Redistributive Politics under Optimally Incomplete Information

by

Tommaso Gabrieli

A thesis submitted in partial fulfilment of the requirements

for the degree of

Doctor of Philosophy in Economics

University of Warwick, Department of Economics

November 2008

Contents

Acknowledgements

I am very grateful to my PhD supervisors Herakles Polemarchakis and Sayantan Ghosal for the guidance and the motivation. I also wish to thank Chuck Blackorby for his work and guidance as my PhD advisor in the first two years.

I found particularly helpful and motivating various feedbacks which I received along the path, for this, in no particular order, I wish to thank Amrita Dhillon, David Austen-Smith, Francis Block, Marco Battaglini, Sandhil Mullainathan, Roland Benabou and Jayasri Dutta.

I greatly acknowledge financial support from Fondazione di Piacenza e Vigevano, Italy and University of Warwick Department of Economics. On this, I wish to thank Norman Ireland for having very actively contributed to make my teaching job absolutely fantastic.

On a more personal note, I am immensely grateful to my parents and family as I evidently perceive that every my achievement is also theirs. There are too many friends and companions to be personally thanked for the support, help, smiles that kept me alive during this journey, but it will be hard to forget those that helped me looking at the Mystery among us. Among those I want to mention in particular Prem Fernando as he has always been there. Finally, I wish to express the greatest love and gratitude to my wife Sherry for keeping me smiling, reminding me of the most important things in life, being exceptionally supportive, loving and patient even as a PhD candidate herself.

Declaration

This dissertation is my own work, and it has not been submitted for a degree at another university.

Tommaso Gabrieli

University of Warwick, Department of Economics

November 2008

Abstract

This thesis wants to contribute to the understanding of the role of collective beliefs and incomplete information in the analysis of the dynamics of inequality, growth and redistributive politics. Extensive 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. The thesis presents a theoretical framework 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 framework is based on standard assumptions. Multiple equilibria originate from multiple welfare-maximizing levels of information for the society. Multiple welfare-maximizing 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 provides a new micro-foundation of incomplete information as an institutional feature and answers various macroeconomic policy questions with different models.

CHAPTER 1

Introduction

This thesis wants to contribute to the understanding of the role of collective beliefs and incomplete information in the analysis of the dynamics of inequality, growth and redistributive politics. More specifically, focusing on the role played by incomplete information, it offers new theoretical insights which can help to answer the following four questions:

(i) Why is the case that very different patterns of redistribution and social contracts are found across countries which are otherwise similar?

(ii) Why is it the case that also the beliefs which people hold about the underlying determinants of wealth and the extent of social mobility are very different across otherwise similar societies?

(iii) It appears that the societies which are characterized by the widespread belief according to which "hard work, self-discipline and other factors under individual control, other than luck, family of origin or other factors outside individual control determine individual wealth"(e.g. the US) redistribute less than the societies which are characterized by the opposite belief (e.g. European countries). Is it possible to describe those outcomes relating beliefs, political and economic outcomes in terms of different equilibria? And which are the driving forces behind those different outcomes?

(iv) Given that beliefs and incomplete information appear to play a crucial role in the determination of political and economy outcomes, which are the relative policy implications? Is there any role for institutions which can affect the degree of information in the economy?

The way in which I address questions (i)-(iii) is by offering new insights within a recently developed theoretical framework which analyzes the role played by beliefs in the determination of political and economic outcomes. This represents a still small but fast-growing field of research which is receiving increasing attention.[1] One of the contributions of my thesis is to take a first step towards a unification of such framework focusing on the role played by incomplete information. The way in which I answer question (iv) is by making a first step towards the development of a new theoretical framework which can link this recent literature which is focused on the role of beliefs to some more traditional strands of theoretical literature in macroeconomics. The most natural link is the one with the dynamic models studying income distribution and intergenerational mobility, but there is also a less obvious link to the literature on optimal taxation, especially in its most recent developments which go under the name of *new dynamic public finance*[2].

---

[1]As the specific references will make clear, the leading scholars of this field can be considered to be Alberto Alesina, Roland Benabou, Edward Glaeser and Thomas Piketty among others. Given the research question, it will be interesting to notice that most of the active scholars in this field are European economists with established working experiences in the US.

[2]See Golosov, Tsyvinski, and Werning (2006) for a review of the this recent literature.

Under a broader perspective, the thesis can be seen as a theoretical contribution to the understanding of the role of wealth inequality in macroeconomics, which has been a research topic under increasing attention since the early 90's. In order to clarify better the contribution of the thesis, in this chapter I present a critical review of the related literature and the main results of thesis, respectively in sections 1.1 and 1.2.

1.1. Historical review of literature

The related literature is very vast and encompasses different fields. Often those different strands of literature have been separated. It is useful to review those different fields in chronological order in order to have a better understanding of the way in which they developed and of the questions which are still unanswered.[3]

1.1.1. Income Distribution and Intergenerational Mobility in Neoclassical Macroeconomics. The interest in the distribution of income used to be central among classical economists. Distributional issues are still paramount for Post-Keynesian models focused on the distribution of income among factor of productions. It is the seminal paper of Stiglitz (1969) to be commonly considered the first modern analysis of the distribution of wealth and income among individuals. The model of Stiglitz presents a strong result of long run convergence in the dynamics of individual income which parallels the seminal result obtained by Solow (1956) in the context

[3]It is inevitable to be reductive, but deeper reviews of those different fields already exist: Bertola, Foellmi, and Zweimuller (2006), Piketty (1998), Aghion, Caroli, and Garcia-Penalosa (1999), Benabou (2005). I will refer to those when necessary.

of country income. In the model of Stiglitz agents are endowed with capital (accumulated factor) and labor (non accumulated factor), markets are competitive and both factors are paid at their marginal return. The assumptions of diminishing returns to capital and of an identical concave saving function across individuals imply that individual wealth increases over time in a concave fashion and eventually converges to a steady state value which does not depend on the initial level of wealth. In other words, in the model inequality across families is solely determined by the differences in the non accumulated factors (i.e. the differences in individual skills) and when all families are equally endowed with the non accumulated factor (i.e. skills are homogenous across families) every family converges to the same level of wealth. It is also shown that redistributive taxation decreases inequality and increases the speed of convergency, but it does not have any effect on aggregate output, specifically because it does not change individual saving behavior which is exogenous. These results rely on the specific assumptions of the model and the following contributions have removed some of the original exogenous elements seeking deeper micro-foundations. For example, one element which can result in the absence of convergency and in the persistence of initial wealth differences is the possibility of different fertility behaviors across families.[4] Another important extension of the analysis of Stigliz is a deeper analysis of the individual saving choices. Bourguignon (1981) shows that a convex saving function can imply that initial wealth differences determine different steady states. Saving choices have also been micro-founded as decisions to leave bequests for the future generation and in this context the analysis gives insights about the dynamics of intergenerational inequality and mobility. In the context of bequests two main formalization have been used in the literature: one in which bequests enter directly into the utility function of the parents (Atkinson (1980)) and one in which parents care about their children's utility per se (Becker and Tomes (1979)). Those two alternative formulations can give different conclusions about the dynamics of accumulation, inequality and the effects of redistributive taxation. Developing this type of analysis, Becker and Tomes (1986) focus on the intergenerational transmission of abilities across generations and study the implications for the investments in human capital and the resulting dynamics of inequality.[5]

It is fair to mention that until recent times, such research on income distribution and intergenerational mobility did not constitute a topic of major interest in modern macroeconomics. Bertola, Foellmi, and Zweimuller (2006) illustrate this point extensively. They explain that the interest in the

[4]This point is already analyzed by Stiglitz (1969) and has a long tradition in economics. James Meade wrote extensively about it (on this see Atkinson (1980)), Chu (1991) presents a modern analysis of the effect of primogeniture on long-run inequality and mobility.
[5]See the review of Piketty (1998) for more on the literature on bequests and intergenerational mobility.

distribution of income used to be central among classical economists. Classical economists were very much concerned with the issue of how the output of an economy (wages, profits and land rents) is divided among the various classes in society, which for David Ricardo was even "the principal problem of Political Economy". However, it appears that income distribution became a topic of minor interest in recent decades. Atkinson and Bourguignon (2001) (page 7265) note that "in the second half of the century, there were indeed times when interest in the distribution of income was at low ebb, economists appearing to believe that differences in distributive outcomes were of second order importance compared with changes in overall economic performance."[6] According to Bertola, Foellmi, and Zweimuller (2006) this appears to be the case especially with regard to growth theories. They explain that early growth models were still strongly concerned with distributional issues: Harrod (1939) and Domar (1946) study how distribution should adjust to support growth and the same question is asked in the post-Keynesian models of Kaldor (1955) and Pasinetti (1962) which consider endogenously determined factor shares.[7] According to Bertola, Foellmi, and Zweimuller (2006), the subsequent neo-classical theoretical developments removed distribution from the set of macroeconomic issues

[6] A similar point is extensively made by Atkinson (1997) showing the relatively small size of the research concerned with income distribution, specifically for what concerns studies on developed economies.

[7] Bertola, Foellmi, and Zweimuller (2006) (chapter 4) present a brief review of a large literature spanning Physiocratic tableaus, Ricardian theory and post Keynesian growth models. Asimakopoulos (1988) offers a more extensive review of this material.

of interest, as micro-founding optimal choices and expectations relied heavily on "representative agent"modeling strategies and the distribution of income and wealth across consumers was viewed as a passive outcome of aggregate dynamics. Nevertheless, Bertola, Foellmi, and Zweimuller (2006) discuss that factor shares affect individual savings and through this channel they affect growth. For this reason they show that the implications of factor shares on growth can be also studied in the framework of neo-classical growth theory, including the recent endogenous growth theory started by Romer (1986), Lucas (1988) and others.[8] Similarly, even if the analysis of income distribution is not the primary concern of the overlapping generations models of Diamond (1965) and the perpetual youth model of Blanchard (1985), Bertola, Foellmi, and Zweimuller (2006) show how those models can give insights about the evolution of inequality and various dynamics in terms of distribution can arise. A specific analysis of problems of income distribution in a neo-classical framework with HARA utility function is done by Chatterjee (1994).[9]

The interest in problems of inequality and income distribution has a substantial come-back in the 90's. One of the main reasons is that the data showed a dramatic increase in inequality, especially in developed economies. Another reason is that both at the empirical and at the theoretical level, the research starts to point to the links from inequality to growth. In those

[8]See chapter 4 of Bertola, Foellmi, and Zweimuller (2006). A review of the literature on endogenous growth is done by Romer (1989).

[9]About those models see chapter 5 of Bertola, Foellmi, and Zweimuller (2006).

years, at the theoretical level there are two main ideas which bring back the interest in inequality in macroeconomics: 1) models with imperfect financial markets and 2) models of political economy.

1.1.2. From inequality to growth: poverty traps and political economy. The basic idea behind the effect of imperfect financial markets is that if poor individuals are prevented from borrowing and hence cannot invest, then in a dynamic context initial inequalities may persist and some dynasties may remain stuck into a poverty trap. The paper of Loury (1981) represents the first example of the introduction of those ideas in the framework of Becker and Tomes (1979). The seminal contribution of Galor and Zeira (1993) develops this type of analysis and discusses the case that, because of credit constraints, people whose wealth is below a certain threshold cannot invest in human capital: this generates persistent inequality and the existence of poverty traps. In the model of Galor and Zeira (1993) the existence of persistent inequality relies on both the assumptions of credit constraints and threshold effects. Without credit constraints everybody would optimally invest in human capital irrespective of one's initial wealth, and all dynasties would converge to the same wealth level. Conversely, without the assumption of a fixed-size investment, poor dynasties could slowly accumulate by starting with small investment levels and eventually catch up with the rich. It is the combination of non-convex technologies and credit constraints that produce non-convexities in transition equations and the possibility of poverty traps. More specifically, this combination gives rise to a dynamic model that is very similar to the Bourguignon (1981) model with a non-convex saving function. The contributions which followed the analysis of Galor and Zeira (1993) explored more sophisticated dynamic implications of the credit constraints. One important finding is that credit constraints can have important long-run effects and produce poverty traps even in the absence of non-convexities and threshold levels in investments. The most relevant contributions with this respect are those of Banerjee and Newman (1993), Piketty (1997) and Aghion and Bolton (1997). Banerjee and Newman (1993) consider a world where moral-hazard induced credit constraints prevent poor agents from investing in large projects but where rich agents can use a technology to monitor poor agents working as wage earners. There are three possible occupations in the model: wage earners (who are too poor to make any investment on their own), self-employed (who finance and run their own investment) and entrepreneurs (who finance large investments and monitor wage earners). The equilibrium wage rate is determined by the equality between the number of agents who do not have other choice but being wage-earners and the number of wage-earners who are required by the entrepreneurs, and thus depends on the entire wealth distribution. This can generate long-run effects of the initial wealth distribution: an initially large mass of poor agents with no other option than becoming a wage-earner leads to a low wage rate and little upward mobility for wage earners, while an initially small mass of poor agents leads to high wage rates and high mobility between wage-earners and self-employed, which reproduces the forces leading to high wage rates. Depending on the initial distribution of wealth, the economy will there converge to different possible long-run distributions associated to different long-run wage rates. In a slightly different context Piketty (1997) considers a model where agents can invest at any level according to a concave production function, but where moral hazard in entrepreneurial effort leads to a credit-rationing curve. Assuming that the interest rate is endogenously determined by the demand and supply of capital, it is possible to show that depending on the exact initial distribution of wealth there will exist different possible long-run distributions of wealth associated to different long-run interest rates. In this model the endogenous interest rate plays a role which is similar to the endogenous interest wage rate in the model of Banerjee and Newman (1993). The intuition is the following: initial distributions with a large population of low-wealth agents lead to a high demand for capital and to high interest rates, which in turn imply that it takes a long time for low-wealth agents to accumulate and rebuild their collateral, so that the initially large mass of poor agents is self-reproducing. Conversely, low initial interest rates lead to high wealth mobility, high accumulation and low equilibrium interest rates. The steady-states with higher interest rates have at the same time less wealth mobility and a lower aggregate output and capital stock. Finally, the two-way interaction between the distribution of wealth and equilibrium factor prices implied by credit constraints can also generate other interesting and empirically plausible development patterns. For instance, Aghion and Bolton (1997) show that this interaction can generate trajectories characterized by a declining price of capital and an endogenous Kuznets curve. During the initial stage of development, little capital is available, the equilibrium interest rate is high and strong credit constraints imply that only the rich can invest, mobility is low and income inequalities tend to widen. The capital accumulation of the rich progressively forces the interest rate to drop, so that credit constraints become less binding, mobility rises and inequality begins to decline. It is important to stress that in these models inequality has negative implications for growth because the assumed concave production function implies that it is ex-ante optimal to equalize investments across agents. A corollary is that, other things equal, redistribution is ex-ante optimal as it achieves equalization of investments in a world where this would be prevented by credit constraints. Nevertheless there could be a standard negative moral hazard effect of redistribution in the case in which optimal individual effort decreases in the level of redistribution. Aghion and Bolton (1997) show that in the presence of moral hazard, effort decreases in the amount borrowed, because the more an individual needs to borrow in order to get production started and the smaller it is the incentive to supply effort, because a larger fraction of the marginal returns from effort are shared with the lenders. Therefore in such case they show that redistribution has only a beneficial effect as the effort exerted by the lenders remains at the first best and the effort exerted by the borrowers increase because of redistribution. Related group of papers specifically focus on the dynamics of human capital accumulation and inequality through the channels of schools' financing [10] or the endogenous sorting of agents into homogeneous communities or other "clubs"[11].

The other influential contribution that has been introduced in the 90's is represented by the literature in which the prevailing level of redistribution is not exogenous but it is the result of a voting process. This idea has been introduced by the seminal paper of Meltzer and Richard (1981) and the seminal contributions which introduced it in dynamic macroeconomic models are those of Perotti (1993), Alesina and Rodrik (1994) and Persson and Tabellini (1994). The main idea behind those models is that, as in Meltzer and Richard (1981), given the median voter theorem, 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 voted level of redistribution in the economy. High levels of redistribution in turn lower individual incentives to accumulate capital and hence the result that inequality lowers growth. Such models have been quite influential, especially in bringing endogenous political choices into the big picture. They also stimulated a great deal of discussion about the relationship between inequality, growth and redistribution. Some empirical evidence challenged the conclusion on the basis of two different observations. The first observation is that it does not always seem to be the case

[10]See Benabou (1996b), Benabou (2002), Glomm and Ravikumar (1992), Gradstein and Justman (1997), Fernandez and Rogerson (1998) among others.

[11]See Benabou (1993), Benabou (1996a), Durlauf (1996), Durlauf (1997), Kremer and Maskin (1996), Fernandez and Rogerson (1996) among others.

that inequality is detrimental to growth, even though the evidence in favor is quite large.[12] The second and major challenge comes from the observation that it does not seem to be the case that more inequality implies higher redistribution. Perotti (1994), Perotti (1996) and most of the other studies reviewed by Benabou (1996c) find no relationship between inequality and the share of government expenditures in GDP. Among advanced countries the effect actually appears to be negative. This is suggested by many examples: among industrial democracies the more unequal ones tend to redistribute less and not more as the Meltzer and Richard (1981) framework would suggest. The archetypal case is that of the United States versus (Western) Europe: pre-tax inequality is higher in the US than in Europe, nevertheless Europe is characterized by more extensive redistributive policies than the US. 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 US (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.[13] The observation holds within Western Europe itself, where Scandinavian countries are both the most equal and the most redistributive. Using panel data for 20 OECD countries, Rodriguez (1998)

[12]This point has represented a hot debate at the empirical level and I refer the reader to the exhaustive reviews of Benabou (1996c), Bertola, Foellmi, and Zweimuller (2006), Banerjee and Duflo (2003), Bourguignon (2004).

[13]More extensive and detailed evidence about this can be found in Alesina, Glaeser, and Sacerdote (2001) and Alesina and Glaeser (2004).

finds that pre-tax inequality has a significative negative effect on every major category of social transfers as a fraction of GDP, as well as on the capital tax rate.

This second challenge inspired a new group of theoretical models whose major focus is to explain the described evidence relating inequality and redistributive politics. These models achieve this result showing the existence of multiple equilibria: a Europe-type equilibrium characterized by relatively lower inequality and higher redistribution versus a US-type equilibrium characterized by relatively higher inequality and lower redistribution. The model of Benabou (2000) is a seminal model which arrives to such conclusions. In the model of Benabou (2000) the prevailing level of redistribution is still a voting outcome, but the relationship between inequality and redistribution is not monotonic. This feature is the result of two assumptions. The first is again the presence of imperfect credit markets which, together with a concave objective function, implies that redistributive policies can have a positive effect on aggregate output and ex-ante welfare (as already discussed in the model of Galor and Zeira (1993) and following ones). The second and novel element is an extension of the standard voting model which aims to capture the idea that some groups have more influence in the political process than others and – more specifically – that in reality the poor vote with lower probability than the rich or – to some extent – money buys political influence. In particular the result of non monotonic relation between inequality and redistribution relies on the case that if the income distribution is right-skewed (and therefore the median is below the mean) – differently from the standard voting model – the pivotal voter can be richer than the median (but not richer than the mean) and over some range increases in inequality imply that the pivotal voter gets richer. Those features imply that the voted rate of redistribution is a U shaped function of inequality. Starting from the case of no inequality where there is unanimous support for redistribution, growing inequality increases the fraction of agents rich enough to prefer lower levels of redistribution, hence the downward sloping part. After a certain point the function becomes upward sloping as the standard Meltzer and Richard (1981) effect dominates: rising number of poor will eventually impose more redistribution. Conversely, since redistribution relaxes the credit constraints bearing on the poor's , long-run inequality is a declining function of the rate of redistribution. The features that redistribution is a U shaped function of inequality and inequality is a declining function of redistribution creates the potential for multiple steady states (fixed points): mutually reinforcing high inequality and low redistribution (US-type equilibrium) or low inequality and high redistribution (Europe-type equilibrium). Other theoretical models which obtain multiple equilibria with similar features are those of Saint-Paul (2001) and Hassler, Rodriguez-Mora, Storesletten, and Zilibotti (2003). [14]

[14]Both these models introduce more complicated problems of dynamic voting. As in the case of Benabou (2000), also the result of Saint-Paul (2001) relies on the possibility that over some range increases in inequality imply that the pivotal voter gets richer. Hassler, Rodriguez-Mora, Storesletten, and Zilibotti (2003) develop a dynamic model with repeated voting. The model may have multiple voting equilibria in which individuals who

The result about multiple equilibria can also be linked to the literature which focuses on the relation between inequality and technological change.[15] Benabou (2005) develops a theoretical model which introduces such link in the framework of Benabou (2000). He develops an equilibrium model of how skill-biased technical change can affect inequality and also how the composition of labor force and therefore inequality of human capital affects the technological choices of firms. Importing this analysis in the one of Benabou (2000) about inequality and political equilibrium, he obtains a multiple equilibria model about the long run determination of redistribution, technology and the inequality.

1.1.3. Beliefs and Redistributive Politics. The beliefs held by people about the underlying determinants of individual wealth and social mobility appear to be strong determinants of the social contract. In the previous section I have already referred to the evidence pointing to the significant differences in the level of redistribution (social contract) across countries and the evidence that across developed countries more unequal countries seem to redistribute less, where the most striking difference is represented by the US versus Europe. There are also striking differences across countries in the beliefs held by people about the underlying determinants of

expect high redistribution in the future invest little in education today, thereby increasing the number of future beneficiaries (and hence supporters) of redistribution. By the same argument large investments in education and low redistribution can be an equilibrium as well.

[15]This literature is reviewed by Acemoglu (2002) and by Benabou (2005).

individual wealth and the causes of poverty. Once again the most striking difference relates to the differences between the US and (Continental Western) Europe. Since De Tocqueville (1835), many have noticed the *exceptionalism* or *dream* characterizing the American society, in other words the dominant belief that everyone can become rich if wants so and mobility is high in the “land of opportunities”. De Tocqueville (1835) first and many other sociologists and political scientists after him have pointed to persistent differences in popular beliefs about social mobility in explaining the persistent differences between US and European redistributive politics. The differences in those beliefs across societies appear to be substantial. 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. It is also important to notice how the described beliefs appear

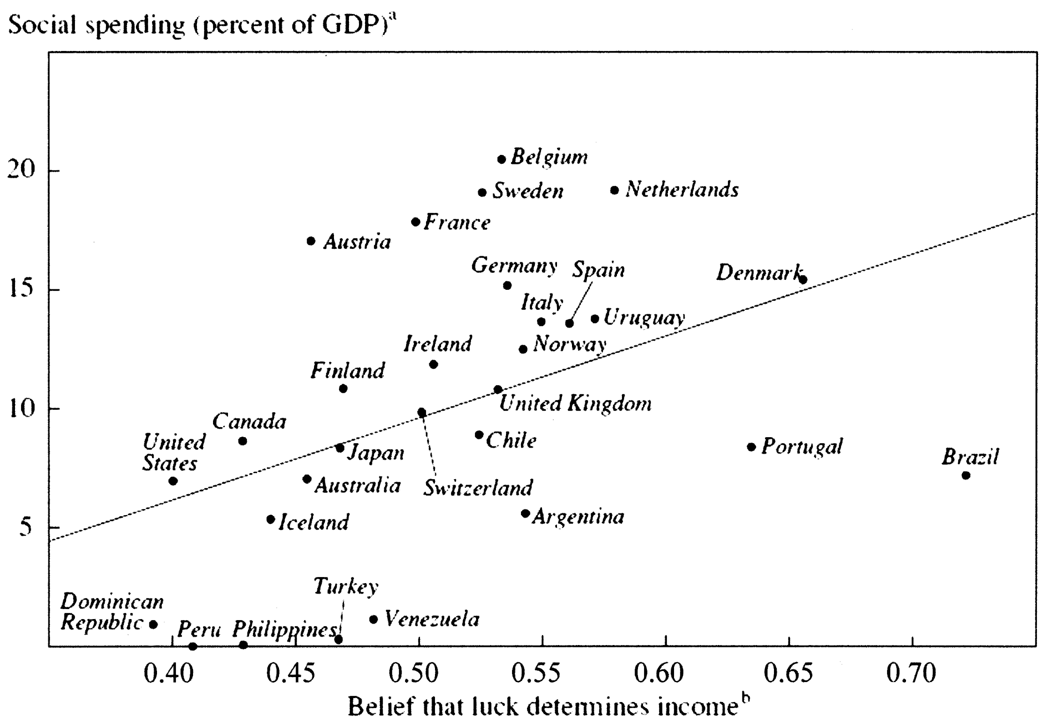

FIGURE 1.1. Beliefs and Policies. Source: Alesina, Glaeser, and Sacerdote (2001).

to be inaccurate and "ideological": there is a significant discrepancy between the view of the American society as an exceptionally mobile one and the actual evidence on intergenerational income or educational mobility.[16] Such massive differences appear to be important, especially since there is a strong correlation between these beliefs and the actual levels of redistribution: see for example figure 1.1 reproduced from Alesina, Glaeser, and Sacerdote (2001). There is also various evidence showing that beliefs are strong determinants of the demand for redistribution and that individual

[16]As pointed out by Piketty (1998), the existing estimates of mobility suffer from methodological controversies. Nevertheless, existing studies show that for some European countries mobility is higher than in the US (see Bjorklund and Jantti (1997) on Scandinavian countries and Couch and Dunn (1997) on Germany), for others it is similar (see Lefranc and Trannoy (2004) on France) and for others it is lower (see Checchi, Ichino, and Rustichini (1999) on Italy).

beliefs determine individual political orientations more than other factors like personal wealth.[17]

The theoretical contributions of Piketty (1995), Alesina and Angeletos (2005) and Benabou and Tirole (2006) have developed insightful models describing how individual beliefs can shape politico-economic outcomes and viceversa and how multiple equilibria (US-type vs Europe-type) are possible. All of those models start from the standard framework of Meltzer and Richard (1981) and extend it. As in the model of Meltzer and Richard (1981), agents vote for the level of redistribution before exerting effort and, because of the moral hazard effect of redistribution, the optimal individual effort decreases in the level of redistribution. For this reason, in the individual ideal level of redistribution the gains from redistribution are traded off the moral hazard effect. As I have already discussed in the previous subsection, the standard Meltzer and Richard (1981) model has a unique equilibrium where the greater is wealth 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. In those models the technology is linear: $y = \theta e + k$, individual wealth y is given by effort e times the return on effort θ plus a fixed endowment k which does not depend on effort and – more importantly – that the individual cannot influence. This is a very convenient formulation which can

[17]See for example Fong (2001), Corneo and Gruner (2002), Alesina and La Ferrara (2005).

capture a number of ideas: the value of the return on effort θ represents the extent to which hard work and other controllable factors contribute to individual wealth, while the value of k represents the extent of the contribution of factors outside individual control such as luck or the family of origin. Incomplete information about the true value of θ versus k creates the premise for different beliefs about the role of the two in the determination of wealth. Moreover a standard moral hazard effect implies that the greater is the expected value of θ versus k 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. The paper of Piketty (1995) is the first one to obtain multiple equilibria in this framework with a dynamic model of imperfect learning. More specifically, 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. Link to some recent contributions in the behavioral economics literature, 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. They discuss two equilibria of the model: in a US-type equilibrium agents believe that effort more than luck determines personal wealth, consequently they vote for low redistribution, incentives are not distorted and the belief is self sustained. Conversely, in the Europe-type equilibrium agents believe that the economic system is not fair and factors as luck, birth, connections, rather than effort, determine personal wealth, hence they vote for high taxes, thus distorting allocations and making the beliefs to be self sustained.[18] 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 θ. 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

[18]Similarly to the paper of Alesina and Angeletos (2005), also in the recent work of Cervellati, Esteban, and Kranich (2006) the individual preferred level of redistribution is not only 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.

and manipulate her own (or her children's in an intergenerational interpretation) 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 workers themselves. Extensive evidence in sociology and psychology seems to suggest this fact.[19] 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 low redistribution the value of a proper motivation (in other words the value of believing that θ is greater than the true value) is higher than with higher redistribution. When redistribution is low (high) everyone thus has greater incentives to believe that effort plays a major (minor) role and consequently more voters finds optimal to hold to such a world-view.[20] 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 θ is high and relatively low redistribution.

