Beliefs and Redistributive Politics under Incomplete Information

Tommaso Gabrieli

No 821

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF
WARWICK

BELIEFS AND REDISTRIBUTIVE POLITICS UNDER INCOMPLETE INFORMATION

TOMMASO GABRIELI*

SEPTEMBER 2007

JOB MARKET PAPER

ABSTRACT. The reason why the social contract is so different in two otherwise comparable societies like the United States and continental Western European countries represents a challenging question. Large empirical evidence shows that the difference in the political support for redistribution appears to reflect a difference in the social perceptions regarding the determinants of individual wealth and the underlying sources of income inequality. I present a model of beliefs and redistribution which explains this evidence through multiple politico-economic equilibria. Differently from the recent literature which obtains multiple equilibria by modeling agents characterized by psychological biases, my model is based on standard assumptions. Multiple equilibria originate from multiple optimal levels of information for the society. Multiple optimal levels of information exist because increasing the informativeness of an economy produces a trade-off between a decrease in adverse selection and an increase in moral hazard. The framework allows the analysis of various comparative statics in order to answer to policy questions.

1. INTRODUCTION

The reason why the social contract is so different in two otherwise comparable societies like the United States and continental Western European countries ("Europe"in short) represents a challenging question which has motivated a large body of research.[1] Such difference in the political support for redistribution appears to reflect a difference in the beliefs which a society holds about the underlying determinants of individual wealth, rather than a substantial difference in the fundamentals of the two economies. As stated by Alesina and Angeletos (2005): "Americans believe that poverty is due to bad choices or lack of effort; Europeans instead view poverty as a trap from which it is hard to escape. Americans perceive wealth and success as the outcome of individual talent, effort, and entrepreneurship; Europeans instead attribute a larger role to luck, corruption, and connections"(Alesina and Angeletos (2005)).[2]

The theoretical contributions of Piketty (1995), Alesina and Angeletos (2005) and Benabou and Tirole (2006) have focused on the link between beliefs and political outcome and have interpreted the empirical evidence through models with multiple politico-economic equilibria. The three cited models extend the standard framework of Meltzer and Richard (1981): agents vote for redistribution first, once that the winning level of redistribution is announced they choose the amount of effort to implement; in the individual choice of the ideal level of redistribution the gains from redistribution are traded off the moral hazard effect of redistribution, namely the fact that the higher is the level of redistribution and the lower is individual effort

---

[1]Pre-tax inequality is higher in the United States than in Europe, nevertheless Europe is characterized by more extensive redistributive policies than the United States. Alesina and Angeletos (2005) report that while the Gini coefficient in the pre-tax income distribution in the United States is 38.5 against 29.1 in Europe, the income tax structure is more progressive in Europe, the overall size of government is about 50 per cent larger in Europe than in the United States (about 30 versus about 45 per cent of GDP) and the largest difference is represented by transfers and other social benefits, where Europeans spend about twice as much as Americans. More extensive and detailed evidence can be found in Alesina, Glaeser, and Sacerdote (2001) and Alesina and Glaeser (2004).

[2]"The data from the World Values Survey reported by Alesina, Glaeser, and Sacerdote (2001) and Keely (2002) show that only 29 percent of Americans believe that the poor are trapped in poverty and only 30 percent that luck, rather than effort or education, determines income. Conversely, the data for Europe are 60 percent and 54 percent, respectively. Ladd and Bowman (1998) show that in a similar way 60 percent of Americans versus 26 percent of Europeans are likely to think that the poor "are lazy or lack willpower"and that 59 percent of Americans versus 34 percent of Europeans are likely to think that "in the long run, hard work usually brings a better life"."(Benabou and Tirole (2006)). The observed correlation between the social beliefs and the actual levels of redistribution is not limited to a comparison of the United States and Europe. See Alesina, Glaeser, and Sacerdote (2001) about the fact that the same correlation can be observed at the cross-country level.

and therefore the lower is the amount of output which is redistributed. Piketty (1995) enriches the standard framework introducing imperfect information and learning.[3] Both Alesina and Angeletos (2005) and Benabou and Tirole (2006) model agents whose preferences are not the standard preferences of Meltzer and Richard (1981) and which present psychological biases. Alesina and Angeletos (2005) model agents who have a concern for the fairness of the economic system, namely for the fact that people should get what they deserve and effort rather than luck should determine economic success.[4] Similarly to the paper of Alesina and Angeletos (2005), also in the recent work of Cervellati, Esteban, and Kranich (2006) the individual preferred level of redistribution is not motivated by purely selfish concerns but also by a social component; nevertheless multiple equilibria do not originate from different beliefs but from different moral sentiments. Differently, in the work of Benabou and Tirole (2006) multiple beliefs are possible because the agents find optimal to deliberately bias their own perception of the truth so as to offset another bias which is imperfect willpower.[5]

The first contribution of my analysis is to use part of the machinery of Benabou and Tirole (2006) in order to answer to some specific policy questions: namely to analyze how the informativeness of an economy does affect the voting on redistribution, the choice of effort and the aggregate output. In the framework of Benabou

---

[3]In the model Piketty (1995), agents have imperfect information about the true return on effort versus the role of predetermined factors and the experimentation of different levels of effort is costly. This implies that the steady-state beliefs resulting from a bayesan learning process over an infinite horizon do not necessarily have to be the correct ones. US- (Europe-) type equilibria characterized by the widespread belief that effort plays a major (minor) role and by low (high) redistribution are possible equilibria.

[4]They discuss two equilibria of the model: in a US-type equilibrium agents believe that effort more than luck determines personal wealth, consequently they vote for low redistribution, incentives are not distorted and the belief is self sustained. Conversely, in the Europe-type equilibrium agents believe that the economic system is not fair and factors as luck, birth, connections, rather than effort, determine personal wealth, hence they vote for high taxes, thus distorting allocations and making the believes self sustained.

[5]Using the words of Benabou and Tirole (2006),"The basic model works as follows. Because of imperfect willpower, people [...] strive to motivate themselves [...] toward effort and given this every agent finds valuable to hold biased beliefs and to think that the return on effort is greater than the true value, but this bias has a cognitive cost. [...] when people anticipate little redistribution, the value of a proper motivation is much higher than with lower redistribution. Everyone thus has greater incentives to believe in self-sufficiency, and consequently more voters finds optimal to hold to such a world-view. Due to these complementarities between individuals ideological choices, there can be two equilibria. A first, "American"equilibrium is characterized by a high prevalence of just-world beliefs and relatively low redistribution. The other, "European"equilibrium is characterized by more pessimism and a more extensive welfare state"(Benabou and Tirole (2006)).

and Tirole (2006) the ideological choice is modeled as the choice to bias an informative signal about the true return on effort in a particular direction. A natural extension of this analysis is to reinterpret the precision of the signal as the degree of informativeness of the economy and to analyze the implications of different degrees of informativeness. This represents an important policy question as the degree of incomplete information about the true return on effort is affected by factors like education, propaganda or other policy variables. In order to derive precise comparative statics I build a "neoclassical" variation of the framework of Benabou and Tirole (2006). In my model agents are fully rational and they vote for redistribution based on exclusively selfish concerns as in the standard framework of Meltzer and Richard (1981). What is added to the standard framework is a simple way to introduce varying degrees of incomplete information about the true return on effort: agents do not know the true value of the return on effort but receive an informative signal about this value and they update their information conditioning on the signal in a Bayesan way. Varying the precision of the signal means to vary the degree of incompleteness of the information in the economy. This framework isolates the effect of incomplete information, as "psychological biases" are not present in the model, and allows a clear analysis of a number of interesting comparative statics. There is a strong link between my object of analysis and those of the literature in growth and in optimal taxation. To my knowledge, in such context, my work is the first attempt to analyze the comparative statics of introducing varying degrees of incomplete information about the return on effort.

The second contribution of my work is to obtain multiple politico-economic equilibria in a framework with standard preference and without psychological biases and to offer a new interpretation for the US- vs Europe-type equilibria. In order to do this I endogenize the level of precision of the signal analyzing the optimal ex-ante precision for the economy. Given that each agent is ex-ante identical, the optimal ex-ante precision can be interpreted as the prevailing degree of information in the economy, no matter wether it is the choice of a benevolent planner, the result of a collective choice as a voting outcome, or the choice of a generation beyond the veil of ignorance for the next generation. I introduce the concept of politico-economic equilibrium as the beliefs, the prevailing level of redistribution and optimal choices of effort which result from the ex-ante optimal precision of the signal. I show that ex-ante there are cases in which multiple optimal values of the signal's precision, and consequently multiple politico-economic equilibria, exist for the society. In the

case in which multiple equilibria do exist, I find a US-type vs a Europe-type politico-economic equilibrium as characterized by a relative (i) low informative signal and high adverse selection – as individual beliefs and effort levels are pooling to similar levels – (ii) low redistribution (iii) low moral hazard – as redistribution is low and this does not distort individual effort much – (iv) high aggregate effort and output. Conversely the Europe-type politico-economic equilibrium is interpreted as an equilibrium characterized by a relative (i) high informative signal and low adverse selection (ii) high redistribution (iii) high moral hazard – as redistribution is high and this diminishes individual effort (iv) low aggregate effort and output. The intuition is that multiple values of the signal's precision are ex-ante optimal because increasing the informativeness of the signal implies a trade off between the positive effect of an increase in the precision of the signal – namely that more information reduces the adverse selection as agents choose effort more optimally with respect to the the true return on effort – and the negative effect – namely that more information increases the prevailing tax rate and this creates a moral hazard effect which reduces the aggregate (or ex-ante individual) output –.

With respect to the existence of multiple politico-economic equilibria, my paper presents a methodological contribution as it shows that multiple equilibria can be obtained without psychological biases. With respect to the interpretation of the multiple politico-economic equilibria, the contribution is to open the door to a different interpretation based on the specific roles of adverse selection and moral hazard. Some empirical evidence from the literature in education goes in the direction of the prediction of my model. For example, Bishop (1996) shows how the American secondary schooling system is less informative than the European about the position of a student in the national distribution of abilities. Moreover, my interpretation of the society's informativeness as the one which is ex-ante optimal for the society could be interesting for the "Neo-marxist"-type of explanation of the differences between US and Europe, which is more common in the literature in political science. A more modern and more symmetric version of this view can be found in the work of Alesina and Glaeser (2004). The authors argue that just as American beliefs result from indoctrination predominantly controlled by the wealthier classes, European beliefs result from indoctrination predominantly controlled by Marxist-influenced intellectuals. Alesina and Glaeser (2004) claim that the process of indoctrination has been achieved through the choice of specific institutions and political systems. For example they show how, in the American political history, factors like federalism,

majority representation and segregation worked towards low cross-ethnic cohesion and the already described beliefs. My analysis shows how certain beliefs can be imposed in a society not only through the choice of particular institutions but also through the choice of a certain level of information. Moreover, my analysis shows that the prevailing informativeness can be actually maintained by the society as an autonomous collective choice and not only as the result of a process of indoctrination. This because it can be the case that behind the veil of ignorance there is no gain in changing to a different level of information. A deeper analysis of the interpretation of my results in relation to the educational and the institutional features of the two societies goes beyond the scope of this paper; nevertheless it could constitute a fertile ground for future research.

The paper is structured as follows. Section 2 introduces the set-up of the model. Section 3 analyzes the voting problem and the relative outcome. Section 4 analyzes the comparative statics considering the precision of the signal as an exogenous policy variable. In section 5 I analyze the optimal ex-ante precision for the economy. In section 6 I introduce the concept of politico-economic equilibrium and investigate the possibility of existence of multiple equilibria. Section 7 analyzes the robustness of the results and section 8 concludes.

