Redistributive Politics under Ambiguity

Donna, Javier

University of Florida

2023

Redistributive Politics under Ambiguity

Javier D. Donna*

November 22, 2023.

Abstract

The conflicting views that agents and voters have about redistributive taxation have been broadly studied. The literature has focused on situations where the counterfactual outcomes that would have occurred had other actions been chosen are observable or point identified. I analyze this problem in a context of ambiguity. The extent to which individuals are responsible for their own fate is partially identified. Agents have partial knowledge of the relative importance of effort in the generation of income inequality and, therefore, the magnitude of the incentive costs. I present a simple model of redistribution and show that multiple equilibria might arise *even* in the presence of ambiguity: One where the rate of redistribution is high, agents are pessimistic, and exert low effort (*Pessimism/Welfare State*), and another where the redistribution tax rate is low, agents are optimistic, and exert high effort (*Optimism/Laissez Faire*).

JEL Codes: D80, H10, H30, P16, E62.

Keywords: Redistributive Politics, Taxes, Ambiguity, Beliefs, Effort, Luck, Multiple Equilibria.

*University of Florida, Department of Economics, 224 Matherly Hall, P.O. Box 117140, Gainesville, FL 32611-7140. Phone: 352.392.0117. Fax: 352.392.7860. Email: jdonna@ufl.edu. I especially thank Charles F. Manski and Dale T. Mortensen for many helpful suggestions. I thank Ann Atwater and Yuchen Zhu for excellent research assistance. I am particularly grateful to the managing editor, Francois Maniquet, the associate editor, Avidit Acharya, and two anonymous referees for insightful comments and suggestions. This paper combines two previously unpublished working papers, "Redistributive Politics in a Context of Ambiguity" and "Optimal Labor Decisions under Partial Knowledge of Effort."

1 Introduction

The relative importance of *effort* and *predetermined factors* (luck) in the production and distribution of income has been extensively studied.[1] There is substantial heterogeneity across countries about individuals' beliefs regarding the causes of poverty, wealth, the rewards of effort, and whether individuals are responsible for their own fate. American exceptionalism and the belief in the American Dream are examples of this phenomenon.

Economists have long argued that markets and capitalist forms of organization outperform socialism. In practice, however, certain capitalist ideas appear to not perform well with the general public in certain world regions.[2] Socialist ideas are more prevalent in the names and platforms of parties in low-income countries. In Latin America, this phenomenon is particularly severe. There has been a political backlash against markets in almost all of the countries in the region after decades of privatizations and deregulation.[3] A natural question that arises is why this polarization has happened.

One explanation that is appealing to economists is that market forces make the income distribution more unequal, leading the median voter to demand higher income taxes, as in the classic models by Romer (1975) and Meltzer and Richard (1981). However, this explanation fails to explain some basic patterns across countries. For example, in the United States, the pretax income distribution is more unequal than in Europe. Yet the United States government redistributes income among their citizens much less than European governments.

Alesina, Glaeser, and Sacerdote (2001, henceforth AGS) emphasize the role of beliefs in explaining these differences. They document that 54 percent of Europeans, yet only 30 percent of Americans, believe that the poor are unlucky, whereas 60 percent of Americans, yet only 26 percent of Europeans, believe the poor are lazy. AGS show that these beliefs are strong predictors of government intervention across countries. Figure 1 shows a strong correlation between social spending as a percent of the GDP and the belief that luck determines income.

Investigating what causes differences in beliefs and the level of effort exerted is paramount. A natural hypothesis is that they might simply reflect individuals' different experiences. In a seminal paper, Piketty (1995) shows that two identical economies evolve along different paths due to differences in the initial shocks. Authors have also analyzed the role of corruption and fairness (Alesina and Angeletos 2005a), the role of beliefs on social mobility (Alesina and Angeletos 2005b), and the incentives for people to engage in belief manipulation (Benabou

[1]See, *e.g.*, Romer (1975), Meltzer and Richard (1981), Piketty (1995), Alesina, Glaeser, and Sacerdote (2001), Alesina and Angeletos (2005a), Alesina and Angeletos (2005b), Benabou and Tirole (2006), and the references therein.

[2]See, *e.g.*, Di Tella and MacCulloch (2009) and the references therein.

[3]See Stokes (2001) and Lora and Olivera (2004).

and Tirole 2006). In these papers, the agents do not face a selection problem: The counterfactual outcomes that would have occurred had other actions been chosen are assumed to be observable. Evidence on the structure of beliefs can be found in, *e.g.*, Hirschman and Rothschild (1973), Hochschild (1981), Ladd and Bowman (1998), Benabou and Ok (2001), Corneo and Grüner (2002), Fong (2001, 2004), Di Tella, Galiani, and Schargrodsky (2007), Inglehart (2018), and the references therein.[4]

I present a simple model of beliefs and redistribution, and show that multiple equilibria along the lines of Benabou and Tirole (2006), might arise *even* in the presence of ambiguity. I use the term ambiguity (or Knightian uncertainty) to refer to a situation where the agent knows the choice set and wants to maximize an unknown objective function, as in Manski (2009).[5]

There is no ambiguity in the models mentioned above. In Piketty (1995), Bayesian voters might have conflicting views about redistributive taxation because rational agents estimate their incentive costs differently due to differences in initial shocks through a Bayesian learning process. In Benabou and Tirole (2006), agents engage in belief manipulation. They model it through a form of *imperfect willpower* that causes the effort choice to be too low compared to the desirable level. The authors assume that agents have prior subjective (alternatively, known) distributions over the objects to be learned in these models. I study the problem faced by the agents when agents have partial knowledge of the relative importance of effort (as opposed to predetermined factors or luck) in the generation of income inequality and, therefore, the magnitude of these incentive costs. The source of ambiguity in my model is the extent to which effort increases output; that is, whether effort pays off as motivated by Figure 1. The treatment is the effort that the agents choose.

I present the model in Section 2. Agents' objective is to maximize individual welfare by choosing the level of effort they exert. Exerting effort is costly for individuals. On the one hand, individuals know their effort cost function; their cost function is point identified for each level of effort. The true return to effort, on the other hand, is unknown and ambiguous.

Decision theorists have proposed various criteria to choose among ambiguous undominated treatment allocations. The Bayes Rule criterion (Savage 1954) states that the agent maximizes the expected utility using a (unique) prior probability and a utility function. The agent is assumed to have a probability distribution over all possible distributions. Several generalizations have been proposed to deal with the consistency conditions, such as the Ellsberg's paradox (Ellsberg 1961), and choosing the priors. Two approaches are using non-additive probabilities and multiple priors. Schmeidler (1989) introduces an extension to situations with non-additive expected utility, using the Choquet integral (Choquet 1954).

[4]See AGS and Alesina and Glaeser (2004) for comprehensive reviews of the literature.

[5]These ideas date back to Keynes (1921) and Knight (1921). The term ambiguity goes back to Ellsberg (1961).

His extension can be used in a social welfare setting to represent an expected utility of a concave Bergson-Samuelson social welfare function using the notion of uncertainty aversion and the interpretation of comonotonic independence, which is consistent with strict uncertainty aversion. Gilboa and Schmeidler (1989) extend the Bayesian criterion using multiple priors (instead of a unique one). In their criterion, the set of priors is the set of possible probability distributions in the statistical decision problem, thus providing axiomatic foundation for Wald's minimax loss criterion (Wald 1950, section 1.4.2). Ghirardato, Maccheroni, and Marinacci (2004) study a generalization of Gilboa and Schmeidler (1989) called the α-maxmin, distinguishing between preferences toward ambiguity and beliefs about the level of ambiguity, and where α corresponds to the degree of the ambiguity aversion. Bossaerts, Ghirardato, Guarnaschelli, and Zame (2010) use it to study ambiguity aversion on equilibrium asset prices. Gilboa and Schmeidler (1993) study the intersection of both types of approaches (non-additive probabilities and multiple priors) to dynamically update ambiguous beliefs. Gul and Pesendorfer (2018) develop a criterion to update beliefs under ambiguity that is recursive (backward induction evaluation of random variables) and consequentialist (the conditional expectation only depends on the values of the random variable on the conditioning event) and where the agent is dynamically consistent. Hurwicz (1951) suggests a criterion that is a weighted average of the minimum and maximum values of the objective function. Thus, it provides a solution where the degree of pessimism encompass other criteria in the corner cases of extreme pessimism (maximin criteria) or extreme optimism (maximax criteria). The minimax regret criterion consists of minimizing the maximum regret of an effort allocation. In Section 3, I present the main specification using an extension of the Hurwicz criterion (discussed next). In Section 4, I discuss extensions using other criteria.

I introduce a cognitive extension to the Hurwicz criterion for choosing among undominated alternatives. This extension allows one to endogenize the degree of optimism/pessimism. The solution has an intuitive interpretation in terms of the endogenous optimism/pessimism regarding whether exerting effort pays off to the individual (relative to luck), as motivated by Figure 1. The Hurwicz criterion also allows to obtain tractable closed-form solutions. I present extensions using other criteria, like the Bayes Rule and the minimax regret, in Section 4, and Appendices D and E.

In Section 3, I examine the behavior of agents in the redistribution game and analyze the qualitative properties of the outcomes and equilibria. To achieve this task, I follow a series of lemmas that describe the solutions and qualitative properties in each stage. I finish with a proposition showing the possibility of multiple equilibria. The intuition is simple. If the pivotal group that determines the equilibrium tax rate is pessimistic, agents optimally set up a high redistribution tax rate which, in turn, makes the non-pivotal group to be pessimistic and exerts a low level of effort. The opposite occurs if the pivotal group is optimistic: The equilibrium tax rate is low, the non-pivotal group is optimistic, and the exerted level of effort is high.

Section 5 concludes and discusses some extensions. The general idea that higher taxation rates would induce individuals to be more pessimistic remains applicable. Proofs and extensions are in the appendix.

2 A Model of Redistribution under Ambiguity

2.1 Setup

The main objective is to understand how multiple equilibria and economies with different degrees of redistribution may arise under ambiguity. The economy is populated by a continuum of agents, $j \in [0, 1]$. Each agent must choose the level of effort (or human capital), e_j, from a choice set, C, and has a response function (or output technology) that maps effort decisions into output, which takes values in space Y:

$$y_j(e_j) : C \to Y.$$

Following Manski (2009), let $z_j \in C$ be the level of effort chosen by agent j. Then person j obtains outcome $y_j := y_j(z_j)$. The outcomes of effort are the source of ambiguity. The counterfactual outcomes, $y_j(c)$ with $c \neq z_j$, that would have occurred had other level of effort been chosen are unobservable.

I consider the case where each agents can obtain one of two pretax incomes, $y \in \{y_0, y_1\}$ with $y_1 > y_0$, by choosing between two treatments of effort, $e_j \in \{e_L = 0, e_H = 1\}$. The agents' objective is to allocate effort treatments to maximize their expected welfare. It is natural to assume that the expected outcome if the agents decide not to exert effort, $e_L = 0$, is known for the individuals. If the agents decide not to work, they know the amount of output that they will obtain and, hence, $\alpha := \mathbb{E}[y_j(e_L)]$ is point identified. But if the agents decide to exert effort, there is ambiguity about the real extent to which individual achievement is responsive to individual effort. Thus, the true expected output that the agents obtain if they decide to exert effort, $e_H = 1$, is unknown (or ambiguous), and is contained in the interval $\beta := \mathbb{E}[y_j(e_H)] \in [\beta_L, \beta_U]$, with $\beta > \alpha$.[6]

[6]Using the law of iterated expectations and $y \in \{y_0, y_1\}$:

$$\begin{aligned}
\beta &:= \mathbb{E}[y_j(e_H)], \\
&= \mathbb{E}[y_j(e_j)|e_j = e_H]\mathbb{P}(e_j = e_H) + \mathbb{E}[y_j(e_j)|e_j \neq e_H]\mathbb{P}(e_j \neq e_H), \\
&\in [\underbrace{\mathbb{E}[y_j(e_j)|e_j = e_H]\mathbb{P}(e_j = e_H) + y_0\mathbb{P}(e_j \neq e_H)}_{\beta_L}, \underbrace{y_j(e_j)|e_j = e_H]\mathbb{P}(e_j = e_H) + y_1\mathbb{P}(e_j \neq e_H)}_{\beta_U}].
\end{aligned}$$

Let $\delta \in [0,1]$ be a treatment allocation that assigns a fraction δ of the effort endowment to treatment e_H. One can interpret δ as the fraction of effort exerted by the agents. Exerting effort is costly for the agents. The agents know how costly is effort for them, so that the convex cost function, $C_j(\cdot)$, is point identified. Let $U(\delta)$ denote the agents' welfare attained by assigning a fraction δ of the effort endowment to treatment e_H and $1-\delta$ to e_L. Then:[7]

$$\begin{aligned} U_j(\delta) &= \mathbb{E}[(1-\tau)y(e_j)] - C_j(\delta), \\ &= (1-\tau)[\alpha + \delta(\beta - \alpha)] - C_j(\delta), \end{aligned}$$

where $\tau \leq 1$ is the income tax rate that faces the agent (see Subsection 2.2 for the model with redistribution). The first term in the last equation reflects the expected disposable income of allocating a fraction δ of effort endowment to e_H. The last term captures the idea that allocating an increasing fraction of effort is costly for the agent.

The agents would like to solve the following optimization problem:

$$\max_{\delta \in [0,1]} U_j(\delta). \tag{1}$$

But agents do not know the true return to effort. They only know the choice set, not the objective function. Thus, the agents cannot compute the expected return to effort because they do not know anything about the counterfactual outcomes. They only know that the response function $y(\cdot) \in Y$, where Y is some set of possible objective functions. In particular, I assumed that α is point identified, but β is only partially identified. Therefore, the agents face a problem of choice under ambiguity.

How should they choose the amount of effort δ? There is not a correct answer to this question; that is, there is no correct answer to the question of how to choose the criterion to solve the problem in (1). However, agents should not choose a *dominated* level of effort. I say that a level of effort $\tilde{\delta}$ is dominated if there exist another feasible level of effort $\hat{\delta}$ that is at least as good as $\tilde{\delta}$ for all $\beta \in [\beta_L, \beta_U]$ and strictly better for some $\beta \in [\beta_L, \beta_U]$. In order for the agents to face an interesting problem of choice under ambiguity, I assume that if the true return to effort is the lower bound, β_L, then no matter how much effort the agents exert, they would be better off by not exerting effort at all and, hence, they would choose e_L. If the true return to effort is the upper bound, β_U, then all treatment allocations leave the agents strictly better off than not exerting effort. The following assumption summarizes these insights.

Assumption 1. *Let $\tau \in (0,1)$ and $(1-\tau)\beta_L - C_j(\delta) < (1-\tau)\alpha - C_j(0) < (1-\tau)\beta_U - C_j(\delta), \quad \forall \delta \in [0,1], \ \forall \tau.$*

[7] The specification of the production function and agents' utility are similar to Piketty (1995) and Benabou and Tirole (2006). The main difference is that I consider the ambiguity context described above.

Given the previous assumption, the agents face a genuine problem of choice under ambiguity.

For the main specification of the model, I use the Hurwicz criterion defined below. The Hurwicz criterion collapses to maximin (maximax) criterion when the individual is extremely pessimistic (optimistic). I therefore begin presenting the maximin and maximax criteria to provide intuition regarding the Hurwicz criterion. In Section 4, I discuss extensions using other criteria.

Maximin Criterion. According to this criterion, the agent chooses the level of effort that maximizes the minimum welfare attainable. Thus, the agent acts as if $\beta = \beta_L$ and solves the following optimization problem:

$$\max_{\delta \in [0,1]} (1-\tau) \Big[\alpha + \delta[\beta_L - \alpha] \Big] - C_j(\delta).$$

The maximin solution is $\delta^*_{MMin} = 0$. The intuition is simple. Because the maximin criterion considers the worse-case scenario to solve the ambiguity problem and because by Assumption 1 effort is not rewarded but costly, the maximin agent chooses not to exert effort. It can be interpreted as a very pessimistic criterion.

Maximax Criterion. Under this criterion, the agent chooses the level of effort that maximizes the maximum welfare attainable by setting $\beta = \beta_U$. Thus, the optimization problem is:

$$\max_{\delta \in [0,1]} (1-\tau) \Big[\alpha + \delta[\beta_U - \alpha] \Big] - C_j(\delta).$$

The maximax solution is $\delta^*_{MMax} = 1$. On the opposite extreme to the maximin, one finds this maximax agent acting extremely optimistic. Again, by Assumption 1 effort is rewarded, so the maximax agent chooses to exert the highest level of effort by setting $\delta_{MMax} = 1$.

Hurwicz Criterion. Hurwicz (1951) suggests a criterion that is a weighted average of the minimum and maximum values of the objective function. Thus, the individual faces the following problem:

$$\begin{aligned} U_j^H(\delta_H^*; \lambda_j) = \max_{\delta \in [0,1]} \lambda_j & \Big\{ \inf_{\beta \in [\beta_L, \beta_U]} \Big(\mathbb{E}[(1-\tau) y(e_j)] - C_j(\delta) \Big) \Big\} \\ + (1-\lambda_j) & \Big\{ \sup_{\beta \in [\beta_L, \beta_U]} \Big(\mathbb{E}[(1-\tau) y(e_j)] - C_j(\delta) \Big) \Big\}, \\ = \max_{\delta \in [0,1]} (1-\tau) & \Big\{ \alpha + \Big[\Big(\lambda_j \beta_L + (1-\lambda_j)\beta_U \Big) - \alpha \Big] \delta \Big\} - C_j(\delta), \end{aligned} \tag{2}$$

where $\lambda_j \in [0, 1]$.

One can interpret λ in this criterion as the degree of pessimism of the individual. In the corner cases where the individual is either extremely pessimistic ($\lambda = 1$) or extremely optimistic ($\lambda = 0$) the Hurwicz criterion collapses, respectively, to the maximin or maximax criteria described above.

2.2 Technology, Preferences, and Timing

Consider a redistribution game using the setup described above. Actions take place according to the timeline in Figure 2.