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

[20] In the words of Benabou and Tirole (2006) this is the Belief in a Just World which captures the idea of American *exceptionalism* versus a more European type of pessimistic belief.

The other, “European”equilibrium is characterized by a high prevalence of the belief that θ is low and relatively high redistribution.

1.2. The contribution of my thesis

In this thesis 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 θ versus k. 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.[21] 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 and Tirole (2006) that I presented in the previous section. The technology is the same linear one that I have described, agents

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

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 modification of the informative set-up of Benabou and Tirole (2006). 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.[22] 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[23] 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

[22]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.

[23]For example there are not the psychological elements which were present in the models of Alesina and Angeletos (2005) and Benabou and Tirole (2006).

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 is the first important result of my analysis, namely that welfare does not increase monotonically in information.[24] A second important result is to show 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 level of information. This 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 important result relates

[24]This result is not entirely dependent on the assumption of linear utility function as I will discuss in chapter 2

to the comparative statics of aggregate output. I show that also aggregate output is not monotonic in the level of information and can also 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, my thesis presents a methodological contribution as it shows that multiple equilibria can be obtained without psychological biases. The contribution in terms of interpretations is to open a door to the role played by possibly different information "cultures". Some empirical evidence in the literature in education goes supports the predictions of my model. For example, there is quite a large strand of literature showing how the American secondary schooling system is less informative than the European about the position of a student in the national distribution of abilities.[25] Another point in support the predictions of the two equilibria comes from the features of the American dream type of belief. If according to this belief "everybody has a chance to become rich"[26], then it means that in such a society people do not perceive that differences in the individual return of effort can be too large. In our model the American-type of equilibrium has precisely this feature that people pool to similar beliefs about the value of the return on effort. Moreover, showing the existence of multiple optimal informative cultures, my analysis can offer new insights with respect to the "Neo-Marxist"type of explanation which is more common in the literature in political science.

[25]John H. Bishop is a leading scholar in this field, among many works on the topic I refer the interested economists to Bishop (1996) which focus on signalling and Bishop (1997).
[26]As it appears from various already cited evidence about social beliefs.

A more modern and more symmetric version of this view can be found in the work of Alesina and Glaeser (2004). The authors argue that just as American beliefs result from indoctrination predominantly controlled by the wealthier classes, European beliefs result from indoctrination predominantly controlled by Marxist-influenced intellectuals. Alesina and Glaeser (2004) claim that the process of indoctrination has been achieved through the choice of specific institutions and political systems. For example they show how, in the American political history, factors like federalism, majority representation and segregation worked towards low cross-ethnic cohesion and the already described beliefs. My analysis shows how certain beliefs can be imposed in a society not only through the choice of the type of institutions but also through the choice of a certain informative structure. More importantly, my analysis shows that the prevailing informative structure can be actually maintained by the society as an autonomous collective choice and not only as the result of a process of indoctrination. This possibility is shown in my model by the result that, behind the veil of ignorance, societies with similar fundamentals, but with different informative structures, can find optimal to maintain such differences. This idea can be further supported by the different historical experiences of the US versus Europe, as it appears that the two societies started their modern histories with different informative structures. Moreover, the interpretation of the two societies as characterized by different political and economic outcomes and at the same time by similar levels of welfare seems particularly appealing when we thinking about anecdotal evidence. A deeper analysis of the interpretation of my results in relation to the educational and the institutional features of the two societies goes beyond the scope of this paper. Nevertheless it could be a fertile ground for future research.

The remaining of the thesis is organized as follows. Chapter 2 presents the main model, the already mentioned results plus other extensions. Chapter 3 develops an intergenerational dynamic model with the same information set up as the model of chapter 2 and takes a first step towards the development of a unifying framework which can allow to study how beliefs about the determinants of wealth can affect the dynamics of inequality, mobility and redistribution. It follows a short conclusion that summarizes the main results and gives some directions for future research. Most of the proofs follow directly the propositions, but two longer proofs are organized in appendix A. Appendix B contains a paper which tackles a completely different theoretical problem of competitive equilibria existence on which I worked in the first phase of my PhD. It is a self contained contribution and has no relation with the rest of the thesis. I aimed to organize the various chapters as much as possible as autonomous articles to be submitted to peer-reviewed journals[27], nevertheless I provide cross references between the various chapters when useful.

[27]For this reason the reader will notice some repetitions across different chapters.

CHAPTER 2

A simple model of Beliefs and Redistributive Politics under Optimally Incomplete Information

2.1. Introduction

The reason behind the observed wide differences in the social contract across similarly developed countries represents a challenging question which has motivated a large body of research across disciplines. The most evident example of such differences is represented by the persistence of Western European-type welfare states versus US-type more laissez-faire social contracts. Pre-tax inequality is higher in the United States than in Western European countries ("Europe"in short), nevertheless it is Europe to be characterized by more extensive redistributive policies.[1] Without denying the importance of some "fundamental"differences across countries which can impact on such redistributive outcomes, various economists looked for explanations of such societal choices without appealing to exogenous differences in tastes, technologies or political systems. Under this perspective,

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

redistributive outcomes are not considered as exogenous but are endogenously determined taking the political process into account.

The seminal paper of Meltzer and Richard (1981) develops the first model in which the prevailing rate of redistribution in an economy is determined endogenously through the median voter theorem. An influential strand of literature which started in the 90's has developed models of inequality, redistribution and growth building on the framework 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 the following models.[2]

[2] Perotti (1994), Perotti (1996) and most of the other studies reviewed in Benabou (1996c) 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 already cited differences between Western Europe and the United States. The observation holds within Wester Europe itself, where Scandinavian countries are both the most equal and the most redistributive. Using panel data for 20 OECD countries, Rodriguez (1998) finds that pre-tax inequality has a significative negative effect on every major category of social transfers as a fraction of GDP, as well as on the capital tax rate. The seminal contributions which introduced the framework of Meltzer and Richard (1981) in dynamic models of growth are those of Perotti (1993), Alesina and Rodrik (1994) and Persson and Tabellini (1994); like Meltzer and Richard (1981), all those models obtain the result that higher inequality translates into higher redistribution.

Despite the existence of solid alternative theoretical explanations [3], 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".[4] De Tocqueville (1835) himself and many

---

[3]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.

[4]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.

other sociologists and political scientists after him[5] 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.[6]

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. All of those models start from the standard framework of Meltzer and Richard (1981) and extend it. As in the model of Meltzer and Richard (1981), agents vote for the level of redistribution before exerting effort and, because of the typical moral hazard of redistribution, the optimal individual effort decreases in the level of redistribution; for this reason in the individual choice of the ideal level of redistribution the gains from redistribution are traded off the moral hazard effect of redistribution. As I have already discussed, the standard Meltzer

[5]For example Lipset wrote extensively about it.

[6]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.

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. The paper of Piketty (1995) is the first one to obtain multiple equilibria in this framework with a dynamic model of imperfect learning. More specifically, 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. [7] 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

[7]Similarly to the paper of Alesina and Angeletos (2005), also in the recent work of Cervellati, Esteban, and Kranich (2006) the individual preferred level of redistribution is not motivated by purely selfish concerns 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.

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.[8]. 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[9] 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

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

[9] In the words of Benabou and Tirole (2006) this is the Belief in a Just World which captures the idea of American *exceptionalism*

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.

The contribution of this chapter.

In this chapter 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 θ versus k. 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.[10] 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 and Tirole (2006) that I presented in the previous section. The technology is the same linear one that I have described, agents

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

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 modification of the informative set-up of Benabou and Tirole (2006). 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.[11] 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[12] 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

[11]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.

[12]For example there are not the psychological elements which were present in the models of Alesina and Angeletos (2005) and Benabou and Tirole (2006).

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 is the first important result of my analysis, namely that welfare does not increase monotonically in information.[13] A second important result is to show 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 level of information. This 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 important result relates

[13]This result is not entirely dependent on the assumption of linear utility function as I will discuss in chapter 2

to the comparative statics of aggregate output. I show that also aggregate output is not monotonic in the level of information and can also 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, my thesis presents a methodological contribution as it shows that multiple equilibria can be obtained without psychological biases. The contribution in terms of interpretations is to open a door to the role played by possibly different information "cultures". Some empirical evidence in the literature in education goes supports the predictions of my model. For example, there is quite a large strand of literature showing how the American secondary schooling system is less informative than the European about the position of a student in the national distribution of abilities.[14] Another point in support the predictions of the two equilibria comes from the features of the American dream type of belief. If according to this belief "everybody has a chance to become rich"[15], then it means that in such a society people do not perceive that differences in the individual return of effort can be too large. In our model the American-type of equilibrium has precisely this feature that people pool to similar beliefs about the value of the return on effort. Moreover, showing the existence of multiple optimal informative cultures, my analysis can offer new insights with respect to the "Neo-Marxist"type of explanation which is more common in the literature in political science.

[14]John H. Bishop is a leading scholar in this field, among many works in the topic I refer the interested economists to Bishop (1996), Bishop (1997) and relative references.

[15]As it appears from various already cited evidence about social beliefs.

A more modern and more symmetric version of this view can be found in the work of Alesina and Glaeser (2004). The authors argue that just as American beliefs result from indoctrination predominantly controlled by the wealthier classes, European beliefs result from indoctrination predominantly controlled by Marxist-influenced intellectuals. Alesina and Glaeser (2004) claim that the process of indoctrination has been achieved through the choice of specific institutions and political systems. For example they show how, in the American political history, factors like federalism, majority representation and segregation worked towards low cross-ethnic cohesion and the already described beliefs. My analysis shows how certain beliefs can be imposed in a society not only through the choice of the type of institutions but also through the choice of a certain informative structure. More importantly, my analysis shows that the prevailing informative structure can be actually maintained by the society as an autonomous collective choice and not only as the result of a process of indoctrination. This possibility is shown in my model by the result that, behind the veil of ignorance, societies with similar fundamentals, but with different informative structures, can find optimal to maintain such differences. This idea can be further supported by the different historical experiences of the US versus Europe, as it appears that the two societies started their modern histories with different informative structures. Moreover, the interpretation of the two societies as characterized by different political and economic outcomes and at the same time by similar levels of welfare seems particularly appealing when we thinking about anecdotal evidence. A deeper analysis of the interpretation of my results in relation to the educational and the institutional features of the two societies goes beyond the scope of this paper. Nevertheless it could be a fertile ground for future research.

The present chapter is structured as follows. Section 2.2 introduces the set-up of the model. Section 2.3 analyzes the voting problem and the relative outcome. Section 2.4 analyzes the comparative statics considering the precision of the signal as an exogenous policy variable. In section 2.5 I analyze the optimal ex-ante precision for the economy. In section 2.6 I introduce the concept of politico-economic equilibrium and investigate the possibility of existence of multiple equilibria. Section 2.7 analyzes the robustness and the generalization of the results. Sections 2.8 and 2.9 respectively introduce heterogenous endowments and risk aversion in the main model as a robustness check. Section 2.10 concludes.

2.2. 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:

$$y^i = k^i + \theta^i e^i, \tag{2.1}$$

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 [16], but I will later consider the possibility of heterogeneous endowments. I assume that θ^i is i.i.d. across agents 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 $\theta_L < \theta_H$. 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:

$$T\left(\begin{bmatrix}\sigma_L \\ \sigma_H\end{bmatrix} \Big| [\theta_L, \theta_H]\right) = \begin{pmatrix}\lambda & 1-\lambda \\ 1-\lambda & \lambda\end{pmatrix}. \tag{2.2}$$

The structure of the economy – including the value of π and matrix (2.2) – 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

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

the Bayes Rule. I introduce the following notation:

$$\mu^i \equiv \Pr[\theta^i = \theta_L | \sigma^i], \tag{2.3}$$

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

$$\mu_{\sigma_L} \equiv (\mu^i | \sigma^i = \sigma_L) = \frac{\pi\lambda}{\pi\lambda + (1-\pi)(1-\lambda)} \tag{2.4}$$

and

$$\mu_{\sigma_H} \equiv (\mu^i | \sigma^i = \sigma_H) = \frac{\pi(1-\lambda)}{\pi(1-\lambda) + \lambda(1-\pi)}. \tag{2.5}$$

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

$$\theta(\mu^i) \equiv \mu^i \theta_L + (1-\mu^i)\theta_H. \tag{2.6}$$

Given the symmetric structure of (2.2) 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. $\mu_{\sigma_L} = 1$, $\mu_{\sigma_H} = 0$.[17] As already explained

[17] 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 $\lambda \in [1/2, 1]$.

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:

$$p_{\sigma_L} \equiv \Pr[\sigma^i = \sigma_L] = \lambda\pi + (1-\lambda)(1-\pi), \tag{2.7}$$

symmetrically

$$p_{\sigma_H} \equiv \Pr[\sigma^i = \sigma_H] = \lambda(1-\pi) + \pi(1-\lambda) = 1 - p_{\sigma_L}. \tag{2.8}$$

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:

$$c^i = (1-\tau)y^i + \tau\bar{y}, \tag{2.9}$$

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 and implies that the government budget constraint is always binding. Throughout the analysis I consider the following individual utility function[18]:

$$u^i(c^i, e^i) = c^i - \frac{a}{2}(e^i)^2. \tag{2.10}$$

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[19].

2.3. Voters' Problem

Plugging expressions (2.1) and (2.9) into (2.10) I obtain the expression of the expected utility of agent i at t:

$$u_t^i = E[(1-\tau)(k + e^i\theta^i) + \tau(k + \overline{e\theta}) - a(e^i)^2/2 | I_t^i], \tag{2.11}$$

[18]In section 2.9 I will consider the possibility of concavity in consumption and discuss the relative implications.

[19]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.

where $E[\cdot|I_t^i]$ is individual i's expectation conditional on the information at time t. As explained in the previous section, the information structure is such that $I_0^i = T$ and $I_1^i = (T, \sigma^i)$. 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 (2.11) conditional on signal σ^i. Solving backwards, each individual i maximizes (2.11) choosing e^i after that the winning tax rate τ is announced. Being (2.11) strictly concave in e^i, by solving the sufficient first order condition I find the optimal individual level of effort:

$$e^i = (1 - \tau)\theta(\mu^i)/a. \tag{2.12}$$

By backward induction, I can plug (2.12) into (2.11) and find the objective function that i maximizes when voting for the tax rate. In order to do this, it is useful to specify the individual i expectation of the output from effort:

$$E[e^i\theta^i|I_1^i] = (1 - \tau)\left(\theta(\mu^i)\right)^2/a \tag{2.13}$$

and of squared effort

$$E[(e^i)^2|I_1^i] = (e^i)^2 = \left(\frac{1 - \tau}{a}\right)^2 \theta(\mu^i)^2. \tag{2.14}$$

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

$$E[\overline{e\theta}|I_1^i] = (1-\tau)\Gamma/a, \tag{2.15}$$

where I define

$$\Gamma \equiv \pi\theta_L\left(\lambda\theta(\mu_{\sigma_L}) + (1-\lambda)\theta(\mu_{\sigma_H})\right) + (1-\pi)\theta_H\left((1-\lambda)\theta(\mu_{\sigma_L}) + \lambda\theta(\mu_{\sigma_H})\right). \tag{2.16}$$

Collecting $\theta(\mu_{\sigma_L})$ and $\theta(\mu_{\sigma_H})$ it is easy to re-write expression (2.16) as

$$\Gamma = p_{\sigma_L}\theta(\mu_{\sigma_L})^2 + (1-p_{\sigma_L})\theta(\mu_{\sigma_H})^2. \tag{2.17}$$

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 (2.13), (2.14) and (2.16) into (2.11), I obtain an indirect form of (2.11) as a function of τ:

$$u_1^i = k^i + (1-\tau)^2\theta(\mu^i)^2/a + \tau(1-\tau)\Gamma/a - (1-\tau)^2\theta(\mu^i)^2/2a. \tag{2.18}$$

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 (2.18) is strictly negative, the ideal tax rate of agent i follows from the first order condition:

$$\tau(\mu^i) = 1 - \frac{1}{2 - \frac{\theta(\mu^i)^2}{\Gamma}}. \tag{2.19}$$

The denominator of (2.19) shows how the subjective prospects of upward mobility reduce the desired tax rate.[20] I introduce an assumption which bounds above the heterogeneity in individual abilities in order to assure the concavity of the objective function (2.18) and thus to use the median voter theorem.

Assumption 2.1: $2\theta_L^2 > \theta_H^2$.

A proposition follows:

PROPOSITION 2.1. *Individual preferences for the rate of redistribution are single peaked and the individual ideal rate of redistribution is given by expression (2.19).*

PROOF. The second derivative of the objective function in problem (2.18) is given by the following expression:

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

[20]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 (2.13) over aggregate output (2.15), noticing that the term $\frac{1-\tau}{a}$ gets canceled out.

The condition stated by Assumption 2.1 is sufficient for (2.20) 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 3.2 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.[21]

2.4. 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 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.

[21]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 (2.19) 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 $\theta(\mu)^2 = \Gamma = (\bar{\theta})^2$.

LEMMA 2.1. *The expected value of individual ability (2.6), 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 (2.16), namely aggregate output from effort net of the distortive effect of redistribution on effort.

LEMMA 2.2. *The expression of Γ (2.16) is (i) increasing and (ii) convex in the level of information λ.*

The proof is in Appendix A.1. The intuition behind this result is very important. Lemma 2.2 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).[22] 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 2.2. *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 (2.19) with $\mu^i = \mu_{\sigma_L}$, the proof follows in a straightforward way from lemmas 2.1 and 2.2. □

From expression (2.19) 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

[22]Not only the assumptions for the result of of Radner and Stiglitz (1984) apply to my set-up, one example of the Radner and Stiglitz (1984) result by Kihlstrom (1984) is based on the same informative structure of this model.

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 (2.12) I define the optimal effort implemented by those who observe σ_L:

$$e|\sigma_L \equiv (1-\tau)\theta(\mu_{\sigma_L})/a, \tag{2.20}$$

and by those who observe σ_H:

$$e|\sigma_H \equiv (1-\tau)\theta(\mu_{\sigma_H})/a. \tag{2.21}$$

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

$$\bar{e} = (1-\tau)(p_{\sigma_L}\theta(\mu_{\sigma_L}) + (1-p_{\sigma_L})\theta(\mu_{\sigma_H}))/a, \tag{2.22}$$

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 2.3. *If $\pi > 1/2$, aggregate effort (2.22) is decreasing in the level of information λ.*

The proof follows trivially from proposition 2.2 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 (2.20), 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 2.1 and 2.2. Instead, the comparative static for the expression of optimal effort which is exerted by those who observe signal σ_H (2.21) 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 2.6 it will turn out to be non monotonic.

I discuss the comparative statics of output. Plugging (2.15) into (2.1) I obtain the expression of aggregate output:

$$\bar{y} = k + (1-\tau)\Gamma/a. \tag{2.23}$$

Notice that for $\pi > 1/2$ the effect of λ is not a-priori clear as given lemma 2.2 and proposition 2.2, λ has opposite effects on $(1-\tau)$ and Γ. Nevertheless I find an interesting property:

PROPOSITION 2.4. *If $\pi > 1/2$, the expression for aggregate output (2.23) is (i)either monotonically decreasing or monotonically decreasing up to a point and then monotonically increasing in the level of information λ* [23]*, (ii) maximized for $\lambda = 1/2$.*

The proof is in Appendix A.2. 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_{σ_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 (2.19) both $\theta(\mu)$ and Γ increase for $\lambda \in [1/2, 1]$ and so the

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

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

PROPOSITION 2.5. *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 (2.19) as

$$\tau = \frac{\Gamma - \theta(\mu_{\sigma_H})^2}{2\Gamma - \theta(\mu_{\sigma_H})^2}. \tag{2.24}$$

Notice that the numerator of (2.24) 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 2.1. 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 2.1 and 2.2, 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 (2.23), lemma 2.2 and proposition 2.5 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 (2.22) 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.

2.5. 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:

$$E[e^i\theta^i|I_0^i] = E[\overline{e\theta}|I_1^i] = (1-\tau)\Gamma/a, \tag{2.25}$$

and the expression for expected individual squared effort:

$$E[(e^i)^2|I_0^i] = \frac{(1-\tau)^2\Gamma}{a^2}. \tag{2.26}$$

Plugging (2.25) and (2.26) into (2.11) and rearranging I obtain the expression for expected utility at $t = 0$:

$$u_0^i = k + (1-\tau^2)\Gamma/2a. \tag{2.27}$$

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 (2.27). The solution of the problem is not a-priori trivial.

Case of $\pi > 1/2$.

In the case of $\pi > 1/2$, lemma 2.2 and proposition 2.2 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 2.6. *If $\pi > 1/2$, the expression of ex-ante utility (2.27) is either monotonically decreasing or monotonically decreasing up to a point and then monotonically increasing*[24].

PROOF. Expression (2.27) 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 2.4. I rename $(1+\tau) \equiv a(\lambda)$ 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 (A.4) in appendix A.2 it can be checked that this expression is strictly negative for $\lambda = 1/2$. Notice that given proposition 2.2 $\frac{da}{d\lambda}b > 0$ and that given proposition 2.4 $a\frac{db}{d\lambda}$ can change sign and become positive at most once. Appendix A.2 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 becomes positive 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 (2.27). 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,

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

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 (2.27) does not increase monotonically in the level of information λ. Secondly, the convexity of the value of information Γ (lemma 2.2) implies that – as stated in proposition 2.6 – that ex-ante utility (2.27) 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 (2.27) not clear. In the case in which the condition for the monotonic behavior given in proposition (2.5) applies, then both $(1 - \tau^2)$ and Γ increase in λ for $\lambda \in [1/2, 1]$. Hence (2.27) 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 (2.25) and (2.26) into (2.11) I obtain the individual problem at $t = 0$:

$$\lambda^i = \arg\max\{(1-\tau)(\bar{k} + (1-\tau(\lambda))\Gamma(\lambda^i)/a) + \tau(\lambda)(\bar{k} + (1-\tau(\lambda))\Gamma(\lambda)/a) - (1-\tau(\lambda))^2\Gamma(\lambda^i)/2a\}. \tag{2.28}$$

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 2.2). 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.