2. SET UP

Consider an economy populated by a continuum of agents $i \in [0, 1]$. Each individual i produces a quantity y^i of output with the following technology:

$$(1) \qquad\qquad y^i = k^i + \theta^i e^i,$$

where k^i is an observable endowment of resources, e^i is the effort implemented by agent i and θ^i is the return to effort or productivity. In this basic version of the model I assume that the endowment is homogeneous across agents, i.e. $k^i = k$ for all i [6], but I will later consider the possibility of heterogeneous endowments. I assume that θ^i takes value θ_L for a fraction π of the population and value θ_H for the remaining fraction $1 - \pi$, with $\theta_L < \theta_H$. Agents have incomplete information: each agent i cannot observe her own or other agents' productivity but only receives a private signal σ^i about the true value of θ^i. Also the signal σ^i is binary. If $\theta^i = \theta_L$ ($\theta^i = \theta_H$), σ^i takes values σ_L (σ_H) or σ_H (σ_L), respectively with probability λ and $1 - \lambda$. In other

---

[6]It will be clear that an homogeneous endowment does not play any role and without loss of generality I could set $k = 0$.

words for each agent i the signal σ^i is independently distributed, it is truthful with probability λ, false with probability $1 - \lambda$ and the transition matrix which takes from the true productivity to the signal is the following:

(2)
$$T\left(\begin{bmatrix} \sigma_L \\ \sigma_H \end{bmatrix} \middle| [\theta_L, \theta_H]\right) = \begin{pmatrix} \lambda & 1 - \lambda \\ 1 - \lambda & \lambda \end{pmatrix}.$$

The structure of the economy – including the value of π and matrix (2) – is common knowledge, the only incomplete information is about the true values of the θ's. Agents are fully rational and agent's i belief of the true value of θ^i, conditional on the observation of the private signal σ^i, is obtained by the Bayes Rule. I introduce the following notation:

(3)
$$\mu^i \equiv \Pr[\theta^i = \theta_L | \sigma^i],$$

(4)
$$\mu_{\sigma_L} \equiv (\mu^i | \sigma_L) = \frac{\pi\lambda}{\pi\lambda + (1 - \pi)(1 - \lambda)}$$

represents the probability that $\theta^i = \theta_L$ conditional on the observation of $\sigma^i = \sigma_L$,

(5)
$$\mu_{\sigma_H} \equiv (\mu^i | \sigma_H) = \frac{\pi(1 - \lambda)}{\pi(1 - \lambda) + \lambda(1 - \pi)}$$

represents the probability that $\theta^i = \theta_L$ conditional on the observation of $\sigma^i = \sigma_H$ and

(6)
$$\theta(\mu^i) \equiv \mu^i \theta_L + (1 - \mu^i)\theta_H$$

represents the expected value of θ^i conditional on the observation of σ^i. It is natural to interpreter the value of λ as the degree of information in the economy. It is straightforward that for $\lambda = 0$ and $\lambda = 1$ the signal is perfectly informative. Instead, for $\lambda = 1/2$ the signal is completely uninformative and the posterior belief μ^i is equal to the prior π.

(7)
$$p_{\sigma_L} \equiv \Pr[\sigma^i = \sigma_L] = \lambda\pi + (1 - \lambda)(1 - \pi)$$

represents the ex-ante probability of observing σ_L,

(8)
$$p_{\sigma_H} \equiv \Pr[\sigma^i = \sigma_H] = \lambda(1 - \pi) + \pi(1 - \lambda) = 1 - p_{\sigma_L}$$

represents the probability of observing σ_H. Over-lined variables stand for average values for the population, hence $\bar{y}$ and $\bar{e}$ are respectively the average, or aggregate,

values of output and effort.

$$\bar{\theta} \equiv \pi\theta_L + (1 - \pi)\theta_H,$$

$$\overline{\theta^2} \equiv \pi\theta_L^2 + (1 - \pi)\theta_H^2,$$

are respectively the average values of productivity and squared productivity. Agents face a linear income tax/redistribution scheme which implies the following expression for individual consumption:

(9) $$c_i = (1 - \tau)y_i + \tau\bar{y},$$

where τ is the tax rate which prevails in the political game with majority voting. Throughout the analysis I consider the following individual utility function:

(10) $$u_i(c_i, e_i) = c_i - \frac{a}{2}(e_i)^2.$$

I consider three periods $t = \{0, 1, 2\}$ and the following timing. In period 0 each agent only knows the values π, λ and the structure of the game. In period 1 each agent i receives the private signal σ^i, then votes over the tax rate τ and once that the prevailing tax rate is revealed, each agent i chooses the effort level e^i. In the final period individual income y^i is realized [7], agents get the net outcome of the production activity plus a net transfer and enjoy consumption.

3. Voters' Problem

Plugging (1) and (9) into (10) I obtain the expression of the expected utility of agent i at t:

(11) $$u_t^i = E[(1 - \tau)(k + e^i\theta^i) + \tau(k + \overline{e\theta}) - a(e^i)^2/2 | I_t^i],$$

where $E[\cdot|I_t^i]$ is individual i's expectation conditional on the information at t, where as explained in the previous section $I_0^i = T$, $I_1^i = (T, \sigma^i)$ and $\overline{e\theta}$ is the average $e^i\theta^i$ of the population. Solving backwards, each individual i maximizes (11) choosing e^i after that the winning tax rate τ is announced. Being (11) strictly concave in e^i, by solving the sufficient first order condition I find the optimal individual level of effort:

(12) $$e^i = (1 - \tau)\theta(\mu^i)/a.$$

---

[7] Therefore the agents get to know the true values of the θ's.

By backward induction, I can plug (12) into (11) and find the objective function that i maximizes when voting for the tax rate. In order to do this, I specify the following terms:

(13)
$$E[e^i\theta^i|I_1^i] = (1-\tau)\left(\theta(\mu^i)\right)^2/a,$$

(14)
$$E[(e^i)^2|I_1^i] = \left(\frac{1-\tau}{a}\right)^2\theta(\mu^i)^2,$$

(15)
$$\overline{e\theta} = (1-\tau)\Gamma/a,$$

where

(16) $\quad \Gamma \equiv \pi\theta_L\left(\lambda\theta(\mu_{\sigma_L}) + (1-\lambda)\theta(\mu_{\sigma_H})\right) +$

$$(1-\pi)\theta_H\left((1-\lambda)\theta(\mu_{\sigma_L}) + \lambda\theta(\mu_{\sigma_H})\right)$$

shows that a fraction π $(1-\pi)$ of the agents have productivity θ_L (θ_H) and that among those a fraction λ chooses the optimal effort after the observation of σ_L (σ_H), whereas a fraction $1-\lambda$ chooses the optimal effort after the observation of σ_H (σ_L). Collecting $\theta(\mu_{\sigma_L})$ and $\theta(\mu_{\sigma_H})$ it is easy to re-write expression (16) as

(17)
$$\Gamma = p_{\sigma_L}\theta(\mu_{\sigma_L})^2 + (1-p_{\sigma_L})\theta(\mu_{\sigma_H})^2.$$

Plugging (13), (14) and (16) into (11), voter i's problem follows:

(18)
$$\max_{\text{w.r.t } \tau}\left[k^i + (1-\tau)^2\theta(\mu^i)^2/a + \tau(1-\tau)\Gamma/a - (1-\tau)^2(\mu^i)^2/2a\right].$$

Assuming for the moment that the second derivative of the objective function in (18) is strictly negative, the preferred tax rate for agent i follows:

(19)
$$\tau(\mu^i) = 1 - \frac{1}{2 - \frac{\theta(\mu^i)^2}{\Gamma}}.$$

As explained by Benabou and Tirole (2006), the denominator of (19) shows how the subjective prospects of upward mobility (POUM) reduce the desired tax rate.[8]

Assumption 1: $2\theta_L^2 > \theta_H^2$.

Proposition 1. *Under Assumption 1 the individual preferences for taxation are single peaked and (19) is the solution to (18).*

---

[8]The term $\frac{\theta(\mu^i)^2}{\Gamma}$ represents POUM as it is the ratio of (13) over (16).

Proof. The second derivative of the objective function in problem (18) is given by the following expression:

$$\frac{d^2 u_1^i}{d\tau} = \frac{-2\Gamma + \theta(\mu)^2}{a}.$$

The condition stated by Assumption 1 is sufficient for (20) to be strictly negative as the maximum value that $\theta(\mu)^2$ can take is θ_H^2 and the minimum value that 2Γ can take is $2\theta_L^2$. $\qquad\square$

Labeling the prevailing tax rate as τ, I analyze the political outcome. There are two groups of voters in the economy: those who observe σ_L and those who observe σ_H, respectively with preferred tax rates $\tau(\mu_{\sigma_L})$ and $\tau(\mu_{\sigma_H})$. Given the majority voting rule, if $p_{\sigma_L} > (<) \, 1/2$, then $\tau = \tau(\mu_{\sigma_L}) \, (\tau = \tau(\mu_{\sigma_H}))$ is the prevailing tax rate in the economy.[9]

4. COMPARATIVE STATICS

I analyze the effect of a change in the value of λ on the endogenous variables of the model: prevailing tax rate, aggregate effort, aggregate output. This is an important exercise in order to understand the effects of policies which change – directly or indirectly – the level of information of an economy.[10] In the following two lemmas I present two important intermediate results which are fundamental for the full analysis of the comparative statics.

Lemma 1. *Expression (6) is a continuous map in λ, where $\theta(\mu_{\sigma_L})$ and $\theta(\mu_{\sigma_H})$ are (i) symmetric to each other with respect to $\lambda = 1/2$ and (ii) respectively decreasing and increasing in λ.*

Proof. Continuity follows immediately from expressions (4) and (5). Property (i) (symmetry) is simply proved by noticing that $\theta(\mu_{\sigma_L})$ and $\theta(\mu_{\sigma_H})$ are equal to each other if in one of the two λ is replaced by $1 - \lambda$. Property (iii) (monotonicity) follows immediately once that the respective first derivative with respect to λ is computed:

$$\theta(\mu_{\sigma_L})_\lambda \equiv \frac{d\theta(\mu_{\sigma_L})}{d\lambda} = -\frac{\pi(1-\pi)(\theta_H - \theta_L)}{(2\pi\lambda + 1 - \lambda - \pi)^2} < 0,$$

---

[9]Obviously when $p_{\sigma_L} = 1/2$ the majority group is undetermined. Notice also that if $\lambda = 1/2$ the signal is uninformative and $\mu_{\sigma_L} = \mu_{\sigma_H} = \pi$ – namely the prior is equal to the posterior – and every agent prefers the same tax rate $\tau(\mu)$, where $\mu = \pi$. From (19) it is immediate to notice that for $\mu = \pi$, $\tau(\mu) = 0$ as for $\mu_{\sigma_L} = \mu_{\sigma_H} = \pi$ it is the case that $\theta(\mu)^2 = \Gamma = \overline{\theta^2}$.

[10]For example policies based on education or policies based on propaganda.

$$\theta(\mu_{\sigma_H})_\lambda \equiv \frac{d\theta(\mu_{\sigma_H})}{d\lambda} = \frac{\pi(1-\pi)(\theta_H - \theta_L)}{(2\pi\lambda - \lambda - \pi)^2} > 0.$$

□

It is easy to understand the intuition behind this result. Looking at the expression of $\theta(\mu_{\sigma_L})$, it can be noticed that for $\lambda = 0$ the signal σ_L is perfectly informative and $\theta(\mu_{\sigma_L}) = \theta_H$. This because $\theta(\mu_{\sigma_L})$ is a weighted average of θ_L and θ_H and in the case of $\lambda = 0$ all the weight is placed on θ_H. Increasing λ up to $\lambda = 1/2$ makes the signal progressively less informative so that $\theta(\mu_{\sigma_L})$ decreases, as the weight placed on θ_L increases[11]. A further increase in λ up to $\lambda = 1$ progressively increases back the informativeness of the signal, as it increases the weight placed on θ_L [12].

