modified if bunching is an issue, see e.g. Hellwig (2007) or Brett and Weymark (2017), the trade-offs that shape best responses are, however, more easily exposed when focusing on the relaxed problem.

ratio

$$\mathcal{G}_W^1(\omega \mid y^1, y^2) := \frac{\bar{g}_W^1(\omega \mid y^1, y^2)}{\bar{g}_W^1(\underline{\omega} \mid y^1, y^2)}$$

relates the gain from making everybody with a type above ω better off to the gain from making everybody better off.

Lemma 3. *Given y^2, the solution to $\max_{y^1} W^1(y^1, y^2)$ is such that, for all ω,*

$$\frac{T^{1'}(y^1(\omega))}{1 - T^{1'}(y^1(\omega))} = -\frac{1 - F(\omega)}{f(\omega)} \left(1 - \mathcal{G}_W^1(\omega \mid y^1, y^2)\right) \frac{v_{21}(y^1(\omega), \omega))}{v_1(y^1(\omega), \omega)} . \qquad (10)$$

The left hand side of equation (10) is an increasing function of the marginal tax rate on the income of type ω individuals, $T^{1'}(y^1(\omega))$. According to the Proposition, this marginal tax is lower the larger is $\mathcal{G}_W^1(\omega \mid y^1, y^2)$, i.e., the larger are the gains that can be generated by making the proposed policy more attractive to individuals with incomes above $y^1(\omega)$. Two other two terms appear on the right hand side of equation (10). The first term, $\frac{1-F(\omega)}{f(\omega)}$, is the inverse of the hazard rate of the type distribution. This term relates the number of people who pay higher taxes after a slight increase of $T^{1'}(y^1(\omega))$, $1 - F(\omega)$, to the mass of people whose incentives to exert effort are worsening after such an increase, $f(\omega)$. Thus, marginal tax rates should be lower the less favorable is the ratio of people who pay more taxes to people whose effort is discouraged. The second term, $-\frac{v_{21}(y^1(\omega), \omega))}{v_1(y^1(\omega), \omega)}$, measures the strength of the behavioral response of those whose effort is discouraged. It is related to the inverse of the elasticity of labor supply with respect to changes in the net wage.[10] The marginal tax rate should be lower the larger is this elasticity.

The optimality condition in (10) could also be derived on the assumption that party 1 seeks to maximize $s_v(y^1) - \mathbb{E}\left[\mathcal{G}(\omega) \, v_2(y^1(\omega), \omega)\right]$ for $\mathcal{G}(\omega) = \mathcal{G}_W^1(\omega \mid y^1, y^2)$, i.e., a utilitarian welfare objective with particular weights on the rents that are realized by types $\omega > \underline{\omega}$. In this maximization problem the weighting function $\mathcal{G}$ is exogenously fixed, albeit at the level that is induced by y^2 and party 1's best response. Thus, equation (10) is akin to Diamond (1998)'s formula in that it provides

---

[10]To see this, interpret ω as a wage rate and $h = \frac{y}{\omega}$ as hours worked. Suppose that the cost of productive effort depends only on hours worked, so that there is an increasing and convex cost function $\tilde{v}$ with $v(y, \omega) = \tilde{v}\left(\frac{y}{\omega}\right)$. Then, $\frac{v_{21}(y,\omega)}{v_1(y,\omega)} = -\frac{1}{\omega}\left(1 + \frac{\tilde{v}''\left(\frac{y}{\omega}\right)\frac{y}{\omega}}{\tilde{v}'\left(\frac{y}{\omega}\right)}\right)$, where $\frac{\tilde{v}''\left(\frac{y}{\omega}\right)\frac{y}{\omega}}{\tilde{v}'\left(\frac{y}{\omega}\right)}$ is the inverse of the elasticity of hours worked with respect to the wage rate. In particular, if the function $\tilde{v}$ is isoelastic so that $\tilde{v}\left(\frac{y}{\omega}\right) = \left(\frac{y}{\omega}\right)^{1+\varepsilon}$, then $\frac{\tilde{v}''\left(\frac{y}{\omega}\right)\frac{y}{\omega}}{\tilde{v}'\left(\frac{y}{\omega}\right)} = \frac{1}{\varepsilon}$.

a characterization of optimal marginal tax rates that distinguishes three components: the inverse hazard rate associated with the income distribution, a weighting of information rents, and an inverse elasticity term. Note that the best response of party 1 is only implicitly defined as the function y^1 appears both on the left-hand side and the right hand side of equation (10). Moreover, the function y^2 also appears on the right hand side of (10). Hence the weighting of rents by party 1 also depends on the policy that is proposed by party 2. This strategic interdependence distinguishes the analysis of political equilibrium policies from the analysis of the policies that a benevolent social planner would choose.

The following Lemma describes the solution to another auxiliary problem for party 1, namely the problem to choose policy with the objective to minimize what is at stake for the supporters of party 2. We omit a proof and discussion of the Lemma as it would involve only a straightforward adjustment of those of Lemma 3. The Lemma involves a weighting function $\mathcal{G}_W^2$ for information rents that is derived from $W^2(y^1, y^2) = \mathbb{E}[G_W^2(u^2(\omega) - u^1(\omega) \mid \omega)]$ along the same lines as $\mathcal{G}_W^1$ is derived from $W^1(y^1, y^2)$.

Lemma 4. *Given y^2, the solution to $\min_{y^1} W^2(y^1, y^2)$ is such that, for all ω,*

$$\frac{T^{1\prime}(y^1(\omega))}{1 - T^{1\prime}(y^1(\omega))} = -\frac{1 - F(\omega)}{f(\omega)} \left(1 - \mathcal{G}_W^2(\omega \mid y^1, y^2)\right) \frac{v_{21}(y^1(\omega), \omega)}{v_1(y^1(\omega), \omega)} . \qquad (11)$$

3.2.2 The main result: Best responses and equilibrium policies

We now turn to party 1's best response problem of interest and then provide an equilibrium characterization. Let y^2 be an arbitrary, but given function, possibly equal to y^{2*}. Given y^2 we analyze the problem to choose y^1 with the objective to maximize $\psi(y^1, y^2) := \frac{W^1(y^1, y^2)}{W^2(y^1, y^2)}$. Suppose that y^{1*} is a solution to that problem. Then, it must also be the case that $\tau^1 = 0$ solves the problem to choose a scalar τ^1 with the objective to maximize

$$\psi(y^{1*} + \tau^1 h^1, y^2) = \frac{W^1(y^{1*} + \tau^1 h^1, y^2)}{W^2(y^{1*} + \tau^1 h^1, y^2)} .$$

21

for any given but arbitrary function h^1. That is, we can characterize y^{1*} by the requirement that, for $\tau^1 = 0$,

$$\left.\frac{d}{d\tau^1}\psi(y^{1*} + \tau^1 \, h^1, y^2)\right|_{\tau^1=0} = 0 \, ,$$

or, equivalently, that

$$\frac{W^1_{\tau^1}(y^{1*}, y^2)}{W^1(y^{1*}, y^2)} - \frac{W^2_{\tau^1}(y^{1*}, y^2)}{W^2(y^{1*}, y^2)} = 0 \, .$$

This expression shows that the best response trades off two considerations. First, mobilizing the supporters of party 1, as represented by the term $W^1_{\tau^1}(y^{1*}, y^2)$. Second, demobilizing the supporters of party 2, as represented by the term $W^2_{\tau^1}(y^{1*}, y^2)$. The weights given to these two considerations ensure that both objectives change by the same amount in percentage terms when the policy reform h^1 is implemented.

We proceed by introducing notation for a weighted average of $\mathcal{G}^1_W$ and $\mathcal{G}^2_W$:

$$\mathcal{G}_{SP}(\omega \mid y^1, y^2) := \lambda^1(y^1, y^2) \, \mathcal{G}^1_W(\omega \mid y^1, y^2) + (1 - \lambda^1(y^1, y^2)) \, \mathcal{G}^2_W(\omega \mid y^1, y^2) \, ,$$

where the weight $\lambda^1(y^1, y^2)$ satisfies

$$\lambda^1(y^1, y^2) = \left[1 + \psi(y^1, y^2)\frac{\bar{g}^2_W(\omega \mid y^1, y^2)}{\bar{g}^1_W(\omega \mid y^1, y^2)}\right]^{-1} \, .$$

Recall that the probability that party 1 wins is a monotonic function of $\psi(y^1, y^2) = \frac{W^1(y^1, y^2)}{W^2(y^1, y^2)}$. Thus, the weighting of $\mathcal{G}^1_W$ and of $\mathcal{G}^2_W$ is such that $\mathcal{G}^1_W$ receives, ceteris paribus, more weight if party 2 is more likely to win and $\mathcal{G}^2_W$ receives more weight if party 1 is more likely to win. Importantly, when both parties propose the same policies we can suppress the dependence of these expressions on y^1 and y^2 and write, e.g., $\bar{g}^{1s}_W(\omega) := \mathbb{E}[g^1_W(0 \mid \omega') \mid \omega' \geq \omega]$ or $\mathcal{G}^{1s}_W(\omega) = \frac{\bar{g}^{1s}_W(\omega)}{\bar{g}^{1s}_W(\omega)}$, where the superscript s stands for symmetry.

Theorem 1.

i) Given y^2, if y^1 is a maximizer of $\psi(y^1, y^2)$ then, for all ω,

$$\frac{T^{1\prime}(y^1(\omega))}{1 - T^{1\prime}(y^1(\omega))} = -\frac{1 - F(\omega)}{f(\omega)}\left(1 - \mathcal{G}_{SP}(\omega \mid y^1, y^2)\right)\frac{v_{21}(y^1(\omega), \omega)}{v_1(y^1(\omega), \omega)} \, . \tag{12}$$

ii) If (y^1, y^2) is a saddle point of ψ, then $y^1 = y^2$, where y^1 is such that, for all ω,

$$\frac{T^{1\prime}(y^1(\omega))}{1 - T^{1\prime}(y^1(\omega))} = -\frac{1 - F(\omega)}{f(\omega)} (1 - \mathcal{G}^s_{SP}(\omega)) \frac{v_{21}(y^1(\omega), \omega))}{v_1(y^1(\omega), \omega)} . \tag{13}$$

The proof of the theorem involves an analysis of the Gateaux differentials of the function $\psi(y^1, y^2)$. Part i) of Theorem 1 characterizes party 1's best response to an arbitrary proposal of party 2, whereas part ii) characterizes subgame perfect equilibrium policies. The characterization of best responses involves the functions y^1 and y^2 that jointly determine the weighting function $\mathcal{G}_{SP}$ that enters the right hand side of (12). In contrast, the second part of the theorem delivers an explicit and closed-form characterization of the equilibrium weights.

To see that equation (13) indeed delivers an equilibrium characterization in terms of the primitives of the model, note first that equilibrium policies have to be symmetric. As we show in the Appendix, the first order condition that y^1 needs to fulfill to qualify as a maximizer of $\psi(y^1, y^2)$ is equal to the first order condition that y^2 needs to fulfill to qualify as a minimizer. Thus, in equilibrium, the weight that party 1 places on mobilizing its own supporters is exactly equal to the weight that party 2 places on demobilizing the supporters of party 1. Analogously, the weight that party 1 places on demobilizing the supporters of party 2 is equal to the weight that party 2 places on mobilizing its own supporters. Second, all the endogenous quantities that appear in the characterizes of best responses, i.e. $\mathcal{G}^1_W$, $\mathcal{G}^1_W$, W^1, W^2, and λ^1, are functions of the utility differences $u^1(\omega) - u^2(\omega)$, $\omega \in \Omega$. With a symmetric equilibrium, we know that $u^1(\omega) - u^2(\omega) = 0$, for all $\omega \in \Omega$. This observation allows us to pin down the equilibrium values of all of these endogenous variables.

Suppose, for concreteness, that idiosyncratic party biases follow a uniform distribution (Assumption 1). In this case,

$$\mathcal{G}^{1s}_W(\omega) = \frac{\mathbb{E}[s^1(\omega') \mid \omega' \geq \omega]}{\bar{s}^1}, \; \mathcal{G}^{2s}_W(\omega) = \frac{\mathbb{E}[s^2(\omega') \mid \omega' \geq \omega]}{\bar{s}^2} \text{ and } \lambda^{1s} = \left[1 + \frac{W^{1s}/\bar{s}^1}{W^{2s}/\bar{s}^2}\right]^{-1},$$

with $\bar{s}^1 := \mathbb{E}[s^1(\omega)]$ and $\bar{s}^2 := \mathbb{E}[s^2(\omega)]$, and with $W^{1s} = \frac{1}{2}\mathbb{E}[s^1(\omega)^2/\alpha(\omega)]$ and $W^{2s} = \frac{1}{2}\mathbb{E}[s^2(\omega)^2/\alpha(\omega)]$. Hence, what drives marginal tax rates here are the functions $s^1 : \Omega \to [0, 1]$ and $s^2 : \Omega \to [0, 1]$ with $s^2(\omega) = 1 - s^1(\omega)$. Suppose that $\alpha(\omega) = \bar{\alpha}$ for all ω so that all types are equally inclined to swing to the party offering a better deal.

The weights λ^{1s} and $\lambda^{2s} := 1 - \lambda^{1s}$ on the weighting functions $\mathcal{G}_W^{1s}(\omega)$ and $\mathcal{G}_W^{2s}(\omega)$ are then entirely shaped by the functions s^1 and s^2. Recall that $s^1(\omega)$ is a measure of party 1's competitive advantage in attracting type ω-voters: if the parties propose the same policies, a fraction $s^1(\omega)$ supports party 1 and the complement $s^2(\omega)$ supports party 2.

For λ^{1s} close to 1, party 1 focuses on the maximization of the benefits to its supporters $W^1(\cdot)$ and weighs the rents of individuals with incomes above $y(\omega)$ by $\mathcal{G}_W^{1s}(\omega)$, with the implication that marginal tax rates for these voters are low if the voters' attachment to party 1 in this subset of the electorate, $\mathbb{E}[s^1(\omega') \mid \omega' \geq \omega]$, is strong. If instead λ^{1s} is close to 0, party 1 focuses on the minimization of the benefits for the supporters of the competing party $W^2(\cdot)$ and weighs the rents of individuals with incomes above $y(\omega)$ by $\mathcal{G}_W^{2s}(\omega)$, with the implication that marginal tax rates for these voters are low if the voters' attachment to party 1 in this subset of the electorate is weak. More generally, an inspection of λ^{1s} shows that it is, ceteris paribus, small if $\psi^s = W^{1s}/W^{2s}$ is large. Recall that party 1's equilibrium probability of winning is given by $\pi^{1s} = \mathcal{P}\left(W^{1s}/W^{2s}\right)$ and hence is a monotonic function of W^{1s}/W^{2s}. Thus, a low probability of winning comes with a large value of λ^{1s}. It implies, moreover, that party 1 puts less weight on reducing $W^2(\cdot)$ than on maximizing $W^1(\cdot)$, i.e., it will then focus on making an attractive offer to its own supporters – that is, its primary goal is to mobilize its voters. By contrast, if party 1 is the front-runner in the campaign it puts more weight on mitigating $W^2(\cdot)$ than on boosting $W^1(\cdot)$ – that is, its primary goal is to demobilize the supporters of the opposite party.

We summarize this discussion in the following Corollary to Theorem 1.

Corollary 1. *Suppose that B_a is a family of distribution functions, one for each type ω, so that $B_a(x \mid \omega) = s_a^1(\omega) + \alpha_a(\omega)\,x$. Denote the corresponding equilibrium quantities by λ_a^{1s}, π_a^{1s}, etc. Let B_b be another family of distribution functions with $B_b(x \mid \omega) = s_b^1(\omega) + \alpha_b(\omega)\,x$ that gives rise to equilibrium quantities λ_b^{1s}, π_b^{1s}, etc. Suppose that $\bar{s}_a^1 = \bar{s}_b^1$ so that either party has, on average, the same voter attachment. Then*

$$\lambda_a^{1s} < \lambda_b^{1s} \quad \text{if and only if} \quad \pi_a^{1s} > \pi_b^{1s} \,.$$

Consider the implications of this equilibrium characterization for one of the most focal aspects of tax design, the marginal tax rates on the rich. If one party has a high chance of winning and, moreover, little attachment from the rich, then equilibrium tax policies involve only small tax distortions on high incomes. Also the disadvantaged

24

party with more attachment from the rich chooses to propose moderate tax rates at the top. The latter does so in an attempt to increase the stakes for its supporters, the former does so in order to undo this attempt. This observation rationalizes why even a more left-leaning party – a party with a comparatively high support from voters with low incomes – may shy away from heavier taxes on the rich. If the party is the likely winner of the election, low taxes on the rich are valuable as they help to discourage the supporters of the other, more right-leaning party. Only a left leaning party with a small probability of winning chooses to propose high taxes on the rich and hence a large amount of lump-sum redistribution to increase the benefits that its own supporters would realize in the unlikely event that it wins the election. The same logic can explain why a more conservative or pro-market party that is supported to a large extent by voters with above average incomes may not propose to abandon a progressive tax system. If the party is the likely winner of the election, its focus is on weakening the support for the competing leftist party. It therefore chooses to propose a platform that is appealing also to voters with below average incomes.

