School of Social Sciences
CITY UNIVERSITY LONDON

Department of Economics

Diverse Societal Beliefs and Redistributive Policies, but Equal Welfare: The Trade-off Effect of Information

Tommaso Gabrieli[1]
Henley Business School, University of Reading
and City University London

Department of Economics
Discussion Paper Series
No. 10/04

[1] Department of Economics, City University London, Social Sciences Bldg, Northampton Square, London EC1V 0HB, UK.
Email: *t.gabrieli@reading.ac.uk*. T: +44 (0)118 378 4257

DIVERSE SOCIETAL BELIEFS AND REDISTRIBUTIVE POLICIES, BUT EQUAL WELFARE: THE TRADE-OFF EFFECT OF INFORMATION

TOMMASO GABRIELI*

THIS VERSION JANUARY 2010, FIRST VERSION OCTOBER 2007

ABSTRACT. 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. This paper presents 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 to analyze various comparative statics in order to answer to policy questions.

Key words and phrases. Politico-Economic Equilibria, Redistribution, Incomplete Information.
JEL Classification: D31, D72, D80, E62, H30, O40.
*Henley Business School University of Reading and Department of Economics City University London. E-mail: t.gabrieli@reading.ac.uk.
The present paper has been completed while at City University London as University Research Fellow in the Department of Economics and at Catholic University of Milan Faculty of Economics as Research Grantee. The present paper is a revised version of Chapter II of my Phd thesis completed at the University of Warwick. An early version of this paper appeared with a different title as Gabrieli (2007).
I thank Herakles Polemarchakis and Sayantan Ghosal for invaluable guidance and motivation as my PhD supervisors. I am grateful for helpful remarks and comments to Marco Battaglini, Enrico Minelli, Roland Benabou, David Austen-Smith, Sendhil Mullainathan, Chuck Blackorby, Francis Bloch, Amrita Dhillon, Jayashri Dutta, Anandi Mani, Lucy White, Peter Hammond, as well as to seminar and conference participants at University of Warwick, Catholic University of Milan, University of Lausanne, University of Essex, City University of London, CSEF University of Salerno, EEA 07, EEA 09, PhD PEUK 08, RES 08, PET 09.

1. INTRODUCTION

Why do similarly developed countries choose widely different social contracts? The archetypical example is represented by the persistence and co-existence of European-type welfare states versus the US-type laissez-faire social contracts. Such question is not new and has motivated a large body of research across disciplines.

Without denying the importance of some "fundamental"differences across countries which can impact on such redistributive outcomes, economists have traditionally looked for explanations of such societal choices without appealing to exogenous differences in tastes, technologies or political systems. Under this perspective, redistributive outcomes are not considered as exogenous but are endogenously determined taking the political process into account.

1.1. **The contribution of the previous literature.**

An influential strand of literature which started in the 90's has developed models of inequality, redistribution and growth building on the seminal contribution of Meltzer and Richard (1981). The basic prediction of the framework of Meltzer and Richard (1981) is that of a unique equilibrium rate of redistribution, where greater inequality translates into a poorer median voter relative to the country's mean income and therefore the greater the inequality and the higher it is the prevailing (or equilibrium) rate of redistribution in the economy. The observation that, especially across developed countries, higher pre-tax inequality does not seem to imply higher redistribution is therefore inconsistent with the predictions of the theory of Meltzer and Richard (1981) and of some following models.[1]

[1] Perotti (1994), Perotti (1996) and most of the other studies reviewed in Benabou (1996) find no relationship between inequality and the share of government expenditures in GDP. Among advanced countries the effect is actually negative. This is suggested by many examples as the differences between (Western) Europe and the US. Western European countries ("Europe"in short) are characterized by more extensive redistributive policies than the US. Nevertheless, contrary to the Meltzer-Richard paradigm of redistribution, pre-tax inequality is higher in the US than in Europe. Alesina and Angeletos (2005) report that while the Gini coefficient in the pre-tax income distribution in the US 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 about this can be found in Alesina, Glaeser, and Sacerdote (2001) and Alesina and Glaeser (2004). Important contributions which introduced the framework of Meltzer and Richard (1981) in dynamic models of growth and inequality are those of Perotti (1993), Alesina and Rodrik (1994) and Persson and Tabellini (1994) among others.

Despite the existence of solid alternative theoretical explanations [2], a small but growing literature has pointed to the fact that the observed differences in the political support for redistribution appear to reflect the differences in the beliefs which different societies hold about the underlying determinants of individual wealth and the extent of social mobility. Not only notable differences exist in the level of redistribution (or social contract) across countries, striking differences appear in the beliefs that different societies hold about the underlying determinants of individual fortunes and poverty and such beliefs appear to be determinant for the observed social contracts. Once again the most striking difference relates to the differences between the United States and Western Europe. Since De Tocqueville (1835), many have noticed the *exceptionalism* or *dream* characterizing the American society, in other words the widespread belief according to which everyone can become rich if wants so and mobility is high in the "land of opportunities".[3] De Tocqueville (1835) himself and many other sociologists and political scientists after him have pointed to persistent differences in the popular beliefs about social mobility in explaining the persistent differences between US and European redistributive politics. Those observed massive differences in the beliefs appear to be important, especially since there is a strong correlation between these beliefs and the actual levels of redistribution and there is also empirical evidence about the fact that beliefs are actually strong determinants of the demand for redistribution.[4]

[2]Departing from the basic Meltzer and Richard (1981) framework but still considering the level of redistribution as endogenously determined, the theories of Benabou (2000), Benabou (2005), Saint-Paul (2001), Hassler, Rodriguez-Mora, Storesletten, and Zilibotti (2003) have been able to show how both European-type welfare states and US-type laissez-faire societies, together with respectively lower and higher levels of inequality, can arise as multiple steady states from the joint dynamics of the wealth distribution and redistributive policies.

[3]Recent 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 cannot escape it 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". Suhrcke (2001) shows that large disparities in attitudes also exist within Europe, especially between OECD and Eastern European countries.

[4]See Alesina, Glaeser, and Sacerdote (2001) about the cross correlation between the belief that "luck determines wealth"and the level of redistribution. See Fong (2001), Corneo and Gruner (2002), Alesina and La Ferrara (2005) about the evidence showing that beliefs are strong determinants of the demand for redistribution and that individual beliefs determine individual political orientations more than other factors like personal wealth.

The theoretical contributions of Piketty (1995), Alesina and Angeletos (2005) and Benabou and Tirole (2006) have developed insightful theoretical models describing how individual beliefs can shape politico-economic outcomes and viceversa and how multiple equilibria (US-type vs Europe-type) with different beliefs are possible. The standard Meltzer and Richard (1981) model has a unique equilibrium where the greater is the inequality and the lower is the wealth of the median voter with respect to the mean and consequently the higher is the prevailing rate of redistribution. In order to introduce a role for beliefs and derive multiple equilibria these models introduce new elements. One common feature is that the economic agents described by these models have incomplete information about the determinants of individual wealth, namely about the value of the return on effort versus the value of the predetermined factors on which the individual has no control. It is this incomplete information to create the premise for different beliefs about the role of controllable (as hard work and discipline) vs. uncontrollable (as luck or family of origin) factors in the determination of wealth. Moreover a standard moral hazard effect implies that the greater is the expected value of the return on effort and the greater is the efficiency loss from redistribution, for this reason different beliefs may lead to different prevailing redistribution rates in the political game. In the model of Piketty (1995) agents have incomplete 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. Making a link to some recent literature in behavioral economics, the models of Alesina and Angeletos (2005) and Benabou and Tirole (2006) introduce different psychological elements in this framework. 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. Also in this model, agents have imperfect information about the true return on effort versus the role of predetermined factors and this allows for different beliefs. The authors 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 beliefs to be self sustained. [5] 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 procrastination. Also in this model agents have incomplete information about the true return on effort versus the role of predetermined factors, but in addition to this each agent receives a signal about the value of the return on effort. The novel feature of the model is that each agent can decide the precision of the signal, in other words each agent can decide how much to be informed and manipulate her own (or her children's) beliefs. Such formalization wants to capture the idea that (false) beliefs about the underlying determinants of wealth and social mobility could derive from a false consciousness which is chosen and valued by the worker themselves. Extensive evidence in sociology and psychology seems to suggest this fact.[6] Given time inconsistent preferences, which captures the idea of procrastination and imperfect willpower, for each agent it is optimal to have imperfect information so to think that the return on effort is greater than the true value. This is because such belief increases the effort implemented by the future self (using the language of behavioral economics, or the future generation according to a more standard interpretation of the model). Nevertheless this bias has a cognitive cost and when people anticipate little redistribution the value of a proper motivation (in other words the value of believing that the return on effort is greater than the true value) is much higher than with higher redistribution. When redistribution is low everyone thus has greater incentives to believe that effort plays a major role[7] 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 the belief that the return on effort is high and relatively low redistribution. The other, "European"equilibrium is characterized by a high prevalence of the belief that the return on effort is low and relatively high redistribution.

[5]Similarly to the paper of Alesina and Angeletos (2005), also in the work of Cervellati, Esteban, and Kranich (2006) the individual preferred level of redistribution is not motivated by purely selfish concerns as in Meltzer and Richard (1981) but also by a social component; in this model, though, multiple equilibria do not originate from different beliefs but from different moral sentiments.

[6]See Benabou and Tirole (2006) for precise references.

[7]In the words of Benabou and Tirole (2006) this is the *Belief in a Just World*

1.2. **The contribution of this paper.** In this paper I present a new theoretical framework which shares some of the underlying features of the last group of models but, unlike those, allows for varying degrees of incomplete information in the economy and focuses on the effect that incomplete information has on the political and economic outcomes. As I have already explained in the previous section, incomplete information is a common element in the models of Piketty (1995), Alesina and Angeletos (2005) and Benabou and Tirole (2006), but it is the addition of other elements on top of that to imply the existence of multiple equilibria in these models. I have already discussed that those additional elements are imperfect learning in Piketty (1995) and psychological elements as preferences for social fairness in Alesina and Angeletos (2005) or as time inconsistent preferences and cognitive dissonance in Benabou and Tirole (2006). Nevertheless, the degree of underlying incomplete information in those models is fixed, in other words a government or another institution could not do anything in order to increase the agents' information about the value of the return on effort versus "luck". In reality it appears to be the case that there are ways in which institutions can influence agents' information about the underlying determinants of wealth and mobility. For example, the type of education can have an impact on the degree of information because the return on effort θ depends on individual ability and an educational system which reveals individual abilities better can be a way to make individual beliefs to be more realistic. Another way for an institution to provide more information could be to provide accurate historical data on the dynamics of mobility. Or again the information of the agents can be influenced by propaganda: a government or a group of people could try to convince others about the importance of effort versus predetermined factors. It is useful to precise that the contribution is not to build a detailed analysis of endogenous propaganda or educational features.[8] It is instead a more abstract exercise which, assuming that there is an institution which can influence the degree of information as indeed seems to be the case, considers the degree of information as a policy variable and answers to a natural policy questions: What is the effect of varying degree of information on individual choices and aggregate outcomes? What is the optimal level of information given different objectives?