Lemma 2. *Expression (16) is a continuous map in λ which is (i) symmetric with respect to $\lambda = 1/2$, (ii) monotonically decreasing (increasing) for $\lambda \in [0, 1/2]$ ($\lambda \in [1/2, 1]$).*

Proof. The proof of (i) (symmetry) is immediate by looking at expression (17), noticing property (i) of lemma 1 and the obvious fact that p_{σ_L} is symmetric to $1 - p_{\sigma_L}$. In order to prove monotonicity, I compute the expression of the derivative of Γ with respect to λ:

$$\begin{aligned}
\Gamma_\lambda \equiv \frac{d\Gamma}{d\lambda} =& \pi\theta_L \frac{d\left(\lambda\theta(\mu_{\sigma_L}^i) + (1-\lambda)\theta(\mu_H^i)\right)}{d\lambda} + \\
& (1-\pi)\theta_H \frac{d\left((1-\lambda)\theta(\mu_{\sigma_L}^i) + \lambda\theta(\mu_H^i)\right)}{d\lambda} = \\
& \pi\theta_L \left(-\frac{\pi(\pi-1)^2(\theta_H - \theta_L)(2\lambda - 1)}{(2\pi\lambda + 1 - \lambda - \pi)^2(2\pi\lambda - \lambda - \pi)^2}\right) + \\
& (1-\pi)\theta_H \left(-\frac{\pi^2(\pi-1)(\theta_H - \theta_L)(2\lambda - 1)}{(2\pi\lambda + 1 - \lambda - \pi)^2(2\pi\lambda - \lambda - \pi)^2}\right) = \\
& \frac{\pi^2(1-\pi)^2(\theta_H - \theta_L)^2(2\lambda - 1)}{(2\pi\lambda + 1 - \lambda - \pi)^2(2\pi\lambda - \lambda - \pi)^2},
\end{aligned}$$

which is ≥ 0 if and only if $\lambda \geq 1/2$. □

The intuition behind this result is very important. (16) is a measure of aggregate output when effort is not diminished by taxation. Lemma 2 shows that when the

---

[11]Up to the point that $\theta(\mu_{\sigma_L}) = \bar{\theta}$ for $\lambda = 1/2$, i.e. when the signal is uninformative and the posterior belief coincides with the prior belief.

[12]Up to the point that $\theta(\mu_{\sigma_L}) = \theta_L$ for $\lambda = 1$.

incentive-distortive effect of taxation is not taken into account, increasing the informativeness of the signal has a positive effect on aggregate output as effort is chosen more optimally given the true values of θ_L and θ_H.

I study the cases of $\pi > 1/2$ and $\pi < 1/2$ separately and I present the results of the comparative statics.

Proposition 2. *If $\pi > 1/2$, τ is a continuous map in λ which is (i) symmetric with respect to $\lambda = 1/2$ and (ii) monotonically decreasing (increasing) for $\lambda \in [0, 1/2]$ ($\lambda \in [1/2, 1]$).*

Proof. Once it is noticed that given $\pi > 1/2$, $p_{\sigma_L} \leq (\geq) 1/2$ for $\lambda \in [0, 1/2]$ ($\lambda \in [1/2, 1]$) and therefore in expression (19) $\theta(\mu) = \theta(\mu_{\sigma_H})$ ($\theta(\mu) = \theta(\mu_{\sigma_L})$) for $\lambda \in [0, 1/2]$ ($\lambda \in [1/2, 1]$), the proof follows in a straightforward way from lemmas 1 and 2. $\square$

Notice that τ is then minimized for $\lambda = 1/2$, when $\tau = 0$ and it is maximized for $\lambda = 0$ and $\lambda = 1$, when $\tau = 1 - \frac{1}{2-(\theta_L^2/\overline{\theta^2})}$. Notice that for $\pi \in [0, 1)$, $\theta_L^2/\overline{\theta^2} < 1$ and hence $\tau \in [0, 1)$.

It is also important to study the comparative statics which are relative to effort. Using (12) I can define the optimal effort implemented by those who observe σ_L:

$$(20) \qquad\qquad e|\sigma_L \equiv (1 - \tau)\theta(\mu_{\sigma_L})/a,$$

and by those who observe σ_H:

$$(21) \qquad\qquad e|\sigma_H \equiv (1 - \tau)\theta(\mu_{\sigma_H})/a.$$

Notice that $e|\sigma_L$ and $e|\sigma_H$ are symmetric to each other as $\theta(\mu_{\sigma_L})$ and $\theta(\mu_{\sigma_H})$ are symmetric to each other as exposed in Lemma 1. Multiplying by the respective weights I obtain the expression of aggregate effort:

$$(22) \qquad\qquad \bar{e} = (1 - \tau)(p_{\sigma_L}\theta(\mu_{\sigma_L}) + (1 - p_{\sigma_L})\theta(\mu_{\sigma_H}))/a,$$

where it is easy to compute that $p_{\sigma_L}\theta(\mu_{\sigma_L}) + (1 - p_{\sigma_L})\theta(\mu_{\sigma_H}) = \overline{\theta}$. A proposition follows:

Proposition 3. *If $\pi > 1/2$, expression (22) is a continuous map in λ which is (i) symmetric with respect to $\lambda = 1/2$ and (ii) monotonically increasing (decreasing) for $\lambda \in [0, 1/2]$ ($\lambda \in [1/2, 1]$).*

The proof follows trivially from proposition 2 and from the fact that $\overline{\theta}$ is a constant. $\bar{e}$ is then maximized for $\lambda = 1/2$, with $\bar{e} = \overline{\theta}$ and it is minimized for $\lambda = 0$ and $\lambda = 1$ with $\bar{e} = \frac{\overline{\theta^2}\overline{\theta}}{2\overline{\theta^2} - \theta_L^2}$. This is an important result as it shows that an uninformative signal

maximizes the aggregate effort. This is because the only way in which the signal enters in the expression of aggregate effort is through the tax rate. Once that this effect on taxation is taken into account, the signal does not play any role on aggregate effort as its effect on the two groups is exactly symmetric.

The effect of λ on $e|\sigma_L$ and $e|\sigma_H$ is instead partially ambiguous. In the case of $\pi > 1/2$, τ increases (decreases) in λ for $\lambda \geq 1/2$ ($\lambda \leq 1/2$) whereas $\theta(\mu_{\sigma_L})$ ($\theta(\mu_{\sigma_H})$) monotonically decreases (increases) in λ. The overall effect depends on how responsive are τ and $\theta(\mu)$ to λ. The only unambiguous path is that $e|\sigma_L$ decreases in λ for $\lambda \geq 1/2$, as both $(1 - \tau)$ and $\theta(\mu_{\sigma_L})$ decrease; and by symmetry that $e|\sigma_H$ increases in λ for $\lambda \leq 1/2$, as both $(1-\tau)$ and $\theta(\mu_{\sigma_L})$ increase. I summarize this last point describing the comparative static of $e|\sigma_L$ and therefore that of $e|\sigma_H$ by symmetry. For $\lambda = 0$, $e|\sigma_L = \frac{\overline{\theta^2}\theta_H^2}{2\overline{\theta^2}-\theta_L^2}$. For $\lambda = 1/2$, $e|\sigma_L = \overline{\theta}^2$. Given different values of the parameters, $e|\sigma_L$ for $\lambda = 0$ can be greater or smaller than $e|\sigma_L$ for $\lambda = 1/2$ and the path between the two values does not have to be monotonic. $e|\sigma_L$ decreases monotonically between $\lambda = 1/2$ and $\lambda = 1$ and $e|\sigma_L = \frac{\overline{\theta^2}\theta_L^2}{2\overline{\theta^2}-\theta_L^2}$ for $\lambda = 1$.

Plugging (15) into (1) I obtain the expression of aggregate output:

(23) $$\bar{y} = k + (1 - \tau)\Gamma/a.$$

Notice that for $\pi > 1/2$ the effect of λ is not a-priori clear as given lemma 2 and proposition 2, λ has opposite effects on $(1 - \tau)$ and Γ.

Proposition 4. *If $\pi > 1/2$, expression (23) is a continuous map in λ which is*

(i) symmetric with respect to $\lambda = 1/2$,

(ii) either monotonically decreasing or monotonically decreasing up to a point and then monotonically increasing for $\lambda \in [1/2, 1]$ (obviously the behavior is symmetric for $\lambda \in [0, 1/2])$ [13] *and*

(iii) maximized for $\lambda = 1/2$.

The proof is in Appendix A.

This is a striking policy result: aggregate output is univocally maximized by a completely uninformative signal. Even if λ has opposite effects on $(1-\tau)$ and Γ, the effect through the tax rate is stronger. The value of the aggregate output for $\lambda = 1/2$ is $\bar{y} = k+\overline{\theta}^2/a$, the value of the aggregate output for $\lambda = 0$ and $\lambda = 1$ is $\bar{y} = k+\frac{\overline{\theta^2}^2}{a(2\overline{\theta^2}-\theta_L^2)} > 0$.

---

[13]Hence in the non-monotonic case, (23) is quasi–convex in the separate sub-domains $\lambda \in [0, 1/2]$ and $\lambda \in [1/2, 1]$.

I analyze the case of $\pi < 1/2$. The only difference with respect to the previous case is that $\tau = \tau(\mu_{\sigma_H})$ for $\lambda \in [1/2, 1]$ and $\tau = \tau(\mu_{\sigma_L})$ for $\lambda \in [0, 1/2]$. Instead, the expressions of τ, $\bar{e}$, $e|\sigma_L$, $e|\sigma_H$, $\bar{y}$ are still given by expressions (19), (22), (20), (21), (23), so they are still symmetric with respect to $\lambda = 1/2$.

In this case the comparative statics of τ with respect to λ are generally non-monotonic. To see this notice that in expression (19) both $\theta(\mu)$ and Γ increase for $\lambda \in [1/2, 1]$ and so the overall effect of λ is not a-priori clear. Nevertheless it is possible to find some properties:

Proposition 5. *The tax rate is always negative and, if* $(2\Gamma \frac{\partial \theta(\sigma_H)}{\partial \lambda}) < \theta(\sigma_H) \frac{\partial \Gamma}{\partial \lambda}$, *it is monotonically decreasing (increasing) for* $\lambda \in [1/2, 1]$ ($\lambda \in [0, 1/2]$).

Proof. I consider the case of $\lambda \in [1/2, 1]$, by symmetry it is sufficient to study this case. It is useful to re-express (19) as

$$(24) \qquad \tau = \frac{\Gamma - \theta(\sigma_H)^2}{2\Gamma - \theta(\sigma_H)^2}.$$

Notice that the numerator of (24) is always negative because $\Gamma = p_{\sigma_L} \theta(\mu_{\sigma_L})^2 + (1 - p_{\sigma_L}) \theta(\mu_{\sigma_H})^2 < \theta(\mu_{\sigma_H})^2$ for $\lambda \in [1/2, 1]$, as it is the case that $\theta(\mu_{\sigma_H}) > \theta(\mu_{\sigma_L})$. The denominator is always positive under assumption 1. This proves the negativity of the expression. To prove the monotonicity I notice that the first derivative of τ with respect to λ is $\tau_\lambda = \frac{\theta(\sigma_H)\left(2\Gamma \frac{\partial \theta(\sigma_H)}{\partial \lambda} - \theta(\sigma_H) \frac{\partial \Gamma}{\partial \lambda}\right)}{(2\Gamma^2 - \theta(\mu_{\sigma_H}))^2}$. Given lemmas 1 and 2, a sufficient condition for τ to be monotonic decreasing is therefore that $(2\Gamma \frac{\partial \theta(\sigma_H)}{\partial \lambda}) < \theta(\sigma_H) \frac{\partial \Gamma}{\partial \lambda}$. $\qquad\square$

When τ decreases (increases) monotonically for $\lambda \in [1/2, 1]$ ($\lambda \in [0, 1/2]$) it is straightforward that expression (23), lemma 2 and proposition 5 imply that aggregate output increases (decreases) monotonically for $\lambda \in [1/2, 1]$ ($\lambda \in [0, 1/2]$). Moreover, given that the tax rate is always negative, then the aggregate output is always greater than in the case of $\pi \geq 1/2$, even when τ does not have any monotonic behavior and therefore the behavior of the output does not have to be monotonic.

