UNIVERSITÀ DEGLI STUDI
DI PERUGIA

ISSN 2385-2275

Working papers of the Department of Economics University of Perugia (IT)

Sophisticated Electoral Accountability: A Political Psychology Agency Theory

Alessandro Belmonte

Working paper No. 13

July 2015

Sophisticated Electoral Accountability: A Political Psychology Agency Theory*

Alessandro Belmonte†
IMT Institute for Advanced Studies Lucca

*I am grateful to Roland Bénabou, Patrick Francois, Edoardo Grillo, Francesco Passarelli, Eugenio Peluso, Michael Rochlitz, Gilles Saint-Paul, Marcello Signorelli, Davide Ticchi, Andrea Vindigni, and the seminar participants at IMT Institute for Advanced Studies Lucca, the 2014 Political Economy Workshop at Catholic University of Milan, the 10th Winter School at Canazei, and the EACES-HSE PhD Seminar (Moscow) for extremely helpful comments and suggestions. The usual disclaimer applies.
†alessandro.belmonte@imtlucca.it

Abstract

I propose a political agency model where rent-maximizer rulers are constrained by sophisticated principals/producers that use an awareness-management model à la Bénabou and Tirole. In the first part of the paper I empirically test the theoretical nexus between education and political sophistication by comparing individuals with different education attainments within more than eighty countries and more that twenty religious groups introduced to capture specific cultural variation in the results. I find elastic (inelastic) political beliefs for respondents with a tertiary (primary) degree according to the quality of political institutions. Motivated by that, I model Political Psychology predictions by introducing heterogeneity on the electoral side: producers are endowed with different levels of education, that increase over time with human capital investments. I allow education to be both the engine of growth and a determinant of political participation; in equilibrium, more educated societies are more able to punish politicians that, in turn, invest more in productive public goods such as infrastructure, roads or legal rules for contracts enforcement. I prove the existence of multiple steady states featuring, respectively, a sophisticated society with congruent politicians in office, and a naive society ruled by dissonant politicians. Finally, I address inequality concerns and show how, for intermediate values, inequality opposingly hits citizens and ruler and only the latter is found to better off; conversely, citizens are averse to inequality, contributing to explain, via sophisticated accountability, why most people dislike living in a society which is too unequal.

Keywords: political economy, voting, signaling, sophistication, naiveté, human capital, economic growth, inequality.

JEL Classification: H30, O43, D21, D72.

1 Introduction

Imagine two hypothetical voters. One is exceedingly well informed about politics, a daily and devout reader of the New York Times, who follows closely the major issues of the day, both national and international. The second, a Daily News fan, is hardly overburdened by the amount of time, or effort, she devotes to public affairs – in fact, looks only at the sports page and cares next to nothing about politics. Is it plausible to suppose that these two voters, asked to make a choice about who should be president of the United States, would make up their minds in the same way?

Sniderman, Brody, and Tetlock (1991) at p. 165

Standard political agency models focuses on elections as an incentive devices through which voters discipline politicians. After the seminal papers by Barro (1973) and Ferejohn (1986), second generation models start to combine hidden action and different types of politicians in order to address political selection issues. Politicians may differ among them in their competence (Austen-Smith and Banks, 1989, Banks and Sundaram, 1993), or in their motivation (Besley and Case, 1995, Coate and Morris, 1995, Fearon, 1999, and Rogoff, 1990)[1]. Besley (2006) proposes the distinction between dissonant and congruent rulers, arguing that the latter are more able/willing to give voters what they want.

Empirical evidence however shows a rather negligible within-country variation with respect to a wide between-country variation in the politicians types distribution[2]. As Figure 1 points out, countries around the World widely differ in terms of WGI Government Effectiveness[3]. On one side, developed countries rank on the first quintile of the worldwide distribution. The first ranked country in 2010 is Finland, followed

[1]For a complete review of the literature see Persson and Tabellini (2000), Grossman and Helpman (2001), Besley (2006), and Ashworth (2012).

[2]There are few recent attempts on political business cycles that combine data from developed and developing countries (Brender and Drazen, 2005; Persson and Tabellini, 2002; Shi and Svensson, 2006). Specifically, Shi and Svensson (2006) find that political budget cycles are large in developing countries but small or nonexistent in developed countries over time.

[3]The WGI Government Effectiveness index captures perceptions of the quality of public services, the quality of the civil service and the degree of its independence from political pressures, the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. The index is constructed upon a number of surveys that summarize the views on the quality of governance provided by a large number of enterprise and expert (survey institutes, think tanks, non-governmental organizations, and international organizations) survey respondents in industrial and developing countries (Kaufmann, Kray, and Mastruzzi, 2010).

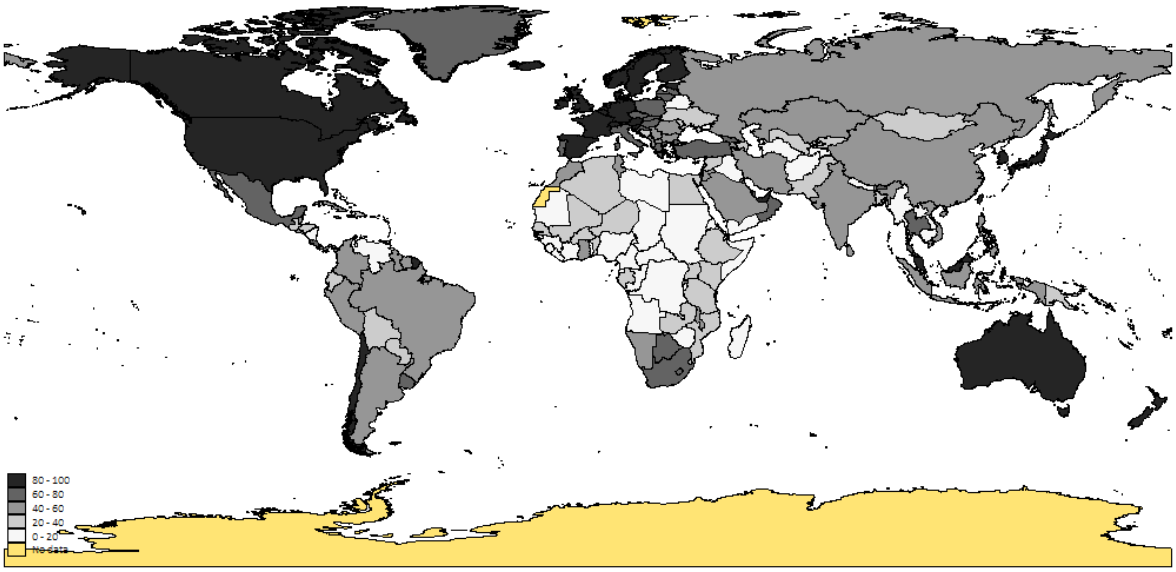


Figure 1: The WGI Government Effectiveness worldwide distribution.

by Singapore, Denmark and Sweden where public and civil services, the degree of its independence from political pressures, the policy formulation and implementation are thought of to be of the highest quality, whereas the government's commitment to such policies the most credible (Kaufmann, Kray, and Mastruzzi, 2010). On the other side, the last ranked countries are mostly developing countries to a great extent located in the sub-Saharan African area (drawn in white in Figure 1).

Adding to the puzzle is the fact that most of developing countries leaders are not less competent than those of developed countries, at least in terms of education background, since most of them graduated in U.S. or European universities. This is the case for example of the longest serving ruler of the African continent, Teodoro Obiang Nguema Mbasogo, ruling over the Equatorial Guinea since the 1979 when he overthrew his uncle in a bloody coup d'etat after having successfully completed his studies in Spain. He is followed, in terms of ruling duration, by the President of Angola, José Eduardo dos Santos who brilliantly graduated in engineering in Soviet Union. *Forbes* also includes the King of Swaziland, King Mswati III, graduated in United Kingdom[4], whereas David Wallechinsky in 'Tyrants: The Worlds 20 Worst Living Dictators' put firmly in discussions the credibility of the elections organized by Paul Biya, President of Cameroon, who years before studied at Sorbonne and Sciences Po in Paris. But many other similar examples involve similar countries across the developing world.

[4]From the article 'The Five Worst Leaders In Africa', *Forbes*, September 2, 2012.

The paper departs from this stylized fact in order to investigate the source of such variation: Why a politician should behave congruently or dissonantly and why in many countries around the World politicians in office are selfish, corrupt, and reckless whereas in others they are not?

In this paper I propose a novel political economy explanation, namely a revised political agency model á la Ferejhon (1986), in which the extent to which the electorate is educated and politically sophisticated affects the way politicians run public business. The channel through which this process is identified is by using new tools from Political Psychology. I argue that rent maximizer politicians rule congruently public business when a sophisticated electorate account for it but the same politicians would have behaved dissonantly had he ruled over a naive electorate.

Despite political sophistication has been generally ignored as a source of heterogeneity in voting in the Political Economy literature, a number of researchers – within the field of Political Psychology – have repeatedly stressed the importance of cognitive sophistication in shaping an individual's ability to make political and economic evaluation (Abramowitz, Lanoue, and Ramesh, 1988; Sniderman, Brody, and Tetlock, 1991; Zallen, 1992; McGraw, 2000; Gomez and Wilson, 2001; Federico and Sidanius, 2002). These works, related to the Political Psychology literature, discussed in Section 2, point out that more educated citizens are less likely to be cheated by politicians. Education provides political sophistication which is meant to give individuals ability in making political evaluations. I model Political Psychology predictions by allowing sophisticated voters to perfectly know which is the type of the politician they are facing of – i.e. whether he is a congruent type or a dissonant one –. In this sense, fully sophisticated voters are not involved in any asymmetric information issues that rather interest the rest of the electorate. On the other side, naive voters are basically unaware of the politicians type and intentions providing to the latter opportunities for the manipulation of the economy. However, as far as sophisticated citizens are the majority, manipulation would be hard to be carried out. In between, a continuum of citizens take economic and political decisions driven by their own political sophistication, the prior belief upon the state of the world, and the signal sent by the politician.

In the first part of the paper, Political Psychology predictions are tested by comparing individuals with different education attainments across more than eighty countries and across more that twenty religious groups introduced to capture specific cultural variation in the results. Using five waves provided by the World Value Survey (WVS), I found that individuals with a primary education show 13 pts less in the scale of interest in politics

than the same peers with a secondary education within each country and Religion group. The most educated individuals, those with a tertiary education, are on the contrary more interested in politics than the ones with a secondary education (+17 pts) suggesting a monotonic relationship between interest in politics and education. Similarly, I found that more educated citizens are more involved in discussing politics with friends (+18 pts). These results are not driven by economic and social status of the respondent and are robust to the inclusion of further individual controls such as gender, age, family status, number of children, and the size of town where they live.

Finally I tested whether education is deputed to affect the extent to which individuals are confident on the Government. Consistently with several Political Psychology predictions (Chappel and Keech, 1984; Abramowitz, Lanoue, and Ramesh, 1988; Sniderman, Brody, and Tetlock, 1991; Zallen, 1992; McGraw, 2000), results robustly suggest that less educated individuals are more confident in the Government in non OECD countries than the more educated ones who are more suspicious despite the economic-social class of belonging, country and religion specific effect. The opposite is rather obtained in OECD countries meaning that differences across individuals with increasing level of education diverge between OECD and non OECD countries. These results pin down the relationship among education, awareness, naivete, and political sophistication and can be interpreted in view of this line of research as the attitude of the less educated persons to be more credulous.

Motivated by that, I propose a dynamic signaling political model where citizens/voters are endowed with different level of education and, according to his own level of education, each of them codifies the signal – the announcement on public investment – sent by the ruler *differently* (Bénabou and Tirole, 2002). In every period, the principal-agent game develops as follow: there are two states of nature about the efficiency level of the State in providing a productive public good. In the low state, a very small amount of the good will be provided. In the high state, productive public investments could support private activities, but this possibility is under ruler's full discretion. Once he/she starts the office comes to know what is the realization of the state, which is then private information and unknown to citizens. After the information is received he/she sends a (costless) signal to citizens that, in turn, use education to screen the plausibility of the announcement. On the basis of posterior beliefs, citizens optimally choose how much invest in human capital and whether to reelect the incumbent or not. Both the accumulation and the voting process depends on political expertise of producers/voters in guessing the correct type of the incumbent. I henceforth show that an electorate ac-

countability is sophisticated when the median citizen is educated enough – or politically sophisticated enough – so as cheating is not the politician's best response.

Despite no costs are involved in signaling, the cheap talk (or babbling) equilibrium is reached only by sophisticated citizens when the sender – the incumbent – behaves dissonantly, always revealing the wrong state. In other words, if the Government is looked at as an untrustworthy one, sophisticated citizens will not pay any attention on the signal it sent. However, in any other cases, the signal still conveys information although the content might be misleading by inducing naive citizens to guess public investment are viable when actually they are not. The incumbent does want to cheat citizens in order to not invest in a not-directly observable productive public project and appropriate of what is left of current tax revenues by rents. On the other side, citizens do prefer more public investments that increase individual productivity and, indirectly, the education attainments. Reducing current public investments would therefore amount to discourage citizens to invest in human capital shrinking expected future tax revenues and, indirectly, his expected future rents – heading the ruler himself toward a binding intertemporal trade-off. The role of citizens as voters is to account for politicians job making sure that they invest when investing is viable. Less investments would mean for citizens an income loss, that they do want to avoid. Once the incumbent ruler is thought to be dishonest by the majority of the voters he will be punished *ex-post* and replaced with another (*ex-ante* identical) politician (see also Ferejohn, 1986). Therefore, in this framework, how democracy works in equilibrium is endogenously determined by the accountability effort pushed by voters that is in turn based on the overall level of political sophistication of the society.

Unlike the existing works in political agency, I take one step further by endogeneizing politicians' types – dissonant versus congruent – and political choices as best responses of the general level of sophistication of the society. In equilibrium I found the ruler's congruence rate to be increasingly dependent on the educational level of the median voter who is found to be pivotal in the political process. This allows us to address the puzzle stated at the beginning explaining away why in some countries rulers perform better than in others. The idea, sketched in Figure 2, is that more educated societies are more able to punish politicians that, in turn, invest more in infrastructure, roads or legal rules for contracts enforcement. These productive public goods foster private investment in education (or human capital) making future electoral accountability more effective. The combination of the accumulation and political mechanisms creates the potential for multiple steady states, one for low-education societies with dissonant rulers and one for high-education societies with congruent rulers.

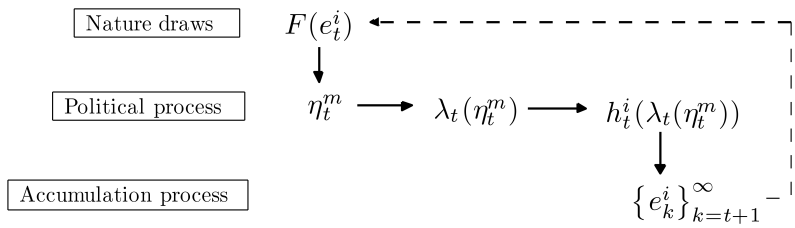


Figure 2: The causal effect mechanisms. In every period t, politicians in charge optimally respond to the political sophistication level of the median citizen, η_t^m. The individual human capital, e_t^i, is inherited from previous periods according to the law of motion $e_{t+1}^i = (1-\delta)e_t^i + h_t^i$. The private investment level in human capital is, in turn, affected by the behavior of the ruler shaping the educational path $\{e_k^i\}_{k=t+1}^{\infty}$ of the society.

Additionally, I show how exogenous shocks can move countries from one equilibrium to another. For example, a financial crisis can make the fiscal budget constraint tighter so as to impede new productive public investments. Anticipating that, citizens reduce human capital investments. If the crisis is persistent enough, the resulting lack of investment will lead the society to loose sophistication – the median citizen moves left –, giving politicians more chances to behave dissonantly. This mechanism contribute to explain several historical events such as the advent of totalitarian regimes in the aftermath of the World War I as the penalties imposed to Germany by the treaty of Versailles – what John Maynard Keynes defined a Carthaginian peace – can be seen as a huge (persistent enough) fiscal shock.

Our main contribution to the existing literature is to provide a general and flexible framework that incorporates into political agency models some aspects of political psychology. Since public investment are not directly observed – as they will be productive in subsequent periods –, both economic and voting decisions are indeed driven by beliefs that are updated through an awareness-management model á la Bénabou and Tirole (2002). According to that citizens are not *equally* aware about what is going on and only some of them are fully Bayesians[5]. I call them sophisticated citizens; sophisticated citizens know what is the congruence rate of the incumbent, though they can be minoritarian in the society. Asymmetric information about the extent to which rulers are

[5]There is a growing literature nowadays, that includes Bénabou and Tirole (2002) among others, based on quasi-Bayesian models: agents generally update beliefs fundamentally in a Bayesian way, but commit a particular error that is inconsistent with rational inference. There are two general mistakes authors assume: systematically incorrect priors, and mistakes in updating beliefs based on information. For a complete review on quasi-Bayesian models see Köszegi (2014).