I develop a model which extensively builds extensively on the framework which is shared by the models of Piketty (1995), Alesina and Angeletos (2005) and Benabou

[8]Nevertheless I will also offer some predictions at this level which are interestingly in line with some existing empirical and anecdotal evidence.

and Tirole (2006). Individual wealth still depends on effort and "luck", agents still have incomplete information about the true value of the return on effort and with such incomplete information they first vote over redistribution and then exert effort. The novel feature is represented by allowing for varying degrees of incomplete of information and consequently derive the comparative statics of the individual and aggregate political and economic outcomes. In order to allow for varying degrees of incomplete information I introduce a simple informative set-up. The economic agents in my model do not know the true value of the individual return on effort on ability but receive informative signals about this individual value. The precision of the signals is the same across agents and this feature wants to capture the idea that the level of information is an institutional feature of the economy.[9] Varying the precision of the signal means to vary the degree of information in the economy. This framework isolates the effect of incomplete information, as other elements[10] are not present in the model and allows a clear analysis of a number of interesting comparative statics. The degree of information impacts on two individual choices: the decision about voting over redistribution and the choice of effort. Increasing the level of information improves the individual choices of effort. For this reason, net of the effect that information has on the voted rate of redistribution, increasing the level of information improves ex-ante welfare. Conversely, increasing the level of information can increase or decrease the prevailing rate of redistribution depending on the identity of the median voter. The most interesting case is when increasing the level of information increases the prevailing rate of redistribution. In such case, in a model with linear utility in wealth and therefore where the ex-ante optimal rate of redistribution is equal to zero, increasing the level of information has a trade-off effect: on one hand it improves the allocation of individual effort but on the other hand it raises inefficient redistributive taxation. This represents the first result of my analysis, namely that welfare does not increase monotonically in information.[11] A second result consists in showing that, generally, the welfare function can be both concave and convex in the level of information. The reason for this is essentially that, net of the effect of the redistributive tax, the welfare function is convex in the

[9]Of course it would be realistic and interesting to allow for the possibility of different precisions across groups of agents or networks and I leave this exercise to the future research.
[10]For example there are not the psychological elements which were present in the models of Alesina and Angeletos (2005) and Benabou and Tirole (2006).
[11]This result is not entirely dependent on the assumption of linear utility function as I will discuss in appendix D

level of information. Such result links back to the seminal contribution of Radner and Stiglitz (1984) who show the convexity of the value of information. The convexity of the welfare function implies that there are cases of multiple optimal levels of information. Considering the level of information as a policy variable, the comparative statics of varying levels of information can address interesting policy questions. In addition to the comparative statics which relate to welfare, I analyze the comparative statics of all the other political and economic outcomes: prevailing rate of redistribution, individual and aggregate effort and output. A third result relates to the comparative statics of aggregate output. I show that also aggregate output is not monotonic in the level of information and in addition can be both concave and convex; nevertheless, in the case in which the prevailing rate of redistribution increases in the level of information I show that output is maximized for the minimum level of information.

Up to this point I have described comparative statics results which are obtained varying the level of information exogenously. The second step of my analysis is to consider an endogenous prevailing level of information. I model it as the result of a collective choice. In the case in which every agent is identical before receiving the informative signal about individual ability, all agents agree on the same ex-ante optimal level of information which also maximizes ex-ante welfare. Fixed the level of information as the ex-ante welfare maximizing, I define a politico-economic equilibrium as the resulting beliefs, prevailing level of redistribution and optimal individual choices of effort. Given multiple optimal values of information, there are multiple politico-economic equilibria which can still be interpreted as US-type versus Europe-type. The two equilibria still present the known macroeconomic features found by the previous literature, namely that the US-type (Europe-type) politico-economic equilibrium is characterized by relatively low (high) redistribution and high (low) aggregate output. What is new in my model is a characterization of the two equilibria in terms of the informative features. I find a US- (Europe-) type politico-economic equilibrium to be characterized by relatively (i) low (high) informative signals and high (low) adverse selection, as individual beliefs and effort levels are pooled (separated) (ii) low (high) moral hazard, as redistribution is low (high) and hence does not (does) distort individual effort to a great extent. With respect to the existence of multiple politico-economic equilibria, this paper presents a methodological contribution as it shows that multiple equilibria can be obtained without psychological biases. For what concerns the interpretation of the two equilibria, the contribution is to show that otherwise similar societies may find optimal to choose different information "cultures"which consequently imply different beliefs' systems and different politico-economic outcomes.

In my model the US-type (Europe-type) equilibrium has signals which are relatively less (more) informative and separate less (more) the beliefs on the return on effort. The two equilibria can still be respectively interpreted as *Belief in a Just World* vs. *Realistic Pessimism* as in the model of Benabou and Tirole (2006): the US-type equilibrium can be interpreted as *Belief in a Just World* because in such equilibrium the majority of the agents deny "bad news"and believe the return on effort to be higher than what it really is. In addition, my model gives another interesting interpretation of the two equilibria: relatively to the EU-type, the US-type equilibrium is characterized by a stronger belief in equal opportunities. Conditional on the signal, in the US-type (EU-type) equilibrium agents believe that opportunities are relatively more (less) equal, as the posterior beliefs about the return on effort are less (more) heterogeneous. Adding this new interpretation seems to be important to understand better the two societies as the concept of equal opportunities seems to be inherently related to the one of American *dream* vs. European *realism or pessimism*. The expression *American Dream* was coined by Adams (1931):

"*The American Dream is that dream of a land in which life should be better and richer and fuller for everyone, with opportunity for each according to ability or achievement. It is a difficult dream for the European upper classes to interpret adequately, and too many of us ourselves have grown weary and mistrustful of it. It is not a dream of motor cars and high wages merely, but a dream of social order in which each man and each woman shall be able to attain to the fullest stature of which they are innately capable, and be recognized by others for what they are, regardless of the fortuitous circumstances of birth or position.*"

For people to believe in the dream of a land in which life should be better and richer and fuller for everyone and "that every child in America has a decent shot at life and that the doors of opportunity remain open to all"[12] it must be the case that opportunities (i.e. the return on effort) cannot be considered too heterogenous across individuals (otherwise some individuals would be coursed from birth to be low achievers). My model therefore predicts the US-type equilibrium, relatively to

[12]Those words have been used by Senator Barack Obama to describe the concept of *American Dream* in his speech on 27th July 2004, commonly referred to as *Reclaiming the American Dream.*

the Europe-type, to be characterized by a stronger belief in equal opportunities, as it seems to be the case.

My model can also shed some light on why two different beliefs systems came into existence and persisted in history. One interesting property of the equilibrium in my model (section 6) shows that two otherwise similar societies with small differences in the extent of the heterogeneity in the prior distribution of the return on effort across agents, may find optimal to choose very different informative cultures: the society with marginally smaller heterogeneity may choose low informative signals and hence US-type beliefs (i.e. *Belief in a Just World* and *belief in equal opportunities*).

Looking at the historical experiences of US and Europe it seems to be the case that in the days of the first pioneers, opportunities were indeed more equal and individual types more homogeneous in the "Land of Opportunities"than in the "Old Continent". Quoting McElroy (2006) (pp 60):

The self-selected immigrants who came to America by their own free choice came as hopeful individuals in search of opportunities to improve their individual lives. Among the things that they fled in leaving Europe were the limitations of the class membership that typified Europe's aristocratic cultures. They wanted to belong to a new society that would encourage their aspirations of self-improvement and their ambition to rise in society.

The properties of the equilibrium in my model (section 6) show how those initial differences can have implied that the societies find two different sets of beliefs to be optimal. Therefore my model can offer an explanation about how certain types of beliefs came to existence and persisted through the collective choice of a certain information structure. In this respect, without denying the importance of other institutional factors, my model offers a complementary institutional explanation to the more traditional one which is common in political science and sociology. [13]

[13] A more modern and more symmetric version of the more traditional Neo-Marxist 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 described beliefs.

Finally, it seems that a natural way to think about how societies have maintained different information "cultures"is to think about different educational systems. Therefore a natural interpretation of the precision of the signals is how much the schooling system informs students and parents about individual skills. As predicted by my model, there exists empirical evidence showing how the American secondary schooling system separates less than the European.[14]

In summary, my model interprets different (US vs Europe type) politico and economic outcomes as originated by different optimal informative cultures, where different (US vs Europe type) informative cultures present the following features: (i) more vs less widespread belief that effort more than luck determines individual wealth (ii) more vs less widespread belief that opportunities are equal (iii) more vs less separating educational systems. As already discussed both existing literature and stylized facts support such interpretation. A deeper empirical test of the interpretation of my theoretical results goes beyond the scope of this paper, nevertheless could be a fertile ground for future research.

The present 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 and the generalization of the results. Section 8 concludes. Appendixes A and B contain some lengthly proofs. Appendixes C and D respectively introduce heterogenous endowments and risk aversion in the model and, testing for the model's robustness, show that the conclusions do not change.

[14]The main reason is that curriculum-based external exit examinations (CBEEE) are relatively more common in Europe. See Bishop (1997) for an introduction to the topic and Bishop (1996) and Woessmann (2003) for comparative studies of US vs. Europe. Another reason is that also early tracking is relatively more common in Europe, see Woessmann (2009). Despite the existing empirical evidence on the different informative power of the two schooling systems, the prediction of my model does not have to be that the European system informs relatively better than the US one, but only that separates more (wether with right or wrong signals) as it indeed appears to be the case.