Each individual has a production technology like the one described in the previous section. For choosing among undominated alternatives, agents use a cognitive version of the Hurwicz criterion as discussed next. Income is taxed at a rate τ determined by majority voting, and tax revenue is redistributed lump-sum. Pretax income is given by $y_j(\cdot)$. Disposable income is $(1 - \tau)y_j(\cdot) + \tau\bar{Y}$, where $\bar{Y}$ is the average level of output.

To endogenize the degree of pessimism, λ, that arises in equilibrium, I use a modified version of the Hurwicz criterion and assume that individuals can manipulate (consciously or not) their degree of pessimism, λ, through a cognitive technology, which might affect their cost of effort. From equation (2), it is straightforward to see that the level of welfare attained for a given choice of effort is a decreasing function of the degree of pessimism, λ. The lower the pessimism, λ, the more weight the agent puts in the upper bound, β_H, and, hence, the higher the *ex ante* utility.

Nevertheless, the true return to effort β is unknown. Being too optimistic may make the individual choose *ex ante* a level of effort (and pay the associate cost) that is too high *ex post* compared to the level of effort that the agent had been chosen if the counterfactual outcomes where actually observable. This difference can be conceptualized as a form of regret: If the agents are too optimistic, they may regret in the future if they choose a level of effort that is too high; if they are too pessimistic, they may regret if they choose a level of effort that is too low. One can take this effect into account by allowing the agents to manipulate their degree of pessimism, λ.

I assume that there is some technology function, $M(\lambda)$, the cognitive technology through which the individuals can manipulate their degree of pessimism. Benabou and Tirole (2006) also use a cognitive technology that agents use to manipulate their own beliefs. Because they are not working in a context of ambiguity, they let the individual form beliefs regarding the true return to effort and use the cognitive technology to manipulate these beliefs. Specifically, they assume that the agents receive a binary signal about the return to effort (whether effort pays off or not), and they have a cognitive technology through which they can manipulate their own beliefs in the case they get the bad news that effort does not pay. They call *pessimism* the probability that this bad-news signal will be recalled. They assume that agents can increase or decrease this probability at some cost $M(\lambda)$. The Hurwicz criterion described above captures this idea by averaging the worst- and best-case scenarios (maximin and maximax, respectively). However, treating λ as an exogenously predetermined parameter is silent regarding how λ should be selected.[8] Assumption 3 introduces the cognitive technology to the Hurwicz criteria and assumes, as discussed above, that the cognitive cost is increasing in λ.

Assumption 2. *Let $M(\lambda) > 0$ and $M'(\lambda) > 0$ be the cognitive technology through which the agents can manipulate their own degree of pessimism.*

The composition of the electorate is straightforward. There are two types of individuals. A minority, $\phi < \frac{1}{2}$, of advantaged agents that have high effort cost parameter, $a_H(\lambda)$, and a majority, $1 - \phi$, of disadvantaged agents that have a lower effort cost parameter, $a_L(\lambda)$. Advantaged individuals have a lower cost, $C_j(\cdot)$, than the disadvantaged; that is, $a_H(\lambda) > a_L(\lambda)$ for a given λ. Optimistic individuals have lower costs than the pessimistic. In Section 4, I present an extension using a continuum of citizen types.

Assumption 3 summarizes this discussion and adds tractable functional forms similar to the cost function in Piketty (1995) and Benabou and Tirole (2006) (specifically, quadratic costs as a function of the effort level and an additively separable function). This specification allows to obtain closed-form solutions for the optimization problems.

Alternatively, advantaged agents can be interpreted as skilled (or rich people) who have a lower cost of exerting a given amount of effort than unskilled (or poor agents). For example, one can think of social origins as determining how costly it is for the agent to exert a given amount of effort and interpret skilled agents as coming from high-income families with access to better opportunities on average. Cunha and Heckman (2010) and Heckman (2006) present evidence showing how parental investments and environments affect the evolution of cognitive and noncognitive traits.

Assumption 3. $C_j(\delta) = \frac{1}{2}\frac{\delta^2}{a_j + a(\lambda)}$, $a(\lambda) > a_L, \forall \lambda \in [0, 1]$, $1 < a_L < a_H$, $\frac{\partial a(\lambda)}{\partial \lambda} := a'(\lambda) < 0$.

The *effort-cost parameter*, $a_j(\lambda) := a_j + a(\lambda)$, measures how costly is for the agent j to exert effort δ. Higher values reflect lower cost for the individual for a given choice of δ. Optimistic individuals have lower costs than the pessimistic, as discussed above.

[8] An alternative approach would be to solve the model using the Hurwicz criterion without introducing the cognitive technology and treating λ as an exogenous parameter that is drawn from some distribution. The conclusions from the following section still hold, but the degree of pessimism, λ, is exogenous.

The solution of the problem in (2) is:

$$\delta_H^* = \begin{cases} 0 & \text{if } \delta_H \leq 0, \\ 1 & \text{if } \delta_H \geq 1, \\ \delta_H & \text{if } \delta_H \in (0,1). \end{cases} \qquad \text{where } \delta_H = a_j(\lambda)(1-\tau)[\bar{\beta}_\lambda - \alpha], \tag{3}$$

In the previous solution, $\bar{\beta}_\lambda := \lambda_j \beta_L + (1-\lambda_j)\beta_U$ is a weighted average of the lowest and highest possible values of the unknown true return to effort, β. For δ_H^* to be fractional, we need, $\lambda_j \in (\underline{\lambda}, \bar{\lambda})$, where $\underline{\lambda} := \max\left\{0, \lambda : a_j(1-\tau)[\bar{\beta}_\lambda - \alpha] < 1\right\}$ and $\bar{\lambda} = \frac{\beta_U - \alpha}{\beta_U - \beta_L}$.

Because no structure has been assumed regarding the distribution of β, we need to say how agents estimate aggregate output, $\bar{Y}$. Assume that it has the same structure as the individual output in equation (2):

$$\begin{aligned} \bar{Y} &:= \mathbb{E}_j[y(e_j)], \\ &= \alpha + \bar{a}(\lambda)(1-\tau)^2 G, \end{aligned} \tag{4}$$

where $\bar{a}(\lambda) := \int_0^1 a_j(\lambda_j)\,\mathrm{d}j$, λ_j is the optimal choice of pessimism from agent j, and G is defined in Appendix B.

From Figure 2, the timing of actions is as follows. At $t = 0$, agents choose the optimal degree of pessimism, λ, using the cognitive technology just described.[9] At the start of period 1, individuals vote over the linear tax rate, τ, that determines how the output will be redistributed in period 2. After voting, the agents choose the optimal level of effort, δ. Finally, in period 2, the outcome, $y(e_j)$, and the redistribution are realized, and agents consume their disposable income. Thus, the agents' welfare is given by, U_j^{CH}, where CH stands for the cognitive Hurwicz criterion:

$$\begin{aligned} U_j^{CH}(\lambda, \delta, \tau) &= \lambda \left\{ \inf_{\beta \in [\beta_L, \beta_U]} \Big(\mathbb{E}[(1-\tau)y(e_j)] - C_j(\delta) \Big) \right\} \\ &\quad + (1-\lambda) \left\{ \sup_{\beta \in [\beta_L, \beta_U]} \Big(\mathbb{E}[(1-\tau)y(e_j)] - C_j(\delta) \Big) \right\} + \tau \bar{Y} + M(\lambda), \\ &= (1-\tau) \left\{ \alpha + \left[\Big(\lambda \beta_L + (1-\lambda)\beta_U \Big) - \alpha \right] \delta \right\} - C_j(\delta) + \tau \bar{Y} + M(\lambda), \\ &= (1-\tau)\left[\alpha + \left(\bar{\beta}_\lambda - \alpha\right)\delta\right] - C_j(\delta) + \tau \bar{Y} + M(\lambda). \end{aligned} \tag{5}$$

2.3 Equilibrium

Definition 1. *A cognitive-politico-economic equilibrium is a triple (λ, τ, δ) such that the following hold:*

[9] I call this period 0 to emphasize that this decision may or may not be conscious.

(i) ***Cognitive optimality:*** *λ is the optimal cognitive decision.*

(ii) ***Majority Tax:*** *τ is the majority tax rate, given the optimal cognitive decision.*

(iii) ***Effort optimality:*** *δ is determined according to the optimal decision of effort.*

3 Solving the Model

I solve the model by backward induction. Proofs are in Appendix B.

3.1 Effort Decisions

Knowing the degree of pessimism that they face, λ, and the tax rate, τ, individuals choose effort optimally using the Hurwicz criterion. Hence, the optimization problem is simplified to equation (2) and the optimal level of effort is given by equation (3).

If the agent is not extremely pessimistic ($\bar{\beta}_\lambda - \alpha > 0$) nor extremely optimistic $\big((\bar{\beta}_\lambda - \alpha)a_j(\lambda) < \frac{1}{(1-\tau)}\big)$, the optimal solution is fractional and $\delta_H^* \in (0,1)$. The optimal level of effort, δ_H^*, decreases monotonically with the tax rate, τ. As the tax rate increases, the true net return to effort (after the tax is paid) decreases, alleviating the ambiguity problem. In the extreme case where $\tau = 1$, the return to effort is negative because the agents cannot increase their level of output, and they still have to pay their effort cost. The optimal level of effort is an increasing function of the effort-cost parameter, $a_j(\lambda)$. Individuals with lower effort cost exert more effort. The next two lemmas summarize these results.

Lemma 1 (Optimal effort decision). *Let Assumptions 1 and 3 hold. Then, the optimal level of effort in period 1 is given by:*

$$\delta_H^* = \begin{cases} 0 & \textit{if } \delta_H \leq 0, \\ 1 & \textit{if } \delta_H \geq 1, \\ \delta_H & \textit{if } \delta_H \in (0,1). \end{cases} \tag{6}$$

where $\delta_H = a_j(\lambda)(1-\tau)[\bar{\beta}_\lambda - \alpha]$ and $\bar{\beta}_\lambda := \lambda\beta_L + (1-\lambda)\beta_U$.

Lemma 2 (Decreasing effort as a function of tax). *Let Assumptions 1 and 3 hold. Then, the optimal level of effort, δ_H^*, decreases as the tax rate, τ, increases.*

3.2 Optimal Desired Tax

Assuming for the moment an interior optimum for effort and substituting equations (6) and (4) into equation (5) yields the individuals' welfare at the time when effort is chosen:

$$\begin{aligned} U_j^\tau(\tau;\lambda) &:= U_j^{CH}(\delta_H^*), \\ &= (1-\tau) + \frac{1}{2}a_j(1-\tau)^2(\bar{\beta}_\lambda - \alpha)^2 + \tau\left[\alpha\bar{a}(\lambda)(1-\tau)(\bar{\beta}_\lambda-\alpha)^2\right] + M(\lambda). \end{aligned} \tag{7}$$

Agent j's desired tax rate is given by maximizing (7) with respect to the tax rate, τ. Lemma 3 shows the ideal tax rate for agent j.

Lemma 3 (Desired tax). *Let Assumptions 1, 2, and 3 hold. Then, the ideal tax rate of agent j is:*

$$\tau_j^*(\lambda) = \begin{cases} \tau^* & \text{if} \quad \tau^* > \underline{\tau}, \\ \underline{\tau} & \text{otherwise}, \end{cases} \tag{8}$$

where $\tau^* = \bar{a}(\lambda)G - (\bar{\beta}-\alpha)^2 a_j(\lambda) / 3\bar{a}(\lambda)G$, G is defined in Appendix B, and $\underline{\tau}$ is the lowest possible tax (e.g., *$\underline{\tau} = 0$ if $\tau \in [0,1)$*).

Four intuitive effects stand out from equation (8). The first is that disadvantaged individuals typically impose a positive redistributive tax rate, whereas advantaged ones prefer a more regressive redistributive policies or the lowest possible tax rate. The second is that the desired tax rate of disadvantaged agents is a decreasing function of the effort-cost parameter. Individuals with higher effort-cost have a higher desired tax. The third is that the higher the pessimism rate, the higher the cost of effort and, thus, the desired tax is higher. Finally, the optimal desired tax rate is an increasing function of the effort-cost parameter.

Now consider how political preferences are aggregated through voting. Voters' preferences over τ are single-peaked $\left(\frac{\partial^2 U_j^\tau(\tau;\lambda)}{\partial \tau^2} < 0\right)$ and, thus, the median voter theorem applies. With the disadvantaged forming a majority, the equilibrium tax outcome, T, is the optimal desired tax from this group, $\tau_L^*(\lambda)$.

The following lemma establishes how the equilibrium tax outcome varies with the pessimism rate. Figure 3 shows the path of the equilibrium tax rate as a function of the pessimism rate.

Lemma 4 (Increasing equilibrium tax as a function of pessimism). *Let Assumptions 1, 2, and 3 hold and $\lambda \in (\underline{\lambda}, \bar{\lambda})$. Then, the equilibrium tax outcome, $T = \tau_L^*(\lambda)$, is increasing in the pessimism rate, λ.*

3.3 Cognitive Decisions

We turn now to the cognitive problem in period 0. Substituting $\tau = T$ from equation (8) into (7) yields the utility of individual j in period 0. The cognitive decision problem is:

$$\max_{\lambda \in [0,1]} U_j^{\tau}(T; \lambda), \tag{9}$$

where T is the equilibrium tax rate.

The solution to the problem in (9), λ^*, is a complicated expression. Lemma 5 identifies conditions for the optimal pessimism rate to be an increasing function of the equilibrium tax rate.

Lemma 5 (Increasing pessimism as a function of equilibrium tax). *Let Assumptions 1, 2 and 3 hold and $\lambda \in (\underline{\lambda}, \bar{\lambda})$. Then, the optimal pessimism rate, λ^*, increases with the equilibrium tax rate, T.*

If the tax rate is low enough, being optimistic pays off because, by reducing the effort cost, it increases the *ex ante* level of utility, provided that the cognitive costs are not too high.

3.4 Pessimism/Welfare State v. Optimism/Laissez Faire

We proceed now to find the equilibria of the redistribution model. Let $\lambda \in \{\hat{\lambda}, \tilde{\lambda}\}$, where $\underline{\lambda} < \hat{\lambda} < \tilde{\lambda} < \bar{\lambda}$. In equilibrium, agents will be either optimistic or pessimisticic. Lemma 6 displays this result.

Lemma 6 (Equilibrium pessimism as a function of equilibrium tax). *Let Assumptions 1, 2 and 3 hold and $\lambda \in \{\hat{\lambda}, \tilde{\lambda}\}$. Then, the equilibrium pessimism rate, λ^*, has the following form:*

$$\lambda^* = \begin{cases} \hat{\lambda} & \text{if } T < \tilde{T}, \\ \tilde{\lambda} & \text{if } T > \tilde{T}. \end{cases}$$

Because the L-types form a majority, the equilibrium tax outcome, T, is the optimal desired tax from this group, $\tau_L^*(\lambda)$. Thus, the welfare of this group is maximized by choosing τ_L^* and λ^* optimally. By Lemma 5, as the equilibrium tax rate (which is the desired tax rate of this group) increases, the optimal pessimism rate also increases. Thus, the disadvantaged agents choose the equilibrium tax rate such that $T = \tau_L^*(\tilde{\lambda})$ or $\big(T = \tau_L^*(\hat{\lambda})\big)$ (because the desired tax rate and the pessimism rate are the solutions to the L-types optimizations problems in period 1 and 0, respectively).

On the one hand, H-types cannot impose their desired tax rate and face τ_L^*. Lemma 6 says that when the tax rate is high, $\tau_L^*(\tilde{\lambda})$, the advantaged agents also prefer to be pessimistic because the cost of effort together with the cognitive technology overcompensate the potential *ex ante* welfare gains of being optimistic. In other words, because the tax rate is too high, the true return to effort has to be high enough to compensate for the associated effort costs. The individual must be very optimistic to believe that the true return to effort is that high but being that optimistic is costly in terms of cognitive technology. Therefore, with a high rate of redistribution, the *ex ante* true return to effort is reduced and generates weak incentives to be optimistic. Thus, agents choose a high λ^* and are pessimistic in this equilibrium. This equilibrium is the *Pessimism/Welfare State.*

On the other hand, a low redistribution tax rate generates strong incentives to be optimistic and, thus, individuals choose a low λ^* and end up being optimistic in equilibrium. This is the *Optimism/Laissez Faire* equilibrium. By Lemma 2, the level of effort exerted in the Pessimism/Welfare State equilibrium is lower than the level exerted in the Optimism/Laissez Faire one.

The following proposition summarizes the discussion in the previous paragraphs and describes the equilibria. Figure 4 graphically depicts the equilibria.

Proposition 1 (Possibility of multiple equilibria). *Let Assumptions 1, 2 and 3 hold,* $\lambda \in \{\hat{\lambda}, \tilde{\lambda}\}$ *and* $\tau \in [0, 1)$. *Then, the following two equilibria are possible:*

(a) Pessimism/Welfare State*: Agents are pessimistic,* $\lambda^* = \tilde{\lambda}$*, they impose a high tax rate,* $T = \tau_L^*(\tilde{\lambda})$*, and they exert low effort,* $\delta_H^*(\tilde{\lambda})$*.*

(b) Optimism/Laissez Faire*: Agents are optimistic,* $\lambda^* = \hat{\lambda}$*, they impose a low tax rate,* $T = \tau_L^*(\hat{\lambda})$*, and they exert high effort,* $\delta_H^*(\hat{\lambda})$*.*

with $\tilde{\lambda} > \hat{\lambda}$, $\tau_L^*(\tilde{\lambda}) > \tau_L^*(\hat{\lambda})$*, and* $\delta_H^*(\tilde{\lambda}) < \delta_H^*(\hat{\lambda})$*.*

4 Extensions

I present two main extensions of the baseline model.