2.6. 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 2.1. 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 (2.4) and (2.5),
(iii) the prevailing rate of redistribution τ is the ideal rate (2.19) 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 (2.27). Plugging $\lambda = 1/2$ and $\lambda = 1$ in (2.27) 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

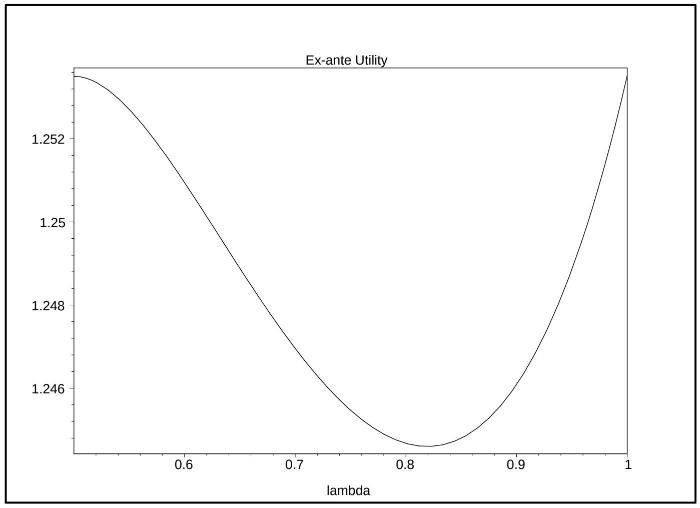

FIGURE 2.1. Ex-ante utility for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

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 (2.4), (2.5), (2.7) (2.19), (2.16) and consequently those expressions in (2.27). I obtain a map of u_0^i in λ, which I plot in figure 2.1. 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

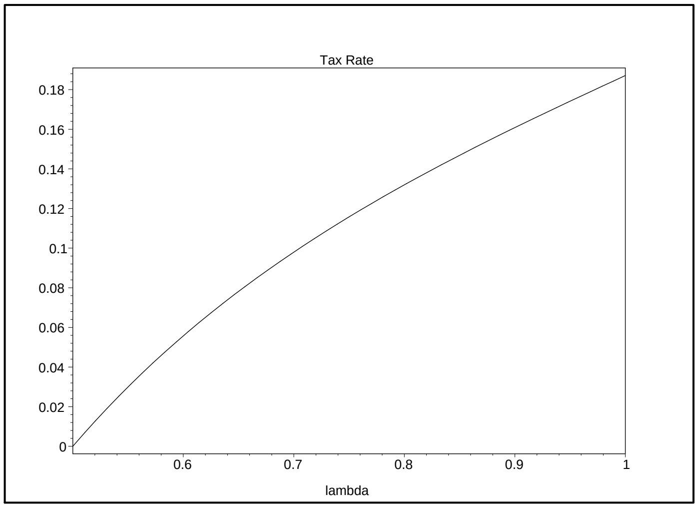

FIGURE 2.2. τ for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

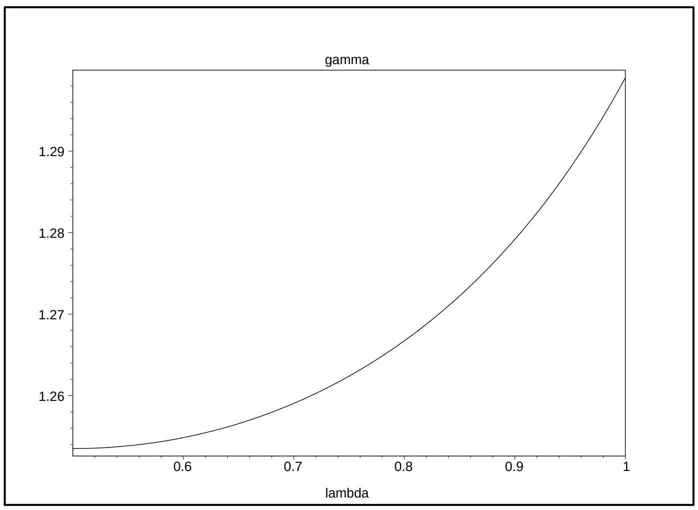

FIGURE 2.3. Γ for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

ex-ante optimal for the society. A society with a higher (lower) value of pi 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.

Interpretation of the multiple Politico-Economic Equilibria.

I plot τ and Γ in figures 2.2 and 2.3 respectively. As explained in the previous section, the uninformative equilibrium is characterized by a lower τ and an higher Γ than the informative equilibrium, the two variables have opposite effects on the ex-ante utility, hence the multiple equilibria. I can further interpret these two equilibria plotting the expressions of the effort exerted by those who observe σ_L (2.20) and σ_H (2.21) as functions of λ, in figures 2.4 and (2.5) respectively. I also plot the expression of aggregate effort (2.22) in figure 2.6. The expression of optimal effort (2.12) shows that the greater is τ and the lower is the optimal effort, hence the informative equilibrium is characterized by a severe moral hazard problem as τ is at the maximum level. It is less immediate to notice an opposite effect of adverse selection. Figures 2.4 and 2.5 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. Figures 2.6 and 2.7 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 agents have wrong beliefs about the real return on effort. Both groups of agents hold the same belief and exert the same levels of effort (pooling equilibrium); in particular the majority of agents (which are those with low ability $_L$) are biased towards optimism as they believe to be more productive than what they truly are. In this equilibrium, the tax rate is at the minimum level, whereas aggregate effort and output are at the

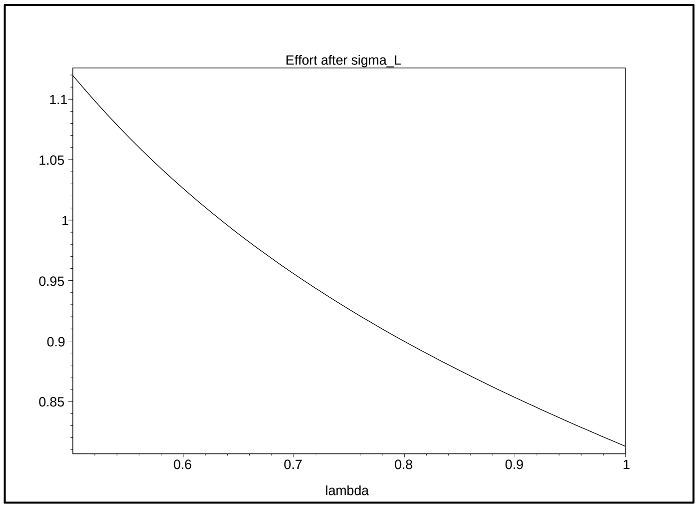

FIGURE 2.4. Effort given σ_L for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

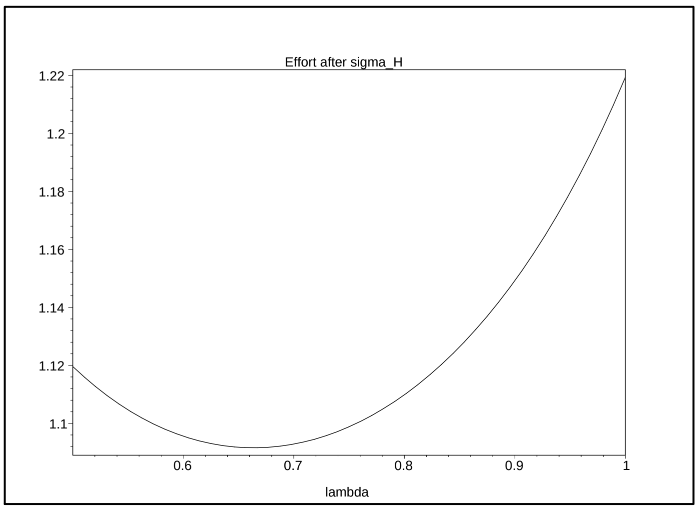

FIGURE 2.5. Effort given σ_H for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

maximum level. The informative equilibrium can instead be interpreted as a Europe-type equilibrium. In this equilibrium the two groups of agents have correct beliefs about the real return on effort. As the low productive agents are the majority, their preferred tax rate is the prevailing in the economy, hence the level of redistribution is higher than in the US-type equilibrium. High redistribution and correct beliefs about the return on effort

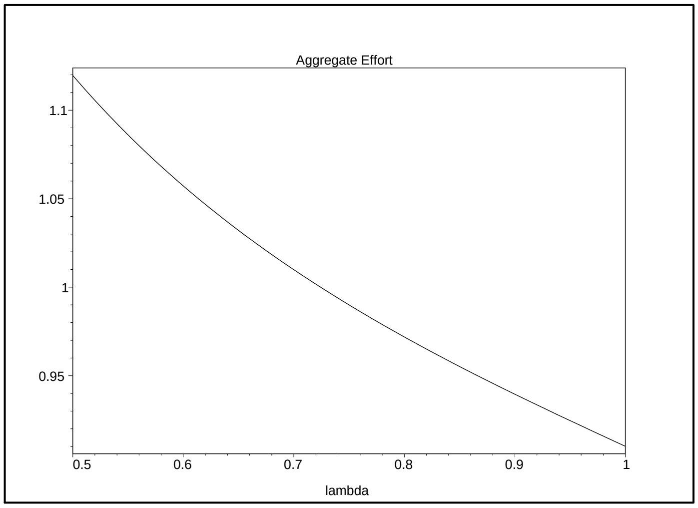

FIGURE 2.6. Aggregate effort for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

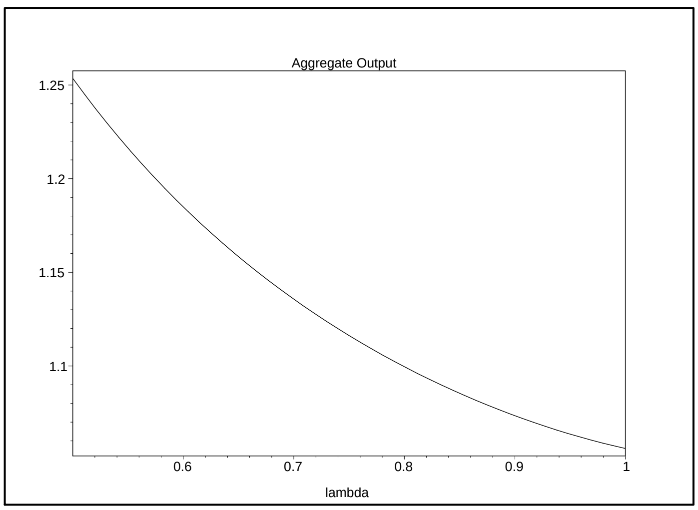

FIGURE 2.7. Aggregate Output for $\pi = 0.761$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 1$, $k = 0$.

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.

2.7. 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 2.8, 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 (2.2), then the problem is the same – this can be seen from the fact that the expressions (2.6) and (2.16) do not change – hence the same results apply. Or again the same results would apply in the case of homogeneous returns and aggregate macroeconomic shocks: $\theta^i = \theta$ for all i and again all that agents know is that and with probability π $(1-\pi)$ the average value of θ is θ_L (θ_H). The fact that the true distribution of the θ's remain unknown shows that a more precise signal in the Europe-type equilibrium does not mean that Europeans get to know the truth whereas Americans do not.

The result should instead be interpreted as showing the possibility and the implications of more versus less separating information structures or cultures. 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 (2.27), 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 (2.2) implies lemma 2.2 and therefore that the more precise is the signal λ and the greater is Γ. This consideration helps to understand the following general result:

THEOREM 1. *Given the ex-ante objective function (2.27), 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 2.2, 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 (2.27) 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. Theorem 1 shows that in the case of multiple equilibria, a US-(Europe-) type equilibrium is relatively characterized by: (i) a less (more) informative signal and therefore (ii) less (more) separated beliefs, (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, because these feature do not change the ex-ante problem, or again on the linearity of the utility function. In the next two

[25]As section 2.8 will show, with heterogenous endowments technical difficulties arise as changing the level of information changes the identity of the median voter and different voters prefer different tax rates given different endowments, so the comparative statics are generally discontinuous; nevertheless the ex-ante optimal λ still has to maximize the object (2.27) and therefore theorem 1 still applies.

sections I show that the main results still hold in both a setup with heterogenous endowments and concave utility in wealth.

2.8. 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 (2.18) – the indirect utility in τ – is given by the following expression:

$$u_t^i = \tau(k^i - \bar{k}) + (1-\tau)^2\theta(\mu^i)^2/a + \tau(1-\tau)\Gamma/a - (1-\tau)^2\theta(\mu^i)^2/2a. \tag{2.29}$$

Assumption 2.1 still assures that expression (2.29) 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:

$$\tau(k^i, \mu^i) = 1 - \frac{1 + \frac{a(k^i - \bar{k})}{\Gamma}}{2 - \frac{\theta(\mu^i)^2}{\Gamma}}. \tag{2.30}$$

As explained by Benabou and Tirole (2006), the numerator of (2.30) 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: $\tau(k_H, \mu_{\sigma_H}) \leq \min\{\tau(k_H, \mu_{\sigma_L}), \tau(k_L, \mu_{\sigma_H})\} \leq \max\{\tau(k_H, \mu_{\sigma_L}), \tau(k_L, \mu_{\sigma_H})\}$ $\leq \tau(k_L, \mu_{\sigma_L})$.[26]

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.

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

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 (2.27), 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 as a function of λ

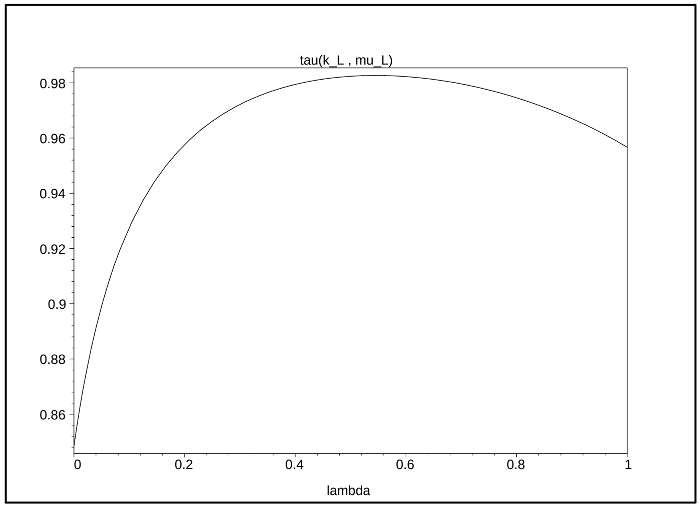

FIGURE 2.8. $\tau(k_L, \mu_L)$ for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

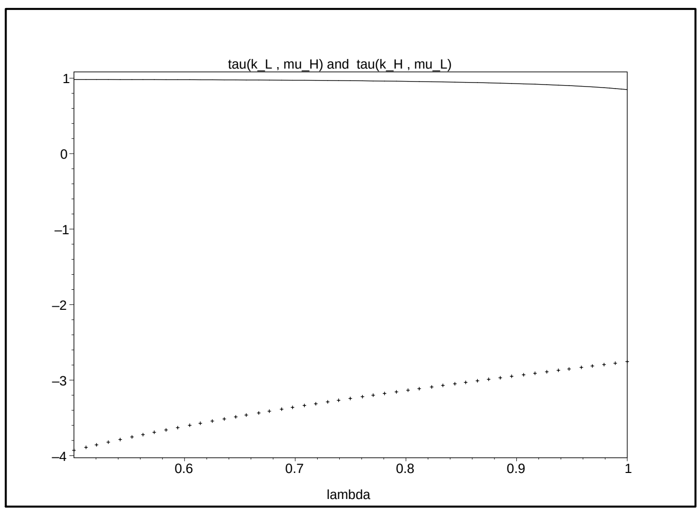

FIGURE 2.9. $\tau(k_L, \mu_H)$ (continuous line) and $\tau(k_H, \mu_L)$ (pointed line) for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

in figure 2.8. 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 two as functions of λ in figure 2.9. 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 2.10. It can

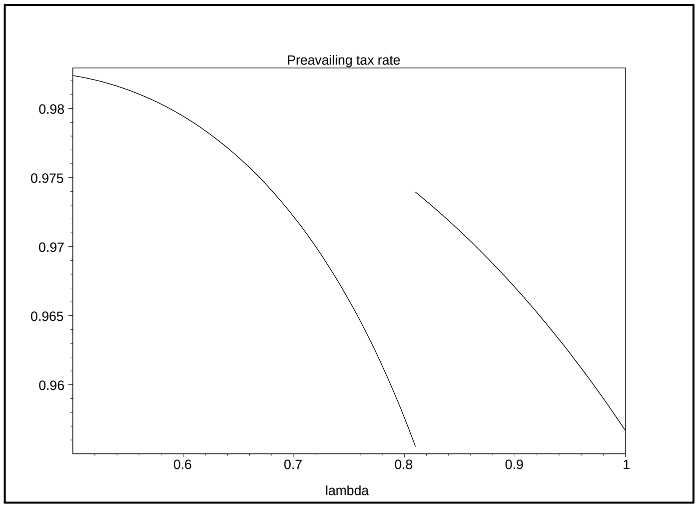

FIGURE 2.10. Prevailing tax rate for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

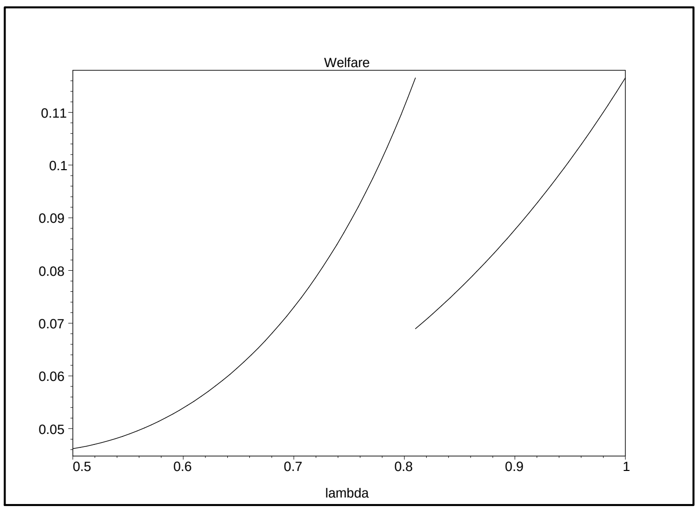

FIGURE 2.11. Welfare for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

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 2.11, 2.12, 2.13, 2.14, 2.15. I also plot figures 2.13 and 2.14 together in figure 2.16, where the thicker line represents figure 2.13.

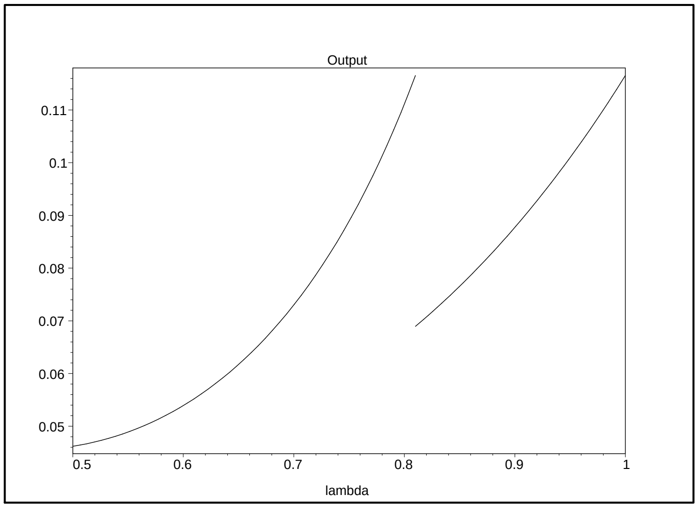

FIGURE 2.12. Aggregate Output for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

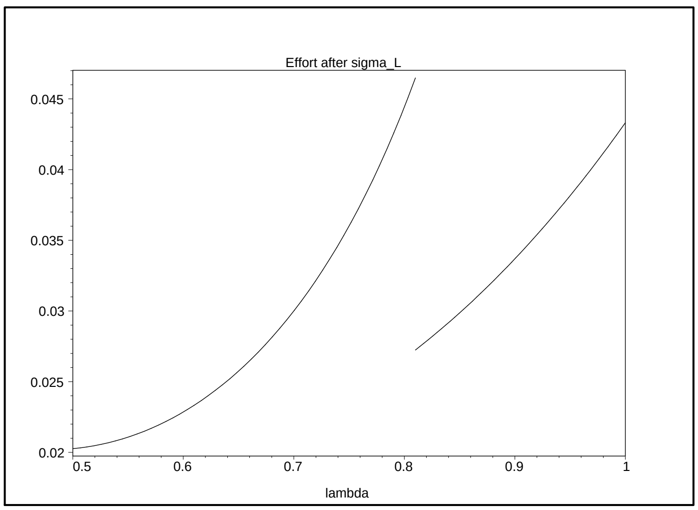

FIGURE 2.13. Effort after the observation of σ_L for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

2.9. Analysis with concave utility

As already explained in section 2.5, 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

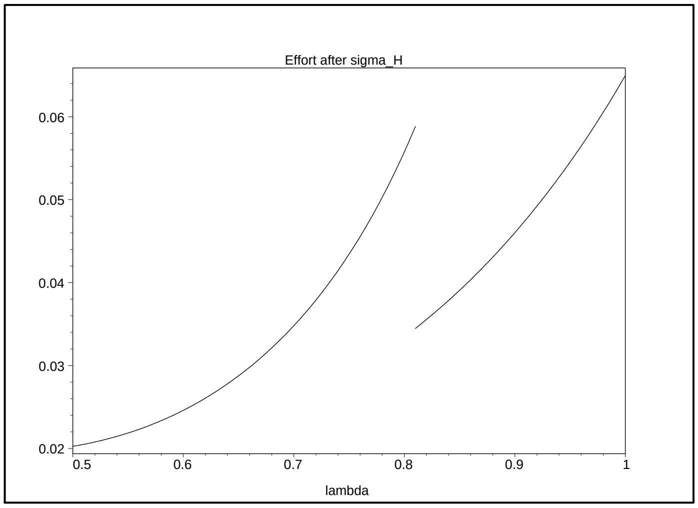

FIGURE 2.14. Effort after the observation of σ_H for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

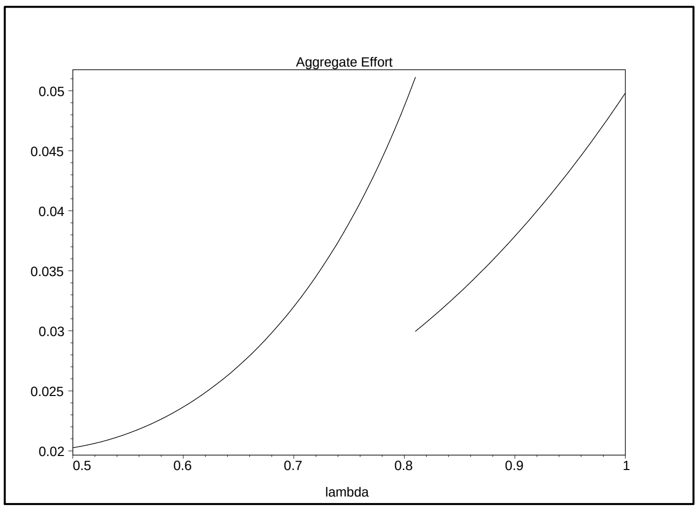

FIGURE 2.15. Aggregate effort for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

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

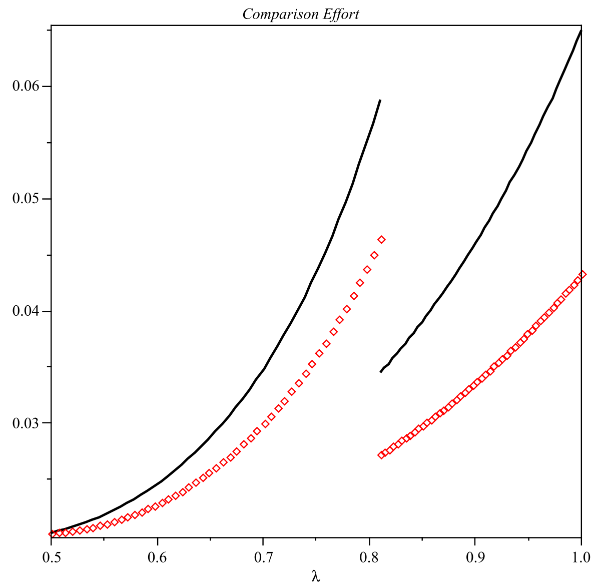

FIGURE 2.16. Comparison of effort levels for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $a = 8$, $k_L = 1$, $k_H = 1.812$.

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. [27]

2.9.1. 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:

$$u_0^i = E_0^i[\frac{1}{\gamma}(c^i)^\gamma - (e^i)^2/2], \tag{2.31}$$

[27]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).

with $\gamma < 1$. The period 1 problem is still the one described in section 2.2, hence the expression for optimal individual effort is still (2.12) and the expression for the aggregate product of effort is still given by (2.25). I plug those expressions together with (2.9) and (2.1) into (2.31) and I obtain the expression for expected utility conditional on the observation of σ_L:

(2.32)
$$u_{0\sigma_L} = \mu_L \frac{1}{\gamma}((1-\tau)\theta_L e_{\sigma_L} + \tau(1-\tau)\Gamma)^\gamma + (1-\mu_L)\frac{1}{\gamma}((1-\tau)\theta_H e_{\sigma_L} + \tau(1-\tau)\Gamma)^\gamma - (e_{\sigma_L})^2/2,$$

and a symmetric expression given the observation of σ_H, where e_{σ^i} is given by expression (2.12) and it represents optimal individual effort conditional on the observation of σ^i. Using those expressions, I can rewrite expression (2.31) as

$$\begin{aligned}(2.33)\quad u_0^i = p_{\sigma_L} u_{\sigma_L} + (1-p_{\sigma_L}) u_{\sigma_H} = \\ \pi\lambda\frac{1}{\gamma}((1-\tau)\theta_L e_{\sigma_L} + \tau(1-\tau)\Gamma)^\gamma + \\ (1-\pi)(1-\lambda)\frac{1}{\gamma}((1-\tau)\theta_H e_{\sigma_L} + \tau(1-\tau)\Gamma)^\gamma + \\ \pi(1-\lambda)\frac{1}{\gamma}((1-\tau)\theta_L e_{\sigma_H} + \tau(1-\tau)\Gamma)^\gamma + \\ (1-\pi)\lambda\frac{1}{\gamma}((1-\tau)\theta_H e_{\sigma_H} + \tau(1-\tau)\Gamma)^\gamma - (1-\tau)^2\Gamma/2.\end{aligned}$$

It is clear from expression (2.33) 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 effect because Γ increases in λ (lemma 2.2) and on the other hand increasing λ separates the levels of effort implemented (lemma 2.1) 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 2.2 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 (2.33) as a function of λ in figure 2.17. The figure shows that there is an interior solution in terms of λ. This example proves the possibility of non monotonicity of information.