2. SET UP

I 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:

$$y^i = k^i + \theta^i e^i, \tag{1}$$

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.[15] I assume that θ^i is i.i.d. across agents and that θ^i takes value θ_L with probability π and value θ_H with probability $1-\pi$, where $\theta_L < \theta_H$. 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 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:

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

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:

$$\mu^i \equiv \Pr[\theta^i = \theta_L | \sigma^i], \tag{3}$$

represents agent i belief that $\theta^i = \theta_L$ conditional on the observation of signal σ^i. From the Bayes rule it follows that:

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

[15]It will be clear that an homogeneous endowment does not play any role and without loss of generality I could set $k = 0$. Nevertheless, this very simple technology is enough to capture the role of "luck"versus effort. Section C (supplemental material) considers the possibility of heterogenous endowments.

and

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

The expected value of θ^i conditional on the observation of σ^i is given by the following expression:

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

Given the symmetric structure of (2) I consider the interval $\lambda \in [1/2, 1]$. For $\lambda = 1/2$ the signal σ^i is completely uninformative and the posterior belief is equal to the prior, i.e. μ_{σ_L}=μ_{σ_H}=π. Increasing λ makes the signal progressively more informative up to the point that $\lambda = 1$ and the signal is perfectly informative, i.e. $\mu_{\sigma_L} = 1$, $\mu_{\sigma_H} = 0$.[16] As already explained in the introduction, the value of λ represents the level of information in the economy and in a rather abstract way I consider it is an institutional feature and a policy variable. The ex-ante probability of observing σ_L is given by the following expression:

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

symmetrically

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

Over-lined variables stand for mean values for the population, hence $\bar{y}$ and $\bar{e}$ are respectively the mean, or aggregate, values of output and effort and

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

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

are respectively the mean 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. Such linear redistribution scheme is due to Romer (1975), it is standard in this literature

[16] I could have alternatively considered the interval $\lambda \in [0, 1/2]$, in this case $\lambda = 0$ implies that the signal is perfectly informative and increasing λ up to $\lambda = 1/2$ makes the signal progressively less informative. It will be clear that given the symmetric structure of the signal the entire analysis would be symmetric to the one obtained considering $\lambda \in [1/2, 1]$.

and implies that the government budget constraint is always binding. Throughout the analysis I consider the following individual utility function[17]:

$$u^i(c^i, e^i) = c^i - \frac{a}{2}(e^i)^2. \tag{10}$$

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, agents get the net outcome of the production activity plus a net transfer and enjoy consumption[18].

3. Voters' Problem

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

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

where $E[\cdot|I^i_t]$ is individual i's expectation conditional on the information at time t. As explained in the previous section, the information structure is such that $I^i_0 = T$ and $I^i_1 = (T, \sigma^i)$. Given that voting and effort choices take place at $t = 1$, after that the signal σ^i is received, what is important to bear in mind is that the objective function that each agent i maximizes when voting and choosing effort is the expected utility (11) conditional on signal σ^i. 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:

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

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, it is useful to specify

[17]In section D (supplemental material) I will consider the possibility of concavity in consumption and discuss the relative implications.

[18]Therefore in the final period the uncertainty regarding the value of θ^i is resolved as agents can infer the true value of θ^i from y^i. Nevertheless, after period 2 the same sequence of events can be repeated for the next generation; in absence of intergenerational transfers and with completely uncorrelated skills, the same results would still apply.

the individual i expectation of the output from effort:

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

and of squared effort

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

In computing the mean (aggregate) product of effort $\overline{e\theta}$, each agent i knows that 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). Therefore it is the case that

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

where I define

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

Collecting $\theta(\mu_{\sigma_L})$ and $\theta(\mu_{\sigma_H})$ it is easy to re-write expression (16) as

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

The term Γ is the expression for aggregate output from effort, net of the distortive effect of redistribution on effort. It will be shown that this term will play a crucial role in the analysis. Plugging (13), (14) and (16) into (11), I obtain an indirect form of (11) as a function of τ:

(18) $$u_1^i = k^i + (1-\tau)^2\theta(\mu^i)^2/a + \tau(1-\tau)\Gamma/a - (1-\tau)^2\theta(\mu^i)^2/2a.$$

This is the object that voter i maximizes voting over the tax rate τ. Assuming for the moment that the second derivative of the objective function (18) is strictly negative, the ideal tax rate of agent i follows from the first order condition:

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

The denominator of (19) shows how the subjective prospects of upward mobility reduce the desired tax rate.[19] I introduce an assumption which bounds above the heterogeneity in individual abilities in order to assure the concavity of the objective function (18) and thus to use the median voter theorem.

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

A proposition follows:

Proposition 1. *Individual preferences for the rate of redistribution are single peaked and the individual ideal rate of redistribution is given by expression (19).*

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

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

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 $2\theta_L^2$. □

Proposition 1 shows that preferences over the tax rate are single peaked and therefore the median voter theorem applies. Labeling the prevailing tax rate in the voting game 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.[20]

4. Comparative Statics

I analyze the effect of a change in the value of the level of information λ on the endogenous variables of the model: prevailing tax rate, individual and aggregate effort, aggregate output. As already explained in the introduction, this is a natural exercise in order to understand the effects of policies which change – directly or

[19] The concept of prospects of upward mobility and its role in the determination of the prevailing rate of redistribution is the focus of the analysis of Benabou and Ok (2001). The term $\frac{\theta(\mu^i)^2}{\Gamma}$ represents the subjective prospects of upward mobility as it is equal to the the ratio of individual output (13) over aggregate output (15), noticing that the term $\frac{1-\tau}{a}$ gets canceled out.

[20] 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 i prefers the same tax rate $\tau(\mu^i)$, where $\mu^i = \pi$. From (19) it is easy to notice that for $\mu^i = \pi$, $\tau(\mu^i) = 0$; this follows from the fact that $\mu_{\sigma_L} = \mu_{\sigma_H} = \pi$ implies that $\theta(\mu)^2 = \Gamma = (\bar{\theta})^2$.

indirectly – the level of information in an economy, for example policies based on education or policies based on propaganda. In the following two lemmas I present two important intermediate results which are fundamental for the full analysis of the comparative statics.

Lemma 1. *The expected value of individual ability (6), conditional on the observation of signal σ_L (σ_H), is decreasing (increasing) in the level of information λ.*

Proof. This property of monotonicity is immediately proved from the computation of the respective first derivative with respect to λ:

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

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

□

It is straightforward to interpret this result: when the level of information in the economy is minimum ($\lambda = 1/2$) everyone maintains the prior belief to be of average ability $\bar{\theta}$. Increasing the the precision of the signals λ implies that the Bayes updating "relies"more on the signal and the expectation of those agents who receive the signal σ_L (σ_H) get progressively closer to the the value θ_L (θ_H). The following result defines the comparative statics which relate to expression (16), namely aggregate output from effort net of the distortive effect of redistribution on effort.

Lemma 2. *The expression of Γ (16) is (i) increasing and (ii) convex in the level of information λ.*

The proof is in Appendix A. The intuition behind this result is very important. Lemma 2 shows that when the incentive-distortive effect of taxation is not taken into account, increasing information has a positive effect on aggregate output, as agents choose effort more optimally given the true values of θ_L and θ_H. Expression Γ measures ex-ante or aggregate output, net from the distortive effect of taxation, and therefore is a measure of the value of information. The result of convexity in the value of information is a known result in economic theory which goes back to the seminal contribution of Radner and Stiglitz (1984). In order to study the comparative statics of the endogenous variables of the model, I study the cases of $\pi > 1/2$ and $\pi < 1/2$ separately.

Comparative statics for the case of $\pi > 1/2$. The case of $\pi > 1/2$ implies, together with the fact that $\lambda \geq 1/2$, that $p_{\sigma_L} \geq 1/2$ and therefore that the majority of the agents observes the signal σ_L and that the prevailing tax rate is $\tau = \tau(\mu_{\sigma_L})$. A proposition follows:

Proposition 2. *If $\pi > 1/2$, the prevailing tax rate τ is increasing in the level of information λ.*

Proof. Given that $\pi > 1/2$ implies that the prevailing tax rate is $\tau = \tau(\mu_{\sigma_L})$, taking the expression for the tax rate (19) with $\mu^i = \mu_{\sigma_L}$, the proof follows in a straightforward way from lemmas 1 and 2. □

From expression (19) it is easy to compute that the minimum value of the tax rate is $\tau = 0$, for $\lambda = 1/2$ and when $\theta(\mu)^2$=Γ=$(\bar{\theta})^2$. In the same way, it is also immediate that the maximum value of the tax rate is $\tau = 1 - \frac{1}{2-(\theta_L^2/\overline{\theta^2})}$, for $\lambda = 1$. Notice also that given that $\pi \in [0, 1)$, $\theta_L^2/\overline{\theta^2} < 1$ and hence $\tau \in [0, 1)$. The intuition behind the comparative static of the prevailing tax rate is easy to understand. Given that the majority group is the one formed by the agents who observe the signal σ_L, their prospects of upper mobility decrease in the level of information and therefore their expected gains from redistribution increase with the level of information.

In order to study the comparative statics which are relative to effort, using (12) I define the optimal effort implemented by those who observe σ_L:

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

and by those who observe σ_H:

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

Multiplying by the respective weights I obtain the expression of aggregate effort:

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

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 $\pi > 1/2$, aggregate effort (22) is decreasing in the level of information λ.*

The proof follows trivially from proposition 2 and from the fact that $\overline{\theta}$ is a constant. The result depends on the fact that the the only effect of information on aggregate effort is through the distortive tax rate. To be more precise, information impacts the expressions of individual effort both through the tax rate and individual beliefs, nevertheless at the aggregate level, information only impacts through the tax rate as the effect on the beliefs of the two groups cancel out. Looking at the expression of the optimal effort which is exerted by those who observe signal σ_L (20), it is immediate to see that it decreases in level of information λ, given that both $(1-\tau)$ and $\theta(\mu_{\sigma_L})$ decrease as suggested by lemmas 1 and 2. Instead, the comparative static for the expression of optimal effort which is exerted by those who observe signal σ_H (21) is ambiguous as $(1-\tau)$ is decreasing but $\theta(\mu_{\sigma_H})$ is increasing. The overall effect depends on the relative responsiveness of the two terms to λ. In a numerical example which I will present in section 6 it will turn out to be non monotonic.