The first extension considers a model with a continuum of citizen types. To that end, I use three additional technical assumptions. The first is that the true return to effort is not too low (that is, $\beta_H > \beta_L > \underline{\beta}$). This is a natural extension of Assumption 1 and ensures that the median voter is not always extremely pessimistic.[10] The second is a piecewise continuity assumption about the distribution of agents' cost parameters. This is a natural extension of Assumption 2 to the continuous case. The final is that $a(\lambda)$ is convex, which ensures that the cost function of advantaged and disadvantaged agents is sufficiently differentiated. Using this

[10]As explained below, if the median voter is always extremely pessimistic, then the optimal effort is always zero, as in the maximin criterion, regardless of the tax and the redistribution.

generalization, I follow a series of lemmas similar to the ones in the baseline model, finishing with an analogue proposition that shows the possibility of multiple equilibria. Appendix C presents the results.

The second type of extensions considers other criteria for choosing among undominated strategies. The Hurwicz criterion provides a solution that is, in general, fractional, and where the degree of pessimism encompasses other criteria in the corner cases of extreme pessimism (maximin criteria) or extreme optimism (maximax criteria).

A fractional solution is necessary for the main result in the paper, the possibility of multiple equilibria. The reason is simple. If individuals are always extremely pessimistic (*e.g.*, maximin criteria or corner solutions from other criteria), then they will not exert any effort, $\delta^* = 0$, because exerting effort is costly but does not pay off. Redistribution cannot affect this outcome because the tax rate does not affect the optimal decisions of effort or pessimism. Similarly for criteria where individuals are always extremely optimistic (*e.g.*, maximax criteria or corner solutions from other criteria).

With the previous discussion in mind, I present extensions using the Bayes rule and the minimax regret criteria in Appendices D and E. These criteria show alternative approaches (to the Hurwicz criterion) where the optimal effort solution is also typically fractional.[11] The Bayes rule consists of using a subjective known distribution for the true return of effort and is similar to Benabou and Tirole (2006). Appendix D presents the results of the Bayes rule criterion using the same steps as in the baseline model. The minimax regret criterion consists of minimizing the maximum regret of an effort allocation. The solution of the minimax-regret is substantially more cumbersome than the Hurwicz and Bayes rule criteria because the former requires to compute the *regret* of an allocation. Appendix E shows that it is also possible to obtain the multiple equilibria result using the minimax regret criterion under some technical simplifying assumptions.

5 Concluding Remarks

When scholars study the conflicting views that agents and voters have about redistributive taxation and the extent to which people are responsible for their own fate, they assume that individuals know the distribution of counterfactual outcomes. The counterfactual outcomes are point identified. A natural question is whether it is possible to reconcile such conflicting views in a context of ambiguity, where the counterfactual outcomes that would have occurred

[11] A natural question that might arise is whether the main result of the paper—the possibility of multiple equilibria—can also be obtained using other criteria, such as the ones discussed in the introduction and concluding remarks. I hypothesize that the answer is "yes" if the optimal effort solution is fractional, although one might have to make some simplifying technical assumptions, such as the ones in the paper. However, I have only proved proposition 1 under the criteria discussed in this section.

had other actions been chosen are not observable. That is, when the counterfactual outcomes are partially identified.

The model presented in this paper provides an answer to this question. In the model, agents have partial knowledge of the relative importance of effort and the counterfactual outcomes that would have occurred had other actions been chosen. The simple model provides a robust framework to understand the political economy of redistribution in a context of ambiguity. It complements the fundamental literature analyzing the link between beliefs and economic institutions. It allows interpretation of stylized facts, such as the ones in Figure 1, in a context of ambiguity.

In the model, I introduce a cognitive extension to the Hurwicz criterion to choose among undominated outcomes. This extension provides a flexible environment that allows endogenizing the degree of pessimism that individuals choose to hold in equilibrium.

To make the problem tractable, I made several simplifying assumptions, such as a tractable functional form for the cost of effort, two types of individuals, and the cognitive Hurwicz criterion. Subject to these simplifications, I have presented a simple model of redistribution and showed that multiple equilibria might arise *even* in the presence of ambiguity: One where the rate of redistribution is high, agents are pessimistic, and exert low effort—*Pessimism/Welfare State*—, and another where the redistribution tax rate is low, agents are optimistic, and exert high effort—*Optimism/Laissez Faire*. I have also studied extensions using a continuum of types and other criteria, such as the Bayes rule and the minimax regret.

Several extensions of the model would be interesting to analyze as avenues for future research. One would be to study the possibility of multiple equilibria using a Bayesian criterion for non-additiive expected utility and using a set of priors instead of a single one, as in the foundational articles of Schmeidler (1989), Gilboa and Schmeidler (1989), and Gilboa and Schmeidler (1993). Another would be to evaluate how agents would behave in a dynamic setting. Of special interest might be to study how a dynamic process to reduce ambiguity might affect outcomes, agents, and equilibria, such as the sequential reduction in ambiguity in Manski (2004), and the recursive and consequentialist belief revision criterion in Gul and Pesendorfer (2018). Another avenue would be to introduce social interactions arising from agents learning from past cohort experiences.

References

Alesina, A., and G.-M. Angeletos (2005a): "Corruption, inequality, and fairness," *Journal of Monetary Economics*, 52(7), 1227–1244. https://www.jstor.org/stable/1209137. (Cited on page 1.)

——— (2005b): "Fairness and redistribution," *American Economic Review. DOI: 10.1257/0002828054825655*, 95(4), 960–980. (Cited on page 1.)

Alesina, A., and E. Glaeser (2004): *Fighting poverty in the US and Europe: A world of difference*. Oxford University Press. (Cited on page 2.)

Alesina, A., E. Glaeser, and B. Sacerdote (2001): "Why Doesn't the United States Have a European-Style Welfare State?," *Brookings Papers on Economic Activity.*, 2001(2), 187–254. DOI: 10.1016/j.jmoneco.2005.05.003. (Cited on pages 1, 2, and 19.)

Benabou, R., and E. A. Ok (2001): "Social mobility and the demand for redistribution: the POUM hypothesis," *The Quarterly Journal of Economics. DOI: 10.1162/00335530151144078*, 116(2), 447–487. (Cited on page 2.)

Benabou, R., and J. Tirole (2006): "Belief in a just world and redistributive politics," *The Quarterly journal of economics*, 121(2), 699–746. DOI: 10.1162/qjec.2006.121.2.699. (Cited on pages 1, 2, 5, 7, 8, and 14.)

Bossaerts, P., P. Ghirardato, S. Guarnaschelli, and W. R. Zame (2010): "Ambiguity in Asset Markets: Theory and Experiment," *The Review of Financial Studies*, 23(4), 1325–1359. (Cited on page 3.)

Choquet, G. (1954): "Theory of capacities," in *Annales de l'institut Fourier*, vol. 5, pp. 131–295. (Cited on page 2.)

Corneo, G., and H. P. Grüner (2002): "Individual preferences for political redistribution," *Journal of Public Economics*, 83(1), 83–107. DOI: 10.1016/S0047–2727(00)00172–9. (Cited on page 2.)

Cunha, F., and J. J. Heckman (2010): "Investing in Our Young People," NBER Working Paper 16201, National Bureau of Economic Research, DOI: 10.3386/w16201. (Cited on page 8.)

Di Tella, R., S. Galiani, and E. Schargrodsky (2007): "The formation of beliefs: evidence from the allocation of land titles to squatters," *The Quarterly Journal of Economics*, 122(1), 209–241. DOI: 10.1162/qjec.122.1.209. (Cited on page 2.)

Di Tella, R., and R. MacCulloch (2009): "Why Doesn't Capitalism Flow to Poor Countries?," *Brookings Papers on Economic Activity*, 2009(1), 285–321. (Cited on page 1.)

Ellsberg, D. (1961): "Risk, ambiguity, and the Savage axioms," *The Quarterly Journal of Economics*, pp. 643–669. DOI: 10.2307/1884324. (Cited on page 2.)

Fong, C. (2001): "Social preferences, self-interest, and the demand for redistribution," *Journal of Public economics*, 82(2), 225–246. DOI: 10.1016/S0047–2727(00)00141–9. (Cited on page 2.)

——— (2004): "Which beliefs matter for redistributive politics? Target-specific versus general beliefs about the causes of income," *Unpublished manuscript. Available at (accessed April 29, 2022): http://www.pubchoicesoc.org:80/papers/fong.pdf*. (Cited on page 2.)

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004): "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory*, 118(2), 133–173. (Cited on page 3.)

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18(2), 141–153. (Cited on pages 3 and 15.)

——— (1993): "Updating Ambiguous Beliefs," *Journal of Economic Theory*, 59(1), 33–49. (Cited on pages 3 and 15.)

Gul, F., and W. Pesendorfer (2018): "Evaluating Ambiguous Random Variables and Updating by Proxy," Working papers, Princeton University. Economics Department. (Cited on pages 3 and 15.)

Heckman, J. J. (2006): "Skill formation and the economics of investing in disadvantaged children," *Science*, 312(5782), 1900–1902. DOI: 10.1126/science.1128898. (Cited on page 8.)

Hirschman, A. O., and M. Rothschild (1973): "The changing tolerance for income inequality in the course of economic development: With a mathematical appendix," *The Quarterly Journal of Economics. DOI: 10.2307/1882024*, 87(4), 544–566. (Cited on page 2.)

Hochschild, J. L. (1981): *What's fair?: American beliefs about distributive justice*. Harvard University Press. (Cited on page 2.)

Hurwicz, L. (1951): "Some specification problems and applications to econometric models," *Econometrica*, 19(3), 343–344. (Cited on pages 3 and 6.)

INGLEHART, R. (2018): *Culture shift in advanced industrial society*. Princeton University Press. (Cited on page 2.)

KEYNES, J. M. (1921): *A treatise on probability*. Macmillan and Company, limited. (Cited on page 2.)

KNIGHT, F. H. (1921): *Risk, uncertainty and profit*, vol. 31. Houghton Mifflin. (Cited on page 2.)

LADD, E. C., AND K. H. BOWMAN (1998): *Attitudes toward economic inequality*. AEI Press. (Cited on page 2.)

LORA, E., AND M. OLIVERA (2004): "What makes reforms likely: Political economy determinants of reforms in Latin America," *Journal of Applied Economics*, 7(1), 99–135. (Cited on page 1.)

MANSKI, C. F. (2004): "Social learning from private experiences: the dynamics of the selection problem," *The Review of Economic Studies*, 71(2), 443–458. DOI: 10.1111/0034–6527.00291. (Cited on page 15.)

——— (2009): *Identification for prediction and decision*. Harvard University Press. (Cited on pages 2 and 4.)

MELTZER, A. H., AND S. F. RICHARD (1981): "A rational theory of the size of government," *Journal of Political Economy*, 89(5), 914–927. DOI: doi.org/10.1086/261013. (Cited on page 1.)

PERSSON, T., AND G. TABELLINI (2005): *The economic effects of constitutions*. MIT press. (Cited on pages 19 and 22.)

PIKETTY, T. (1995): "Social mobility and redistributive politics," *The Quarterly Journal of Economics*, 110(3), 551–584. DOI: 10.2307/2946692. (Cited on pages 1, 2, 5, and 8.)

ROMER, T. (1975): "Individual welfare, majority voting, and the properties of a linear income tax," *Journal of Public Economics*, 4(2), 163–185. (Cited on page 1.)

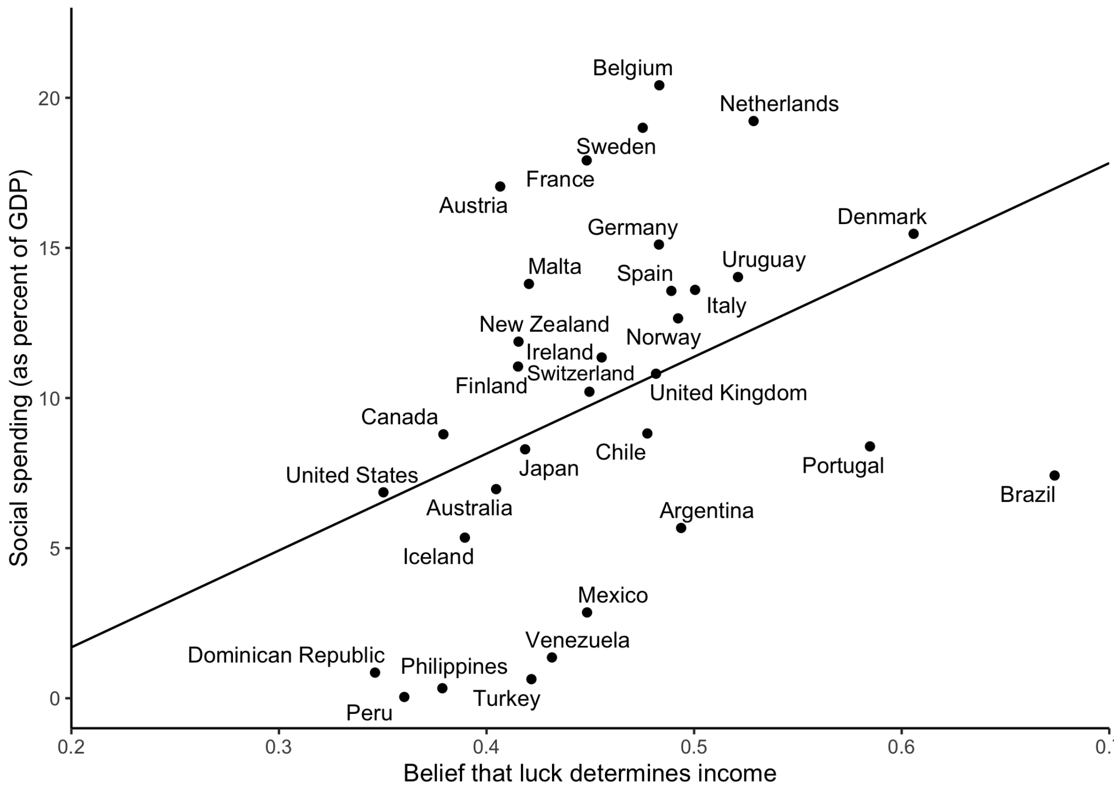
SAVAGE, L. J. (1954): *The foundations of statistics*. John Wiley & Sons. (Cited on page 2.)

SCHMEIDLER, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57(3), 571–587. (Cited on pages 2 and 15.)

STOKES, S. C. (2001): *Mandates and democracy: Neoliberalism by surprise in Latin America*. Cambridge University Press. (Cited on page 1.)

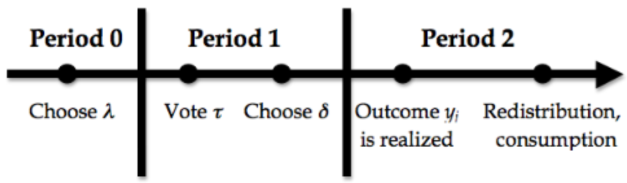
WALD, A. (1950): *Statistical decision functions*. Wiley, New York. (Cited on page 3.)

Figure 1: Social spending and beliefs that luck determines income.



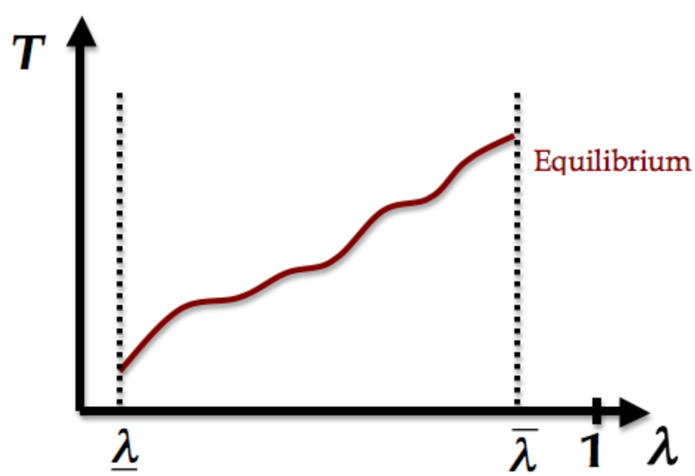
Notes: Author's calculation based on data from the OECD Economic Outlook and the World Values Survey described in Appendix A. The figure displays, in the vertical axis, *Social Spending* (as a percent of the GDP), obtained from the OECD Economic Outlook (for the period 1960-1998) obtained from Persson and Tabellini (2005) and, in the horizontal axis, the *Belief that Luck Determines Income* (mean value for each country), obtained from the World Values Survey (for the period 1981-1997), indexed from 0.1 to 1, with 1 indicating the strongest belief. The figure replicates Figure 6 from Alesina, Glaeser, and Sacerdote (2001). See Appendix A for details.

Figure 2: Timeline.



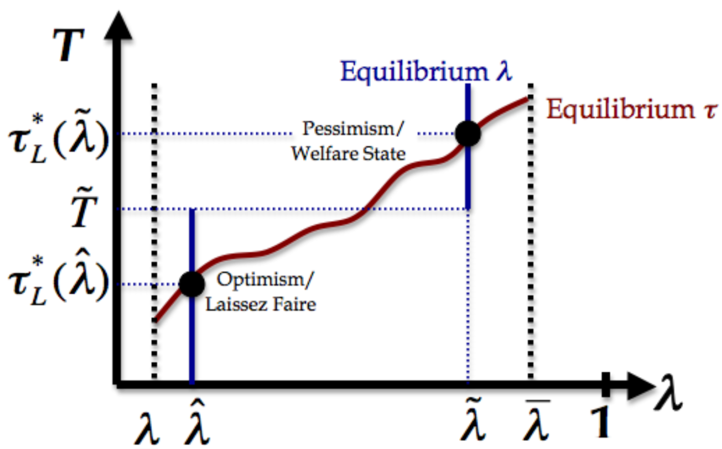
Notes: The figure depicts the timeline of the actions. At $t = 0$, each agent chooses the optimal degree of pessimism, λ, using the cognitive technology described in the text. At the start of period 1, individuals vote over the linear tax rate, τ, that determines how the output will be redistributed in period 2. After voting, the agents choose the optimal level of effort, δ. Finally, in period 2, the outcome $y(e_j)$ and the redistribution are realized, and agents consume their disposable income.

Figure 3: Equilibrium tax *vs.* pessimism.