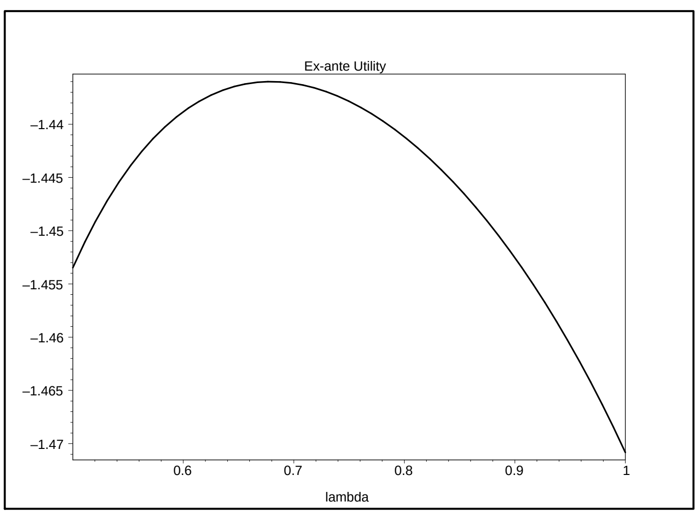

FIGURE 2.17. Welfare for $\pi = 0.8$, $\theta_L = 1$, $\theta_H = 1.5$, $\gamma = -1$.

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 (2.33) as a function of λ in figure 2.18. 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

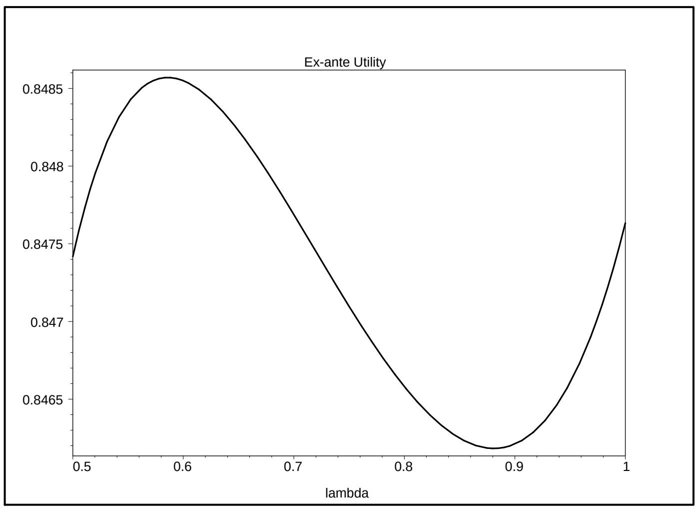

FIGURE 2.18. Welfare for $\pi = 0.8$, $\theta_L = 1$, $\theta_H = 1.5$, $\gamma = 0.8$.

also possible to have a case with multiple optimal values of information. I consider the case of $\gamma = 0.845$, I plot expression (2.33) as a function of λ in figure 2.19. 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.

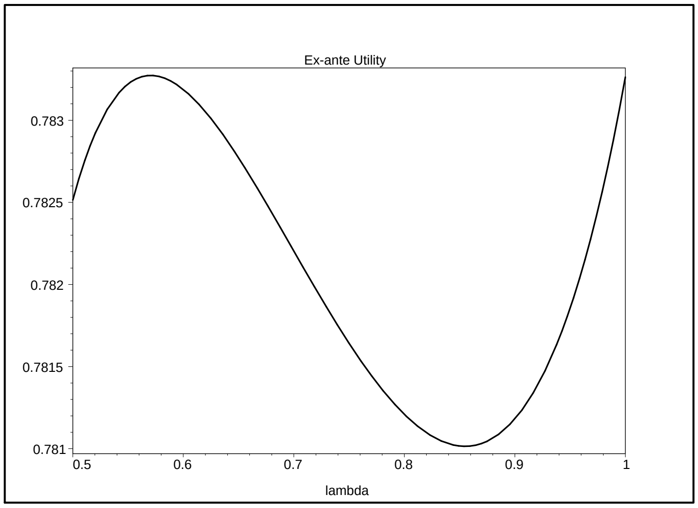

FIGURE 2.19. Welfare for $\pi = 0.8$, $\theta_L = 1$, $\theta_H = 1.5$, $\gamma = 0.845$.

2.10. Conclusion

This chapter developed a simple but rich 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 welfare or output maximizing level of information.

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 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, as shown by figure 2.16. In conclusion, the presented model focuses 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 initial endowments (both directly and through their role in affecting voted redistribution) and in the presented model there is no specification of the wealth generating process (which should naturally be seen as dynamic). Therefore the contribution of the next chapter will be precisely at this level.

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

CHAPTER 3

Redistributive Politics under Optimally Incomplete Information in an intergenerational model with bequests

3.1. Introduction

In the previous chapter I presented a model which studies how incomplete information about the determinants of individual wealth (i.e. incomplete information about the value of return to effort versus the role of luck or other predetermined factors) can affect, through individual beliefs, the individual preferences over redistribution and the individual optimal decisions of how much effort to exert. Consequently the model analyzes also the impact of incomplete information on the prevailing level of redistribution under majority voting and on aggregate outcomes as aggregate effort, aggregate output and welfare. The model offers policy results in terms of comparative statics and insights about the various observed differences between laissez-faire versus welfare state type of economies. The previous model does not have a full dynamic setting, because it describes a one period economy where individuals vote, exert effort and receive income. Thus, the fact that the model does not analyze savings, wealth accumulation or intergenerational transfers implies that the model cannot say anything about the dynamics of inequality and wealth mobility over time. It is a natural further step to study how incomplete information can impact beliefs, political and economic outcomes in a dynamic set up in order to gain insights about the dynamics of inequality and wealth mobility. The first reason is that if individual beliefs about the underlying determinants of wealth are important determinants of individual voting and effort choices, then it is also the case that the same beliefs are also important determinants of the dynamics of inequality and mobility. The second reason is that the individual beliefs about the determinants of wealth are intrinsically very much related to the individual beliefs about the determinants and the extent of mobility, which is a dynamic process, therefore it is important to study such beliefs in a dynamic setting. For this reason, in order to gain insights about the role of beliefs in affecting the dynamics of inequality and mobility, as well as in shaping the preferences for redistribution, I introduce the information set up of the previous chapter in an intergenerational model with bequests and stochastically evolving skills. There are three main elements in the model. The first element is the dynamic set-up with with bequests and stochastic skills which allows to analyze intergenerational inequality and mobility. The second element is represented by the political economy side: as in the model of the previous chapter, I introduce a linear redistribution scheme, where the prevailing rate of redistribution is set by the median voter. The third element is the information structure which builds on the model of the previous chapter. In every period the level of information about the true value of individual skills can vary in a continuous way from a completely uninformative structure in which every individual only has the same prior information about the distribution of skills to the case that each individual is perfectly informed about her own skills. As it was the case with the model of the previous chapter the level of information should be interpreted as an institutional feature of the economy which governments or other institutions can possibly affect through various policies: educational policies, the release of information or propaganda. The present chapter introduces a dynamic set-up with bequests in the previous model. As in the previous chapter, one exercise consists in analyzing how different levels of incomplete information affect the endogenous outcomes (individual voting and effort, redistribution, aggregate effort, output and welfare) and another exercise consists in considering also the level of information as endogenous and analyzing which levels of information are optimal for the society and can arise in an equilibrium.

The model of this chapter constitutes a first attempt to link three different strands of theoretical literature: models of intergenerational inequality, models on the political economy of redistribution, models which analyze the role of the beliefs about the underlying determinants of wealth. The related literature has already been extensively reviewed in chapter 1.

(i) Neoclassical models of intergenerational inequality. It is the seminal paper of Stiglitz (1969) to be commonly considered the first modern analysis of the distribution of wealth and income among individuals. The model of Stiglitz presents a strong result of long run convergence in the dynamics of individual income which parallels the seminal result obtained by Solow (1956) in the context of country income. In the model of Stiglitz agents are endowed with capital (accumulated factor) and labor (non accumulated factor), markets are competitive and both factors are paid at their marginal return. The assumptions of diminishing returns to capital and of an identical concave saving function across individuals imply that individual wealth increases over time in a concave fashion and eventually converges to a steady state value which does not depend on the initial level of wealth. In other words, in the model inequality across families is solely determined by the differences in the non accumulated factors (i.e. the differences in individual skills) and when all families are equally endowed with the non accumulated factor (i.e. skills are homogenous across families) every family converges to the same level of wealth. Building on this seminal model, other authors have extended the basic set-up in order to study intergenerational inequality and mobility. In this context, saving choices have been micro-founded as decisions to leave bequests for the future generation. For what concerns bequests, two main formalization have been used in the literature: one in which bequests enter directly into the utility function of the parents (Atkinson (1980)) and one in which parents care about their children's utility per se (Becker and Tomes (1979)). Those two alternative formulations can give different conclusions about the dynamics of accumulation, inequality and the effects of redistributive taxation. Developing this type of analysis, Becker and Tomes (1986) focus on the intergenerational transmission of abilities across generations and study the implications for the investments in human capital and the resulting dynamics of inequality.[1] More recently there has been extensive work on models with credit market imperfections and poverty traps. The basic idea behind the effect of imperfect financial markets is that if poor individuals are prevented from borrowing and hence cannot invest, then in a dynamic context initial inequalities may persist and some dynasties remain stuck into a poverty trap. Therefore these models can produce persistent inequality across dynasties abstracting from the effect of skills' differences. The most relevant models in this group are the models of Galor and Zeira (1993), Banerjee and Newman (1993), Piketty (1997), Aghion and Bolton (1997).

(ii) Models of political economy. The contribution of these models is represented by the fact that the prevailing level of redistribution is not exogenous but it is the result of a voting process. This idea has been introduced by the seminal paper of Meltzer and Richard (1981) and the seminal contributions which introduced it in dynamic macroeconomic models are those of Perotti (1993), Alesina and Rodrik (1994) and Persson and Tabellini (1994). The main idea behind those models is that, as in Meltzer and Richard (1981), given the median voter theorem, 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 voted level of redistribution in the economy. High levels of redistribution in turn

[1]See the review of Piketty (1998) for more on the literature on bequests and intergenerational mobility.

lower individual incentives to accumulate capital and hence the result that inequality lowers growth. In those models inequality derives from the fact that skills are fixed and persistently different across dynasties. Therefore such models are focused on the study of the determinants and the implications of redistribution but cannot give insights about the dynamics of mobility. Such models have been quite influential, especially in bringing endogenous political choices into the big picture. They also stimulated a great deal of discussion about the relationship between inequality, growth and redistribution. Some empirical evidence challenged the conclusion on the basis of two different observations. The first observation is that it does not always seem to be the case that inequality is detrimental to growth, even though the evidence in favor is quite large. The second and major challenge comes from the observation that it does not seem to be the case that more inequality implies higher redistribution. This second challenge inspired a new group of theoretical models whose major focus is to explain the described evidence relating inequality and redistributive politics. These models achieve this result showing the existence of multiple equilibria: a Europe type equilibrium characterized by relatively lower inequality and higher redistribution versus a US type equilibrium characterized by relatively higher inequality and lower redistribution. The model of Benabou (2000) is a seminal model which arrives to such conclusions. In the model of Benabou (2000) the prevailing level of redistribution is still a voting outcome, but unlike in the previously mentioned models the relationship between inequality and redistribution is not monotonic. Other theoretical models which obtain multiple equilibria with similar features are those of Saint-Paul (2001) and Hassler, Rodriguez-Mora, Storesletten, and Zilibotti (2003). Despite the fact that in the model of Benabou (2000) skills evolve stochastically across generations, every generation exerts effort before knowing the realization of ability and because of the same prior on the value of abilities, at each period all individuals exert the same value of effort. Therefore in the model there are no dynasties which remain stuck in poverty and the model cannot give insights about the dynamics of mobility. (iii) Models which focus on the role of beliefs. Starting from the evidence that the beliefs held by people about the underlying determinants of individual wealth and social mobility appear to be strong determinants of the social contract, the theoretical contributions of Piketty (1995), Alesina and Angeletos (2005) and Benabou and Tirole (2006) have developed insightful models describing how individual beliefs can shape politico-economic outcomes and viceversa and how multiple equilibria (US-type vs Europe-type) are possible. These models with beliefs can explain how beliefs affect redistribution and effort choices in static set-ups, but not how beliefs can affect intergenerational inequality and mobility.[2]

[2] The paper of Alesina and Angeletos (2005) contains a dynamic version of the main model. In this dynamic version skills are fixed and permanently different across generations. For this reason the model can show how beliefs can persist and societies can remain stuck in different equilibria as a result of different initial conditions, but the model cannot give insights about mobility. Also the model of Piketty (1995) is dynamic, but does not allow

In the model that I present in this chapter the amount of wealth left as bequest enter in the utility function of the parents as in the model of Atkinson (1980). This is a convenient formalization that I take from Alesina and Angeletos (2005) and which simplifies the dynamic problem to a great extent because it implies that every generation only wants to maximize present wealth and avoids issues of inter-temporal optimization and dynamic voting. The ultimate aim of this chapter is to build a dynamic model with bequests, stochastic skills, endogenous voting and endogenous information. Stochastic skills are an important ingredient to study mobility and to allow for incomplete information about the determinants of wealth but will imply main technical problems which I will discuss. Given those technical issues, I present to the full model by steps. Section 3.3 analyzes the case of exogenous political outcome and exogenous information, section 3.4 introduces voting but maintains information exogenous, section 3.5 allows for voting and endogenous information.

3.2. Set Up with Bequests

I consider an economy populated by a continuum of non-overlapping generations $i \in [0, 1]$. Each generation (or agent) lives for one period t and is labeled by i_t. Each generation i_t produces output y_t^i with the following

for intergenerational transfers. The dynamic aspect only concerns the learning about the determinants of wealth. The model can give insights about the beliefs on mobility but not about the actual mobility process.

technology:

$$y_t^i = \theta_t^i e_t^i + k_{t-1}^i, \tag{3.1}$$

where k_{t-1}^i represents the bequest or other parental investment received by the previous generation, e_t^i is the effort implemented by generation i_t and θ_t^i is the return to effort or productivity. With respect to abilities, I consider both the case in which abilities are i.i.d across different dynasties i but persistent over the life of a dynasty ($\theta_t^i = \theta^i$) and the more interesting case that θ_t^i is random and i.i.d. across i and t. As standard in this literature, agents face the linear tax/redistribution scheme introduced by Romer (1975). The individual budget constraint is given by

$$c_t^i + k_t^i = w_t^i = (1 - \tau_t) y_t^i + \tau_t \bar{y}_t, \tag{3.2}$$

where c_t^i denotes own consumption, k_t^i is the bequest left to the next generation, w_t^i denotes disposable wealth, τ_t is the tax rate, $\tau_t \bar{y}_t$ is the lump-sum transfer and $\bar{y}_t$ is the mean output in generation t. In each period t the timing of the actions follows the model of chapter 2. Each agent votes for the tax rate τ_t and exerts effort after that the tax rate is announced. Individuals receive neat wealth according to (3.2) and then decide how much to consume and to leave in bequests out of it. The private utility of each agent is given by the following function:

$$u_t^i(c_t^i, k_t^i, e_t^i) = \frac{1}{(1-\alpha)^{1-\alpha}\alpha^\alpha}(c_t^i)^{1-\alpha}(k_t^i)^\alpha - \frac{be_t^{i^2}}{2}. \tag{3.3}$$

The first term in (3.3) represents the utility from own consumption and bequests, whereas the second term is the disutility of effort.

Assuming a Cobb-Douglas aggregator over consumption and bequests with $\alpha \in (0,1)$, together with the constant $\frac{1}{(1-\alpha)^{1-\alpha}\alpha^\alpha}$ implies that $\alpha \in (0,1)$ denotes the fraction of wealth allocated to bequests and maintains the dynamic problem very tractable. Agent i_t chooses consumption, bequest and effort (c_t^i, k_t^i, e_t^i) so as to maximize utility subject to the budget constraint, taking the political outcome (τ_t) as given. It follows that the optimal individual consumption and bequest are respectively

$$c_t^i = (1-\alpha)w_t^i \tag{3.4}$$

and

$$k_t^i = \alpha w_t^i. \tag{3.5}$$

The indirect utility function in terms of wealth thus reduces to

$$u_t^i = w_t^i - \frac{be_t^{i^2}}{2}, \tag{3.6}$$

which is exactly the same utility function as in chapter 2. Given that the objective is the same as in chapter 2 the solutions to the problems of voting and choosing effort will be the same. This set up allows to avoid intertemporal optimization and problems of dynamic voting.

3.3. Step 1: exogenous tax rate and exogenous information.

In this first version of the model I maintain two assumptions: (i) I abstract from voting over τ considering this political outcome to be exogenously determined and constant over time ($\tau_t = \tau$), (ii) I consider the case that each individual i is fully informed about the value of θ^i. This section shows results which are already known by the previous literature but it is important for building the rest of the analysis.

Plugging the expression for pre-tax wealth (3.1) into the utility function (3.6) and solving the f.o.c. I find the expression for the individual optimal effort:

$$e_t^i = (1 - \tau_t)\theta_t^i / b. \tag{3.7}$$

Plugging the expression for optimal effort (3.7) and the expression for pre-tax wealth (3.1) into the individual budget constraint (3.2) I find the law of motion of bequests:

$$k_{t+1}^i = \alpha\big((1 - \tau)k_t^i + (1 - \tau)^2 \theta_t^{i^2} / b + \tau \bar{k}_t + \tau(1 - \tau)\bar{\theta_t^2} / b\big), \tag{3.8}$$

which determines also the law of motion of wealth given (3.5). In the case in which abilities are i.i.d. across different dynasties but constant over the life of a dynasty i ($\theta_t^i = \theta^i$), the law of motion (3.8) describes a convergent auto-regressive process.[3] It is immediate to derive the steady-state bequest of dynasty i

$$k^i = \frac{\alpha((1-\tau)^2\theta^{i^2}/b + \tau(1-\tau)\bar{\theta^2}/b + \tau\bar{k})}{1-\alpha(1-\tau)} \tag{3.9}$$

and the steady-state mean (or aggregate) bequest

$$\bar{k} = \frac{\alpha(1-\tau)\bar{\theta^2}}{b(1-\alpha)}. \tag{3.10}$$

Not surprisingly, given that the only source of heterogeneity is in the abilities θ^i, expression (3.9) shows that the greater is θ^i and the greater is the steady-state wealth. It is also obvious that redistribution has an equalizing effect. From (3.9) and (3.10) it can be computed that the difference between mean and individual i bequest equals

$$\bar{k} - k^i = \frac{\alpha(1-\tau)^2(\bar{\theta^2} - \theta^{i^2})}{b(1-\alpha(1-\tau))} \tag{3.11}$$

and decreases in the tax rate τ. At the same time expression (3.10) shows that redistribution diminishes aggregate bequest and hence aggregate wealth.

[3]The process is convergent as the coefficient of k_t^i is $\alpha(1-\tau) < 1$ and the rest of expression (3.8) is constant.

This trade-off between redistribution and growth is due to the fact that individual effort (3.7) decreases in the tax rate and redistribution has no other effect on output[4].

In order to have some more insights about the intergenerational dynamics of inequality and mobility it is interesting to explore the case that abilities are not persistent over the life of a dynasty. Assuming that abilities are drawn at random for each generation t and that $\theta_t^{i^2} = \bar{\theta}^2 + \epsilon_t^i$, where ϵ_t^i is an i.i.d. error term across i and t, with 0 mean, variance equal to σ^2 and zero serial correlation, the law of motion (3.8) still describes a convergent auto-regressive process.[5] The steady state mean bequest is still given by expression (3.10) and the variance is equal to

$$\text{var}(k^i) = \frac{\alpha^2(1-\tau)^4\sigma^2}{b^2(1-\alpha^2(1-\tau)^2)}. \tag{3.12}$$

As in the case with persistent abilities, increasing the rate of redistribution reduces inequality across agents.[6]

This model predicts convergency to a steady state value of wealth which does not depend on the initial level of wealth. The version with persistent abilities is qualitatively very similar to the seminal model of Stiglitz (1969)

[4]This is due to the linear production function and it is different from other models (reviewed in the introduction) with concave production functions, where redistribution can improve the efficiency of the inputs' allocations and hence increase output.

[5]In order to prove convergency, it is enough to cite the result of Hellwig (1980) which applies to Markov processes of this type.

[6]It is easy to compute that expression (3.12) monotonically decreases in τ. $d(var(k^i))/d\tau = \frac{\alpha^2\sigma^2}{b^2}\left[\frac{-4(1-\tau)^3(1-\alpha^2(1-\tau)^2)-2\alpha^2(1-\tau)^5}{(1-\alpha^2(1-\tau)^2)^2}\right] < 0$, given $\tau < 1$.

which is widely known for a Solow-type convergency in the context of individual wealth. Stiglitz (1969) model shows that with a concave saving function – in this case bequest function – if abilities are identical for everybody, everybody will converge to the same wealth, regardless of the amount of initial wealth. With heterogenous abilities, inequality is driven by abilities but once again initial wealth inequality does not matter in the long run.[7] Plugging the expression of optimal effort (3.7) and the expression of pre-tax wealth (3.1) into the utility function (3.6) gives the expression for expected utility as a function of the tax rate τ:

$$\bar{u} = \bar{y} - \frac{c\bar{e^2}}{2} = \bar{k} + \frac{(1-\tau^2)\bar{\theta^2}}{2b}. \tag{3.13}$$

As in the analysis of chapter 2, expected utility is maximized by a zero tax rate. The reason is that the utility function is linear in wealth and therefore there are no ex-ante gains from redistribution; on the other hand effort and output decrease in the tax rate.

Case of incomplete information.

In order to build the analysis of the next sections, it is useful to analyze the case in which agents have incomplete information and expect to be of

[7] A detailed analysis of the results of Stiglitz (1969) in the context of intergenerational transmission of wealth is done contained in the review of Piketty (1998); Piketty (1998) explains that, as shown by Bourguignon (1981), in the case of convex bequest function the result about convergency does not generally hold. Other causes that can imply non convergency are different fertility behaviours across households and credit market imperfections. He also discusses the case of random abilities.

average ability $\bar{\theta}$. In this case optimal effort is equal to

$$e_t^i = (1-\tau_t)\bar{\theta}/b. \tag{3.14}$$

Plugging this expression and the expression for pre-tax wealth (3.1) into the individual budget constraint (3.2) gives the law of motion of bequest:

$$k_{t+1}^i = \alpha\big((1-\tau)k_t^i + (1-\tau)^2(\bar{\theta})^2/b + \tau\bar{k}_t + \tau(1-\tau)(\bar{\theta})^2/b\big). \tag{3.15}$$

The steady-state individual bequest, mean bequest and variance follow:

$$k^i = \frac{\alpha((1-\tau)^2\bar{\theta}\theta^i/b + \tau(1-\tau)(\bar{\theta})^2/b + \tau\bar{k})}{1-\alpha(1-\tau)}, \tag{3.16}$$

$$\bar{k} = \frac{\alpha(1-\tau)(\bar{\theta})^2}{b(1-\alpha)}, \tag{3.17}$$

$$\text{var}(k) = \frac{(1-\tau)^4\sigma^2}{b^2(1-\alpha^2(1-\tau)^2)}. \tag{3.18}$$

Given that $\bar{\theta^2} > (\bar{\theta})^2$, the mean wealth with complete information (3.10) is strictly greater than the mean wealth under incomplete information (3.17). As explained in chapter 2 information improves the efficiency of effort allocations and increases output. The same beneficial effect of information appears by comparing the expression of expected utility in the case of complete information (3.13) with the respective expression in the case of incomplete information:

$$\bar{u} = \bar{k} + \frac{(1-\tau^2)(\bar{\theta})^2}{2b}. \tag{3.19}$$

From (3.16) and (3.17) it can be computed that under incomplete information the difference between mean and individual equals

$$\bar{k} - k^i = \frac{\alpha(1-\tau)^2\bar{\theta}(\bar{\theta}-\theta^i)}{b(1-\alpha(1-\tau))} \tag{3.20}$$

and decreases in the tax rate τ.

It is easy to notice that the steady state with incomplete information is characterized by lower inequality than the steady state with incomplete information, as the difference between expression (3.11) and (3.20) is equal to $\alpha(1-\tau)^2(1-\pi)(\theta_H-\theta_L)(\pi\theta_H+(1-\pi)\theta_L)b(1-\alpha(1-\tau))$ which is a positive term given $\tau < 1$.