I discuss the comparative statics of output. Plugging (15) into (1) I obtain the expression of aggregate output:

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

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. Nevertheless I find an interesting property:

Proposition 4. *If $\pi > 1/2$, the expression for aggregate output (23) is (i)either monotonically decreasing or monotonically decreasing up to a point and then monotonically increasing in the level of information λ* [21]*, (ii) maximized for $\lambda = 1/2$.*

The proof is in Appendix B. This represents a striking policy result: aggregate output is univocally maximized by the minimum level of information. Even if the level of information λ has opposite effects on aggregate output, as it increases distortive taxes τ but at the same time it increases output from effort Γ through a better allocation of effort, the distortive effect through the tax rate is always dominant. 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 $\lambda = 1$ is $\bar{y} = k + \frac{\overline{\theta^2}^2}{a(2\overline{\theta^2}-\theta_L^2)} > 0$.

Comparative statics for the case of $\pi < 1/2$. The case of $\pi < 1/2$ implies, together with the fact that $\lambda \geq 1/2$, that $p_{\sigma_L} \leq 1/2$ and therefore that the majority of the agents observes the signal σ_H and that the prevailing tax rate is $\tau = \tau(\mu_{\sigma_H})$. 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

[21]This behavior can be described as single peaked from below and it is a form of quasi-convexity.

the overall effect of λ is not a-priori clear. Nevertheless it is possible to find some properties:

Proposition 5. *In the case that $\pi < 1/2$, the prevailing tax rate τ is (i) always negative and (ii) if $(2\Gamma\frac{\partial\theta(\sigma_H)}{\partial\lambda}) < \theta(\sigma_H)\frac{\partial\Gamma}{\partial\lambda}$, it is decreasing in the level of information λ.*

Proof. It is useful to re-express (19) as

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

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 monotonicity I notice that the first derivative of τ with respect to λ is $\frac{d\tau}{d\lambda} = \frac{\theta(\mu_{\sigma_H})\left(2\Gamma\frac{\partial\theta(\mu_{\sigma_H})}{\partial\lambda} - \theta(\mu_{\sigma_H})\frac{\partial\Gamma}{\partial\lambda}\right)}{(2\Gamma^2 - \theta(\mu_{\sigma_H}))^2}$. Given lemmas 1 and 2, a sufficient condition for τ to be monotonic decreasing is therefore that $(2\Gamma\frac{\partial\theta(\mu_{\sigma_H})}{\partial\lambda}) < \theta(\mu_{\sigma_H})\frac{\partial\Gamma}{\partial\lambda}$. □

The negativity of the prevailing tax rate is easily interpretable. Given that the majority group which sets the tax rate is formed by the agents who observe the signal σ_H, whenever λ is greater than $1/2$ they expect to produce more than the average individual and therefore to loose out from redistribution. When τ decreases monotonically in λ it is straightforward that expression (23), lemma 2 and proposition 5 imply that aggregate output increases monotonically in λ. Moreover, given that the tax rate is always negative, then the aggregate output is always greater than in the case of $\pi \geq 1/2$.

Aggregate effort still depends exclusively on the tax rate, when τ decreases monotonically in λ it is straightforward that expression (22) implies that aggregate effort increases monotonically in λ. Moreover, given that the tax rate is always negative, also aggregate effort is always greater than in the case of $\pi \geq 1/2$. In the case in which τ decreases monotonically in λ, $e|\sigma_H$ increases in λ, as both $(1-\tau)$ and $\theta(\mu_{\sigma_H})$ increase. The effect of λ on $e|\sigma_L$ is instead partially ambiguous, as $(1-\tau)$ increases in λ whereas $\theta(\mu_{\sigma_L})$ decreases. The overall effect depends on how responsive are τ and $\theta(\mu_L)$ to λ.

5. Optimal Information

In the previous section I studied different comparative statics and the results offered insights for policy questions such as the level of information which maximizes output or how the level of information does affect the voted tax rate. It is now a natural question to explore the comparative statics in terms of welfare. More precisely, it is a natural question to explore the level of information which maximizes the ex-ante utility, namely the utility function at time 0 before that the agents receive the signal. In other words I investigate whether someone behind the veil of ignorance desires to remove the veil.

In order to compute the expression of the expected utility at time 0, I notice that at time 0, before receiving the signal, everyone is identical and expects to have mean ability $\bar{\theta}$. It follows the expression of expected individual output from effort:

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

and the expression for expected individual squared effort:

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

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

$$u_0^i = k + (1-\tau^2)\Gamma/2a. \tag{27}$$

Given that at time 0 everyone is identical this is the expression of both ex-ante individual utility and aggregate welfare. If an agent had to choose an optimal value of λ for the society at $t = 0$, she would choose a value of λ which maximizes (27). The solution of the problem is not a-priori trivial.

Case of $\pi > 1/2$. 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 that the overall effect is not a-priori clear. Nevertheless I find an interesting property:

Proposition 6. *If $\pi > 1/2$, the expression of ex-ante utility (27) is either monotonically decreasing or monotonically decreasing up to a point and then monotonically increasing*[22].

Proof. Expression (27) 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 $(1+\tau) \equiv a(\lambda)$

[22]This behavior can be described as single peaked from below and it is a form of quasi-convexity.

and $(1-\tau)\Gamma \equiv b(\lambda)$, where $a(\lambda)$ and $b(\lambda)$ are functions of λ. I study the sign of $\frac{d(a(\lambda)b(\lambda))}{d\lambda} = \frac{da}{d\lambda}b + a\frac{db}{d\lambda}$ in the interval $\lambda \in [1/2, 1]$. Using expression (32) in appendix B it can be checked that this expression is strictly negative for $\lambda = 1/2$. Notice that given proposition 2 $\frac{da}{d\lambda}b > 0$ and that given proposition 4 $a\frac{db}{d\lambda}$ can change sign and become positive at most once. Appendix B also shows that $a\frac{db}{d\lambda}$ starts negative and increases monotonically. Therefore it follows that the entire expression for the derivative $\frac{d(a(\lambda)b(\lambda))}{d\lambda}$ starts negative for $\lambda = 1/2$ and if it becomes positive, then it will continue to be positive. □

This result implies that in the case of $\pi > 1/2$, there are corner solutions: either $\lambda = 1/2$ or $\lambda = 1$ maximize ex-ante utility (27). 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 $\lambda = 1$. In other words agents either want to stay behind the veil of ignorance or want to remove it completely. It is important to stress the economic intuition behind this result. Firstly, increasing the level of information has a trade-off effect, on one hand it improves the allocation of individual effort but on the other hand it raises inefficient redistributive taxation, therefore ex-ante utility (27) does not increase monotonically in the level of information λ. Secondly, the convexity of the value of information Γ (lemma 2) implies that – as stated in proposition 6 – that ex-ante utility (27) is either monotonically decreasing or quasi-convex and therefore the corner solutions.

Case of $\pi < 1/2$. As explained with the analysis of the comparative static in the previous section, 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 (27) not clear. In the case in which the condition for the monotonic behavior given in proposition (5) applies, then both $(1-\tau^2)$ and Γ increase in λ for $\lambda \in [1/2, 1]$. Hence (27) is maximized for $\lambda = 0$ or $\lambda = 1$, i.e. for a perfectly informative signal.

Case of individual information λ^i. In order to gain further insights, I temporarily depart from the original 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. Plugging (25) and (26) into (11) I obtain the individual problem at $t = 0$:

$$\lambda^i = \arg\ \max\{(1-\tau)(\bar{k} + (1-\tau(\lambda))\Gamma(\lambda^i)/a) + \tau(\lambda)(\bar{k} + (1-\tau(\lambda))\Gamma(\lambda)/a) - (1-\tau(\lambda))^2\Gamma(\lambda^i)/2a\}. \quad (28)$$

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 $\Gamma(\lambda^i)$, which monotonically increases in the level of information (lemma 2). Therefore if individuals were free to autonomously choose the individual level of information, then everyone would choose to be perfectly informed and therefore the economy would be in a state which is identical to the case of perfect information $\lambda = 1$ in the original set-up. This result helps to approach the analysis of the next section.

6. Politico-Economic Equilibrium

In this section I consider the level of information λ in the economy to be an endogenous outcome and I introduce the concept of Politico-Economic equilibrium. Using the analysis of the previous section, I consider that the prevailing level of information λ in the economy is the one which maximizes the ex-ante utility. Such value of λ could be chosen 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 would agree on the same value of information. Once the level of information has been fixed as the ex-ante optimal, individual and aggregate choices and outcomes follow as already described in the previous sections. A definition follows:

Definition 1. I define a **Politico–Economic Equilibrium** as the prevailing level of information, beliefs and voted rate of redistribution $(\lambda, \mu_{\sigma_L}, \mu_{\sigma_H}, \tau)$ such that
(i) the prevailing level of information is ex ante optimal, i.e. λ = arg max u_0,
(ii) beliefs are bayesan-rational, i.e. beliefs μ_{σ_L} and μ_{σ_H} are respectively given by (4) and (5),
(iii) the prevailing rate of redistribution τ is the ideal rate (19) of is the median voter.

Analyzing the case of $\pi > 1/2$, the results of the previous section show that both minimum information ($\lambda = 1/2$) and perfect information ($\lambda = 1$) can be ex-ante optimal. It is easy to construct numerical examples in which both $\lambda = 1/2$ and $\lambda = 1$ are global maxima of the ex-ante utility function (27). Plugging $\lambda = 1/2$ and $\lambda = 1$

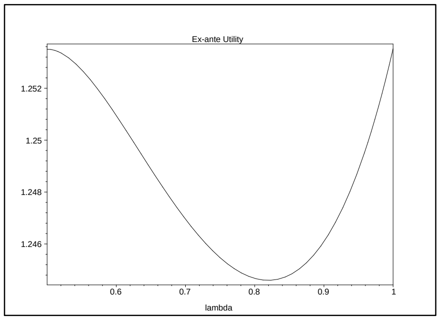
in (27) it can be easily computed that $u_0|_{\lambda=\frac{1}{2}} = (\bar{\theta})^2$ and that $u_0|_{\lambda=1} = \frac{(\bar{\theta^2})^2(3\bar{\theta^2}-2\theta_L^2)}{(2\bar{\theta^2}-\theta_L^2)^2}$. While the set of parameters such that both $\lambda = 1/2$ and $\lambda = 1$ are global maxima of u_0 has zero measure, it follows that there are sets with positive measure such that both $\lambda = 1/2$ and $\lambda = 1$ are local maxima. It is not immediate to find inequality relations on the parameters stating which of the two maxima is the global one, but numerical exercises show clearly that increasing π or the difference $\theta_H - \theta_L$ will imply that the value of $u_0|_{\lambda=1}$ increases relatively to $u_0|_{\lambda=\frac{1}{2}}$. Therefore the multiplicity of equilibria can be interpreted as saying that societies with minimal differences in the parameters π, θ_L, θ_H may find very different levels of information to be optimal.