Notes: The figure shows that the path of the equilibrium tax outcome, T, increases with the pessimism rate, λ, following Lemma 4.

Figure 4: Multiple Equilibria.



Notes: The figure summarizes the two equilibria from Proposition 1. In the *Pessimism/Welfare State* equilibrium, agents are pessimistic, $\lambda^* = \tilde{\lambda}$, they impose a high tax rate, $T = \tau_L^*(\tilde{\lambda})$, and they exert low effort, $\delta_H^*(\tilde{\lambda})$. In the *Optimism/Laissez Faire* equilibrium, agents are optimistic, $\lambda^* = \hat{\lambda}$, they impose a low tax rate, $T = \tau_L^*(\hat{\lambda})$, and they exert high effort, $\delta_H^*(\hat{\lambda})$, with $\tilde{\lambda} > \hat{\lambda}$, $\tau_L^*(\tilde{\lambda}) > \tau_L^*(\hat{\lambda})$, and $\delta_H^*(\tilde{\lambda}) < \delta_H^*(\hat{\lambda})$.

Appendix

A Data

This data appendix describes the data sources and variables used in Figure 1, *Social Spending* and *Belief that Luck Determines Income*.

Social spending data

The dataset was obtained from Persson and Tabellini (2005),[12] for the period 1960-1998. The variable used is SSW, which is defined in Persson and Tabellini's data appendix as:

> **Social spending.** Consolidated central government expenditures on social services and welfare as a percentage of GDP as reported in the IMF GFS Yearbook divided by GDP and multiplied by 100. Source: IMF - GFS Yearbook 2000 and IMF-IFS CD-Rom.

Survey data

The dataset was obtained from the World Values Survey (WVS),[13] for the period 1981-1997. The question used corresponds to E040, where responders were asked whether they agree (higher value) or disagree (lower value) with the statement that: "Hard work doesn't generally bring success." Specifically, the variable is constructed as follows.

> **Belief that luck determines income.** Average responses by country to the WVS question:
>
> *Now I'd like you to tell me your views on various issues. How would you place your views on this scale? 1 means you agree completely with the statement there; 10 means you agree completely with the statement there; and if your views fall somewhere in between, you can choose any number in between. "In the long run, hard work usually brings a better life"* vs. *"Hard*

[12] Downloaded from Enrico Tabellini's webpage available at (accessed on April 21, 2022): https://didattica.unibocconi.it/mypage/index.php?IdUte=48805&idr=4273. File name: 60panel_26maj.dta.

[13] Downloaded from the World Values Survey's webpage available at (accessed on April 21, 2022): https://www.worldvaluessurvey.org/WVSEVStrend.jsp and https://search.gesis.org/research_data/ZA7503?doi=10.4232/1.13736. File names: WVS.dta and EVS.dta, respectively, under datasets on the webpages.

work doesn't generally bring success—it's more a matter of luck and connections:

1: *In the long run, hard work usually brings a better life.*

2: . . .

⋮

9: . . .

10: *Hard work doesn't generally bring success—it's more a matter of luck and connections.*

The WVS data in Figure 1 correspond to the average E040 responses by country divided by 10.

B Proofs

B.1 Maximin Criterion

Proof optimal effort Maximin Criterion. The agent acts as if $\beta = \beta_L$ and solves:

$$\max_{\delta\in[0,1]} (1-\tau)\Big[\alpha + \delta[\beta_L - \alpha]\Big] - C_j(\delta).$$

Assumption 1 implies that $\beta_L < \alpha$ and, thus, the solution to the previous problem is trivially solved by setting $\delta^*_{MMin} = 0$. □

Proof that effort does not depend on tax Maximin Criterion. The agent's welfare under Maximin criterion is given by:

$$\begin{aligned} U_j^{MMin}(\lambda,\delta,\tau) &= (1-\tau)\Big[\alpha + \delta[\beta_L - \alpha]\Big] - C_j(\delta) + \tau\bar{Y} + M(\lambda), \\ &= \alpha + M(\lambda), \end{aligned}$$

where the last equality is obtained by replacing maximin solution $\delta^*_{MMin} = 0$.

Then:

$$\frac{\partial U_j^{Mmin}(\lambda,\delta,\tau)}{\partial\tau} = 0.$$

□

B.2 Maximax Criterion

Proof optimal effort Maximax Criterion. Now the agent sets $\beta = \beta_U$ and solves:

$$\max_{\delta\in[0,1]} (1-\tau)\Big[\alpha + \delta[\beta_U - \alpha]\Big] - C_j(\delta).$$

Assumption 1 implies that $\beta_U > \alpha$. Therefore, the maximax solution is $\delta^*_{MMax} = 1$.

□

Proof that effort does not depend on tax Maximax Criterion. The agent's welfare under Maximax criterion is given by:

$$\begin{aligned} U_j^{MMax}(\lambda,\delta,\tau) &= (1-\tau)\Bigg[\alpha+\delta[\beta_U-\alpha]\Bigg] - C_j(\delta)+\tau\bar{Y}+M(\lambda), \\ &= (1-\tau)\Bigg[\alpha+[\beta_U-\alpha]\Bigg] - \frac{1}{2a_j(\lambda)} - \tau\Bigg[\alpha+[\beta_U-\alpha]\Bigg] + M(\lambda), \end{aligned}$$

where the last equality is obtained by replacing maximin solution $\delta^*_{MMax} = 1$ and $\bar{Y} = \alpha + (\beta_U - \alpha)$.

Then:

$$\frac{\partial U_j^{MMax}(\lambda,\delta,\tau)}{\partial\tau} = 0.$$

□

B.3 Hurwicz Criterion

Proof optimal effort Hurwicz Criterion. According to this criteria the agent solves the following problem:

$$\begin{aligned} U_j^H(\delta^*_H;\lambda_J) &= \max_{\delta\in[0,1]} \lambda_j\Bigg\{\inf_{\beta\in[\beta_L,\beta_U]}\Bigg(\mathbb{E}[(1-\tau)y(e_j)]-C_j(\delta)\Bigg)\Bigg\} \\ &+ (1-\lambda_j)\Bigg\{\sup_{\beta\in[\beta_L,\beta_U]}\Bigg(\mathbb{E}[(1-\tau)y(e_j)]-C_j(\delta)\Bigg)\Bigg\}, \\ &= \max_{\delta\in[0,1]}(1-\tau)\Bigg\{\alpha+\Bigg[\Bigg(\lambda_j\beta_L+(1-\lambda_j)\beta_U\Bigg)-\alpha\Bigg]\delta_j\Bigg\} - C_j(\delta), \\ &= \max_{\delta\in[0,1]}(1-\tau)\big\{\alpha+\big[\bar{\beta}_\lambda-\alpha\big]\delta_j\big\} - C_j(\delta), \end{aligned}$$

where $\bar{\beta}_\lambda := \lambda_j\beta_L + (1-\lambda_j)\beta_U$.

The first-order necessary condition for an interior solution is:

$$\begin{aligned} (1-\tau)[\bar{\beta}_\lambda-\alpha] &= C_j'(\delta_H), \\ &= \frac{\delta_H}{a_j}. \end{aligned}$$

Solving for δ_H yields $\delta_H = a_j(1-\tau)[\bar{\beta}_\lambda - \alpha]$. Thus, the solution is:

$$\delta^*_H = \begin{cases} 0 & \text{if } \delta_H \leq 0, \\ 1 & \text{if } \delta_H \geq 1, \\ \delta_H & \text{if } \delta_H \in (0,1). \end{cases}$$

The second-order sufficient condition is satisfied because $\frac{\partial^2 U_j^H(\delta)}{\partial \delta^2} = -\frac{\partial^2 C_j(\delta)}{\partial \delta^2} := -C_j''(\delta) = -\frac{1}{a_j(\lambda)} < 0$ by Assumption 3. (The second-order sufficient condition only requires $C_j(\delta)$ to be convex.)

□

PROOF OF INTERIOR CONDITION FOR PESSIMISM RATE, λ. For $\overline{\lambda}$, $\partial\delta_H/\partial\lambda < 0$, then $\overline{\lambda} = \{\lambda : \delta_H = a_j(1-\tau)[\lambda\beta_L + (1-\lambda)\beta_U - \alpha] = 0\}$. Then, $\overline{\lambda} = \frac{\beta_U - \alpha}{\beta_U - \beta_L} \in (0,1)$. For $\underline{\lambda}$, let $\underline{\lambda}^* = \{\lambda : a_j(1-\tau)[\lambda\beta_L + (1-\lambda)\beta_U - \alpha] = 1\}$, yielding $\underline{\lambda}^* = \frac{\beta_U - \alpha}{\beta_U - \beta_L} - \frac{1}{a_j(1-\tau)} \lesseqgtr 0$. Then, $\underline{\lambda} = \max\{0, \lambda : \underline{\lambda}^* < 1\}$. □

PROOF OF LEMMA 1. By backwards induction, the optimal level of effort is the value of δ that maximizes equation (5) because each agent chooses effort optimally using the Hurwicz criterion. Knowing the degree of pessimism, λ, and the equilibrium tax rate, τ, this optimization problem is simplified to equation (2) and the optimal level of effort is given by equation (3). That is, $\arg\max_{\delta\in[0,1]} U_j^{CH}(\lambda,\delta,\tau) = \arg\max_{\delta\in[0,1]} U_j^H(\delta;\lambda_j)$. Then, equation (6) follows from optimal effort under the Hurwicz criterion. □

PROOF OF LEMMA 2. By Lemma 1, the optimal level of effort is given by equation (6). If $(\bar{\beta}_\lambda - \alpha) \leq 0$, then $\delta_H \leq 0$ and $\delta_H^* = 0$ (corner solution), so $\frac{\partial \delta_H}{\partial \tau} = 0$. If $(\bar{\beta}_\lambda - \alpha) > 0$, there are two possibilities. If $(\bar{\beta}_\lambda - \alpha) \geq \frac{1}{a_j(1-\tau)}$, then $\delta_H \leq 1$ and $\delta_H^* = 1$ (corner solution), so $\frac{\partial \delta_H}{\partial \tau} = 0$. If $0 < (\bar{\beta}_\lambda - \alpha) < \frac{1}{a_j(1-\tau)}$, we have an interior solution because $\delta_H \in (0,1)$ and $\delta_H^* = \delta_H$. Taking the derivative with respect to the tax rate yields $\frac{\partial \delta_H}{\partial \tau} = -(\bar{\beta}_\lambda - \alpha) < 0$. □

PROOF OF LEMMA 3. The agent's welfare is given by equation (5):

$$U_j^{CH}(\lambda,\delta,\tau) = (1-\tau)(\alpha + (\bar{\beta}_\lambda - \alpha)\delta) - C(\delta) + \tau\bar{Y} + M(\lambda).$$

Maximizing the previous expression with respect to τ:

$$\frac{\partial U_j^{CH}(\lambda,\delta,\tau)}{\partial \tau} = -(1-\tau)(\bar{\beta} - \alpha)^2 a_j(\lambda) + \bar{a}(\lambda)(1-\tau)^2 G - 2\tau(1-\tau)\bar{a}(\lambda)G,$$

where $G := [(1-\phi)a_L(\bar{\beta}_{\lambda_L} - \alpha) + \phi a_H(\bar{\beta}_{\lambda_H} - \alpha)][(1-\phi)(\bar{\beta}_{\lambda_L} - \alpha) + \phi(\bar{\beta}_{\lambda_H} - \alpha)]$.

In an interior solution:

$$\tau^* = \frac{\bar{a}(\lambda)G - (\bar{\beta} - \alpha)^2 a_j(\lambda)}{3\bar{a}(\lambda)G}.$$

The second-order sufficient condition is $\frac{\partial^2 U_j^{CH}(\lambda,\delta,\tau^*)}{\partial \tau^2} = -5(\bar{\beta}-\alpha)^2 a_L(\lambda) - 2\bar{a}(\lambda)G < 0$ because $G > 0$ and $a_L(\lambda) > 0$.

□

Proof of Lemma 4. Voters' preferences over τ are single peaked and, thus, the median voter theorem applies. Because $\phi < \frac{1}{2}$, the equilibrium tax outcome, T, is the optimal desired tax from the disadvantaged group, $\tau_L^*(\lambda)$. Denote $\bar{\beta}_L = \lambda_L \beta_L + (1-\lambda_L)\beta_U$, where λ_L is the pessimism rate of disadvantaged agents; similarly, for advantaged agents. The lemma holds trivially if the equilibrium tax rate is 0. If the equilibrium tax rate is interior, then by Lemma 3, $\tau^* = \frac{1}{3} - \frac{(\bar{\beta}_L-\alpha)^2 a_j(\lambda)}{3\bar{a}(\lambda)P}$, where $P := a_L(\lambda_L)(1-\phi)^2(\bar{\beta}_{\lambda_L}-\alpha)^2 + \phi(1-\phi)(a_L(\lambda_L) + a_H(\lambda_A))(\bar{\beta}_{\lambda_A}-\alpha)(\bar{\beta}_{\lambda_L}-\alpha) + a_H(\lambda_A)\phi^2(\bar{\beta}_{\lambda_A}-\alpha)^2$. Then:

$$\frac{\partial T}{\partial \lambda} = \frac{\partial \tau_L^*(\lambda)}{\partial \lambda},$$
$$= -\left[\underbrace{(3\bar{a}(\lambda)P)^{-1}}_{:=A}\underbrace{2a_j(\lambda)(\bar{\beta}_L-\alpha)(\beta_L-\beta_U) + (\bar{\beta}_L-\alpha)^2 a'(\lambda)}_{:=B}\right].$$

Note that:

- $A > 0$ because $\bar{a}(\lambda) > 0$ and $P > 0$.
- $B < 0$ because $\beta_L - \beta_U < 0$, $a'(\lambda) < 0$, $2a_j(\lambda) > 0, \bar{\beta}_L - \alpha > 0$.

Then, $\frac{\partial T}{\partial \lambda} > 0$.

Next, we note that pessimism is decreasing in agents' cost parameters. Let Assumptions 1, 2, 3 hold and $\lambda \in (\underline{\lambda}, \bar{\lambda})$. Consider the problem of choosing λ in period 0:

$$\max_{\lambda\in(\underline{\lambda},\bar{\lambda})} (1-\tau)\big[\alpha + \big(\bar{\beta}_\lambda - \alpha\big)\delta\big] - C_j(\delta) + \tau\bar{Y} + M(\lambda).$$

Or, after replacing δ_H^* and the equilibrium tax rate:

$$\max_{\lambda\in(\underline{\lambda},\bar{\lambda})} F_j(\lambda, T), \tag{B.1}$$

where $F_j(\lambda, T) := (1-T)\big[\alpha + \big(\bar{\beta}_\lambda - \alpha\big)\delta_H^*\big] - C_j(\delta_H^*) + T\bar{Y} + M(\lambda)$.

Let $J_j(\Lambda, a_j) := (1-T)\alpha + \frac{1}{2}(a_j + a(-\Lambda))(1-T)^2(\bar{\gamma}_\Lambda - \alpha)^2 + T\bar{Y} + M(-\Lambda)$, where $\Lambda := -\lambda$, and $\bar{\gamma}_\Lambda := -\beta_L\Lambda + (1+\Lambda)\beta_U = \beta_L\lambda + (1-\lambda)\beta_U = \bar{\beta}_\lambda$.

Then:

$$\frac{\partial J_j(\Lambda, a_j)}{\partial a_j} = \frac{1}{2}(1-T)^2(\bar{\gamma}_\Lambda - \alpha)^2.$$

$$\frac{\partial^2 J_j(\Lambda, a_j)}{\partial a_j \partial \Lambda} = \underbrace{(1-T)(\bar{\gamma}_\Lambda - \alpha)^2 \frac{\partial T}{\partial \lambda}}_{:=C} + \underbrace{(1-T)\frac{\delta_H^*}{a_j + a(\lambda)}(\beta_U - \beta_L)}_{:=D}.$$

Note that:

- $C > 0$ because $(1-T) \in (0,1)$ and T is increasing in λ.
- $D > 0$ because $\delta_H^* \in (0,1)$.

Therefore, $J_j(\Lambda, a_j)$ has increasing differences in (Λ, a_j), and pessimism being decreasing in agents' cost parameters is established using monotone comparative statics.

Finally, note that the preferred tax rate of the disadvantaged agents is strictly increasing with the pessimism rate. It then follows that disadvantaged agents prefer an interior tax rate. Let assumptions 1, 2, and 3 hold, and $\lambda \in (\underline{\lambda}, \bar{\lambda})$. Agent j preferred tax is $\tau^* = \frac{1}{3}[1 - \frac{(\bar{\beta}-\alpha)^2 a_j(\lambda)}{\overline{a(\lambda)}P}]$, where $P := a_L(\lambda_L)(1-\phi)^2(\bar{\beta}_{\lambda_L} - \alpha)^2 + \phi(1-\phi)(a_L(\lambda_L) + a_H(\lambda_A))(\bar{\beta}_{\lambda_A} - \alpha)(\bar{\beta}_{\lambda_L} - \alpha) + a_H(\lambda_A)\phi^2(\bar{\beta}_{\lambda_A} - \alpha)^2$. The disadvantaged agents choose an interior tax rate if $(\bar{\beta}_L - \alpha)^2 a_j(\lambda) < \overline{a(\lambda)}P$. Note that $\overline{a(\lambda)}$ is a weighted average of $a_L(\lambda_L)$ and $a_H(\lambda_H)$, where $\lambda_H > \lambda_L$, $\overline{a(\lambda)} > a_L(\lambda)$ (decreasing pessimism as a function of cost). Note also that $a_H > a_L$ and $(\bar{\beta}_{\lambda_H} - \alpha) > (\bar{\beta}_{\lambda_L} - \alpha)$. Then:

$$\begin{aligned}
&a_L(\lambda_L)(1-\phi)^2(\bar{\beta}_{\lambda_L} - \alpha)^2 + \phi(1-\phi)(a_L(\lambda_L) + \\
&a_H(\lambda_H))(\bar{\beta}_{\lambda_H} - \alpha)(\bar{\beta}_{\lambda_L} - \alpha) + a_H(\lambda_H)\phi^2(\bar{\beta}_{\lambda_H} - \alpha)^2 \\
&> a_L(\lambda_L)(1-\phi)^2(\bar{\beta}_{\lambda_L} - \alpha)^2 + \phi(1-\phi)(2a_L(\lambda_L))(\bar{\beta}_{\lambda_L} - \alpha)^2 + a_L(\lambda_L)\phi^2(\bar{\beta}_{\lambda_L} - \alpha)^2, \\
&> (\bar{\beta}_{\lambda_L} - \alpha)^2,
\end{aligned}$$

by Assumption 3. Then, the result follows because $P > (\bar{\beta} - \alpha)^2$. □

Proof of Lemma 5. As before, the problem of choosing λ in period 0 is:

$$\max_{\lambda \in (\underline{\lambda}, \bar{\lambda})} (1-T)\left[\alpha + (\bar{\beta}_\lambda - \alpha)\delta_H^*\right] - C_j(\delta_H^*) + T\bar{Y} + M(\lambda).$$

Or:

$$\max_{\lambda \in (\underline{\lambda}, \bar{\lambda})} F_j(\lambda, T),$$

where $F_j(\lambda, T) := (1-T)\alpha + \frac{1}{2}a_j(\lambda)(1-T)^2(\bar{\beta}_\lambda - \alpha)^2 + T\bar{Y} + M(\lambda)$.