3.3   Existence of a saddle point

Subgame perfect equilibria give rise to a non-trivial strategic interdependence of the parties' proposals. We noted above that the weighting of information rents that is associated with a best response of party 1 depends on the policy proposed by party 2. The possibility that the parties choices are strategic substitutes or complements raises the question whether the existence of pure strategy equilibria can be ensured. Matching pennies is the prototype zero-sum game with strategic complements and it is well-known that it does not have a pure strategy equilibrium. The following Proposition shows that upon imposing Assumption 1 we can also state sufficient conditions for the existence of a subgame perfect equilibrium.

Proposition 3. *Suppose that Assumption 1 holds and moreover that $s^1(\omega) = \frac{1}{2}$, for all ω. Also suppose that there is a number ζ so that $\alpha(\omega) \leq \zeta$, for all ω. Then, a pair of policies that satisfies (13) is a saddle point of the function ψ.*

Proposition 3 looks at a symmetric case in which each party attracts fifty percent of the electorate in the case that both parties make the same proposal. Moreover, it assumes that the distribution of idiosyncratic party preferences has a wide support, as captured by the assumption that $\alpha(\omega)$ is, for all ω, sufficiently small. The proof in the

Appendix is of a stand alone interest as it shows how to use functional derivatives to check second order conditions that ensure equilibrium existence. The distributional assumptions primarily serve to simplify the analysis. They are stronger than needed as there is a range of close enough distributions for which second order conditions are also fulfilled.

4 Nash equilibrium vs subgame perfect equilibrium

4.1 Nash equilibrium policies

We now compare subgame perfect equilibrium policies to Nash equilibrium policies. Nash equilibrium policies do not take the endogeneity of turnout into account. The comparison therefore allows us to highlight how this endogeneity shapes equilibrium policies. The following Proposition provides a characterization of Nash equilibrium policies. A proof would involve only a straightforward adaptation of the arguments in the proof of Lemma 3 and is therefore omitted.

Proposition 4. *If (y^1, y^2) is a saddle point of the function $\bar{\Gamma}^{1u}$, then*

$$\frac{T^{1\prime}(y^1(\omega))}{1 - T^{1\prime}(y^1(\omega))} = -\frac{1 - F(\omega)}{f(\omega)} (1 - \mathcal{G}_N(\omega \mid y^1, y^2)) \frac{v_{21}(y^1(\omega), \omega))}{v_1(y^1(\omega), \omega)} , \qquad (14)$$

where

$$\mathcal{G}_N(\omega \mid y^1, y^2) := \frac{\bar{b}^u(\omega \mid y^1, y^2)}{\bar{b}^u(\underline{\omega} \mid y^1, y^2)} ,$$

and

$$\bar{b}^u(\omega \mid y^1, y^2) := \mathbb{E}[\bar{q}^u(\omega') \, b(u^1(\omega') - u^2(\omega') \mid \omega') \mid \omega' \geq \omega] .$$

Recall that $b(u^1(\omega') - u^2(\omega') \mid \omega')$ measures the mass of type ω'-individuals that party 1 can attract into its base by offering them slightly more utility. Potential voters, however, are only the rule-utilitarian members of the base so that $\bar{q}^u(\omega') \, b(u^1(\omega') - u^2(\omega') \mid \omega')$ is the mass of potential voters that party 1 can attract by proposing a tax policy that generates more utility for type ω'-individuals. Therefore, $\bar{b}^u(\omega \mid y^1, y^2)$ is the mass of potential voters that party 1 can add to its base by offering all agents with types above ω slightly more utility. The mass of potential voters that party 1 can attract by slightly raising everybody's utility is given by $\bar{b}^u(\underline{\omega} \mid y^1, y^2)$ and the

ratio $\mathcal{G}_N(\omega \mid y^1, y^2)$ relates the political return of making everybody with a type above ω better off to the political return of making everybody better off.

The Proposition provides a swing voter explanation for what shapes political equilibrium tax policies. When considering, say, higher marginal taxes on high incomes, parties take into account the extent to which "the rich" withdraw political support in response to higher taxes. Thereby they also take into account the turnout rate among "the rich". To see these forces most clearly, suppose that idiosyncratic party biases follow a uniform distribution (Assumption 1)). Then,

$$\mathcal{G}_N(\omega \mid y^1, y^2) = \frac{\mathbb{E}[\bar{q}^u(\omega')\alpha(\omega') \mid \omega' \geq \omega]}{\overline{q^u \alpha}} \ , \tag{15}$$

where $\overline{q^u \alpha} := \mathbb{E}[\bar{q}^u(\omega)\alpha(\omega)]$. Moreover, if all types swing in equal proportions, so that for any pair ω and ω', $\alpha(\omega) = \alpha(\omega')$, then the weighting of the different types' information rents is entirely driven by their participation rates in elections. For instance, if individuals with higher incomes are more likely to participate in the election, then $\bar{q}^u : \omega \mapsto \bar{q}^u(\omega)$ is an increasing function so that $\mathcal{G}_N(\omega \mid y^1, y^2)$ is also an increasing function. The formula in (14) would then stipulate that marginal tax rates are, ceteris paribus, decreasing with income.

Recall that subgame perfect equilibrium policies were shaped by the functions $s^1 : \Omega \to [0, 1]$ and $s^2 : \Omega \to [0, 1]$ that determine the distribution of voter attachments to parties 1 and 2, respectively, among type ω-citizens. The weighting functions $\mathcal{G}_W^1$ and $\mathcal{G}_W^2$ that are relevant for subgame perfect equilibrium policies depend only on these two functions. Nash equilibrium policies, by contrast, are shaped by the function $\omega \mapsto \bar{q}^u(\omega)\alpha(\omega)$. Taking into account the endogeneity of turnout turns a swing voter model into a model that emphasizes voter attachments.[11]

---

[11]The participation rates $\{\bar{q}^u(\omega)\}_{\omega \in \Omega}$ play no role for subgame perfect equilibrium policies. This is a consequence of a modeling choice due to Feddersen and Sandroni (2006) and Coate and Conlin (2004): When the rule-utilitarian supporters of party 1 formulate the rule for participating in an election they weigh the benefits of everybody who is a supporter of party 1. In particular, they include the benefits of the free-riders who are made better off if party 1 wins but would never participate in an election. If they weighted only the benefits of those with a sense of civic duty, the above formulas would become $\mathcal{G}_W^{1s}(\omega) = \frac{\mathbb{E}[\bar{q}^u(\omega')s^1(\omega')|\omega' \geq \omega]}{\overline{q^{1u}s}}$ and $G_W^{2s}(\omega) = \frac{\mathbb{E}[\bar{q}^u(\omega')s^2(\omega')|\omega' \geq \omega]}{\overline{q^{2u}s}}$, where $\overline{q^{1u}s} := \mathbb{E}[q^u(\omega)s^1(\omega)]$ and $\overline{q^{2u}s} := \mathbb{E}[q^u(\omega)s^2(\omega)]$, with similar adjustments for the weight λ^{1s}. With this alternative modeling choice, differential turnout among poor, rich and middle class voters also comes into play as an additional determinant of subgame perfect equilibrium policies.

4.2 Strategic substitutes and complements

4.2.1 Preliminaries: using tax schedules as the policy domain

We briefly sketch an alternative equilibrium characterization in which the policy domain is taken to be the set of tax schedules – rather than the set of non-decreasing earnings functions. This makes it possible to relate our analysis more directly to standard concepts in public finance such as, e.g., the marginal cost of public funds. More importantly, this turns out to be particularly useful for the analysis of whether the two parties' policies are strategic substitutes or complements.

As we explained in Section 3.1, we can directly let parties choose nonlinear tax schedules, i.e., twice continuously differentiable functions $T^j : \mathbb{R}_+ \to \mathbb{R}$, that are such that the government budget remains balanced – that is, any extra revenue (or loss) generated by the tax reform that we are considering, taking into account the behavioral responses it induces on labor supply, is rebated (or taxed away) lump-sum and uniformly to the population. We can replace the arguments (y^1, y^2) and ω of all the functions we consider with (T^1, T^2) and $y^1 \equiv y^1(\omega)$, respectively. Consider for instance the problem of choosing T^1 so as to maximize $W^1(T^1, T^2) = \mathbb{E}[G_W^1(u(T^1, y^1) - u(T^2, y^2) \mid y^1)]$, where the expectation is over the incomes y^1 that are realized if party 1 is elected. The details of the argument and the derivations are in the Appendix.

Consider a reform of the platform T^1 in the direction h^1 along with a lump-sum rebate R^1, so that the tax function proposed by party 1 becomes $y \mapsto T^1(y) + \tau^1[h^1(y) - R^1]$. By the envelope theorem, the utility of this agent decreases one-for-one with the change in the total tax payment, that is, its Gateaux differential is given by

$$u_{\tau^1}(T^1, y^1(\omega)) = -h^1(y^1(\omega)) + R^1. \tag{16}$$

Moreover, the net revenue generated by the tax reform h^1, and hence the lump-sum tax rebate, is the sum of a mechanical effect and a behavioral effect and is given by[12]

$$R^1 = \int_{\underline{y}^1}^{\bar{y}^1} \left[(1 - F_{Y^1}(y^1)) - \frac{T^{1\prime}(y^1)}{1 - T^{1\prime}(y^1)} y^1 \tilde{\varepsilon}(y^1) f_{Y^1}(y^1) \right] h^{1\prime}(y^1) dy^1,$$

where $\tilde{\varepsilon}(\omega)$ is the taxable income elasticity along the nonlinear budget constraint. Collecting these terms, we can easily obtain the change in the expected gain of party

---

1's supporters, $W^1_{\tau^1}(T^1, T^2)$. The optimum platform T^1 is such that $W^1_{\tau^1}(T^1, T^2) = 0$ for all tax reforms h^1. This implies:

$$\frac{T^{1\prime}(y^1)}{1 - T^{1\prime}(y^1)} = \frac{1 - F_{Y^1}(y^1)}{y^1 f_{Y^1}(y^1)} \left(1 - \mathcal{G}^1_W(y^1 \mid T^1, T^2)\right) \frac{1}{\tilde{\varepsilon}(y^1)}, \qquad (17)$$

where the weights $\mathcal{G}^1_W(y^1 \mid T^1, T^2)$ are the analogue of those defined in Lemma 3.

In the standard social welfare-maximization problem, the marginal value of public funds is given by $\lambda_W := \mathbb{E}[g(u(T, y(\omega)))]$, where G is the concave social welfare function and g is its derivative. That is, λ_W is the value of rebating lump-sum one unit of tax revenue uniformly to the entire population. Here, the corresponding variable is $\lambda_{G^1} := \mathbb{E}[g^1_W(u(T^1, y^1(\omega)) - u(T^2, y^2(\omega)) \mid \omega)]$ which appears in the denominator of $\mathcal{G}^1_W(y \mid T^1, T^2)$. It is the total political advantage that party 1 can secure over party 2 by rebating lump-sum one unit of tax revenue uniformly to the entire population.

4.2.2 Analyzing the strategic complementarities

Subgame perfect and Nash equilibria differ also in the way in which the best response of party 1 depends on the proposal of party 2 and vice versa. For a systematic analysis of this strategic interdependence we use the observation that the best response of party 1 proposes a tax schedule analogous to (17), where the weights $\mathcal{G}^1_W$ in this formula are replaced by $\mathcal{G}_N$ if its best response seeks to maximize $\bar{\Gamma}^{1u}$ and by $\mathcal{G}_{SP}$ if it seeks to maximize ψ. We focus on $\mathcal{G}_N$ in what follows and then comment how the analysis would have to be altered for $\mathcal{G}_{SP}$.

The dependence of $\mathcal{G}_N$ on both y^1 and y^2 allows for an analysis of whether the tax policies of different parties are strategic substitutes or complements. For instance, if party 2 moves towards a tax schedule that has higher tax rates at the top, does this imply that party 1's best response also changes in the direction of higher tax rates at the top? Functional derivatives can be used to operationalize questions of this type. If the Gateaux differential of $\mathcal{G}_N(\omega \mid y^1, y^2)$ with respect to a change of y^2 into a direction of higher marginal tax rates at the top is positive, this implies that party 1's best response will involve lower marginal tax rates on individuals of type ω. If it is negative, it involves higher marginal taxes.

To analyze this question it is particularly convenient to adopt the alternative equilibrium characterization developed in Section 4.2.1, whereby parties 1 and 2 choose tax functions $T^1, T^2 : \mathbb{R}_+ \to \mathbb{R}$ directly, rather than earnings functions y^1, y^2. We saw

in equation (16) that by the envelope theorem, the indirect utility of agents moves one-for-one with the change in their total tax payment, so that the Gateaux differential of utility when the tax schedule proposed by party 2 is perturbed in the direction h^2 is given by $u_{\tau^2}(T^2, \omega) = -h^2\left(y^2\left(\omega\right)\right)$. The following Proposition provides a characterization of the Gateaux differential of the the weights $\mathcal{G}_N(y^1 \mid T^1, T^2)$ in response to a tax reform h^2 proposed by party 2.

Proposition 5. *The Gateaux differential of $\mathcal{G}_N(y^1 \mid T^1, T^2)$ in a fixed but arbitrary direction h^2, evaluated at T^2, equals*

$$\frac{\mathcal{G}_{N,\tau^2}(y^1\left(\omega\right) \mid T^1, T^2)}{\mathcal{G}_N(y^1\left(\omega\right) \mid T^1, T^2)} = \frac{\mathbb{E}\left[h^2\left(y^2\left(\omega'\right)\right)\bar{q}^u\left(\omega'\right)b'\left(u^1\left(\omega'\right) - u^2\left(\omega'\right)\mid\omega'\right)\mid\omega'\geq\omega\right]}{\mathbb{E}\left[\bar{q}^u\left(\omega'\right)b\left(u^1\left(\omega'\right) - u^2\left(\omega'\right)\mid\omega'\right)\mid\omega'\geq\omega\right]}$$
$$- \frac{\mathbb{E}\left[h^2\left(y^2\left(\omega'\right)\right)\bar{q}^u\left(\omega'\right)b'\left(u^1\left(\omega'\right) - u^2\left(\omega'\right)\mid\omega'\right)\right]}{\mathbb{E}\left[\bar{q}^u\left(\omega'\right)b\left(u^1\left(\omega'\right) - u^2\left(\omega'\right)\mid\omega'\right)\right]},$$

(18)

where $u^j\left(\omega'\right) := u\left(T^j, y^j\left(\omega'\right)\right)$. In particular, suppose that the initial equilibrium is symmetric, so that $T^1 = T^2$ and hence $u\left(T^1, y^1\left(\omega\right)\right) = u\left(T^2, y^2\left(\omega\right)\right)$ for all $\omega \in \Omega$. For simplicity, suppose moreover that the expected fraction of ethical voters and the number of swing voters is the same at all income levels, so that $\bar{q}^u\left(\omega\right)$ and $b\left(0 \mid \omega\right)$ are constant, and that $b'\left(0 \mid \omega\right) := b'\left(0\right) \lessgtr 0$ is constant. We obtain

$$\mathcal{G}_{N,\tau^2}(y^1\left(\omega\right) \mid T^1, T^2) = \frac{b'\left(0\right)}{b\left(0\right)}\left\{\mathbb{E}\left[h^2\left(y^2\left(\omega'\right)\right) \mid \omega' \geq \omega\right] - \mathbb{E}\left[h^2\right]\right\}. \quad (19)$$

Thus, if the tax reform proposed by party 2 is increasing, i.e., it raises the statutory tax payments by a larger amount for higher incomes, then party 1's best response does the same if $b'\left(0\right) < 0$ (strategic substitutes) and does the opposite if $b'\left(0\right) > 0$ (strategic complements).

The sign of these Gateaux differentials depends in particular on the curvature of the functions $\{B(\cdot \mid \omega)\}_{\omega \in \Omega}$, i.e., on the distribution of idiosyncratic party preferences. With uniform distributions (Assumption 1), there are neither complementarities nor substitutabilities in the parties tax schedules as $b'\left(x \mid \omega\right) = 0$, for any x and ω. Hence, also $\mathcal{G}_{N,\tau^2}(y^1\left(\omega\right) \mid T^1, T^2) = 0$ by equation (18). Thus, with uniform distributions of party preferences, the parties's choices are strategically neutral. The best response of party 1 is then independent of what party 2 is proposing and vice versa. Put differently, the Nash equilibrium is then also a dominant strategy equilibrium. Equation (19) shows that, starting from a symmetric equilibrium, the two parties' policies are

strategic substitutes if B is locally concave at 0 (i.e., $b'(0) < 0$), and strategic comple­ments otherwise (if $b'(0) > 0$). Importantly, note that formula (18) is a closed-form expression. Given the tax schedules T^1 and T^2 proposed by the two parties, and given the tax reform h^2, all of the variables of the equation are known; (18) therefore delivers immediately the best response of party 1 to any arbitrary change in policy by party 2, given any (possibly off-equilibrium) initial status quo (T^1, T^2).