Example of Multiple Politico-Economic Equilibria. I present a numerical example with multiple Politico-Economic Equilibria. I consider the following values of the parameters: $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 0.5$, $k = 0$. I plug those values in (4), (5), (7) (19), (16) and consequently those expressions in (27). I obtain a map of the ex-ante utility (27) in λ, which I plot in figure 1a.

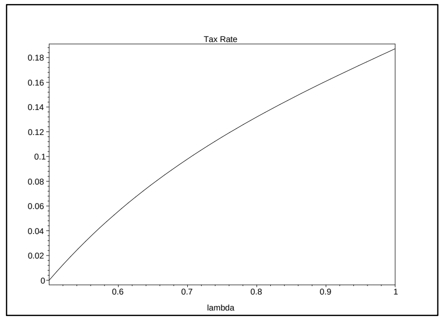
I verify that the function has two global maxima for $\lambda = 1/2$ and $\lambda = 1$ with value 1.25352. Therefore given the value of the parameters, both perfect information and minimum information are ex-ante optimal for the society. Numerical exercises show that a society with a higher (lower) value of π or higher (lower) value in the difference $\theta_H - \theta_L$ than the specified ones would find $\lambda = 1$ ($\lambda = 1/2$) to be optimal.[23]

Interpretation of the multiple Politico-Economic Equilibria. I plot τ (19) and Γ (16) as functions of γ in figures 1b and 1c respectively. As shown by lemma 2 and proposition 2, the uninformative equilibrium is characterized by a lower τ and an higher Γ than the informative equilibrium. The two variables have opposite effects on ex-ante utility (27), hence as shown by proposition 6 increasing the level of information λ has a trade-off effect and multiple equilibria are possible. I further interpret the two equilibria in terms of effort and output. I plotting the expressions of the level of effort exerted by the agents who observe σ_L (20) and σ_H (21) as functions of λ, in figures 1d and 1e respectively. I also plot the expression of aggregate effort (22) in figure 1f. The expression of optimal effort (12) shows that the greater is τ and the lower is the optimal effort, hence the informative equilibrium is characterized by a

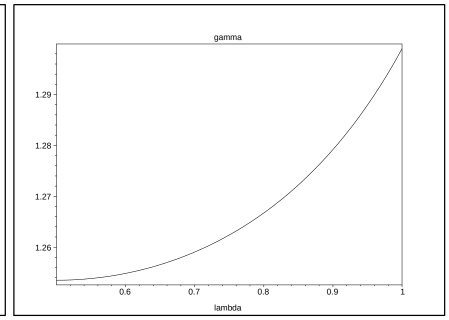
[23]As already stressed, while the set of parameters such that both $\lambda = 1/2$ and $\lambda = 1$ are global maxima of (27) has zero measure, there are sets with positive measure such that both $\lambda = 1/2$ and $\lambda = 1$ are local maxima.



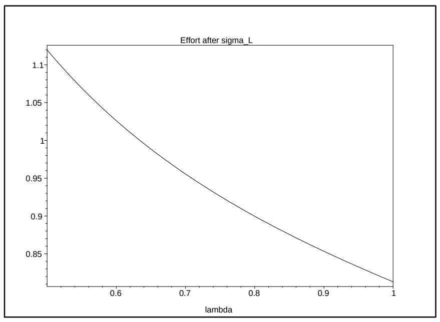
(A) Ex-ante welfare u_0



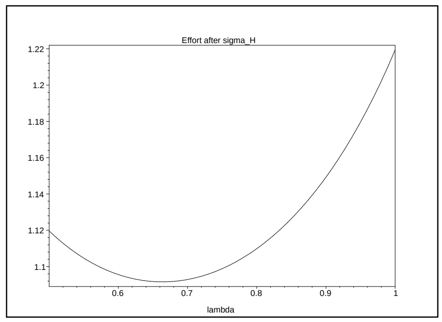
(B) Prevailing tax rate τ



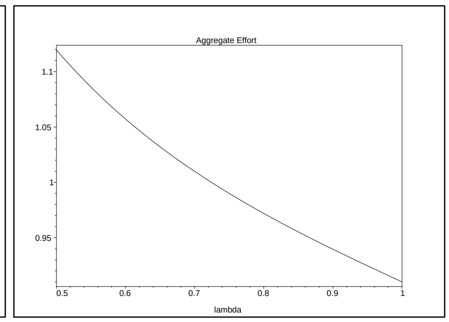
(C) Γ



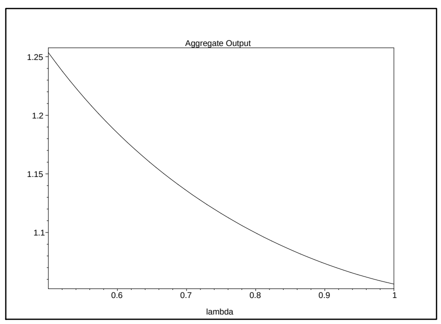
(D) Effort given σ_L



(E) Effort given σ_H



(F) Aggregate effort



(G) Aggregate Output

FIGURE 1. Plots of endogenous variables in λ, for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 0.5$, $k = 0$.

severe moral hazard problem as τ is at the maximum level. It is less immediate to notice an opposite effect of adverse selection. Figures 1d and 1e show that the greater is the precision of the signal and the more separated is the level of effort exerted by the two groups. When the signal is completely uninformative everyone pools to the same level of effort, 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. As shown by propositions 3 and 4 the moral hazard effects dominates the comparative statics of aggregate effort and output hence figures 1f and 1g respectively show that both aggregate effort and output are maximized at the uninformative equilibrium.

The uninformative equilibrium can be interpreted as a US-type equilibrium. In this equilibrium both groups of agents hold the same belief and exert the same levels of effort (pooling equilibrium). In this equilibrium, the tax rate is at the minimum level, whereas aggregate effort and output are at the maximum level. The informative equilibrium can instead be interpreted as a Europe-type equilibrium. In this equilibrium the two groups of agents have separated beliefs about the value of the 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). The distortive effect of taxation results in lower aggregate effort and aggregate output than those at the uninformative equilibrium. As mentioned in the introduction, my model interprets different (US vs Europe type) politico and economic outcomes as originated by different optimal informative cultures, where different (US vs Europe type) informative cultures present the following features: (i) more vs less widespread belief that effort more than luck determines individual wealth (ii) more vs less widespread belief that opportunities are equal (iii) more vs less separating educational systems.

7. Interpretation and Generalization of the Results

The result of the last section should not be interpreted as stating that Europeans are perfectly informed whereas Americans are not. The first consideration to be made is that in a more general set-up it does not have to be the case that the more informative equilibrium is characterized by perfect information: in section C, taking into account the possibility of heterogeneous endowments, I show that interior values of λ can be optimal. 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 can get to know is the average ability and a fraction π $(1-\pi)$ of agents has average ability 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).[24]

The result should instead be interpreted as showing the possibility and the implications of different optimal information cultures (more versus less separating). 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 (27), 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:

Proposition 7. *Given the ex-ante objective function (27), if $\tau \in [0, 1]$ is part of a politico-economic equilibrium, then the higher the rate of redistribution τ and the higher the level of information λ in the equilibrium.*

Proof. In a politico-economic equilibrium, λ= 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 $\tau(\lambda') > \tau(\lambda'')$. □

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 (27) 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. It is important to notice that the theorem applies independently from the type of comparative statics. Proposition 7 shows that in the case of multiple equilibria, a US-(Europe-) type equilibrium is relatively characterized by: (i) a less (more) informative signal

[24] The fact that in such cases the true distribution of the θ's remain unknown shows that a more precise signal in the Europe-type equilibrium does not imply that Europeans get to know the truth.

and therefore (ii) less (more) separated beliefs, (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. The introduction has motivated how this interpretation is supported by empirical and anecdotal evidence and how the features of the two equilibria seems to offer new insights about the observed political and economic features of different societies. This result is general and robust as it does not depend on the assumption of homogenous endowments[25], or on the underlying distribution of the abilities[26], or again on the linearity of the utility function[27].

8. Conclusion

This paper presented a simple theoretical model in order to analyze the role of incomplete information in the determining 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 answer natural policy questions as the level of information which maximizes welfare, output and other politico-economic outcomes.

The theoretical model presented in the chapter interprets a US-type vs a Europe-type politico-economic equilibrium as characterized by relative (i) higher adverse selection – individual beliefs and effort levels are pooling to similar levels despite underlying heterogeneity in the true distribution of the return on effort – (ii) lower redistribution (iii) lower moral hazard – redistribution is low and this does not distort individual effort much (iv) higher aggregate effort and output. Conversely the Europe-type politico-economic equilibrium is interpreted as an equilibrium characterized by relative (i) lower adverse selection (ii) higher redistribution (iii) higher

[25]In appendix C I show that heterogenous endowments introduce technical complications because, in such case, changing the level of information changes the identity of the median voter, different voters prefer different tax rates given different endowments, hence the comparative statics are generally discontinuous; nevertheless the ex-ante optimal λ still has to maximize the object (27) and therefore proposition 7 still applies.

[26]The objective function (27) would be the same with a different underlying distribution of the abilities.

[27]As a further test of the model's robustness, appendix D introduces risk aversion in the model and shows that the conclusions do not change.

moral hazard – taxation is high and this diminish individual effort (iv) low aggregate effort and output. The two equilibria are both ex ante optimal. The results are robust to variations of the basic framework, as the introduction of heterogenous endowments and a concave utility function.