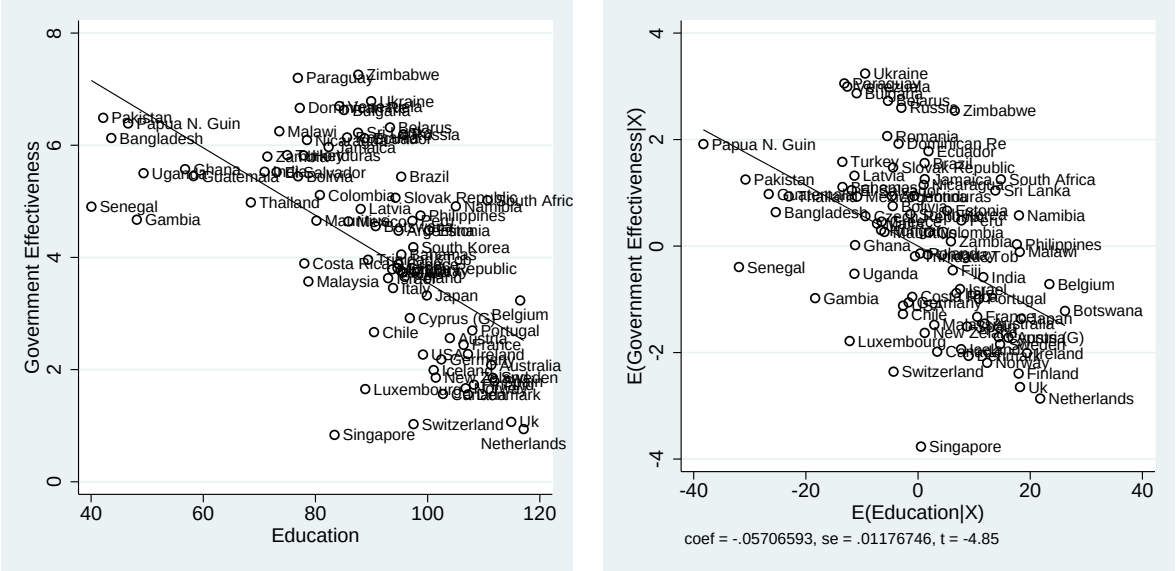
dissonant only involves economic and political decisions of naive citizens, that, when majoritarian, give rent maximizer politicians the chance to manipulate the economy.

The second contribution relies on the introduction into political agency models of heterogeneity on the principal side. Though heterogeneity complicates things a bit, the model leads to tractable analytic results. In particular, in line with Glomm and Ravikumar (1992) and Bénabou (1996, 2000, 2002), I allow citizens to have different education endowments lognormally distributed. Also the distribution of wealth remains lognormal and closed-form solutions are obtained. Contrary to what has been proposed before in political agency literature, no representative voter has been characterized but, in equilibrium, the median citizen is found to be pivotal in the political process, consistently with the median voter theorem. The median citizen also provides a measure of the general level of sophistication (or naivité) of the society leading to a straight intuition of political results and comparative statics.

Our results are partially consistent with the modernization theory that emphasizes the role of education in promoting democracy[6]. On the one hand, education is found to be crucial in shaping democratic institutions via political accountability. On the other hand, however, initially low educated societies fail in providing democratic institutions, and, even worst, bad governments are found to be persistent due to a persistent low level of accountability. This endogenous nexus – theoretically developed in this paper – is captured by Figure 3, that scatters countries' educational level over the WGI Government Effectiveness index, as measure of good Government, for 80 democracies. In Panel (a) the cross-country unconditional correlation is shown outlining several clusters of countries: 'consolidated democracies' with high levels of education, in the South-East corner, that more than a century ago have embarked a joint virtuous evolution of Institutions and political sophistication; 'minimalist democracies' with low level of education, in the North-West corner, that started the democratization process accompanied by a low sophistication ending up in the worst equilibrium (Bidner, Francois, and Trebbi, 2014); countries in between that are still in the middle of the democratization process tending to either equilibria. In Panel (b) I show that the relationship also holds after controlling for (the logarithm of) the GDP per capita.

There is, in addiction, a wide selection of empirical works that have documented the link between the distribution of education and democracy. Some of them, notably

[6]Lipset (1959) identifies two mechanisms by which education promotes democracy: (a) education enables a culture of democracy and, at the same time, (b) it leads to greater prosperity, which is also thought to cause political development.



(a) *WGI Government Effectiveness index.*

(b) *WGI Government Effectiveness index conditional to the GDP per capital level.*

Figure 3: Education attainments and Government effectiveness for $N = 80$ democracies over the period 1990-98 (observations are averaged over the considered period). Data are taken from Persson and Tabellini (2000). Education attainments are defined as the total enrollment in primary and secondary education, measured as a percentage of the relevant age group in the population, computed dividing the number of pupils (or students) enrolled in a given level of education regardless of age by the population of the age-group which officially corresponds to the given level of education, and multiplying the result by 100. The governance indicators are from Kaufmann et al. (2010). It ranges from around 0 to around 10 (lower values correspond to better outcome). In Panel (a) I scatter the Government Effectiveness index over the education attainments. In Panel (b) the relationship is investigated once conditioning to the logarithm of the real GDP per capita. The relationship remains statistically significant.

Barro (1999) and Przeworsky et al. (2000), have provided evidence consistent with the view popularized by Lipset (1959), whereas Glaeser et al. (2004) further investigated the empirical nexus arguing that differences in schooling are a major causal factor explaining not only differences in democracy, but more generally in political institutions. Introducing country fixed effects, Acemoglu et al. (2005) have challenged the view that high educational standard is a prerequisite for a country to become a democracy. This conclusion has been recently reverted by two subsequent papers, Bobba and Coviello (2007) and Castelló-Climent (2008). In particular, Castelló-Climent stresses that what really matters for the implementation and sustainability of democracy is an increase in the education attained by the majority of the population rather than the average years of schooling. A measure of the distribution of education has been included in the regres-

sion making sure that education attainments are yielded by the less educated fraction of population. This last work is much closer to what we do in this work, given that the median of the distribution defines the general level of sophistication of the society whereas in equilibrium it is found to be pivotal.

The paper is also related to other strands of literature. Beside the aforementioned literature on political agency, there is a growing literature on signaling in elections that draws attention to the role of the politician's platform choice to signal to voters his type (Banks (1990), Harrington (1993)). Kartik and McAfee (2007) study a Hotelling-Downs model of electoral competition where a fraction of candidates have character and are exogenously committed to a campaign platform (different from the median voter position). Callender and Wilkie (2007) develop a general electoral framework in which the willingness to lie varies across candidates and discuss the implications of cheap-talking on signaling equilibria. More recently, Acemoglu, Egorov and Sonin (2013) expand this idea arguing that honest politicians, in order to get reelection, choose populist policies (defined as policies to the left of the median voter) as a way of signaling that he is not beholden to the interest of the rich elite. None of these papers discuss or derive the politicians' credibility or attitude to lie as a best response of the electorate sophistication.

Secondly, the paper relates to the literature that emphasizes the complementarity of the investment by the State and the investment by the citizens. Contributions include Barro (1990), Barro and Sala-i-Martin (1992), Benhabib et al. (2001). We use that framework to shed lights on the importance, in democracy, of good politicians in driving economy through public investment. Bad politicians, on the other side, draw private rents wiping out private incentive to invest. The endogenous growth mechanism and the accumulation of human capital has been widely studied by Lucas (1988), Galor and Zeira (1993), Durlauf (1996), Gradstein and Justman (1997), Saint-Paul (1994). Most directly related are the models in Bénabou (1996, 2000, 2002) where producers have different level of human (or physical) capital, lognormally distributed; the accumulation and the redistributive mechanisms dynamically interact pushing different unequal societies to different equilibria (social contracts). Here I build on Bénabou's framework to clarify dynamic interaction between sophisticated electoral accountability and the accumulation process. In this sense, our work is also close in spirit to Bourguignon and Verdier (2000) in which an oligarquic society is split up into an initially uneducated poor class that do not participate to political decisions that are only taken by a rich elite; more equal societies democratize sooner because the higher are the incentive for an educated elite to subsidize the poor's education, that, in turn, gain political control. Similarly to

Bourguignon and Verdier (2000), I allow education to be both the engine of growth and a determinant of political participation. By doing that we show how more educated societies are more able to punish politicians that, in turn, invest more in productive public goods. In the stationary state, countries initially educated reach an upper bound in education and wealth. Others persist in a ignorance trap[7].

The paper is organized as follows. In Section 2 I review the political psychology literature. World Value Survey (WVS) data are also used to test political psychology predictions that represent the main assumptions of the theoretical model. In Section 3 I shall introduce the main features of the model, namely preferences and beliefs, and voting rules. The Markov Perfect Bayesian Equilibrium will be characterized in Section 4, firstly discussing how the accumulation process is affected by the institutional setting – the incumbent's congruence rate – and then endogeneizing political choices as best responses of the sophistication rate of the electorate – via sophisticated electoral accountability. I then discuss the effect of inequality on the players' payoffs. Section 5 discusses the dynamic implication of the political economic interaction and multiple equilibria are characterized. Section 6 proposes an alternative scenario where citizens are also allowed to choose the optimal taxation rate and, accordingly, to punish politicians ex-ante. Section 7 concludes.

2 An introduction to political psychology: an empirical test to its main predictions

Individual political behavior has been widely studied by political psychologists. Many of these works aim at pointing out that political behavior is hardly rational. First of all, people are motivated to act in accordance with their own personality characteristics, values, beliefs, and attachments to groups. Secondly, individuals are imperfect information processors: people employ logical, but often faulty, perceptions of others when deciding how to act, and they often are unaware of the causes of their own behavior (Zaller, 1992; Delli Carpini and Keeter, 1996).

In order to overcome the second issue, political actors require political sophistication – defined by one of the father of political psychology, Robert C. Luskin, as '*the quantity*

[7]Ashworth, Bueno de Mesquita and Friedenberg (2013) similarly speak of accountability traps that are driven by bad expectations. In an accountability trap a polity is caught in a self-reinforcing pattern of behavior with low accountability and – without changing institutions – another self-reinforcing pattern of behavior with greater accountability and higher voter welfare exists.

and organization of a person's political cognitions' (Luskin, 1987). The individual level of political sophistication is therefore of paramount importance to allow decision maker to undertake correct political and economic decisions. Education is one of the forces that make political decision maker more expert in taking political decisions: '*The uneducated man or the man with limited education is a different political actor from the man who has achieved a higher level of education*' (Almond and Verba, 1963).

The work by Almond and Verba received a great deal of attention in Sociology and Political Psychology but has been generally ignored by economists who have been rather attracted by the modernization theory that emphasized the role of education in promoting democracy (Barro, 1999; Przeworsky et al., 2000; Glaeser et al., 2004; Acemoglu et al., 2005; Bobba and Coviello, 2007; Castelló-Climent, 2008). The papers by Glaeser, Ponzetto, and Shleifer (2007) and Botero, Ponce, and Shleifer (2013) represent exceptions – along with the theoretical paper written by Bourguignon and Verdier (2000) – investigating the connection between education and political participation. In Glaeser, Ponzetto, and Shleifer (2007) the connection is identified according to the socialization hypothesis (Helliwell and Putnam, 2007): schooling teaches people to interact with others reducing the costs of civic participation – or increasing the 'social capital' – including voting and organizing and by creating and preserving a culture of democracy. Conversely, Botero, Ponce, and Shleifer (2013) empirically test the accountability hypothesis (Verba and Nie, 1972) according to which better educated people are more likely to complain and report official misconduct[8]. None of them consider the role of education as a cognitive tool which increases the level of political sophistication. In the rest of this section, I provide novel results in favor of the political sophistication hypothesis.

Table 1 and 2 present the main results of the empirical analysis. Table 1 aims at testing the Almond and Verba's hypothesis, i.e. whether different educated individuals are indeed different political actors; Table 2 brings support in favor of the Political Psychology hypothesis. It shows that much of the connection between education and political participation can be explained in terms of political sophistication of the political actor rather than in term of socialization. As far as I know the results shown in Table 2 are entirely original.

I use the five available waves of the World Values Survey (1980, 1990, 1995, 2000, and 2005) and I focus on the following three different measures of political participation/sophistication:

[8]In the accountability hypothesis education affects political participation via a pure human capital channel: complaining is like any other activity for which productivity rises with education.

- A004 – *Indicate how important is politics in your life: very important (A004 = 1), rather important (A004 = 2), not very important (A004 = 3), not at all important (A004 = 4)*;

- A062 – *How often discusses political matters: When you get together with your friends, would you say you discuss political matters frequently (A062 = 1), occasionally (A062 = 2) or never (A062 = 3)?*

- E069_11 – *Confidence in the Government. Could you tell me how much confidence you have in the Government: is it a great deal of confidence (E069_11 = 1), quite a lot of confidence (E069_11 = 2), not very much confidence (E069_11 = 3) or none at all (E069_11 = 4)?*

WVS does not permit to follow same individuals over time but within each country the sample of respondents varies among different waves of studying. Nevertheless, for each of them the WVS Questionnaire collects several individual characteristics such as the gender, education, age, family status, income status, whether she/he has a son and what is the size of the town where she/he is currently living in. We particularly focus on question X025 – *Highest educational level attained*, that, in line with Gleaser, Ponzetto, and Shleifer (2007), I codify in three dummies: $D_{PRIMARY_{icrt}}$, $D_{SECONDARY_{icrt}}$, and $D_{TERTIARY_{icrt}}$. The rest of the individual characteristics are gathered in vector $\mathbf{X}_{icrt}$. i indicates an individual who lives in country c, with religion (or culture[9]) r, at time t. This informational structure allows us to control for unobserved but fixed confounders that operates at a country level in each time t, $\mathbf{\Gamma}_{ct}$, and the ones that operates at a cultural level in each time t, $\mathbf{\Phi}_{rt}$. It also allows to account for aggregate shocks that might hit a country or more countries simultaneously in time t.

Conditional on $\mathbf{\Gamma}_{ct}$, $\mathbf{\Phi}_{rt}$, and observed covariates we can identify the effect of having a primary (tertiary) degree with respect to have a secondary one on political beliefs and behaviors within a country c, a culture r, and in a given point in time t. The interest of

[9]The profession of a religion is highly correlated with the culture of the respondent. In countries where a confession is predominantly professed, as for example the Catholicism in Italy, religion is also a good proxy for national culture (see Guiso et al., 2003). For example, the Italian philosopher Benedetto Croce stated that the Christian tradition has affected the Italian culture so much that Italians cannot be considered non-Christian even if they are atheists. Countries with the same predominant religions, as for example Italy and Spain, may share common cultural traits. The introduction of a religion fixed effect then captures variation within the culture of the respondents, part of which is also captured by specific country features. To avoid bias in the estimation I categorize all the religions with less than 200 members in a common box tagged as 'other'. As a results 24 religion fixed effect are included in the model, including the residual one tagged as 'other'.

	(1) A004	(2) A004	(3) A062	(4) A062	(5) E069_11	(6) E069_11
PRIMARY	0.138***	0.151***	0.196***	0.167***	-0.129***	-0.0689***
	(0.00533)	(0.00708)	(0.00426)	(0.00586)	(0.00537)	(0.00691)
TERTIARY	-0.170***	-0.161***	-0.180***	-0.168***	0.0620***	0.0239***
	(0.00578)	(0.00742)	(0.00468)	(0.00612)	(0.00564)	(0.00710)
Individual Controls		yes		yes		yes
Country FE		yes		yes		yes
Religion FE		yes		yes		yes
Year FE		yes		yes		yes
Constant	2.651***	2.525***	2.135***	2.162***	2.608***	2.924***
	(0.00363)	(0.0244)	(0.00289)	(0.0316)	(0.00358)	(0.0235)
N	219685	125106	148737	83538	198144	112170
N_c	86	64	71	47	82	61
adj. R^2	0.013	0.118	0.044	0.120	0.007	0.156

Robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1: Different models describing the relationship between education and political behavior (columns 1-4) and beliefs (columns 5-6).

the analysis at this stage is the estimation of parameter α, β_1, and β_2 in the following model:

$$\mathbb{E}[y_{icrt}|\cdot] = \alpha + \beta_1 D_{PRIMARY_{icrt}} + \beta_2 D_{TERTIARY_{icrt}} + \phi' \mathbf{X}_{icrt} + \delta_c + \mu_r + \theta_t \qquad (1)$$

where y is A004, *Interest in politics*, in column (1) and (2) of Table 1, A062, *How often discusses political matters with friends*, in column (2) and (3), and E069_11, *Confidence in the Government*, in column (4) and (5).

The uneven columns in Table 1 report OLS estimations for the basic model where $(\hat{\alpha}+\hat{\beta}_1)$, $\hat{\alpha}$, and $(\hat{\alpha}+\hat{\beta}_2)$ stand for the conditional expected value of y for the distribution of respondents all around the world with a primary education, secondary education, and tertiary education respectively. Estimations suggest that respondents with a tertiary degree are more interested in politics and discuss more frequently political matters with friends than respondents with a secondary education. The former are also found to be less confident to the Government than the latter. The respondents with a secondary

degree are in turn more interested in politics, discuss more frequently political matters with friends, and are less confident to the Government than respondents with a primary education.

Similar estimations are reported in the even columns where I control for individual characteristics and look at the variation within countries, religions, and years. In this specifications the constant $\alpha = \mathbb{E}[y_{icrt}|D_{PRIMARY_{icrt}} = 0; D_{TERTIARY_{icrt}} = 0; \mathbf{X}_{icrt} = \mathbf{0}, \mathbf{\Gamma}_{ct}, \mathbf{\Phi}_{rt}]$ has a straight interpretation in terms of the average value in the scale y for a grown-up single female with no child with a secondary education and average income (lower middle class) living in 2005 in a US city with more than hundred thousands inhabitants. In both scales results confirm a monotonic relationship between political attitudes and beliefs and education bringing support to the Almond and Verba's hypothesis.

Similar individuals who live in different countries are however expected to have different beliefs over the Government as the quality of politicians might be sensibly different between countries. In developing countries, where democracy is not fully assessed, Governments are on average less effective than what we have in reacher countries. As a result, similar individuals with same education are expected to be less confident in the Government in the former countries provided they have enough political sophistication.