In the general model with endogenous turnout, analyzing the substitutability between the two parties policies again requires evaluating Gateaux differentials to evaluate how $\mathcal{G}_{SP}(y^1(\omega) \mid T^1, T^2)$ depends on the proposal of party 2. We focus on a simple special case here (see Appendix for the more general expressions). Assume that initially, both parties propose the same platform $T^1 = T^2$, and that Assumption 1 holds with $\alpha(\omega) = \bar{\alpha}$ for all ω. In response to a reform h^2 of the platform T^2 proposed by party 2, the weights $\mathcal{G}_W^1(\cdot \mid T^1, T^2)$ that party 1 would use if its objective were to mobilize its own supporters (i.e., maximize $W^1(y^1, y^2)$) change by, for all $\omega \in \Omega$,

$$\frac{\mathcal{G}_{W,\tau^2}^1(y^1(\omega) \mid T^1, T^2)}{\bar{\alpha}\mathcal{G}_W^1(y^1(\omega) \mid T^1, T^2)} = \frac{\mathbb{E}\left[h^2(y^2(\omega')) \mid \omega' \geq \omega\right]}{\mathbb{E}\left[s^1(\omega') \mid \omega' \geq \omega\right]} - \frac{\mathbb{E}\left[h^2(y^2(\omega'))\right]}{\mathbb{E}\left[s^1(\omega')\right]} . \quad (20)$$

Similarly, the weights $\mathcal{G}_W^2(\cdot \mid T^1, T^2)$ that Party 1 would use if its objective was to de-mobilize Party 2's supporters (i.e., minimize $W^2(y^1, y^2)$) change by, for all $\omega \in \Omega$,

$$-\frac{\mathcal{G}_{W,\tau^2}^2(y^1(\omega) \mid T^1, T^2)}{3\bar{\alpha}\mathcal{G}_W^2(y^1(\omega) \mid T^1, T^2)} = \frac{\mathbb{E}\left[h^2(y^2(\omega')) \mid \omega' \geq \omega\right]}{\mathbb{E}\left[s^2(\omega') \mid \omega' \geq \omega\right]} - \frac{\mathbb{E}\left[h^2(y^2(\omega'))\right]}{\mathbb{E}\left[s^2(\omega')\right]} . \quad (21)$$

Finally, the relative weight $\lambda^1(T^1, T^2)$ that Party 1 puts on the first of these objectives changes by a constant $\lambda_{\tau^2}^1(T^1, T^2)$ independent of income y, which can generally be positive or negative depending on the mean and variance of s^1 and its correlation with the reform h^2 (see the Appendix). Using these expressions, it is then straightforward to obtain the adjustments in the weights $\mathcal{G}_{SP}(\cdot \mid T^1, T^2)$ that determine party 1's proposed tax schedule in the subgame perfect equilibrium:

$$\mathcal{G}_{SP,\tau^2}(y^1(\omega) \mid T^1, T^2) = \lambda^1 \mathcal{G}_{W,\tau^2}^1 + (1 - \lambda^1)\mathcal{G}_{W,\tau^2}^2 + \lambda_{\tau^2}^1\left(\mathcal{G}_W^1 - \mathcal{G}_W^2\right) ,$$

for all $\omega \in \Omega$. The first two terms of this expression account for the adjustments in the weights $\mathcal{G}_W^1, \mathcal{G}_W^2$ associated with both objectives of party 1 (mobilizing its own

base and demobilizing that of party 2), while the third term accounts for the change in the relative importance λ^1 of each of these objectives in response to the reform h^2. For concreteness, suppose that $s^1(\omega) = s^2(\omega) = \frac{1}{2}$ for all ω, so that both parties have an equal attachment of the electorate at each income level. In this case, we have $\mathcal{G}_W^1 = \mathcal{G}_W^2$ and $\lambda^1 = \frac{1}{2}$. For high enough incomes, we have $\mathcal{G}_{W,\tau^2}^1 > 0$ and $\mathcal{G}_{W,\tau^2}^2 < 0$ in response to a reform by party 2 that raises taxes on the rich, i.e., such that $\mathbb{E}[h^2(y^2(\omega')) \mid \omega' \geq \omega]$ increases with ω. That is, in order to mobilize its supporters (resp., demobilize those of the opponent), party 1 tends to decrease (resp., increase) the marginal tax rates on the rich. Since $\mathcal{G}_{W,\tau^2}^2 = -3\mathcal{G}_{W,\tau^2}^1$, we get $\mathcal{G}_{SP,\tau^2} < 0$ for high incomes, so that the second objective dominates and the two parties' policies are strategic complements. Note moreover that both $\mathcal{G}_{W,\tau^2}^1$ and $\mathcal{G}_{W,\tau^2}^2$ are increasing with income y^1, so that the best response of party 1 is to decrease the progressivity of its tax schedule.

5   Additional considerations

5.1   Political failures

We now discuss to what extent political equilibrium policies are appealing from a normative perspective. We therefore relate them to the tax policy that maximizes a social welfare function $SWF = \mathbb{E}[G(u(\omega))]$, where G is a concave transformation and $u(\omega)$ is the utility level that a type ω individual realizes under a given tax policy. Let g denote the derivative of G. Then $\bar{g}(\omega \mid y) := \mathbb{E}[g(u(\omega') \mid \omega' \geq \omega]$ is the welfare gain that is realized if all individuals with incomes above $y(\omega)$ are made slightly better off, and the ratio $\mathcal{G}_D(\omega \mid y) := \frac{\bar{g}(\omega \mid y)}{\bar{g}(\omega \mid y)}$ relates this welfare gain to the one that would be obtained under a tax reform that raises everybody's utility – i.e., the denominator is the marginal value of public funds.

Proposition 6. (Diamond, 1998) *A welfare maximizing tax system is such that, for all ω,*

$$\frac{T'(y(\omega))}{1 - T'(y(\omega))} = -\frac{1 - F(\omega)}{f(\omega)}(1 - \mathcal{G}_D(\omega \mid y))\frac{v_{21}(y(\omega), \omega)}{v_1(y(\omega), \omega)}. \qquad (22)$$

We omit a formal proof of Proposition 6 as it requires only a straightforward adaptation of the proof of Lemma 3. With a concave social welfare function, the

weighting function $\mathcal{G}_D(\cdot \mid y)$ is decreasing. Ceteris paribus, i.e. keeping the inverse hazard rate and the elasticity of labor supply constant – marginal tax rates should therefore increase with income.

This observation is useful for a welfare analysis of political equilibrium outcomes. If a weighting function for information rents $\mathcal{G}$ can be derived from a social welfare function, then $\omega' < \omega$ implies that $\mathcal{G}(\omega' \mid y) \geq \mathcal{G}(\omega \mid y)$ as individuals with lower income cannot receive a lower weight in the social welfare function. If the weighting function associated with a political equilibrium violates this monotonicity property then we can conclude that the political outcome is incompatible with the objective of welfare-maximization. We then say that there is a *political failure.*

Depending on the primitives, both Nash and subgame perfect equilibria may give rise to such political failures. For instance, with a uniform distribution of idiosyncratic party preferences (Assumption 1), $\mathcal{G}_N(\omega \mid y^1, y^2)$ is given by expression (15), so that if α is constant across types but the inclination to vote increases with income, so that $\bar{q}^u$ is an increasing function, then Nash equilibrium policies give rise to a political failure. Now consider subgame perfect equilibrium policies, again under Assumption 1. Recall that $\mathcal{G}^s_{SP}(\omega)$ is a convex combination of two weighting functions $\mathcal{G}^1_W$ and $\mathcal{G}^2_W$, where the weight on $\mathcal{G}^1_W$ is low if party 1 has a high chance of winning. Suppose that s^1 is a decreasing function and s^2 an increasing function, then the weighting function $\mathcal{G}^1_W$ is decreasing and the weighting function $\mathcal{G}^2_W$ is increasing. In this scenario, if party 1's winning probability is sufficiently close to one, subgame perfect equilibrium policy gives rise to a political failure.

5.2 Equilibrium turnout

Parties in our model do not care about turnout by itself. They care about turnout only to the extent that it increases their probability of winning an election. For this purpose, only the turnout advantage over the competing party is important; i.e., parties prefer a large share in small set of voters over a small share in a large set of voters. The following Proposition enables us to formalize this claim.

Proposition 7. *Under Assumptions 2 and 3,*

$$\sigma^{1*}(y^1, y^2) = \frac{W^1(y^1, y^2)}{\kappa \, \bar{\Gamma}^{1u}(y^1, y^2)} \, \psi(y^1, y^2) \, \rho(\psi(y^1, y^2)) \,, \tag{23}$$

where $\rho(\cdot)$ is the density function of the random variable $\frac{1-\eta^2}{1-\eta^1}$.

According to Proposition 7, if party 1 wanted to maximize its number of voters – rather than its probability of winning – it would choose y^1 so as to maximize the right hand side of (23). In a subgame perfect equilibrium, by contrast, it chooses y^1 with the objective to maximize $\psi(y^1, y^2)$ – recall that, under Assumptions 2 and 3, party 1's probability of winning is a monotonic function of $\psi(y^1, y^2)$. That is, in equilibrium parties have opportunities to increase overall turnout, but choose not to use them as this would be detrimental for their probability of winning.

More formally, let $\mathcal{BC}^1(y^1, y^2) = \frac{W^1(y^1, y^2)}{\kappa \, \bar{\Gamma}^{1u}(y^1, y^2)}$ be a shorthand for the benefit-cost ratio that drives the turnout of the rule-utilitarian supporters of party 1, and let $\Pi^1(\psi(y^1 y^2)) := \psi(y^1, y^2) \, \rho(\psi(y^1, y^2))$ capture how the probability of a victory of party 1 affects this turnout ratio. To show that party 1 could deviate – from a subgame perfect equilibrium allocation or merely a platform that is a best response in the sense of maximizing party 1's probability of winning for a given platform of party 2 – in a way that increases turnout, we look at the Gateaux-differential of $\ln(\sigma^{1*}(y^1, y^2))$ for an arbitrary direction h^1,

$$\frac{\sigma_{\tau^1}^{1*}(y^1, y^2)}{\sigma^{1*}(y^1, y^2)} = \frac{\mathcal{BC}_{\tau^1}^1(y^1, y^2)}{\mathcal{BC}^1(y^1, y^2)} + \frac{\Pi^{1'}(\psi(y^1 y^2)) \, \psi_{\tau^1}(y^1 y^2)}{\Pi^1(\psi(y^1 y^2))} \, .$$

If y^1 is a best response for party 1, then for any direction h^1, $\psi_{\tau^1}(y^1 y^2) = 0$ and hence $\frac{\sigma_{\tau^1}^{1*}(y^1, y^2)}{\sigma^{1*}(y^1, y^2)} = \frac{\mathcal{BC}_{\tau^1}^1(y^1, y^2)}{\mathcal{BC}^1(y^1, y^2)}$. Thus, any deviation into a direction that increases $\mathcal{BC}^1(y^1, y^2)$ will have a negligible effect on party 1's probability of winning, and a first order effect on the turnout of its supporters.

Previous analyses of ethical voter models by Coate and Conlin (2004) and Feddersen and Sandroni (2006) have shown that overall turnout increases if elections are close or if preferences over alternatives are more polarized. The observation that closeness is conducive to overall turnout has a resemblance to the implication of equation (23) that turnout is increasing in the semi-elasticity of the probability of winning with respect to ψ^1, $\psi(y^1, y^2) \, \rho(\psi(y^1, y^2))$. If we interpret a close race as one in which small changes in the proposed policies have large consequences for relative turnout and hence also for each party's probability of winning, then it is indeed an implication of (23) that turnout tends to be larger if elections are close.

Our analysis has no straightforward analogue to the finding that polarization drives turnout. The analysis of polarization in Feddersen and Sandroni (2006) is based on the premise that the expected benefits to the supporters of party 1 and the

expected benefits to the supporters of party 2 add up to a constant which is, moreover, independent of the probability of winning. In our setting, this would require the existence of a number $\bar{W}$ so that, for all (y^1, y^2) and for all $\pi^1 \in [0, 1]$, $\pi^1 W^1(y^1, y^2) + (1 - \pi^1)W^2(y^1, y^2) = \bar{W}$. Feddersen and Sandroni provide comparative statics results on how turnout varies with $\bar{W}$. In our setting the expected benefits of the supporters of parties 1 and 2 do not add up to a constant, so that we cannot perform a similar exercise. The observations that equilibria are symmetric and that a deviation from equilibrium policies would increase overall turnout is therefore a related finding. It shows that a lack of polarization leaves room for an increase of overall turnout.

6   Conclusion: endogenous social welfare weights

The theory of optimal taxation characterizes the tax policy that would be chosen by a benevolent social planner with the objective to maximize social welfare. However, there are many "reasonable" social welfare functions. The literature then often refers to "society's social welfare function" as the relevant one but typically does not provide a precise formulation of how society would come to agree on a given social welfare function. Still there are attempts to identify society's social welfare function empirically. The idea is to use the ABC formula for welfare-maximizing taxation (22), data on the tax system, the earnings distribution and labor supply elasticities to map out society's social welfare weights $\mathcal{G}$, see e.g. Blundell et al. (2009). Bargain et al. (2011), Zoutman, Jacobs and Jongen (2016) or Lockwood and Weinzierl (2016).

Propositions 4 and 6 and Theorem 1 contain different ABC-formulas. Proposition 4 gives the tax rates that are predicted by a swing voter model with exogenous turnout, Proposition 6 characterizes welfare-maximizing taxes, and Theorem 1 takes into account the endogeneity of turnout. Each of these formulas provides a different interpretation of the $\mathcal{G}$-term that can be mapped out given data on the tax system, the earnings distribution and labor supply elasticities. The data cannot tell us, however, how plausible these different interpretations are. This requires criteria external to the formal framework developed here. What we can say, however, is that it makes little sense to let the data speak about what society's social welfare function looks like without at the same time having an explanation for how society would come to settle on a specific social welfare function.

In fact, Arrow's impossibility shows that it may be entirely infeasible to aggregate

individual preferences over tax policies into a social welfare function. What is feasible, by contrast, is to reinterpret any process that selects a tax schedule from the set of Pareto-efficent tax schedules as if it was the choice of an agent with preferences over the set of Pareto-optima, see Samuelson (1967). Saez and Stantcheva (2016) recognize that observed tax policies are not entirely driven by welfare considerations, but also by non-welfarist value judgments or political economy forces. They propose to capture these forces by means of generalized social welfare functions. Thus, observed tax policies are interpreted as if there had been an agent who made a selection from the set of Pareto optima and who used a generalized social welfare function for this purpose.

In the preceding analysis we have analyzed a strategic game between political parties who propose non-linear tax schedules to an electorate that consists of individuals with heterogenous incomes. We have shown that political equilibrium tax policies can be represented by ABC-formulas akin to those that are used to characterize welfare-maximizing tax systems. Thus our analysis provides a detailed and microfounded formalization that political economy forces can indeed be captured by generalized social welfare weights. Moreover, we emphasize that these generalized social welfare weights are objects which are endogenous to the political process and, in particular, that alternative models of political competition give rise to alternative social welfare weights. We have also shown that political equilibrium outcomes may be incompatible with the maximization of a concave social welfare function. Generalized social welfare weights are therefore needed. The class of traditional social welfare weights cannot capture such political failures.

References

Acemoglu, D., M. Golosov, and A. Tsyvinski, "Political Economy of Mechanisms," *Econometrica*, 2008, *76*, 619–641.

_ , _ , and _ , "Dynamic Mirrlees Taxation under Political Economy Constraints," *Review of Economic Studies*, 2010, *77*, 841–881.

Alesina, A. and G.-M. Angeletos, "Fairness and Redistribution," *American Economic Review*, 2005, *95* (4), 960–980.

Bargain, O., M. Dolls, D. Neumann, A. Peichl, and S. Siegloch, "Tax-Benefit Systems in Europe and the US: Between Equity and Efficiency," *IZA Discussion Papers 5440*, 2011.

Barro, R., "The Control of Politicians: An economic model," *Public Choice*, 1973, *14*, 19–42.

Besley, T. and S. Coate, "A Model of Representative Democracy," *Quarterly Journal of Economics*, 1997, *112*, 85–114.

Bierbrauer, Felix and P. Boyer, "Politically feasible reforms of non-linear tax systems," *CESifo Working Paper No. 6573*, 2017.

Bierbrauer, Felix J and Martin F Hellwig, "Robustly coalition-proof incentive mechanisms for public good provision are voting mechanisms and vice versa," *The Review of Economic Studies*, 2016, *83* (4), 1440–1464.

Bierbrauer, F.J. and P.C. Boyer, "Efficiency, Welfare, and Political Competition," *Quarterly Journal of Economics*, 2016, *131* (1), 461–518.