Then:

$$\frac{\partial F_j(\lambda,T)}{\partial T} = -\alpha - a_j(\lambda)(1-T)(\bar{\beta}_\lambda-\alpha)^2 + \bar{Y} + T\frac{\partial \bar{Y}}{\partial T}.$$

$$\frac{\partial^2 F_j(\lambda,T)}{\partial T\,\partial\lambda} = \underbrace{-a_j'(\lambda)(1-T)(\bar{\beta}_\lambda-\alpha)^2}_{:=E} + \underbrace{(-2)a_j(\lambda)(1-T)(\bar{\beta}_\lambda-\alpha)(\beta_L-\beta_U)}_{:=F}$$

- $E > 0$ because $a'(\lambda) < 0$ by Assumption 3 and $1-T>0$ (interior solution).
- $F > 0$ because $\bar{\beta}_\lambda - \alpha > 0$.

Therefore, $F_j(\lambda,T)$ has increasing differences in (λ,T) and the result follows from monotone comparative statics. □

Proof of Lemma 6. It is sufficient to show that $\exists\ \bar{T} = \tau_L^*(\bar{\lambda}_L^*) \wedge\ \underline{T} = \tau_L^*(\underline{\lambda}_L^*)$ with $\bar{\lambda} = \bar{\lambda}_L^* > \underline{\lambda} = \underline{\lambda}_L^*$ such that:

$$F_L(\bar{\lambda}_L^*,\bar{T}) > F_L(\underline{\lambda}_L^*,\bar{T}), \tag{B.2a}$$

$$F_H(\bar{\lambda}_H^*,\bar{T}) > F_H(\underline{\lambda}_H^*,\bar{T}), \tag{B.2b}$$

$$F_L(\bar{\lambda}_L^*,\underline{T}) < F_L(\underline{\lambda}_L^*,\underline{T}), \tag{B.2c}$$

$$F_H(\bar{\lambda}_H^*,\underline{T}) < F_H(\underline{\lambda}_H^*,\underline{T}). \tag{B.2d}$$

Expressions (B.2a) and (B.2c) hold because L-types are optimizing and they are the pivotal group. We know that $F_H(\bar{\lambda},\bar{T}) > F_L(\bar{\lambda},\bar{T})$ because H-types have lower costs than L-types and $F_L(\bar{\lambda},\bar{T}) > F_L(\underline{\lambda},\bar{T})$ (equation B.2a). Using the H-types desired tax for $\bar{\lambda}_H^*$, $F_H(\bar{\lambda}_H^*,\bar{\tau}^*) > F_H(\underline{\lambda}_H^*,\bar{\tau}^*)$. Then, $F_H(\bar{\lambda}_H^*,\bar{T}) > F_H(\underline{\lambda}_H^*,\bar{T})$ because of Lemma 4 and because the L-types are the pivotal group. An analogue argument shows that $F_H(\bar{\lambda}_H^*,\underline{T} > F_H(\underline{\lambda}_H^*,\underline{T})$ holds. Then, the result follows by Lemma 5 because the solution to (B.2), $\lambda_j^*(T)$, is a continuous function of the equilibrium tax rate, T.

□

Proof of Proposition 1. We know that $\tau_L^*(\tilde{\lambda}) > \tau_L^*(\hat{\lambda})$ because $\underline{\lambda} < \hat{\lambda} < \tilde{\lambda} < \bar{\lambda}$. Take $\tilde{T} \in (\bar{T} = \tau_L^*(\bar{\lambda}), \underline{T} = \tau_L^*(\underline{\lambda}))$. Such a $\tilde{T}$ exists by continuity. Then, the result follows by Lemmas 2, 4 and 6. □

Online Appendix
(Not For Publication):

Redistributive Politics under Ambiguity

Javier D. Donna[§]

November 22, 2023

Contents (Appendix)

[§]University of Florida, Department of Economics, 224 Matherly Hall, P.O. Box 117140, Gainesville, FL 32611-7140. Phone: 352.392.0117. Fax: 352.392.7860. Email: jdonna@ufl.edu.

C Model with continuum of citizen types

C.1 Assumptions and setup

There is a contnuum of agents types, indexed by $j \in [0,1]$, with a a minority, $\phi < \frac{1}{2}$, of advantaged agents that have high effort cost parameter, a_H, and a majority, $1-\phi$, of disadvantaged agents that have a lower effort cost parameter, a_L.

Assumption 1C. *Let $\tau \in (0,1)$, $\underline{\beta} < \beta_L < \beta_U$, and $(1-\tau)\beta_L - C_j(\delta) < (1-\tau)\alpha - C_j(0) < (1-\tau)\beta_U - C_j(\delta)$, $\forall \delta \in [0,1], \forall \tau$.*

Assumption 2C. *Let $M(\lambda) > 0$ and $M'(\lambda) > 0$ be the cognitive technology through which the agents can manipulate their own degree of pessimism.*

Assumption 3C. *Let $C_j(\delta) = \frac{1}{2}\frac{\delta^2}{a_j + a(\lambda)}$, where $a(\lambda) > a_0, \forall \lambda \in [0,1], \frac{\partial a(\lambda)}{\partial \lambda} := a'(\lambda) < 0$, $a(\lambda)$ is convex, and agent's j cost is given by:*

$$a_j = \begin{cases} 1 + j a_L & j \leq \frac{1}{2}, \\ 1 + j a_H & j > \frac{1}{2}, \end{cases}$$

where $1 < a_L < a_H$.

As in the two-type case, agents estimate $\bar{Y}$, the average aggregate income, as:

$$\bar{Y} := \alpha + (1-\tau) \int_0^1 [a_j + a(\lambda_j)]\, dj \int_0^1 [\beta_{\lambda_j} - \alpha] d_j \int_0^1 \delta_j^* dj.$$

C.2 Results

Lemma 1C. *Let assumptions 1C and 3C hold. Then, agent j's optimal level of effort in period 1 is given by*

$$\delta^*_{H,j} = \begin{cases} 0 & \text{if } \delta_{h,j} \leq 1, \\ 1 & \text{if } \delta_{h,j} \geq 1, \\ \delta_{h,j} & \text{if } \delta_{h,j} \in (0,1), \end{cases} \tag{C.1}$$

where $\delta_{h,j} = (1-\tau)(a_j + a(\lambda_j))[\bar{\beta}_{\lambda_j} - \alpha]$ and $\bar{\beta}_{\lambda_j} := \lambda_j \beta_L + (1-\lambda_j)\beta_U$.

Proof. The optimal level of effort for agent j is the level that maximizes:

$$F(\delta_{H,j}) = (1-\tau)[\alpha + (\bar{\beta}_{\lambda_j} - \alpha)\delta_{H,j}] - C_j(\delta_{H,j}) + \tau\bar{Y} + M(\lambda).$$

By backwards induction, maximizing the objective function with respect to $\delta_{H,j}$ yields:c

$$\begin{aligned} 0 = \frac{\partial F(\cdot)}{\partial \delta_{H,j}} &= (1-\tau)(\bar{\beta}_{\lambda_j} - \alpha) - C_j'(\delta_{H,j}), \\ &= (1-\tau)(\bar{\beta}_{\lambda_j} - \alpha) - \frac{\delta_{H,j}}{a_j + a(\lambda_j)}, \\ c \implies \frac{\delta_{H,j}}{a_j + a(\lambda_j)} &= (1-\tau)(\bar{\beta}_{\lambda_j} - \alpha), \\ \implies \delta^*_{H,j} &= (1-\tau)(a_j + a(\lambda_j))(\bar{\beta}_{\lambda_j} - \alpha). \end{aligned}$$

As $\frac{\partial F(\cdot)}{\partial \delta_{H,j}} < 0$, we have that this is agent j's optimal effort allocation. Equation (C.1) follows from effort being restricted to between 0 and 1.

□

Lemma 2C. *Let assumptions 1C and 3C hold. Then, agent j's optimal level of effort $\delta^*_{H,j}$ decreases as the tax rate τ increases.*

Proof. By Lemma 1C, agent j's optimal effort allocation is given by equation (C.1). At an interior solution, the optimal allocation is $\delta^*_{H,j} = (1-\tau)(a_j + a(\lambda_j))(\bar{\beta}_{\lambda_j} - \alpha)$. Differentiating this quantity with respect to τ yields:

$$\frac{\partial \delta^*_{H,j}}{\partial \tau} = -(a_j + a(\lambda_j))(\bar{\beta}_{\lambda_j} - \alpha).$$

Then, we have that $(a_j + a(\lambda_j)) > 0$ by Assumption 3C, and $(\bar{\beta}_{\lambda_j} - \alpha) > 0$ by Assumption 1C. So the agent's effort choice is decreasing in the tax rate if effort is fractional. The corner cases are trivial. Then, $\frac{\partial \delta^*_{H,j}}{\partial \tau} \leq 0$. □

Lemma 3C. *Let assumptions 1C through 3C hold. Then, agent j's ideal tax rate is*

$$\tau_j^*(\lambda) = \begin{cases} \tau^* & \textit{if} \quad \tau^* \in (0,1), \\ 0 & \textit{otherwise}, \end{cases} \tag{C.2}$$

where:

$$\tau^* = \frac{1}{3} - \frac{(\bar{\beta}_{\lambda_j} - \alpha)^2(a_j + a(\lambda_j))}{3\int_0^1 [a_j + a(\lambda_j)]\,dj \int_0^1 [\beta_{\lambda_j} - \alpha] dj \int_0^1 \left[(a_j + a(\lambda_j)(\bar{\beta}_{\lambda_j} - \alpha))\right] dj}.$$

Proof. The agent's objective is to maximize:

$$U(\cdot) = (1-\tau)(\alpha + (\bar{\beta}_\lambda - \alpha)\delta) - C(\delta) + \tau\bar{Y} + M(\lambda).$$

Let $a_j(\lambda) := a_j + a(\lambda)$. Taking the derivative with respect to τ yields:

$$\begin{aligned}
\frac{\partial U(\cdot)}{\partial \tau} &= \frac{\partial}{\partial \tau}[(1-\tau)(\alpha + (\bar{\beta}_{\lambda_j} - \alpha)\delta^*_{H,j}(\tau)] - \frac{\partial}{\partial \tau}C(\delta^*_{H,j}(\tau)) + \frac{\partial}{\partial \tau}[\tau\bar{Y}] + 0, \\
&= -\alpha + (\bar{\beta}_{\lambda_j} - \alpha)[-\delta^*_{H,j}(\tau) - (1-\tau)a_j(\lambda)(\bar{\beta}_{\lambda_j} - \alpha)] + \delta^*_{H,j}(\tau)(\bar{\beta}_{\lambda_j} - \alpha) + \bar{Y} + \tau\frac{\partial \bar{Y}}{\partial \tau}, \\
&= -\alpha - (\bar{\beta}_{\lambda_j} - \alpha)^2(1-\tau)a_j(\lambda) + \bar{Y} + \tau\frac{\partial \bar{Y}}{\partial \tau}.
\end{aligned}$$

Let $A := \int_0^1 [a_j + a(\lambda_j)]\, d_j$, $B := \int_0^1 [\beta_{\lambda_j} - \alpha] d_j$ and $C := \int_0^1 \left[(a_j + a(\lambda_j)(\bar{\beta}_{\lambda_j} - \alpha))\right] dj$. Note that $\bar{Y} = \alpha + (1-\tau)^2 ABC$. Then:

$$\frac{\partial \bar{Y}}{\partial \tau} = -2(1-\tau)ABC + (1-\tau)^2 BC\frac{\partial A}{\partial \tau} + (1-\tau)^2 AC\frac{\partial B}{\partial \tau} + (1-\tau)^2 AB\frac{\partial C}{\partial \tau}.$$

By construction, $\frac{\partial A}{\partial \tau} = \frac{\partial B}{\partial \tau} = \frac{\partial C}{\partial \tau} = 0$. Then:

$$\frac{\partial \bar{Y}}{\partial \tau} = -2(1-\tau)ABC.$$

Solving for τ yields:

$$\begin{aligned}
\frac{\partial U(\cdot)}{\partial \tau} &= -\alpha - (\bar{\beta}_{\lambda_j} - \alpha)^2(1-\tau)a_j(\lambda) + \bar{Y} + \tau\frac{\partial \bar{Y}}{\partial \tau}, \\
&= -\alpha - (\bar{\beta}_{\lambda_j} - \alpha)^2(1-\tau)a_j(\lambda) + \bar{Y} - 2\tau(1-\tau)ABC, \\
&= -\alpha - (\bar{\beta}_{\lambda_j} - \alpha)^2(1-\tau)a_j(\lambda) + \alpha + (1-\tau)^2 ABC - 2\tau(1-\tau)ABC, \\
&= -(\bar{\beta}_{\lambda_j} - \alpha)^2(1-\tau)a_j(\lambda) + (1-\tau)^2 ABC - 2\tau(1-\tau)ABC, \\
\implies 0 &= -(\bar{\beta}_{\lambda_j} - \alpha)^2 a_j(\lambda) + (1-\tau)ABC - 2\tau ABC, \\
\implies 3\tau ABC &= -(\bar{\beta}_{\lambda_j} - \alpha)^2 a_j(\lambda) + ABC, \\
\implies \tau^* &= \frac{ABC - (\bar{\beta}_{\lambda_j} - \alpha)^2 a_j(\lambda)}{3ABC}, \\
&= \frac{1}{3} - \frac{(\bar{\beta}_{\lambda_j} - \alpha)^2 a_j(\lambda)}{3ABC}, \\
&= \frac{1}{3} - \frac{(\bar{\beta}_{\lambda_j} - \alpha)^2 a_j(\lambda)}{3\int_0^1 [a_j + a(\lambda_j)]\, d_j \int_0^1 [\beta_{\lambda_j} - \alpha] d_j \int_0^1 \left[(a_j + a(\lambda_j)(\bar{\beta}_{\lambda_j} - \alpha))\right] dj}.
\end{aligned}$$

□

Lemma 4C. *Let assumptions 1C through 3C hold and let $\lambda \in (\underline{\lambda}, \overline{\lambda})$. Then, each agent's preferred tax rate is increasing in λ, the pessimism rate.*

Proof. Let $j \in [0,1]$. If $\tau_j^* = 0$ for any pessimism level, then the claim holds trivially. So assume that at least for some pessimism levels the desired tax rate is interior. Note that $\int_0^1 [a_n + a(\lambda_n)]dn \int_0^1 [\beta_{\lambda_n} - \alpha]dn \int_0^1 \left[(a_n + a(\lambda_n)(\bar{\beta}_{\lambda_n} - \alpha))\right] dn$ does not depend on λ_j nor a_n. Call this product D and note that by construction, $D \geq 0$.

$$\begin{aligned}\frac{\partial \tau_j^*}{\partial \lambda_j} &= -\frac{1}{D}\frac{\partial}{\partial \lambda_j}\left[(\bar{\beta}_{\lambda_j} - \alpha)^2(a_j + a(\lambda_j))\right], \\ &= -\frac{1}{D}\left[\underbrace{2(a_j + a(\lambda_j))(\bar{\beta}_{\lambda_j} - \alpha)(\beta_L - \beta_U)}_{:=E_1} + \underbrace{a'(\lambda_j)(\bar{\beta}_{\lambda_j} - \alpha)^2}_{:=E_2}\right].\end{aligned}$$

Note that:

- $E_1 < 0$ because $a_j + a(\lambda_j) > 0$ by Assumption 3C, $\bar{\beta}_{\lambda_j} - \alpha > 0$ by Assumption 1 and $\lambda \in (\underline{\lambda}, \overline{\lambda})$, and $\beta_L - \beta_U < 0$ as $\beta_U > \beta_L$.
- $E_1 < 0$ because $(\bar{\beta}_{\lambda_j} - \alpha)^2 > 0$ and $a'(\lambda_j) < 0$ by Assumption 3C.

As in the two-type case, note that pessimism is decreasing in agents' cost parameters. In period 0, agent j faces the problem of choosing λ given by:

$$\max_{\lambda \in (\underline{\lambda}, \overline{\lambda})} (1-T)\alpha + \frac{1}{2}a_j(\lambda)(1-T)^2(\bar{\beta}_\lambda - \alpha)^2 + T\bar{Y} + M(\lambda).$$

Let $\Lambda = -\lambda$. Then:

$$\max_{\lambda \in (\underline{\lambda}, \overline{\lambda})} (1-T)\alpha + \frac{1}{2}(a_j + a(-\Lambda))(1-T)^2(\bar{\gamma}_\Lambda - \alpha)^2 + T\bar{Y} + M(-\Lambda),$$

where $\bar{\gamma}_\Lambda = -\beta_L\Lambda + (1+\Lambda)\beta_U = \beta_L\lambda + (1-\lambda)\beta_U = \bar{\beta}_\lambda$.