In models such as the one presented in Table 1, differences at country level have been captured by the inclusion of a country fixed effect δ_c, which allows us to estimate the variation of the dependent variable within N_c countries[10]. Results have been interpreted starting from United States and they show significant differences between countries. For example in column (6), a grown-up single female with no child and secondary education and average income (lower middle class) living in a Russian city with more than hundred thousands inhabitants in 2005 is significantly less confident in the Government than the same American peer.

In Table 2 I use the OECD membership as a source of exogenous variation in the quality of politicians and political institutions[11]. Between OECD and non-OECD countries several differences arise in the WVS sample. First of all differences are significant with respect to the shares of individuals with at most a primary diploma. These shares are 37.48 ppts and 33.50 ppts in developing countries and in OECD countries respectively. Conversely, more individuals get a tertiary degree within the OECD area (24.77 ppts) than in developing countries (19.93 ppts).

[10]The cardinality of the countries sample is reported in Table 1 for each specification as N_c.

[11]OECD includes many of the World's most advanced countries. They have in common that they all rank at top positions in the World Government Indicators cross-country distribution (Kaufmann, Kray, and Mastruzzi, 2010) and in the Polity IV index of democracy (Jaggers and Marshall, 2003).

In Table 2 I rather investigate the effect of the individual level of education on political attitudes and beliefs in order to test two hypotheses related with political psychology predictions:

i) Within an OECD country individuals with tertiary education are more confident in the Government than the same peers living in a developing country.

ii) Such a difference is far greater than what we have for individuals with a primary education.

Column (3) of Table 2 reports estimations for the following model:

$$\begin{aligned}\mathbb{E}[E069\_11_{icrt}|\cdot] = \alpha_1 + \alpha_2\overline{OECD} + \beta_1 D_{PRIMARY_{icrt}} + \beta_2 D_{TERTIARY_{icrt}} \\ + \gamma_1 D_{PRIMARY_{icrt}} * \overline{OECD} + \gamma_2 D_{TERTIARY_{icrt}} * \overline{OECD} \\ + \phi' \mathbf{X}_{icrt} + \delta_c + \mu_r + \theta_t\end{aligned} \tag{2}$$

where $\overline{OECD} = 1 - OECD = \{0, 1\}$ is null if the respondent lives in a country member of the OECD. Such a models not only introduce an OECD status but additionally interact education with such economic/political status.

The parameters of interest now become six, including the constant. We are now able to reexpress the two hypotheses as follow:

i)

$$\begin{aligned}\mathbb{E}[E069\_11|D_{TERTIARY_{icrt}} = 1, \overline{OECD} = 0, \mathbf{X}_{icrt}, \mathbf{\Gamma}_{ct}, \mathbf{\Phi}_{rt}] - \\ \mathbb{E}[E069\_11|D_{TERTIARY_{icrt}} = 1, \overline{OECD} = 1, \mathbf{X}_{icrt}, \mathbf{\Gamma}_{ct}, \mathbf{\Phi}_{rt}] < 0\end{aligned}$$

$$\alpha_1 + \beta_2 < \alpha_1 + \alpha_2 + \beta_2 + \gamma_2 \quad \Longleftrightarrow \quad \alpha_2 + \gamma_2 > 0$$

ii)

$$\begin{aligned}\mathbb{E}[E069\_11|D_{TERTIARY_{icrt}} = 1, \overline{OECD} = 0, \mathbf{X}_{icrt}, \mathbf{\Gamma}_{ct}, \mathbf{\Phi}_{rt}] - \\ \mathbb{E}[E069\_11|D_{TERTIARY_{icrt}} = 1, \overline{OECD} = 1, \mathbf{X}_{icrt}, \mathbf{\Gamma}_{ct}, \mathbf{\Phi}_{rt}] > \\ \mathbb{E}[E069\_11|D_{PRIMARY_{icrt}} = 1, \overline{OECD} = 0, \mathbf{X}_{icrt}, \mathbf{\Gamma}_{ct}, \mathbf{\Phi}_{rt}] - \\ \mathbb{E}[E069\_11|D_{PRIMARY_{icrt}} = 1, \overline{OECD} = 1, \mathbf{X}_{icrt}, \mathbf{\Gamma}_{ct}, \mathbf{\Phi}_{rt}]\end{aligned}$$

$$\alpha_2 + \gamma_2 > \alpha_2 + \gamma_1 \quad \Longleftrightarrow \quad \gamma_2 > \gamma_1$$

	(1) A004	(2) A062	(3) E069_11
$PRIMARY$	0.164***	0.167***	-0.0236
	(0.0128)	(0.0106)	(0.0122)
$TERTIARY$	-0.207***	-0.198***	-0.0794***
	(0.0130)	(0.0115)	(0.0120)
$PRIMARY*\overline{OECD}$	-0.0175	-0.000695	-0.0629***
	(0.0148)	(0.0122)	(0.0142)
$TERTIARY*\overline{OECD}$	0.0643***	0.0429**	0.151***
	(0.0155)	(0.0133)	(0.0145)
$\overline{OECD}$	0.653***	-0.0876***	0.143***
	(0.0353)	(0.0216)	(0.0333)
Individual Controls	yes	yes	yes
Country FE	yes	yes	yes
Religion FE	yes	yes	yes
Year FE	yes	yes	yes
Constant	2.546***	2.177***	2.968***
	(0.0252)	(0.0321)	(0.0242)
N	125106	83538	112170
N_c	64	47	61
adj. R^2	0.118	0.120	0.157

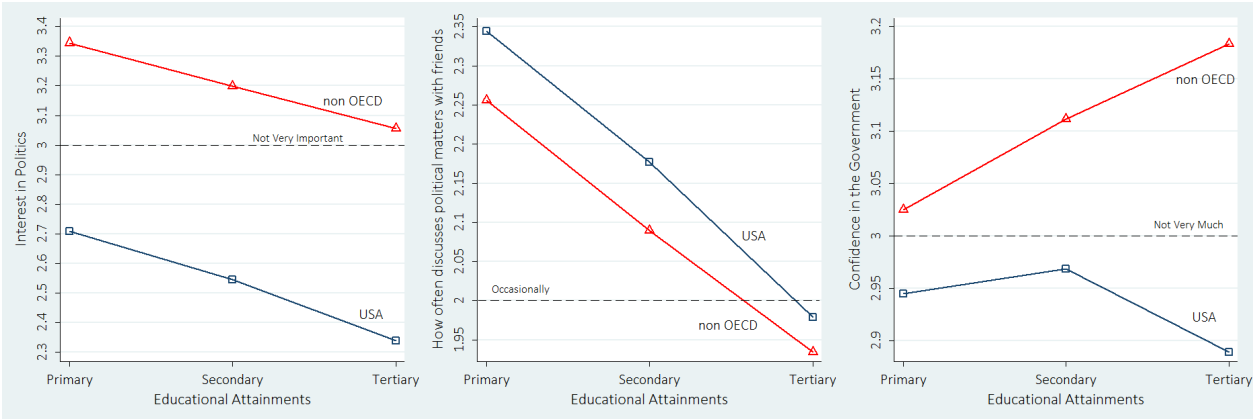
Robust Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Education and political behavior (columns 1-2) and beliefs (columns 3) once introduced the OECD status.

Table 2 also reports estimations for A004 and A062 in column (1) and (2) respectively. Figure 4 draws the patterns of the expected value of the scale y across different level of education and between OECD, in blue, and non OECD countries, in red.

The most interesting results come from combining WVS question E069_11 with the OECD status. Column (3) brings support for both the two hypotheses that political psychology theorizes. In particular, $\hat{\alpha_2} + \hat{\gamma_2} = 0.294 > 0$ and $\hat{\gamma_2} - \hat{\gamma_1} = 0.213 > 0$[12].

[12]The F-test associated to the event that α_2 and γ_2 are both zero is 77.86. The F-test associated to



(a) *Importance of politics.* (b) *How often discusses political matters.* (c) *Confidence in Government.*

Figure 4: Different patterns of political behavior (panel (a) and (b)) and political beliefs (panel (c)) across different educated individuals between OECD countries, in blue, and developing countries, in red.

In Figure 1(c) I sum up the parameters of interest. The general picture that emerges tells us that political beliefs are very likely to be driven by both endogenous factors, *in primis* the level of political sophistication (education) of the respondent, and exogenous factors, such as the quality of political institutions. According to that I find that any differences between individuals living in developing and developed countries but with a primary diploma are based on a location effect operating at the regional level ($\hat{\alpha}_2 > 0$, whereas $\hat{\gamma}_1 < 0$). Differences between individuals living in developing and developed countries but with a tertiary degree are mainly based on a scale effect operating at individual educational level ($\hat{\gamma}_2 > 0$). The differences across individuals with increasing level of education therefore diverge between OECD and non OECD countries. In the former countries more sophisticated individuals realize that political institutions work satisfactorily and respond to the Questionnaire by revealing that they are more confident to the Government than the same peers with a primary degree. In the latter instead the more educated individuals are more aware of the bad political situation and respond to be less confident to the Government in office. Less educated individuals are barely aware of what is going on so as their beliefs do not move in accordance to the quality of political institutions. In other words, I find inelastic political beliefs for respondents with at most a primary degree – politically naive people – with respect to the quality of political institutions. Conversely, respondents with at least a tertiary degree – politically

the event that $\gamma_2 - \gamma_1 = 0$ is 178.21.

sophisticated people – show elastic political beliefs.

3 The model

3.1 Technology and preferences

An incumbent politician competes against the opponent to stay in office. In absence of term limits, office guarantees to incumbents a lifetime flow of rents

$$U_0^r = T_0 - BG_0 + \mathbb{E}_0 \sum_{t=1}^{\infty} \beta^t (T_t - BG_t)\varphi_{t-1} \tag{3}$$

conditionally of being in office in time t. φ_t is a state variable of the economy, given at time $t = 0$ and equals to 1 if r is the incumbent. Therefore, $\varphi_0 = 1$ if r is the incumbent. Yet, φ_t endogenously evolves over time: at the end of every period t election are held and voters are called for to retain the incumbent or to replace him with an identical challenger, that is to choose an action $\varphi_t = \{0, 1\}$. Rents are composed by tax revenues T_t the citizens pay to benefit from a productive public good that costs B to the administration, plus future discounted incomes, conditionally of being reelected, i.e. $\varphi_{t-1} = 1$. $G_t = \{0, 1\}$ is an indicator function equals to 1 if the public investment is made in time t.

The opponent running against the incumbent is identical in all respects from the viewpoint of the voters. Thus the only reason for not reappointing the incumbent is to punish him ex post by taking off future rents, and since the opponent is identical it is indeed (weakly) optimal for the voters to carry out this punishment[13].

The politician in office rules over a continuum of unit mass of risk neutral infinitely-lived citizens endowed with different initial educational levels, e_0^i. The distribution of citizens is initially exogenous according to $F(e_0^i)$ and evolves (endogenously) across periods t on the basis of the following law of motions:

$$e_{t+1}^i = (1 - \delta)e_t^i + h_t^i \tag{4}$$

where h_t^i is investment in human capital carried out to accumulate human capital that in turns persists over time with rate $1 - \delta$. Despite educational differences, all citizens

[13] In the Ferejhon (1986) setting all politicians are thought to be a priori equally untrusty. As a result, elections work as a referendum over the incumbent based on his past actions. This basically amounts to rule out the case in which the ruler provides the public good ($G_t = 1$) without being reelected. Furthermore, since challengers are *a priori* of the *same* type imposing that, when indifferent, voters reelect the incumbent is costless.

use the same (Cobb-Douglas) technology to transform individual effort in the unique final good in the economy to be consumed according to (6):

$$y_t^i = G_{t-1}^{\gamma}(e_t^i)^{\alpha} \tag{5}$$

$$c_t^i \leq (1-\tau)y_t^i - \frac{1}{\phi}(h_t^i)^{\phi} \tag{6}$$

where $(1/\phi)(h_t^i)^{\phi}$ are convex costs in investing h_t^i in education, given that $\phi \geq 2$. In order to get analytical results, we allow investing to be equally costly for every skilled citizens despite the most skilled ones are expected to exert less effort in learning. Yet, $\gamma, \alpha > 0$ are the elasticity of the public and private investment on the output, respectively[14]: in this fashion, investment by the state is complementary to the investments of citizens (accordingly to Barro (1990), Barro and Sala-i-Martin (1992), Benhabib et al. (2001)). Production is carried out by citizens, but it depends on their investments as well as on the quality of the infrastructure, the strength of the law and order, or yet on the legal rules for contract enforcement. All these factors are determined by the public good investments made by the politician and his decision to not carry out public investments sharply leads citizens to not plan any private investments. A proportional taxation scheme is levied by the ruler to collect tax revenues, $T_t = \tau y_t$, aimed to invest in a productive public good (i.e. $G_t = 1$) and to remunerate the politician in charge according to (3)[15]. The investment made in period t will be productive in the subsequent period $t+1$ so that the political accountability effort carried by citizens for having a properly use of public money falls together with an increase in future consumption flows.

Finally, all citizens discount the future with factor β (the same of the politician) and have the same additive (across states and across time) lifetime utility function

$$U_0^i = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t c_t^i \tag{7}$$

that only depends on consumption c_t^i, defined above.

The incumbent decision to implement the public project, i.e. whether $G_t = 0$ or $G_t = 1$, is unobserved in time t but will be productive (and therefore observed to citizens)

[14]In principle, any values of γ could be allowed albeit for $\gamma \leq 0$ citizens do not care in public sector weakening the attention toward politician announcements (for example, $\gamma = 0$ means that public investments are no longer productive). Furthermore, negative values of γ commute G_t to be a bad that is not so far from what has been observing in many developing countries (e.g. by running wars).

[15]In Section 6 an alternative scenario is discussed in which citizens – in every period t – decide the optimal level of taxation $\tau_t^* = \tau(e_t^i)$.

in the subsequent period $t+1$. As it would be clear in the following Section, citizens only know (imperfectly) the type of the ruler and on the basis of such (incomplete) information infer the value of G_t. All citizens' economic and political decisions are then driven by beliefs that are described below. The framework is relaborated from Bénabou and Tirole (2002).

3.2 Information and beliefs

There are two state of nature about the efficiency level of the State in providing a productive public good, $\sigma_t = \{H, L\}$. In state L, a very small amount of the good will be provided ($G_t^L \approx 0$). On the other side, productive investments could support private activities in state H, but this possibility is under ruler's full discretion (i.e. $G_t^H = \{0, 1\}$). Once the policymaker starts the office comes to know what is the realization of σ_t, which is then private information and unknown to citizens who have common prior $P(\sigma_t = H) = q$[16]. After the information is received the incumbent sends a signal to citizens, $\hat{\sigma}_t = \{H, L\}$, assessing the state of the world. As equation (3) makes clear, the ruler can have the incentive to cheat citizens by signaling L when the true state is $\sigma_t = H$. In such a way, he can pocket public savings for himself by larger rents[17]. This reasoning also implies that in state L he never has the incentive to cheat, that amounts to say that $P(\hat{\sigma}_t = L | \sigma_t = L) = 1$.

Due to the asymmetric source of information over G_t, the role of citizens as voters is to account for the incumbent announcement making sure that he invests less in state L and more in state H[18]. Preventing to be cheated requires a minimum level of sophistication and awareness about the politicians' purposes that not all of them possess. In what

[16]The assumption of common prior is not crucial, and is only done to highlight the effect of educational level on the inferential process. Furthermore, q might also be interpreted as the extent to which citizens are optimistic over the economy. Introducing heterogeneity on the prior would amount to allow citizens to be differently optimistic.

[17]This is the short run gain obtained by choosing $G_t = 0$ so as the cost of public investment, BG_t, is zero and, as a result, current rents equal current tax revenues T_t. In addiction, there is a more subtle gain which operates in the long run: due to imperfect credit markets, not investing today in G_t impedes citizens to invest, in turn, in human capital which, for every $\delta > 0$, amounts to move the median voter to the left of the distribution becoming more credulous.

[18]The consequences of temporary information asymmetries are investigated also by Rogoff and Sibert (1988) and Rogoff (1990). They argue that politicians use to manipulate certain not-observable macroeconomic policy variables in the imminence of the vote, to increase the probability of being reelected. Herrington (1993) goes further stressing the role of voters' uncertainty on which policy is best, though no one as far as I know shape the role of voters' naivité.

follows, I exploit the political psychology predictions that have been tested in Section 2: more sophisticated voters are the most educated in society whereas naive people are the least educated[19]. Then, according to his own level of education e_t^i each citizen codifies the signal sent by the ruler *differently*. The idea is that education helps us to convey all the essential informations to make inference. $\eta_t^i = e_t^i/\bar{e}$, where $\bar{e}$ can be larger or equal to the maximum of the education distribution, captures this intuition:

- $\eta_t^i = 1 \Leftrightarrow e_t^i = \bar{e}$ stands for an individual i who behaves as a perfect bayesian agent with all the information in hands;
- $\eta_t^i = 0 \Leftrightarrow e_t^i = 0$ stands for an individual i who is fully naive and believes that the ruler tell the truth whatever the state of the world[20].

In $\sigma_t = H$, the ruler can either tell the truth to citizens (and investing $G_t^H = 1$) or tell a lie ($G_t^H \approx 0$). Each citizen i knows the congruence rate $0 \leq \lambda_t \leq 1$ of the policymaker according to his own level of education. It follows that only the perfect bayesian citizen (i.e. with $\eta_t^i = 1$) can predict with certainty the true rate of the politician; less educated citizens can only know a fraction of his intentions, whereas fully naive agents (i.e. with $\eta_t^i = 0$) believe what the ruler tells whatever the state of the world. The probability that he is cheating (or dissonance rate) is *for citizen i* then equal to:

$$P(\hat{\sigma}_t^i = L|\sigma_t = H) = (1 - \lambda_t)\eta_t^i \tag{8}$$

whereas the probability that the government says the truth (or congruency rate) is:

$$P(\hat{\sigma}_t^i = H|\sigma_t = H) = 1 - (1 - \lambda_t)\eta_t^i \tag{9}$$

where $\lambda_t = 1$ stands for a politician that always tells the truth, and $\lambda_t = 0$ for one that always tells a lie.