Aggregate effort still depends exclusively on the tax rate, therefore I conclude that when τ decreases (increases) monotonically for $\lambda \in [1/2, 1]$ ($\lambda \in [0, 1/2]$) it is straightforward that expression (22) implies that aggregate effort increases (decreases) monotonically for $\lambda \in [1/2, 1]$ ($\lambda \in [0, 1/2]$). Moreover, given that the tax rate is always negative, then the aggregate effort is always greater than in the case of $\pi \geq 1/2$, even when τ does not have any monotonic behavior and therefore the behavior of the aggregate effort does not have to be monotonic.

The effect of λ on $e|\sigma_L$ and $e|\sigma_H$ is instead partially ambiguous as in the case of $\pi > 1/2$ τ increases (decreases) in λ for $\lambda \geq 1/2$ ($\lambda \leq 1/2$) whereas $\theta(\mu_{\sigma_L})$ ($\theta(\mu_{\sigma_H})$) monotonically decreases (increases) in λ. The overall effect depends on how responsive are τ and $\theta(\mu)$ to λ. The only unambiguous path is that $e|\sigma_L$ decreases in λ for $\lambda \geq 1/2$ as both $(1-\tau)$ and $\theta(\mu_{\sigma_L})$ decrease and that $e|\sigma_H$ increases in λ for $\lambda \leq 1/2$ as both $(1-\tau)$ and $\theta(\mu_{\sigma_L})$ increase. I summarize describing the comparative static of $e|\sigma_L$ and by symmetry that of $e|\sigma_H$. For $\lambda = 0$, $e|\sigma_L = \frac{\overline{\theta^2}\theta_H^2}{2\overline{\theta^2}-\theta_L^2}$. For $\lambda = 1/2$, $e|\sigma_L = \overline{\theta}^2$. Given different values of the parameters, $e|\sigma_L$ for $\lambda = 0$ can be greater or smaller than $e|\sigma_L$ for $\lambda = 1/2$ and the path between the two values does not have to be monotonic. $e|\sigma_L$ decreases monotonically between $\lambda = 1/2$ and $\lambda = 1$ and $e|\sigma_L = \frac{\overline{\theta^2}\theta_L^2}{2\overline{\theta^2}-\theta_L^2}$ for $\lambda = 1$.

In the case in which τ decreases (increases) monotonically for $\lambda \in [1/2, 1]$ ($\lambda \in [0, 1/2]$), the only unambiguous paths is that $e|\sigma_L$ decreases in λ for $\lambda \leq 1/2$, as both $(1-\tau)$ and that $\theta(\mu_{\sigma_L})$ decrease; by symmetry, $e|\sigma_H$ increases in λ for $\lambda \geq 1/2$, as both $(1-\tau)$ and $\theta(\mu_{\sigma_H})$ increase. The effect of λ on $e|\sigma_L$ and $e|\sigma_H$ is instead partially ambiguous, as τ increases (decreases) in λ for $\lambda \leq 1/2$ ($\lambda \geq 1/2$), whereas $\theta(\mu_{\sigma_L})$ ($\theta(\mu_{\sigma_H})$) monotonically decreases (increases) in λ. The overall effect depends on how responsive are τ and $\theta(\mu)$ to λ.

5. OPTIMAL INFORMATION

In the previous section I studied different comparative statics and the results offer insights for policy questions such as which is the level of information which maximizes output or how does the level of information affect the prevailing the tax rate. It is now a natural question to ask which is the level of information preferred by agents.

Before analyzing the preferred value of λ for the society as a whole, I temporarily depart from the set up assuming that each agent i at $t = 0$ can individually chose the optimal precision λ^i of the signal to be observed at $t = 1$ by herself. In this case the optimal value of λ^i would maximize the expected utility at $t = 0$ taking the choices of the other agents as given. Notice that at $t = 0$ everyone is identical. In order to compute the expression of the expected utility at period 0, I compute the following terms:

(25) $$E[e^i\theta^i|I_0^i] = E[\overline{e\theta}|I_0^i] = E[\overline{e\theta}|I_1^i] = (1-\tau)\Gamma/a,$$

(26)
$$E[(e^i)^2|I_0^i] = \frac{(1-\tau)^2\Gamma}{a}.$$

Plugging (25) and (26) into (11) I obtain the problem at $t = 0$:

(27) $\quad \lambda^i = \arg \max\{(1-\tau)(\bar{k} + (1-\tau(\lambda))\Gamma(\lambda^i)/a)+$
$$\tau(\lambda)(\bar{k} + (1-\tau(\lambda))\Gamma(\lambda)/a) - (1-\tau(\lambda))^2\Gamma(\lambda^i)/2a\}.$$

As a single individual cannot influence the prevailing tax rate, this it taken as given when the optimal λ^i is chosen. The problem has an easy solution because λ^i only influences the object through $\Gamma(\lambda^i)$. Therefore given lemma 2, the solution which is a perfect informative signal, either $\lambda = 0$ or $\lambda = 1$. Going back to the basic set-up, another question is more interesting: what is the optimal value of λ for the society as a whole at $t = 0$? In other words I investigate whether someone behind the veil of ignorance desires to leave in a world beyond or behind the veil.

Plugging (25) and (26) into (11) and rearranging I obtain the expected utility at $t = 0$:

(28)
$$u_0^i = k + (1-\tau^2)\Gamma/2a.$$

Notice that (28) is symmetric with respect to $\lambda = 1/2$ as both τ and Γ are symmetric with respect to $\lambda = 1/2$. If an agent had to choose an optimal value of λ for the society at $t = 0$, he would choose a value of λ which maximizes (28). The solution of the problem is not a-priori trivial. In the case of $\pi < 1/2$ it is possible that τ does not have a monotonic behavior and this makes the effect of λ on (28) not clear. In the case in which the condition for the monotonic property explained in proposition (5) applies, then both $(1-\tau^2)$ and Γ increase in λ for $\lambda \in [1/2, 1]$. Hence (28) is maximized for $\lambda = 0$ or $\lambda = 1$, i.e. for a perfectly informative signal.

In the case of $\pi > 1/2$, lemma 2 and proposition 2 show that λ has opposite effects on $(1-\tau^2)$ and Γ so the overall effect is not a-priori clear. Nevertheless I find an interesting property:

Proposition 6. *Given $\pi > 1/2$, expression (28) is either monotonically decreasing or monotonically decreasing up to a point and then monotonically increasing for $\lambda \in [1/2, 1]$ (and obviously the symmetric behavior applies for $\lambda \in [0, 1/2])$ [14].*

---

[14]Hence in the non-monotonic case, (23) is quasi–convex in the sub-sets $\lambda \in [0, 1/2]$ and $\lambda \in [1/2, 1]$.

Proof. (28) can be rewritten as $k + (1 + \tau)(1 - \tau)\Gamma/2a$. Notice that the derivative of $(1 - \tau)\Gamma$ has already been studied in proposition (4). I rename $(1 + \tau) = a(\lambda)$ and $(1 - \tau)\Gamma = b(\lambda)$, where $a(\lambda)$ and $b(\lambda)$ are functions of λ. I analyze the interval $\lambda \in [1/2, 1]$ and I compute $\frac{d(a(\lambda)b(\lambda))}{d\lambda} = \frac{da}{d\lambda}b + a\frac{db}{d\lambda}$. Notice that $\frac{da}{d\lambda}b < 0$ and that $a\frac{db}{d\lambda}$ can change sign and become positive at most once, hence the entire expression is either always negative or it can change sign and become positive at most once.   □

This result implies that in the case of $\pi > 1/2$, the solution is either ($\lambda = 1/2$) or ($\lambda = 0$, $\lambda = 1$). The result is interesting because it shows that the ex-ante optimal level of information for the economy is either a completely uninformative signal ($\lambda = 1/2$) or a completely informative signal ($\lambda = 0$ or $\lambda = 1$). In other words agents either want to stay behind the veil of ignorance or to remove it completely.

6. POLITICO–ECONOMIC EQUILIBRIUM

In this section I endogenize the value of λ. I use the concept introduced in the previous section and I consider that the prevailing value of λ is the one which maximizes the ex-ante utility. Such value of the parameter λ could be implemented by a benevolent planner, it could be a voting outcome or it could be the outcome of any other collective choice. Being everyone ex-ante identical, as long as the optimal λ is computed at $t = 0$, everyone chooses λ in order to maximize the same object. A definition follows:

Definition 1. I define a **Politico–Economic Equilibrium** as a vector $(\lambda, \mu_{\sigma_L}, \mu_{\sigma_H}, \tau)$ such that

(i) λ = arg max u_0,
(ii) Beliefs μ_{σ_L} and μ_{σ_H} are respectively given by (4) and (5),
(iii) The prevailing tax rate τ is given by (19) where i is the median voter.

I analyze the case of $\pi > 1/2$. The results of the previous section show that for $\lambda > 1/2$ ($\lambda < 1/2$) τ increases (decreases), while Γ decreases (increases) in λ, therefore the overall effect of λ on (28) is not a-priori clear. It is easy to construct numerical examples of different comparative statics and it is easy to present an example in which u_0 has three global maxima for $\lambda = 0$, $\lambda = 1/2$, $\lambda = 1$.[15] In the case in which the parameters of the model are changed and τ is made more (less) responsive to λ

---

[15]Plugging $\lambda = 1/2$ and $\lambda = 1$ in (28) it is straightforward to compute that the set of parameters such that both $\lambda = 1/2$ and $\lambda = 0$, $\lambda = 1$ are optimal is given by those parameters such that $1 - \frac{(\bar{\theta}^2 - \theta_L^2)^2}{(2\theta^2 - \theta_L^2)^2} =$
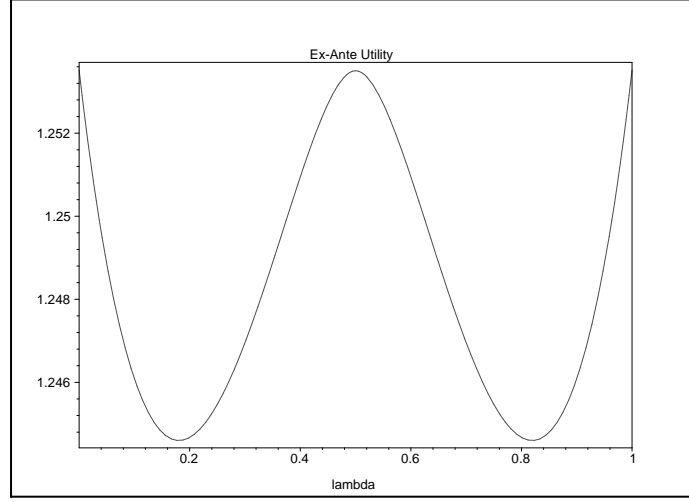
FIGURE 1. Welfare for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

than Γ,[16] the effect of τ (Γ) becomes dominant and therefore u_0 is maximized when τ (Γ) is minimized (maximized), namely for $\lambda = 1/2$ ($\lambda = 0$, $\lambda = 1$).

Example of Multiple Equilibria. Using `Maple`® I present an example of multiple Politico-Economic Equilibria.

Consider $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 0.5$, $k = 0$. I plug those values in (4), (5), (7) (19), (16) and consequently those expressions in (28). I obtain a map of u_0^i in λ, which I plot in figure 1.