It is worthy to stress that the presented model does not give clear predictions about the heterogeneity of exerted effort and the levels of inequality in two different equilibria. In the basic version with homogenous endowment k across agents, the non informative equilibrium ($\lambda = 1/2$) is a pooling equilibrium where every agent exerts the same effort, whereas in the full informative equilibrium ($\lambda = 1$) effort levels are separated and hence output is relatively more heterogenous (pre-tax inequality is higher). This should not lead to conclude that the model predicts that the Europe-type equilibrium is characterized by higher inequality and more separated effort levels than the US-type equilibrium which, as discussed in the introduction, would contrast some empirical evidence. In the more general exposition of the model with heterogenous endowments, where interior values of λ can be welfare maximizing, it can be the case that despite the fact that the more informative equilibrium is characterized by more separated beliefs the fact that redistribution is higher implies that effort levels are less separated and that output before taxes is less heterogenous. In such case, the driving force behind the fact that effort levels are less separated in the Europe-type equilibrium would be the distortive effect of taxation.[28] Moreover, in the model with heterogenous endowments the fact that the rate of redistribution does not change continuously and monotonically in λ implies that the separation of effort levels does not have to increase in the level of information.[29] In conclusion, the model focused on the determinants of different beliefs and rates of redistribution but cannot say much about the levels of inequality as those depend on the values of the endowments – both directly and through their role in affecting the voted redistribution – and in the model there is no specification of the wealth generating process, which should be naturally modeled as dynamic. Therefore, the study of dynamic inequality and mobility in a framework with endogenous incomplete information could be an interesting problem for future research.[30]

---

[28]This point seems to have some robust empirical support, see for example Prescott (2004).

[29]This possibility is shown in the example in appendix C, see figure 2i.

[30]A first study of such problem can be found in Gabrieli (2009).

References

Adams, J. T. (1931): *The epic of America*. Boston MA: Little, Brown and company.

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.

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.

Alesina, A., and E. La Ferrara (2005): "Preferences for redistribution in the land of opportunities," *Journal of Public Economics*, 89(5-6), 897–931.

Alesina, A., and D. Rodrik (1994): "Distributive Politics and Economic Growth," *The Quarterly Journal of Economics*, 109(2), 465–490.

Benabou, R. (1996): "Inequality and Growth," in *NBER Macroeconomics Annual 1996*, ed. by B. S. Bernanke, and J. Rotemberg. Cambridge, MA: MIT Press.

——— (2000): "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, 90(1), 96–129.

——— (2005): "Inequality, Technology and the Social Contract," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. Amsterdam: Elsevier.

Benabou, R., and E. A. Ok (2001): "Social Mobility And The Demand For Redistribution: The POUM Hypothesis," *The Quarterly Journal of Economics*, 116(2), 447–487.

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 *Improving America's Schools: The Role of Incentives*, ed. by E. A. Hanushek, and D. W. Jorgenson, pp. 111–146. National Academy Press.

——— (1997): "The Effect of National Standards and Curriculum-Based Exams on Achievement," *American Economic Review*, 87(2), 260–264.

Cervellati, M., J. Esteban, and L. Kranich (2006): "Redistributive Taxation with Endogenous Sentiments," IZA Discussion Papers 2312, Institute for the Study of Labor (IZA).

Corneo, G., and H. P. Gruner (2002): "Individual preferences for political redistribution," *Journal of Public Economics*, 83(1), 83–107.

De Tocqueville, A. (1835): *De la Democratie en Amerique*. Paris.

Farmer, R. E. A. (1990): "Rince Preferences," *The Quarterly Journal of Economics*, 105(1), 43–60.

Fong, C. (2001): "Social preferences, self-interest, and the demand for redistribution," *Journal of Public Economics*, 82(2), 225–246.

Gabrieli, T. (2007): "Beliefs and Redistributive Politics under Incomplete Information," The Warwick Economics Research Paper Series (TWERPS) 821, University of Warwick, Department of Economics.

——— (2009): "Redistributive Politics under Optimally Incomplete Information," Ph.D. thesis, University of Warwick, UK.

Hassler, J., J. V. Rodriguez-Mora, K. Storesletten, and F. Zilibotti (2003): "The Survival of the Welfare State," *American Economic Review*, 93(1), 87–112.

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.

McElroy, J. H. (2006): *Divided We Stand: The Rejection of American Culture since the 1960s*. Lanham, MD: Rowman and Littlefield Publishers.

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

Perotti, R. (1993): "Political Equilibrium, Income Distribution, and Growth," *Review of Economic Studies*, 60(4), 755–76.

——— (1994): "Income distribution and investment," *European Economic Review*, 38(3-4), 827–835.

——— (1996): "Growth, Income Distribution, and Democracy: What the Data Say," *Journal of Economic Growth*, 1(2), 149–87.

Persson, T., and G. Tabellini (1994): "Is Inequality Harmful for Growth?," *The American Economic Review*, 84(3), 600–621.

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

Prescott, E. C. (2004): "Why do Americans work so much more than Europeans?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1), 2–13.

Radner, R., and J. Stiglitz (1984): "A non-concavity in the value of information," in *Bayesian Models in Economic Theory*, ed. by Bayor, and Kihlstrom. Elsevier.

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

SAINT-PAUL, G. (2001): "The Dynamics of Exclusion and Fiscal Conservatorism," *Review of Economic Studies*, 4, 275–302.

SUHRCKE, M. (2001): "Preferences for Inequality: East vs. West," Innocenti Working Papers inwopa01/16, UNICEF Innocenti Research Centre.

WOESSMANN, L. (2003): "Central Exit Exams and Student Achievement: International Evidence," in *No Child Left Behind? The Politics and Practice of School Accountability*, ed. by P. E. Peterson, and M. R. West, pp. 292–323. Washington, D.C.: Brookings Institution Press.

——— (2009): "International Evidence on School Tracking: a review," Discussion paper, CESifo DICE Report 1/2009.

APPENDIX A. PROOF OF LEMMA 2

In order to prove (i) (monotonicity), I compute the expression of the first derivative of Γ with respect to λ:

$$\begin{aligned}\frac{d\Gamma}{d\lambda} =&\pi\theta_L \frac{d\left(\lambda\theta(\mu^i_{\sigma_L}) + (1-\lambda)\theta(\mu^i_H)\right)}{d\lambda} + \\ &(1-\pi)\theta_H \frac{d\left((1-\lambda)\theta(\mu^i_{\sigma_L}) + \lambda\theta(\mu^i_H)\right)}{d\lambda} = \\ &\pi\theta_L\left(-\frac{\pi(\pi-1)^2(\theta_H-\theta_L)(2\lambda-1)}{(2\pi\lambda+1-\lambda-\pi)^2(2\pi\lambda-\lambda-\pi)^2}\right) + \\ &(1-\pi)\theta_H\left(-\frac{\pi^2(\pi-1)(\theta_H-\theta_L)(2\lambda-1)}{(2\pi\lambda+1-\lambda-\pi)^2(2\pi\lambda-\lambda-\pi)^2}\right) = \\ &\frac{\pi^2(1-\pi)^2(\theta_H-\theta_L)^2(2\lambda-1)}{(2\pi\lambda+1-\lambda-\pi)^2(2\pi\lambda-\lambda-\pi)^2},\end{aligned} \quad econometrica \quad econometrica$$

which is ≥ 0 for $\lambda \geq 1/2$.

In order to prove (ii) (convexity), I compute the expression of the second derivative of Γ with respect to λ:

(29)
$$\frac{\partial^2\Gamma}{(\partial\lambda)^2} = \frac{2\pi^2(1-\pi)^2(\theta_H-\theta_L)^2(1+12\pi\lambda(1-\lambda)(1-\pi)-3\pi(1-\pi)-3\lambda(1-\lambda))}{\left(\pi\lambda+(1-\pi)(1-\lambda)\right)^3\left(\pi(\lambda-1)+\lambda(\pi-1)\right)^3}.$$

The expression is positive as it can be proved that the term $(1+12\pi\lambda(1-\lambda)(1-\pi)-3\pi(1-\pi)-3\lambda(1-\lambda))$ (call this X) 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 therefore positive. Hence the term X increases in λ; it is immediate that X is equal to zero for the smallest value of λ, $\lambda = 1/2$. Therefore for any value of π and λ, X is positive.■

APPENDIX B. PROOF OF PROPOSITION 4

It is useful to plug (19) into (23) and re-express this as

$$k + \frac{\Gamma^2}{a(2\Gamma-\theta(\mu)^2)}, \tag{30}$$

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:

$$\frac{2\Gamma^2\frac{\partial\Gamma}{\partial\lambda} - 2\theta(\mu_{\sigma_L})^2\Gamma\frac{\partial\Gamma}{\partial\lambda} + 2\theta(\mu_{\sigma_L})\Gamma^2\frac{\partial\theta(\mu_{\sigma_L})}{\partial\lambda}}{a^2\left(2\Gamma-\theta(\mu_{\sigma_L})^2\right)^2} \tag{31}$$

where

$$\frac{\partial\theta(\mu_{\sigma_L})}{\partial\lambda} = -\frac{\pi(1-\pi)(\theta_H-\theta_L)}{(\pi\lambda+(1-\lambda)(1-\pi))^2} \leq 0$$
$$\frac{\partial\Gamma}{\partial\lambda} = \frac{\pi^2(1-\pi)^2(2\lambda-1)(\theta_H-\theta_L)^2}{(\pi\lambda+(1-\pi)(1-\lambda))^2(\pi(\lambda-1)+\lambda(\pi-1))^2} \geq 0.$$
(32)

The denominator of (31) 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

$$(33)\qquad (\Gamma - \theta(\mu_{\sigma_L})^2)\frac{\partial\Gamma}{\partial\lambda} + \theta(\mu_{\sigma_L})\Gamma\frac{\partial\theta(\mu_{\sigma_L})}{\partial\lambda}.$$

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 (31) is negative for $\lambda = 1/2$. I compute the second derivative of (33):

$$(34)\quad (\Gamma - \theta(\mu_{\sigma_L})^2)d^2\Gamma + (d\Gamma)^2 - \theta(\mu_{\sigma_L})d\theta(\mu_{\sigma_L})d\Gamma + \Gamma(d\theta(\mu_{\sigma_L}))^2 + \theta(\mu_{\sigma_L})\Gamma d^2\theta(\mu_{\sigma_L}),$$

where

$$\frac{\partial^2\theta(\mu_{\sigma_L})}{(\partial\lambda)^2} = \frac{2\pi(1-\pi)(2\pi-1)(\theta_H-\theta_L)}{(\pi\lambda+(1-\pi)(1-\lambda))^3} \geq 0$$
$$\frac{\partial^2\Gamma}{(\partial\lambda)^2} = \frac{2\pi^2(1-\pi)^2(\theta_H-\theta_L)^2(1+12\pi\lambda(1-\lambda)(1-\pi)-3\pi(1-\pi)-3\lambda(1-\lambda))}{(\pi\lambda+(1-\pi)(1-\lambda))^3(\pi(\lambda-1)+\lambda(\pi-1))^3}.$$
(35)

Notice that $\frac{\partial^2\Gamma}{(\partial\lambda)^2} \geq 0$ as it has already been proved in Appendix A.