Using an awareness-management model à la Bénabou-Tirole (Bénabou and Tirole, 2002) each citizen i assesses the plausibility of the incumbent's claiming: Is it plausible that the state is in the low state? Is the ruler cheating us investing less in public goods and increasing office rents? The probability that the ruler will cheat is for citizen i equal

[19]Education alone does not explain the entire variation associated to the set of individual political attitudes and beliefs. The literature on political psychology suggests other sources of political naivite such as inexperience, innocence, or simplicity.

[20]See Bénabou and Tirole (2002) and Bénabou (2013) for a discussion about bayesian behavior and naivete and the way to model it in a standard microeconomic model.

to[21]:

$$
\begin{aligned}
p_t^i(\lambda_t, \eta_t^i, q) &\equiv P(\sigma_t = H | \hat{\sigma}_t^i = L; \lambda_t, \eta_t^i) \\
&= \frac{(1-\lambda_t)\eta_t^i q}{1 - q + (1-\lambda_t)\eta_t^i q}
\end{aligned}
\tag{10}
$$

It turns out that $(1-\lambda_t)\eta_t^i \leq 1$ implies that $p^i \leq q$. The babbling equilibrium where $p^i = q$ is reached iff $\lambda_t = 0$ and $\eta_t^i = 1$: if the government is looked at as a untrustworthy one, citizens will not pay any attention on the signal it sent and classify it as a cheap talkers[22]. However, such a level of awareness can be caught by only the most educated agents in society. If λ_t is still null but $\eta_t^i < 1$, the agent will be led to throw away some degree of awareness (or sophistication) in the inferential process. In this sense, the government has some interests in decreasing η_t^i, to make its moves wider.

3.3 Voting

Citizens vote retrospectively according to the evidence they have collected on political announces. Since the government's strategy is realized only after the elections, only individual beliefs are involved in the inferential process. Every citizen i processes all the informations collected and votes again for the incumbent if she has no evidence of the fraud, i.e. iff the evidence E_t in favor of the hypothesis p_t^i is not positive[23]:

$$
E(p_t^i) \equiv \log\left(\frac{p_t^i}{1-p_t^i}\right) \leq 0 \tag{11}
$$

which occurs where $p_t^i \leq 1/2$. It in turn means that if the majority of them has no evidence about the cheating move of the policymaker he will be reelected, contingency that occurs when $\int_0^{1/2} f(p_t^i)dp_t^i \geq 1/2$. Now, given $0 \leq p_t^m \leq q$ with $\int_0^{p_t^m} f(p_t^i)dp_t^i \equiv 1/2$, we require that

$$
\int_0^{1/2} f(p_t^i)dp_t^i \geq \frac{1}{2} = \int_0^{p_t^m} f(p_t^i)dp_t^i \iff p_t^m \leq \frac{1}{2} \tag{12}
$$

by monotonicity of $F(\cdot)$. In other words, it turns out that the incumbent won't be reappointed if the median citizen thinks that he is plausibly cheating them. Therefore, if politician cares about reelection he would be willing to push down p_t^m at least to 1/2.

[21] A number of posterior distributions are drawn in Appendix – Section A.2 – for different $F(\eta^i)$, priors, and λ.

[22] See Callender and Wilkie (2007) for a discussion on credible and cheap talkers politicians.

[23] Straightforward computation shows that the logit function, $E_t^i = \log(1-\lambda) + \log \eta^i + \log(q/1-q)$, is an increasing function of η^i and q, whereas decreases with λ. Once again, more educated people collect more evidence upon the job of politicians.

This is of course easier in a society where people can easily be made fools, i.e. in one with a skewed distribution of η_t^i. I summarize this result in Proposition 1:

PROPOSITION 1: *Let* $p_t^m(\lambda_t, \eta_t^m, q) \equiv P(\sigma_t = H|\hat{\sigma}_t^m = L; \lambda_t, \eta_t^m)$.

(*i*) *If* $p_t^m(\lambda_t, \eta_t^m, q) \leq \frac{1}{2}$ *the optimal strategy in the stage game is to play* $\varphi_t = 1$.

(*ii*) *If* $p_t^m(\lambda_t, \eta_t^m, q) > \frac{1}{2}$ *the incumbent will not be retained (i.e.* $\varphi_t = 0$*).*

3.4 Timing of events

The timing of events within every period is as follows:

T1 Nature draws $\sigma_t = \{H, L\}$, that is private information of the ruler. Each citizen inherits e_t^i from the private investment made at time $t-1$ and benefits from the public investment made by the former government, G_{t-1}.

T2 Politician in office chooses the action λ_t, the congruence rate, and, accordingly, invest in a public good (that will be productive in $t+1$), i.e. chooses $G_t = \{0, 1\}$.

T3 Citizens plan to invest h_t^i in human capital based on their beliefs on the ruler's type.

T4 Elections are held (the median citizen chooses $\varphi_t = \{0, 1\}$ based on posterior beliefs).

T5 Payoffs are given by rents and consumptions to politician and citizens respectively.

4 The Markov Perfect Bayesian Equilibrium

4.1 The political agency problem and the definition of equilibrium

The model features a typical agency problem where politicians in office maximize private rents, expressed by (3). In every period t he must decide whether to appear pleasant to voters and being reelected or not, by choosing which type of politician being (i.e. an admissible value of λ_t). Behaving congruently, for a given distribution of education, raises the chance of being reelected in the next period, but *nothing* says on political choices that the incumbent will take on in the future. In fact, despite the strong incentive the ruler has, any promises cannot be credible, since, in every period t, the ruler has the

chance of disregarding the announcement made on G_t and nevertheless being reelected. This commitment problem impedes politician to build a reputation over time[24] and allows us to solve the dynamic game using the Markov Perfect Bayesian Equilibrium (MPBE) concept. The MPBE is defined as a set of Markovian strategies which only depends on the current payoff-relevant states of the economy, $e_t^i \in \mathbb{R}_+ \cup \{0\}$, $\sigma_t \in \Sigma \equiv \{L, H\}$, $G_{t-1} \in \{0, 1\}$, and $\varphi_{t-1} \in \{0, 1\}$, and on prior actions within the same date, according to the timing of events in 3.4, denoted by $k_t \in \mathcal{K}$; for every possible history, $l^{t,k} \in L^{t,k}$, of the dynamic game up to time t and stage k of the stage game of time t, such a strategies are best responses to each other. The Markovian strategies are also optimal given beliefs, and beliefs are updated using Bayes' rule, according to (10).

More formally, for every $q \in [0, 1]$ and for each value of the state variable and each combination of prior moves in the stage game given by $\mathcal{K}$, a Markovian strategy mapping

$$s : [0, 1] \times \Sigma \times \{0, 1\}^2 \times \mathbb{R}_+ \cup \{0\} \times \mathcal{K} \to [0, 1] \times \mathbb{R}_+ \cup \{0\} \times \{0, 1\}$$

assigns a value for each of the actions: the congruence rate taken on by the ruler, $\lambda_t \in [0, 1]$, the amount of private investment made by each citizen, $h_t^i \in \mathbb{R}_+ \cup \{0\}$, and the decision of reelecting the incumbent, $\varphi_t \in \{0, 1\}$. We then proceed to determine the equilibrium within each period by backward induction, given e_t^i, σ_t, G_{t-1}, and φ_{t-1} and the beliefs.

4.2 Elections and sophisticated political accountability

At the end of each period t, elections are held based on posterior beliefs that each citizen i has. Elections predict that the incumbent will be reelected if the median of the education distribution guesses that he is not cheating them. Reexpressing (12) I get the set of admissible strategies for an incumbent who cares about reelection:

$$\lambda_t \geq 1 - \frac{1}{\eta_t^m}\Big(\frac{1-q}{q}\Big) \equiv \lambda_t^*(\eta_t^m, q) \tag{13}$$

According to (13), the optimal strategy in the stage game of the ruler is, if the public investment is not too costly, to play in time T2

$$\max\Big[0, \lambda_t^*(\eta_t^m, q)\Big] \leq \lambda_t \leq 1 \tag{14}$$

[24]The impossibility of building up a reputation roots with the seminal work by Barro (1973) and it is a milestone of political agency models with no types differences. According to that campaign promises are meaningless, given that lying is costless, and policies are determined only once a candidate is installed in office.

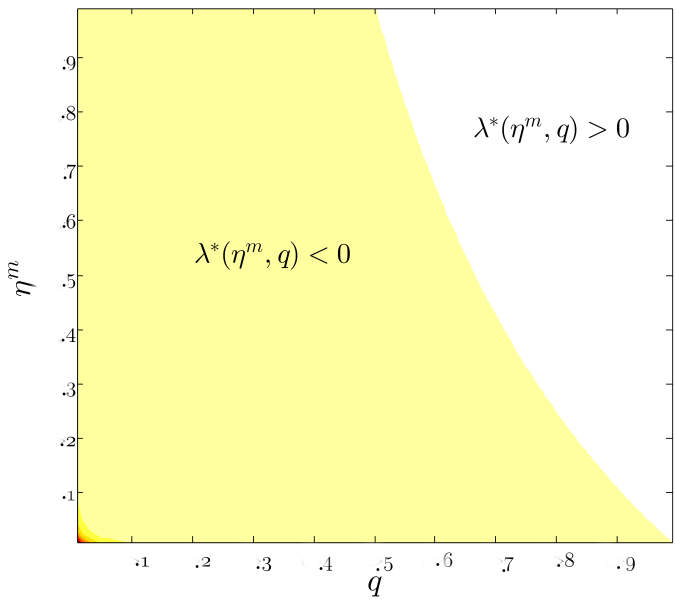


Figure 5: Set of admissible strategies for a politician who cares in reelection as a function of η_t^m and q. The function $\lambda_t^*(\eta_t^m, q)$ monotonically increases in η_t^m and q. The white section of the contour plot depicts all the points where $\lambda_t^*(\eta_t^m, q) > 0$ so as the pure strategy of cheating is not admissible. On left side – the colored section – $0 \leq \lambda_t \leq 1$ is always greater than $\lambda_t^*(\eta_t^m, q)$ as it is now negative (see inequality (13)). Any combination of q and η_t^m in this area makes viable for the government to cheat citizens. Note that even though voters think that the state $\sigma_t = H$ occurs almost surely (i.e. for q very close to one), a distribution of education collapsed around zero allows the government to cheat them.

Note that the fact that $\lambda_t^*(\eta_t^m, q)$ is increasing in η_t^m strongly put into the government's business the issue of accountability. Finally, if $q \leq 1/2$ the inequality (13) is always true, for all λ_t (see Figure 5).

4.3 The accumulation process

We now engage with the accumulation process carried out by each citizen given the ruler's strategy λ_t. In T3 citizens invest in human capital on the basis of the informations they have in mind. Accordingly, no private investments would be planned if the state is thought to be low because $G_t^L \approx 0$ certainly and so will be the output tomorrow. Trivially, each citizen i would carried out $h_t^i = 0$ in any MPBE no matter the education level he has and, in that case, the human capital stock would be firmly the same than period $t-1$ unless depreciation takes place at rate $\delta > 0$. On the other side, with probability q, the state is high and public investments are thought to be viable; suddenly, each agent's task becomes to puzzle out whether the ruler is telling the truth or a lie (for

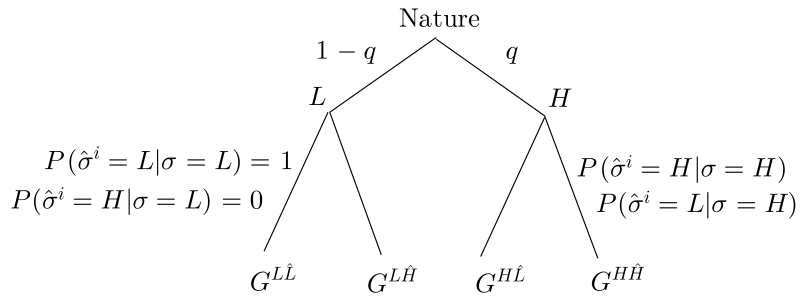


Figure 6: Public investments subgame. Note that when the state is thought to be low (L) every citizens know that the ruler will play $G_t = G^{L\hat{L}} = 0$, no matter the educational level of him. On the contrary, when the state is thought to be high (H) public investments can be positive ($G^{H\hat{H}} = 1$) or null ($G^{H\hat{L}} = 0$) with probabilities given respectively by equations (8) and (9).

an exemplification of the cognitive process see Figure 6). Using (8) and (9), we allow politicians to play the following mixed strategy in the stage game:

$$G_t^H = \begin{cases} 1 & \text{with} \quad 1-(1-\lambda_t)\eta_t^i \\ 0 & \text{with} \quad (1-\lambda_t)\eta_t^i \end{cases} \tag{15}$$

Thereby the current and the expected future individual output in state H will be

$$y_t^i = G_{t-1}^{\gamma}(e_t^i)^{\alpha} \tag{16}$$

$$\mathbb{E}_t(y_{t+1}^i) = q\Big[1-(1-\lambda_t)\eta_t^i\Big](e_{t+1}^i)^{\alpha} \tag{17}$$

Recursively, each citizen maximizes the expected current period return that will be consumed according to (6) and the agent i's intertemporal utility at time t is

$$\begin{aligned} V(e_t^i) = \max_{h_t^i} \Big\{ &(1-\tau)y_t^i - \frac{1}{\phi}(h_t^i)^{\phi} + \beta q\Big[\mathbb{E}_t\big[V^H(e_{t+1}^i)|\hat{\sigma}^i = H\big] + \\ &+ \mathbb{E}_t\big[V^H(e_{t+1}^i)|\hat{\sigma}^i = L\big]\Big] + \beta(1-q)\Big[\mathbb{E}_t\big[V^L(e_{t+1}^i)|\hat{\sigma}^i = H\big] \\ &+ \mathbb{E}_t\big[V^L(e_{t+1}^i)|\hat{\sigma}^i = L\big]\Big]\Big\} \end{aligned} \tag{18}$$

where income is given by (17) and (16) and private investments by (4). Note that for any $\beta > 0$ each agent cares both in today and tomorrow, and he would like to invest today for consuming tomorrow too. Maximization[25] gives the individual i's optimal investment

[25]To get analytical results we drop the learning effect by setting investment costs to be equal to $(h_t^i)^{\phi}/\phi$. Furthermore, to simplify notations we will keep all the results as functions of e_t^i, dropping G_{t-1}, φ_{t-1} and σ_t as arguments though both are state variables of the economy.

effort in each period t as a function of the state variable e_t^i and the ruler congruence rate λ_t.

PROPOSITION 2: *Citizens optimally respond to politicians' congruence rate by lowering private investments when λ_t decreases. The reaction is as strong as larger is the level of sophistication η_t^i:*

$$h_t^i(e_t^i; \lambda_t) = (1-\delta)\left[\frac{\Omega(1-(1-\lambda_t)\eta_t^i)}{1-\Omega(1-(1-\lambda_t)\eta_t^i)}\right]e_t^i \tag{19}$$

with $\Omega(q) \equiv 2\beta q(1-\tau) \leq 1$[26].

Consistently to the endogenous growth literature, Proposition 2 emphasizes the welfare enhancing role of public good provision and thereby the importance to select perfect agents, who always choose $\lambda_t = 1$, to rule public business. As soon as the congruence rate decreases, citizens optimally respond by lowering private investment. However, despite everyone faces the same technology (17), the extent of the reaction is heterogeneous. In the Appendix – where all the proofs are gathered – we show how $\partial^2 h_t^i/\partial\lambda_t\partial\eta_t^i \geq 0$, meaning that more sophisticated citizens react faster than naives. The former indeed collect all the information required to screen the ruler being aware of what is going on and therefore would be willing to pay more for a given increase in the congruence rate than the naive types.

Education helps them to be informed and sophisticated. However, as Figure 7 shows the effect of education on private investments is hill-shaped conditional on the politician behavior. Naive citizens always invest increasingly with the human capital stock e_t^i no matter the ruler's strategy unless $e_t^i = 0$ which forces unskilled agents to not invest due to a liquidity constraint for imperfect credit markets. As soon as more information is acquired – people who are more politically sophisticated moving on the right tail of the distribution –, citizens start to be aware of rulers' moves dropping investments if something wrong is thought to be done[27]. The identified cutoff $\tilde{e}_t(\lambda_t)$ is an increasing function of λ_t; as we showed more formally in the Appendix, a congruent political environment wipes out the implications of decreasing educational effect because, for

[26]The assumption of $\Omega(q) \leq 1$ is necessary to ensure the uniqueness of the solution and will be clarified in the Appendix.

[27]Although the absence of learning effects on the accumulation of education is unrealistic, it has the merit to emphasize the dropping effect (evenly unrealistically missing in other former models) which encourages more sophisticated citizens to not invest when politicians engage in *per se* rent seeking policies.