3.4. Step 2: endogenous tax rate and exogenous information.

The second step of the analysis is to introduce voting. Being the utility function (3.6) identical to utility function (2.10) of chapter 2, the voting problem and its solution at each t is also identical. Agents vote for the tax rate before exerting effort; solving backward I find the objective function of voter i_t by plugging the expression of optimal effort (3.7) and the expression of net wealth (3.2) into the utility function (3.6) and then maximizing the obtained expression with respect to τ. As shown in the analysis of chapter 2, each voter i maximizes the following indirect utility function in τ:

$$u_t^i = \tau(k_t^i - \bar{k}) + (1-\tau)^2\theta^{i^2}/b + \tau(1-\tau)\bar{\theta^2}/b - (1-\tau)^2\theta^{i^2}/2b. \tag{3.21}$$

Assuming for the moment that the second derivative of expression (3.21) with respect to τ is strictly negative, the first order condition gives the ideal tax rate of voter i:

$$\tau_t^i(k_t^i, \theta^i) = 1 - \frac{1 + \frac{b(k_t^i - \bar{k}_t)}{\bar{\theta^2}}}{2 - \frac{\theta^{i2}}{\bar{\theta^2}}}. \tag{3.22}$$

Both the numerator and the denominator of (3.22) show that the gains from redistribution are traded off the moral hazard effect of taxation. I introduce the following assumption in order to assure the concavity of the objective function (2.18) and therefore in order to use the median voter theorem:

Assumption 3.1: $2\theta_L^2 > \theta_H^2$.

A proposition follows:

PROPOSITION 3.1. *The individual preferences for taxation are single-peaked and the individual ideal tax rate is given by expression (3.22).*

PROOF. The second derivative of the objective function (3.21) is given by the following expression:

$$\frac{d^2 u_t^i}{d\tau} = \frac{(\theta^i)^2 - 2\bar{\theta^2}}{b}.$$

The condition stated by assumption 3.1 is sufficient for (3.23) to be strictly negative as the maximum value that $(\theta^i)^2$ can take is θ_H^2 and the minimum value that $2\bar{\theta}^2$ can take is $2\theta_L^2$. □

Proposition 3.1 shows that preferences over the tax rate are single peaked and therefore the median voter theorem applies. In order to analyze a steady state of the dynamic model, I look for a steady state such that given a stationary history $\tau_s = \tau$ for all generations $s \leq t-1$, then $\tau_t = \tau$ is optimal for generation t.

Persistent abilities across generations. I start with the case in which abilities are persistent over dynasties, namely $\theta_t^i = \theta^i$. Consider a stationary history $\tau_s = \tau$ for all generations $s \leq t-1$. For each dynasty i the low of motion (3.8) implies that the value of the bequest converges to expression (3.9) and that the mean value converges to expression (3.10). Given persistent abilities ($\theta_t^i = \theta^i$) and assuming that every dynasty starts life with no endowment ($k_0^i = 0$ for all i), in every period t the median and prevailing tax rate τ_t is that one of the dynasty with median ability θ^m. Plugging the expressions of the steady state median bequest k^m (obtained through (3.9)) and of the steady state mean bequest (3.10) into the expression for the individual ideal tax rate (3.22) and using (3.11), I find the tax rate τ_t which follows a given stationary history:

$$\tau_t = \frac{\bar{\theta}^2 - \theta^{m^2} + \frac{\alpha(1-\tau)^2(\bar{\theta}^2 - \theta^{m^2})}{1-\alpha(1-\tau)}}{2\bar{\theta}^2 - \theta^{m^2}}. \tag{3.23}$$

This expression is decreasing in τ, hence there is a unique fixed point. This result of a unique steady state is found also by Alesina and Angeletos (2005). It is the standard result of Meltzer and Richard (1981) in a dynamic context. Redistribution is driven by the difference between mean and median wealth and this is the case in every period. The dynamic implication is that present redistribution depends on the history of past redistributive outcomes and more precisely the tax rate declines over time.[8]

Random abilities across generations. I explore the case of random abilities. As I have done in the previous section with exogenous voting I assume that abilities are drawn at random for each generation and that $\theta_t^{i^2} = \bar{\theta}^2 + \epsilon_t^i$, where ϵ_t^i is i.i.d. across i and t with mean = 0 and variance equal to σ^2. Considering a stationary history $\tau_s = \tau$ for all generations $s \leq t-1$, the law of motion (3.8) implies that the mean bequest converges to expression (3.10) and that the variance is equal to expression (3.12).

In general, it is complicated to identify the median voter at time t because this depends on both the distributions of k and θ. To see this take expression (3.22) and re-express it as $\tau_t^i(k_t^i, \theta_t^i) = \frac{\bar{\theta}^2 - \theta^{i2} + b(k_t^i - \bar{k}_t)}{2\bar{\theta}^2 - \theta^{i2}}$. This shows that the distribution of τ_t^i is given by the ratio of two distributions, respectively the numerator and the denominator of expression (3.22). Take for example the case in which ϵ_t^i is normally distributed and each generation starts with

[8]The same result is also found by Bertocchi (2007) in a model which specifically analyzes the evolution of bequest taxation over time.

no endowment at time 0 ($k_0^i = 0$ for all i). In this case both k^i and θ^i are normally distributed at time t and the distribution of $\tau_t^i(k_t^i, \theta_t^i) = \frac{\bar{\theta^2} - \theta^{i2} + b(k_t^i - \bar{k_t})}{2\bar{\theta^2} - \theta^{i2}}$, is given by the ratio of two normal distributions. It is quite complicated to identify such ratio distribution where the two normal variables have different means and to find the median.[9] This difficulty to deal with dynamic models of voting in which abilities change over time has been recognized in the literature.[10]

Random abilities across generations and incomplete information. In order to skip this technical issue it is useful to consider the case of incomplete information in which everyone expects to be of average ability $\bar{\theta}$. In this case the ideal tax rate of each individual i is given by expression (3.22) once that θ^{i2} and $\bar{\theta^2}$ are replaced by $(\bar{\theta})^2$, obtaining

$$\tau^i(k^i, \bar{\theta}) = \frac{b(k_t^i - \bar{k_t})}{(\bar{\theta})^2}. \tag{3.24}$$

Given that the distribution of the ideal tax rates only depends on the distribution of k^i it is immediate to identify the median voter to be the voter with median endowment k^m. I consider two specific distributions as examples.

[9] A technical analysis of the properties of this type of ratio distribution is done by Hinkley (1969).

[10] To my knowledge, the only dynamic model of voting in which abilities change over time is the one of Persson and Tabellini. (1991), but in this paper there is no identification of the median voter in a steady state. In the published version of the paper (Persson and Tabellini (1994)) it is only considered the case of persistent abilities. All the other papers in the literature only consider the case of persistent abilities: Alesina and Rodrik (1994), Das and Ghate (2004), Hassler, Rodriguez-Mora, Storesletten, and Zilibotti (2003), Saint-Paul (2001), Bertocchi (2007).

Example 1: binomial distribution of abilities. In every period each individual i has ability θ_L with probability π and ability θ_H with probability $1-\pi$ and I assume that every dynasty starts life with no endowment ($k_0^i = 0$ for all i). Given a stationary history, the distribution of k will converge to a normal distribution, because the distribution of k comes from repeated independent Bernuolli trials over ability realizations and the distribution of those realizations converges to a normal distribution. The fact that the distribution of k converges to the normal implies that at time t the median and the mean endowment coincide and therefore that $\tau_t = 0$. Thus in the case of binomial distribution of abilities there is a unique steady-state with zero tax. This is again a case in which inequality progressively decreases. Being the steady state with zero tax, there is convergency to the same steady state as in a model without redistribution. On the other hand taxation may increase the speed of convergency.[11]

Example 2: Log-Normal distribution of abilities. I consider the case in which $\theta_t^{i^2} = \bar{\theta}^2 + \epsilon_t^i$, with ϵ_t^i i.i.d for all i and t and log-normally distributed. I also assume that in period 0 everyone starts life with no endowment, $k_0^i = 0$ for all i. Given that agents have incomplete information about θ^i and believe to be of average ability $\bar{\theta}$, a stationary history is still specified by expressions (3.15), (3.17), (3.18) and k is log-normally distributed in every period. Using the properties of the log-normal distribution, given $\bar{k}$ and $\text{var}(k^i)$, the steady state median bequest is equal to $k^m = \frac{\bar{k}^2}{\sqrt{\bar{k}^2 + \text{var}(k^i)}}$.

[11]Also this point has been early discussed by Stiglitz (1969)

Given such stationary history, the voted tax rate at time t is still given by $\tau_t = \frac{b(\bar{k}(\tau) - k^m(\tau))}{\bar{\theta}^2}$. The difference $\bar{k} - k^m$ decreases in τ, hence also in this case there is a unique fixed point. Noticing that k^m decreases in var(k^i), therefore the greater it is the underlying inequality σ and the greater it is the steady state level of redistribution τ. I can solve for $\tau_t = \tau$ and characterize the steady state. Given that the distribution is skewed to the right, the steady state level of redistribution can be different from zero, for example the computations for the case of $\alpha = 0.2, b = 1, \bar{\theta^2} = 1, \sigma^2 = 2$ imply that $\tau = 0.17$.

3.5. Step 3: endogenous tax rate and endogenous information.

The previous section clarified the difficulties implied by the introduction of voting in a model with heterogenous abilities. Using the set up of chapter 2, in this section I allow for varying levels of information in the present dynamic model with bequests and I consider the level of information as an endogenous variable in the economy. The main contribution is to shed some light on the technical difficulties which are implied. The main technical problem with varying levels of information is still represented by the identification of the median voter. The individual ideal tax rate is determined by the individual value of wealth and the expected ability. A steady state with endogenous information must have the feature that the level of information is optimal. In order to check for this it is necessary to verify that there are no gains in changing the level of information. This is difficult because changing the level of information changes expectations, hence changes the distribution of ideal tax rates and such change is difficult to address. I provide a numerical examples with multiple politico economic equilibria. In this example both complete information ($\lambda = 1$) and minimum information ($\lambda = 1/2$) are optimal; the two equilibria have different macroeconomic features and can be interpreted as Europe-type versus American-type equilibria.

3.5.1. Set up with varying information. Each generation i_t has ability θ_L with probability π and ability θ_H with probability $1-\pi$, for all i and t. I maintain the assumption that $\pi > 1/2$. In each period t each agent i cannot observe her own or other agents' productivity but only receives a private signal σ_t^i about the true value of θ_t^i. Also the signal σ_t^i is binary. If $\theta_t^i = \theta_L$ ($\theta_t^i = \theta_H$), σ_t^i takes values σ_L (σ_H) or σ_H (σ_L), respectively with probability λ_t and $1-\lambda_t$. In other words for each agent i_t the signal σ_t^i is independently distributed, it is truthful with probability λ_t, false with probability $1-\lambda_t$ and the transition matrix which takes from the true productivity to the signal is the following:

$$T\left(\begin{bmatrix}\sigma_L\\ \sigma_H\end{bmatrix}\Bigg|[\theta_L,\theta_H]\right)=\begin{pmatrix}\lambda_t & 1-\lambda_t\\ 1-\lambda_t & \lambda_t\end{pmatrix}. \tag{3.25}$$

Agent's i belief of the true value of θ_t^i, conditional on the observation of the private signal σ_t^i, is obtained by the Bayes Rule. I introduce the following notation:

$$\mu_t^i \equiv \Pr[\theta_t^i = \theta_L | \sigma_t^i], \tag{3.26}$$

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

$$\mu_{t_{\sigma_L}} \equiv (\mu_t^i | \sigma_L) = \frac{\pi \lambda_t}{\pi \lambda_t + (1-\pi)(1-\lambda_t)} \tag{3.27}$$

and

$$\mu_{t_{\sigma_H}} \equiv (\mu_t^i | \sigma_H) = \frac{\pi (1-\lambda_t)}{\pi (1-\lambda_t) + \lambda_t (1-\pi)}. \tag{3.28}$$

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

$$\theta(\mu_t^i) \equiv \mu_t^i \theta_L + (1-\mu_t^i)\theta_H. \tag{3.29}$$

Given the symmetric structure of (2.2) I consider the interval $\lambda_t \in [1/2, 1]$, for all t. For $\lambda_t = 1/2$ the signal σ_t^i is completely uninformative and the posterior belief is equal to the prior, i.e. $\mu_{t_{\sigma_L}}$=$\mu_{t_{\sigma_H}}$=π. Increasing λ_t makes the signal progressively more informative up to the point that $\lambda_t = 1$ and the signal is perfectly informative with $\mu_{t_{\sigma_L}} = 1$, $\mu_{t_{\sigma_H}} = 0$. As already explained in the previous chapter, 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 for each generation alive at t is given by the following expression:

$$p_{\sigma_L t} \equiv \Pr[\sigma_t^i = \sigma_L] = \lambda_t \pi + (1 - \lambda_t)(1 - \pi), \tag{3.30}$$

symmetrically

$$p_{\sigma_H t} \equiv \Pr[\sigma_t^i = \sigma_H] = \lambda_t(1 - \pi) + \pi(1 - \lambda_t) = 1 - p_{t_{\sigma_L}} \tag{3.31}$$

represents the probability of observing σ_H. 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. The timing of the model is such that each individual i who is alive at time t starts life receiving a signal σ_t^i and being aware of the level of precision λ_t. The game proceeds as before: agents vote on tax, exert effort, receive net wealth, consume and leave bequests. The last action of the agents alive at t is to collectively decide the level of information for the offspring, namely λ_{t+1}. I assume that the future level of information λ_{t+1} is determined by majority voting.

In the case in which abilities are random in each period t I would face the technical difficulties implied by the determination of the median voter which I explained in section 3.4. In order to keep the model tractable I restrict the analysis to the case of persistent abilities and I assume that there are only two dynasties: the low ability dynasty with ability θ_L in each period t and the high dynasty with ability θ_H in each period t, respectively a fraction π and $1-\pi$ of the population. Nevertheless agents ignore completely the persistence of abilities and they "falsely"believe their prior: namely that in each period t abilities are i.i.d across agents and agents are of type θ_L with probability π and of type θ_H with probability $1-\pi$ in each period t. As explained before, on top of this prior, agents receive a signal about their ability. I stress the fact that the assumption that abilities are persistent and agents ignore this fact is very strong and it is not done for sake of realism but exclusively to maintain the model tractable. [12] Nevertheless the main result of existence of multiple optimal levels of information and multiple steady-states is not driven by this assumption.

3.5.2. Politico Economic Equilibrium with perfect information. Given a stationary history $s \leq t-1$ such that the voted tax rate is always $\tau_s = \tau$ and the voted information is always full information $\lambda_{s+1} = 1$, I describe a steady state such that at time t, the voted tax rate is still $\tau_t = \tau$ and the

[12]It is clear that the existence of bequests would imply that at some point agents learn that abilities are persistent. This because in each period t agents would either belong to a group who received either the low or the high level of bequests. Therefore for the prior not to be updated across generations it should be also the case that agents should ignore or forget what happened to the previous generations.

voted information is still $\lambda_{t+1} = 1$. Given a stationary history $\tau_s = \tau$ and $\lambda_{s+1} = 1$ for $s \leq t-1$, at time t the bequest for dynasty L (H) converges to the steady state value k_L (k_H) given by (3.9) and the mean bequest converges to the steady state value (3.10). Given the history with persistent abilities and perfect information, at time t there are two groups of voters, respectively with preferred tax rates $\tau(k_L, \theta_L)$ and $\tau(k_H, \theta_H)$ given by expression (3.22). The prevailing tax rate at time t is $\tau_t = \tau(k_L, \theta_L)$, because $\pi > 1/2$ implies that the majority of the population belongs to the dynasty with low ability and low endowment. In a steady state it must be the case that $\tau_t = \tau$ and that the bequest that agents with endowments k_L (k_H) leave for the $t+1$ offsprings is still k_L (k_H).

3.5.3. Politico Economic Equilibrium with incomplete information. Given a stationary history $s \leq t-1$ such that the voted tax rate is always $\tau_s = \tau$ and the voted level of information is always the minimum $\lambda_{s+1} = 1/2$, I describe a steady state such that at time t, the voted tax rate is still $\tau_t = \tau$ and the voted information is still $\lambda_{t+1} = 1/2$. Given a stationary history $\tau_s = \tau$ and $\lambda_{s+1} = 1/2$ for $s \leq t-1$, at time t the bequest for dynasty L (H) converges to the steady state value k_L (k_H) given by (3.16) and the mean bequest converges to the steady state value (3.17). Given the history with persistent abilities and perfect information, at time t there are two groups of voters, respectively with preferred tax rates $\tau(k_L, \bar{\theta})$ and $\tau(k_H, \bar{\theta})$ given by expression (3.24). The prevailing tax rate at time t is $\tau_t = \tau(k_L, \bar{\theta})$, because $\pi > 1/2$ implies that the majority of the population belongs to the dynasty with low ability and low endowment. In a steady state it must be the case that $\tau_t = \tau$ and that the bequest that agents with endowments k_L (k_H) leave for the $t+1$ offsprings is still k_L (k_H).

3.5.4. Solution of the individual problem. In addition to leaving bequests, at time t agents decide by majority voting the future information λ_{t+1}. Each agent i_t votes on λ_{t+1} in order to maximize the utility of the offspring i_{t+1}. In order to find the utility of an offspring as a function of λ_{t+1} it is necessary to solve backwards the choices of effort and voting of the agents alive at $t+1$. Conditional on the signal, the expected utility that the generations who are alive at time t maximize when they vote and exert effort is the following:

$$E[u_t^i|\sigma_t^i] = E[(1-\tau_t)(k_t^i + e_t^i\theta_t^i) + \tau_t(\bar{k}_t + \overline{e_t\theta_t}) - b(e_t^i)^2/2|\sigma_t^i]. \tag{3.32}$$

Solving the sufficient first order condition, the optimal level of effort exerted by individual i is

$$e_t^i = (1-\tau_t)\theta(\mu_t^i)/b. \tag{3.33}$$

By backward induction, I can plug (3.33) into (3.32) and find the objective function that i maximizes when voting for the tax rate. In order to do this, it is useful to specify the individual i expectation of the output from effort:

$$E[e_t^i\theta_t^i|\sigma_t^i] = (1-\tau_t)\left(\theta(\mu_t)\right)^2/b \tag{3.34}$$

and of squared effort

$$E[(e_t^i)^2|\sigma_t^i] = (e_t^i)^2 = \left(\frac{1-\tau_t}{a}\right)^2\theta(\mu_t)^2. \tag{3.35}$$

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 the individual expectation of the aggregate output from effort is given by the following expression:

$$E[\overline{e_t\theta_t}|\sigma_t^i] = (1-\tau_t)\Gamma/b, \tag{3.36}$$

where I define

$$\Gamma \equiv \pi\theta_L\left(\lambda\theta(\mu_{\sigma_L}) + (1-\lambda)\theta(\mu_{\sigma_H})\right) + \\ (1-\pi)\theta_H\left((1-\lambda)\theta(\mu_{\sigma_L}) + \lambda\theta(\mu_{\sigma_H})\right). \tag{3.37}$$

Collecting $\theta(\mu_{\sigma_L})$ and $\theta(\mu_{\sigma_H})$ it is easy to re-write expression (3.41) as

$$\Gamma = p_{\sigma_L}\theta(\mu_{\sigma_L})^2 + (1-p_{\sigma_L})\theta(\mu_{\sigma_H})^2. \tag{3.38}$$

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 will play a crucial role in the analysis. Plugging (3.34), (3.35) and (3.41) into (3.32), I obtain an indirect form of (3.32) as a function of τ_t:

(3.39)

$$u_t^i(\tau_t, \mu_t^i) = \tau_t(k_t^i - \bar{k}_t) + (1-\tau_t)^2\theta(\mu_t^i)^2/b + \tau_t(1-\tau_t)\Gamma/b - (1-\tau_t)^2\theta(\mu_t^i)^2/2b.$$

This is the object that voter i_t maximizes voting over the tax rate τ_t. Assuming for the moment that the second derivative of the obtained indirect utility in τ is strictly negative, the first order condition gives the ideal tax rate of voter i:

$$\tau(k_t^i, \mu_t^i) = 1 - \frac{1 + \frac{b(k_t^i - \bar{k}_t)}{\Gamma}}{2 - \frac{\theta(\mu_t^i)^2}{\Gamma}}, \tag{3.40}$$

where

$$\Gamma \equiv \pi\theta_L\left(\lambda\theta(\mu_{\sigma_L}) + (1-\lambda)\theta(\mu_{\sigma_H})\right) + (1-\pi)\theta_H\left((1-\lambda)\theta(\mu_{\sigma_L}) + \lambda\theta(\mu_{\sigma_H})\right). \tag{3.41}$$

As already explained in the previous chapter, the numerator of (3.40) shows that the gains from redistribution are traded off the distortive effect of redistribution and the denominator of (3.40) shows how the subjective prospects of upward mobility reduce the desired tax rate.[13]

PROPOSITION 3.2. *The individual preferences for taxation are single peaked and the individual ideal tax rate is given by expression (3.40) .*

PROOF. The second derivative of the objective function in problem (2.18) is given by the following expression: $\frac{d^2u_1^i}{d\tau} = \frac{-2\Gamma+\theta(\mu)^2}{b}$. The condition stated by Assumption 2.1 is sufficient for this expression 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 3.2 shows that preferences over the tax rate are single peaked and therefore the median voter theorem applies. Plugging (3.33) into (3.32) and taking expectations conditional on the information at time t, I obtain generation i_t expectation of generation i_{t+1} utility:

$$E_t^i[u_{t+1}^i] = \tau_{t+1}(\bar{k} - k^i) + (1 - (\tau_{t+1})^2)\Gamma_{t+1}/2b. \tag{3.42}$$

I model the collective choice of λ_{t+1} as a choice by majority voting. Given that the agents with k_L are the majority, their choice of λ_{t+1} will determine the prevailing one. Therefore for a specific level of information λ' to be part

[13]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 (3.34) over aggregate output (3.36), noticing that the term $\frac{1-\tau}{a}$ gets canceled out.

of an equilibrium it is necessary that λ' is the arg max $\{\tau_{t+1}(\bar{k}-k_L)+(1-(\tau_{t+1})^2)\Gamma/2b.\}$ In order to check for this it is necessary to know how τ_{t+1} changes in λ. At time $t+1$ there are four groups of voters, respectively with preferred tax rates $\tau(k_L,\theta_{\sigma_L})$, $\tau(k_H,\theta_{\sigma_L})$, $\tau(k_L,\theta_{\sigma_H})$, $\tau(k_H,\theta_{\sigma_H})$. I claim that in this case of $\pi>1/2$ the prevailing tax rate under majority voting is either $\tau(k_L,\theta_{\sigma_L})$, or the greater between $\tau(k_H,\theta_{\sigma_L})$ and $\tau(k_L,\theta_{\sigma_H})$, depending on the value of λ. This claim can be easily proved. The fraction of agents who prefer $\tau(k_L,\theta_{\sigma_L})$ is equal to πp_{σ_L}, where p_{σ_L} is given by (3.30). When $\lambda=1$ the fraction π of agents with endowment k_L knows to be of type θ_L. In the case in which $\pi>1/2$, this implies that they are the majority group and impose their favorite tax rate. Decreasing λ implies that agents can have two types of beliefs, namely θ_{σ_L} and θ_{σ_H}, and there are four group of voters, respectively with preferred tax rates $\tau(k_L,\theta_{\sigma_L})$, $\tau(k_H,\theta_{\sigma_L})$, $\tau(k_L,\theta_{\sigma_H})$, $\tau(k_H,\theta_{\sigma_H})$. Decreasing λ implies that πp_{σ_L} decreases. There is a value $\lambda^*\in(1/2,1)$ such that $\pi p_{\sigma_L}=1/2$, namely $\lambda^*=\frac{-2\pi+2\pi^2+1}{2\pi(2\pi-1)}$. For $\lambda\in[1/2,\lambda^*)$ it happens that the group which prefers $\tau(k_L,\theta_{\sigma_L})$ is not the majority group, and the pivotal group will be either $\tau(k_H,\theta_{\sigma_L})$ or $\tau(k_L,\theta_{\sigma_H})$ depending on which is the greater tax rate of the two. This is because in the case in which $\tau(k_H,\theta_{\sigma_L})\geq\tau(k_L,\theta_{\sigma_H})$, then it is the case that the total ranking of tax rates is $\tau(k_L,\theta_{\sigma_L})>\tau(k_H,\theta_{\sigma_L})\geq\tau(k_L,\theta_{\sigma_H})$ and the fact that the fraction p_{σ_L} of agents with belief θ_{σ_L} is greater than 1/2 (this because $\pi>1/2$ and $\lambda>1/2$) implies that the median voter must belong to the group with $\tau(k_H,\theta_{\sigma_L})$. Otherwise in the case in which $\tau(k_L,\theta_{\sigma_H})\geq\tau(k_H,\theta_{\sigma_L})$, it is the case that the total ranking of tax rates is $\tau(k_L, \theta_{\sigma_L}) > \tau(k_L, \theta_{\sigma_H}) \geq \tau(k_H, \theta_{\sigma_L})$ and the fact that the fraction of agents with k_L is greater than $1/2$ (this because $\pi > 1/2$) implies that the median voter must belong to the group with $\tau(k_L, \theta_{\sigma_H})$.