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, (33) is strictly positive and therefore (31) can change sign at most once in the range $\lambda \in [1/2, 1]$. Therefore in the range $\lambda \in [1/2, 1]$ (31) is either always negative or negative up to a point and then always positive, this implies the quasi-convexity.

In order to prove (ii) notice that 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 $\bar{y} = k + \overline{\theta}^2/a$, the value of the aggregate output for

$\lambda = 0$ and $\lambda = 1$ is $\bar{y} = k + \frac{\overline{\theta^2}^2}{a(2\overline{\theta^2} - \theta_L^2)}$. For the output to be greater at $\lambda = 1/2$ than $\lambda = 1$, the condition to be satisfied is the following:

$$(\pi\theta_L + (1-\pi)\theta_H)^2(2\pi\theta_L^2 + 2(1-\pi)\theta_H^2 - \theta_L^2) - (\pi\theta_L^2 + (1-\pi)\theta_H^2)^2 \geq 0 \tag{36}$$

i.e.

$$(\theta_L - \theta_H)(-1+\pi)\left(-2\,\theta_H{}^2\pi^2\theta_L + 2\,\theta_H{}^3\pi^2 - 2\,\theta_H\pi^2\theta_L{}^2 + 2\,\pi^2\theta_L{}^3 - 3\,\theta_H{}^3\pi + \pi\,\theta_L\theta_H{}^2 + 2\,\theta_H\pi\,\theta_L{}^2 + \theta_H{}^3 + \theta_L\theta_H{}^2\right) \geq 0 \tag{37}$$

Observe that

$$-2\,\theta_H{}^2\pi^2\theta_L + 2\,\theta_H{}^3\pi^2 - 2\,\theta_H\pi^2\theta_L{}^2 2\,\pi^2\theta_L{}^3 - 3\,\theta_H{}^3\pi + \pi\,\theta_L\theta_H{}^2 + 2\,\theta_H\pi\,\theta_L{}^2 + \theta_H{}^3 + \theta_L\theta_H{}^2 = \tag{38}$$

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

Observe that

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

Hence, after a factorization condition (36) can be rewritten as

$$(\theta_L - \theta_H)(-1+\pi)\big(2\,\pi^2\theta_L{}^3 + 2\,\pi(1-\pi)\theta_H\theta_L{}^2 + (1-\pi)\theta_H^2(2\pi\theta_L - 2\pi\theta_H + \theta_L + \theta_H\big), \tag{41}$$

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 (36) is satisfied.■

APPENDIX C. 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 version of (18) – the indirect utility in τ – is given by the following expression:

$$u_t^i = \tau(k^i - \bar{k}) + (1-\tau)^2\theta(\mu^i)^2/a + \tau(1-\tau)\Gamma/a - (1-\tau)^2\theta(\mu^i)^2/2a. \tag{42}$$

Assumption 1 still assures that expression (42) is strictly concave as the variable k does not enter the second order conditions. The new expression for the ideal tax rate of agent i follows:

$$\tau(k^i, \mu^i) = 1 - \frac{1 + \frac{a(k^i - \bar{k})}{\Gamma}}{2 - \frac{\theta(\mu^i)^2}{\Gamma}}. \tag{43}$$

As explained by Benabou and Tirole (2006), the numerator of (43) indicates that a lower relative endowment $(k^i - \bar{k})$ naturally increases the desired tax rate and that wether progressive or regressive, such distributive goals must be traded off against distortions to the effort-elastic component of the tax base (moral hazard problem). As before, the denominator indicates that increases in the prospects of upper mobility decrease the ideal tax rate. 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 $\alpha \in (1/2, 1]$ $\big(\alpha \in [0, 1/2)\big)$ the majority of the agents has an endowment $k^i = k_L$ $\big(k^i = k_H\big)$. If $p_{\sigma_L} > 1/2$ $\big(p_{\sigma_L} < 1/2\big)$, the majority of the agents holds a belief μ_{σ_L} $\big(\mu_{\sigma_H}\big)$ at $t = 1$. I analyze the voting outcome analyzing the various possible cases.

Voting outcome with heterogenous endowments. Before proceeding with the various cases notice that the fact that $\lambda \geq 1/2$ implies that $\mu_{\sigma_H} \geq \mu_{\sigma_L}$ and therefore the following ranking of preferred tax rates: $\tau(k_H, \mu_{\sigma_H}) \leq \min\{\tau(k_H, \mu_{\sigma_L}), \tau(k_L, \mu_{\sigma_H})\} \leq \max\{\tau(k_H, \mu_{\sigma_L}), \tau(k_L, \mu_{\sigma_H})\} \leq \tau(k_L, \mu_{\sigma_L})$.[31]

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

Case 1: $\alpha \geq 1/2$ and $\pi \geq 1/2$. $\alpha \geq 1/2$ 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 $\tau(k_L, \cdot)$ 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: $\alpha \geq 1/2$ and $\pi \leq 1/2$. $\pi \leq 1/2$ and $\lambda \geq 1/2$ 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 2** has the same outcome of Case 1.2.

Case 3: $\alpha \leq 1/2$ and $\pi \geq 1/2$. $\alpha \leq 1/2$ 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.

Case 3.1: $(1-\alpha)p_{\sigma_L} > 1/2$. 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.

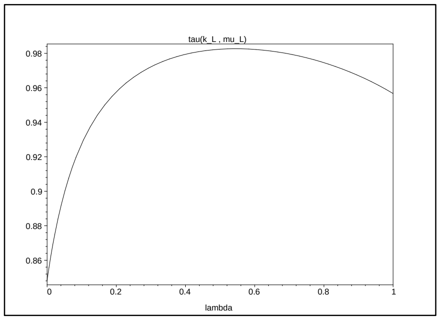
Case 3.2: $(1-\alpha)p_{\sigma_L} < 1/2$. If $\tau(k_L, \mu_{\sigma_H}) > \tau(k_H, \mu_{\sigma_L})$ then the pivotal group is the one who prefers $\tau(k_H, \mu_{\sigma_L})$ whereas if $\tau(k_H, \mu_{\sigma_L}) > \tau(k_L, \mu_{\sigma_H})$ then the pivotal group is the one who prefers $\tau(k_L, \mu_{\sigma_H})$.

Case 4: $\alpha \leq 1/2$ and $\pi \leq 1/2$. $\pi \leq 1/2$ and $\lambda \geq 1/2$ together imply that $p_{\sigma_L} \leq 1/2$, therefore $(1-\alpha)p_{\sigma_L} > 1/2$ is never verified and hence Case 3.1 is never verified. Therefore **Case 4** has the same outcome as case 3.2.

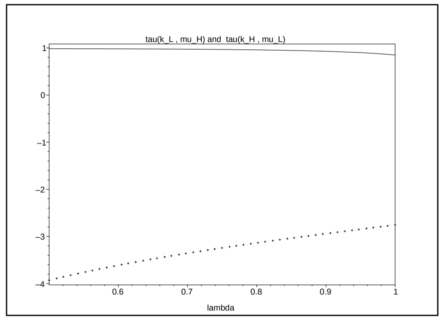
It is important to notice how heterogenous endowments can imply discontinuous comparative statics. In order to see this assume to be in case 1.1 where the pivotal tax rate is $\tau(k_L, \mu_{\sigma_L})$. If λ increases the pivotal tax rate remains $\tau(k_L, \mu_{\sigma_L})$ and increases monotonically till $\tau(k_L, \theta_L)$ 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 and 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 in a discontinuous way. Discontinuous comparative statics imply the possibility of interior welfare maximizing values of λ even if the comparative statics are monotonic. The following numerical example shows this possibility.

Example of multiple equilibria with discontinuous comparative statics. In the case of heterogenous endowments the ex-ante objective function is still given by (27), where $k = \bar{k} = \alpha k_L + (1-)k_H$. Consider the following values: $\alpha = 0.8$, $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$. Such values imply that $p_{\sigma_L} = 0.4\lambda + 0.3$. If there is a value λ^* such that $\alpha(0.4\lambda^* + 0.3) > 1/2$, then λ^* is a point of discontinuity. 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 $\lambda^* \simeq 0.81$.

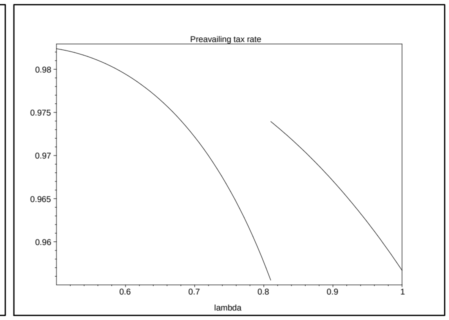
I analyze the object u_0^i as a function of λ. For $\lambda > \lambda^*$ the voted tax rate is $\tau = \tau(k_L, \mu_{\sigma_L})$. I plot this function of λ in figure 2a. For $\lambda < \lambda^*$ the voted tax rate is the greater between $\tau(k_L, \mu_{\sigma_H})$ and $\tau(k_H, \mu_{\sigma_L})$. I plot both functions of λ in figure 2b. The figure shows that $\tau(k_L, \mu_{\sigma_H})$ is greater throughout the interval, therefore for $\lambda < \lambda^*$ the voted tax rate is $\tau(k_L, \mu_{\sigma_H})$. I plot the voted tax rate in figure 2c. 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. I plot u_0^i, τ, Γ and the optimal values of individual and aggregate effort respectively in figures 2d, 2e, 2f, 2g, 2h. I also plot figures 2f and 2g together in figure 2i, where the thicker line represents figure 2f.