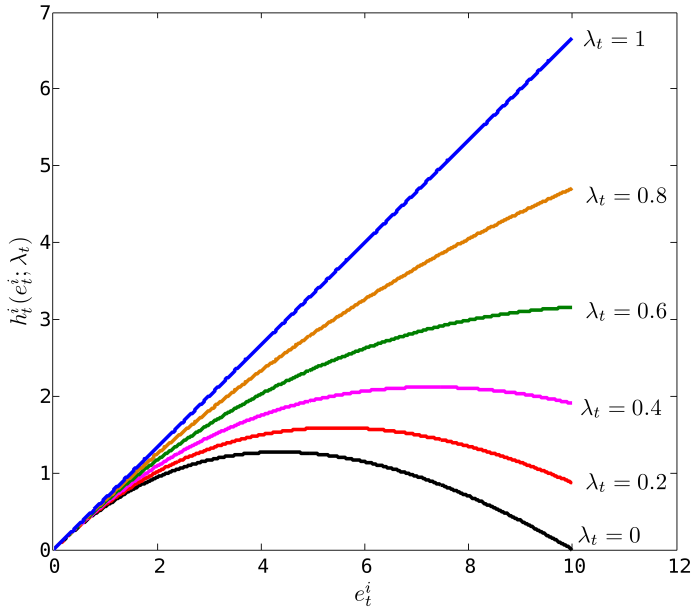


Figure 7: Private investments h_t^i as a function of human capital stock e_t^i for different values of ruler's congruence rate λ_t. Parameters are: $\delta = 0$, $\Omega = 0.4$, and $\bar{e} = 10$. The plots underlines the inelastic response of private investments with respect to politics for naive persons who lie on the left tail of the distribution.

$\lambda_t \to 1$, $\tilde{e}_t$ converges to $\bar{e}$ – the maximum level of the distribution – making investments increasing in human capital for the most skilled agents too. Institutions thus shape the agents' attitudes of investing but the most naive's. The latter always invest more and more even though a dishonest politician has been facing. In this sense, as we demonstrated in Section 2, the investment reaction of the naive citizens is inelastic with respect to politics: $\tilde{e}_t(0)$ is still positive so as investments are increasing for $e_t^i < \tilde{e}_t(0)$. This is consistent with a large literature dealing with political *extractive* institutions and economic incentives in developing countries (see among others Acemoglu, Johnson, and Robinson 2001), though heterogeneity and naiveté have never been investigated before.

Besides, complete comparative statics on the subgame equilibrium has been investigated:

(a) First of all, depreciation δ discourages agents to invest more given that much of it will be destroyed in future times ($\partial h_t^i / \partial \delta \leq 0$). Similarly taxation τ does, casting down the accumulation process ($\partial h_t^i / \partial \tau \leq 0$).

(b) On the other side, more optimistic agents clearly invest more ($\partial h_t^i / \partial q \geq 0$) and evenly do the most patient ones ($\beta \to 1$) in order to consume more in the future.

Once the private optimal investment has been characterized, the law of motion of education $e^i_{t+1}(e^i_t;\lambda_t)$, for every citizen i, is easily yielded by substituting (19) into (4):

$$e^i_{t+1}(e^i_t;\lambda_t) = \frac{1-\delta}{1-\Omega(1-(1-\lambda_t)\eta^i_t)} e^i_t \tag{20}$$

State equation (20) is an increasing concave function which implies that some mechanisms will lead all the citizens to a common steady state human capital level[28].

4.4 The political process

In section 4.3 we have demonstrated how the accumulation process and the wealth of a society depends on the institutions and on the extent to which the ruler is willing to be congruent (both summarized by λ). We now move on to consider the reverse interaction going through political mechanisms: how do sophisticated voters bind politician's attitude to be dissonant. This amounts to endogenize political choices as a best response of the sophistication rate of the electorate.

According to the timing of events depicted in section 3.4, in T2 the ruler anticipates what is the level of private investments made by each citizen i and chooses the optimal congruence rate $\lambda_t = P(\hat{\sigma}_t = H|\sigma_t = H)$ ranged according to (14) – the sophisticated electoral accountability constraint. If he is prone to be congruent to the announcement made in state H he will make the claimed investments. Otherwise, he will not carry any public investment out. More generally, we allow politicians to play the following mixed strategy in the stage game[29]:

$$G_t = \begin{cases} 1 & \text{with} \quad \lambda_t \\ 0 & \text{with} \quad 1-\lambda_t \end{cases} \tag{21}$$

so that the MPBE is obtained by solving, according to equation (3), the following recursive optimization problem:

$$\max_{\lambda_t} \quad V^r_t(\lambda_t) = T_t - qB\lambda_t + \beta\mathbb{E}_t\big[V^r_{t+1}(\lambda_t)\big] \quad \text{s.t. (14)} \tag{22}$$

[28]As we have argued, this is a straight consequence of the absence of learning effect (see footnote 25). Nevertheless, different common steady state human capital levels would be easily obtained allowing citizens to have different priors. In this fashion, more optimistic citizens (i.e. with q^i higher) would get in equilibrium more.

[29]Note that in equation (15) the same mixed strategy has been described from the citizens' point of view whom know what is the level of λ_t according to his own level of sophistication. That makes up the *guessed* probability about whether politician are telling a lie or the truth.

Therefore, political choices, undertaken at time t, shape both current rents, $T_t - qB\lambda_t$, and the income flow that is expected in $t+1$ from reelection. Since credit markets are imperfect, the government cannot spend more than what has been collected by taxing the electorate. It implies that the cost of the project $B \leq T_t$. In what follows, we just express B as a fraction $b \in [0,1]$ of the current tax revenues, i.e. $B = bT_t$, which in turn are equal to $T_t = \tau y_t$, that involve aggregated outcome level $y_t = \int_0^1 y_t^i di$. To keep things easy, in line with Glomm and Ravikumar (1992) and Benabou (1996, 2000, 2002), we suppose education to be initially distributed as a log-normal random variable with mean μ_0 and variance Δ_0^2, i.e. $\ln e_0^i \sim \mathcal{N}(\mu_0, \Delta_0^2)$. However, it is easy to note from (20) that the distribution of e_t^i, which is endogenous, remains log-normally distributed over time with mean μ_t and variance Δ_t^2. It follows from (16) that also income remains log-normally distributed over time with mean $m_t = 2\mu_t + \Delta_t^2$, i.e. $\ln y_t^i \sim \mathcal{N}(2\mu_t + \Delta_t^2, v_t^2)$.

At the same way, from (20) we obtain the difference equation which governs the evolution of the economy, i.e. the law of motion of the aggregate level of human capital:

$$\mu_{t+1} = \mu_t + \Delta_t^2/2 + (\Omega - \delta) - \Omega(1-\lambda_t)\eta_t \tag{23}$$

for small values of $\Omega(1-\lambda_t)\eta_t$ and with $\eta_t \equiv \exp(\mu_t + \Delta_t^2/2)/\bar{e}$. Substituting (23) into (17) yields the expected output in $t+1$ of the economy:

$$m_{t+1} = m_t + 2(\Omega - \delta) - (1+2\Omega)(1-\lambda_t)\eta_t \tag{24}$$

The first two terms describe, respectively, the positive effect of the initial condition of the economy and of exogenous parameters that feature preferences (β), beliefs (q), policies ($-\tau$), and the obsoleteness of the capital stock ($-\delta$). The social cost of cheating, in terms of future income loss, is instead showed in the last term. In particular, equation (24) makes clear how dissonant politicians, that always play low values of λ_t, are detrimental to citizens reducing future wealth. Interestingly, the social cost increases with Ω and μ_t by pushing politicians' incentive to draw more rents.

The incumbent's expected rents from being reelected are then equal to:

$$\ln V_t^r(\lambda_t) = \max_{\lambda_t} \Big\{ \ln \tau + 2\mu_t + \Delta_t^2 - qb\lambda_t + \beta \mathbb{E}_t \big[\ln V_{t+1}^r(\lambda_t) \big] \Big\} \quad \text{s.t. (14)} \tag{25}$$

Maximizing rents amounts to choose an optimal rate of congruence λ_t, ranged according to (14) – the sophisticated electoral accountability constraint. By doing that the incumbent trades off expected future tax revenues with current rents coming from smaller public investment. Due to the functional form of rents, it is easy to note that $\partial \ln V_t^r / \partial \lambda_t \leq 0$ iff

$$b \geq \frac{\beta(1+2\Omega)\eta_t}{q} \equiv \underline{b}(\beta, q, \tau, \eta_t) \tag{26}$$

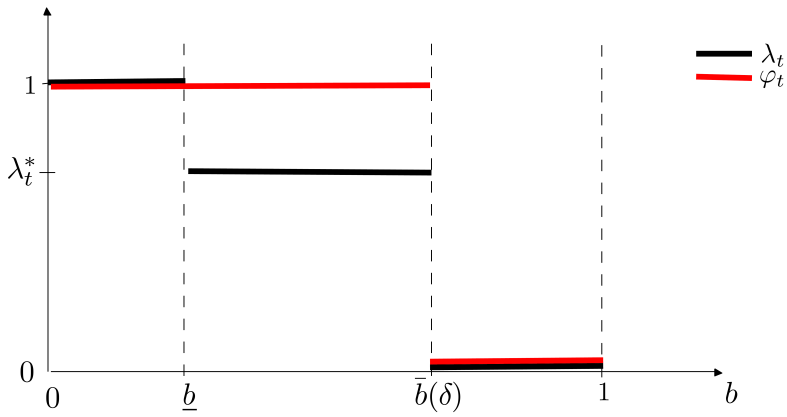


Figure 8: Political Equilibrium as a function of b.

with $\eta_t \equiv \exp(\mu_t + \Delta_t^2/2)/\bar{e}$. In other words, rents are found to be decreasing with his congruence rate provided that the cost of the public investment, relative to tax revenues, is high enough. The threshold $\underline{b}(\beta, q, \tau, \eta_t)$ is found to be increasing with β, q, and $-\tau$, meaning that better economic conditions reduce the incumbent's incentives to behave dissonantly. So does a more sophisticated electorate. When condition (26) holds, incumbent rulers maximize office rents by pushing down λ_t as much as they can. However, given the accountability effort exerted by voters, the lowest still optimal value is the maximum between $\lambda_t^* = 1 - 1/\eta_t^m((1-q)/q)$ and zero, according to (14). We summarize this result in Proposition 3.

Proposition 3: *Political equilibrium. Assume* $\ln e_0^i \sim \mathcal{N}(\mu_0, \Delta_0^2)$. *There exists* $\underline{b}$ *and* $\bar{b}(\eta_t^m)$, *decreasing in* η_t^m, *with* $0 \leq \underline{b} \leq \bar{b}(\eta_t^m) \leq 1$, *such that:*

(i) If $b < \underline{b}$ *the incumbent plays* $\lambda_t = 1$ *and the median voter plays* $\varphi_t = 1$.

(ii) If $b \in [\underline{b}, \bar{b}(\delta)]$ *the incumbent plays* $\lambda_t = \lambda_t^*(\eta_t^m)$ *and the median voter plays* $\varphi_t = 1$.

(iii) If $b > \bar{b}(\delta)$ *the incumbent plays* $\lambda_t = 0$ *and the median voter plays* $\varphi_t = 0$ *(go-for-broke).*

When public investments are costly enough it is optimal for rents-maximizing policymakers to set the rate of congruence λ_t to be the lowest possible value. However, due to accountability effort, the optimal congruence rate is:

$$\lambda_t^* = 1 - \frac{1}{\eta_t^m}\left(\frac{1-q}{q}\right) \tag{27}$$

that is increasing both in q and in the overall level of sophistication of the society. In particular, it is worth to note that for every $\eta_t^m \leq (1-q)/q$ the optimal strategy in the stage game for the incumbent is to be fully dissonant (i.e. $\lambda_t^* = 0$).

As Figure 8 shows, the cost of the public good drives political decisions for a given distribution of education in the society. When b is very small (i.e. $b < \underline{b}$) there is no incentive to cheat the electorate because investing increases future rents more than what he would have obtained today by choosing $G_t = 0$. However, the cost of the project, while feasible (i.e. $b \leq 1$), could be so high to incentive the ruler to go-for-broke. If $b > \bar{b}$ playing go-for-broke by extracting all the tax revenues strictly dominates $\lambda_t = \lambda_t^*$ (and *a fortiori* any $\lambda_t > \lambda_t^*$). The median voter anticipates that for every $b > \bar{b}$ the incumbent behaves dissonantly (i.e. $\lambda_t < \lambda_t^*$) assigning probability zero on the event that the ruler plays a congruence rate greater than λ_t^*. Consequently, the median voter plays $\varphi_t = 0$ and the ruler, that anticipates this move, goes-for-broke.

In the Appendix we also show that $\bar{b}(\delta)$ is a decreasing function of the depreciation rate of the human capital of the producers so as for high level of δ go-for-broke is more likely to be the optimal strategy in the stage game. In particular, there exists a threshold δ^* such that for any $\delta < \delta^*$ going-for-broke is not an admissible strategy for the ruler, i.e. $\bar{b}(\delta) > 1$.

In what follows we characterize and compare the first two regions where $b < \underline{b}$ and where $b \in [\underline{b}, \bar{b}(\delta)]$. We show how – for any distribution of education, $F(e_t^i)$, and for every prior $q < 1$ – the incentive of the ruler to behave dissonantly implies a loss for every citizen. We then show, in Section 4.6, how this loss is higher in more unequal societies.

4.5 The characterization of the MPBE

4.5.1 The MPBE in $b < \underline{b}$ (the baseline framework)

The region where $b < \underline{b}$ stands for the baseline/standard framework: the cost of the public project is small so as the ruler has no incentive to cheat the electorate. As a results, the ruler acts as a perfect agent and the principal – the median voter – retains the 'high quality' politician. In other words, $\lambda_t = 1$ and $\varphi_t = 1$.

This political equilibrium gives to each citizens the following investment effort in human capital

$$h_t^i(e_t^i) = (1-\delta)\frac{\Omega}{1-\Omega}e_t^i \tag{28}$$

and stock

$$e^i_{t+1}(e^i_t) = \frac{1-\delta}{1-\Omega} e^i_t, \tag{29}$$

where both are obtained by substituting $\lambda_t = 1$ into (19) and (20) respectively.

As in any standard endogenous growth model, both the optimal investment level and the human capital stock depend on the initial condition from where the accumulation had its start and, additionally, from exogenous parameters such as the level of depreciation rate δ (negatively), the discount rate of future returns β (positively), and taxation τ (negatively). Finally, since the economy proceeds through a stochastic environment, also beliefs q over the state of the world are involved in the accumulation process of each individual.

4.5.2 The MPBE in $b \in [\underline{b}, \bar{b}(\delta)]$ (Political-agency framework with potentially dissonant politicians)

In the region characterized by a high cost of the public project – i.e. where $b \in [\underline{b}, \bar{b}(\delta)]$ – the model features a political-agent relationship with potentially dissonant politicians, that is with politicians that might play $\lambda_t < 1$ according to the general level of sophistication of the society, η^m_t.

Substituting (27) into (19) and (20) yields the optimal individual level of investment and stock of human capital under potentially dissonant politicians:

$$h^i_t(e^i_t; q, e^m_t) = (1-\delta)\left[\frac{\Omega(1 - \frac{e^i_t}{e^m_t}\frac{1-q}{q})}{1-\Omega(1 - \frac{e^i_t}{e^m_t}\frac{1-q}{q})}\right] e^i_t \tag{30}$$

$$e^i_{t+1}(e^i_t; q, e^m_t) = \frac{1-\delta}{1-\Omega(1 - \frac{e^i_t}{e^m_t}\frac{1-q}{q})} e^i_t \tag{31}$$

All the MPBE values not only depend on the own level of education – and on exogenous parameters that additionally affect (28) and (29) – but also on the general level of sophistication of the society, e^m_t. This positive external effect works through the electoral accountability effort pushed by the median citizen, that makes more likely a fair political environment with higher investment returns. This is in line with what Bidner and Francois (2013) define dynamic complementarity between the willingness to vote out today's transgressing leader with a higher expectation that citizens will vote out future transgressors. In our model, dynamic complementarity among voters intertemporal strategies naturally emerge: sophisticated societies force indeed rulers to invest more in public goods making them more sophisticated in the future.

Beliefs also play a crucial role, and they are updated according to (10); in particular, a strong belief in favor of state L will bring citizens to not carry any investment out. Combining the two effects shows that agents are willing to produce and invest iff $e_t^i \leq (q/1-q)e_t^m$, that is very likely to occur for high values of q (for $q \to 1$ the right hand side diverges to $+\infty$ but a $q = 0$ leads agents to inactivity) and e_t^m.

It is straightforward to see that the level of investments in human capital is lower under dissonant politicians than in the baseline framework and so is the human capital stock in (31). This is true in every economies and for every citizens. In the economies that are expected to be persistently high performed (with $q = 1$) – where politicians always prefer to be fully congruent – and for unskilled citizens (with $e_t^i = 0$) the two contingencies are equivalent.

Below, we characterize the MBPE in Proposition 4:

Proposition 4: *Assume* $\ln e_0^i \sim \mathcal{N}(\mu_0, \Delta_0^2)$*, there exists a unique Markov Bayesian Perfect Equilibrium such that*

(i) If the state is low ($\sigma_t = L$) the incumbent ruler does not carry out any public investment ($G_t = 0$). Citizens know that and respond by not investing ($h_t^i = 0$) regardless of the education they have and by retaining the incumbent ($\varphi_t = 1$); future stocks of human capital are only driven by past level and depreciation (i.e. $e_{t+1}^i = (1-\delta)e_t^i$).

(ii) If the state is high ($\sigma_t = H$) and if the cost of the public good is sufficiently high, the incumbent ruler has the incentive to cheat citizens that in turn bind politicians via electoral accountability. The ruler cares to be reappointed and the optimal strategy is given by (27) when condition (26) holds so as citizens relect him ($\varphi_t = 1$). Citizens, on the other side, guess the ruler does not invest ($G_t = 0$) with probability $(1-\lambda_t)\eta_t^i$ and respond by investing and accumulating according to (30) and (31), respectively.