Blundell, R., M. Brewer, P. Haan, and A. Shephard, "Optimal income taxation of lone mothers: an empirical comparison of the UK and Germany," *Economic Journal*, 2009, *119*, F101–F121.

Brett, C. and J.A. Weymark, "Voting over selfishly optimal nonlinear income tax schedules," *Games and Economic Behavior*, 2017, *101*, 172–188.

Coate, Stephen and Michael Conlin, "A Group Rule-Utilitarian Approach to Voter Turnout: Theory and Evidence," *The American Economic Review*, 2004, *94* (5), 1476–1504.

Coughlin, Peter and Shmuel Nitzan, "Electoral outcomes with probabilistic voting and Nash social welfare maxima," *Journal of Public Economics*, 1981, *15* (1), 113 – 121.

Diamond, P.A., "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 1998, *88*, 83–95.

Downs, A., *An Economic Theory of Democracy*, New York, Harper and Row., 1957.

Farhi, E., C. Sleet, I. Werning, and S. Yeltekin, "Non-linear Capital Taxation Without Commitment," *Review of Economic Studies*, 2012, *79* (4), 1469–1493.

Feddersen, Timothy and Alvaro Sandroni, "A Theory of Participation in Elections," *The American Economic Review*, 2006, *96* (4), 1271–1282.

Feddersen, Timothy J., "Rational Choice Theory and the Paradox of Not Voting," *The Journal of Economic Perspectives*, 2004, *18* (1), 99–112.

Ferejohn, J., "Incumbent Performance and Electoral Control," *Public Choice*, 1986, *50*, 5–25.

Golosov, M., A. Tsyvinski, and N. Werquin, "A Variational Approach to the Analysis of Tax Systems," Working Paper 20780, National Bureau of Economic Research December 2014.

Harsanyi, John C., "Rule utilitarianism, rights, obligations and the theory of rational behavior," *Theory and Decision*, Jun 1980, *12* (2), 115–133.

Hellwig, M.F., "A Contribution to the Theory of Optimal Utilitarian Income Taxation," *Journal of Public Economics*, 2007, *91*, 1449–1477.

Lindbeck, A. and J. Weibull, "Balanced-budget Redistribution as the Outcome of Political Competition," *Public Choice*, 1987, *52*, 273–297.

Lockwood, Benjamin B and Matthew Weinzierl, "Positive and normative judgments implicit in US tax policy, and the costs of unequal growth and recessions," *Journal of Monetary Economics*, 2016, *77*, 30–47.

Meltzer, A. and S. Richard, "A Rational Theory of the Size of Government," *Journal of Political Economy*, 1981, *89*, 914–927.

Mirrlees, J., "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 1971, *38*, 175–208.

Myerson, Roger B., "Optimal Auction Design," *Mathematics of Operations Research*, 1981, *6* (1), 58–73.

Osborne, M. and A. Slivinski, "A Model of Political Competition With Citizen-Candidates," *Quarterly Journal of Economics*, 1996, *111*, 65–96.

Piketty, T., "La redistribution fiscale face au chômage," *Revue française d'économie*, 1997, *12*, 157–201.

Riker, William H. and Peter C. Ordeshook, "A Theory of the Calculus of Voting," *The American Political Science Review*, 1968, *62* (1), 25–42.

Roberts, K., "Voting over Income Tax Schedules," *Journal of Public Economics*, 1977, *8.*

Rochet, J.-C. and L.A. Stole, "Nonlinear Pricing with Random Participation," *Review of Economic Studies*, 2002, *69* (1), 277–311.

Roell, A.A., "Voting over Nonlinear Income Tax Schedules," *Working Paper*, 2012.

Saez, E., "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001, *68*, 205–229.

_ **and S. Stantcheva**, "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, 2016, *106* (1), 24–45.

Samuelson, P., "Arrow's Mathematical Politics," in S. Hook, ed., *Human Values and Economic Policy*, New York University Press, 1967.

Weinzierl, Matthew, "The promise of positive optimal taxation: normative diversity and a role for equal sacrifice," *Journal of Public Economics*, 2014, *118*, 128–142.

Zoutman, F.T., B. Jacobs, and E.L.W. Jongen, "Redistributive Politics and the Tyranny of the Middle Class," *Tinbergen Institute Discussion Paper 16-032/VI*, 2016.

A Appendix

Derivation of equation (1)

Given two party proposals p^1 and p^2, and given the turnout for party 2, σ^2, the best response problem of the rule-utilitarian supporters of party 1 is to choose σ^1 so as to maximize the expected value of the following expression

$$\mathbf{1}\left(\tilde{V}^1(p^1,p^2,\sigma^1,\tilde{q}^1) \geq \tilde{V}^2(p^1,p^2,\sigma^2,\tilde{q}^2)\right) \ E[B(u(p^1,\omega) - u(p^2,\omega) \mid \omega) \ u(p^1,\omega)]$$

$$+ \left(1 - \mathbf{1}\left(\tilde{V}^1(p^1,p^2,\sigma^1,\tilde{q}^1) \geq \tilde{V}^2(p^1,p^2,\sigma^2,\tilde{q}^2)\right)\right) \times$$

$$E[B(u(p^1,\omega) - u(p^2,\omega) \mid \omega) \ u(p^2,\omega) + \mathcal{E}^1(u(p^1,\omega) - u(p^2,\omega) \mid \omega)]$$

$$-k(\sigma^1) \ E[\tilde{q}^{1u}(\omega) \ B(u(p^1,\omega) - u(p^2,\omega) \mid \omega)] - \kappa \ E[\tilde{q}^{1v}(\omega) \ B(u(p^1,\omega) - u(p^2,\omega) \mid \omega)] \ .$$

In this expression, $\mathbf{1}$ is an indicator function and

$$\mathbf{1}\left(\tilde{V}^1(p^1,p^2,\sigma^1,\tilde{q}^1) \geq \tilde{V}^2(p^1,p^2,\sigma^2,\tilde{q}^2)\right) \ E[B(u(p^1,\omega) - u(p^2,\omega) \mid \omega) \ u(p^1,\omega)]$$

is utilitarian welfare realized by the supporters of party 1 in the event that their wins. Analogously,

$$(1 - \mathbf{1}(\cdot)) \, E[B(u(p^1,\omega) - u(p^2,\omega) \mid \omega) \ u(p^2,\omega) + \mathcal{E}^1(u(p^1,\omega) - u(p^2,\omega) \mid \omega)]$$

is utilitarian welfare if party 2 wins, where

$$\mathcal{E}^1(u(p^1,\omega) - u(p^2,\omega) \mid \omega) := \int \varepsilon_i \, \mathbf{1}(\varepsilon_i \leq u(p^1,\omega) - u(p^2,\omega)) \, b(\varepsilon_i \mid \omega) \, d\varepsilon_i \ ,$$

is the sum of the gains (or losses) that the supporters of party 1 realize in this event because of their idiosyncratic party preference. Voting costs do not depend on which party wins the election.

Upon exploiting the linearity of the expectations operator and dropping terms that do not depend on σ^1, we can equivalently write this optimization problem as follows: Choose $\sigma^1 \in [0,1]$ to maximize

$$\pi^1(p^1,p^2,\sigma^1,\sigma^2) \, W^1(p^1,p^2) - \kappa \, \sigma^1 \, \bar{\Gamma}^{1u}(p^1,p^2) \ , \tag{24}$$

where $\pi^1(p^1,p^2,\sigma^1,\sigma^2)$ is the probability of the event $\tilde{V}^1(p^1,p^2,\sigma^1,\tilde{q}^1) \geq \tilde{V}^2(p^1,p^2,\sigma^2,\tilde{q}^2)$, and

$$W^1(p^1,p^2) = E[G^1_W(u(p^1,\omega) - u(p^2,\omega) \mid \omega)] \ ,$$

where, for any $x \in \mathbb{R}$,

$$G^1_W(x \mid \omega) := B(x \mid \omega) \, x \ - \ \mathcal{E}^1(x \mid \omega) \ ,$$

is the welfare gain that is realized by the supporters of party 1 if their party wins and

$$\bar{\Gamma}^{1u}(p^1, p^2) = E[\bar{q}^{1u}(\omega)\, B(u(p^1, \omega) - u(p^2, \omega) \mid \omega)]$$

is the expected value of the mass of rule-utilitarian supporters of party 1.

Proof of Lemma 1

Under Assumption 2 the parties' vote numbers are, respectively, given by

$$\tilde{V}^1(p^1, p^2, \sigma^1, \sigma^2) \;=\; \sigma^1\, E[\tilde{q}^{1u}(\omega)B(u(p^1, \omega) - u(p^2, \omega) \mid \omega)] \;=:\; \sigma^1\, \tilde{\Gamma}^{1u}(p^1, p^2)\,,$$

and

$$\tilde{V}^2(p^1, p^2, \sigma^1, \sigma^2) \;=\; \sigma^2\, E[\tilde{q}^{2u}(\omega)(1 - B(u(p^1, \omega) - u(p^2, \omega) \mid \omega))] \;=:\; \sigma^2\, \tilde{\Gamma}^{2u}(p^1, p^2)\,.$$

Given p^1 and p^2, the probability that party 1 wins the election is therefore equal to the probability of the event

$$\frac{\sigma^1}{\sigma^2} \geq \frac{\tilde{\Gamma}^{2u}(p^1, p^2)}{\tilde{\Gamma}^{1u}(p^1, p^2)}\,.$$

Given p^1 and p^2, denote by $\mathcal{P}(\cdot \mid p^1, p^2)$ the *cdf* and by $\rho(\cdot \mid p^1, p^2)$ the density of the random variable $\frac{\tilde{\Gamma}^{2u}(p^1, p^2)}{\tilde{\Gamma}^{1u}(p^1, p^2)}$. Thus,

$$\pi^1(p^1, p^2, \sigma^1, \sigma^2) = \mathcal{P}\left(\frac{\sigma^1}{\sigma^2} \mid p^1, p^2\right)\,. \tag{25}$$

We take the party platforms p^1 and p^2 as given and characterize equilibrium turnout. We say that the turnout subgame has an interior equilibrium if $0 < \sigma^{1*}(p^1, p^2) < 1$ and $0 < \sigma^{2*}(p^1, p^2) < 1$. An interior equilibrium is characterized by the first order conditions

$$\pi^1_{\sigma^1}(\cdot)\, W^1(p^1, p^2) - \kappa\, \bar{\Gamma}^{1u}(p^1, p^2) = 0\,, \tag{26}$$

and

$$-\, \pi^1_{\sigma^2}(\cdot)\, W^2(p^1, p^2) - \kappa\, \bar{\Gamma}^{2u}(p^1, p^2) = 0\,. \tag{27}$$

Using equation (25), these first order conditions can also be written as

$$\rho\left(\frac{\sigma^1}{\sigma^2} \mid p^1, p^2\right) \frac{\sigma^1}{\sigma^2} \frac{1}{\sigma^1}\, W^1(p^1, p^2) - \kappa\, \bar{\Gamma}^{1u}(p^1, p^2) = 0\,, \tag{28}$$

and

$$\rho\left(\frac{\sigma^1}{\sigma^2} \mid p^1, p^2\right) \frac{\sigma^1}{\sigma^2} \frac{1}{\sigma^2}\, W^2(p^1, p^2) - \kappa\, \bar{\Gamma}^{2u}(p^1, p^2) = 0\,. \tag{29}$$

Equations (28) and (29) allow us to pin down the equilibrium value of relative turnout,

$$\chi^*(p^1, p^2) := \frac{\sigma^{1*}(p^1, p^2)}{\sigma^{2*}(p^1, p^2)} = \frac{W^1(p^1, p^2)/\bar{\Gamma}^{1u}(p^1, p^2)}{W^2(p^1, p^2)/\bar{\Gamma}^{2u}(p^1, p^2)}\,. \tag{30}$$

as had to be shown.

Proof of Lemma 2

Under Assumption 3,

$$\tilde{\Gamma}^{1u}(p^1, p^2) = (1 - \eta^1)\,\bar{\Gamma}^{1u}(p^1, p^2) \quad \text{and} \quad \tilde{\Gamma}^{2u}(p^1, p^2) = (1 - \eta^2)\,\bar{\Gamma}^{2u}(p^1, p^2)\,.$$

The probability of the event that party 1 wins the election is therefore equal to the probability of the event

$$\sigma^1\,(1 - \eta^1)\,\bar{\Gamma}^{1u}(p^1, p^2) \geq \sigma^2\,(1 - \eta^2)\,\bar{\Gamma}^{2u}(p^1, p^2)$$

or, equivalently,

$$\frac{\sigma^1}{\sigma^2}\,\frac{\bar{\Gamma}^{1u}(p^1, p^2)}{\bar{\Gamma}^{2u}(p^1, p^2)} \geq \frac{1 - \eta_2}{1 - \eta^1}\,.$$

By the definition of $\mathcal{P}$, this probability can be written as

$$\mathcal{P}\left(\frac{\sigma^1}{\sigma^2}\,\frac{\bar{\Gamma}^{1u}(p^1, p^2)}{\bar{\Gamma}^{2u}(p^1, p^2)}\right)\,,$$

where, by Lemma 1, with an interior equilibrium of the participation subgame,

$$\frac{\sigma^1}{\sigma^2} = \chi^*\,.$$

Proof of Proposition 2

Part i) of Proposition 2 follows from the observation that party 1 seeks to maximize $\chi\,M(\bar{\Gamma}^{1u})$ and that party 2 seeks to minimize this expression. With χ viewed as an exogenous quantity and since M is an increasing function, this is equivalent to party 1 choosing p^1 with the objective to maximize $\bar{\Gamma}^{1u}(p^1, p^2)$ and party 2 choosing p^2 with the objective to minimize this expression. Hence, if (p^{1*}, p^{2*}) is a pair of Nash equilibrium policies then these constitute a saddle point of the function $\bar{\Gamma}^{1u} : (p^1, p^2) \mapsto \bar{\Gamma}^{1u}(p^1, p^2)$.

Part ii) follows from taking the endogeneity of turnout into account so that

$$\begin{aligned}
\chi\,M(\bar{\Gamma}^{1u}) &= \chi^*(p^1, p^2)\,\frac{\bar{\Gamma}^{1u}(p^1, p^2)}{\bar{\Gamma}^{2u}(p^1, p^2)} \\
&= \frac{W^1(p^1, p^2)}{W^2(p^1, p^2)}\,.
\end{aligned}$$

where the second equality follows from Lemma 1. Hence, if (p^{1*}, p^{2*}) is a pair of subgame perfect equilibrium policies then these constitute a saddle point of the function $\psi : (p^1, p^2) \mapsto \psi(p^1, p^2) = \frac{W^1(p^1, p^2)}{W^2(p^1, p^2)}$.

Proof of Lemma 3

We begin by stating party 1's best response problem in a way that enables an analysis using a Gateaux differential. Let $y^1 = y^{1*} + \tau^1 h^1$, be a perturbed version of party 1's best response y^{1*}, in which τ^1 is a scalar and $h^1 : \Omega \to \mathbb{R}$ is a function. If y^{1*} is a best response, then, for any perturbation (τ^1, h^1),

$$
\begin{aligned}
&E\left[G_W^1\left(s_v(y^{1*}) - \int_{\underline{\omega}}^{\omega} v_2(y^{1*}(z), z)\, dz - u^2(\omega) \mid \omega\right)\right] \\
&\geq E\left[G_W^1\left(s_v(y^{1*} + \tau^1 h^1) - \int_{\underline{\omega}}^{\omega} v_2(y^{1*}(z) + \tau^1 h^1(z), z)\, dz - u^2(\omega) \mid \omega\right)\right] .
\end{aligned}
\tag{31}
$$

Equivalently, for any function h^1, $\tau^1 = 0$ must be a maximizer of the auxiliary function

$$
A(\tau^1 \mid y^{1*}, y^2) = E\left[G_W^1\left(s_v(y^{1*} + \tau^1 h^1) - \int_{\underline{\omega}}^{\omega} v_2(y^{1*}(z) + \tau^1 h^1(z), z)\, dz - u^2(\omega) \mid \omega\right)\right] .
$$

In the following, we will characterize y^{1*} by analyzing the implications of the requirement that the derivative of this expression with respect to τ^1, evaluated at $\tau^1 = 0$, is equal to zero, i.e.

$$
A_{\tau^1}(y^{1*}, y^2) = 0 ,
\tag{32}
$$

for all functions h^1. Formally, $A_{\tau^1}(y^{1*}, y^2)$ is the Gateaux differential of

$$
E\left[G_W^1\left(s_v(y^1) - \int_{\underline{\omega}}^{\omega} v_2(y^1(z), z)\, dz - u^2(\omega) \mid \omega\right)\right]
$$

in direction h^1 evaluated at $y^1 = y^{1*}$.