I verify that the function has three global maxima for $\lambda = 0$, $\lambda = 1/2$, $\lambda = 1$ with value 1.25352. For $\lambda = 0$ and $\lambda = 1$, the signal is perfectly informative. For $\lambda = 1/2$ the signal is completely uninformative. Both perfect information and minimum information are ex-ante optimal for the society. In order to interpret the result I plot τ and Γ in figures 2 and 3 respectively. As figure 2 shows, τ is maximized for $\lambda = 0$ and $\lambda = 1$ and it is minimized for $\lambda = 1/2$. As figure 3 shows, Γ is maximized for $\lambda = 0$ and $\lambda = 1$ and it is minimized for $\lambda = 1/2$. Expressions (25) and (26) show that Γ measures aggregate output or aggregate squared effort, if optimal effort is not diminished by τ. Expression (28) increases in Γ and decreases in τ. The uninformative equilibrium is characterized by a lower τ and an higher Γ than the informative

---

$\overline{\theta^2}$, i.e. $2\pi\theta_L{}^2 + 2\theta_L\theta_H - 4\pi\theta_L\theta_H - 2\theta_H{}^2 + 2\theta_H{}^2\pi =$
$\left(1 - \frac{\left(\pi\theta_L{}^2 + (1-\pi)\theta_H{}^2 - \theta_L{}^2\right)^2}{\left(2\pi\theta_L{}^2 + 2(1-\pi)\theta_H{}^2 - \theta_L{}^2\right)^2}\right)\left(\pi\theta_L{}^2 + (1-\pi)\theta_H{}^2\right)$.

[16]This can be done increasing (decreasing) π or increasing (decreasing) the difference between θ_H and θ_L.
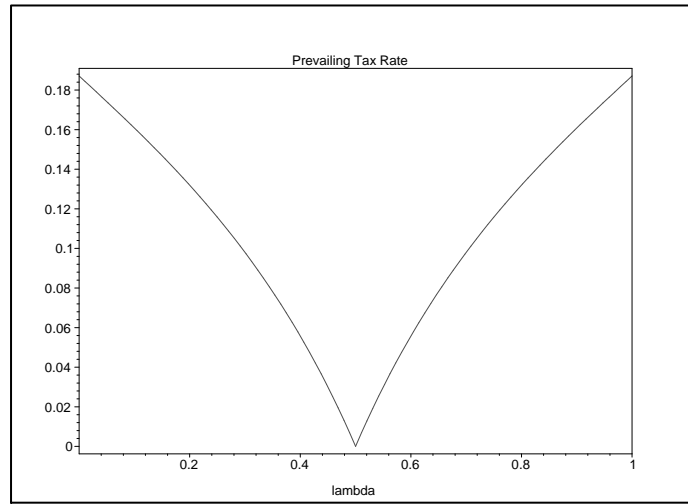
FIGURE 2. τ for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.
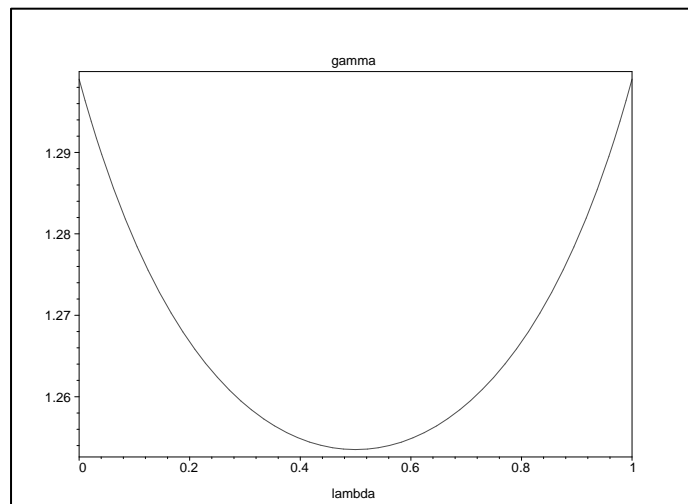

FIGURE 3. Γ for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

equilibrium. The two effects work in opposite directions, hence the multiple equilibria. I can further interpret this result plotting the expressions of the effort exerted by those who observe σ_L and σ_H, as functions of λ, in figures 4 and 5 respectively. I also plot the expression of aggregate effort[17] in figure 6. The expression of optimal effort (12) shows that the greater is τ and the lower is the optimal effort, hence a problem of moral hazard follows: the informative equilibrium is characterized by

---

[17] Namely the weighted sum of effort exerted by those who observe σ_L plus effort exerted by those who observe σ_H with weights p_{σ_L} and $1 - p_{\sigma_L}$ respectively.
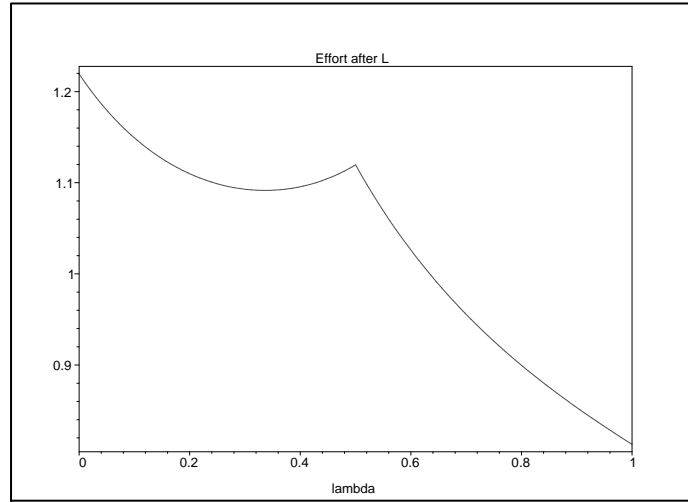
20

FIGURE 4. Effort given σ_L for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

a severe moral hazard problem as τ is at the maximum level. Instead, the uninformative equilibrium is characterized by a non severe moral hazard problem as τ is at the minimum level. It is less immediate to notice an opposite effect of adverse selection. Figures 4 and 5 show that the greater is the precision of the signal and the more different is the level of effort exerted by the two groups. When the signal is completely uninformative everyone chooses the same level of effort (pooling equilibrium), whereas when the signal is perfectly informative the highly productive choose the maximum level of effort and the low productive choose the minimum value of effort (separating equilibrium).[18] Figure 6 shows that aggregate effort is maximized at the uninformative equilibrium which is characterized by severe adverse selection and no moral hazard.

I plot the expression of aggregate output (23) as a function of λ in figure 7 and this shows that also aggregate output is maximized at the uninformative equilibrium.

The uninformative equilibrium can be interpreted as a US-type equilibrium. In this equilibrium agents have wrong beliefs about the real return on effort. Both groups of agents hold the same belief and exert the same effort (pooling equilibrium); in particular the low productive ones are biased towards optimism as they believe to be more productive than what they truly are. The tax rate is at the minimum level, aggregate effort and output are at the maximum level. The informative equilibrium can be interpreted as a Europe-type equilibrium. In this equilibrium the

---

[18]Notice that when $\lambda = 0$ those that observe the signal σ_L are those with productivity θ_H and viceversa.
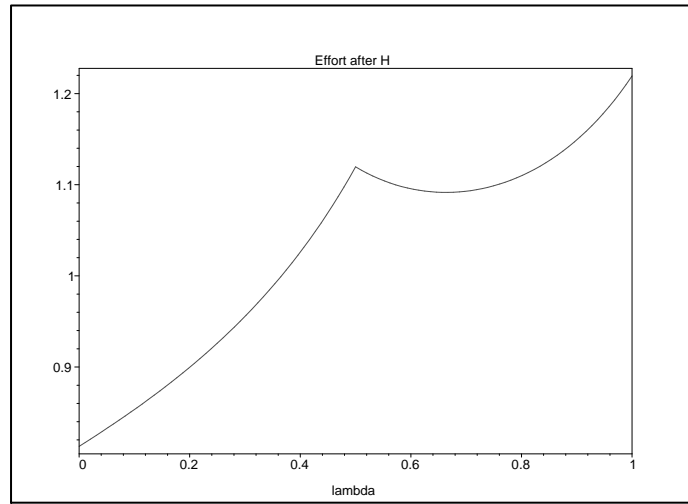
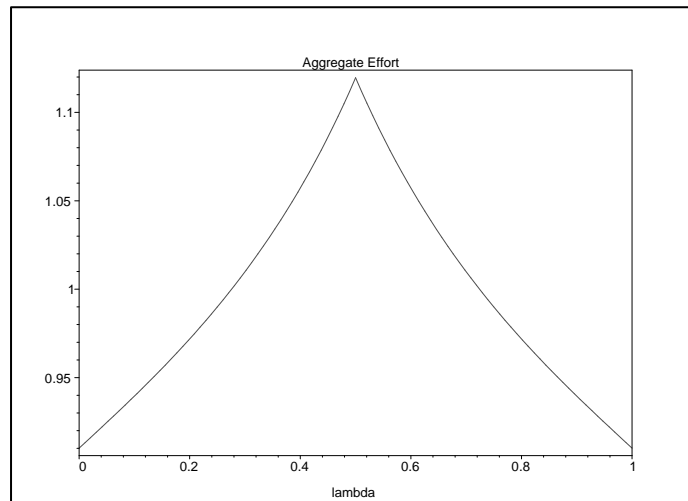FIGURE 5. Effort given σ_H for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

FIGURE 6. Aggregate effort for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

two groups of agents have correct beliefs about the real return on effort. As the low productive agents are the majority, their preferred tax rate is the prevailing in the economy, hence the level of redistribution is higher than in the US-type equilibrium. High redistribution and correct beliefs about the return on effort imply that the low productive ones minimize the effort whereas the high productive ones maximize it (separating equilibrium). This results in lower aggregate effort and aggregate output than those at the uninformative equilibrium.
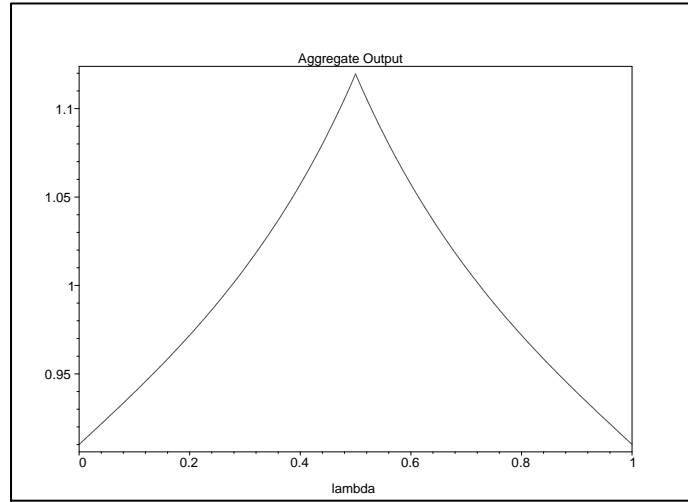
FIGURE 7. Aggregate Output for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

7. INTERPRETATION AND GENERALIZATION OF THE RESULTS

The result of the last section does not have to be interpreted as stating that Europeans are perfectly informed or that Americans are poorly informed. The first consideration to be made is that it does not have to be the case that in the Europe-type equilibrium the optimal value of λ is either 0 and 1 so that the signal is fully informative; in Appendix B, taking into account the possibility of heterogeneous endowments, I show that in both types of equilibria it is possible to have intermediate optimal λ's. The second consideration to be made is that the same results in terms of multiple equilibria would follow with a different underlying true distribution of the θ's. For example, take a case in which the true distribution of the θ's is very complicated and all that agents know is only that with probability π $(1 - \pi)$ the average value of θ^i is θ_L (θ_H). If the structure of the signal is still the one in (2), then the problem is the same – this can be seen from the fact that the expressions (6) and (16) do not change – hence the same results apply. Or again the same results would apply in the case of homogeneous returns and aggregate macroeconomic shocks: $\theta^i = \theta$ for all i and again all that agents know is that and with probability π $(1 - \pi)$ the average value of θ is θ_L (θ_H). The fact that the true distribution of the θ's can be unknown shows that a more precise signal in the Europe-type equilibrium does not mean that Europeans get to know the truth whereas Americans do not.