3.5.5. Example of multiple politico-economic equilibria. I consider the following numerical example: $\theta_L = 1, \theta_H = 1.5, \pi = 0.7, \alpha = 0.2, b = 1$. I first consider the case of complete information $\lambda = 1$. Evaluating k_L, $\bar{k}$ and τ_t for the given values and solving for $\tau_t = \tau$ with I obtain that $\tau = 0.243$. I plug this value back into expressions (3.9) and (3.10) and I find the steady state values $k_L = 0.209$ and $\bar{k} = 0.26$. The majority of the agents at time t have endowment k_L. Under majority voting, they determine λ_{t+1} in order to maximize the expected welfare of their offspring:

$$u_L = \tau_{t+1}(\bar{k} - k_L) + (1 - (\tau_{t+1})^2)\Gamma/2. \tag{3.43}$$

If $\pi p_{\sigma_L} \geq 1/2$ then the prevailing tax rate at $t+1$ is $\tau = \tau(k_L, \theta(\mu_{\sigma_L}))$. Given $\pi = 0.7$ and using the definition of $p_{\sigma_L t+1} = \pi\lambda_{t+1} + (1-\pi)(1-\lambda_{t+1})$, $\pi p_{\sigma_L t+1} \geq 1/2$ for $\lambda_{t+1} \geq 1.03$, therefore it is never the case that $\tau(k_L, \theta(\mu_{\sigma_L}))$ is the pivotal. The prevailing tax rate is the greater one between $\tau(k_L, \theta(\mu_{\sigma_H}))$ and $\tau(k_H, \theta(\mu_{\sigma_L}))$. I plot both those two tax rates, respectively in figure 3.1 and 3.2.

Computations show that they equal each other for 0.584. Hence in the considered interval $\tau(k_L, \mu_{\sigma_H})$ is greater and it is the prevailing tax rate. It is also important to notice that $\tau(k_L, \mu_{\sigma_H}) \leq 0$, for $\lambda \geq 0.55$ and that

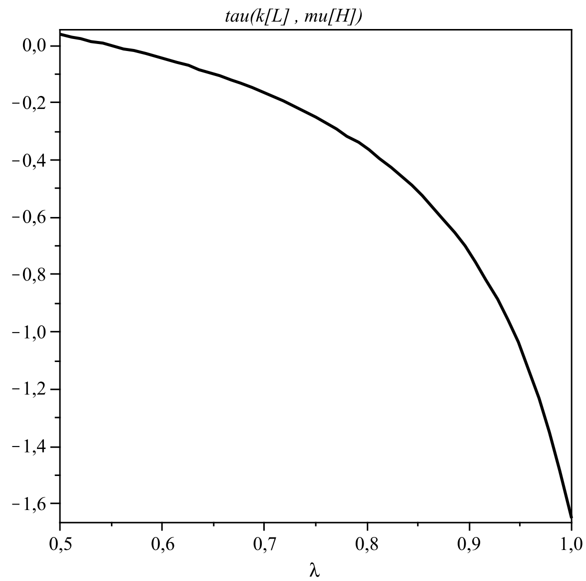

FIGURE 3.1. $\tau(k_L, \mu_{\sigma_H})$ for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $b = 1$, $\alpha = 0.2$.

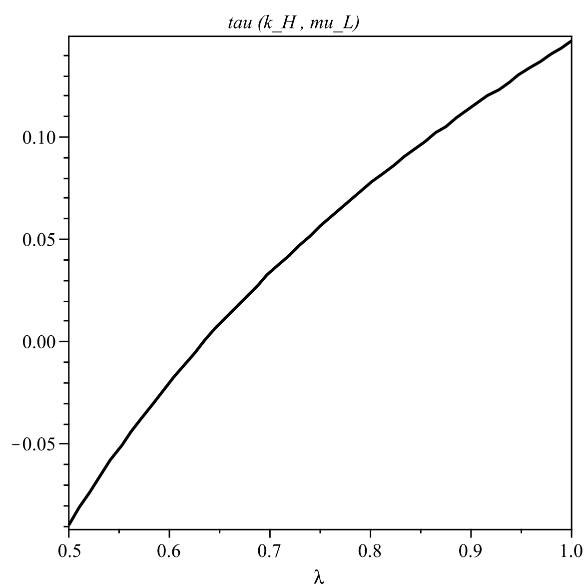

FIGURE 3.2. $\tau(k_H, \mu_{\sigma_L})$ for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $b = 1$, $\alpha = 0.2$.

$\tau(k_H, \mu_{\sigma_L}) \leq 0$, for $\lambda \leq 0.63$. Therefore I consider $\tau = 0$ in the interval $\lambda \in [0.55, 0.63]$. I plot the objective function (3.43) in figure 3.3, where

$\tau_{t+1} = \tau(k_L, \mu_{\sigma_H})$ for $\lambda_{t+1} \in [0.5, 0.55]$, $\tau_{t+1} = 0$ for $\lambda_{t+1} \in [0.55, 0.63]$, $\tau_{t+1} = \tau(k_H, \mu_{\sigma_L})$ for $\lambda_{t+1} \in [0.63, 1]$. $\lambda_{t+1} = 1$ is a global maximum.

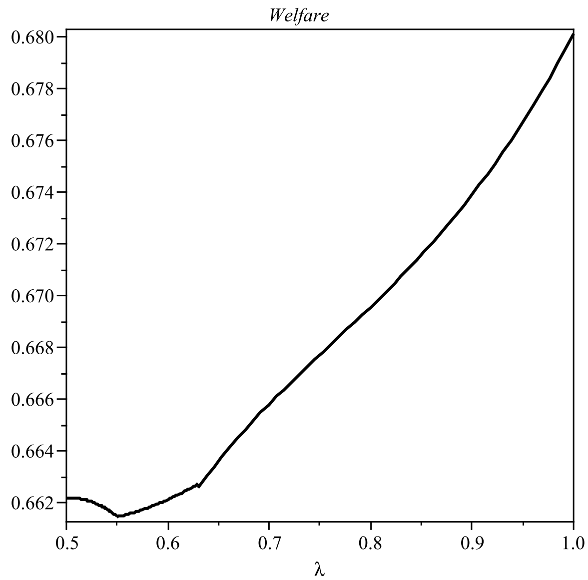

FIGURE 3.3. Objective function (3.43) for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $b = 1$, $\alpha = 0.2$.

I now consider the case of completely uninformative signals $\lambda = 1/2$. Given a stationary history $s \leq t-1$ such that the voted tax rate is $\tau_s = \tau$ and the voted information is the minimum information $\lambda_{s+1} = 1/2$, I describe a steady state such that at time t, the voted tax rate is still $\tau_t = \tau$ and the voted information is still $\lambda_{t+1} = 1/2$. Given a stationary history $\tau_s = \tau$ and $\lambda_{s+1} = 1/2$ for $s \leq t-1$, at time t the bequest for dynasty L (H) converges to the steady state value k_L (k_H) given by (3.16) and the mean bequest converges to the steady state value (3.17). Given the history with persistent abilities and minimum information, at time t there are two groups of voters, respectively with preferred tax rates $\tau(k_L, \bar{\theta})$ and $\tau(k_H, \bar{\theta})$ given by expression (3.22). The prevailing tax rate at time t is $\tau_t = \tau(k_L, \bar{\theta})$, because $\pi > 1/2$ implies that the majority of the population belongs to the dynasty with low ability and low endowment. In a steady state it must be the case that $\tau_t = \tau$ and that the bequest that agents with endowments k_L (k_H) leave for the $t+1$ offsprings is still k_L (k_H). Following the previous example for $\lambda = 1/2$ to be part of a steady state it is necessary that $\lambda = 1/2$ is the arg max expression (3.43).

I consider the same numerical example: $\theta_L = 1, \theta_H = 1.5, \pi = 0.7, \alpha = 0.2, b = 1$. Evaluating k_L, $\bar{k}$ and τ_t for the given values and solving for $\tau_t = \tau$ with Maple I obtain $\tau = 0.03$. I plug this value back into expressions (3.16) and (3.17) and I find the steady state values $k_L = 0.28$ and $\bar{k} = 0.32$. If $\pi p_{\sigma_L t+1} \geq 1/2$ then the prevailing tax rate at $t+1$ is $\tau = \tau(k_L, \theta(\mu_{\sigma_L}))$. Given $\pi = 0.7$ and using the definition of $p_{\sigma_L t+1} = \pi\lambda_{t+1} + (1-\pi)(1-\lambda_{t+1})$, $\pi p_{\sigma_L t+1} \geq 1/2$ for $\lambda_{t+1} \geq 1.03$, therefore it is never the case that $\tau(k_L, \theta(\mu_{\sigma_L}))$ is the pivotal. The prevailing tax rate is the greater one between $\tau(k_L, \theta(\mu_{\sigma_H}))$ and $\tau(k_H, \theta(\mu_{\sigma_L}))$. I plot both those two tax rates for $\lambda \in [1/2, 1]$, respectively in figure 3.4 and 3.5. They equal each other for 0.567, therefore the prevailing tax rate is $\tau(k_L, \theta(\mu_{\sigma_H}))$ for $\lambda \in [0.5, 0.567]$ and $\tau(k_H, \theta(\mu_{\sigma_L}))$ for $\lambda \in [0.567, 1]$. Hence in the considered interval $\tau(k_L, \mu_H)$ prevails. It is also important to notice that $\tau(k_L, \mu_H) \leq 0$, for $\lambda \geq 0.551$, and that $\tau(k_H, \mu_L) \leq 0$, for $\lambda \geq 0.6$, therefore I consider $\tau = 0$ in the interval $\lambda \in [0.567, 0.6]$. I plot the objective function (3.43) in figure 3.3, where $\tau_{t+1} = \tau(k_L, \mu_H)$ for $\lambda_{t+1} \in [0.5, 0.567]$, $\tau_{t+1} = 0$ for $\lambda_{t+1} \in [0.567, 0.6]$, $\tau_{t+1} = \tau(k_H, \mu_L)$ for

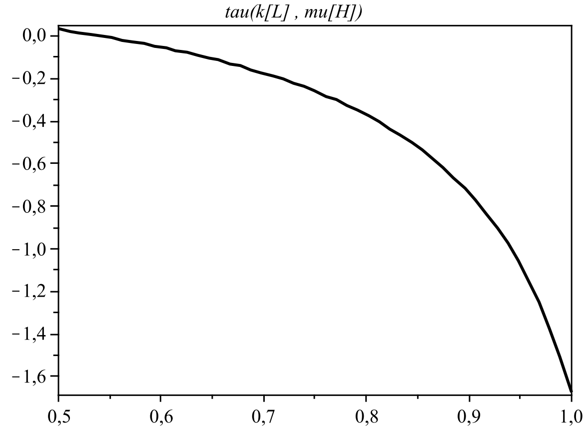

FIGURE 3.4. $\tau(k_L, \mu_{\sigma_H})$ for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $b = 1$, $\alpha = 0.2$.

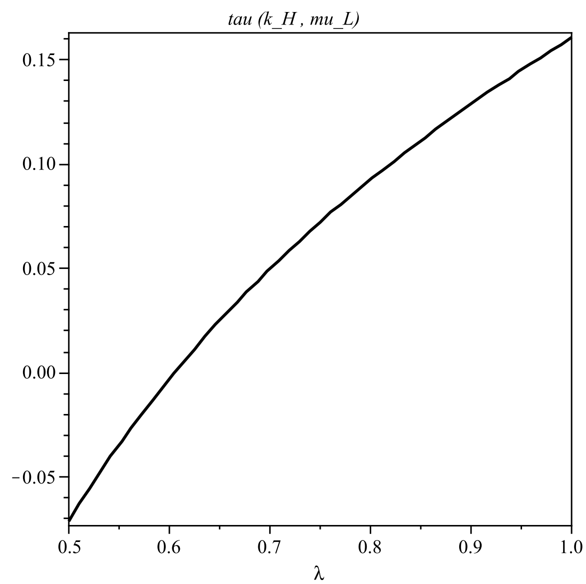

FIGURE 3.5. $\tau(k_H, \mu_{\sigma_L})$ for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $b = 1$, $\alpha = 0.2$.

$\lambda_{t+1} \in [0.6, 1]$. $\lambda_{t+1} = 1/2$ is a local maximum. [14] The two equilibria can be further characterized in terms of the other endogenous outcomes.

[14]Therefore with a linear cost of changing information $C(\lambda' - 1/2)$ greater than the slope of the line connecting the welfare function at $\lambda = 1/2$ and at $\lambda = 1$ it is an equilibrium, as shown in figure 3.6.

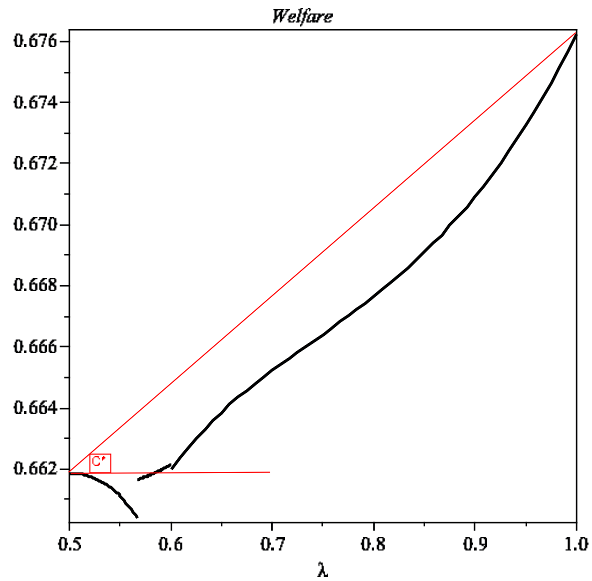

FIGURE 3.6. Objective function (3.43) for $\pi = 0.7$, $\theta_L = 1$, $\theta_H = 1.5$, $b = 1$, $\alpha = 0.2$.

Values of the endogenous variables in the equilibrium with perfect information.

Wealth of dynasty with low ability before taxes (I label it with the apex B):

$w_L^B = k_L + (1-\tau)\theta_L^2 = 0.965.$

Wealth of dynasty with high ability before taxes (I label it with the apex B):

$w_H^B = k_H + (1-\tau)\theta_H^2 = 2.080.$

Wealth of dynasty with low ability after taxes: $w_L = (1-\tau)(k_L + (1-\tau)\theta_L^2) + \tau(\bar{k} + (1-\tau)\bar{\theta}^2) = 1.047.$

Wealth of dynasty with high ability after taxes: $w_H = (1-\tau)(k_H + (1-\tau)\theta_H^2) + \tau(\bar{k} + (1-\tau)\bar{\theta}^2) = 1.891.$

Effort exerted by low ability individuals: $e_L = (1-\tau)\theta_L = 0.757.$

Effort exerted by high ability individuals: $e_H = (1-\tau)\theta_H = 1.135.$

Aggregate effort: $\bar{e} = (1-\tau)\bar{\theta} = 0.870.$

Aggregate output: $\bar{y} = \bar{k} + (1-\tau)\bar{\theta^2} = 1.300$.

Values of the endogenous variables in the equilibrium with imperfect information.

Wealth of dynasty with low ability before taxes (I label it with the apex B): $w_L^B = k_L + (1-\tau)\theta_L\bar{\theta} = 1.395$.

Wealth of dynasty with high ability before taxes (I label it with the apex B): $w_H^B = k_H + (1-\tau)\theta_H\bar{\theta} = 2.083$.

Wealth of dynasty with low ability after taxes: $w_L = (1-\tau)(k_L+(1-\tau)\theta_L\bar{\theta}) + \tau(\bar{k} + (1-\tau)\bar{\theta^2}) = 1.401$.

Wealth of dynasty with high ability after taxes: $w_H = (1-\tau)(k_H + (1-\tau)\theta_H\bar{\theta}) + \tau(\bar{k} + (1-\tau)\bar{\theta^2}) = 2.069$.

Effort exerted by both low and high ability individuals: $e_L = e_H = \bar{e} = (1-\tau)\bar{\theta} = 1.115$.

Aggregate output: $\bar{y} = \bar{k} + (1-\tau)\bar{\theta^2} = 1.603$.

Interpretation of the result of multiple equilibria. The numerical example shows the existence of two equilibria, respectively with complete ($\lambda = 1$) and with minimum ($\lambda = 1/2$) information. In this numerical example the multiplicity arises because $\lambda = 1$ happens to be a global maximum and $\lambda = 1/2$ is a local maximum.[15] In this numerical example the non-monotonicity of the objective function with respect to the level of information which implies the possibility that $\lambda = 1/2$ is a local maximum is driven by the change in the identity of the median voter. In the interval in which the median voter is the group with preferred tax rate $\tau(k_L, \mu_H)$ the tax rate decreases in λ. As figure 3.6 suggests, in this interval this effect dominates the sign of the derivative of the objective function (3.43) and implies that in such interval the objective function decreases. Information affects both τ and Γ. There could be other cases of non monotonicity. For example as it was the case in the previous chapter, in principle it can still be possible that over some interval $-\tau^2$ dominates the sign of the derivative and the objective function decreases in λ, whereas over other intervals Γ dominates the sign of the derivative and the objective function increases in λ.

3.6. Conclusion

This chapter developed a dynamic model with bequests, stochastic skills, endogenous voting and endogenous information. Given the technical issues involved, I proceeded by steps adding one element at the time. The intermediate steps showed results already known by the previous literature but offered a unified a framework. The result of the chapter is to have taken a first step towards the development of a unifying framework which

[15]Therefore for $\lambda = 1/2$ to be optimal it is necessary to introduce the additional assumption that there is a cost of increasing the level of information which is large enough. It is worthy to mention that in the model of Benabou and Tirole (2006) for multiple equilibria to exist the same assumption is necessary.

allows to study how beliefs about the determinants of wealth can affect the dynamics of inequality, mobility and redistribution.

Given the technical difficulties that stochastic skills imply in the determination of the median voter, I can only characterize equilibria with endogenous information for the case of persistent abilities over the life of dynasties ($\theta_t^i = \theta^i$). In this case, with a numerical example, I show the possibility of equilibria with complete ($\lambda = 1$) and with minimum ($\lambda = 1/2$) information.

This example essentially shows that societies with similar fundamentals can find optimal to be permanently "stuck"at different informative structures. Such different informative structures imply different steady states in terms of beliefs, redistribution, aggregate effort, aggregate output, separation of effort choices across individuals and inequality of wealth across individuals. Why, in this example, does it happen that a society remain stuck at a particular steady state? In terms of interpretations, the equilibria of the example describe a society which is characterized by a certain level of information long enough such that wealth is distributed so that for the majority group there is no ex-ante gain in changing the information structure. This happens because that particular level of information maximizes the welfare of the majority group. Given that in the uninformative equilibrium of the numerical example $\lambda = 1/2$ is a local maximum, the interpretation of such equilibrium should be that of a society for which the welfare gain in increasing the level of information would be off-set by some implied structural costs of doing so. As explained at the end of the previous section, it is important to notice that for this to be the case, a necessary condition is that the welfare function does not increase monotonically in the level of information.

The analysis of this chapter does not completely characterize the equilibria, in the sense that it does not generally show how the economic and redistributive outcomes of an equilibrium depend on the level of information. The numerical example only shows the possibility of equilibria with complete ($\lambda = 1$) and with minimum ($\lambda = 1/2$) information. Comparing the endogenous variables in the two equilibria of the example, the equilibrium with $\lambda = 1$ is characterized by relatively higher redistribution, greater inequality before taxes, greater inequality after taxes, more separated levels of effort, lower aggregate effort and lower aggregate output. Other numerical examples that I tried with different values of the parameters did not seem to change these features of the two steady states. As it was the case with the model of the previous chapter, the informative (uninformative) equilibrium presents Europe-type (US-type) features in terms of redistribution, aggregate effort and output. At the same time, the feature that the Europe-type equilibrium is characterized by relatively higher inequality is not entirely satisfactory because it contradicts some of the empirical evidence.[16] This feature that the more informative equilibrium is characterized

[16]See the discussion in section 1.1.2 of chapter 1.

by higher inequality is not so surprising given that in the example the uninformative equilibrium is characterized by minimum information and levels of effort which are identical across different types of agents. One would expect such pooling equilibrium to be associated with low inequality. In principle it does not have to be the case that more informative equilibria are always associated with higher inequality. The opposite could happen if in the more informative equilibrium the higher rate of redistribution distorts effort so much that the levels of exerted effort are less separated, despite the fact that beliefs are more separated. Interestingly enough given the already mentioned empirical evidence about redistribution and inequality, in such a case the driving force behind higher taxation would not be the actual level of inequality, even though this would still impact on the ideal tax rate, but the beliefs about the determinants of wealth. Moreover, in such a 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.[17]

In order to verify this possibility it is necessary to look for equilibria with interior solutions in terms of information because the extreme cases do not seem to give that result. Even in the case of persistent abilities, with interior solutions the identification of the median voter is problematic. Allowing for the case in which abilities evolve stochastically across generations could give interesting results about the study of mobility and inequality, because it would be possible to characterize how different levels

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

of information impact on effort decisions and taxation and obtain interesting comparative statics. Therefore, given the implied technical issues in the identification of the median voter, it seems a promising direction for future research to investigate those issues with numerical methods.

Conclusions of the Thesis

This thesis aimed to contribute to the understanding of the role of collective beliefs and incomplete information in the analysis of inequality, growth and redistributive politics. More specifically it aimed to offer insights to answer open questions relating societal beliefs and the level of redistribution:
(i) Why is the case that very different patterns of redistribution and social contracts are found across countries which are otherwise similar?
(ii) Why is it the case that also the beliefs which people hold about the underlying determinants of wealth and the extent of social mobility are very different across otherwise similar societies?
(iii) It appears that the societies which are characterized by the widespread belief according to which "hard work, self-discipline and other factors under individual control more than luck, family of origin or other factors outside individual control determine individual wealth"(e.g. the US) redistribute less than the societies which are characterized by the opposite belief (e.g. European countries). Is it possible to describe those outcomes relating beliefs, political and economic outcomes in terms of different equilibria? And which are the driving forces behind those different outcomes?

Presenting an extensive review of various strands of literature, chapter 1 motivated why such questions are of great relevance for the understanding of macroeconomic problems. Simplifying the argument to the core, the existing literature shows the great importance of studying the determinants of redistributive politics as the latter have a major impact on inequality and growth. In addition, beliefs seem to influence to a great extent the shape of redistributive politics and therefore ignoring beliefs means to forget about an important channel. Models which can originate multiple equilibria seem to be particularly insightful as they are able to rationalize the observed economic and political outcomes as endogenous outcomes of societies which are otherwise equal. The theoretical contributions of Piketty (1995), Alesina and Angeletos (2005) and Benabou and Tirole (2006) have been the first to develop insightful models describing how individual beliefs can shape politico-economic outcomes and viceversa and how multiple equilibria (US-type vs Europe-type) are possible. These models offer different explanations for the observed outcomes but they all share the feature that economic agents have incomplete information about the underlying determinants of wealth – i.e. about the role of factors under individual control versus the role of factors outside individual control. Chapter 2 develops a theoretical model which shares the same underlying features of these models, but focuses on the sole role of incomplete information – without adding other elements as psychological factors – and considers the level of incomplete information in the economy as an institutional feature

that policies can in principle change. The model can generate multiple equilibria with US-type versus Europe-type features. One main contribution is therefore to show that incomplete information is a driving force behind the existence of multiple equilibria. Moreover it offers a different economic interpretation of the two equilibria which points to the existence of different informative cultures and which is supported by evidence from the literature in education and sociology. Specifically it shows that otherwise similar societies can find optimal to pearmain at different levels of information, where relatively more (less) separating informative structures should are associated to higher (lower) redistribution and lower aggregate output and effort. My model also wanted to address specific policy questions:
(iv) Given that beliefs and incomplete information appear to play a crucial role in the determination of political and economy outcomes, which are the relative policy implications? Is there any role for institutions which can affect the degree of information in the economy?

This question could not be answered by the existing models. In my model information is instead considered to be an institutional feature and it is possible to change the degree of incomplete information in a continuous way. In this way I offer various results in terms of comparative statics of how individual and aggregate political and economic outcomes change with degree of incomplete information.

Chapter 3 develops a dynamic version of the model of chapter 2, introducing the same information set-up in an intergenerational economy with bequests and voting. In this way, the chapter develops a dynamic model with bequests, stochastic skills, endogenous voting and endogenous information. The model shows that also in a dynamic setting there can be multiple optimal welfare maximizing levels of information, and therefore otherwise similar societies can remain "stuck"in steady states with different levels of information, where relatively more (less) separating informative structures should are accompanied by higher (lower) redistribution and lower aggregate output and effort. Given some technical issues which I extensively discuss, the model is not able to give a general characterization of the steady states, nevertheless this could be done by numerical methods and seems to be a promising direction for the analysis of the dynamics of inequality and intergenerational mobility.

Future research should continue the analysis in the following directions. My theoretical framework interprets different political and economic outcomes in otherwise similar societies as driven by different informative cultures/structues. Future empirical research should characterize such different informative features. At the theoretical level, it seems interesting to link my rather abstract information set up with theories focused on how political groups or other institutions can influence information and beliefs through education-financing[18], propaganda[19] or mass media[20]. The idea that separation across individuals impacts their information and beliefs

[18]See references in footnote 10, chapter 1.
[19]See for example Alesina, Glaeser, and Sacerdote (2001).
[20]See for example Prat and Strömberg (2005)

seems also quite related to the models which analyze the role of segregation or discrimination on inequality.[21] Another natural link seems to be with theoretical models which study the role of aspirations in the dynamics of poverty[22] as such concepts are strongly related to the impact of beliefs about the determinants of wealth on effort choices. Finally, as already pointed out, numerical methods seem to be a promising direction for the analysis of comparative statics and optimal information policies in dynamic settings with intergenerational inequality and mobility.