Denote the objective function as $F(\cdot)$ and differentiate it with respect to a_j, and then with respect to Λ.

$$\frac{\partial F(\cdot)}{\partial a_j} = \frac{1}{2}(1-T)^2(\bar{\gamma}_\Lambda - \alpha)^2.$$

$$\frac{\partial^2 F(\cdot)}{\partial \Lambda \partial a_j} = 2(1-T)(\bar{\gamma}_\Lambda - \alpha)^2\frac{\partial T}{\partial \lambda} + 2(1-T)\frac{\delta_j}{a_j(\lambda)}(\beta_U - \beta_L).$$

The first term is zero because $\frac{\partial T}{\partial \lambda} = 0$. The second term is positive as $\delta_j \in (0,1)$. As before, $J_j(\Lambda, a_j)$ has increasing differences in (Λ, a_j) and the result is established using monotone comparative statics.

The preferred tax rate of the disadvantaged agents is strictly increasing with the pessimism rate, so the disadvantaged agents prefer an interior tax rate. To see this, note that the median agent, agent $j = 0.5$, has interior preferred tax rate if, and only if:

$$(\bar{\beta}_{\lambda_{0.5}} - \alpha)^2(a_{0.5} + a(\lambda_{0.5})) < F,$$

where

$$F := \int_0^1 [a_j + a(\lambda_j)]\, d_j \int_0^1 [\beta_{\lambda_j} - \alpha] d_j \int_0^1 \left[(a_j + a(\lambda_j)(\bar{\beta}_{\lambda_j} - \alpha))\right] dj.$$

We proceed in two steps. The first is to note that $a_{0.5} + a(\lambda_{0.5}) < \int_0^1 (a_j + a(\lambda_j))dj$. To that end note that:

$$\begin{aligned}
\int_0^1 (a_j + a(\lambda_j))dj &= 1 + 0.5a_L + 0.5a_H + \int_0^1 a(\lambda_j)dj, \\
&> 1 + 0.5a_L + 0.5a_H + \int_0^1 a(\lambda_j)dj, \\
&= a_{0.5} + 0.5a_H + \int_0^1 a(\lambda_j)dj.
\end{aligned}$$

Fix $\gamma > 0.5$. Then, $\lambda_{0.5} < t\lambda_0 + (1-t)\lambda_\gamma$ for some $t > \gamma$ as shown above. Then, by convexity $a(\lambda_{0.5}) < ta(\lambda_0) + (1-t)a(\lambda_\gamma)$. Next, note that $\int_0^1 a(\lambda_j)dj > a(\lambda_{0.5})$:

$$\begin{aligned}
\int_0^1 a(\lambda_j)dj &= \int_0^t a(\lambda_j)dj + \int_t^1 a(\lambda_j)dj, \\
&\geq \int_0^t a(\lambda_0)dj + \int_t^1 a(\lambda_\gamma)dj, \\
&= ta(\lambda_0) + (1-t)a(\lambda_\gamma), \\
&> a(\lambda_{0.5}).
\end{aligned}$$

which concludes the first step.

The second is that $(\bar{\beta}_{\lambda_{0.5}} - \alpha)^2 < \int_0^1 [\beta_{\lambda_j} - \alpha] d_j \int_0^1 \left[(a_j + a(\lambda_j)(\bar{\beta}_{\lambda_j} - \alpha))\right] dj$. First, note that $\int_0^1 \left[(a_j + a(\lambda_j)(\bar{\beta}_{\lambda_j} - \alpha)\right] dj \geq 2\int_0^1 (\bar{\beta}_{\lambda_j} - \alpha)dj$. It is sufficient to show that $2[\int_0^1 (\bar{\beta}_{\lambda_j} - \alpha)dj]^2 \geq (\bar{\beta}_{\lambda_{0.5}} - \alpha)^2$. Because $\lambda \in (\underline{\lambda}, \overline{\lambda})$, the condition is equivalent to $(+\sqrt{2})\int_0^1 (\bar{\beta}_{\lambda_j} - \alpha)dj > \bar{\beta}_{\lambda_{0.5}} - \alpha$, which in turn is equivalent to $(+\sqrt{2})\int_0^1 \bar{\beta}_{\lambda_j} dj > \bar{\beta}_{\lambda_j}$. Then:

$$\begin{aligned}
(+\sqrt{2})\int_0^1 \bar{\beta}_{\lambda_j} dj &\geq \frac{(+\sqrt{2})}{2}\bar{\beta}_{\lambda_0} + \frac{(+\sqrt{2})}{2}\bar{\beta}_{\lambda_{0.5}} > \bar{\beta}_{\lambda_{0.5}}, \\
\iff \frac{(+\sqrt{2})}{2}\bar{\beta}_{\lambda_0} &> \bar{\beta}_{\lambda_{0.5}} - \frac{(+\sqrt{2})}{2}\bar{\beta}_{\lambda_{0.5}} = (\frac{2 - (+\sqrt{2})}{2})\bar{\beta}_{\lambda_{0.5}}, \\
\iff \bar{\beta}_{\lambda_0} &> (\frac{2 - (+\sqrt{2})}{(+\sqrt{2})})\bar{\beta}_{\lambda_{0.5}} = ((+\sqrt{2}) - 1)\bar{\beta}_{\lambda_{0.5}}.
\end{aligned}$$

The left hand side is bounded from below by β_L. The right hand side is bounded from above by $((+\sqrt{2})-1)\beta_U$, and $\beta_L > \underline{\beta} := 0.5\beta_U > ((+\sqrt{2})-1)\beta_U$ by Assumption 1C. □

Lemma 5C. *Let Assumptions 1C through 3C hold and $\lambda \in (\underline{\lambda}, \bar{\lambda})$. Then, the optimal pessimism rate, λ^*, increases with the equilibrium tax rate, T.*

Proof. As before, the problem of choosing λ in period 0 is:

$$\max_{\lambda \in (\underline{\lambda}, \bar{\lambda})} (1-T)\left[\alpha + (\bar{\beta}_\lambda - \alpha)\delta^*_{H,j}\right] - C_j(\delta^*_{H,j}) + T\bar{Y} + M(\lambda).$$

Or:

$$\max_{\lambda \in (\underline{\lambda}, \bar{\lambda})} F(\lambda, T),$$

where $F(\lambda, T) := (1-T)\alpha + \frac{1}{2}[a_j + a(\lambda)](1-T)^2(\bar{\beta}_\lambda - \alpha)^2 + T\bar{Y} + M(\lambda)$.
Then:

$$\frac{\partial F(\lambda, T)}{\partial T} = -\alpha - [a_j + a(\lambda)](1-T)(\bar{\beta}_\lambda - \alpha)^2 + \bar{Y} + T\frac{\partial \bar{Y}}{\partial T}.$$

$$\begin{aligned}\frac{\partial^2 F(\lambda, T)}{\partial T\,\partial \lambda} &= -a'(\lambda)(1-T)(\bar{\beta}_\lambda - \alpha)^2 + (-2)[a_j + a(\lambda)](1-T)(\bar{\beta}_\lambda - \alpha)(\beta_L - \beta_U)\\ &+ \frac{\partial \bar{Y}}{\partial T}\frac{\partial T}{\partial \lambda} + T\frac{\partial^2 \bar{Y}}{\partial \lambda\,\partial T},\\ &= \underbrace{-a'(\lambda)(1-T)(\bar{\beta}_\lambda - \alpha)^2}_{:=G} + \underbrace{(-2)[a_j + a(\lambda)](1-T)(\bar{\beta}_\lambda - \alpha)(\beta_L - \beta_U)}_{:=H}.\end{aligned}$$

Note that:

- $G > 0$ because $a'(\lambda) < 0$ by Assumption 3C and $1 - T > 0$ (interior solution).
- $H > 0$ because $\bar{\beta}_\lambda - \alpha > 0$ due to $\lambda \in (\underline{\lambda}, \overline{\lambda})$ and $\beta_L < \beta_U$.

Therefore, $F(\lambda, T)$ has increasing differences in (λ, T) and the result follows from monotone comparative statics. □

Lemma 6C. *Let Assumptions 1C through 3C hold and $\lambda \in \{\hat{\lambda}, \tilde{\lambda}\}$. Then, the equilibrium pessimism rate, λ^*, has the following form:*

$$\lambda^* = \begin{cases} \hat{\lambda} & \text{if } T < \tilde{T}, \\ \tilde{\lambda} & \text{if } T > \tilde{T}. \end{cases}$$

Proof. Let $L \in [0, 0.5)$ and $H \in (0.5, 1]$ be arbitrary individuals from the disadvantaged and advantaged populations, respectively. It is sufficient to show that there exist $\bar{T} = \tau_L^*(\bar{\lambda}_L^*) \wedge \underline{T} = \tau_L^*(\underline{\lambda}_L^*)$ with $\bar{\lambda} = \bar{\lambda}_L^* > \underline{\lambda} = \underline{\lambda}_L^*$ such that:

$$F_L(\bar{\lambda}_L^*, \bar{T}) > F_L(\underline{\lambda}_L^*, \bar{T}), \tag{C.3a}$$
$$F_H(\bar{\lambda}_H^*, \bar{T}) > F_H(\underline{\lambda}_H^*, \bar{T}), \tag{C.3b}$$
$$F_L(\bar{\lambda}_L^*, \underline{T}) < F_L(\underline{\lambda}_L^*, \underline{T}), \tag{C.3c}$$
$$F_H(\bar{\lambda}_H^*, \underline{T}) < F_H(\underline{\lambda}_H^*, \underline{T}). \tag{C.3d}$$

Note that by construction, the additional equations hold, $F_{0.5}(\bar{\lambda}_L^*, \bar{T}) > F_{0.5}(\underline{\lambda}_L^*, \bar{T})$ and $F_{0.5}(\bar{\lambda}_L^*, \underline{T}) < F_{0.5}(\underline{\lambda}_L^*, \underline{T})$. First, we show that expressions (C.3a) and (C.3c) hold. Note that agent L has desired pessimism higher than the pivotal agent, $j = 0.5$, as shown in the proof of Lemma 4C. Additionally, the tax rates and pessimism are positively associated by Lemma 5C.

We know that $F_H(\bar{\lambda}, \bar{T}) > F_{0.5}(\bar{\lambda}, \bar{T})$ because agent H has lower costs than agent $j = 0.5$ and $F_{0.5}(\bar{\lambda}, \bar{T}) > F_{0.5}(\underline{\lambda}, \bar{T})$. Using agent H's desired tax for $\bar{\lambda}_H^*$, $F_H(\bar{\lambda}_H^*, \bar{\tau}^*) > F_H(\underline{\lambda}_H^*, \bar{\tau}^*)$. Then, $F_H(\overline{\lambda}_H^*, \bar{T}) > F_H(\underline{\lambda}_H^*, \bar{T})$ because of Lemma 3C and because agent 0.5 is pivotal.

An analogous argument shows that $F_H(\bar{\lambda}_H^*, \underline{T}) > F_H(\underline{\lambda}_H^*, \underline{T})$. Because L and H were arbitrarily picked, the result holds. □

Proposition 1C. *Let Assumptions 1C through 3C hold, $\lambda \in \{\hat{\lambda}, \tilde{\lambda}\}$ and $\tau \in [0, 1)$. Then, the following two equilibria are possible:*

(a) Pessimism/Welfare State*: Agents are pessimistic, $\lambda^* = \tilde{\lambda}$, they impose a high tax rate, $T = \tau_{0.5}^*(\tilde{\lambda})$, and they exert low effort, $\delta_H^*(\tilde{\lambda}, j)$.*

(b) Optimism/Laissez Faire*: Agents are optimistic, $\lambda^* = \hat{\lambda}$, they impose a low tax rate, $T = \tau_{0.5}^*(\hat{\lambda})$, and they exert high effort, $\delta_H^*(\hat{\lambda}, j)$.*

with $\tilde{\lambda} > \hat{\lambda}$, $\tau_{0.5}^*(\tilde{\lambda}) > \tau_{0.5}^*(\hat{\lambda})$, and $\delta_H^*(\tilde{\lambda}) < \delta_H^*(\hat{\lambda})$.

Proof. As before, we know that $\tau_L^*(\tilde{\lambda}) > \tau_L^*(\hat{\lambda})$ because $\underline{\lambda} < \hat{\lambda} < \tilde{\lambda} < \bar{\lambda}$. Take $\tilde{T} \in (\bar{T} = \tau_L^*(\bar{\lambda}), \underline{T} = \tau_L^*(\underline{\lambda}))$. Such a $\tilde{T}$ exists by continuity. Then, the result follows by Lemmas 2C, 4C and 6C. □

D Bayes Rule

D.1 Assumptions and setup

Assumption 1B. *Let* $\tau \in (0,1)$ *and* $(1-\tau)\beta_L - C_j(\delta) < (1-\tau)\alpha - C_j(0) < (1-\tau)\beta_U - C_j(\delta), \quad \forall \delta \in [0,1],\ \forall \tau.$

Assumption 2B. *Let* $M(\lambda) > 0$ *and* $M'(\lambda) > 0$ *be the cognitive technology through which the agents can manipulate their own degree of pessimism.*

Assumption 3B. $C_j(\delta) = \frac{1}{2}\frac{\delta^2}{a_j + a(\lambda)}$, $a(\lambda) > a_L, \forall \lambda \in [0,1]$, $1 < a_L < a_H$, $\frac{\partial a(\lambda)}{\partial \lambda} := a'(\lambda) < 0$.

The Bayes Rule shows that the maximization problem takes a standard form using a subjective (alternatively, known) distribution for the true return to effort. A Bayesian decision-maker places a subjective distribution on the ambiguous parameter and maximizes the subjective expected utility function. Let π be a specific probability distribution over β. I follow the same steps as with the Hurwicz criterion and omit the proofs that use the same arguments. The Bayesian solves the following optimization problem:

$$\begin{aligned} U_j(\delta_B^*) &= \max_{\delta \in [0,1]} \mathbb{E}_\pi[(1-\tau)y(e_j)] - C_j(\delta), \\ &= \max_{\delta \in [0,1]} (1-\tau)\Big[\alpha + \delta[\mathbb{E}_\pi(\beta) - \alpha]\Big] - C_j(\delta). \end{aligned}$$

D.2 Results

Lemma 1B. *The Bayesian solution is:*

$$\delta_B^* = \begin{cases} 0 & \text{if } \delta_B \leq 0, \\ 1 & \text{if } \delta_B \geq 1, \\ \delta_B & \text{if } \delta_B \in (0,1). \end{cases}$$

where $\delta_B = (a_j + a(\lambda))(1-\tau)[\mathbb{E}_\pi(\beta) - \alpha]$.

Proof. The first-order necessary condition for an interior solution is:

$$\begin{aligned} (1-\tau)[\mathbb{E}_\pi(\beta) - \alpha] &= C_j'(\delta_B), & \text{where } C_j'(\delta) &:= \frac{\partial C_j(\delta)}{\partial \delta}, \\ &= \frac{\delta_B}{a_j + a(\lambda)}, & \text{provided } C_j(\delta) &= \frac{\delta^2}{2(a_j + a(\lambda))}. \end{aligned}$$

Solving for δ_B yields $\delta_B = (a_j + a(\lambda))(1-\tau)[\mathbb{E}_\pi(\beta) - \alpha]$ and, hence, the Bayesian solution is:

$$\delta_B^* = \begin{cases} 0 & \text{if } \delta_B \leq 0, \\ 1 & \text{if } \delta_B \geq 1, \\ \delta_B & \text{if } \delta_B \in (0,1). \end{cases}$$

The second-order sufficient condition is satisfied using the same argument as in the Hurwicz criterion.

□

Because the subjective distribution, π, determines the subjective mean, $\mathbb{E}_\pi(\beta)$, the treatment choice resulting from the Bayesian criteria depends critically on this subjective distribution. According to Bayesian theorists, π should reflect the agent's personal beliefs about whether effort pays off or not; that is, where β lies in the interval $[\beta_L, \beta_U]$.

Lemma 2B. *Let Assumptions 1B and 3B hold. Then, the optimal level of effort decreases as the tax rate increases.*

Proof. The result follows from Lemma 1B.

□

Lemma 3B. *Let Assumptions 1B through 3B hold. Then, the ideal tax rate of agent j is:*

$$\tau_j^*(\lambda) = \begin{cases} \tau_j^* & \text{if} \quad \tau^* > \underline{\tau}, \\ \underline{\tau} & \text{otherwise}, \end{cases} \tag{D.1}$$

where $\tau_j^* := \frac{a_j + a(\lambda_j) - [(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))]}{a_j + a(\lambda_j) - 2[(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))]}$ *and* $\underline{\tau}$ *is the lowest possible tax* (e.g., $\underline{\tau} = 0$ *if* $\tau \in [0,1)$).

Proof. The agent's welfare under Bayes Rule is given by:

$$U_j^B(\lambda, \delta, \tau) = (1-\tau)\Big[\alpha + [\mathbb{E}_\pi(\beta) - \alpha]\Big] - C_j(\delta) + \tau\bar{Y} + M(\lambda),$$

where, $\bar{Y} = \alpha + (1-\tau)[\mathbb{E}_\pi(\beta) - \alpha]^2[(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))]$.