We note that

$$
A_{\tau^1}(y^{1*}, y^2) = E\left[g_W^1(\omega \mid y^{1*}, y^2)\left(s_{v,\tau^1}(y^{1*}) - \int_{\underline{\omega}}^{\omega} h^1(z)\, v_{21}(y^{1*}(z), z)\, dz\right)\right] ,
\tag{33}
$$

where

$$
g_W^1(\omega \mid y^{1*}, y^2) := b\left(s_v(y^{1*}) - \int_{\underline{\omega}}^{\omega} v_2(y^{1*}(z), z)\, dz - u^2(\omega) \mid \omega\right)
$$

and

$$
s_{v,\tau^1}(y^{1*}) = E\left[h^1(\omega)\left(1 - v_1(y^{1*}(\omega), \omega) + \tfrac{1-F(\omega)}{f(\omega)} v_{21}(y^{1*}(\omega), \omega)\right)\right]
$$

is the Gateaux differential of the virtual surplus $s_v(y^1)$ in direction h^1 evaluated at $y^1 = y^{1*}$.

Equation (33) can now be rewritten as

$$
\begin{aligned}
A_{\tau^1}(y^{1*}, y^2) = &\ \bar{g}_W^1(\underline{\omega} \mid y^{1*}, y^2) E\left[h^1(\omega)\left(1 - v_1(y^{1*}(\omega), \omega) + \tfrac{1-F(\omega)}{f(\omega)} v_{21}(y^{1*}(\omega), \omega)\right)\right] \\
&- E\left[g_W^1(\omega \mid y^{1*}, y^2) \int_{\underline{\omega}}^{\omega} h^1(z)\, v_{21}(y^{1*}(z), z)\, dz\right]
\end{aligned}
\tag{34}
$$

where, for any $\omega \in \Omega$, $\bar{g}_W^1(\omega \mid y^{1*}, y^2) := E[g_W^1(\omega' \mid y^{1*}, y^2) \mid \omega' \geq \omega]$. Moreover, an integration by

43

parts shows that

$$E\left[g_W^1(\omega \mid y^{1*}, y^2) \int_{\underline{\omega}}^{\omega} h^1(z)\, v_{21}(y^{1*}(z), z)\, dz\right] = E\left[h^1(\omega)\, \bar{g}_W^1(\omega \mid y^{1*}, y^2) \frac{1 - F(\omega)}{f(\omega)} v_{21}(y^{1*}(\omega), \omega)\right] \tag{35}$$

so that condition (32) can equivalently be written as the requirement that, for all functions h^1,

$$E\left[h^1(\omega)\left(1 - v_1(y^{1*}(\omega), \omega) + (1 - \mathcal{G}_W^1(\omega \mid y^{1*}, y^2)) \frac{1 - F(\omega)}{f(\omega)} v_{21}(y^{1*}(\omega), \omega)\right)\right] = 0 \tag{36}$$

where $\mathcal{G}_W^1(\omega \mid y^{1*}, y^2) = \frac{\bar{g}_W^1(\omega \mid y^{1*}, y^2)}{\bar{g}_W^1(\omega \mid y^{1*}, y^2)}$. Condition (36) can hold only if, for all ω,

$$1 - v_1(y^{1*}(\omega), \omega) + (1 - \mathcal{G}_W^1(\omega \mid y^{1*}, y^2)) \frac{1 - F(\omega)}{f(\omega)} v_{21}(y^{1*}(\omega), \omega) = 0\ , \tag{37}$$

or, equivalently, if

$$\frac{1 - v_1(y^{1*}(\omega), \omega)}{v_1(y^{1*}(\omega), \omega)} = -(1 - \mathcal{G}_W^1(\omega \mid y^{1*}, y^2)) \frac{1 - F(\omega)}{f(\omega)} \frac{v_{21}(y^{1*}(\omega), \omega)}{v_1(y^{1*}(\omega), \omega)}\ . \tag{38}$$

Using $T^{1\prime}(y^{1*}(\omega)) = 1 - v_1(y^{1*}(\omega), \omega)$ we can rewrite this equation as

$$\frac{T^{1\prime}(y^{1*}(\omega))}{1 - T^{1\prime}(y^{1*}(\omega))} = -(1 - \mathcal{G}_W^1(\omega \mid y^{1*}, y^2)) \frac{1 - F(\omega)}{f(\omega)} \frac{v_{21}(y^{1*}(\omega), \omega)}{v_1(y^{1*}(\omega), \omega)}\ , \tag{39}$$

which is what had to be shown.

Proof of Theorem 1

Best responses

Let y^2 be an arbitrary, but given function, possibly equal to y^{2*}. Given y^2 we look at the problem to choose y^1 with the objective to maximize

$$\psi(y^1, y^2) = \frac{W^1(y^1, y^2)}{W^2(y^1, y^2)}\ .$$

Suppose that y^{1*} is a solution to that problem. The, it must also be that case that $\tau = 0$ solves the problem to choose a scalar τ^1 with the objective to maximize

$$\psi(y^{1*} + \tau^1\, h^1, y^2) = \frac{W^1(y^{1*} + \tau^1\, h^1, y^2)}{W^2(y^{1*} + \tau^1\, h^1, y^2)}\ .$$

for any given but arbitrary function h^1. That is, we can characterize y^{1*} be the requirement that, for $\tau^1 = 0$,

$$\psi_{\tau^1}(y^{1*} + \tau^1\, h^1, y^2) = 0\ ,$$

44

or, equivalently, that

$$W^1_{\tau^1}(y^{1*}, y^2) \, W^2(y^{1*}, y^2) - W^1(y^{1*}, y^2) \, W^2_{\tau^1}(y^{1*}, y^2) = 0 \;.$$

The following Lemma provides a characterization of $W^1_{\tau^1}$ and $W^2_{\tau^1}$ and of the analogous expressions $W^1_{\tau^2}$ and $W^2_{\tau^2}$ that are relevant for party 2's best response problem. We omit a proof of the Lemma as it would require only an easy adaptation of the arguments in the proof of Lemma 3.

Lemma 5. *The derivative of W^1 with respect to τ^1 evaluated at (y^{1*}, y^2) equals*

$$W^1_{\tau^1}(y^{1*}, y^2) =$$

$$\bar{g}^1_W(\underline{\omega} \mid y^{1*}, y^2) E\left[h^1(\omega)\left\{1 - v_1(y^{1*}(\omega), \omega) + \tfrac{1 - F(\omega)}{f(\omega)}\left(1 - \mathcal{G}^1_W(\omega \mid y^{1*}, y^2)\right) v_{21}(y^{1*}(\omega), \omega)\right\}\right] \;.$$
(40)

The derivative of W^2 with respect to τ^1 evaluated at (y^{1*}, y^2) equals

$$W^2_{\tau^1}(y^{1*}, y^2) =$$

$$-\bar{g}^2_W(\underline{\omega} \mid y^{1*}, y^2) E\left[h^1(\omega)\left\{1 - v_1(y^{1*}(\omega), \omega) + \tfrac{1 - F(\omega)}{f(\omega)}\left(1 - \mathcal{G}^2_W(\omega \mid y^{1*}, y^2)\right) v_{21}(y^{1*}(\omega), \omega)\right\}\right] \;.$$
(41)

The derivative of W^1 with respect to τ^2 evaluated at (y^1, y^{2*}) equals

$$W^1_{\tau^2}(y^1, y^{2*}) =$$

$$-\bar{g}^1_W(\underline{\omega} \mid y^1, y^{2*}) E\left[h^2(\omega)\left\{1 - v_1(y^{2*}(\omega), \omega) + \tfrac{1 - F(\omega)}{f(\omega)}\left(1 - \mathcal{G}^1_W(\omega \mid y^1, y^{2*})\right) v_{21}(y^{2*}(\omega), \omega)\right\}\right] \;.$$
(42)

The derivative of W^2 with respect to τ^2 evaluated at (y^1, y^{2*}) equals

$$W^2_{\tau^2}(y^1, y^{2*}) =$$

$$\bar{g}^2_W(\underline{\omega} \mid y^1, y^{2*}) E\left[h^2(\omega)\left\{1 - v_1(y^{2*}(\omega), \omega) + \tfrac{1 - F(\omega)}{f(\omega)}\left(1 - \mathcal{G}^2(\omega \mid y^1, y^{2*})\right) v_{21}(y^{2*}(\omega), \omega)\right\}\right] \;.$$
(43)

We introduce additional notation: For any pair (y^1, y^2), let

$$\mathcal{Q}(y^1, y^2) := \bar{g}^1_W(\underline{\omega} \mid y^1, y^2) \, W^2(y^1, y^2)$$

and

$$\mathcal{R}(y^1, y^2) := \bar{g}^2_W(\underline{\omega} \mid y^1, y^2) \, W^1(y^1, y^2) \;.$$

Also let

$$\mathcal{S}(y^1, y^2) = \mathcal{Q}(y^1, y^2) + \mathcal{R}(y^1, y^2) \;,$$

$$\mathbf{q}(y^1, y^2) := \frac{\mathcal{Q}(y^1, y^2)}{\mathcal{S}(y^1, y^2)} \;,$$

45

and

$$\mathbf{r}(y^1, y^2) := \frac{\mathcal{R}(y^1, y^2)}{\mathcal{S}(y^1, y^2)} \, .$$

Finally, let

$$\mathcal{G}_{SP}(\omega \mid y^1, y^2) := \mathbf{q}(y^1, y^2) \, \mathcal{G}_W^1(\omega \mid y^1, y^2) + \mathbf{r}(y^1, y^2) \, \mathcal{G}_W^2(\omega \mid y^1, y^2) \, .$$

Using this notation and collecting the terms in Lemma 5 yields the following Lemma.

Lemma 6. *The sign of the derivative of ψ with respect to τ^1 evaluated at (y^{1*}, y^2) is equal to the sign of*

$$W_{\tau^1}^1(y^{1*}, y^2) \, W^2(y^{1*}, y^2) - W^1(y^{1*}, y^2) \, W_{\tau^1}^2(y^{1*}, y^2) =$$

$$\mathcal{S}(y^{1*}, y^2) E \left[h^1(\omega) \left\{ 1 - v_1(y^{1*}(\omega), \omega) + \tfrac{1-F(\omega)}{f(\omega)} \left(1 - \mathcal{G}_{SP}(\omega \mid y^{1*}, y^2) \right) v_{21}(y^{1*}(\omega), \omega) \right\} \right] \, . \tag{44}$$

The sign of the derivative of ψ with respect to τ^2 evaluated at (y^1, y^{2*}) is equal to the sign of

$$W_{\tau^2}^1(y^1, y^{2*}) \, W^2(y^1, y^{2*}) - W^1(y^1, y^{2*}) \, W_{\tau^2}^2(y^1, y^{2*}) =$$

$$\mathcal{S}(y^1, y^{2*}) E \left[h^2(\omega) \left\{ 1 - v_1(y^{2*}(\omega), \omega) + \tfrac{1-F(\omega)}{f(\omega)} \left(1 - \mathcal{G}_{SP}(\omega \mid y^1, y^{2*}) \right) v_{21}(y^{2*}(\omega), \omega) \right\} \right] \, . \tag{45}$$

Subgame perfect equilibrium

If y^{1*} and y^{2*} are mutually best responses, then it must be the case that

$$W_{\tau^1}^1(y^{1*}, y^{2*}) \, W^2(y^{1*}, y^{2*}) - W^1(y^{1*}, y^{2*}) \, W_{\tau^1}^2(y^{1*}, y^{2*}) = 0 \, ,$$

for all functions h^1 and

$$W_{\tau^2}^1(y^{1*}, y^{2*}) \, W^2(y^{1*}, y^{2*}) - W^1(y^{1*}, y^{2*}) \, W_{\tau^2}^2(y^{1*}, y^{2*}) = 0 \, ,$$

for all functions h^2. By Lemma 6 this requires that, for all ω,

$$1 - v_1(y^{1*}(\omega), \omega) + \tfrac{1-F(\omega)}{f(\omega)} \left(1 - \mathcal{G}_{SP}(\omega \mid y^{1*}, y^{2*}) \right) v_{21}(y^{1*}(\omega), \omega) = 0 \, , \tag{46}$$

and

$$1 - v_1(y^{2*}(\omega), \omega) + \tfrac{1-F(\omega)}{f(\omega)} \left(1 - \mathcal{G}_{SP}(\omega \mid y^{1*}, y^{2*}) \right) v_{21}(y^{2*}(\omega), \omega) = 0 \, . \tag{47}$$

An inspection of (46) and (47) yields the following Lemma.

Lemma 7. *In any interior subgame perfect equilibrium, $y^{1*} = y^{2*}$.*

To complete the proof of Theorem 1 we note that that

$$\mathcal{G}_{SP}(\omega \mid y^{1*}, y^{2*}) = \lambda^1 \, \mathcal{G}_W^{1s}(\omega) - \lambda^2 \, \mathcal{G}_W^{2s}(\omega) \, ,$$

46

where the expressions on the right hand side of this equality have been defined in the body of the text. Using this observation, rearranging the terms in (46), and using that $T'(y(\omega)) = v_1(y(\omega), \omega)$ yields equation (13) in the body of the text.

Proof of Proposition 3

Suppose that a pair of policies (y^{1*}, y^{2*}) satisfies the first order conditions (46) and (46) in the proof of Proposition 1. We seek to show that (y^{1*}, y^{2*}) is a saddle point of the function

$$\psi(y^1, y^2) = \frac{W^1(y^1, y^2)}{W^2(y^1, y^2)} .$$

We now state this saddle point condition in a way that enables an analysis using functional derivatives. Let $y^1 = y^{1*} + \tau^1 \, h^1$, be a perturbed version of y^{1*}, in which τ^1 is a scalar and $h^1 : \Omega \to \mathbb{R}$ is a function. Analogously, let $y^2 = y^{1*} + \tau^1 \, h^1$, be a perturbed version of y^2. The saddle point condition according to which, for all (y^1, y^2),

$$\psi(y^1, y^{2*}) \leq \psi(y^{1*}, y^{2*}) \leq \psi(y^{1*}, y^2)$$

can therefore be written as: For any pair of perturbations (τ^1, h^1) and (τ^2, h^2),

$$\psi(y^{1*} + \tau^1 \, h^1, y^{2*}) \leq \psi(y^{1*}, y^{2*}) \leq \psi(y^{1*}, y^{*2} + \tau^2 \, h^2) . \tag{48}$$

Equivalently, for all functions (h^1, h^2), the point $(\tau^1, \tau^2) = (0, 0)$ must be a saddle-point of

$$\psi(y^{1*} + \tau^1 \, h^1, y^{2*} + \tau^2 \, h^2) = \frac{W^1(y^{1*} + \tau^1 \, h^1, y^{*2} + \tau^2 \, h^2)}{W^2(y^{1*} + \tau^1 \, h^1, y^{*2} + \tau^2 \, h^2)} .$$

Having a saddle point requires that all entries of the Jacobi-matrix

$$J_\psi(y^{1*}, y^{2*}) = \begin{pmatrix} \psi_{\tau^1}(y^{1*}, y^{2*}) \\ \psi_{\tau^2}(y^{1*}, y^{2*}) \end{pmatrix}$$

are equal to zero and that the Hessian

$$H_\psi(y^{1*}, y^{2*}) = \begin{pmatrix} \psi_{\tau^1, \tau^1}(y^{1*}, y^{2*}) & \psi_{\tau^1, \tau^2}(y^{1*}, y^{2*}) \\ \psi_{\tau^1, \tau^2}(y^{1*}, y^{2*}) & \psi_{\tau^2, \tau^2}(y^{1*}, y^{2*}) \end{pmatrix}$$

is indefinite. The policies (y^{1*}, y^{2*}) satisfying the first order conditions (46) and (46) is equivalent to all entries of the Jacobi-matrix being equal to zero. Hence, what remains to be shown is that $H_\psi(y^{1*}, y^{2*})$ is indefinite. To this end, it suffices to show that $\psi_{\tau^1, \tau^1}(y^{1*}, y^{2*}) < 0$, and $\psi_{\tau^2, \tau^2}(y^{1*}, y^{2*}) > 0$. These two inequalities can be shown to hold provided that

$$\frac{\partial}{\partial \tau^1} \left\{ W^1_{\tau^1}(y^{1*}, y^{2*}) \, W^2(y^{1*}, y^{2*}) - W^1(y^{1*}, y^{2*}) \, W^2_{\tau^1}(y^{1*}, y^{2*}) \right\} < 0 ,$$

and

$$\frac{\partial}{\partial \tau^2} \left\{ W_{\tau^2}^1(y^{1*}, y^{2*}) \, W^2(y^{1*}, y^{2*}) - W^1(y^{1*}, y^{2*}) \, W_{\tau^2}^2(y^{1*}, y^{2*}) \right\} > 0 \,,$$

or, equivalently, if

$$W_{\tau^1, \tau^1}^1(y^{1*}, y^{2*}) \, W^2(y^{1*}, y^{2*}) - W^1(y^{1*}, y^{2*}) \, W_{\tau^1, \tau^1}^2(y^{1*}, y^{2*}) < 0 \,, \tag{49}$$

and

$$W_{\tau^2, \tau^2}^1(y^{1*}, y^{2*}) \, W^2(y^{1*}, y^{2*}) - W^1(y^{1*}, y^{2*}) \, W_{\tau^2, \tau^2}^2(y^{1*}, y^{2*}) > 0 \,. \tag{50}$$