In order to interpret the result about the existence of multiple equilibria correctly it is necessary to understand the key-driver of the result. Going back to expression

(28), it is clear that the fact that there may be multiple optimal values of λ – and therefore multiple equilibria – comes from the non-monotonic effect of λ on $(1 - \tau^2)\Gamma$. In particular the information structure in (2) implies lemma 2 and therefore that the more precise is the signal λ and the greater is Γ. This consideration helps to understand the following general result:

Theorem 1. *Given the ex-ante objective function (28), if $\tau \in [0, 1]$ is part of a politico-economic equilibrium, then the greater is τ and the greater is the precision of the signal in the equilibrium.*

Proof. In a politico-economic equilibrium, $\lambda = \arg \max (1 - \tau^2)\Gamma$. Assume without loss of generality that two different $\lambda's$ are part of a different equilibria with $\lambda' > \lambda'' > 1/2$. Given lemma 2, this implies that $\Gamma(\lambda') > \Gamma(\lambda'')$ and therefore that $\tau(\lambda') > \tau(\lambda'')$. $\qquad\qquad\square$

The result shows that if multiple equilibria exist then it must be the case that ex-ante there is a trade-off in increasing the precision of the signal: increasing the precision of the signal increases Γ, but increasing the precision of the signal can also increase τ. Hence, when the effect of λ on the object (28) is non-monotonic, then multiple equilibria are possible. In economic terms the trade-off is between the positive effect of an increase in the precision of the signal, namely that more information reduces adverse selection as agents choose effort more optimally given their abilities, and the negative effect, namely that more information can increase the prevailing tax rate and this creates a moral hazard effect which reduces aggregate effort. Theorem 1 shows that in the case of multiple equilibria, a US-(Europe-) type equilibrium is relatively characterized by: (i) a less (more) informative signal and therefore (ii) less (more) separated beliefs and individual levels of effort implemented, (iii) lower (higher) redistribution and therefore (iv) higher (lower) aggregate effort and output. In other words, the result states that the case of multiple equilibria is a case in which an economy relatively characterized by more adverse selection and less moral hazard is ex-ante equally optimal to another one characterized by less adverse selection and more moral hazard. This result is general and robust. It does not depend on the heterogeneity of endowments[19] or on the underlying distribution of the abilities

---

[19]As it is shown in appendix B, with heterogenous endowments technical difficulties arise because changing the level of information changes the identity of the median voter, different voters prefer different tax rates given different endowments and hence the comparative statics are generally discontinuous. Nevertheless the ex-ante optimal λ still has to maximize the object (28) and therefore theorem 1 still applies.

because those feature do not change the ex-ante problem. Even if the the set up of the model does not allow further generalizations, it can be conjectured that given that what drives the possible ex-ante optimality of different level of information is the trade off between more hazard effect and adverse selection, then the same multiple equilibria could in principle exist also in a different set-up with concave utility in wealth[20], or with a completely different information structure[21].

8. CONCLUSION

The aim of the paper is to provide a simple theoretical model to analyze the role of incomplete information in the determination of heterogeneous beliefs and different politico-economic equilibria.

Different comparative statics can be studied with this model and the results can be used in order to find the optimal level of information for different objectives as the maximization of aggregate output or of aggregate welfare.

The theoretical model presented in the paper interprets the US-type vs the Europe-type politico-economic equilibrium as characterized by relative (i) high adverse selection – individual beliefs and effort levels are pooling to similar levels despite underlying heterogeneity in the true distribution of the return on effort and this creates inefficiencies – (ii) low redistribution (iii) low moral hazard – redistribution is low and this does not distort individual effort much (iv) high aggregate effort and output. Conversely the Europe-type politico-economic equilibrium is interpreted as an equilibrium characterized by relative (i) low adverse selection (ii) high redistribution (iii) high moral hazard – taxation is high and this diminish individual effort (iv) low aggregate effort and output. The two equilibria are both ex ante optimal. This result is robust to variations of the basic framework.

REFERENCES

ALESINA, A., AND G.-M. ANGELETOS (2005): "Fairness and Redistribution: US versus Europe," *American Economic Review*, 95, 960–980.

ALESINA, A., AND E. GLAESER (2004): *Fighting Poverty in the US and Europe: A World of Difference*. Oxford University Press, Oxford, UK.

---

[20]The ex-ante preferred taxation rate would not be any longer equal to zero but it would not even have to be as large as $\tau = 1$, hence still creating room for a moral hazard effect of an increase in information.

[21]Provided of course, as it seems logical, that more information still reduces the adverse selection problem.

ALESINA, A., E. GLAESER, AND B. SACERDOTE (2001): "Why Doesn't the US Have a European-Type Welfare State?," *Brookings Papers on Economic Activity*, 2, 187–277.

BENABOU, R., AND J. TIROLE (2006): "Belief in a Just World and Redistributive Politics," *Quarterly Journal of Economics*, 121, 699–746.

BISHOP, J. H. (1996): *Signaling, Incentives, and School Organization in France, the Netherlands, Britain, and the United States.* In: Hanushek E. A. and D. W. Jorgenson, eds., Improving America's Schools: The Role of Incentives.chap. 7, pp. 111–146. National Academy Press, Washington DC.

CERVELLATI, M., J. ESTEBAN, AND L. KRANICH (2006): "Redistributive Taxation with Endogenous Sentiments," IZA Discussion Papers 2312, Institute for the Study of Labor (IZA).

KEELY, L. C. (2002): "Mobility and Fertility," University of Wisconsin-Madison, mimeo.

LADD, E., AND K. BOWMAN (1998): *Attitudes Towards Economic Inequality*. American Enterprise Institute Studies on Understanding Economic Inequality, Washington, DC.

MELTZER, A. H., AND S. F. RICHARD (1981): "A Rational Theory of the Size of Government," *Journal of Political Economy*, 89, 914–927.

PIKETTY, T. (1995): "Social Mobility and Redistributive Politics," *Quarterly Journal of Economics*, 110, 551–584.

APPENDIX A. PROOF OF PROPOSITION 4

The symmetry follows trivially from the symmetry of Γ and τ, respectively proved in lemma 2 and proposition 2. Given the symmetry of $\bar{y}$ it is enough to study the quasi-convexity in the interval $\lambda \in [1/2, 1]$. It is useful to plug (19) into (23) and re-express this as

$$
(29) \qquad k + \frac{\Gamma^2}{a(2\Gamma - \theta(\mu)^2)},
$$

where, given that $\lambda \in [1/2, 1]$, $\theta(\mu) = \theta(\mu_{\sigma_L})$. I compute the first derivative of this expression with respect to λ:

$$
(30) \qquad \frac{2\Gamma^2 \frac{\partial\Gamma}{\partial\lambda} - 2\theta(\mu_{\sigma_L})^2\Gamma\frac{\partial\Gamma}{\partial\lambda} + 2\theta(\mu_{\sigma_L})\Gamma^2\frac{\partial\theta(\mu_{\sigma_L})}{\partial\lambda}}{a^2\left(2\Gamma - \theta(\mu_{\sigma_L})^2\right)^2}
$$

where

$$
\frac{\partial\theta(\mu_{\sigma_L})}{\partial\lambda} = -\frac{\pi(1-\pi)(\theta_H - \theta_L)}{(\pi\lambda + (1-\lambda)(1-\pi))^2} \leq 0
$$

$$
\frac{\partial\Gamma}{\partial\lambda} = \frac{\pi^2(1-\pi)^2(2\lambda-1)(\theta_H - \theta_L)^2}{(\pi\lambda + (1-\pi)(1-\lambda))^2(\pi(\lambda-1) + \lambda(\pi-1))^2} \geq 0.
$$

(31)

The denominator of (30) is positive, so the sign of the numerator determines the sign of the entire expression. I can divide the numerator by 2Γ which is a positive quantity and the numerator reduces to

$$
(32) \qquad (\Gamma - \theta(\mu_{\sigma_L})^2)\frac{\partial\Gamma}{\partial\lambda} + \theta(\mu_{\sigma_L})\Gamma\frac{\partial\theta(\mu_{\sigma_L})}{\partial\lambda}.
$$

The value of this last expression for $\lambda = 1/2$ is $-4\pi(1-\pi)(\theta_H - \theta_L)(\pi\theta_L + (1-\pi)\theta_H)$ which is negative, hence I conclude that (30) is negative for $\lambda = 1/2$. I compute the second derivative of (32):

$$
(33) \quad (\Gamma - \theta(\mu_{\sigma_L})^2)d^2\Gamma + (d\Gamma)^2 - \theta(\mu_{\sigma_L})d\theta(\mu_{\sigma_L})d\Gamma + \Gamma(d\theta(\mu_{\sigma_L}))^2 + \theta(\mu_{\sigma_L})\Gamma d^2\theta(\mu_{\sigma_L}),
$$

where

$$\frac{\partial^2 \theta(\mu_{\sigma_L})}{(\partial \lambda)^2} = \frac{2\pi(1-\pi)(2\pi-1)(\theta_H - \theta_L)}{(\pi\lambda + (1-\pi)(1-\lambda))^3} \geq 0$$

$$\frac{\partial^2 \Gamma}{(\partial \lambda)^2} = \frac{2\pi^2(1-\pi)^2(\theta_H - \theta_L)^2(1 + 12\pi\lambda(1-\lambda)(1-\pi) - 3\pi(1-\pi) - 3\lambda(1-\lambda))}{(\pi\lambda + (1-\pi)(1-\lambda))^3(\pi(\lambda-1) + \lambda(\pi-1))^3}.$$

(34)

Notice that $\frac{\partial^2 \Gamma}{(\partial \lambda)^2}$ is positive as it can be proved that $(1 + 12\pi\lambda(1-\lambda)(1-\pi) - 3\pi(1-\pi) - 3\lambda(1-\lambda))$ is strictly positive. To see this, compute the first derivative with respect to λ which is equal to $3(2\pi-1)^2(2\lambda-1)$ and so positive in the interval of $\pi \in (1/2, 1]$ which it is considered. Therefore the expression increases in λ; it is immediate that it is equal to zero for the smallest value of λ in the interval which is considered, i.e. $\lambda = 1/2$, for any value of π, therefore it is positive for all $\lambda \in (1/2, 1]$.

Given the signs of $d\theta(\mu_{\sigma_L})$, $d^2\theta(\mu_{\sigma_L})$, $d\Gamma$, $d^2\Gamma$ and the fact that $\Gamma - \theta(\mu_{\sigma_L})$ is positive in the range considered, (32) is strictly positive and therefore (30) can change sign at most once in the range $\lambda \in [1/2, 1]$. Therefore in the range $\lambda \in [1/2, 1]$ (30) is either always negative or negative up to a point and then always positive, this implies the quasi-convexity.

The quasi-convexity implies that in the range $\lambda \in [1/2, 1]$, the maximum must be either for $\lambda = 1/2$ or for $\lambda = 1$. The value of the aggregate output for $\lambda = 1/2$ is $\bar{y} = k + \bar{\theta}^2/a$, the value of the aggregate output for $\lambda = 0$ and $\lambda = 1$ is $\bar{y} = k + \frac{\bar{\theta^2}^2}{a(2\bar{\theta^2} - \theta_L^2)}$. For the output to be greater at $\lambda = 1/2$ than $\lambda = 1$, the condition to be satisfied is the following:

$$(\pi\theta_L + (1-\pi)\theta_H)^2(2\pi\theta_L^2 + 2(1-\pi)\theta_H^2 - \theta_L^2) - (\pi\theta_L^2 + (1-\pi)\theta_H^2)^2 \geq 0$$

(35)

i.e.