---

[21] See references in sections 5 and 6 of Piketty (1998) and references in footnote 11, chapter 1.

[22] See for example the review article of Banerjee (2006) and the models of Heifetz and Minelli (2006) and Dalton and Ghosal (2008)

APPENDIX A

Appendix to Chapter 2

A.1. Proof of lemma 2.2

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

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

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:

(A.1)

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

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.■

A.2. Proof of proposition 2.4

It is useful to plug (2.19) into (2.23) and re-express this as

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

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

$$\frac{2\Gamma^2 \frac{\partial \Gamma}{\partial \lambda} - 2\theta(\mu_{\sigma_L})^2 \Gamma \frac{\partial \Gamma}{\partial \lambda} + 2\theta(\mu_{\sigma_L})\Gamma^2 \frac{\partial \theta(\mu_{\sigma_L})}{\partial \lambda}}{a^2 \left(2\Gamma - \theta(\mu_{\sigma_L})^2\right)^2} \tag{A.3}$$

where

$$\begin{aligned} \frac{\partial \theta(\mu_{\sigma_L})}{\partial \lambda} &= -\frac{\pi(1-\pi)(\theta_H - \theta_L)}{(\pi\lambda + (1-\lambda)(1-\pi))^2} \leq 0 \\ \frac{\partial \Gamma}{\partial \lambda} &= \frac{\pi^2(1-\pi)^2(2\lambda-1)(\theta_H-\theta_L)^2}{(\pi\lambda+(1-\pi)(1-\lambda))^2(\pi(\lambda-1)+\lambda(\pi-1))^2} \geq 0. \end{aligned} \tag{A.4}$$

The denominator of (A.3) 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

$$(\Gamma - \theta(\mu_{\sigma_L})^2)\frac{\partial \Gamma}{\partial \lambda} + \theta(\mu_{\sigma_L})\Gamma\frac{\partial \theta(\mu_{\sigma_L})}{\partial \lambda}. \tag{A.5}$$

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 (A.3) is negative for $\lambda = 1/2$. I compute the second derivative of (A.5):

(A.6)

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

where

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

(A.7)

Notice that $\frac{\partial^2 \Gamma}{(\partial \lambda)^2} \geq 0$ as it has already been proved in Appendix A.1.

Given the signs of $d\theta(\mu_{\sigma_L})$, $d^2\theta(\mu_{\sigma_L})$, $d\Gamma$, $d^2\Gamma$ and the fact that $\Gamma - \theta(\mu_{\sigma_L})$ is positive in the range considered, (A.5) is strictly positive and therefore (A.3) can change sign at most once in the range $\lambda \in [1/2, 1]$. Therefore in the range $\lambda \in [1/2, 1]$ (A.3) 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:

$$(\pi\theta_L + (1-\pi)\theta_H)^2(2\pi\theta_L^2 + 2(1-\pi)\theta_H^2 - \theta_L^2) - (\pi\theta_L^2 + (1-\pi)\theta_H^2)^2 \geq 0 \tag{A.8}$$

i.e.

$$(\theta_L - \theta_H)(-1+\pi)\left(-2\,\theta_H{}^2\pi^2\theta_L + 2\,\theta_H{}^3\pi^2 - 2\,\theta_H\pi^2\theta_L{}^2 + 2\,\pi^2\theta_L{}^3 - 3\,\theta_H{}^3\pi + \pi\,\theta_L\theta_H{}^2 + 2\,\theta_H\pi\,\theta_L{}^2 + \theta_H{}^3 + \theta_L\theta_H{}^2\right) \geq 0 \tag{A.9}$$

Observe that

$$-2\,\theta_H{}^2\pi^2\theta_L + 2\,\theta_H{}^3\pi^2 - 2\,\theta_H\pi^2\theta_L{}^2 2\,\pi^2\theta_L{}^3 - 3\,\theta_H{}^3\pi + \pi\,\theta_L\theta_H{}^2 + 2\,\theta_H\pi\,\theta_L{}^2 + \theta_H{}^3 + \theta_L\theta_H{}^2 = \tag{A.10}$$

(A.11)

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

Observe that

$$\text{(A.12)}\quad \left(-2\,\pi^2+\pi+1\right){\theta_H}^2\theta_L + \left(2\,\pi^2+1-3\,\pi\right){\theta_H}^3 = (1-\pi)\theta_H^2(2\pi\theta_L - 2\pi\theta_H + \theta_L + \theta_H).$$

Hence, after a factorization condition (A.8) can be rewritten as

$$\text{(A.13)}\quad (\theta_L-\theta_H)\,(-1+\pi)\left(2\,\pi^2{\theta_L}^3 + 2\,\pi(1-\pi)\theta_H{\theta_L}^2 + (1-\pi)\theta_H^2(2\pi\theta_L - 2\pi\theta_H + \theta_L + \theta_H\right),$$

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 (A.8) is satisfied.■

APPENDIX B

Non-Existence of Competitive Equilibria with Dynamic Inconsistent Preferences

This chapter represents an extended version of a joint mimeograph with S. Ghosal which go under the same title.

B.1. Introduction

This paper examines the existence of competitive equilibria in dynamic general equilibrium models when agents have dynamically inconsistent preferences. Such preferences imply that the individual optimal consumption-saving plan – i.e. the inter-temporal marginal rate of substitution between future dates – changes over time.

For given individual preferences (whether consistent or not) and endowments, a competitive equilibrium of an exchange economy is defined as the prices and the allocations such that (i) allocations are optimal and feasible and (ii) markets clear. Time inconsistent preferences imply two problems for the existence of competitive equilibria. The first problem is represented by the existence of an optimal solution to the individual consumption-saving problem, given that the optimal plan changes over time. The second problem lies in the market clearing requirement aspect, in other words to find prices such that optimal choices clear the markets.

As we will show, given that the optimal consumption-saving solution does not present standard features this is not a standard existence problem.

Most of the literature has only dealt with the first aspect. The solution to the choice problem of a time-inconsistent consumer is the object of an early literature: Strotz (1956), Pollak (1968), Blackorby, Nissen, Primont, and Russell (1973), Peleg and Yaari (1973), Goldman (1980) are commonly cited as the major contributions. Strotz (1956) is the seminal paper that introduces the idea of time-inconsistent preferences and proposes a first solution to the problem of a time-inconsistent consumer. Doing this Strotz (1956) distinguishes between *naive* and *sophisticated* decision makers, i.e. between a consumer who is respectively not aware or aware of the fact that her preferences change over time. Pollak (1968) corrects part of the analysis of Strotz (1956) and introduces the idea of *intra-personal* game. Given that the same consumer has different preferences in different periods, the same consumer is considered as a different decision maker in each period. If the consumer is aware that her preferences will change, the best she can do is to maximize the present utility taking as given the optimal choices of her future self. The solution of this *intra-personal* game is the optimal and consistent consumption plan of a consumer whose preferences change over time. As pointed out by Peleg and Yaari (1973), such an intra-personal game need not to have a Markov perfect equilibrium, but as introduced by Goldman (1980), the optimal consumption path can be found as a Subgame Perfect Nash Equilibrium (SPNE) which is proved to exist.

The interest in time-inconsistent preferences is resurrected by Laibson (1997). Motivated by evidence from psychology and introspection, Laibson and many others consider a special case of time-inconsistent preferences, the quasi-hyperbolic discounting and apply it to a number of problems. There is a now a large literature which applied time inconsistent preferences (not only quasi-hyperbolic) to different topics. In terms of existence, most of those papers take a quasi-equilibrium approach in other words they only investigate the consumer solution without worrying about the market side. Nevertheless, even the existence of the inconsistent consumer solution is not a closed matter and there are recent papers which consider the issue in more complicated set ups.[1]

The second aspect, which is the market aspect, in other words to find prices such that consumer choices clear the markets has been much less discussed. Luttmer and Mariotti (2003), Kocherlakota (2001), Barro (1999) are equilibrium models where quasi-hyperbolic and time separability guarantees homoteticity of preferences, they can characterize equilibria easily and existence is not an issue. More related to our work, for special classes of preferences including quasi-hyperbolic discounting, existing work by Luttmer and Mariotti (2007), Luttmer and Mariotti (2006) and Herings and Rohde (2006) have shown that equilibria exist. In contrast, in this paper we

[1]See for example Harris and Laibson (2001), Caplin and Leahy (2006) and Ekeland and Lazrak (2008).

show, by example, robust non-existence of equilibria. Our set up is a representative agent deterministic version of the seminal model of Lucas (1978) which is equivalent to theirs. [2]

In this note we show through an example that competitive equilibria fail to exist because time-inconsistent preferences induce both non-convexity of demand correspondences and satiation. Especially the presence of satiation represents a new issue which has not been discussed by the previous literature. Those two features combined together imply that non existence is robust, in other words methods that are able to re-establish existence in the presence of non convexities or satiation alone and have been used by the previous literature do not assure existence with a general class of time inconsistent preferences.

Following the previous literature, also in our set-up given the time inconsistent preferences the described agent plays an intra-personal game and maximize present utility taking as given future choices. Just as in the case of time-consistent preferences, this future behavior can be summarized by a value function defined over wealth saved by the consumer at the initial date. However, because of time- inconsistency, the value function which is induced by the intra-personal game need not be concave, even if the underlying period utility functions are. This is because the value function which is induced by the intra-personal game is affected by the best responses of the future versions of the same consumer, in other words the future optimal

[2]We will clarify precisely what is the role played by the assumption of a single representative agent.

consumption choices as functions of the wealth saved in the present. Without specific assumptions, even if the underlying period utility functions are monotone and concave the shape of the future best responses can imply that the present value function is not concave. As a result, the demand correspondences of individual consumers may not be convex-valued. This non convexity is the first problem that time-inconsistent preferences induce for the existence of a competitive equilibrium. Such non-convexity has already been noticed by many. Herings and Rohde (2006) do not deal with the possibility of non-convexities as they impose by assumption that the preferences induced by the intra-personal game are convex and prove existence of a competitive equilibrium for this case. Luttmer and Mariotti (2007), Luttmer and Mariotti (2006) allow for induced non-convex preferences but prove existence allowing for a large number of consumers. With one agent non-convexity may imply non existence because it may imply that demand correspondences are non continuous it does not always a price vector such that for the consumer is optimal to demand the market clearing bundle. Nevertheless assuming a large economy, it is a known result[3] that existence of equilibrium is re-established despite of individual non-convex preferences. The reason is that with a large economy, for an equilibrium to exist is enough to find a price such that the market clearing bundle is in the convex hull of the demand function and such a price vector can indeed

[3]See for example Mas-Colell, Whinston, and Green (1995) chapter 17.

be found as a fixed point. Therefore also in a case with a single representative agent like the one that we examine, induced non-convexity alone does not imply non existence if the representative agent stands for a large number of identical consumers. The second problem that time-inconsistent preferences induce for the existence of a competitive equilibrium is represented by satiation. This issue has not been discussed by the previous literature. We show that even if the underlying period utility functions are monotone and concave the shape of the future best responses can imply that the present value function is characterized by satiation points. In the example that we construct this happens because we allow for specific non time-separable preferences such that some goods are normal from the perspective of today but become inferior in the future. For this reason, the consumer can have a satiation point in the level of savings (i.e. future wealth), because due to the presence of inferior goods higher savings today can imply a level of future consumption which is too low given the preferences of today. In our example such induced satiation implies non-existence of a competitive equilibrium because the market clearing quantity of some good is greater than the satiation quantity for all positive vectors of prices. In other words – as the specifics of the example will show – in our case the value function has a special non convex behavior which implies that there is no price such that the market clearing quantity of some good lies in the individual demand function or not even in the its convex hull. Therefore in this example re-convexification through a large economy does not re-establishes existence. Free disposal and the possibility of negative prices are standard ways to get around to deal with satiation in existence problems.[4] Nevertheless, our set up consists of a pure exchange economy where the only way to transfer wealth across periods is to demand assets and free disposal does not apply to such asset market. When free disposal fails negative prices could in principle help, but it is not the case in this set-up because if asset prices go negative then demand grows to infinity. Technically, the example that we propose is very simple. It is a three periods economy with a representative agent and no uncertainty about prices and dividends. The only technical challenge is represented by the need of a well behaved utility function which allows for inferior goods. We find such a utility function introducing a class of preferences known as Addilog preferences which have been introduced by Houthakker (1960).

B.2. The Economy

We consider a simple "Robinson Crusoe"economy over three periods; each period is labeled by t and $t \in \{1, 2, 3\}$. The economy is populated by one representative agent and there is a unique asset (the tree) which delivers units of a consumption good (dividends or fruit) in every period. In each period t the agent maximizes a life-time utility function choosing present consumption (c_t) and a non negative[5] fraction of the asset (θ_{t+1}) to be own in the next period. The consumption good is non storable, hence

[4]See the classic references on satiation: Bergstrom (1976), Hart and Kuhn (1975), Nielsen (1990), Polemarchakis and Siconolfi (1993) among others .

[5]In this set-up short selling on the asset is not allowed.

the asset provides the only way to transfer wealth across periods. The ownership of a fraction θ_{t+1} of the asset at the beginning of period $t+1$ yields $\theta_{t+1}d_{t+1}$ units of the consumption good in $t+1$. The agent decides the fraction of asset θ_{t+1} to be carried to the next period at the end of each period t, after consumption takes place. The consumption/investment decisions takes place on competitive markets for the consumption good and for the shares in the asset. We consider that prices are normalized such that the price of one unit of the consumption good is fixed to 1 in each period, hence the (real) price of the asset is p_t in each period. The model is completely deterministic: the values of all the prices and the dividends are known from the beginning by the agent. At the beginning of period 1, the agent is endowed with the entire asset ($\theta_1 = 1$) and the entire paid dividend d_1.

Assumption B.1: In each period t, the utility (objective) function $u_t(c_t, ...c_3)$ is defined on all non negative present and future consumption sets and is strictly increasing, strictly concave, and twice continuously differentiable.

We assume that preferences change inconsistently between period 1 and 2. We formally define time-inconsistency as follows.

DEFINITION B.1. **Time Inconsistent Preferences**. Preferences are **Time Inconsistent** if it does not exist an integrable map $f'() > 0$ such that

proj $u_1(c_1, c_2, c_3)$ on $(c_2, c_3) \in \mathbb{R}^2_+ = f(u_2(c_2, c_3))$.

Time-inconsistency implies that optimal consumption savings plan change between period 1 and period 2 as the optimal inter-temporal rates of substitution change. There is no time-inconsistency between period 2 and period 3, because period 3 is the last and preferences are monotonic and therefore, from the perspective of every period the optimal period 3 choice is to consume the entire endowment. Given time-inconsistent preferences, following Pollak (1968), Laibson (1997) and many others we define two possible solutions to the consumer problem.

DEFINITION B.2. **Naive Solution (NS)**. A Naive Solution consists in a vector of consumption-saving choices $(c_1^*, \theta_2^*, c_2^*, \theta_3^*, c_3^*)$ such that

$$(c_1^*, \theta_2^*) = \text{arg max } u_1(c_1, c_2, c_3),$$

$$\text{subject to:} \tag{B.1}$$

$$c_1 + p_1\theta_2 \leq p_1 + d_1.$$

$$(c_2^*, \theta_3^*) = \text{arg max } u_2(c_2, c_3),$$

$$\text{subject to:} \tag{B.2}$$

$$c_2 + p_2\theta_3 \leq (p_2 + d_2)\theta_2^*.$$

$$(c_3^*) = \text{arg max } u_3(c_3),$$

(B.3) $$\text{subject to:}$$

$$c_3 \leq d_3\theta_3^*.$$

DEFINITION B.3. **Sophisticated Solution (SS)**. A Sophisticated Solution consists in a vector of consumption-saving choices $(c_1^*, \theta_2^*, c_2^*, \theta_3^*, c_3^*)$ such that

$$(c_3^*) = \text{arg max } u_3(c_3),$$

(B.4) $$\text{subject to:}$$

$$c_3 \leq d_3\theta^{3\cdot}$$

$$(c_2^*, \theta_3^*) = \text{arg max } u_2(c_2, c_3^*),$$

(B.5) $$\text{subject to:}$$

$$c_2 + p_2\theta_3 \leq (p_2 + d_2)\theta^{2\cdot}$$

$$(c_1^*, \theta_2^*) = \text{arg max } u_1(c_1, c_2^*, c_3^*),$$

(B.6) $$\text{subject to:}$$

$$c_1 + p_1\theta_2 \leq p_1 + d_1.$$

Notice that the assumption that every period utility function is strictly monotone in consumption implies that both the **NS** and the **SS** imply that all the three inter-temporal budget constraints are satisfied with the equality. The market clearing condition for this economy is trivial: the agent must hold the entire unit of the asset in each period ($\theta_1 = \theta_2 = \theta_3 = 1$) and consumption must be equal to the entire paid dividend in each period ($c_1 = d_1, c_2 = d_2, c_3 = d_3$). It follows the definition of competitive equilibrium for the economy.

DEFINITION B.4. **Competitive Equilibrium**. A competitive equilibrium with Naive (Sophisticated) agents for the economy is given by prices (p_1^*, p_2^*) and allocations $(\theta_1^*, c_1^*, \theta_2^*, c_2^*, \theta_3^*, c_3^*)$ such that:

(i) $(\theta_1^*, c_1^*, \theta_2^*, c_2^*, \theta_3^*, c_3^*)$ is **NS** (**SS**).

(ii) $(c_1^* = d_1, \theta_2^* = 1, c_2^* = d_2, \theta_3^* = 1, c_3^* = d_3)$.

PROPOSITION B.1. *A Competitive Equilibrium with naive agents does exist if and only if markets are allowed to re-open.*

PROOF. In each period the objective function is concave. For the Separating Hyperplane Theorem it is possible to find prices which support the unique market clearing bundle in every period. Time-inconsistent preferences imply that inter-temporal marginal rate of substitutions change from period to period, therefore also the relative price which sustains a competitive equilibrium changes from one period to the other. □

PROPOSITION B.2. *A Competitive Equilibrium with Sophisticated Agents does not always exist.*

We prove the proposition with an example

B.3. An Example of Non Existence

In order to simplify the computations we fix $d_1 = d_2 = d_3 = 1$. We specify the utiliy functions for each period.
First period utility function:

$$U_1(c_1, c_2, c_3) = a\ln(c_1) + b\ln(c_2) + c\ln(c_3), \tag{B.7}$$

where a, b, c are strictly positive and smaller than 1.
We present the second period utility function through its indirect form:

$$V_2(p_2, w_2) = \alpha_2 \frac{(w_2/p_{c_2})^{\beta_2}}{\beta_2} + \alpha_3 \frac{(w_2/p_2)^{\beta_3}}{\beta_3}, \tag{B.8}$$

where w_2 is the period 2 wealth ($\theta_2(p_2 + d_2) = w_2$) and p_{c_2} is the price of c_2 which is normalized to $p_{c_2} = 1$. Expression (B.8) constitutes an example of Indirect Addilog Utility Function. This class of indirect utility functions has been introduced by Houthakker (1960). Murthy (1982) shows that the general formulation can be written as follows: $V(p, w) = \Sigma_{i=1}^{n} \frac{\alpha_i (w/p_i)^{\beta_i}}{\beta_i}$, where w/p_i is the expenditure share for commodity i. de Boer, Bröcker, Jensen, and van Daal (2006) formally prove that for appropriate values of α's and

β's the correspondent direct utility is well behaved, although not analytically specified; the paper also contains an historical review of the Indirect Addilog System. For appropriate values of the parameters β_2 and β_3, the use of the Indirect Addilog Utility Function allows us to impose that either c_2 or c_3 is an inferior good for the period 2 consumer, while assuring that the direct form satisfies the properties of assumption B.1. We impose the following values of the parameters:

(B.9)

$$\beta_2 = 1, \beta_3 = -0.5, \alpha_2 = .6297714880, \alpha_3 = 1 - .6297714880, d_1 = d_2 = d_3 = 1.$$

Given the value of the parameters, it is immediate to verify that (B.8) presents all the regularity properties[6]:

(i) homogeneous of degree zero in prices p_2, p_{c_2} and income w_2.

(ii) non-increasing in prices and nondecreasing in income.

(iii) quasi-convex in prices.

(iv) differentiable in all positive prices and in positive income.

The fact that the indirect utility function is strictly convex in prices implies that the direct utility function, i.e. the dual of (B.8), is strictly quasi-concave by a well known result in duality theory (Mas-Colell et al., page 66). Therefore it satisfies assumption 1.

Period 3 utility function: any $u(c_3)$ satisfying assumption 1.

[6]De Boer et al 2006 show that the regularity properties are satisfied for $-1 < \beta_i < 1$. The β's are parameters which affect the urgency of consumption and allow for the possibility of inferior goods. The α's are conveniently fixed to sum to one in order to prove the regularity properties.

LEMMA B.1. *A sophisticated competitive equilibrium does not exist if the* ***SS*** *implies the following conditions:*

(i) There is a unique p_2^* *such that* $\theta_2^* = \theta_3^* = 1$.

(ii) For any $p_1 \geq 0$, $\frac{\partial V_1(p_1,p_2^*)}{\partial \theta_2}|_{\theta_2 \geq 1} < 0$

(iii) For any $p_1 < 0$, $\lim_{\theta_2 \to +\infty} V_1(p_1, p_2^*) = +\infty$.

PROOF. Condition (i) implies that with sophisticated agents there is a unique p_2^* candidate equilibrium price at period 2. For an equilibrium to exist, given p_2^*, there must be a p_1^* such that for the representative agent $\theta_2^* = 1$ is a **SS**. Condition (ii) implies that there is no positive p_1 such that $\theta_2 = 1$ is a solution to (B.6) and therefore cannot be part of the individual **SS**. In the case in which we allow for a large economy and the representative agent “stands in”for a large number of identical consumers, for a competitive equilibrium to exist it would be sufficient to find a p_1 such that $\theta_2 = 1$ belongs to the convex hull of the individual demand function. In this case, the market clearing condition would be satisfied assigning appropriate fractions of the consumers to different bundles of the demand function (see Mas-Colell et al. Chap. 17). Lotteries can achieve the same results in a decentralized way. For this to be the case it is necessary that there exists some value $\theta_2 > 1$ which is a solution to (B.6), but this is excluded by condition (ii). Condition (iii) implies that not even a negative period 1 price p_1 can assure the existence of an equilibrium, because if prices are negative the unique solution to (B.6) consists in choosing $\theta_2 \to +\infty$. □

PROPOSITION B.3. *Given the period 1 utility (B.7) and the period 2 utility (B.8), the* ***SS*** *implies that all the conditions of lemma B.1 are satisfied.*

Proof of condition (i).

We first compute the period 2 demand functions. Given that the period 2 utility function is quasi-concave, we can apply Roy's Lemma and obtain:

$$c_2^* = -\frac{dV_2(\cdot)}{dp_{c_2}} \Big/ \frac{dV_2(\cdot)}{dw_2} = \frac{\alpha_2(w_2)^{\beta_2+1}}{\alpha_2(w_2)^{\beta_2} + \alpha_3(w_2/p_2)^{\beta_3}}. \tag{B.10}$$

Given c_2^* so obtained, θ_3^* is immediately found through the period 2 budget constraint satisfied with the equality, hence

$$\theta_3^* = \frac{w_2 - c_2}{p_2}. \tag{B.11}$$

Given that the period 2 objective function is strictly quasi-concave and the fact that at the period 2 equilibrium price p_2^* it must be optimal for the period 2 consumer to demand $\theta_3^* = 1$, the period 2 equilibrium price is positive and unique. Given that by the definition B.4 follows that at the equilibrium price vector it must be optimal for the consumer to demand $\theta_2^* = \theta_3^* = 1$, the period 2 equilibrium price p_2^* is that price such that $\theta_3^* = 1$ is optimal in period 2 given that $\theta_2^* = 1$ has been chosen in period 1. Hence the equilibrium period 2 price p_2^* can be computed by imposing the market clearing conditions $c_2 = d_2 = 1$ and $\theta_2 = 1$ (which implies $w_2 = p_2 + 1$) in (B.10) and then solving for p_2. Given the specified values of β_2 and β_3, we obtain the following identity:

$$(p_2^* + 1)^3 p_2^* = (\alpha_3/\alpha_2)^2. \tag{B.12}$$

Plugging the specified values of α_2 and α_3 in (B.12), we can compute that there exists only one real positive solution to (B.12), namely $p_2^* = 0.2$ and this is the period 2 equilibrium price. □

Proof of condition (ii).

Given the **SS** θ_3^* can be expressed as functions of θ_2 and can be interpreted as the reaction functions of the period 2 consumer to the choice of the period 1 consumer. Plugging (B.10) in (B.11) and considering the expression which we obtain for the specified values of β_2, β_3, d_2 and for $p_2 = p_2^*$, we find the demand of θ_3 as a function of θ_2, namely

$$\theta_3(\theta_2) = \frac{.9068709427\sqrt{\theta_2}}{.75572578856\theta_2 + .1511451571/\sqrt{\theta_2}}. \tag{B.13}$$

It is important to notice that for values of $\theta_2 \geq 1$, θ_3 decreases in θ_2. This can be shown immediately by computing the first derivative of B.13: $\theta_3'(\theta_2) = \frac{-\frac{xy}{2}\theta_2^{3/2}+xz}{y\theta_2^{3/2}+z}$, where $x = .9068709427$, $y = .7557257856$, $z = .1511451571$. A sufficient condition for this expression to be strictly negative when θ_2 is $2z < y$ as it is indeed the case. Hence, θ_3 and c_3 are inferior commodities for the period 2 consumer, over some range of his income.