Exploiting Assumption 1. Recall that, under Assumption 1,

$$W^1(y^1, y^2) = E\left[s^1(\omega) \left(u^1(\omega) - u^2(\omega) \right) + \frac{1}{2} \, \alpha(\omega) \left(u^1(\omega) - u^2(\omega) \right)^2 + \frac{1}{2} \left(\frac{s^1(\omega)^2}{\alpha(\omega)} \right) \right]$$

and

$$W^2(y^1, y^2) = E\left[s^2(\omega) \left(u^2(\omega) - u^1(\omega) \right) + \frac{1}{2} \, \alpha(\omega) \left(u^2(\omega) - u^1(\omega) \right)^2 + \frac{1}{2} \left(\frac{s^2(\omega)^2}{\alpha(\omega)} \right) \right] \,.$$

Also note that for $u^1(\omega) - u^2(\omega) = z$

$$\bar{g}_W^1(\omega \mid y^1, y^2) := s^1(\omega) + \alpha(\omega) \, z \,,$$

and for $u^2(\omega) - u^1(\omega) = z'$,

$$\bar{g}_W^2(\omega \mid y^1, y^2) := s^2(\omega) + \alpha(\omega) \, z \,.$$

It follows from Theorem 1 that $y^{1*} = y^{2*}$, hence $W^1(y^{1*}, y^{2*}) > 0$ and $W^2(y^{1*}, y^{2*}) > 0$. Sufficient conditions for the validity of (49) and (50) are therefore that

$$W_{\tau^1, \tau^1}^1(y^{1*}, y^{2*}) < 0 \quad \text{and} \quad W_{\tau^1, \tau^1}^2(y^{1*}, y^{2*}) > 0 \,, \tag{51}$$

and

$$W_{\tau^2, \tau^2}^1(y^{1*}, y^{2*}) > 0 \quad \text{and} \quad W_{\tau^2, \tau^2}^2(y^{1*}, y^{2*}) < 0 \,. \tag{52}$$

We can now use the expressions for $W_{\tau^1}^1$, $W_{\tau^1}^2$, $W_{\tau^2}^1$ and $W_{\tau^2}^2$ in Lemma 5 to compute W_{τ^1, τ^1}^1, W_{τ^1, τ^1}^2, W_{τ^2, τ^2}^1 and W_{τ^2, τ^2}^2. If we evaluate the resulting expressions in the limit case as $\alpha(\omega)$ arbitrarily close to zero for all ω, we can verify that (51) and (52) indeed hold. For instance, we then find that

$$W_{\tau^1, \tau^1}^1(y^{1*}, y^{2*}) = \bar{s}^1(\underline{\omega}) E\left[h^1(\omega)^2 \left(-v_{11}(y^*(\omega), \omega) + \frac{1 - F(\omega)}{f(\omega)} \left(1 - \frac{\bar{s}^1(\omega)}{\bar{s}^1(\underline{\omega})} \right) v_{211}(y^*(\omega), \omega) \right) \right] \,,$$

where $\bar{s}^1(\omega) := E\left[s^1(\omega') \mid \omega' \geq \omega \right]$. With $s^1(\omega) = \frac{1}{2}$, for all ω,

$$1 - \frac{\bar{s}^1(\omega)}{\bar{s}^1(\underline{\omega})} = 0 \,,$$

for all ω, so that

$$W^1_{\tau^1,\tau^1}(y^{1*}, y^{2*}) = \bar{s}^1(\underline{\omega})E\left[h^1(\omega)^2\left(-v_{11}(y^*(\omega),\omega)\right)\right] < 0\,.$$

Variational approach: using tax schedules as the policy domain

We briefly sketch an alternative equilibrium characterization in which the policy domain is taken to be the set of tax schedules – rather than the set of non-decreasing earnings functions. This makes it possible to relate our analysis more directly to standard concepts in public finance such as, e.g., the marginal cost of public funds. More importantly, this turns out to be particularly useful for the analysis of whether the two parties' policies are strategic substitutes or complements.

As we explained in Section 3.1, we can directly let parties choose nonlinear tax schedules, i.e., twice continuously differentiable functions $T^j : \mathbb{R}_+ \to \mathbb{R}$. To prevent parties from arbitrarily raising agents' utilities by lowering their taxes, we must impose that any tax reform that a party can propose is budget-neutral. Without loss of generality, this is equivalent to imposing that any extra revenue (or loss) generated by the tax reform that we are considering, taking into account the behavioral responses it induces on labor supply, is rebated (or taxed away) lump-sum and uniformly to the population. Since there is a one-to-one map between earnings functions and tax schedules, we can replace the arguments (y^1, y^2) of all the functions we consider with (T^1, T^2). Moreover, since there is a one-to-one map between types and incomes for a given tax schedule, we can use interchangeably the arguments ω or $y^1 \equiv y^1(\omega)$ for all of our variables.

Consider for instance the problem of choosing T^1 so as to maximize $W^1(T^1, T^2)$, where

$$
\begin{aligned}
W^1(T^1, T^2) &= \mathbb{E}[G^1_W(u(T^1, y^1(\omega)) - u(T^2, y^2(\omega)) \mid \omega)] \\
&= \mathbb{E}[G^1_W(u(T^1, y^1) - u(T^2, y^2) \mid y^1)]\,,
\end{aligned}
$$

where the indirect utility of an agent with type ω and income $y^j(\omega)$ under the policy T^j is given by $u\left(T^j, y^j(\omega)\right) = y^j(\omega) - T^j\left(y^j(\omega)\right) - v\left(y^j(\omega), \omega\right)$. The second equality follows from a change of variables from types ω to incomes $y^1(\omega)$, and the expectation is over the incomes y^1 that are realized if party 1 is elected.

Consider a reform of the platform T^1 in the direction h^1 along with a lump-sum rebate R^1, so that the tax function proposed by party 1 becomes $y \mapsto T^1(y) + \tau^1\left[h^1(y) - R^1\right]$. By the envelope theorem, the utility of this agent decreases one-for-one with the change in the total tax payment, that is, its Gateaux differential is given by

$$u_{\tau^1}(T^1, y^1(\omega)) = -h^1(y^1(\omega)) + R^1. \tag{53}$$

Moreover, the net revenue generated by the tax reform h^1, and hence the lump-sum tax rebate, is

given by[13]

$$R^1 = \int_{\underline{\omega}}^{\bar{\omega}} \left[h^1\left(y^1\left(\omega\right)\right) - \frac{T^{1\prime}\left(y\left(\omega\right)\right)}{1 - T^{1\prime}\left(y\left(\omega\right)\right)} y\left(\omega\right) \tilde{\varepsilon}\left(\omega\right) h^{1\prime}\left(y\left(\omega\right)\right) \right] f\left(\omega\right) d\omega$$

$$= \int_{\underline{y}^1}^{\bar{y}^1} \left[\left(1 - F_{Y^1}(y^1)\right) - \frac{T^{1\prime}(y^1)}{1 - T^{1\prime}(y^1)} y^1 \tilde{\varepsilon}(y^1) f_{Y^1}(y^1) \right] h^{1\prime}(y^1) dy^1,$$

where we define the taxable income elasticity along the nonlinear budget constraint by $\tilde{\varepsilon}\left(\omega\right) = \frac{\varepsilon(\omega)}{1+\varepsilon(\omega)\varphi(y(\omega))}$ with $\varepsilon\left(\omega\right) = \frac{v_1(y(\omega),\omega)}{y(\omega)v_{11}(y(\omega),\omega)}$ and $\varphi\left(y\right) = \frac{yT^{1\prime\prime}(y)}{1-T^{1\prime}(y)}$. The second equality follows from a change of variables from types ω to incomes $y^1\left(\omega\right)$, using the distribution of incomes F_{Y^1} that arises if party 1 implements the policy T^1, and an integration by parts. The change in tax revenue is composed of the sum of two terms: a mechanical effect, coming from the fact that each agent ω pays a higher tax bill $h^1\left(y^1\left(\omega\right)\right)$, absent any behavioral responses; and a behavioral effect, according to which an increase in the marginal tax rate by $h^{1\prime}\left(y\left(\omega\right)\right)$ reduces the labor supply and hence the income of agent ω by $\frac{y(\omega)}{1-T^{j\prime}(y(\omega))}\tilde{\varepsilon}\left(\omega\right)$, of which a share $T^{1\prime}\left(y\left(\omega\right)\right)$ accrues to the government.

Collecting these terms, we obtain that in response to a reform of the platform T^1 in the direction h^1, along with the lump-sum rebate that ensures budget neutrality, the change in the expected gain $W^1(T^1, T^2)$ of the supporters of Party 1 if they win the election is equal to

$$W^1_{\tau^1}(T^1, T^2) = \mathbb{E}\left[g^1_W(u(T^1, y^1) - u(T^2, y^2) \mid y^1) \times u_{\tau^1}(T^1, y^1)\right]$$

$$= -\mathbb{E}\left[g^1_W(u(T^1, y^1) - u(T^2, y^2) \mid y^1)\, h^1(y^1)\right]$$

$$+ \mathbb{E}\left[g^1_W(u(T^1, y^1) - u(T^2, y^2) \mid y^1)\right] R^1 \,.$$

Finally, an integration by parts implies that we can rewrite the first term in the right hand side of the previous expression as

$$-\int_{\underline{y}^1}^{\bar{y}^1} \left[\int_{y^1}^{\bar{y}^1} g^1_W(u(T^1, y^{1\prime}) - u(T^2, y^{2\prime}) \mid y^{1\prime}) f_{Y^1}(y^{1\prime}) dy^{1\prime} \right] h^{1\prime}\left(y^1\right) dy^1 \,,$$

where in this expression $y^{2\prime}$ denotes the income that an agent who would earn $y^{1\prime}$ under party 1's platform gets if party 2 is elected instead, given the two tax schedules $\left(T^1, T^2\right)$.

The optimum platform T^1 is such that $W^1_{\tau^1}(T^1, T^2) = 0$ for all tax reforms h^1. Straightforward algebra shows that this implies

$$\frac{T^{1\prime}(y^1)}{1 - T^{1\prime}(y^1)} = \frac{1 - F_{Y^1}\left(y^1\right)}{y^1 f_{Y^1}\left(y^1\right)} \left(1 - \mathcal{G}^1_W(y^1 \mid T^1, T^2)\right) \frac{1}{\tilde{\varepsilon}\left(y^1\right)}, \quad (54)$$

where we let $\bar{g}^1_W(y^1 \mid T^1, T^2) = \mathbb{E}[g^1_W(u(T^1, y^{1\prime}) - u(T^2, y^{2\prime}) \mid y^{1\prime}) \mid y^{1\prime} \geq y^1]$ and

$$\mathcal{G}^1_W(y^1 \mid T^1, T^2) := \frac{\bar{g}^1_W(y^1 \mid T^1, T^2)}{\bar{g}^1_W(\underline{y}^1 \mid T^1, T^2)} \,.$$

This formula is equivalent to that we derived using mechanism design arguments in Lemma 3.

---

[13]See, e.g., Saez (2001), Golosov, Tsyvinski and Werquin (2014).

In the standard social welfare-maximization problem, the marginal value of public funds is given by $\lambda_W := \mathbb{E}[g(u(T, y(\omega)))]$, where G is the concave social welfare function and g is its derivative. That is, λ_W is the value of rebating lump-sum one unit of tax revenue uniformly to the entire population. Here, the corresponding variable is $\lambda_{G^1} := \mathbb{E}[g_W^1(u(T^1, y^1(\omega)) - u(T^2, y^2(\omega)) \mid \omega)]$ which appears in the denominator of $\mathcal{G}_W^1(y \mid T^1, T^2)$. It is the total political advantage that party 1 can secure over party 2 by rebating lump-sum one unit of tax revenue uniformly to the entire population.

Proof of Proposition 5

We start by the following lemma, the proof of which is straightforward algebra.

Lemma 8. *Under Assumption 1,*

$$\mathcal{E}^1(x \mid \omega) = \frac{1}{2} \, \alpha(\omega) \left(x^2 - \left(\frac{s^1(\omega)}{\alpha(\omega)} \right)^2 \right) \quad and \quad \mathcal{E}^2(x \mid \omega) = \frac{1}{2} \, \alpha(\omega) \left(\left(\frac{s^2(\omega)}{\alpha(\omega)} \right)^2 - x^2 \right),$$

$$W^1(y^1, y^2) = E\left[s^1(\omega) \left(u^1(\omega) - u^2(\omega) \right) + \frac{1}{2} \, \alpha(\omega) \left(u^1(\omega) - u^2(\omega) \right)^2 + \frac{1}{2} \left(\frac{s^1(\omega)^2}{\alpha(\omega)} \right) \right]$$

and

$$W^2(y^1, y^2) = E\left[s^2(\omega) \left(u^2(\omega) - u^1(\omega) \right) + \frac{1}{2} \, \alpha(\omega) \left(u^2(\omega) - u^1(\omega) \right)^2 + \frac{1}{2} \left(\frac{s^2(\omega)^2}{\alpha(\omega)} \right) \right] .$$

In the case of exogenous turnout, we have

$$\bar{b}^u \left(y^1 \mid T^1, T^2 \right) \quad = \quad \int_{y^1}^{\bar{y}^1} \bar{q}^u \left(y^{1\prime} \right) b \left(u \left(T^1, y^{1\prime} \right) - u \left(T^1, y^{2\prime} \right) \mid y^{2\prime} \right) \frac{f_{Y^1}\left(y^{1\prime} \right)}{1 - F_{Y^1}(y^1)} dy^{1\prime},$$

where in the integral we used a change of variables from types ω' to incomes $y^{1\prime} = y^1 (\omega')$ under party 1's policy T^1, and where $y^{2\prime}$ denotes the income that the corresponding type ω' would earn under party 2's policy T^2. The Gateaux differential of $\bar{b}_{\tau^2}^u \left(y^1 \mid T^1, T^2 \right)$ in the direction h^2 is thus given by

$$\bar{b}_{\tau^2}^u \left(y^1 \mid T^1, T^2 \right) \quad = \quad \int_{y^1}^{\bar{y}^1} h^2 \left(y^{2\prime} \right) \bar{q}^u \left(y^{1\prime} \right) b \left(u \left(T^1, y^{1\prime} \right) - u \left(T^2, y^{2\prime} \right) \mid y^{1\prime} \right) \frac{f_{Y^1}\left(y^{1\prime} \right)}{1 - F_{Y^1}(y^1)} dy^{1\prime},$$

since we argued in Section 4.2.1 that $u_{\tau^2} \left(T^2, y^{2\prime} \right) = -h^2 \left(y^{2\prime} \right)$. The weights in party 1's best

response are given by $\mathcal{G}_N\left(y^1 \mid T^1, T^2\right) = \frac{\bar{b}^u\left(y^1 \mid T^1, T^2\right)}{\bar{b}^u\left(\underline{y}^1 \mid T^1, T^2\right)}$, so that their Gateaux differential is equal to

$$
\begin{aligned}
\mathcal{G}_{N,\tau^2}\left(y^1 \mid T^1, T^2\right) &= \frac{\int_{y^1}^{\bar{y}^1} h^2\left(y^{2\prime}\right)\bar{q}^u\left(y^{1\prime}\right) b\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right)\mid y^{1\prime}\right)\frac{f_{Y^1}\left(y^{1\prime}\right)}{1 - F_{Y^1}\left(y^1\right)} dy^{1\prime}}{\int_{\underline{y}^1}^{\bar{y}^1} \bar{q}^u(y^{1\prime}) b(u(T^1, y^{1\prime}) - u(T^1, y^{2\prime})\mid y^{2\prime}) f_{Y^1}(y^{1\prime}) dy^{1\prime}} \\
&\quad - \mathcal{G}_N\left(y^1 \mid T^1, T^2\right)\frac{\int_{\underline{y}^1}^{\bar{y}^1} h^2\left(y^{2\prime}\right)\bar{q}^u\left(y^{1\prime}\right) b\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right)\mid y^{1\prime}\right) f_{Y^1}\left(y^{1\prime}\right) dy^{1\prime}}{\int_{\underline{y}^1}^{\bar{y}^1} \bar{q}^u(y^{1\prime}) b(u(T^1, y^{1\prime}) - u(T^1, y^{2\prime})\mid y^{2\prime}) f_{Y^1}(y^{1\prime}) dy^{1\prime}} \\
&= \mathcal{G}_N\left(y^1 \mid T^1, T^2\right)\left[\frac{\int_{y^1}^{\bar{y}^1} h^2\left(y^{2\prime}\right)\bar{q}^u\left(y^{1\prime}\right) b\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right)\mid y^{1\prime}\right)\frac{f_{Y^1}\left(y^{1\prime}\right)}{1 - F_{Y^1}\left(y^1\right)} dy^{1\prime}}{\int_{y^1}^{\bar{y}^1} \bar{q}^u(y^{1\prime}) b(u(T^1, y^{1\prime}) - u(T^2, y^{2\prime})\mid y^{1\prime})\frac{f_{Y^1}\left(y^{1\prime}\right)}{1 - F_{Y^1}\left(y^1\right)} dy^{1\prime}}\right. \\
&\quad \left. - \frac{\int_{\underline{y}^1}^{\bar{y}^1} h^2\left(y^{2\prime}\right)\bar{q}^u\left(y^{1\prime}\right) b\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right)\mid y^{1\prime}\right) f_{Y^1}\left(y^{1\prime}\right) dy^{1\prime}}{\int_{\underline{y}^1}^{\bar{y}^1} \bar{q}^u(y^{1\prime}) b(u(T^1, y^{1\prime}) - u(T^1, y^{2\prime})\mid y^{2\prime}) f_{Y^1}(y^{1\prime}) dy^{1\prime}}\right].
\end{aligned}
$$

This proves equation (18). Suppose now that the initial equilibrium is symmetric, so that $T^1 = T^2$ and hence $u\left(T^1, y^1\left(\omega\right)\right) = u\left(T^2, y^2\left(\omega\right)\right)$ for all $\omega \in \Omega$, and that $\bar{q}^u\left(\omega\right)$, $b\left(0 \mid \omega\right)$, and $b'\left(0 \mid \omega\right) := b'\left(0\right)$ are constant. We get

$$
\frac{\mathcal{G}_{N,\tau^2}\left(y^1 \mid T^1, T^2\right)}{\mathcal{G}_N\left(y^1 \mid T^1, T^2\right)} = \frac{b'\left(0\right)\mathbb{E}\left[h^2\left(y^2\left(\omega'\right)\right) \mid \omega' \geq \omega\right]}{b\left(0\right)} - \frac{b'\left(0\right)\mathbb{E}\left[h^2\left(y^2\left(\omega'\right)\right)\right]}{b\left(0\right)},
$$

which leads to equation (19).