(36) $\quad (\theta_L - \theta_H)(-1 + \pi)\left(-2\theta_H^2\pi^2\theta_L + 2\theta_H^3\pi^2 - 2\theta_H\pi^2\theta_L^2 + \right.$

$$\left. 2\pi^2\theta_L^3 - 3\theta_H^3\pi + \pi\theta_L\theta_H^2 + 2\theta_H\pi\theta_L^2 + \theta_H^3 + \theta_L\theta_H^2\right) \geq 0$$

Observe that

$$(37) \quad -2\,\theta_H{}^2\pi^2\theta_L + 2\,\theta_H{}^3\pi^2 - 2\,\theta_H\pi^2\theta_L{}^2 2\,\pi^2\theta_L{}^3 - 3\,\theta_H{}^3\pi + \pi\,\theta_L\theta_H{}^2 +$$

$$2\,\theta_H\pi\,\theta_L{}^2 + \theta_H{}^3 + \theta_L\theta_H{}^2 =$$

$$(38) \quad 2\,\pi^2\theta_L{}^3 + 2\,\pi(1 - \pi)\theta_H\theta_L{}^2 + \left(-2\,\pi^2 + \pi + 1\right)\theta_H{}^2\theta_L + \left(2\,\pi^2 + 1 - 3\,\pi\right)\theta_H{}^3$$

Observe that

$$(39) \quad \left(-2\,\pi^2 + \pi + 1\right)\theta_H{}^2\theta_L + \left(2\,\pi^2 + 1 - 3\,\pi\right)\theta_H{}^3 =$$

$$(1 - \pi)\theta_H^2(2\pi\theta_L - 2\pi\theta_H + \theta_L + \theta_H).$$

Hence, after a factorization condition (35) can be rewritten as

$$(40) \quad (\theta_L - \theta_H)\,(-1 + \pi)\left(2\,\pi^2\theta_L{}^3 + 2\,\pi(1 - \pi)\theta_H\theta_L{}^2 +\right.$$

$$(1 - \pi)\theta_H^2(2\pi\theta_L - 2\pi\theta_H + \theta_L + \theta_H),$$

which is positive. Notice that $2\pi\theta_L - 2\pi\theta_H + \theta_L + \theta_H \geq 0$ IFF $\frac{2\theta_L}{2\pi - 1} \geq \theta_H - \theta_L$, which is always verified in the case $\pi \geq 1/2$ which I am considering.

This proves that condition (35) is satisfied.∎

APPENDIX B. ANALYSIS WITH WITH HETEROGENOUS ENDOWMENTS

In this section I explore the possibility of heterogeneous endowments as I assume that k^i takes value k_L for a fraction α of the population and value k_H for the remaining fraction $1 - \alpha$, with $k_L < k_H$, and that θ^i takes value θ_L for a fraction π of the population and value θ_H for the remaining fraction $1 - \pi$, with $\theta_L < \theta_H$. The two distributions are independent. This last assumption and the low of large numbers which applies to this large economy together imply that $(\theta^i, k^i) = (\theta_L, k_L)$ for a fraction $\pi\alpha$, $(\theta^i, k^i) = (\theta_H, k_L)$ for a fraction $(1 - \pi)\alpha$, $(\theta^i, k^i) = (\theta_L, k_H)$ for a fraction $\pi(1 - \alpha)$ and $(\theta^i, k^i) = (\theta_H, k_H)$ for a fraction $(1 - \pi)(1 - \alpha)$ of the population. The new expression for the optimal tax rate preferred by agent i follows:

(41)
$$\tau(k^i, \mu^i) = 1 - \frac{1 + \frac{a(k^i - \bar{k})}{\Gamma}}{2 - \frac{\theta(\mu^i)^2}{\Gamma}}.$$

As explained by Benabou and Tirole (2006), the numerator of (41) indicates that a lower relative endowment $(k^i - \bar{k})$ naturally increases the desired tax rate and that whether progressive or regressive, such distributive goals must be traded off against distortions to the effort-elastic component of the tax base (moral hazard problem).

The tuple (k^i, μ^i) identifies the preferred tax rate by voter i and given α, π and λ, there are four groups of voters in the economy. If $\alpha \in (1/2, 1]$ $(\alpha \in [0, 1/2))$ the majority of the agents has an endowment $k^i = k_L$ $(k^i = k_H)$. If $p_{\sigma_L} > 1/2$ $(p_{\sigma_L} < 1/2)$, the majority of the agents holds a belief μ_{σ_L} (μ_{σ_H}) at $t = 1$.

Voting Outcome with Heterogenous Endowments.
Case 1: $\alpha \geq 1/2$, $\pi \geq 1/2$, $\lambda \geq 1/2$.
$\lambda \geq 1/2$ implies that $\mu_{\sigma_H} \geq \mu_{\sigma_L}$ and therefore the following ranking of preferred tax rates: $\tau(k_H, \mu_{\sigma_H}) \leq \min\{\tau(k_H, \mu_{\sigma_L}), \tau(k_L, \mu_{\sigma_H})\} \leq \max\{\tau(k_H, \mu_{\sigma_L}), \tau(k_L, \mu_{\sigma_H})\} \leq \tau(k_L, \mu_{\sigma_L})$.[22] $\alpha \geq 1/2$ implies that the majority of the agents has $k^i = k_L$. $\lambda \geq 1/2$ and $\pi \geq 1/2$ together imply that $p_{\sigma_L} \geq 1/2$. There are two possible sub-cases.

Case 1.1: $\alpha p_{\sigma_L} > 1/2$. The pivotal group is the one who prefers $\tau(k_L, \mu_{\sigma_L})$; this because more than half of the population belongs to this group.

Case 1.2: $\alpha p_{\sigma_L} < 1/2$. If $\tau(k_L, \mu_{\sigma_H}) > \tau(k_H, \mu_{\sigma_L})$ then the pivotal group is the one who prefers $\tau(k_L, \mu_{\sigma_H})$, this because the ranking implies that the group with $\tau(k_L, \cdot)$

---

[22]This because $\tau(k, \mu)$ monotonically decreases in both k and μ.

includes the median voter but this does not belong to the group with $\tau(k_L, \mu_{\sigma_L})$. If $\tau(k_H, \mu_{\sigma_L}) > \tau(k_L, \mu_{\sigma_H})$ then the pivotal group is the one who prefers $\tau(k_H, \mu_{\sigma_L})$, this because the ranking implies that the group with $\tau(\cdot, \mu_{\sigma_L})$ includes the median voter but this does not belong to the group with $\tau(k_L, \mu_{\sigma_L})$.

Case 2: $\alpha \geq 1/2$, $\pi \geq 1/2$, $\lambda \leq 1/2$. $\lambda \leq 1/2$ implies $\mu_{\sigma_L} \geq \mu_{\sigma_H}$ and therefore the following ranking of preferred tax rates: $\tau(k_H, \mu_{\sigma_L}) \leq \min\{\tau(k_H, \mu_{\sigma_H}), \tau(k_L, \mu_{\sigma_L})\} \leq \max\{\tau(k_H, \mu_{\sigma_H}), \tau(k_L, \mu_{\sigma_L})\} \leq \tau(k_L, \mu_{\sigma_H})$. $\alpha \geq 1/2$ implies that the majority of the agents has an endowment $k^i = k_L$ and $\pi \geq 1/2$, $\lambda \leq 1/2$ imply that $p_{\sigma_L} \leq 1/2$, hence $p_{\sigma_H} \geq 1/2$. Two possible sub-cases follow.

Case 2.1: $\alpha p_{\sigma_H} > 1/2$. It is immediate that the pivotal group is the one who prefers $\tau(k_L, \mu_{\sigma_H})$; this because more than half of the population belongs to this group.

Case 2.2: $\alpha p_{\sigma_H} < 1/2$. If $\tau(k_L, \mu_{\sigma_L}) > \tau(k_H, \mu_{\sigma_H})$ then the pivotal group is the one who prefers $\tau(k_L, \mu_{\sigma_L})$, whereas if $\tau(k_H, \mu_{\sigma_H}) > \tau(k_L, \mu_{\sigma_L})$ then the pivotal group is the one who prefers $\tau(k_H, \mu_{\sigma_H})$.

Case 3: $\alpha \geq 1/2$, $\pi \leq 1/2$, $\lambda \geq 1/2$. $\lambda \geq 1/2$ implies $\mu_{\sigma_H} \geq \mu_{\sigma_L}$ and therefore the ranking of preferred tax rates is the same as in **Case 1**. $\pi \leq 1/2$, $\lambda \geq 1/2$ imply that $p_{\sigma_L} \leq 1/2$, therefore $\alpha p_{\sigma_L} > 1/2$ is never verified and hence Case 1.1 is never verified. Therefore **Case 3** has the same outcome of Case 1.2.

Case 4: $\alpha \geq 1/2$, $\pi \leq 1/2$, $\lambda \leq 1/2$. $\lambda \leq 1/2$ implies $\mu_{\sigma_H} \leq \mu_{\sigma_L}$ and therefore the ranking of preferred tax rates is the same as in **Case 2**. $\pi \leq 1/2$, $\lambda \leq 1/2$ imply that $p_{\sigma_H} \leq 1/2$, therefore $\alpha p_{\sigma_H} > 1/2$ is never verified and hence Case 2.1 is never verified. Therefore **Case 4** has the same outcome of Case 2.2.

Case 5: $\alpha \leq 1/2$, $\pi \geq 1/2$, $\lambda \geq 1/2$.
$\lambda \geq 1/2$ implies that $\mu_{\sigma_H} \geq \mu_{\sigma_L}$ and therefore the following ranking of preferred tax rates: $\tau(k_H, \mu_{\sigma_H}) \leq \min\{\tau(k_H, \mu_{\sigma_L}), \tau(k_L, \mu_{\sigma_H})\} \leq \max\{\tau(k_H, \mu_{\sigma_L}), \tau(k_L, \mu_{\sigma_H})\} \leq \tau(k_L, \mu_{\sigma_L})$.[23] $\alpha \leq 1/2$ implies that the majority of the agents has $k^i = k_H$. $\lambda \geq 1/2$ and $\pi \geq 1/2$ together imply that $p_{\sigma_L} \geq 1/2$. There are two possible sub-cases.

---

[23]This because $\tau(k, \mu)$ monotonically decreases in both k and μ.

Case 5.1: $(1-\alpha)p_{\sigma_L} > 1/2$. The pivotal group is the one who prefers $\tau(k_H, \mu_{\sigma_L})$; this because more than half of the population belongs to this group.

Case 5.2: $(1-\alpha)p_{\sigma_L} < 1/2$. If $\tau(k_L, \mu_{\sigma_H}) > \tau(k_H, \mu_{\sigma_L})$ then the pivotal group is the one who prefers $\tau(k_H, \mu_{\sigma_L})$ whereas if $\tau(k_H, \mu_{\sigma_L}) > \tau(k_L, \mu_{\sigma_H})$ then the pivotal group is the one who prefers $\tau(k_L, \mu_{\sigma_H})$.