Re-expressing c_1, c_2 and c_3 through the three inter-temporal budget constraints satisfied with the equality, we obtain the period 1 indirect utility function:

(B.14) $$V_1 = a\ln(p_1 + d_1 - p_1\theta_2) + b\ln((p_2 + d_2)\theta_2 - p_2\theta_3(\theta_2)) + c\ln(d_3\theta_3(\theta_2)).$$

Plugging (B.13) and the specified values $d_1 = d_2 = 1$, $p_2^* = 0.2$ into (B.14), we obtain the period 1 indirect utility as a function of p_1 and θ_2

(B.15)
$$\begin{aligned} V_1 =& a\ln(p_1 + 1 - p_1\theta_2) + b\ln(1.2\theta_2 - \frac{.1813741885\sqrt{\theta_2}}{.75572578856\theta_2 + .1511451571/\sqrt{\theta_2}}) + \\ & c\ln(\frac{.9068709427\sqrt{\theta_2}}{.75572578856\theta_2 + .1511451571/\sqrt{\theta_2}}). \end{aligned}$$

In order to study the sign of the first derivative of (B.15) with respect to θ_2, we rename the addends of (B.15) as follows:

$a\ln(p_1 + 1 - p_1\theta_2) \equiv A$,

$b\ln(1.2\theta_2 - \frac{.18137418885\theta_2}{.75572578856\theta_2^{3/2} + .1511451571}) \equiv b\ln(k\theta_2 - \frac{x\theta_2}{y\theta_2^{3/2} + z}) \equiv B$, where $k \equiv 1.2$, $x \equiv .1813741885$, $y \equiv .7557257856$, $z \equiv .1511451571$.

$c\ln(\frac{.90687094427\theta_2}{.75572578856\theta_2^{3/2} + .1511451571}) \equiv c\ln(\frac{h\theta_2}{y\theta_2^{3/2} + z}) \equiv C$, where $h \equiv .9068709427$, $y \equiv .7557257856$, $z \equiv .1511451571$.

It is immediate to compute that $\frac{\partial A}{\partial \theta_2} \geq 0$, for all θ_2 and non negative values of p_1. Computing the derivative of $B + C$, we obtain the following condition on c and b for it to be strictly negative:

$$\frac{d(B+C)}{d\theta_2} < 0 \text{ iff } \frac{c}{b} > \frac{\theta_2^4 ky^2 + \theta_2^{5/2}(2kyz + xy) + \theta_2(kz^2 - xz)}{\theta_2^4 ky + \theta_2^{5/2}(kyz - xy) + \theta_2(xz - kz^2 - kz)}.$$

The condition above can be satisfied by appropriate finite values of b and c for all $\theta_2 \geq 1$, notice that asymptotic results imply that the right hand side approach zero for θ_2 that goes to $+\infty$. $\square$

Proof of condition (iii).

Using the definitions of A, B, C it is immediate to notice that asymptotic results imply that a sufficient condition for condition (iii) to be satisfied is $a > c$. $\square$

B.4. Remarks

In order to facilitate the comprehension of the example, I plot in figure B.1 V_1 as a function of θ_2, for $p_1 = 0.1, a = 0.7, b = 0.1, c = 0.7$, over the entire positive domain. Figure B.2 zooms the same function in a neighborhood of 1. It is clear from those two figures that there is a satiation quantity of θ_2 which is smaller than one. Increasing the quantity of θ_2 above this value strictly decreases utility. I had to plot V_1 for a given value of p_1 but changing this value in the positive domain would not change the monotonically increasing behavior, given values of b and c satisfying the condition in the proof of condition (ii) of proposition B.3.

Figure B.1 shows a period 1 objective function which is not concave throughout the entire domain; this implies a non convexity in the period 1

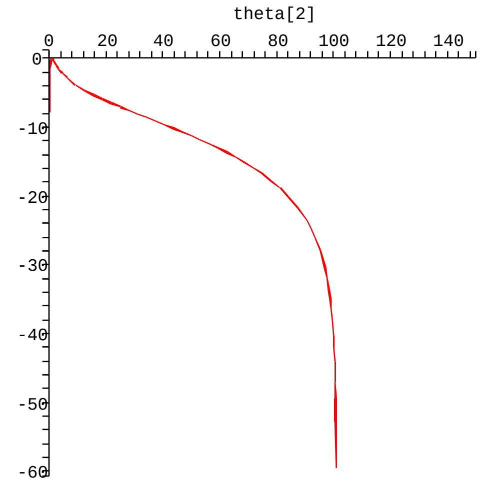

FIGURE B.1. $V_1(\theta_2)$, given $p_1 = 0.1, a = 7, b = 1, c = 7$.

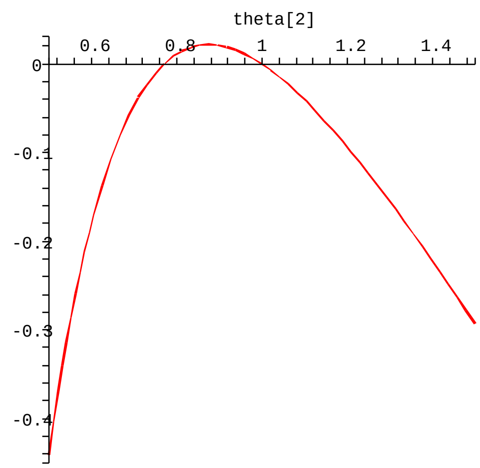

FIGURE B.2. $V_1(\theta_2)$, given $p_1 = 0.1, a = 7, b = 1, c = 7$.

preferences. It can be easily shown that the possible non-convexity of the period 1 preferences, which would not be found in the standard problem with time consistent preferences, originates from the fact that the objective function of the period 1 consumer is induced by the backward induction solution, through the reaction function $\theta_3(\theta_2)$. The reaction function $\theta_3(\theta_2)$

may fail to be concave over the entire domain and this can imply, as it happens in our example, that the period 1 objective function fails to be concave over the entire domain. Nevertheless, in our example the equilibrium fails to exist not because of the presence of the non-convexity alone, but because of the combination of the non-convexity with a satiation quantity in the demand for θ_2. The contribution of our paper is to show that in a very standard and well behaved problem the only presence of time inconsistent preferences and the consequent solution by backward induction may generate both non-convexities and satiation in the preferences such that a new case of non existence of a competitive equilibrium arises. In what follows we explain and motivate this point.

The non-convexity induced by the backward induction solution could in principle be treated with the standard convexifying methods used in large economies, once the representative agent is interpreted as a "stand in"for a large number of identical consumers. With respect to this, the most important reference is the paper of Luttmer and Mariotti (2007). This paper analyzes a three periods model which is qualitatively similar to our model, in the case in which we interpret the representative agent as a "stand in"for a continuum of identical agents. Imposing quasi-hyperbolic discounting *a la* Laibson (1997), which constitutes a particular type of time-inconsistent preferences that are time-separable, Luttmer and Mariotti (2007) prove the existence of a competitive equilibrium.

The result of Luttmer and Mariotti (2007) can be fully understood with our model. With a continuum of identical consumers, by the law of large numbers, the aggregate demand is given by the convex hull of the individual demand correspondence. In the case of time separable preferences it can be shown that there exists a p_1 such that the period 1 market clearing quantity $\theta_2^* = 1$ belongs to the convex hull of the demand correspondence and such a p_1 is a competitive equilibrium, even though non decentralized. Time separability of preferences is a fundamental assumption for the existence of such a p_1. Time separability of preferences and assumption B.1 together imply that c_2 and c_3 are both normal goods for the period 2 consumer and this is enough to avoid satiation in period 1, because if both c_2 and c_3 increase in θ_2, then for the period 1 consumer utility always increases in θ_2. In other words, if c_2 and c_3 are both normal goods in period 2, then the period 1 preference for θ_2 is monotonic. In such a situation the existence of a p_1 such that the period 1 market clearing quantity $\theta_2 = 1$ belongs to the convex hull of the demand correspondence can be proved with a standard fixed point argument, as the convex hull of the individual excess of demand of period 1 consumer is convex (trivial) and has a closed graph (implied by the upper hemicontinuity and compact-valuedness of the demand function and the monotonicity of the preferences). Using a geometric intuition we can say that a necessary condition to be able to re-establish the existence of the competitive equilibrium through re-convexity methods is that there is at least one optimal quantity θ_2 smaller than the market clearing quantity $\theta_2 = 1$ and at least one optimal quantity θ_2 which is greater.

In our example, convexifying methods do not successfully re-establish the existence of a competitive equilibrium because, for any positive p_1, (B.15) decreases in θ_2 for all $\theta_2 \geq 1$ and this implies that $\theta_2 = 1$ cannot belong to the convex hull of the demand function. This happens because c_3 is an inferior commodity for all $\theta_2 \geq 1$ and the chosen values of b, c imply that, given an increase in θ_2, the consequent decrease in c_3 hurts the period 1 consumer more than how much the parallel increase in c_2 benefits him.

Negative prices can generally re-establish the existence of a competitive equilibria which fails to exist because of non satiation when free disposal is not allowed for every single good, as it happens in our set-up. We now analyze weather negative prices are able to re-establishes the existence of the equilibrium in our example. In the case of negative p_1, the period 1 budget constraint is the following: $\theta_2 \geq 1 + d_1/p_1 - c_1/p_1$, i.e. with a negative p_1 there is no upper bound on the quantity of θ_2 which the consumer can demand and there is instead a lower bound. We proved that any negative p_1 cannot re-establish the competitive equilibrium in our example. We proved this point showing that, given any negative p_1, the unique optimal choice for the period 1 consumer is to demand a quantity of θ_2 that goes to $+\infty$ implying an excess of demand. We therefore showed that for any negative value of p_1 and for a large enough, (B.15) reaches its global maximum for θ_2 that goes to $+\infty$; hence in this case no consumer will demand the market

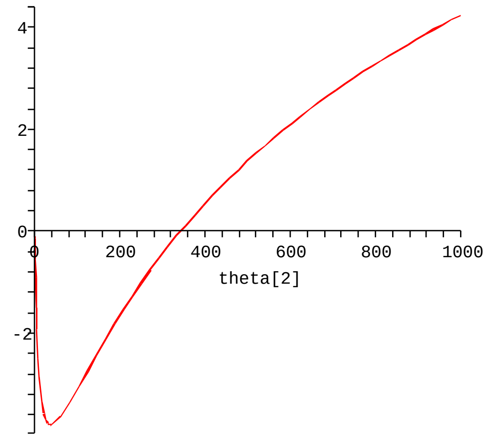

FIGURE B.3. $V_1(\theta_2)$, given $p_1 = -0.1, a = 7, b = 1, c = 7$.

clearing quantity $\theta_2 = 1$. Figure B.3 and B.4 offer an illustration of this situation, the figures show the graph of V_1 as a function of θ_2, given $p_1 = -0.1$, $a = 0.7$, $b = 0.1$, $c = 0.7$. Figure B.3 shows the graph of V_1 over a large subset the positive domain, whereas figure B.4 zooms around $\theta_2 = 1$. Figure B.3 shows that V_1 grows unboundedly for large values of θ_2 and figure B.4 shows that even though we can find a negative p_1 such that $\theta_2 = 1$ is a local maximum, this can never be the global maximum.

B.5. Conclusion

We presented a simple model of consumption saving based on standard assumptions and we showed, through an example, that a competitive equilibrium may fail to exist once a broad class of time inconsistent preferences is allowed. In the example, despite the underlying standard assumptions of the model, allowing for time inconsistency generates induced non-convexities and satiation which cause the competitive equilibrium to

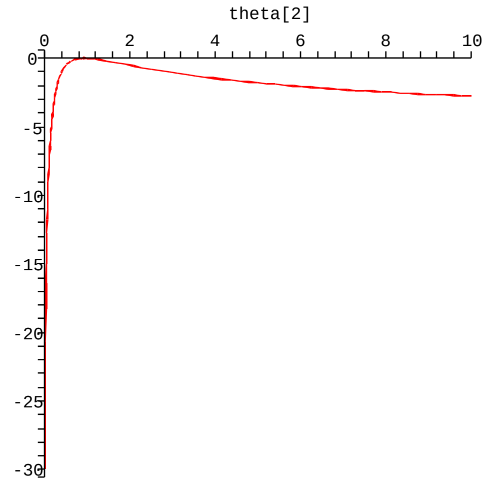

FIGURE B.4. $V_1(\theta_2)$, given $p_1 = -0.1, a = 7, b = 1, c = 7$.

fail. This failure cannot be solved by the standard methods generally applied in case of non-convexities and satiation. It is worthy to mention that our example of non existence does hold for other values of the dividends and of the parameters because given any value of the dividends we can always find appropriate values of a, b and c in (B.7) such that we fall in the non-existence case described in the example.

Bibliography

ACEMOGLU, D. (2002): "Technical Change, Inequality, and the Labor Market," *Journal of Economic Literature*, 40(1), 7–72.

AGHION, P., AND P. BOLTON (1997): "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies*, 64(2), 151–72.

AGHION, P., E. CAROLI, AND C. GARCIA-PENALOSA (1999): "Inequality and Economic Growth: The Perspective of the New Growth Theories," *Journal of Economic Literature*, 37(4), 1615–1660.

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.

ASIMAKOPOULOS, A. (1988): *Theories of Income Distribution*. Boston: Kluver Academic Publishers.

ATKINSON, A. (1980): "Inheritance and the Redistribution of Wealth," in *Public Policy and the Tax System: Essays in honour of James Meade*, ed. by G. Hughes, and G. Heal. London: Allen and Unwin.

ATKINSON, A., AND F. BOURGUIGNON (2001): "Income Distribution," in *International Encyclopedia of the Social and Behavioral Sciences (Economics/Public and Welfare Economics)*, pp. 7265–71. Amsterdam: Elsevier.

ATKINSON, A. B. (1997): "Bringing Income Distribution in From the Cold," *The Economic Journal*, 107(441), 297–321.

BANERJEE, A. V. (2006): "Aspirations, Poverty and Economic Change," in *Understanding Poverty*, ed. by A. V. Banerjee, R. Benabou, and D. Mookherjee. Oxford University Press.

BANERJEE, A. V., AND E. DUFLO (2003): "Inequality and Growth: What Can the Data Say?," *Journal of Economic Growth*, 8(3), 267–99.

BANERJEE, A. V., AND A. F. NEWMAN (1993): "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101(2), 274–98.

BARRO, R. J. (1999): "Ramsey Meets Laibson in the Neoclassical Growth Model," *The Quarterly Journal of Economics*, 114(4), 1125–1152.

BECKER, G. S., AND N. TOMES (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87(6), 1153–89.

——— (1986): "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 4(3), 1–39.

BENABOU, R. (1993): "Workings of a City: Location, Education, and Production," *The Quarterly Journal of Economics*, 108(3), 619–52.

——— (1996a): "Equity and Efficiency in Human Capital Investment: The Local Connection," *Review of Economic Studies*, 63(2), 237–64.

——— (1996b): "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance," *American Economic Review*, 86(3), 584–609.

——— (1996c): "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.

——— (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 70(2), 481–517.

——— (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.

BERGSTROM, T. (1976): "How to discard free disposability - at no cost," *Journal of Mathematical Economics*, 3, 131–134.

BERTOCCHI, G. (2007): "The Vanishing Bequest Tax: The Comparative Evolution of Bequest Taxation in Historical Perspective," IZA Discussion Papers 2578, Institute for the Study of Labor (IZA).

BERTOLA, G., R. FOELLMI, AND J. ZWEIMULLER (2006): *Income distribution in macroeconomic models*. Princeton University Press, Princeton, New Jersey.

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.

BJORKLUND, A., AND M. JANTTI (1997): "Intergenerational Income Mobility in Sweden Compared to the United States," *American Economic Review*, 87(5), 1009–18.

BLACKORBY, C., D. NISSEN, D. PRIMONT, AND R. RUSSELL (1973): "Consistent Intertemporal Decision Making," *Review of Economic Studies*, 40, 239–248.

BLANCHARD, O. J. (1985): "Debt, Deficits, and Finite Horizons," *The Journal of Political Economy*, 93(2), 223–247.

BOURGUIGNON, F. (1981): "Pareto Superiority of Unegalitarian Equilibria in Stiglitz' Model of Wealth Distribution with Convex Saving Function," *Econometrica*, 49(6), 1469–1475.

——— (2004): "The Poverty-growth-inequality triangle," Indian Council for Research on International Economic Relations, New Delhi Working Papers 125, Indian Council for Research on International Economic Relations, New Delhi, India.

CAPLIN, A., AND J. LEAHY (2006): "The recursive approach to time inconsistency," *Journal of Economic Theory*, 131(1), 134–156.

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

CHATTERJEE, S. (1994): "Transitional dynamics and the distribution of wealth in a neoclassical growth model," *Journal of Public Economics*, 54(1), 97–119.

CHECCHI, D., A. ICHINO, AND A. RUSTICHINI (1999): "More equal but less mobile?: Education financing and intergenerational mobility in Italy and in the US," *Journal of Public Economics*, 74(3), 351–393.

CHU, C. Y. C. (1991): "Primogeniture," *Journal of Political Economy*, 99(1), 78–99.

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

COUCH, K., AND T. DUNN (1997): "Intergenerational Correlations in Labor Market Status: A Comparison of the United States and Germany," *Journal of Human Resources*, 32, 210–32.

DALTON, P., AND S. GHOSAL (2008): "Behavioural Decisions and Welfare," The Warwick Economics Research Paper Series (TWERPS) 834, University of Warwick, Department of Economics.

DAS, S., AND C. GHATE (2004): "Endogenous Distribution, Politics, and the Growth-Equity Tradeoff," *Contributions to Macroeconomics*, 4(1), 1199–1199.

DE BOER, P., J. BRÖCKER, B. S. JENSEN, AND J. VAN DAAL (2006): "Theoretical restrictions on the parameters of the indirect addilog system revisited," Econometric Institute Report EI 2006-11, Erasmus University Rotterdam: http://people.few.eur.nl/pmdeboer/research/Restrictions6-11.pdf.

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

DIAMOND, P. A. (1965): "National Debt in a Neoclassical Growth Model," *The American Economic Review*, 55(5), 1126–1150.

DOMAR, E. D. (1946): "Capital Expansion, Rate of Growth and Employment," *Econometrica*, 14, 137–47.

DURLAUF, S. N. (1996): "A Theory of Persistent Income Inequality," *Journal of Economic Growth*, 1(1), 75–93.

DURLAUF, S. N. (1997): "The Memberships Theory of Inequality: Ideas and Implications," Research in Economics 97-05-047e, Santa Fe Institute.

EKELAND, I., AND A. LAZRAK (2008): "Equilibrium policies when preferences are time inconsistent," Mimeo, University of British Columbia, http://arxiv.org/abs/0808.3790v1.

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

FERNANDEZ, R., AND R. ROGERSON (1996): "Income Distribution, Communities, and the Quality of Public Education," *The Quarterly Journal of Economics*, 111(1), 135–64.

——— (1998): "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform," *American Economic Review*, 88(4), 813–33.

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

GALOR, O., AND J. ZEIRA (1993): "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60(1), 35–52.

GLOMM, G., AND B. RAVIKUMAR (1992): "Public versus Private Investment in Human Capital Endogenous Growth and Income Inequality," *Journal of Political Economy*, 100(4), 813–34.

GOLDMAN, S. M. (1980): "Consistent Plans," *Review of Economic Studies*, 47, 533–537.

GOLOSOV, M., A. TSYVINSKI, AND I. WERNING (2006): "New Dynamic Public Finance: A User's Guide," in *NBER Macroeconomics Annual 2006*, ed. by D. Acemoglu, K. Rogoff, and M. Woodford.

GRADSTEIN, M., AND M. JUSTMAN (1997): "Democratic Choice of an Education System: Implications for Growth and Income Distribution," *Journal of Economic Growth*, 2(2), 169–83.

HARRIS, C., AND D. LAIBSON (2001): "Dynamic Choices of Hyperbolic Consumers," *Econometrica*, 69, 935–957.

HARROD, R. F. (1939): "An Essay in Dynamic Theory," *The Economic Journal*, 49, 14–33.

HART, ., AND H. KUHN (1975): "A proof of the existence of equilibrium without the free disposal assumption," *Journal of Mathematical Economics*, 2, 335–344.

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.

HEIFETZ, A., AND E. MINELLI (2006): "Aspiration Traps," Working Papers ubs0610, University of Brescia, Department of Economics.

HELLWIG, M. F. (1980): "Stochastic processes of temporary equilibria : A note," *Journal of Mathematical Economics*, 7(3), 287–299.

HERINGS, P., AND K. ROHDE (2006): "Time-inconsistent preferences in a general equilibrium model," *Economic Theory*, 29(3), 591–619.

HINKLEY, D. V. (1969): "On the Ratio of Two Correlated Normal Random Variables," *Biometrika*, 56(3), 635–639.

HOUTHAKKER, H. S. (1960): "Additive Preferences," *Econometrica*, 28, 244–257.

KALDOR, N. (1955): "Alternative Theories of Distribution," *The Review of Economic Studies*, 23(2), 83–100.

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

KIHLSTROM, R. (1984): "A simple example of the Radner-Stigliz nonconcavity in the value of information," in *Bayesian Models in Economic Theory*, ed. by Bayor, and Kihlstrom. Elsevier.

KOCHERLAKOTA, N. R. (2001): "Looking for Evidence of Time-Inconsistent Preferences in Asset-Market Data," *Federal Reserve Bank of Minneapolis Quarterly Review*, 25, 13–24.

KREMER, M., AND E. MASKIN (1996): "Wage Inequality and Segregation," Harvard Institute of Economic Research Working Papers 1777, Harvard - Institute of Economic Research.

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

LAIBSON, D. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112, 443–477.

LEFRANC, A., AND A. TRANNOY (2004): "Intergenerational earnings mobility in France : Is France more mobile than the US ?," IDEP Working Papers 0401, Institut d'economie publique (IDEP), Marseille, France.

LOURY, G. C. (1981): "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49(4), 843–67.

LUCAS, R. J. (1988): "On the mechanics of economic development," *Journal of Monetary Economics*, 22(1), 3–42.

LUTTMER, E. G. J., AND T. MARIOTTI (2003): "Subjective Discounting in an Exchange Economy," *Journal of Political Economy*, 111, 959–989.

——— (2006): "Competitive Equilibrium when Preferences Change over Time," *Economic Theory*, 27, 679–690.

LUTTMER, E. G. J., AND T. MARIOTTI (2007): "Efficiency and Equilibrium when Preferences are Time-Inconsistent," *Journal of Economic Theory*, 132(1), 493–506.

MAS-COLELL, A., M. D. WHINSTON, AND J. R. GREEN (1995): *Microeconomic Theory*. Oxford University Press.

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

MURTHY, K. N. (1982): "Theoretcal restrictions on the parameters of the indirect addilog demand equations – a comment.," *Econometrica*, 50, 225–227.

NIELSEN, L. (1990): "Equilibrium in CAPM without a riskless asset," *Review of Economics Studies*, 57, 315–324.

PASINETTI, L. L. (1962): "Rate of Profit and Income Distribution in Relation to the Rate of Economic Growth," *The Review of Economic Studies*, 29(4), 267–279.

PELEG, B., AND M. E. YAARI (1973): "On the Existence of a Consistent Course of Action when Tastes Are Changing," *Review of Economic Studies*, 40, 391–401.

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. (1991): "Is Inequality Harmful for Growth? Theory and Evidence," Economics Working Papers 91-155, University of California at Berkeley.

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.

——— (1997): "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing," *RES*, 64(2), 173–190.

——— (1998): "Theories of persistent inequality and intergenerational mobility," in *Handbook of Income Distribution*, ed. by A. Atkinson, and F. Bourguignon. Amsterdam: North Olland.

POLEMARCHAKIS, H. M., AND P. SICONOLFI (1993): "Competitive equilibria without free disposal or nonsatiation," *Journal of Mathematical Economics*, 22, 85–99.

POLLAK, R. A. (1968): "Consistent planning," *Review of Economic Studies*, 35, 201–208.

PRAT, A., AND D. STRÖMBERG (2005): "Commercial Television and Voter Information," CEPR Discussion Papers 4989, C.E.P.R. Discussion Papers.

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.

RODRIGUEZ, F. (1998): "Inequality, Redistribution and Rent-Seeking," Ph.D. thesis, Harvard University.

ROMER, P. M. (1986): "Increasing Returns and Long-run Growth," *Journal of Political Economy*, 94(5), 1002–37.

ROMER, P. M. (1989): "Capital Accumulation in the Theory of Long Run Growth," in *Modern Business Cycle Theories*, ed. by R. J. Barro. Cambridge, MA: Harvard University Press.

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.

SOLOW, R. M. (1956): “A Contribution to the Theory of Economic Growth,” *The Quarterly Journal of Economics*, 70(1), 65–94.

STIGLITZ, J. E. (1969): “Distribution of Income and Wealth among Individuals,” *Econometrica*, 37(3), 382–97.

STROTZ, R. (1956): “Myopia and Inconsistency in Dynamic Utility Maximization,” *Review of Economic Studies*, 23, 165–180.

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