Now suppose that turnout is endogenous. In this case, we first have

$$
g_W^1\left(x \mid \omega\right) = \frac{d}{dx}\left[B\left(x \mid \omega\right) x - \int_{-\infty}^x \varepsilon_i b\left(\varepsilon_i \mid \omega\right) d\varepsilon_i\right] = B\left(x \mid \omega\right),
$$

$$
\mathcal{G}_W^1\left(y^1 \mid T^1, T^2\right) = \frac{E\left[B\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right) \mid y^{1\prime} \geq y^1\right]}{E\left[B\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right)\right]}.
$$

The Gateaux differential of $\mathcal{G}_W^1$ in the direction h^2 reads

$$
\begin{aligned}
\mathcal{G}_{W,\tau^2}^1\left(y^1 \mid T^1, T^2\right) &= \frac{E\left[h^2\left(y^{2\prime}\right) b\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right) \mid y^{1\prime} \geq y^1\right]}{E\left[B\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right)\right]} \\
&\quad - \mathcal{G}_W^1\left(y^1 \mid T^1, T^2\right)\frac{E\left[h^2\left(y^{2\prime}\right) b\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right)\right]}{E\left[B\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right)\right]} \\
&= \mathcal{G}_W^1\left(y^1 \mid T^1, T^2\right)\left\{\frac{E\left[h^2\left(y^{2\prime}\right) b\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right) \mid y^{1\prime} \geq y^1\right]}{E\left[B\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right) \mid y^{1\prime} \geq y^1\right]}\right. \\
&\quad \left. - \frac{E\left[h^2\left(y^{2\prime}\right) b\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right)\right]}{E\left[B\left(u\left(T^1, y^{1\prime}\right) - u\left(T^2, y^{2\prime}\right) \mid y^{1\prime}\right)\right]}\right\}
\end{aligned}
$$

Second, we have

$$g_W^2 (x \mid \omega) = \frac{d}{dx} \left[(1 - B(-x \mid \omega)) x + \int_{-x}^{\infty} \varepsilon_i b(\varepsilon_i \mid \omega) d\varepsilon_i \right]$$

$$= 1 - B(-x \mid \omega) + 2xb(-x \mid \omega),$$

$$\mathcal{G}_W^1 \left(y^1 \mid T^1, T^2 \right) = \frac{1 - E\left[B\left(\Delta u' \mid y^{1\prime} \right) \mid y^{1\prime} \geq y^1 \right] - 2E\left[(\Delta u') b\left(\Delta u' \mid y^{1\prime} \right) \mid y^{1\prime} \geq y^1 \right]}{1 - E\left[B\left(\Delta u' \mid y^{1\prime} \right) \right] - 2E\left[(\Delta u') b\left(\Delta u' \mid y^{1\prime} \right) \right]},$$

where we let $\Delta u' := u\left(T^1, y^{1\prime} \right) - u\left(T^2, y^{2\prime} \right)$, and hence

$$\frac{\mathcal{G}_{W,\tau^2}^2 \left(y^1 \mid T^1, T^2 \right)}{\mathcal{G}_W^2 \left(y^1 \mid T^1, T^2 \right)} = - \frac{3E\left[h^2 \left(y^{2\prime} \right) b\left(\Delta u' \mid y^{1\prime} \right) \mid y^{1\prime} \geq y^1 \right] + 2E\left[h^2 \left(y^{2\prime} \right) (\Delta u') b'\left(\Delta u' \mid y^{1\prime} \right) \mid y^{1\prime} \geq y^1 \right]}{1 - E\left[B\left(\Delta u' \mid y^{1\prime} \right) \mid y^{1\prime} \geq y^1 \right] - 2E\left[(\Delta u') b\left(\Delta u' \mid y^{1\prime} \right) \mid y^{1\prime} \geq y^1 \right]}$$

$$+ \frac{3E\left[h^2 \left(y^{2\prime} \right) b\left(\Delta u' \mid y^{1\prime} \right) \right] + 2E\left[h^2 \left(y^{2\prime} \right) (\Delta u') b\left(\Delta u' \mid y^{1\prime} \right) \right]}{1 - E\left[B\left(\Delta u' \mid y^{1\prime} \right) \right] - 2E\left[(\Delta u') b\left(\Delta u' \mid y^{1\prime} \right) \right]}.$$

Third, we have

$$\mathrm{d}_{\tau^2} \frac{W^1 \left(T^1, T^2 \right)}{W^2 \left(T^1, T^2 \right)} = \frac{W^1 \left(T^1, T^2 \right)}{W^2 \left(T^1, T^2 \right)} \left\{ \frac{E\left[h^2 \left(y^{2\prime} \right) g_W^1 \left(\Delta u' \mid y^{1\prime} \right) \right]}{E\left[G_W^1 \left(\Delta u' \mid y^{1\prime} \right) \right]} + \frac{E\left[h^2 \left(y^{2\prime} \right) g_W^2 \left(\Delta u' \mid y^{1\prime} \right) \right]}{E\left[G_W^2 \left(\Delta u' \mid y^{1\prime} \right) \right]} \right\},$$

$$\mathrm{d}_{\tau^2} \frac{\bar{g}_W^2 \left(\underline{y}^1 \mid T^1, T^2 \right)}{\bar{g}_W^1 \left(\underline{y}^1 \mid T^1, T^2 \right)} = \frac{-3E\left[h^2 \left(y^{2\prime} \right) b\left(\Delta u' \mid y^{1\prime} \right) \right] - 2E\left[h^2 \left(y^{2\prime} \right) (\Delta u') b'\left(\Delta u' \mid y^{1\prime} \right) \right]}{E\left[B\left(\Delta u' \mid y^{1\prime} \right) \right]}$$

$$+ \frac{\bar{g}_W^2 \left(\underline{y}^1 \mid T^1, T^2 \right)}{\bar{g}_W^1 \left(\underline{y}^1 \mid T^1, T^2 \right)} \frac{E\left[h^2 \left(y^{2\prime} \right) b\left(\Delta u' \mid y^{1\prime} \right) \right]}{E\left[B\left(\Delta u' \mid y^{1\prime} \right) \right]},$$

$$\frac{\mathrm{d}\lambda^1 \left(T^1, T^2 \right)}{\left(\lambda^1 \left(T^1, T^2 \right) \right)^2} = - \left[\frac{\bar{g}_W^2 \left(\underline{y}^1 \mid T^1, T^2 \right)}{\bar{g}_W^1 \left(\underline{y}^1 \mid T^1, T^2 \right)} \mathrm{d}_{\tau^2} \frac{W^1 \left(T^1, T^2 \right)}{W^2 \left(T^1, T^2 \right)} + \frac{W^1 \left(T^1, T^2 \right)}{W^2 \left(T^1, T^2 \right)} \mathrm{d}_{\tau^2} \frac{\bar{g}_W^2 \left(\underline{y}^1 \mid T^1, T^2 \right)}{\bar{g}_W^1 \left(\underline{y}^1 \mid T^1, T^2 \right)} \right],$$

and finally,

$$\mathcal{G}_{SP,\tau^2} \left(y^1 \mid T^1, T^2 \right) = \lambda^1 \mathcal{G}_{W,\tau^2}^1 + \left(1 - \lambda^1 \right) \mathcal{G}_{W,\tau^2}^2 + \lambda_{\tau_2}^1 \left[\mathcal{G}_W^1 - \mathcal{G}_W^2 \right].$$

Starting from a symmetric equilibrium where Assumption 1 holds with $\alpha(\omega) = \bar{\alpha}$ and $E\left[s^1(\omega) \right] = E\left[s^2(\omega) \right] = \frac{1}{2}$, we get

$$\frac{\mathcal{G}_{W,\tau^2}^1 \left(y^1 \mid T^1, T^2 \right)}{\mathcal{G}_W^1 \left(y^1 \mid T^1, T^2 \right)} = \frac{\bar{\alpha}}{E\left[s^1(\omega) \right]} \left\{ \frac{E\left[s^1(\omega) \right]}{E\left[s^1(\omega) \mid y^1(\omega) \geq y^1 \right]} E\left[h^2 \left(y^{2\prime} \right) \mid y^{1\prime} \geq y^1 \right] - E\left[h^2 \left(y^{2\prime} \right) \right] \right\},$$

$$\frac{\mathcal{G}_{W,\tau^2}^2 \left(y^1 \mid T^1, T^2 \right)}{\mathcal{G}_W^2 \left(y^1 \mid T^1, T^2 \right)} = - \frac{3\bar{\alpha}}{E\left[s^2(\omega) \right]} \left\{ \frac{E\left[s^2(\omega) \right]}{E\left[s^2(\omega) \mid y^1(\omega) \geq y^1 \right]} E\left[h^2 \left(y^{2\prime} \right) \mid y^{1\prime} \geq y^1 \right] - E\left[h^2 \left(y^{2\prime} \right) \right] \right\},$$

$$\frac{\lambda_{\tau^2}^1 \left(T^1, T^2 \right)}{\lambda^1 \left(T^1, T^2 \right)} = - \bar{\alpha} \left(1 - \lambda^1 \right) \left\{ \left(\frac{1}{E\left[s^1(\omega) \right]} - \frac{3}{E\left[s^2(\omega) \right]} + \frac{2}{E\left[s^1(\omega)^2 \right]} \right) E\left[h^2 \left(y^2(\omega) \right) s^1(\omega) \right] \right.$$

$$\left. + \left(\frac{1}{E\left[s^1(\omega) \right]} - \frac{3}{E\left[s^2(\omega) \right]} + \frac{2}{E\left[s^2(\omega)^2 \right]} \right) E\left[h^2 \left(y^2(\omega) \right) s^2(\omega) \right] \right\}.$$

This proves in particular equations (20) and (21).

53

Proof of Proposition 7

Using Lemmas 1 and 2 and the first order conditions of the optimization problems in (24), we can write

$$\sigma^{1*}(y^1, y^2) = \frac{W^1(y^1, y^2)}{\kappa \, \bar{\Gamma}^{1u}(y^1, y^2)} \chi^*(y^1, y^2) \, \frac{\bar{\Gamma}^{1u}(y^1, y^2)}{\bar{\Gamma}^{1u}(y^1, y^2)} \, \rho\left(\chi^*(y^1, y^2) \frac{\bar{\Gamma}^{1u}(y^1, y^2)}{\bar{\Gamma}^{2u}(y^1, y^2)}\right) \tag{55}$$

$$\chi^*(y^1, y^2) = \frac{W^1(y^1, y^2)/\bar{\Gamma}^{1u}(y^1, y^2)}{W^2(y^1, y^2)/\bar{\Gamma}^{2u}(y^1, y^2)} \, . \tag{56}$$

Substituting (56) into (55) yields (7).

B   A model that includes ethical voters who always vote

We assume that the random variables $\tilde{q}^1$ and $\tilde{q}^2$ are driven by aggregate shocks that affect the shares of definite and rule-utilitarian voters one the one hand and of definite abstainers on the other so that the following two properties are satisfied: First, the ratio of definite and rule-utilitarian voters is not subject to randomness; i.e., shocks affect the ratio of potential voters to definite abstainers without affecting the internal composition of the set of potential voters. Second, among the supporters of party j, the ratio of definite to rule-utilitarian voters is the same for all types.

Assumption 4. *There is a pair of independent random variables, η_1 and η_2, so that, for all ω,*

$$\tilde{q}^{1v}(\omega) = \bar{q}^{1v}(\omega) \, \eta_1 \quad and \quad \tilde{q}^{1u}(\omega) = \bar{q}^{1u}(\omega) \, \eta_1$$

and

$$\tilde{q}^{2v}(\omega) = \bar{q}^{2v}(\omega) \, \eta_2 \quad and \quad \tilde{q}^{2u}(\omega) = \bar{q}^{2u}(\omega) \, \eta_2.$$

In addition, there are numbers q^{1v}, q^{1u}, q^{2v} and q^{2u} so that, for all ω,

$$\bar{q}^{1v}(\omega) = q^{1v} \quad and \quad \bar{q}^{1u}(\omega) = q^{1u}$$

and

$$\bar{q}^{2v}(\omega) = q^{2v} \quad and \quad \bar{q}^{2u}(\omega) = q^{2u} \, .$$

Under Assumption 4 the total number of votes for party 1 can be written as

$$\tilde{V}^1(p^1, p^2, \sigma^1, \tilde{q}^1) = \eta^1 \, V^1(p^1, p^2, \sigma^1)$$

where $V^1(p^1, p^2, \sigma^1) := m^1(\sigma^1) \, \Gamma^1(p^1, p^2)$ and $m^1(\sigma^1) := q^{1v} + \sigma^1 \, q^{1u}$ is a multiplier that determines how party 1's base $\Gamma^1(p^1, p^2)$ is transformed into actual votes. Analogously, the votes for party 2 are given by $\tilde{V}^2(p^1, p^2, \sigma^2, \tilde{q}^2) = \eta^2 \, V^2(p^1, p^2, \sigma^2)$, where $V^2(p^1, p^2, \sigma^2) := m^2(\sigma^2) \, \Gamma^2(p^1, p^2)$ and $m^2(\sigma^2) = q^{2v} + \sigma^2 \, q^{2u}$. Armed with this notation, we can express the probability that party 1 wins

as

$$\pi^1(p^1, p^2, \sigma^1, \sigma^2) = \mathcal{P}\left(\frac{V^1(p^1, p^2, \sigma^1)}{V^2(p^1, p^2, \sigma^2)}\right) = \mathcal{P}\left(\frac{m^1(\sigma^1)}{m^2(\sigma^2)}\frac{\Gamma^1(p^1, p^2)}{\Gamma^2(p^1, p^2)}\right) , \tag{57}$$

where $\mathcal{P}$ is the *cdf* of the random variable $\frac{\eta^2}{\eta^1}$. Its density function is denoted by ρ. Note that imposing Assumption 4 implies a multiplicative separability between the term

$$\chi(\sigma^1, \sigma^2) = \frac{m^1(\sigma^1)}{m^2(\sigma^2)} , \tag{58}$$

that is shaped by the rule-utilitarian voter's participation thresholds and the ratio of their bases

$$\gamma(p^1, p^2) = \frac{\Gamma^1(p^1, p^2)}{\Gamma^2(p^1, p^2)}$$

so that we can write

$$\pi^1(p^1, p^2, \sigma^1, \sigma^2) = \mathcal{P}\left(\chi(\sigma^1, \sigma^2)\,\gamma(p^1, p^2)\right) . \tag{59}$$

Turnout. For now, we take the party platforms p^1 and p^2 as given and characterize the parties' equilibrium turnout. We say that the turnout game has an interior equilibrium if $0 < \sigma^{1*}(p^1, p^2) < 1$ and $0 < \sigma^{2*}(p^1, p^2) < 1$. If the function $\mathcal{P}$ is continuously differentiable then an interior equilibrium is characterized by the first order conditions

$$\pi^1_{\sigma^1}(\cdot)\,W^1 - \kappa\,q^{1u}\,\Gamma^1 = 0 , \tag{60}$$

and

$$-\pi^1_{\sigma^2}(\cdot)\,W^2 - \kappa\,q^{2u}\,\Gamma^2 = 0 . \tag{61}$$

Using Assumption 4 we can rewrite these conditions as

$$\frac{\rho(\cdot)\chi(\sigma^1, \sigma^2)}{q^{1v} + \sigma^1\,q^{1u}}\,W^1 - \kappa\,\Gamma^1 = 0 , \tag{62}$$

and

$$\frac{\rho(\cdot)\chi(\sigma^1, \sigma^2)}{q^{2v} + \sigma^2\,q^{2u}}\,W^2 - \kappa\,\Gamma^2 = 0 . \tag{63}$$

Equations (62) and (63) imply that

$$\chi(\sigma^1, \sigma^2) = \frac{W^1/\kappa\,\Gamma^1}{W^2/\kappa\,\Gamma^2} = \frac{W^1/\Gamma^1}{W^2/\,\Gamma^2} , \tag{64}$$

which is the same expression as (3) in the body of the text.