Case 6: $\alpha \leq 1/2$, $\pi \geq 1/2$, $\lambda \leq 1/2$. $\lambda \leq 1/2$ implies $\mu_{\sigma_L} \geq \mu_{\sigma_H}$ and therefore the following ranking of preferred tax rates: $\tau(k_H, \mu_{\sigma_L}) \leq \min\{\tau(k_H, \mu_{\sigma_H}), \tau(k_L, \mu_{\sigma_L})\} \leq \max\{\tau(k_H, \mu_{\sigma_H}), \tau(k_L, \mu_{\sigma_L})\} \leq \tau(k_L, \mu_{\sigma_H})$. $\alpha \leq 1/2$ implies that the majority of the agents has an endowment $k^i = k_H$ and $\pi \geq 1/2$, $\lambda \leq 1/2$ imply that $p_{\sigma_L} \leq 1/2$, hence $p_{\sigma_H} \geq 1/2$. Two possible sub-cases follow.

Case 6.1: $(1-\alpha)p_{\sigma_H} > 1/2$. It is immediate that the pivotal group is the one who prefers $\tau(k_H, \mu_{\sigma_H})$, because more than half of the population belongs to this group.

Case 6.2: $(1-\alpha)p_{\sigma_H} < 1/2$. If $\tau(k_L, \mu_{\sigma_L}) > \tau(k_H, \mu_{\sigma_H})$ then the pivotal group is the one who prefers $\tau(k_H, \mu_{\sigma_H})$, whereas if $\tau(k_H, \mu_{\sigma_H}) > \tau(k_L, \mu_{\sigma_L})$ then the pivotal group is still the one who prefers $\tau(k_H, \mu_{\sigma_H})$, this because agents with $\tau(;\mu_{\sigma_H})$ are the more than half of the population but given that $(1-\alpha)p_{\sigma_H} < 1/2$ those with preferred tax equal to $\tau(k_L, \mu_{\sigma_H})$ cannot include the median voter.

Case 7: $\alpha \leq 1/2$, $\pi \leq 1/2$, $\lambda \geq 1/2$. $\lambda \geq 1/2$ implies $\mu_{\sigma_H} \geq \mu_{\sigma_L}$ and therefore the ranking of preferred tax rates is the same as in **Case 5**. $\pi \leq 1/2$, $\lambda \geq 1/2$ imply that $p_{\sigma_L} \leq 1/2$, therefore $(1-\alpha)p_{\sigma_L} > 1/2$ is never verified and hence Case 5.1 is never verified. Therefore **Case 7** has the same outcome as case 5.2.

Case 8: $\alpha \geq 1/2$, $\pi \leq 1/2$, $\lambda \leq 1/2$. $\lambda \leq 1/2$ implies $\mu_{\sigma_H} \leq \mu_{\sigma_L}$ and therefore the ranking of preferred tax rates is the same as in **Case 6**. $\pi \leq 1/2$, $\lambda \leq 1/2$ imply that $p_{\sigma_H} \leq 1/2$, therefore $(1-\alpha)p_{\sigma_H} > 1/2$ is never verified and hence Case 6.1 is never verified. Therefore **Case 8** has the same outcome as case 6.2.

Example of Discontinuous Comparative Statics. I explore how a change in the value of λ does affect the prevailing tax rate in the economy. Referring to the cases explored in the previous section, I start by looking at Case 1.1., i.e. $\alpha \geq 1/2, \pi \geq$

$1/2, \lambda \geq 1/2, \alpha p_{\sigma_L} > 1/2$. The pivotal tax rate is $\tau(k_L, \mu_{\sigma_L})$. If λ increases the pivotal tax rate remains $\tau(k_L, \mu_{\sigma_L})$ and it goes to $\tau(k_L, \theta_L)$ for $\lambda = 1$. If λ decreases it is certain that there will be a $\lambda^* \in (1/2, 1)$ small enough such that the condition $\alpha p_{\sigma_L} > 1/2$ is not satisfied. This because for $\lambda = 1/2$ the condition is not satisfied and therefore for the continuity of αp_{σ_L} in λ there will be a value λ^* arbitrarily close to $\lambda = 1/2$ (the greater is α the smaller is λ^*) such that the condition does not hold. For this λ^*, either $\tau(k_L, \mu_{\sigma_H})$ or $\tau(k_H, \mu_{\sigma_L})$ becomes pivotal and hence the pivotal tax rate jumps downwards.

Example of Continuous Comparative Statics. If the condition $\alpha p_{\sigma_L} > 1/2$ does not hold, then the pivotal tax rate is $\tau(k_L, \mu_{\sigma_H})$ for $\lambda \in [1/2, 1]$, i.e. Case 1.2, and $\tau(k_L, \mu_{\sigma_L})$ for $\lambda \in [0, 1/2]$, i.e. Case 2.2. This implies that the prevailing tax rate is still symmetric with respect to $\lambda = 1/2$. For $\lambda \in [1/2, 1]$ ($\lambda \in [0, 1/2]$), the prevailing tax rate decreases (increases) monotonically from $\overline{\theta}^2$ (θ_L^2) to θ_L^2 ($\overline{\theta}^2$) and therefore in this case $\tau(\lambda)$ is a quasi-concave function. I have thus shown that if it does not exist a λ^* such that $\alpha p_{\sigma_L}(\lambda^*) > 1/2$ then τ is a quasi-concave symmetric function of λ.

Example of Multiple Equilibria with discontinuous comparative statics. In the case of heterogenous endowments the ex-ante objective function is still given by (28), where $k = \bar{k} = \alpha k_L + (1-)k_H$. Consider $\alpha = 0.8$, $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$. $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$ imply $p_{\sigma_L} = 0.4\lambda + 0.3$. If there is a value λ^* such that $\alpha(0.4\lambda^* + 0.3) > 1/2$, then λ^* is point of discontinuity for $\tau(\lambda)$ as shown in the example of discontinuous comparative statics. For such a λ^* to exist it must be that $0.7\alpha > 1/2$, i.e. $\alpha > 5/7$. I take the case of $\alpha = 0.8$, which implies $\lambda^* \simeq 0.81$. I analyze the object u_0^i as a function of λ. It can be computed that u_0^i is maximized and equal to 1.63128 for both $\lambda = 0.81$ and $\lambda = 1$, hence the multiple equilibria. This example shows that heterogeneous endowments imply that intermediate values of λ can be optimal. I plot u_0^i, τ, Γ and the optimal values of individual and aggregate effort in figures 8, 9, 10, 11, 12.
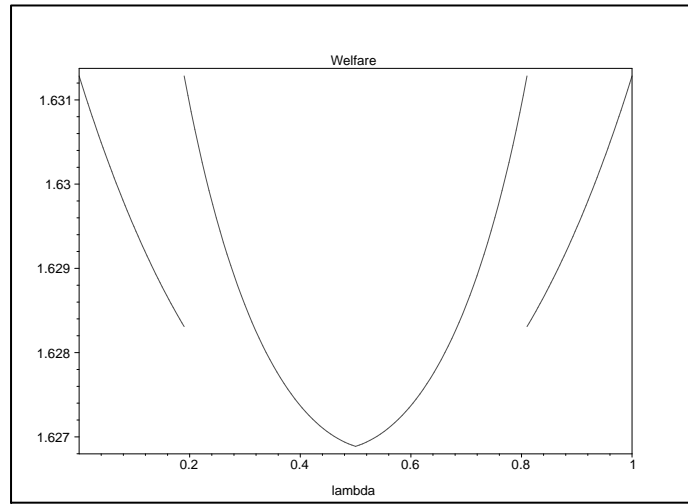
FIGURE 8. Welfare for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.
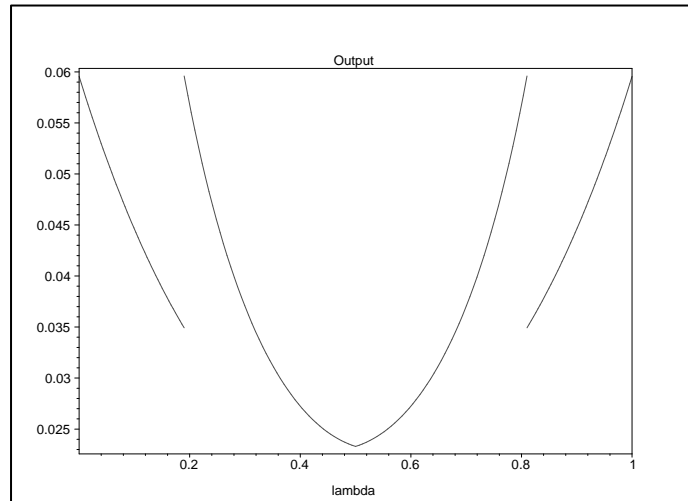

FIGURE 9. Aggregate Output for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.
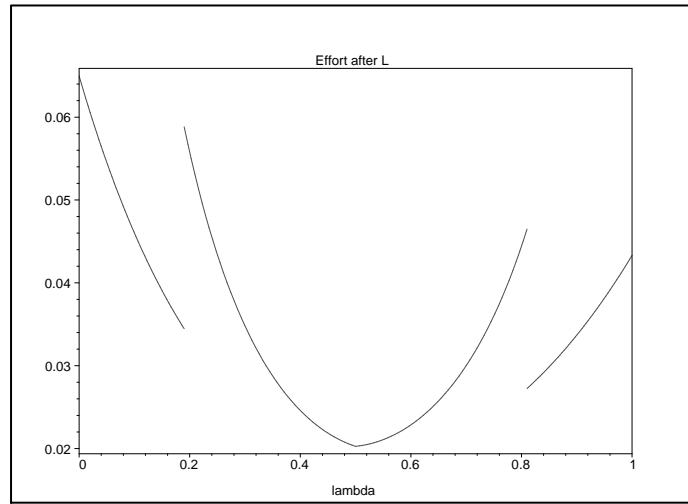
FIGURE 10. Effort after the observation of σ_L for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.
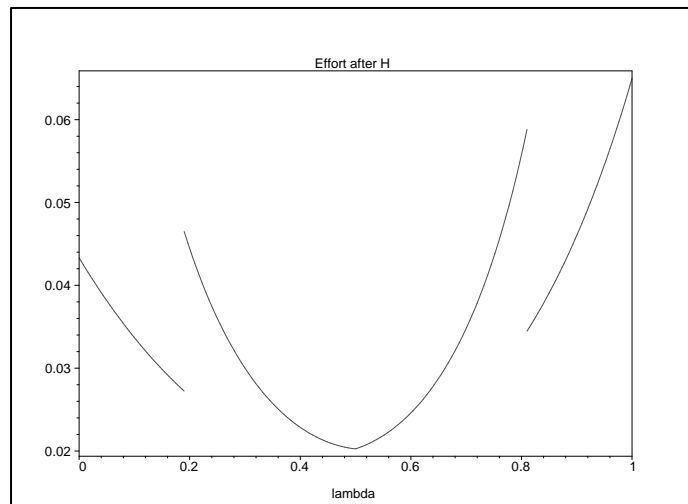

FIGURE 11. Effort after the observation of σ_H for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.
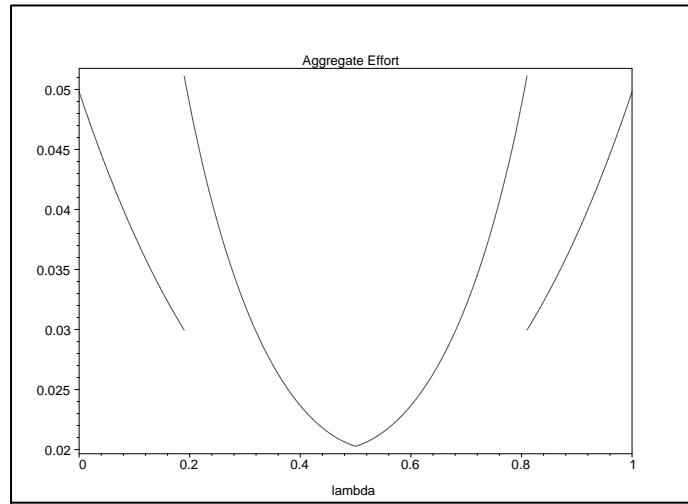
FIGURE 12. Aggregate effort for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.