4.6 Equilibrium payoffs and concerns for inequality

State H is the most interesting one, conveying all the insights here presented. Given the optimal strategies, in equilibrium, payoffs are given by future income and rents to citizens and politicians, that in state H are respectively:

$$m_{t+1} = 2\mu_t + \Delta_t^2 + 2(\Omega - \delta) - (1 + 2\Omega)\frac{1-q}{q}\exp\left(\frac{\Delta_t^2}{2}\right) \tag{32}$$

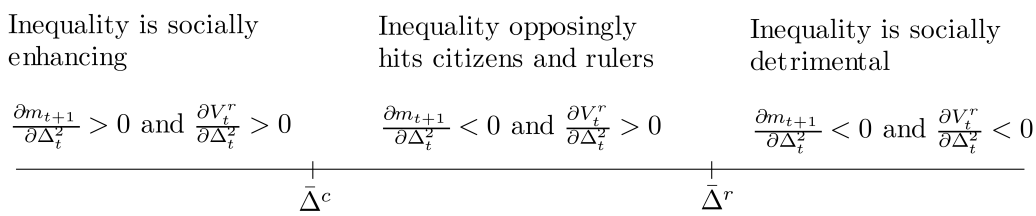


Figure 9: Heterogeneous effects of inequality upon future income and politician rents.

$$\begin{aligned}\mathcal{V}_t^r \equiv \ln V_t^r = (1+\beta)\ln\tau + 2(1+\beta)\mu_t + (1+\beta)\Delta_t^2 - qb + b(1-q)\frac{1}{\eta_t^m} \\ + \beta\ln q + 2\beta(\Omega-\delta) - \beta(1+2\Omega)(1-q)\exp\left(\frac{\Delta_t^2}{2}\right)\end{aligned} \tag{33}$$

where (32) stands for the average producer. Note that both citizens' and ruler's payoffs depend on peculiar features of the distribution of human capital, $\ln(e_t/e_t^m) = \Delta_t^2/2$. In particular, Δ_t^2 describes the extent to which human capital is unequally distributed among different citizens and, limited to the case of the log-normal distribution, it increases with the mean but declines with the median level of education owned by the political pivotal citizen.

The global effect of inequality on both income and rents turns to be non linear and hill-shaped, meaning that little inequality is tolerated by citizens. The levels of inequality tolerated by citizens, however, are smaller than those preferred by politicians and higher levels hit opposingly citizens and politicians, and only the latter benefit for that. The idea is illustrated in Figure 9, where $\bar{\Delta}^c$ and $\bar{\Delta}^r$ are respectively the bliss points of citizens and rulers, with $\bar{\Delta}^c < \bar{\Delta}^r$. We collect these results in Proposition 5 established in the Appendix:

Proposition 5: *Let $\ln e_0^i \sim \mathcal{N}(\mu_0, \Delta_0^2)$, there exists $\bar{\Delta}^c$ and $\bar{\Delta}^r$ with $0 < \bar{\Delta}^c < \bar{\Delta}^r$, such that:*

(i) for each $\Delta_t^2 \in [\bar{\Delta}^c, \bar{\Delta}^r]$ future income declines with inequality Δ_t^2 such that, given the accountability effort, citizens, on average, worst off. Conversely, inequality increases ruler's rents, manipulating poor and extracting rents from taxes of the wealthiest.

(ii) for $\Delta_t^2 < \bar{\Delta}^c$ ($\Delta_t^2 > \bar{\Delta}^r$) both citizens and politicians better (worst) off with inequality.

The role of inequality can be easily interpreted into our framework. Positive skewed and unequal (right-tailed) distributions characterize societies with most naive agents and scarcely sophisticated citizens. Naive agents always invest in human capital and vote for incumbents whatever the latter do. Sophisticated agents, conversely, invest more but under the condition that the incumbent invest too. In this case, they do not reappoint the incumbent either. Because of the positive externalities generated, via electoral accountability, by the median voter, citizens give a positive weight to the median and henceforth to equality, on the region on the right of $\bar{\Delta}^c$, contributing to explain why people dislike living in a society which is too unequal, beside altruism, or aversion to social tension, crime, or civil wars. Pushing down the median value of the distribution instead permits the ruler being more independent from the electorate control. That explains why inequality, on the region on the left of $\bar{\Delta}^r$, is strictly preferred by the rulers. On the other hand, citizens' production is required to bring up tax revenues that, in turns, mainly constitute rents. Too unequal societies, like most developing countries, fail in accumulate human capital that mainly constitutes tax revenues, explaining why political rents are hill-shaped with respect to inequality. The interesting social conflict that characterizes well-functioning democracies, with intermediate inequality (i.e. $\Delta_t^2 \in [\bar{\Delta}^c, \bar{\Delta}^r]$), is resolved when inequality is too low or high, and only in the former case it is found to be socially enhancing, since either the median is already high to avoid electorate manipulation or the mean is low, such that more accumulation is wished for.

5 Dynamics and multiple steady states

The initial distribution of human capital strongly shapes the dynamics of the economy, political choices, and agents' incentives. Right tailed distributions are mostly composed by naive agents who invest a small but positive amount in human capital and, at the same time, barely account for ruler's duties that, in turn, are allowed to extradraw private rents (by playing small values of λ_t). Conversely, we found that rulers are constrained by more sophisticated societies that impose high values of λ_t as the price for reappointment.

The model thus predicts multiple steady-states, one for sophisticated societies with congruent politicians in charge and one for naive societies ruled by dissonant politicians, and we found the median agent to be pivotal in determining the dynamics of the whole society. To demonstrate that we need to solve the following recursive dynamical system

which describes the joint evolution of education and policy:

$$\begin{cases} e^i_{t+1} = \chi(e^i_t; \lambda_t, e^m_t) \\ \lambda_t = \Lambda(e^m_t) \end{cases} \tag{34}$$

The solution, at the intersection of the two loci, gives us the long-run human capital level of equilibrium:

$$e_\infty = \left(1 - \frac{\delta}{\Omega}\right)\left(\frac{q}{1-q}\right)e^m_\infty \tag{35}$$

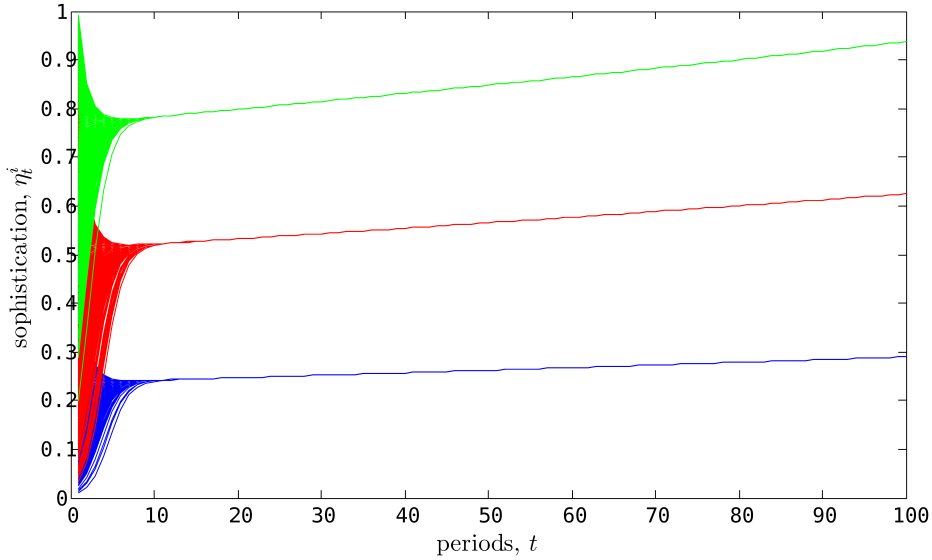
where, by equation (20), e^m_∞ is a function of the initial condition e^m_0, that is $e^m_\infty(e^m_0)$. The model thus predicts convergence of all agents – but the unskilled ones, with $e^i_t = 0$ – to a common educational value, that increases with the median and the prior q. All the parameters that compose Ω (β, q, and $(1-\tau)$) contribute to push up e_∞, whereas it decreases with depreciation. That means that initially less skewed societies, i.e. with higher e^m_0, will be more educated and richer in the long run (see Figure 10a)[30].

However, as illustrated in the two panels of Figure 10, two societies with the same initial conditions, and in particular with the same initial distribution of education at time $t = 0$, can nonetheless be driven toward two different steady-states. A consistent raise in depreciation or taxation can in fact undermine the effect of optimism (q) and faith in the future (β) persistently discouraging agents to invest to push down society to a zero-level educational state.

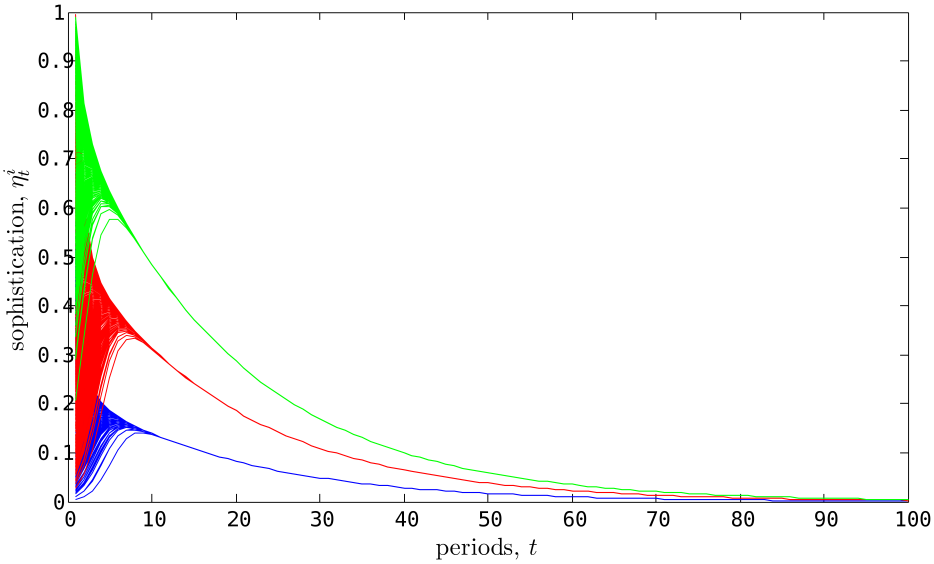
The role of the median is then pivotal in that. We show that if $q < \bar{q}(\beta, \delta, \tau)$ the median agent takes a decreasing trajectory and the rest of the society will do the same, firstly the more sophisticated agents and, at last, the naive ones[31]. In this sense, the interaction between political sophistication and institutions acts as a centripetal force with respect to the accumulation process through time leading to a degenerate ergodic distribution in the steady-state. In fact, if the median decreases over time makes easy the politician to push down λ, that in turn discourage private agents to carry out any investments, firstly sophisticated agents – that are the first to be aware of the deteriorated political environment – and then the naive ones (see Figure 10b). More general results are collected in Proposition 6.

[30] In the region where $b < \underline{b}$ – the standard case where politicians act as perfect agents – the dynamics of the system do not depend on endogenous factors such as the general level of sophistication, e^m_t, but only on exogenous parameters, i.e. on q, β, τ, and δ. Arguably, $e_t \to \infty$ if $\delta < \Omega$; $e_t \to 0$ if $\delta > \Omega$. Finally, if $\delta = \Omega$ the system is characterized by a continuum of steady states.

[31] It turns out that $\bar{q}(\beta, \delta, \tau)$ is a U-shaped function of β, bell-shaped in τ and linearly increasing in δ. A society is then less likely to be pushed down to $e_\infty = 0$ for high values of β, small taxes and depreciation.



(a) *Multiple steady-states* ($q \geq \bar{q}(\beta, \delta, \tau)$).



(b) *A unique steady-state* ($q < \bar{q}(\beta, \delta, \tau)$).

Figure 10: Different initial distributions of education follow different paths in panel (a) whereas converge to a unique steady-state equilibrium ($e_\infty = 0$) in panel (b). Each line describes the trajectory of an individual $i = 1, \ldots, 10^3$ for $t = 1, \ldots, 100$. Green lines draw a negative skewed distribution ($\eta_0 \sim Beta(6, 2)$), red lines a symmetric one ($\eta_0 \sim Beta(2, 2)$), and blue lines a positive skewed one ($\eta_0 \sim Beta(2, 6)$). Same colors describe same distributions at $t = 1$ among panel (a) and (b). Note that convergence is caught up firstly by the sophisticated agents (that for first know what is going on) whereas the last to catch up are the most naive citizens.

Proposition 6: *When citizens are pessimists enough, i.e. if $q < \bar{q}(\beta, \delta, \tau)$, (or equivalently when the depreciation rate is above a critical value $\bar{\delta}(\beta, q, \tau)$), society converges to a zero-level of education, no matter initial conditions. Conversely, if $q = \bar{q}(\beta, \delta, \tau)$ multiple stable steady-states arise ($e_\infty = e_0^m$), and rulers are more congruent in societies with higher initial educational achievements. Finally, if $q > \bar{q}(\beta, \delta, \tau)$ society gets richer over time and the speed of the human capital growth is determined by the initial distribution of education.*

Where multiple steady-states occur *history matters* (Bénabou, 2000). Temporary shocks to the state variables of the economy can permanently move the overall society from one equilibrium to another or produce persistent effects on the system only recovered after several lags. Despite idiosyncratic shocks can be neglected, as we are interested in the whole distribution of citizens, aggregate country-specific shocks must be taking into account.

In what follows I briefly discuss the (confounding) effect of several aggregate shocks on the political mechanism outlined in the model. I distinguish among shocks that hit the distribution of human capital – even the ones involving a small segment of the distribution – and the ones hitting the state of the world, σ_t. Among the first ones are:

- *immigration:* though immigration is a highly perceived phenomenon, that media and politics use in order to influence political beliefs of the citizenship, it is generally a limited phenomenon, as the number of new comers is humble with respect to the number of the whole population. Adding to that is the fact that obtaining the citizenship – and henceforth the voting right – is, in several cases, a long process. In view of that immigration is expected to have a modest effect on the general level of sophistication of the society and on the political equilibrium, moving the median voter only slightly on the left – if the new comers are on a great extent poorly educated – or on the right – if the immigrants are more educated than the citizens of the country that welcomes them, as generally is the high skilled workers migration.

- *IT technology:* a new technology is expected to produce ambiguous effects and certainly non-linear. Despite the direct effect of technology on the screening process plays a role if media are not able to convey political signals, it is not determinant for the paper story. Conversely, a new technology may improve the learning process making the electorate more sophisticated (indirect effect). On the other hand, an excessive exposition to IT technology can divert the attention of students to the contents of the teachings, making the effect hill-shaped.

- *social discrimination:* the exposition to the IT technologies and, more generally, to the physical capital required to learn and get more sophisticated depends on social discrimination. Social discrimination can therefore enlarge the segment of population binded by a liquidity constraint, those with $e_t^i = 0$, and, in view of imperfect credit markets, it is expected to produce persistent poverty and political naivete.

Beside shocks to the distribution of education, there are several shocks on the state of the world σ_t that are expected to perturb the system along its pathway. For example, a financial crisis can make the fiscal budget constraint tighter so as to impede new productive public investments. Anticipating that, citizens reduce human capital investments. If the crisis is persistent enough, the resulting lack of investment will lead the society to loose political sophistication – the median citizen moves left –, giving politicians more chances to behave dissonantly.

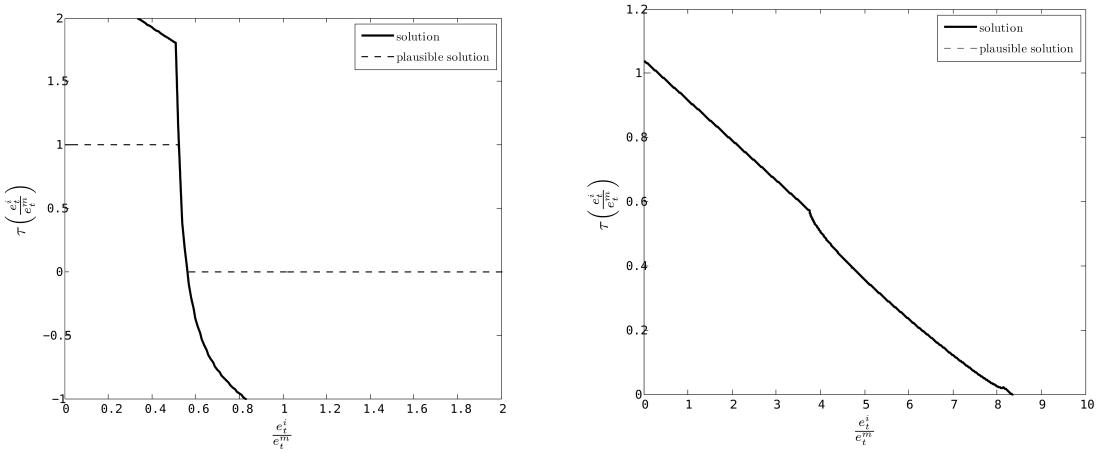
This mechanism contributes to explain several historical events such as the advent of totalitarian regimes in the aftermath of the World War I as the penalties imposed to Germany by the treaty of Versailles – what John Maynard Keynes defined a Carthaginian peace[32] – can be seen as a huge (persistent enough) fiscal shock that for many years constrained σ_t to L.

6 Endogenous fiscal choices

In a framework á la Ferejhon (1986) the only instrument that citizens/voters have to punish a bad politician is to vote him out by replacing him with a challenger. Voters do not decide the most preferred tax rate but whether to retain the incumbent or not, on the basis of unobservable past actions undertaken by the incumbent. The only unobservable past action that incumbents control, in this work, is their attitude to signal the right state of the world – namely, the congruence rate λ_t – which in turn determines the choice of implementing a productive public project, G_t, that entails costs B. All these decisions take the tax rate τ as given.