Then:

$$
\begin{aligned}
\frac{\partial U_j^B}{\partial \tau} &= \frac{\partial}{\partial \tau}\left[(1-\tau)\left[\alpha + \delta[\mathbb{E}_\pi(\beta) - \alpha]\right] - C_j(\delta)\right] + \frac{\partial}{\partial \tau}[\tau \bar{Y}] + \frac{\partial}{\partial \tau}[M(\lambda)], \\
&= \frac{\partial}{\partial \tau}\left[(1-\tau)\alpha + \frac{1}{2}(a_j + a(\lambda_j))(1-\tau)^2[\mathbb{E}_\pi(\beta) - \alpha]^2\right] + \bar{Y} + \tau\frac{\partial \bar{Y}}{\partial \tau}, \\
&= -\alpha - (a_j + a(\lambda_j))(1-\tau)(\mathbb{E}_\pi(\beta) - \alpha)^2 \\
&\qquad + \bar{Y} - \tau[\mathbb{E}_\pi(\beta) - \alpha]^2[(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))], \\
&= -(a_j + a(\lambda))(1-\tau)(\mathbb{E}_\pi(\beta) - \alpha)^2 \\
&\qquad + (1-2\tau)[\mathbb{E}_\pi(\beta) - \alpha]^2[(1-\phi)(a_L + a(\lambda_j)) + \phi(a_H + a(\lambda_H))].
\end{aligned}
$$

In an interior solution:

$$
\begin{aligned}
a_j + a(\lambda_j) - [(1-\phi)(a_L + a(\lambda_j)) + \phi(a_H + a(\lambda_H))] = t(a_j + a(\lambda_j)) \\
- 2t[(1-\phi)(a_L + a(\lambda_j)) + \phi(a_H + a(\lambda_H))]
\end{aligned}
$$

Then:

$$
\tau^* = \frac{a_j + a(\lambda_j) - [(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))]}{a_j + a(\lambda_j) - 2[(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))]}.
$$

We now turn our attention to the median voter. Let $X_j := a_j + a(\lambda_j)$ and let $X_{j/k} := \frac{X_j}{X_k}$. Them for the median voter:

$$
\begin{aligned}
\tau_L^* &= \frac{X_L - (1-\phi)X_L - \phi X_H}{X_L - 2(1-\phi)X_L - 2\phi X_H}, \\
&= \frac{1 - (1-\phi) - \phi X_{H/L}}{1 - 2(1-\phi) - 2\phi X_{H/L}}, \\
&= \frac{\phi - \phi X_{H/L}}{2\phi - 1 - 2\phi X_{H/L}}, \\
&= \phi\frac{1 - X_{H/L}}{2\phi - 1 - 2\phi X_{H/L}}.
\end{aligned}
$$

Then, $X_{H/L} > 1$ because $X_H > X_L$. So the desired tax rate is positive. Next, we show that it is less than 1:

$$
\begin{aligned}
\tau_L^* &= \phi\frac{1 - X_{H/L}}{2\phi - 1 - 2\phi X_{H/L}}, \\
&= \phi\frac{1 - X_{H/L}}{2\phi(1 - X_{H/L}) - 1}, \\
&= \phi\frac{1}{2\phi - \frac{1}{1 - X_{H/L}}}.
\end{aligned}
$$

Here, $2\phi > 1$ by construction and $-\frac{1}{1-X_{H/L}} > 0$, so the fraction is less than 1.

We now show that the second-order condition holds. The second-order condition holds if:

$$\begin{aligned}\frac{\partial^2 U_j^B}{\partial \tau^2} &= (a_j + a(\lambda_j))(\mathbb{E}_\pi(\beta) - \alpha)^2 \\ &\qquad - 2[\mathbb{E}_\pi(\beta) - \alpha]^2[(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))], \\ &< 0.\end{aligned}$$

Note that:

$$\begin{aligned}0 &> (a_j + a(\lambda_j)) - 2[(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))], \\ &= (a_L + a(\lambda_L)) - 2[(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))], \\ &= -1 + 2\phi(a_L + a(\lambda_L)) - 2\phi(a_H + a(\lambda_H))], \\ &= -1 + 2\phi[(a_L + a(\lambda_L) - (a_H + a(\lambda_H)), \\ &= -1 + 2\underbrace{[X_L - X_H]}_{<0}.\end{aligned}$$

□

Lemma 4B. *Let Assumptions 1B through 3B hold and $\lambda \in (\underline{\lambda}, \bar{\lambda})$. Then, the equilibrium tax outcome, $T = \tau_L^*(\lambda)$, is strictly increasing with the pessimism rate, λ.*

Proof. By Lemma 3B, the equilibrium tax rate is:

$$\tau_j^* = \frac{a_j + a(\lambda_j) - [(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))]}{a_j + a(\lambda_j) - 2[(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))]}.$$

where j is the median voter.

Let $P := [(1-\phi)(a_L + a(\lambda_L)) + \phi(a_H + a(\lambda_H))]$ and let $D := a_j + a(\lambda_j) - 2P$. Note that $D < 0$ and $\frac{\partial D}{\partial \lambda_j} = a'(\lambda_j)$. Then, the equilibrium tax rate is $\tau_j^* = \frac{a_j + a(\lambda_j) - P}{D}$.

$$
\begin{aligned}
\frac{\partial \tau_j^*}{\partial \lambda_j} &= \frac{\partial}{\partial \lambda_j}\left[\frac{a_j + a(\lambda_j) - P}{D}\right], \\
&= \frac{(P - a_j - a(\lambda_j))a'(\lambda_j)}{D^2} + \frac{a'(\lambda_j)}{D}, \\
&= \frac{(P - a_j - a(\lambda_j))a'(\lambda_j) + a'(\lambda_j)D}{D^2}, \\
&= \frac{(P - a_j - a(\lambda_j))a'(\lambda_j) + (a_j + a(\lambda_j))a'(\lambda_j) - 2Pa'(\lambda_j)}{D^2}, \\
&= \frac{Pa'(\lambda_j) - 2Pa'(\lambda_j)}{D^2}, \\
&= P \times \frac{-a'(\lambda_j)}{D^2}, \\
&> 0.
\end{aligned}
$$

where the last inequality follows from Assumption 3B. □

Lemma 5B. *Let Assumptions 1B through 3B hold and* $\lambda \in (\underline{\lambda}, \bar{\lambda})$. *Then, the optimal pessimism rate,* λ^*, *increases with the equilibrium tax rate,* T.

Proof. Agent j's problem of choosing λ in period 0 is:

$$
\max_{\lambda \in (\underline{\lambda}, \bar{\lambda})} (1 - T)\left[\alpha + [\mathbb{E}_\pi(\beta) - \alpha]\delta_H^*\right] - C_j(\delta_H^*) + T\bar{Y} + M(\lambda).
$$

Or:

$$
\max_{\lambda \in (\underline{\lambda}, \bar{\lambda})} F_j(\lambda, T),
$$

where $F_j(\lambda, T) := (1 - T)\alpha + \frac{1}{2}a_j(\lambda)(1 - T)^2(\mathbb{E}_\pi(\beta) - \alpha)^2 + T\bar{Y} + M(\lambda)$.

Then:

$$
\frac{\partial F_j(\lambda, T)}{\partial T} = -\alpha - a_j(\lambda)(1 - T)(\bar{\beta}_\lambda - \alpha)^2 + \bar{Y} + T\frac{\partial \bar{Y}}{\partial T}.
$$

$$
\begin{aligned}
\frac{\partial^2 F_j(\lambda, T)}{\partial T\, \partial \lambda} &= -a_j'(\lambda)(1 - T)(\mathbb{E}_\pi(\beta) - \alpha)^2, \\
&> 0,
\end{aligned}
$$

where the inequality follows from $a'(\lambda) < 0$ by Assumption 3B and $1 - T > 0$ (interior solution).

Therefore, $J_j(\lambda, T)$ has increasing differences in (λ, T) and the result is established using monotone comparative statics.

□

Lemma 6B. *Let Assumptions 1B through 3B hold and $\lambda \in \{\hat{\lambda}, \tilde{\lambda}\}$. Then, the equilibrium pessimism rate, λ^*, has the following form:*

$$\lambda^* = \begin{cases} \hat{\lambda} & \text{if } T < \tilde{T}, \\ \tilde{\lambda} & \text{if } T > \tilde{T}. \end{cases}$$

Proof. It is sufficient to show that $\exists\ \bar{T} = \tau_L^*(\bar{\lambda}) \wedge\ \underline{T} = \tau_L^*(\underline{\lambda})$ with $\bar{\lambda} > \underline{\lambda}$ such that:

$$F_L(\bar{\lambda}, \bar{T}) > F_L(\underline{\lambda}, \bar{T}), \tag{D.2a}$$
$$F_H(\bar{\lambda}, \bar{T}) > F_H(\underline{\lambda}, \bar{T}), \tag{D.2b}$$
$$F_L(\bar{\lambda}, \underline{T}) < F_L(\underline{\lambda}, \underline{T}), \tag{D.2c}$$
$$F_H(\bar{\lambda}, \underline{T}) < F_H(\underline{\lambda}, \underline{T}). \tag{D.2d}$$

Expressions (D.2a) and (D.2c) hold because L-types are optimizing and they are the pivotal group. We know that $F_H(\bar{\lambda}, \bar{T}) > F_L(\bar{\lambda}, \bar{T})$ because H-types have lower costs than L-types, and $F_L(\bar{\lambda}, \bar{T}) > F_L(\underline{\lambda}, \bar{T})$ by (D.2a). Evaluating $\Delta_{H,\bar{T}} := F_H(F < \bar{\lambda}, \bar{T}) - F_H(\underline{\lambda}, \bar{T})$ yields:

$$\Delta_{H,\bar{T}} = \underbrace{\left[M(\bar{\lambda}) - M(\underline{\lambda})\right]}_{:=A} + \underbrace{(1-\bar{T})(\mathbb{E}_\pi(\beta) - \alpha)\left[\overline{\delta_H} - \underline{\delta_H}\right] + [C_H(\underline{\delta^*}) - C_H(\overline{\delta^*})]}_{:=B}.$$

Note that:

- $A > 0$ because $M'(\lambda) > 0$ (Assumption 2B) and $\overline{\lambda} > \underline{\lambda}$.
- Regarding B, first, observe that:

$$\overline{\delta_H} - \underline{\delta_H} = (1-\bar{T})(\mathbb{E}_\pi(\beta) - \alpha)(a(\overline{\lambda}) - a(\underline{\lambda})),$$
$$C_H(\underline{\delta^*}) - C_H(\overline{\delta^*}) = \frac{1}{2}(1-\bar{T})^2[\mathbb{E}_\pi(\beta) - \alpha]^2[a(\underline{\lambda}) - a(\overline{\lambda})].$$

As such, B simplifies to:

$$\frac{1}{2}(1-\bar{T})^2[\mathbb{E}_\pi(\beta) - \alpha]^2[a(\overline{\lambda}) - a(\underline{\lambda})] < 0.$$

To compare A and B, define $\Delta_{L,\bar{T}} := F_L(\bar{\lambda}, \bar{T}) - F_L(\underline{\lambda}, \bar{T})$. Then:

$$\Delta_{L,\bar{T}} = \underbrace{\left[M(\bar{\lambda}) - M(\underline{\lambda})\right]}_{A_L} + \underbrace{(1-\bar{T})(\mathbb{E}_\pi(\beta) - \alpha)\left[\overline{\delta_L} - \underline{\delta_L}\right] + [C_L(\underline{\delta^*}) - C_L(\overline{\delta^*})]}_{B_L}.$$

Note that $A_L > 0$ as before. Regarding B_L, we have that $\overline{\delta_L} - \underline{\delta_L} = (1 - \bar{T})(\mathbb{E}_\pi(\beta) - \alpha)(a(\overline{\lambda}) - a(\underline{\lambda}))$ and $C_L(\underline{\delta^*}) - C_L(\overline{\delta^*}) = \frac{1}{2}(1-\bar{T})^2[\mathbb{E}_\pi(\beta) - \alpha]^2 \left[a(\underline{\lambda}) - a(\overline{\lambda})\right]$, which is equivalent to B. Then, (D.2b) holds using (D.2a). An analogous argument shows that (D.2d) holds. Finally, because the solution to (10), $\lambda_j^*(T)$, is a continuous function of the equilibrium tax rate, T, the result follows by an application of Lemma 5B.

□

Proposition 1B. *Let Assumptions 1B through 3B hold, $\lambda \in \{\hat{\lambda}, \tilde{\lambda}\}$ and $\tau \in [0, 1)$. Then, the following two equilibria are possible:*

(a) Pessimism/Welfare State*: Agents are pessimistic, $\lambda^* = \tilde{\lambda}$, they impose a high tax rate, $T = \tau_L^*(\tilde{\lambda})$, and they exert low effort, $\delta_B^*(\tilde{\lambda})$.*

(b) Optimism/Laissez Faire*: Agents are optimistic, $\lambda^* = \hat{\lambda}$, they impose a low tax rate, $T = \tau_L^*(\hat{\lambda})$, and they exert high effort, $\delta_B^*(\hat{\lambda})$.*

with $\tilde{\lambda} > \hat{\lambda}$, $\tau_L^*(\tilde{\lambda}) > \tau_L^*(\hat{\lambda})$, and $\delta_B^*(\tilde{\lambda}) < \delta_B^*(\hat{\lambda})$.

Proof. The argument is similar to the proof of Proposition 1.

□

E Minimax-Regret

E.1 Assumptions and setup

Assumption 1MR. *Let* $\tau \in (0,1)$ *and* $(1-\tau)\beta_L - C_j(\delta) < (1-\tau)\alpha - C_j(0) < (1-\tau)\beta_U - C_j(\delta), \quad \forall \delta \in [0,1], \ \forall \tau.$

Assumption 2MR. *Let* $M(\lambda) > 0$ *and* $M'(\lambda) > 0$ *be the cognitive technology through which the agents can manipulate their own degree of pessimism.*

Assumption 3MR. *Let* $C_j(\delta) = \frac{1}{2}\frac{\delta^2}{a_j + a(\lambda)}$*, where* $a(\lambda) > a_L, \forall \lambda \in [0,1]$*,* $\frac{\partial a(\lambda)}{\partial \lambda} := a'(\lambda) < 0$*,* $0.55 < a_L < a_H < 1$*.*

The minimax regret criterion shows an alternative approach to the Hurwicz criterion that is typically fractional, as discussed in Section 2. We follow the same steps as with the Hurwicz criterion and omit the proofs that use the same arguments. The solution of the minimax-regret is substantially more cumbersome because the criterion requires to compute the *regret* of an allocation. Consider an agent choosing an allocation δ of effort. The resulting regret of that allocation is defined as:

$$R_\delta(\beta) = \max_{\tilde{\delta} \in [0,1]} \left\{ (1-\tau)\left[\alpha + (\beta - \alpha)\tilde{\delta}\right] - C_j(\tilde{\delta}) \right\} - \left\{ (1-\tau)\left[\alpha + (\beta - \alpha)\delta\right] - C_j(\delta) \right\}.$$

The first term in the above expression is the maximum value of the welfare function. The second subtracted term is the value of the welfare attained by choosing δ. The maximum regret is obtained by maximizing the above regret, $R_\delta(\beta)$, across ambiguity, *i.e.*, across β. Thus, the maximum regret is:

$$MR(\delta) = \max_{\beta \in [\beta_L, \beta_H]} R_\delta(\beta).$$

The minimax regret criterion chooses δ to minimize the maximum regret $MR(\delta)$. Therefore, the optimization problem is:

$$\min_{\delta \in [0,1]} MR(\delta).$$

E.2 Results

Lemma 1MR. *The minimax-regret effort allocation is:*

$$\delta^*_{MR} = \begin{cases} 0 & \textit{if } \delta_{MR} \leq 0, \\ 1 & \textit{if } \delta_{MR} \geq 1, \\ \delta_{MR} & \textit{if } \delta_{MR} \in (0,1), \end{cases}$$

where $\delta_{MR} = \frac{\beta_U - \alpha - \frac{1}{2a_j(\lambda)(1-\tau)}}{\beta_U - \beta_L}$.

Proof. Consider an individual choosing δ. The resulting regret for a given value of β is given by $R_\delta(\beta)$. To compute the regret, we need to solve the above maximization problem that has the following first-order condition:

$$\begin{aligned}(1-\tau)[\beta - \alpha] &= C'_j(\delta^*), \\ &= \frac{\delta^*}{a_j(\lambda)}.\end{aligned}$$

Solving for δ^* yields $\delta^* = a_j(\lambda)(1-\tau)[\beta - \alpha]$ and, thus, the solution is:

$$\bar{\delta} = \begin{cases} 0 & \text{if } \delta^* \leq 0, \\ 1 & \text{if } \delta^* \geq 1, \\ \delta^* & \text{if } \delta^* \in (0,1). \end{cases}$$

Let $M_0 := (1-\tau)\alpha - C_j(0)$ be the maximum value attained when $\bar{\delta} = 0$. Similarly, let M_1 and M_{δ^*} be the maximum values when $\bar{\delta} = 1$ and $\bar{\delta} = \delta^*$, respectively. Thus, $M_1 := (1-\tau)\beta - C_j(1)$ and:

$$\begin{aligned} M_{\delta^*} &:= (1-\tau)[\alpha + (\beta - \alpha)\underbrace{a_j(\lambda)(1-\tau)(\beta-\alpha)}_{\delta^*}] - C_j(\delta^*), \\ &= (1-\tau)[\alpha + (\beta-\alpha)^2 a_j(\lambda)(1-\tau)] - \underbrace{\frac{\left[a_j(\lambda)(1-\tau)(\beta-\alpha)\right]^2}{2a_j(\lambda)}}_{C_j(\delta^*)}, \\ &= (1-\tau)\left[\alpha + \frac{1}{2}(\beta-\alpha)^2 a_j(\lambda)(1-\tau)\right]. \end{aligned}$$

Plugging this expression back into the regret function yields the resulting regret:

$$R_\delta(\beta) = \max\{M_0, M_1, M_{\delta^*}\} - F(\delta),$$

where $F(\delta) := (1-\tau)\big[\alpha + (\beta-\alpha)\delta\big] - C_j(\delta)$.