Probability of winning. Suppose first that parties seek to maximize the probability of winning. Let

$$\chi^*(p^1, p^2) = \chi(\sigma_1^*(p^1, p^2), \sigma_2^*(p^1, p^2))$$

be the ratio of the parties support multipliers m^1 and m^2 that is induced by a pair of policies p^1 and p^2. Thus, with the solution concept of subgame perfect equilibrium, party 1 seeks to maximize

$$\mathcal{P}\left(\chi^*(p^1,p^2)\frac{\Gamma^1(p^1,p^2)}{\Gamma^2(p^1,p^2)}\right)$$

and party 2 seeks to minimize this expression. As $\mathcal{P}$ is a non-decreasing function we can as well assume that party 1 maximizes

$$\chi^*(p^1,p^2)\frac{\Gamma^1(p^1,p^2)}{\Gamma^2(p^1,p^2)}$$

or any monotone transformation of it such as e.g.

$$\ln\left(\chi^*(p^1,p^2)\right)+\ln\left(\Gamma^1(p^1,p^2)\right)-\ln\left(\Gamma^2(p^1,p^2)\right) . \tag{65}$$

Remark 1. The "conventional" probabilistic voting model can be viewed as a special case of this that is defined by two properties. First, since turnout is exogenous and universal, $\chi^*(p^1,p^2)=1$, for all (p^1,p^2) and hence $\ln\left(\chi^*(p^1,p^2)\right)=0$. Second, and again for the reason that turnout is exogenous and universal, $V^1=\Gamma^1(p^1,p^2)$ and $V^2=\Gamma^2(p^1,p^2)=1-\Gamma^1(p^1,p^2)$. In the probabilistic voting model, the objective of party 1 can therefore be taken to be $\ln\left(\Gamma^1(p^1,p^2)\right)-\ln\left(1-\Gamma^1(p^1,p^2)\right)$ or simply $V^1=\Gamma^1(p^1,p^2)$. I.e. maximizing the probability of winning is the same as maximizing the number of votes.

Moreover:

Remark 2. With Nash equilibrium as the solution concept, the parties view $\chi(\sigma^1,\sigma^2)$ as exogenously fixed, albeit at the level that is induced by the equilibrium policies. Party 1 then seeks to maximize

$$\ln\left(\Gamma^1(p^1,p^2)\right)-\ln\left(\Gamma^2(p^1,p^2)\right)$$

and party 2 seeks to minimize this expression. Since $\Gamma^2(p^1,p^2)=1-\Gamma^1(p^1,p^2)$, party 1's objective can as well simply taken to be $\Gamma^1(p^1,p^2)$ and $\Gamma^2(p^1,p^2)$ can be taken to be the objective of party 2. Nash equilibrium then requires that p^1 solves $\max_{\hat{p}^1\in P}\ \Gamma^1(\hat{p}^1,p^2)$ and that p^2 solves $\max_{\hat{p}^1\in P}\ \Gamma^2(p^1,\hat{p}^2)$. Note that these equilibrium are also the equilibrium conditions in the "conventional" probabilistic voting model. Thus, equilibrium existence in the "conventional" probabilistic voting model implies the existence of a Nash equilibrium in the given setup.

If the turnout subgame has an interior equilibrium, then the probability of winning for party 1 can be written in a reduced form that no longer involves an explicit reference to the participation thresholds σ^1 and σ^2. Specifically, equation (64) implies that the winning probability in (59) becomes

$$\bar{\pi}^1(p^1,p^2)=\mathcal{P}\left(\psi(p^1,p^2)\right) \quad\text{for}\quad \psi(p^1,p^2):=\frac{W^1(p^1,p^2)}{W^2(p^1,p^2)} . \tag{66}$$

Thus, as in the main body of the text (Proposition 2), under Assumption 4, if (p^1,p^2) is a pair of interior subgame perfect equilibrium policies, then it it is a saddle point of the function ψ.

C   Quadratic cost functions

An adaptation of the arguments in the Proof of Lemma 1 to the case of a quadratic cost function reveals that

$$\chi^*(p^1, p^2) := \frac{\sigma^{1*}(p^1, p^2)}{\sigma^{2*}(p^1, p^2)} = \left(\frac{W^1(p^1, p^2)/\bar{\Gamma}^{1u}(p^1, p^2)}{W^2(p^1, p^2)/\bar{\Gamma}^{2u}(p^1, p^2)} \right)^{\frac{1}{2}} . \tag{67}$$

To get to this equation one has to follow the same steps as in the proof of this lemma, while taking account of the fact that the first order conditions (26) and (27) have to be adjusted. With a quadratic cost function they are

$$\pi^1_{\sigma^1}(\cdot)\, W^1(p^1, p^2) - \kappa\, \sigma^1\, \bar{\Gamma}^{1u}(p^1, p^2) = 0 , \tag{68}$$

and

$$- \pi^1_{\sigma^2}(\cdot)\, W^2(p^1, p^2) - \kappa\, \sigma^2\, \bar{\Gamma}^{2u}(p^1, p^2) = 0 . \tag{69}$$

Following the same steps as in the proof of Lemma 2 reveals that party 1's probability of winning is, again, given by

$$\bar{\pi}^1(p^1, p^2) = \mathcal{P}\left(\chi^*(p^1, p^2) \frac{\bar{\Gamma}^{1u}(p^1, p^2)}{\bar{\Gamma}^{2u}(p^1, p^2)} \right) .$$

Upon plugging the expression for $\chi^*(p^1, p^2)$ in (67) into this expression we find that

$$\bar{\pi}^1(p^1, p^2) = \mathcal{P}\left(\left(\frac{W^1(p^1, p^2)\, \bar{\Gamma}^{1u}(p^1, p^2)}{W^2(p^1, p^2)\, \bar{\Gamma}^{2u}(p^1, p^2)} \right)^{\frac{1}{2}} \right)$$

Thus, the probability that party 1 wins the election is a monotonic function of $\frac{W^1(p^1, p^2)}{W^2(p^1, p^2)} \frac{\bar{\Gamma}^{1u}(p^1, p^2)}{\bar{\Gamma}^{2u}(p^1, p^2)}$. Consequently, with a quadratic cost function, an interior subgame perfect equilibrium is associated with a saddle point of the function

$$\psi^{qu}(p^1, p^2) = \frac{W^1(p^1, p^2)\, \bar{\Gamma}^{1u}(p^1, p^2)}{W^2(p^1, p^2)\, \bar{\Gamma}^{2u}(p^1, p^2)} .$$

Recall that, with a linear cost function, it is associated with a saddle point of

$$\psi(p^1, p^2) = \frac{W^1(p^1, p^2)}{W^2(p^1, p^2)} .$$

In both cases a Nash equilibrium is associated with a saddle point of

$$\bar{\Gamma}^{1u}(p^1, p^2)$$

or, equivalently, a saddle point of

$$M(\bar{\Gamma}^{1u}(p^1, p^2)) = \frac{\Gamma^{1u}(p^1, p^2)}{E[\bar{q}^u(\omega)] - \Gamma^{1u}(p^1, p^2)} = \frac{\Gamma^{1u}(p^1, p^2)}{\Gamma^{2u}(p^1, p^2)} .$$

Thus, both with a linear and a quadratic cost function, a difference between subgame perfect and Nash equilibrium policies arises because the function $\psi(p^1, p^2)$ is relevant for the former but not for

the latter. The relevant functional forms differ, however.

D   A characterization of admissible tax systems

We use a mechanism design approach to characterize the set of admissible tax systems. By the taxation principle, see e.g. Hammond (1979) or Guesnerie (1995), an allocation (c, y) consisting of a consumption schedule $c : \Omega \to \mathbb{R}_+$ and an earnings schedule $y : \Omega \to \mathbb{R}_+$ can be induced by an income tax if and only if it satisfies the resource constraint,

$$E[y(\omega)] \geq E[c(\omega)] \tag{70}$$

and incentive compatibility constraints: For all ω and ω',

$$u(\omega) \geq c(\omega') - v(y(\omega'), \omega) . \tag{71}$$

where

$$u(\omega) := c(\omega) - v(y(\omega), \omega) . \tag{72}$$

gives the utility that a type ω individual realizes under allocation (c, y).

It is also well known how to obtain a characterization of incentive compatible allocations in models with quasilinear preferences, see e.g. Myerson (1981). An application of the envelope theorem makes it possible to show that incentive compatibility holds if and only if two conditions are satisfied: First, for all ω,

$$u(\omega) = \underline{u} - \int_{\underline{\omega}}^{\omega} v_2(y(z), z) \, dz , \tag{73}$$

where $\underline{u} := u(\underline{\omega})$ is a shorthand for the lowest type's utility and $-\int_{\underline{\omega}}^{\omega} v_2(y(z), z) \, dz$ is the information rent that higher types realize in the presence of incentive compatibility constraints. Second, y is a non-decreasing function, i.e. individuals with higher productive abilities must not earn less than individuals with lower productive abilities.

We can use these insights to derive a representation of preferences over tax polices in a reduced form that only depends on the income function y and no longer involves a reference to the consumption function c. This will enable us to represent a party's problem of choosing a tax policy as an optimization problems that no longer involve resource and incentive constraints.

Suppose that (c, y) is incentive compatible, then using (72), (73) and an integration by parts we obtain

$$E[c(\omega)] = \underline{u} + E\left[v(y(\omega), \omega) - \frac{1 - F(\omega)}{f(\omega)} \, v_2(y(\omega), \omega)\right] .$$

Plugging this expression into the public sector budget constraint $E[y(\omega)] - E[c(\omega)] = 0$ yields an expression for $\underline{u}$; it is equal to the virtual surplus:

$$\underline{u} := s_v(y) := E\left[y(\omega) - v(y(\omega), \omega) + \frac{1 - F(\omega)}{f(\omega)} \, v_2(y(\omega), \omega)\right] . \tag{74}$$

The virtual surplus is a surplus measure that takes account of the information rents that tax-payers realize. To arrive at the virtual surplus, the surplus of aggregate output over costs of effort

$$s(y) := E\left[y(\omega) - v(y(\omega), \omega)\right]$$

is reduced by the aggregate information rent

$$-E\left[\int_{\underline{\omega}}^{\omega} v_2(y(z), z) \, dz\right] = -E\left[\frac{1 - F(\omega)}{f(\omega)} \, v_2(y(\omega), \omega)\right] \ .$$

Indirect utility induced by an incentive compatible allocation can now be written as a sum of virtual surplus and information rents

$$u(\omega) := s_v(y) - \int_{\underline{\omega}}^{\omega} v_2(y(z), z) \, dz \ . \tag{75}$$

With this characterization, the utility realized by a type ω individual depends on the whole earnings schedule $y : \Omega \to \mathbb{R}_+$ but no longer on the consumption schedule $c : \Omega \to \mathbb{R}_+$.

When we study political competition over income tax schedules, we will denote party 1's proposal by $y^1 : \Omega \to \mathbb{R}_+$ and party 2's proposal by $y^2 : \Omega \to \mathbb{R}_+$ and we denote by $u^1(\omega)$ and $u^2(\omega)$ the associated utility levels for a type ω individual.

We are particularly interested in the marginal tax rates that are associated with the tax systems that the parties propose. To get from an incentive compatible allocation to the associated tax schedule T we use the first order condition of the utility-maximization problem that individuals face in the presence of this tax system. If tax system T induces an incentive compatible allocation (c, y), then

$$1 - T'(y(\omega)) = v_1(y(\omega), \omega) \ .$$

Hence, $1 - v_1(y(\omega), \omega)$ is interpreted as the marginal tax rates that type ω agents face.

E   Public goods

Our framework for studying endogenous turnout and endogenous platforms in political competition is developed for a generic policy domain. We have emphasized that the set of non-linear income tax systems is a policy domain of particular interest. That said, our framework can also be applied to study the implications of endogenous turnout for political competition over other policy domains. In this section, we briefly summarize the results from such an analysis. Specifically, we report on the implications of our framework for public goods provision.

Individuals have preferences over public goods that are given by $u(\omega, p) = \omega \, p - k(p)$, where $p \in \mathbb{R}_+$ denotes the quantity of the public good, $\omega \in \Omega$ is an individual's public goods preference and the cost function k captures the per capita cost of public goods provision.[14] We begin with a

---

[14]In an economy with a continuum of individuals and private information on public goods preferences, equal cost sharing is the only way of satisfying robust incentive compatibility, budget balance

characterization of the public good provision level that party 1 would choose if its sole objective was to mobilize its supporters. In this case, it would choose q^1 with the objective to maximize

$$W^1(p^1, p^2) = \mathbb{E}[G_W^1(\omega \, p^1 - k(p^1) - u(p^2, \omega) \mid \omega)] \, .$$

Given p^2, the first order condition characterizing the optimal choice of p^1 is

$$\mathbb{E}\left[\mathcal{G}_W^1(\omega \mid p^1, p^2) \, \omega\right] = k'(p^1) \quad \text{where} \quad \mathcal{G}_W^1(\omega \mid p^1, p^2) = \frac{g_W^1(u(p^1, \omega) - u(p^2, \omega) \mid \omega)}{\mathbb{E}[g_W^1(u(p^1, \omega) - u(p^2, \omega) \mid \omega)]} \, .$$

This first order condition is a political economy analogue to the Samuelson rule for first-best public good provision. For the given setup, the Samuelson rule stipulates that $\mathbb{E}[\omega] = k'(p)$, i.e., it requires equal weights for all public goods preferences. For the purpose of mobilizing its supporters, party 1 does not apply equal weights. Instead the public good preferences of different individuals are weighted according to the function $\mathcal{G}_W^1$. The public good provision level that party 1 would choose if it only wanted only to demobilize the supporters of party 2 is such that

$$\mathbb{E}\left[\mathcal{G}_W^2(\omega \mid p^1, p^2) \, \omega\right] = k'(p^1) \, ,$$

and the policy that maximizes $\frac{W^1(p^1 p^2)}{W^2(p^1, p^2)}$ satisfies

$$\mathbb{E}[\mathcal{G}_{SP}(\omega \mid p^1, p^2) \, \omega] = k'(p^1)$$

where

$$\mathcal{G}_{SP}(\omega \mid p^1, p^2) := \lambda^1(p^1, p^2) \, \mathcal{G}_W^1(\omega \mid p^1, p^2) + (1 - \lambda^1(p^1, p^2)) \, \mathcal{G}_W^2(\omega \mid p^1, p^2) \, .$$

Again, the party compromises between mobilizing its own supporters and demobilizing the supporters of the other party – with the weight on the own supporters being smaller if the party is more likely to win.

Strategic substitutes and complements: Public goods vs non-linear taxation

For public good provision, it is particularly easy to determine whether the provision levels of different parties are substitutes or complements. Suppose that

$$\frac{\partial}{\partial p^2} \mathbb{E}[\mathcal{G}_{SP}(\omega \mid p^1, p^2) \, \omega] > 0 \, ,$$

i.e., that an increase of party 2's provision level leads to an increase of the $\mathcal{G}_{SP}$-weighted average of public goods preferences. Then the best response of party 1 is to also increase its provision level,

---

and anonymity, see Bierbrauer and Hellwig (2016).

i.e., the parties choices are then strategic complements. If, by contrast,

$$\frac{\partial}{\partial p^2} \mathbb{E}[\mathcal{G}_{SP}(\omega \mid p^1, p^2) \, \omega] < 0 \ ,$$

they are strategic substitutes.

For non-linear income taxation, this analysis is more complicated as it has to be applied separately for each level of income. There may be a complementarity in marginal tax rates for high incomes and a substitutability in marginal tax rates for low incomes, etc. A public good is jointly consumed by all individuals. One cannot have a high public good provision level for some people and a low public good provision level for others. This simplifies the analysis of the parties' strategic interdependence.

For non-linear income taxation, the formal expression that allows us to analyze how strategic substitutabilities and complementarities vary with incomes or productive abilities is the Gateaux differential of the weighting functions $\mathcal{G}$. This Gateaux differential is a function, it gives a different value for each type ω. For public good provision, we only have to look at how a population average of public goods preferences changes with the parties proposals. The formal analysis therefore does not require functional derivatives but only standard calculus.