We now extend the model to allow citizens to decide the optimal fiscal rate, that so far has been held exogenous. A reduced form framework á la Meltzer and Richard (1981) is incorporated into the main model: each citizen has a preferred tax rate, as a function of the own level of political sophistication, but eventually is the median, by majority rule, the agent who decides the size of the government.

[32] See '*The Economic Consequences of the Peace*' (1919).



(a) $q = 0.25$. (b) $q = 0.75$.

Figure 11: Optimal taxation as a function of political sophistication. We set parameters $\beta = 1$ and $\delta = 0$. In panel (a), a pessimistic scenario is depicted with $q = 1/4$. Here, only naive citizens with $e_t^i \leq (1/3)e_t^m$ would be glad to contribute to a public project that is unlikely to be done. τ_t^* is not a maximum for $e_t^i \in ((1/3)e_t^m, (9/10)e_t^m)$, but a corner solution has been characterized in the Appendix where these citizens always claim for $\tau_t^* = 0$. Finally, in panel (b) a more optimistic scenario is depicted with $q = 3/4$ and accordingly more sophisticated citizens are more inclined to participate to the public project.

Fiscal choices are trivial in state $\sigma_t = L$ and the most preferred rate is $\tau_t = 0$. Conversely, in state H redistribution is gladly accepted and necessary to provide investment G_t in a productive public good. We then assume that before the ruler announces his decision to carry out public investment in time T2, citizens are called to decide the most preferred level of redistribution, $\tau_t^i(e_t^i, e_t^m)$. By majority rule, $\tau_t^m(e_t^m) \equiv \tau_t^*$ is then applied. All the other timing is unchanged.

The individual optimal tax rate is obtaining by solving the MPBE by backward induction. Each citizen i, endowed with e_t^i in time t, anticipates which is the optimal congruence rate of the ruler, λ_t^*, and accordingly solves the following quadratic maximization problem, obtained by substituting (27), (31), and (30) into (18):

$$\begin{aligned}\mathcal{V}^*(\tau_t) = \max_{\tau_t} \Big\{ &(1-\tau_t)(e_t^i)^2 - \frac{1}{2}h_t^i(\tau_t, \lambda_t^*) + \beta q \Big[\mathbb{E}_t\big[V^H(e_{t+1}^i(\tau_t, \lambda_t^*))|\hat{\sigma}^i = H\big] + \\ &+ \mathbb{E}_t\big[V^H(e_{t+1}^i(\tau_t, \lambda_t^*))|\hat{\sigma}^i = L\big]\Big] + \beta(1-q)\Big[\mathbb{E}_t\big[V^L(e_{t+1}^i(\tau_t, \lambda_t^*))|\hat{\sigma}^i = H\big] + \\ &+ \mathbb{E}_t\big[V^L(e_{t+1}^i(\tau_t, \lambda_t^*))|\hat{\sigma}^i = L\big]\Big]\Big\}.\end{aligned} \tag{36}$$

In Figure 11 we show the solution of the maximization problem (36), $\tau_t(e_t^i, e_t^m)$, for $\beta = 1$ and $\delta = 0$. In panel (a) we plot $\tau_t(e_t^i, e_t^m; q = 0.25)$: such restriction describes

a pessimistic scenario where despite human capital does not depreciate over time and citizens give the greatest importance and weight to future payoffs, the state H is really unlikely to occur. In such scenario, only naive citizens would claim more taxes for doing the public project. The rest of them, and in particular citizens with $e_t^i \geq (1/3)e_t^m$, would like to not contribute to a project that they anticipate will never be done.

However, as we pointed out in Section 4.4, the optimal congruence rate $\lambda_t^*(q)$ increases with q. Sophisticated citizens anticipate that and, accordingly, choose the optimal tax rate that increases with q. In Figure 11(b) the optimal taxation is depicted for $q = 0.75$ as a function of the individual relative level of sophistication, e_t^i/e_t^m. Though naive citizens still would like to contribute at a higher rate to the public project, sophisticated ones anticipate that the politician in office might be a congruent type and therefore increase τ_t^* with respect to the former case. Finally, when $q = 1$ all the citizens, regardless of the own relative individual level of sophistication, vote for the highest contribution to the project.

We collect this result in Proposition 7:

Proposition 7: *Citizens punish politicians by choosing the optimal contribution rate τ_t^* according to their own relative individual level of sophistication e_t^i/e_t^m. When they expect politicians to cheat they react by lowering the contribution rate, provided they are sophisticated enough.*

7 Conclusions

This paper discusses the importance of education for the success of democracy, as a cognitive tool that citizens/voters can use to decode the information content of political signal and to keep rulers in charge accountable. Remarkably, productivity and citizens welfare may increase with education, via sophisticated electoral accountability, even if education has no direct effect on productivity, via human capital accumulation. This second nexus, tough realistic, is not necessary for the paper story but it only strengthens results by defining virtuous or wicked political paths, depending on the initial distribution of sophistication in the society.

The paper originally gives support in favor of the political sophistication channel – linking education and democracy – proposed by many researches in the political psychology field. In the empirical part we estimate inelastic political beliefs with respect to the quality of political institutions for low educated individuals, whereas elastic beliefs are estimated for high educated individuals. This channel represents the strongest

assumption of the theoretical model based on the premise that not all political actors are fully Bayesian but some of them commit mistakes on the basis of his own level of education (or political sophistication).

The model is only partially consistent with the modernization theory. On the one hand, education is found to be crucial in shaping democratic institutions via sophisticated electoral accountability. More educated societies are indeed more able to punish politicians that, in turn, invest more in infrastructure, roads or legal rules for contracts enforcement. These productive public goods foster private investment in education making future accountability more sophisticated and more effective over time. On the other hand, however, initially low educated societies fail in providing democratic institutions, and, even worst, bad governments are found to be persistent due to a persistent low level of electoral accountability. This point turns to be crucial in young democracies where voters are (on a median level) poorly educated, giving rooms to rent-maximizer politicians to easily cheat them.

The responsiveness of the politicians – captured in the model by the congruence rate –, and more generally of political institutions, to the general level of sophistication of the society frames the democratization process as an *endogenous* process, hardly imposed from outer forces.

A Supporting Details

A.1 Evolution of beliefs upon the state of the world

Let $\xi = P(\sigma_{t+1}|\sigma_t)$ the persistence rate of the Markov process. The evolution of beliefs upon the state of the world is then given by

$$q_{t+1} = P(\sigma_{t+1} = H) = \xi q_t + (1-\xi)(1-q_t).$$

It therefore depends on ξ (see Figure 12) for a given level of q_t. One may see that $\xi = 1$ generates a Bernoulli scheme that is a process characterized by fully persistent beliefs, i.e. $P(\sigma_{t+1}) = P(\sigma_t)$. For $\xi = 0.5$ the process is a random walk where future beliefs upon the state of the world are independent from past beliefs but for $\xi = 0$ citizens behave fully irrationally guessing that the state of the world changes time by time. Henceforth, for $\xi < 0.5$ the process is behavioral whereas for $\xi = 0.5$ citizens are totally incapable to make any prediction given q_t[33]. To keep things easy we allow $\xi = 1$ throughout all the paper so as $q_{t+1} = q_t = q$.

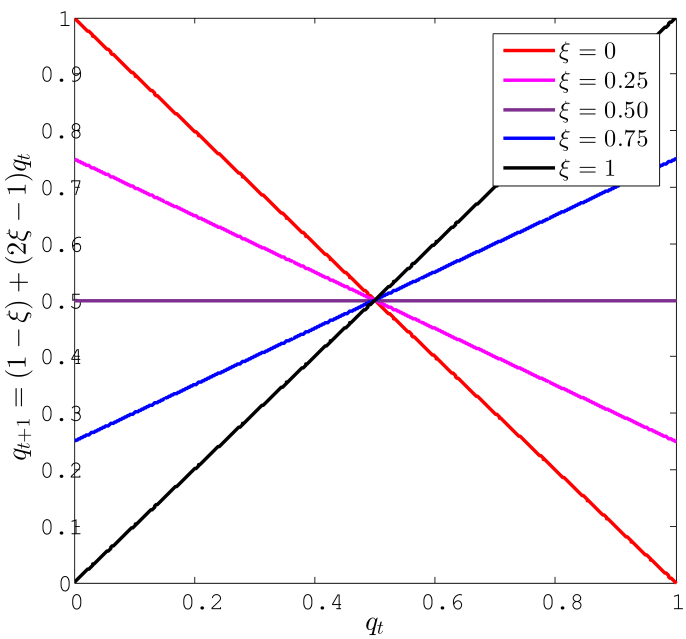


Figure 12: Law of motion of $q_{t+1}(q_t)$ for different values of ξ .

[33]That may be a good model for society characterized by high uncertain about future state of the world.

A.2 Beliefs over politicians in differently politically sophisticated societies

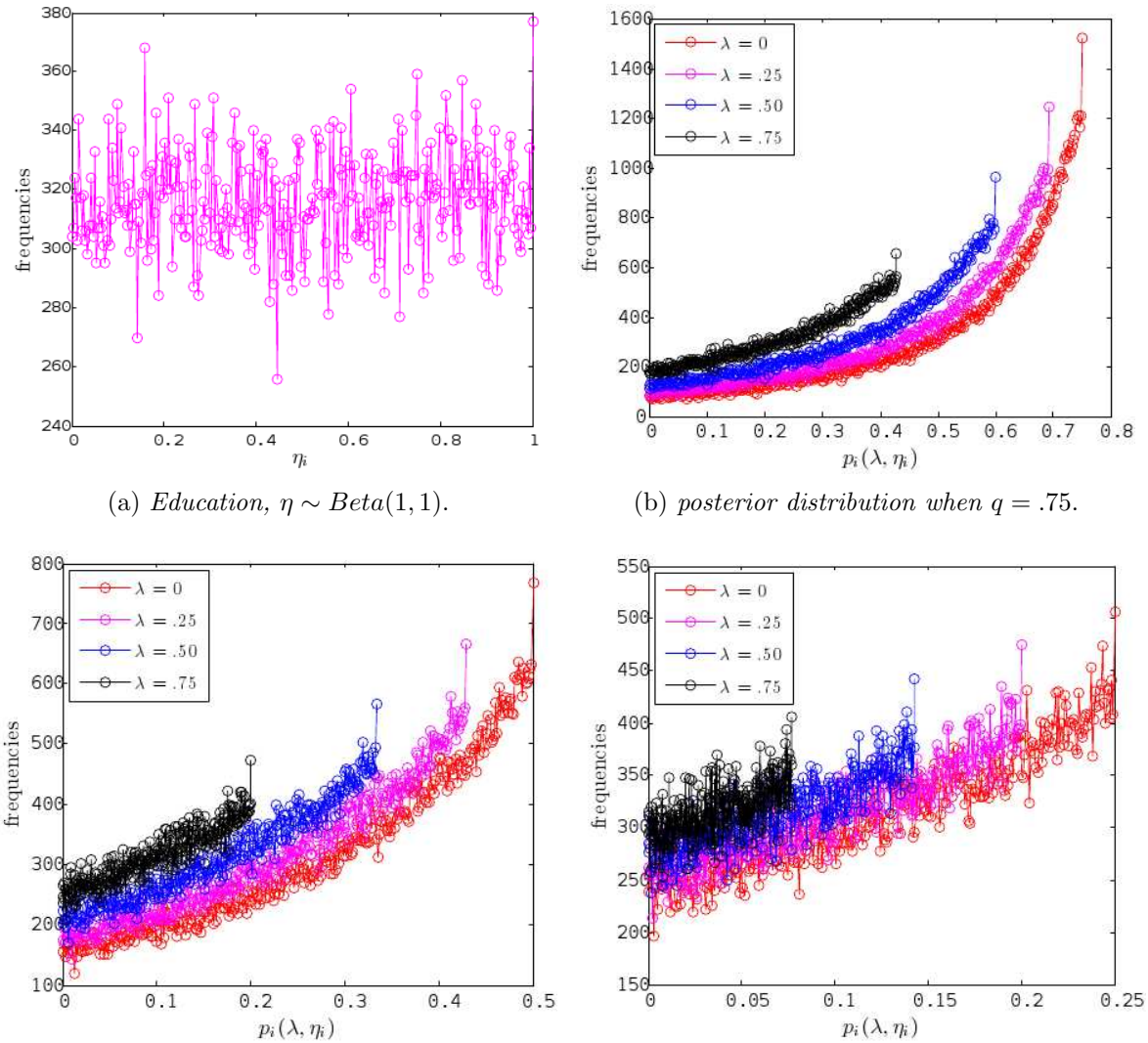
In this Section we show different posterior distributions originated from different *types* of societies, that is societies characterized by different distribution of political sophistication η^i. We start considering a society where naive citizens are of the same numerosity than sophisticated citizens. We then move to one where, though education distribute symmetrically, fully naive and sophisticated citizens are unlikely to occur. Finally we show Bayesian inference in very different societies: one in which the distribution is positive skewed – with more naive citizens than sophisticated ones – and one in which the distribution is negative skewed – a vey sophisticated society.

For every initial distribution, we compute posteriors $p_t^i(\lambda_t, \eta_t^i, q)$, for different values of q – the prior over the state of the world – and λ_t, using the awareness-management model stated in (10). From $f(p_t^i)$ we can then compute the voting equilibrium – as stated in Proposition 1 – that does not require the assumption of lognormality.

This Section aims at showing how different societies, with a different distribution of citizens in term of political sophistication, $f(p_t^i)$, can feature different electoral outcomes despite the high state of the world is equally likely to occur (q) and politicians show the same attitude to lie (λ_t). In Figure 16 – where sophisticated citizens are more numerous than naives – more people are in favor of replacing the incumbent who lies at the rate λ_t, because they are more politically aware. In other words, $p_t^m > 1/2$ is more likely to occur in a society where education is drawn from a negative skewed distribution, i.e. where $\eta_t^i \sim Beta(4, 2)$, than in one where education is drawn from a positive skewed distribution, i.e. where $\eta_t^i \sim Beta(2, 4)$, as the one depicted in Figure 15.

Secondly, we show here that the model is fully able to replicate empirical findings discussed in Section 2: $f(p_t^i)$ is crucially shaped by the extent to which individuals are politically naive with inelastic responses with respect to the quality of political institutions – here approximated by λ_t. Bayesian inference then establishes scale effects in view of the fact the part of the decision makers make mistakes in the inferential process. Naive citizens namely remain stuck in the left tail of the distribution $f(p_t^i)$ even when politically institutions are of a bad quality – i.e. $\lambda_t = 0$, the case depicted in red in panels (b), (c), and (d) in Figures 13-16. Sophisticated citizens conversely move according to λ_t.

A.2.1 Bayesian inference in a society of $n = 10^5$ citizens where education distributes uniformly



(a) *Education*, $\eta \sim Beta(1, 1)$.

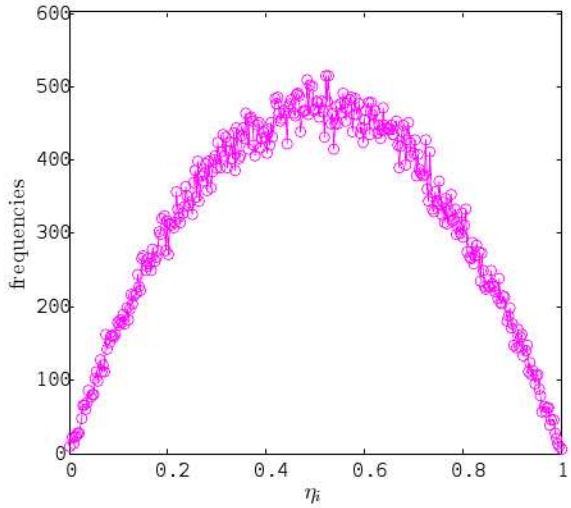
(b) *posterior distribution when* $q = .75$.

(c) *posterior distribution when* $q = .50$.

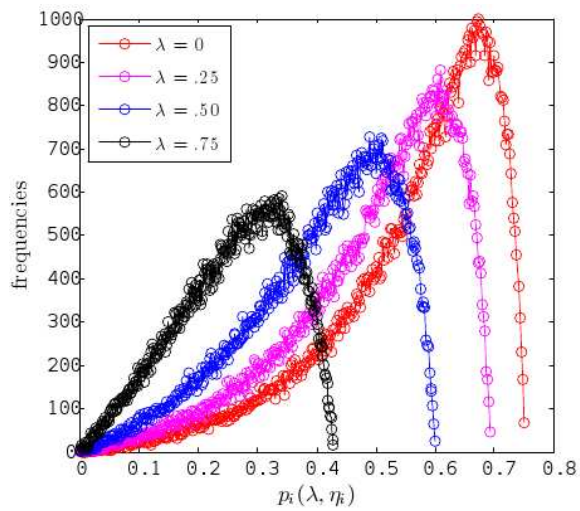
(d) *posterior distribution when* $q = .25$.

Figure 13: Bayesian inference in a society of $n = 10^5$ citizens where education distributes uniformly.