The maximum regret across all feasible values of β is:

$$
\begin{aligned}
MR(\delta) &= \max_{\beta\in[\beta_L,\beta_H]} R_\delta(\beta), \\
&= \max_{\beta\in[\beta_L,\beta_H]} \Big\{ [M_0 - F(\delta)]\mathbf{1}\{\delta^* \leq 0\} + [M_1 - F(\delta)]\mathbf{1}\{\delta^* \geq 1\} \\
&+ [M_{\delta^*} - F(\delta)]\mathbf{1}\{\delta^* \in (0,1)\}\Big\},
\end{aligned}
$$

where $\mathbf{1}\{\cdot\}$ is an indicator function. Let:

$$
\begin{aligned}
a(\beta;\delta) := M_0 - F(\delta) &= (1-\tau)\alpha - C_j(0) - (1-\tau)\alpha - (1-\tau)(\beta-\alpha)\delta + C_j(\delta), \\
&= -C_j(0) - (1-\tau)(\beta-\alpha)\delta + C_j(\delta).
\end{aligned}
$$

Thus, $\beta_L = \arg\max_{\beta\in[\beta_L,\beta_H]} a(\beta;\delta)$. Using Assumption 3MR: $a^*(\beta_L;\delta) = -(1-\tau)(\beta_L - \alpha)\delta + \frac{\delta^2}{2a_j(\lambda)}$. Note that $\frac{\partial a^*(\beta_L;\delta)}{\partial\delta} = -(1-\tau)\underbrace{(\beta_L-\alpha)}_{<0} + \frac{1}{a_j(\lambda)} > 0$.

Similarly, let:

$$
\begin{aligned}
b(\beta;\delta) := M_1 - F(\delta) &= (1-\tau)\beta - C_j(1) - (1-\tau)\alpha - (1-\tau)(\beta-\alpha)\delta + C_j(\delta), \\
&= (1-\delta)(1-\tau)(\beta-\alpha) - C_j(1) + C_j(\delta).
\end{aligned}
$$

Now, $\beta_U = \arg\max_{\beta\in[\beta_L,\beta_H]} b(\beta;\delta)$. Using Assumption 3MR: $b^*(\beta_U;\delta) = (1-\delta)(1-\tau)(\beta_U - \alpha) - \frac{1}{2a_j(\lambda)} + \frac{\delta^2}{2a_j(\lambda)}$ and $\frac{\partial b^*(\beta_L;\delta)}{\partial\delta} = -(1-\tau)\underbrace{(\beta_U-\alpha)}_{>0} + \frac{1}{a_j(\lambda)} < 0$, because $\mathbf{1}\{\delta^* \geq 1\}$ holds and the last condition implies: $\beta > \frac{1}{a_j(\lambda)(1-\tau)} + \alpha$.

Finally let:

$$
\begin{aligned}
c(\beta;\delta) := M_{\delta^*} - F(\delta) &= (1-\tau)\left[\alpha + \frac{1}{2}(\beta-\alpha)^2 a_j(\lambda)(1-\tau)\right] - (1-\tau)\alpha \\
&- (1-\tau)(\beta-\alpha)\delta + C_j(\delta), \\
&= (1-\tau)(\beta-\alpha)\left[\frac{1}{2}(1-\tau)(\beta-\alpha)a_j(\lambda) - \delta\right] \\
&+ \frac{\delta^2}{2a_j(\lambda)}, \quad \text{by Assumption 3MR.}
\end{aligned}
$$

Taking the derivative of the previous expression with respect to β yields:

$$\begin{aligned}\frac{\partial c(\beta;\delta)}{\partial\beta} &= (1-\tau)\left[\frac{1}{2}(1-\tau)(\beta-\alpha)a_j(\lambda)-\delta\right]\\ &+(1-\tau)(\beta-\alpha)\frac{1}{2}(1-\tau)>0,\\ &\iff (\beta-\alpha)(1-\tau)a_j(\lambda)>\delta.\end{aligned}$$

Thus, $(\beta-\alpha)(1-\tau)a_j(\lambda)>\delta$, implies $\frac{\partial c(\beta;\delta)}{\partial\beta}>0$ and, hence, $\beta_U=\arg\max_{\beta\in[\beta_L,\beta_H]}c(\beta;\delta)$. Similarly, $(\beta-\alpha)(1-\tau)a_j(\lambda)<\delta$, implies $\frac{\partial c(\beta;\delta)}{\partial\beta}<0$ and, hence, $\beta_L=\arg\max_{\beta\in[\beta_L,\beta_H]}c(\beta;\delta)$.

Finally, note that:

$$c^*(\beta_U;\delta)=(1-\tau)(\beta_U-\alpha)\left[\frac{1}{2}(1-\tau)(\beta-\alpha)a_j(\lambda)-\delta\right]+\frac{\delta^2}{2a_j(\lambda)},$$

$$c^*(\beta_L;\delta)=(1-\tau)(\beta_L-\alpha)\left[\frac{1}{2}(1-\tau)(\beta-\alpha)a_j(\lambda)-\delta\right]+\frac{\delta^2}{2a_j(\lambda)}.$$

Thus, the maximum regret is:

$$\begin{aligned}MR(\delta)=\max\{&a^*(\beta_L;\delta),b^*(\beta_L;\delta),c^*(\beta_U;\delta)\mathbf{1}\{0<\delta<(\beta-\alpha)(1-\tau)a_j(\lambda)\},\\ &c^*(\beta_L;\delta)\mathbf{1}\{(\beta-\alpha)(1-\tau)a_j(\lambda)<\delta<1\}\}\end{aligned}$$

And provided $a^*(\beta_L;\delta)\geq c^*(\beta_L;\delta)$ $\forall\delta\in\big((\beta-\alpha)(1-\tau)a_j(\lambda),1\big)$ and $b^*(\beta_U;\delta)\geq c^*(\beta_U;\delta)$ $\forall\delta\in\big(0,(\beta-\alpha)(1-\tau)a_j(\lambda)\big)$ we obtain:

$$\begin{aligned}MR(\delta)&=\max\{a^*(\beta_L;\delta),b^*(\beta_L;\delta)\},\\ &=\max\{\underbrace{-(1-\tau)(\beta_L-\alpha)\delta+\frac{\delta^2}{2a_j(\lambda)}}_{a^*(\beta_L;\delta)},\underbrace{(1-\delta)(1-\tau)(\beta_U-\alpha)-\frac{1}{2a_j(\lambda)}+\frac{\delta^2}{2a_j(\lambda)}}_{b^*(\beta_L;\delta)}\}.\end{aligned}$$

The objective is to choose δ to minimize $MR(\delta)$. Because, as noted above, $a^*(\beta_L;\delta)$ is increasing in δ and $b^*(\beta_L;\delta)$ is decreasing in δ, the solution to $\min_{\delta\in[0,1]}MR(\delta)$ is obtained by choosing δ to equalize both quantities. Thus, we obtain the minimax-regret treatment allocation by solving:

$$-(1-\tau)(\beta_L-\alpha)\delta_{MR}+\frac{\delta_{MR}^2}{2a_j(\lambda)}=(1-\delta_{MR})(1-\tau)(\beta_U-\alpha)-\frac{1}{2a_j(\lambda)}+\frac{\delta_{MR}^2}{2a_j(\lambda)}.$$

Solving this expression yields $\delta_{MR} = \frac{\beta_U - \alpha - \frac{1}{2a_j(\lambda)(1-\tau)}}{\beta_U - \beta_L}$ and, therefore, the minimax-regret allocation is:

$$\delta^*_{MR} = \begin{cases} 0 & \text{if } \delta_{MR} \leq 0, \\ 1 & \text{if } \delta_{MR} \geq 1, \\ \delta_{MR} & \text{if } \delta_{MR} \in (0,1). \end{cases}$$

□

Lemma 2MR. *Let Assumptions 1MR and 3MR hold. Then, the optimal level of effort decreases as the tax rate increases.*

Proof. Note that:

$$\begin{aligned} \frac{\partial \delta_{MR}}{\partial \tau} &= \frac{-\left[\frac{-2a_j(\lambda)(-1)}{[2a_j(\lambda)(1-\tau)]^2}\right](\beta_U - \beta_L)}{(\beta_U - \beta_L)^2}, \\ &= -\frac{1}{2a_j(\lambda)(1-\tau)^2}, \\ &< 0. \end{aligned}$$

Thus, $\frac{\partial \delta^*_{MR}}{\partial \tau} \leq 0$.

□

One can see from the above lemmas that the minimax-regret allocation is, in general, fractional. The level of effort decreases as the tax rate τ increases ($\frac{\partial \delta_{MR}}{\partial \tau} < 0$) and increases with the effort-parameter $a_j(\lambda)$ ($\frac{\partial \delta_{MR}}{\partial a_j(\lambda)} > 0$). Skilled or wealthy individuals who have higher $a_j(\lambda)$ (and, hence, lower cost) exert more effort.

Lemma 3MR. *Let Assumptions 1MR through 3MR hold. Then, the ideal tax rate of agent j is:*

$$\tau^*_j(\lambda) = \begin{cases} \tau^* & \textit{if} \quad \tau^* > \underline{\tau}, \\ \underline{\tau} & \textit{otherwise}, \end{cases} \tag{E.1}$$

where τ^* is defined below and $\underline{\tau}$ is the lowest possible tax (e.g., $\underline{\tau} = 0$ *if* $\tau \in [0,1)$).

Proof. Mathematica shows that:

$$\frac{\partial U_j^{MR}(\lambda, \delta, \tau)}{\partial \tau} = N_1 \times [N_2 + N_3(\tau(\lambda)) + N_4(\tau(\lambda) + N_5(\tau(\lambda)],$$

where:

$$N_1 := \frac{1}{4\left(\beta_u - \beta_l\right)^2},$$

$$N_2 := 4\left(\beta_u - \beta_l\right)\left(2\alpha^2 + \left(\beta + \beta_l\right)\beta_u - \alpha\left(\beta + 2\beta_l + \beta_u\right)\right),$$

$$N_3 := -\frac{1}{a_j^3(\lambda)(\tau^* - 1)^3},$$

$$N_4 := \frac{2a_H(\lambda)\left(a_l(\lambda)\left(\alpha - \beta_n\right) - a_j^2(\lambda)(\beta - \alpha)\left(\beta_u - \beta_l\right)(1 - \phi)\right)}{a_H(\lambda)a_j(\lambda)(1 - \tau^*)^2},$$

$$N_5 := -\frac{2a_j^2(\lambda)a_L(\lambda)(\beta - \alpha)\left(\beta_u - \beta_l\right)\phi}{a_H(\lambda)a_j(\lambda)(1 - \tau^*)^2}.$$

At an interior solution, the optimal tax rate is given by:

$$\tau^* := \{\tilde{\tau} : N_2 + N_3(\tilde{\tau}) + N_4(\tilde{\tau}) + N_5(\tilde{\tau}) = 0\}. \tag{E.2}$$

□

Lemma 4MR. *Let Assumptions 1MR through 3MR hold and $\lambda \in (\underline{\lambda}, \bar{\lambda})$. Then, the equilibrium tax outcome, $T = \tau_L^*(\lambda)$, is increasing with the pessimism rate, λ.*

Proof. The results follows from Lemma 3MR by applying the implicit function theorem to (E.2) with $T = \tau_L^*(\lambda)$.

□

Lemma 5MR. *Let Assumptions 1MR through 3MR hold and $\lambda \in (\underline{\lambda}, \bar{\lambda})$. Then, the optimal pessimism rate, λ^*, increases with the equilibrium tax rate, T.*

Proof. Let:

$$F_j(\lambda, T) := (1 - T(\lambda))[\alpha - \beta_L]\delta^*_{MR} + \frac{\delta^2_{MR}}{2a_j(\lambda)} + \tau\bar{Y} + M(\lambda),$$
$$J_j(\lambda, T) := \frac{1}{T(1-T)} F_j(\lambda, T).$$

It is sufficient to show that $\frac{\partial^2 J_j(\lambda,T)}{\partial T(\lambda)\partial\lambda} > 0$.

Mathematica shows that the derivative of $J_j(\lambda, T)$ with respect to $T(\lambda)$ is:

$$\frac{\partial J_j(\lambda, T)}{\partial T(\lambda)} = M_1 \times M_2,$$

where:

$$M_1 := \left[\frac{1}{8(\beta_L - \beta_U)^2(T(\lambda) - 1)^2 T(\lambda)^2}\right],$$

$$\begin{aligned} M_2 := \Bigg(& \frac{8(\beta_L - \beta_U)(a_H(\lambda)(-a_L(\lambda)(T(\lambda) - 1)T(\lambda)^2(2\alpha^2 - \alpha(\beta + 2\beta_L + \beta_U) + \beta_U(\beta + \beta_L))}{a_H(\lambda)a_L(\lambda)(T(\lambda) - 1)} \\ & - 2(\alpha - \beta_L)(\alpha - \beta_U)T(\lambda) + (\alpha - \beta_L)(\alpha - \beta_U) - (\beta_L - \beta_U)M(\lambda)(2T(\lambda) - 1) \\ & - (\phi(\alpha - b)T(\lambda)^2)) + \phi a_L(\lambda)(\alpha - b)T(\lambda)^2) \\ & + \frac{4(2T(\lambda) - 1)(\alpha^2 + \alpha(\beta_L - 3\beta_U) - \beta_L^2 + \beta_L\beta_U + \beta_U^2)}{a_j(\lambda)} \\ & - \frac{4(\alpha - \beta_U)(3T(\lambda) - 1)}{a_j(\lambda)^2(T(\lambda) - 1)} + \frac{4T(\lambda) - 1}{a_j(\lambda)^3(T(\lambda) - 1)^3}\Bigg). \end{aligned}$$

Mathematica also reveals that $\frac{\partial^2 J_j(\lambda,T)}{\partial T(\lambda)\partial\lambda} > 0$ with:

$$\frac{\partial^2 J_j(\lambda, T)}{\partial T(\lambda)\partial\lambda} = M_3 \times M_4,$$

where:

$$M_3 := [-8(\beta_U - \beta_L)^2(1 - T(\lambda))^2 T(\lambda)^3]^{-1},$$

$$M_4 := m_1 + m_2 + m_3 + m_4 + m_5 + m_6,$$

with:

$$m_1 := \frac{-8\phi(\beta-\alpha)(\beta_U-\beta_L)T^3(\lambda)a'_H(\lambda)}{a_H^2(\lambda)},$$

$$m_2 := \frac{3(4T(\lambda)-1)T(\lambda)a'_j(\lambda)}{-a_j^4(\lambda)(1-T(\lambda))},$$

$$m_3 := \frac{-24\phi(\beta-\alpha)(\beta_U-\beta_L)T^3(\lambda)T'(\lambda)}{-a_H(\lambda)(1-T(\lambda))}$$
$$\frac{8(-\alpha^2+\beta_L^2-\alpha(\beta_L-3\beta_U)+\beta_L\beta_U+\beta_U^2)(1-3(1-T(\lambda))T(\lambda)T'(\lambda)}{a_j(\lambda)},$$

$$m_4 := \frac{2\times(-4(\beta_U-\alpha)T(\lambda)(3T(\lambda)-1)T'(\lambda)+\frac{-1+5(1-2T(\lambda)T(\lambda)T'(\lambda)}{(1-T(\lambda))^2})}{a_j(\lambda)^3},$$

$$m_5 := [-4(\alpha^2-\beta_L^2-\alpha(3\beta_U-\beta_L)+\beta_L\beta_U+\beta_U^2)(1-T(\lambda))^2T(\lambda)(2T(\lambda)-1)a'_j(\lambda)$$
$$+8(-\beta_U+\alpha)(1-4T(\lambda)+6T(\lambda)^2)T'(\lambda)]\frac{1}{-a_j(\lambda)(1-T(\lambda))},$$

$$m_6 := \frac{-1}{a_j^2(\lambda)(1-T(\lambda))}8(\beta_U-\beta_L)\Big((1-T(\lambda))T(\lambda)\Big((\beta-\alpha)(1-\phi)T^2(\lambda)a'_j(\lambda)$$
$$+(\beta_U-\beta_L)a_L^2(\lambda)(1-T(\lambda))(2T(\lambda)-1)M'(\lambda)\Big)$$
$$+a_L(\lambda)\Big(3(\beta-\alpha)(1-\phi)\ T^3(\lambda)-2a_L(\lambda)(1-T(\lambda))$$
$$\Big(-(\beta_L-\alpha)(\beta_U-\alpha)+3(\beta_L-\alpha)(\beta_U-\alpha)T(\lambda)$$
$$-3(\beta_L-\alpha)(\beta_U-\alpha)T(\lambda^2)+(2\alpha^2+(\beta+\beta_L)\beta_U-\alpha(\beta+2\beta_L+\beta_U))T^3(\lambda)$$
$$+(\beta_U-\beta_L)M(\lambda)\Big(1-3(1-T(\lambda))(T(\lambda))\Big)\Big)\Big)T'(\lambda)\Big).$$

□

Lemma 6MR. *Let Assumptions 1MR through 3MR hold and* $\lambda \in \{\hat{\lambda}, \tilde{\lambda}\}$*. Then, the equilibrium pessimism rate,* λ^**, has the following form:*

$$\lambda^* = \begin{cases} \hat{\lambda} & \text{if } T < \tilde{T}, \\ \tilde{\lambda} & \text{if } T > \tilde{T}. \end{cases}$$

Proof. The argument is similar to the proof of Lemma 6. □

Proposition 1MR. *Let Assumptions 1MR through 3MR hold,* $\lambda \in \{\hat{\lambda}, \tilde{\lambda}\}$ *and* $\tau \in [0, 1)$*. Then, the following two equilibria are possible:*

(a) Pessimism/Welfare State*: Agents are pessimistic,* $\lambda^* = \tilde{\lambda}$*, they impose a high tax rate,* $T = \tau_L^*(\tilde{\lambda})$*, and they exert low effort,* $\delta_{MR}^*(\tilde{\lambda})$*.*

(b) Optimism/Laissez Faire*: Agents are optimistic,* $\lambda^* = \hat{\lambda}$*, they impose a low tax rate,* $T = \tau_L^*(\hat{\lambda})$*, and they exert high effort,* $\delta_{MR}^*(\hat{\lambda})$*.*

with $\tilde{\lambda} > \hat{\lambda}$*,* $\tau_L^*(\tilde{\lambda}) > \tau_L^*(\hat{\lambda})$*, and* $\delta_{MR}^*(\tilde{\lambda}) < \delta_{MR}^*(\hat{\lambda})$*.*

Proof. The argument is similar to the proof of Proposition 1. □