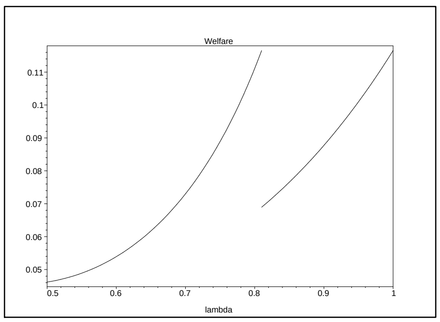
(A) $\tau(k_L, \mu_L)$



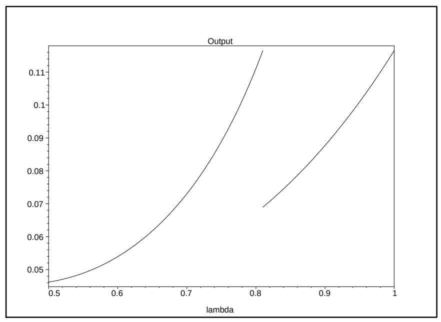
(B) $\tau(k_L, \mu_H)$ (continuous line) and $\tau(k_H, \mu_L)$ (pointed line)



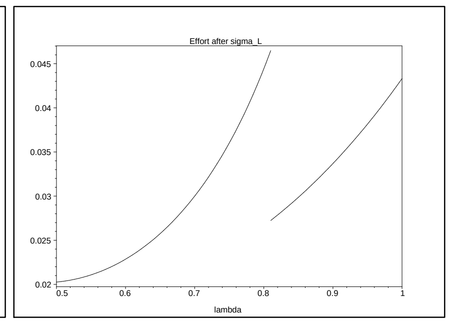
(C) Prevailing tax rate



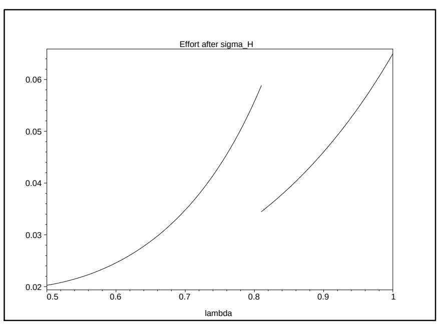
(D) Ex-ante welfare function



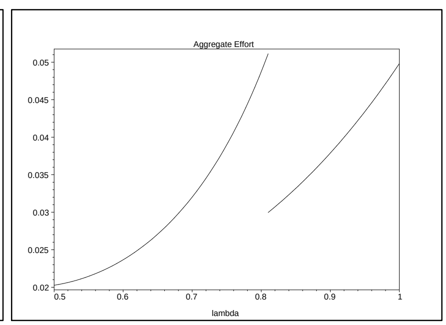
(E) Aggregate Output



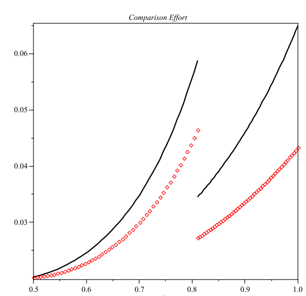
(F) Effort after the observation of σ_L



(G) Effort after the observation of σ_H



(H) Aggregate effort



(I) Comparison

FIGURE 2. Endogenous variables with heterogenous endowments, for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

Appendix D. Analysis with concave utility

As already explained in section 5, in the model that I have presented the trade-off effect of information – and hence the possibility of multiple ex-ante optimal levels of information – arise because on one hand information increases ex-ante inefficient taxes (hence increases the moral hazard problem) and on the other hand information improves the efficiency of effort's allocations (hence reduces the adverse selection problem). One natural question to ask is wether this trade-off – and hence the result of multiple equilibria – is robust to the introduction of risk aversion in the problem. With a concave ex-ante utility function in consumption, ceteris paribus, redistribution is ex-ante efficient. On the other hand redistribution still decreases individual effort and therefore decreases the amount of output which is redistributed, hence the overall effect of taxation on ex-ante utility is not clear a priori. Moreover, with a concave ex-ante utility function in consumption not even the overall effect of information is a-priori clear. This is the case even when it is ignored the effect that information has on the prevailing tax rate, in other words when the level of redistribution is fixed. The reason for this is that on one hand information separates the levels of exerted effort implemented (which is ex-ante un-optimal given the concavity of the utility function) but on the other hand information improves effort's allocations and therefore it increases the amount of output which is redistributed. Hence in the case in which information increases the prevailing tax rate, a concave utility function implies that increasing the level of information has two positive effects: to increase ex -ante optimal taxes and to increase aggregate effort and output (which will be redistributed). Increasing the level of information has also two negative effects: to separate the levels of effort and output (which is ex ante un-optimal given concavity) and to decrease individual effort through higher taxation and therefore to decrease the output which will be redistributed.

In order to gain insights about the overall effect of information and to check the robustness of the result that ex-ante utility is nor monotonic nor concave in the level of information I present some numerical examples. I introduce a utility function which is concave in consumption at time 0, when welfare is evaluated, but I maintain the same linear utility function for the rest of the problem, namely when both taxes are voted for and when effort is chosen. The reason for doing this is to maintain the tractability. As it is shown in the analysis of Meltzer and Richard (1981) and followers, using a concave function for the choice of effort and voting implies that the prevailing tax rate is not an explicit function. [32]

Numerical examples with concave utility. Without loss of generality I fix that $k^i = 0$ for all i and that $a = 1$ in order to simplify the computations. Agent i utility function at time 0 is concave in consumption and it is given by the following expression:

$$u_0^i = E_0^i[\frac{1}{\gamma}(c^i)^\gamma - (e^i)^2/2], \tag{44}$$

with $\gamma < 1$. The period 1 problem is still the one described in section 2, hence the expression for optimal individual effort is still (12) and the expression for the aggregate product of effort is still given by (25). I plug those expressions together with (9) and (1) into (44) and I obtain the expression for expected utility conditional on the observation of σ_L:

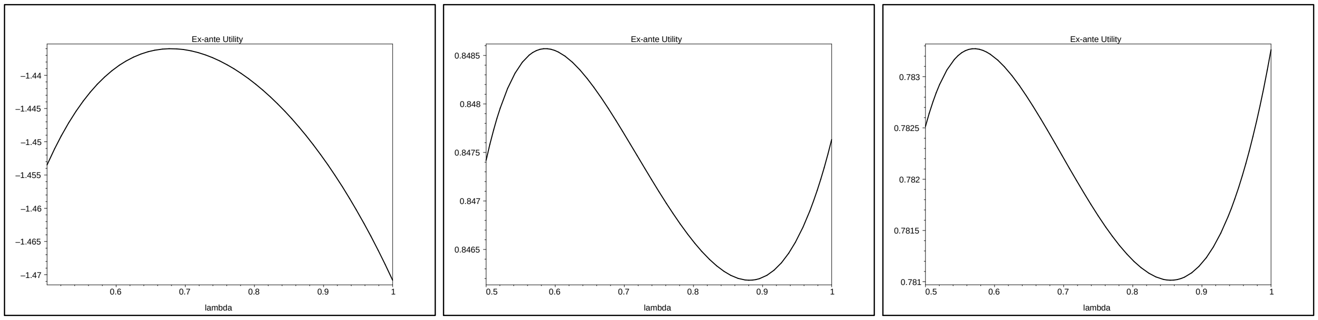
$$u_{0\sigma_L} = \mu_L \frac{1}{\gamma}((1-\tau)\theta_L e_{\sigma_L} + \tau(1-\tau)\Gamma)^\gamma + (1-\mu_L)\frac{1}{\gamma}((1-\tau)\theta_H e_{\sigma_L} + \tau(1-\tau)\Gamma)^\gamma - (e_{\sigma_L})^2/2, \tag{45}$$

and a symmetric expression given the observation of σ_H, where e_{σ^i} is given by expression (12) and it represents optimal individual effort conditional on the observation of σ^i. Using those expressions, I can rewrite expression (44) as

$$\begin{aligned} u_0^i = p_{\sigma_L} u_{\sigma_L} + (1 - p_{\sigma_L}) u_{\sigma_H} = \\ \pi\lambda\frac{1}{\gamma}((1-\tau)\theta_L e_{\sigma_L} + \tau(1-\tau)\Gamma)^\gamma + \\ (1-\pi)(1-\lambda)\frac{1}{\gamma}((1-\tau)\theta_H e_{\sigma_L} + \tau(1-\tau)\Gamma)^\gamma + \\ \pi(1-\lambda)\frac{1}{\gamma}((1-\tau)\theta_L e_{\sigma_H} + \tau(1-\tau)\Gamma)^\gamma + \\ (1-\pi)\lambda\frac{1}{\gamma}((1-\tau)\theta_H e_{\sigma_H} + \tau(1-\tau)\Gamma)^\gamma - (1-\tau)^2\Gamma/2. \end{aligned} \tag{46}$$

It is clear from expression (46) that on one hand τ has the positive effect of redistributing and on the other hand τ has the negative effect to decrease the amount of resources which are redistributed. It is also clear that also increasing the level of information λ has two opposite effects on welfare. One one hand there is a positive

[32]The type of utility function which I introduce can be shown to belong to the class of RINCE Preferences introduced by Farmer (1990). This class of preferences imply that the utility is concave over non-stochastic outcomes (like the period 0 utility function), but it becomes risk neutral over stochastic outcomes (like the period 1 utility function).



(A) Ex-ante welfare for $\gamma = -1$

(B) Ex-ante welfare for $\gamma = 0.8$.

(C) Ex-ante welfare for $\gamma = 0.845$.

FIGURE 3. Endogenous variables with concave utility, for $\pi = 0.8$, $\theta_L = 1$, $\theta_H = 1.5$ and different values of γ.

effect because Γ increases in λ (lemma 2) and on the other hand increasing λ separates the levels of effort implemented (lemma 1) which is non optimal given the concavity. In order to check for the overall effect of the level of information on ex-ante utility I proceed with some numerical examples. I consider the case in which $p_{\sigma_L} \geq 1/2$ and the majority of agents observe σ_L so that the prevailing tax rate is τ_{σ_L}. It follows from proposition 2 that in this case the voted tax rate τ increases in the level of information.

I consider a numerical example with a coefficient of risk aversion of $\gamma = -1$. I plot expression (46) as a function of λ in figure 3a. The figure shows that there is an interior solution in terms of λ. This example proves the possibility of non monotonicity of information.

Decreasing the coefficient of risk aversion implies that the beneficial effect of information through the tax rate is less valued. I Consider the case of $\gamma = 0.8$. I plot expression (46) as a function of λ in figure 3b. The figure shows that the optimal value of information is still unique but smaller than before. This example shows that also with a concave utility in consumption, ex-ante utility does not have to be concave in information.

It is also possible to have a case with multiple optimal values of information. I consider the case of $\gamma = 0.845$, I plot expression (46) as a function of λ in figure 3c. As before, the equilibrium with relatively higher (lower) information has higher (lower) taxes and more (less) separated effort choices, while the solution with lower (higher) information has lower (higher) taxes but less separated effort choices.