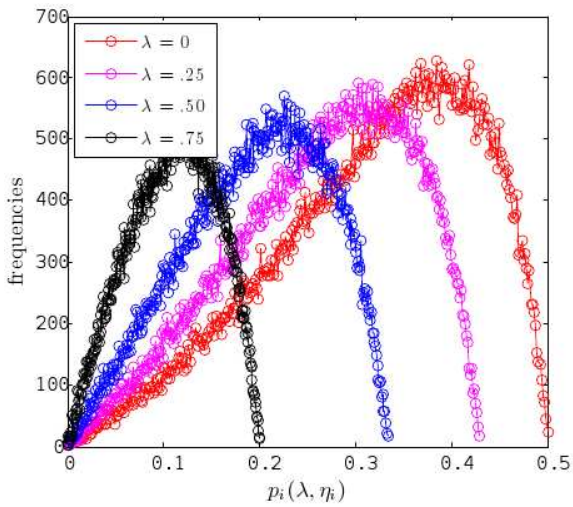
A.2.2 Bayesian inference in a society of $n = 10^5$ citizens where education distributes symmetrically



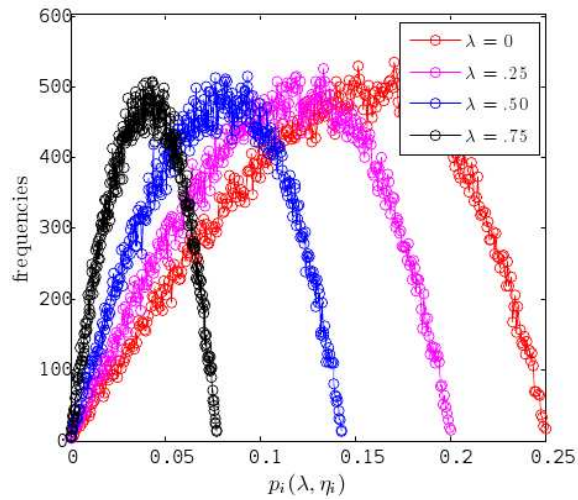
(a) *Education,* $\eta \sim Beta(2, 2)$.



(b) *posterior distribution when* $q = .75$.



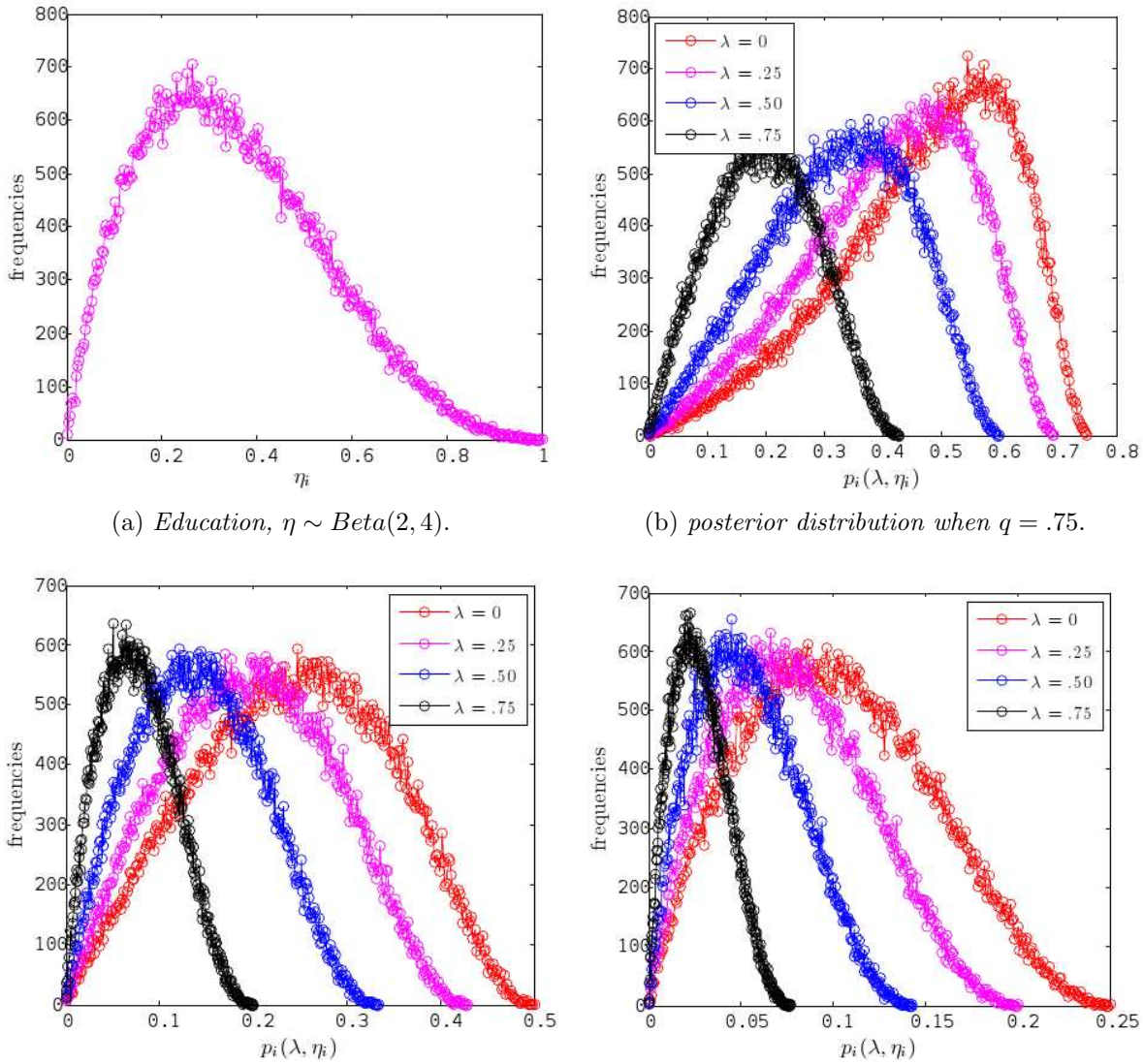
(c) *posterior distribution when* $q = .50$.



(d) *posterior distribution when* $q = .25$.

Figure 14: Bayesian inference in a society of $n = 10^5$ citizens where education distributes symmetrically.

A.2.3 Bayesian inference in a society of $n = 10^5$ citizens where the distribution of education is positive skewed (right-tailed)



(a) *Education,* $\eta \sim Beta(2, 4)$.

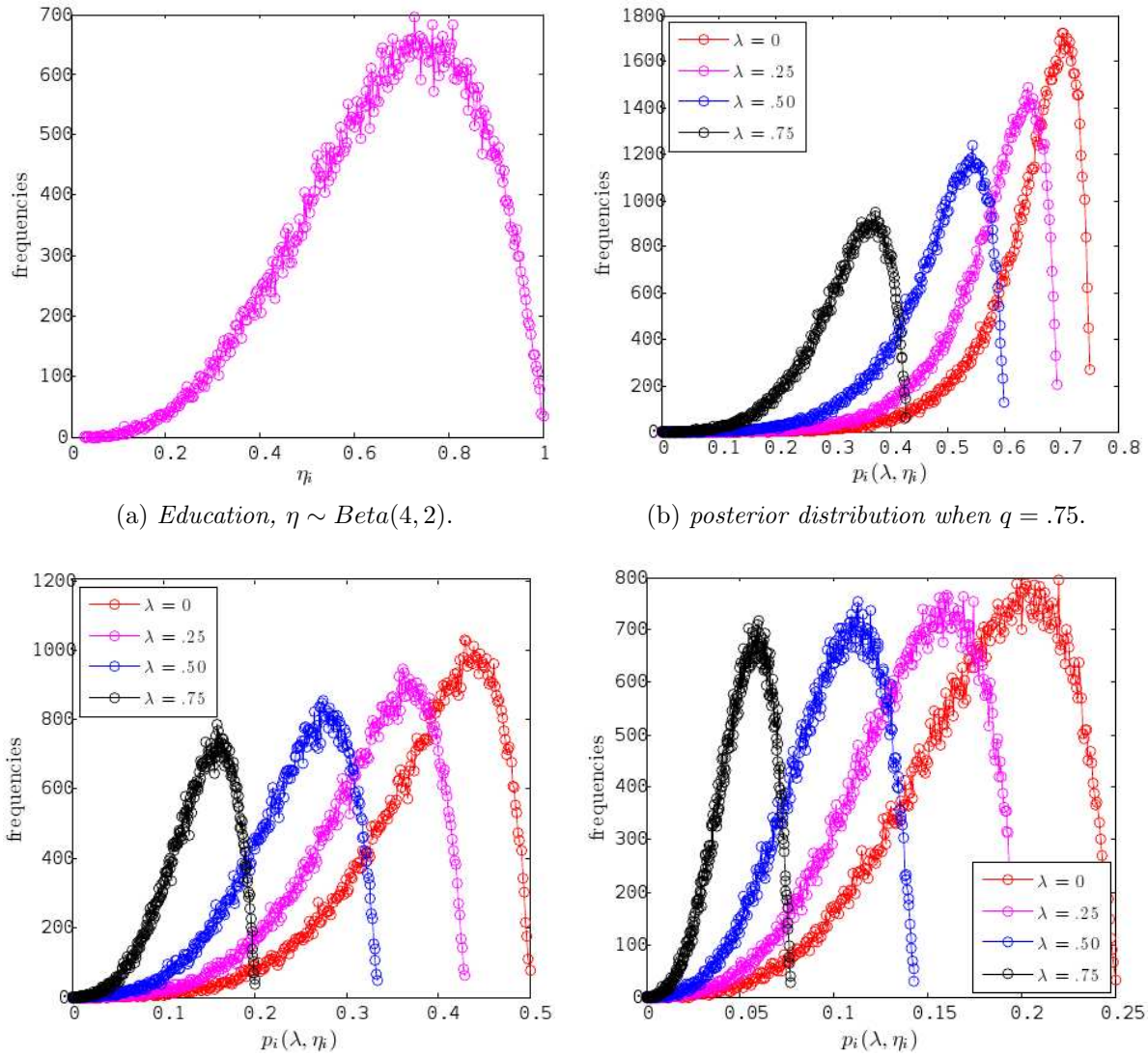
(b) *posterior distribution when* $q = .75$.

(c) *posterior distribution when* $q = .50$.

(d) *posterior distribution when* $q = .25$.

Figure 15: Bayesian inference in a society of $n = 10^5$ citizens where the distribution of education is positive skewed (right-tailed).

A.2.4 Bayesian inference in a society of $n = 10^5$ citizens where the distribution of education is negative skewed (left-tailed)



(a) *Education*, $\eta \sim Beta(4, 2)$.

(b) *posterior distribution when* $q = .75$.

(c) *posterior distribution when* $q = .50$.

(d) *posterior distribution when* $q = .25$.

Figure 16: Bayesian inference in a society of $n = 10^5$ citizens where the distribution of education is negative skewed (left-tailed).

B Proofs

B.1 Proof of Proposition 2

Claim 1: *there exists a unique optimal level of private investment as a function of the own level of education and the congruence rate of the ruler.*

Proof: The first order condition of maximization (18) is

$$(h_t^i)^{\phi-1} = \beta q V'(e_{t+1}^i)$$

that equals the costs of investing one unit more today, on the left hand side, to the expected marginal benefits from getting more educated in the future, on the right hand side: these are namely an increasing in future output and the catch up of higher political sophistication. The solution of the FOC uses the standard envelope condition to compute the expected future marginal benefits:

$$V'_{t+1} = \alpha(1-\tau^H)(1-(1-\lambda_t)\eta_t^i)(e_{t+1}^i)^{\alpha-1}$$

Hence, we are left with

$$(h_t^i)^{\phi-1} = \alpha\beta q(1-\tau^H)(1-(1-\lambda_t)\eta_t^i)(e_{t+1}^i)^{\alpha-1}$$

To get an analytical solution we constrain the human capital elasticity $\alpha = 2$ and investment costs to be quadratic (i.e. $\phi = 2$). This yields equation (19). To prove that solution (19) is also unique we compute the second order condition that is satisfied iff

$$1 - \Omega(1-(1-\lambda_t)\eta_t^i) \geq 0$$

Since $(1-(1-\lambda_t)\eta_t^i) \leq 1$, a sufficient condition for the SOC to hold is that $\Omega \equiv 2\beta q(1-\tau^H) \leq 1$.

Claim 2: *Citizens optimally respond to politicians' congruence rate by lowering private investments when λ_t decreases.*

Proof: To prove that, we need to demonstrate that $\partial h_t^i/\partial\lambda_t \geq 0$. Let us define $A \equiv \Omega(1-(1-\lambda_t)\eta_t^i)$. Differentiation of (19) yields:

$$\begin{aligned}
\frac{\partial h_t^i}{\partial \lambda_t} &= \frac{(1-\delta)e_t^i}{1-A}\frac{\partial A}{\partial \lambda_t} + \frac{(1-\delta)Ae_t^i}{(1-A)^2}\frac{\partial A}{\partial \lambda_t} \\
&= \frac{(1-\delta)e_t^i}{1-A}\frac{\partial A}{\partial \lambda_t}\Big(1+\frac{A}{1-A}\Big) \\
&\geq 0
\end{aligned}$$

Claim 3: *The reaction is as strong as larger is the level of sophistication η_t^i (Spence-Mirlees condition).*

Proof: The statement requires that $\partial^2 h_t^i / \partial\lambda_t \partial e_t^i \geq 0$. Straight differentiation yields:

$$\frac{\partial^2 h_t^i}{\partial\lambda_t \partial e_t^i} = \frac{1-\delta}{1-A}\left[2 - \Omega\frac{(1-\lambda_t)\eta_t^i}{1-A}\right]\frac{\partial A}{\partial\lambda_t} \geq 0$$

iff $2(1-A) \geq \Omega(1-\lambda_t)\eta_t^i$. Rearranging we get

$$(1-\Omega) + (1-A) \geq 0$$

where both the addends are positive by SOC.

Claim 4: *Private investment are hill-shaped with respect to human capital stock conditional on the political environment.*

Proof: Straight differentiation shows that $\partial h_t^i / \partial e_t^i \geq 0$ iff

$$3\Omega(1-\lambda_t)^2(\eta_t^i)^2 - 2(1-\Omega)(1-\lambda)\eta_t^i + (1-\Omega) \geq 0$$

Solving with respect to e_t^i yields a cut-off value, $\tilde{e}_t(\lambda_t)$, according to that two different patterns of propensity to invest have been found:

- $\forall\, e_t^i \in [0, \tilde{e}_t(\lambda_t)] \Longrightarrow \frac{\partial h_t^i}{\partial e_t^i} \geq 0$
- $\forall\, e_t^i \in [\tilde{e}_t(\lambda_t), \bar{e}] \Longrightarrow \frac{\partial h_t^i}{\partial e_t^i} \leq 0$

Note that the identified cut-off

$$\tilde{e}_t(\lambda_t) = \frac{1}{3}\frac{\bar{e}}{\Omega(1-\lambda_t)}(1-\Omega)^{1/2}\left[(1-\Omega)^{1/2} + (1-4\Omega)^{1/2}\right]$$

is increasing to the congruence rate of the ruler, meaning that a fair political environment wipes out the implications of decreasing educational effect because, for $\lambda_t \to 1$, $\tilde{e}_t$ converges to $\bar{e}$ making investments increasing in human capital for the most skilled agents too. Finally, note that $\tilde{e}_t(0)$ is still positive and increasing for $e_t^i < \tilde{e}_t(0)$, i.e. even with a dissonant politician naive citizens still go through private investment.

B.2 Proof of Proposition 3

Assume $\ln e_0^i \sim \mathcal{N}(\mu_0, \Delta_0^2)$. To show that e_t^i keeps distributing lognormally over time let us call

$$\begin{aligned}\xi_0^i &= 1 - \Omega(1 - (1 - \lambda_0)\eta_0^i) \\ &= (1 - \Omega) + (\Omega(1 - \lambda_0)/\bar{e})e_0^i\end{aligned}$$

Since e_0^i distribute lognormally, with mean μ_0 and variance Δ_0^2, and the other terms in ξ_0^i are constants, ξ_0^i must distribute lognormally as well with mean $\mu_0 + \ln\Omega + \ln(1 - \lambda_0) - \ln\bar{e} + (1 - \Omega)$ and variance Δ_0^2. Finally, note that the sum of log-normal distributions yields, under certain conditions, another lognormal distribution.

If $\ln e_t^i \sim \mathcal{N}(\mu_t, \Delta_t^2)$ then $y_t^i = (e_t^i)^2$ is a transformation of e_t^i and must distribute as a lognormal too: $\ln y_t^i \sim \mathcal{N}(m_t, v_t^2)$, with $m_t = 2\mu_t + \Delta_t^2$.

The law of motion of aggregated level of human capital is obtained from (20). Taking a logarithmic transformation in both sides we are left with

$$\begin{aligned}\ln e_{t+1}^i &= \ln e_t^i + \ln(1 - \delta) - \ln(1 - \Omega(1 - (1 - \lambda_t)\eta_t)) \\ &\approx \ln e_t^i - \delta + \Omega - \Omega(1 - \lambda_t)\eta_t\end{aligned}$$

for small values of δ and $\Omega(1 - \lambda_t)\eta_t$. Averaging yields:

$$\mu_{t+1} = \mu_t + \frac{\Delta_t^2}{2} + (\Omega - \delta) - \Omega(1 - \lambda_t)\eta_t,$$

with $\eta_t \equiv \exp(\mu_t + \Delta_t^2/2)/\bar{e}$.

At the same way, we can compute the expected output in $t + 1$ of the economy; taking a logarithmic transformation of (17) yields:

$$\begin{aligned}\ln y_{t+1}^i &= 2\ln e_{t+1}^i + \ln(1 - (1 - \lambda_t)\eta_t) + \ln(q) \\ &\approx 2\ln e_{t+1}^i - (1 - \lambda_t)\eta_t + \ln(q)\end{aligned}$$

Taking the average:

$$\begin{aligned}m_{t+1} &= 2\mu_{t+1} - (1 - \lambda_t)\eta_t \\ &= 2\mu_t + \Delta_t^2 + 2(\Omega - \delta) - 2\Omega(1 - \lambda_t)\eta_t - (1 - \lambda_t)\eta_t \\ &= m_t + 2(\Omega - \delta) - (1 + 2\Omega)(1 - \lambda_t)\eta_t\end{aligned}$$

In time T2 the incumbent ruler faces the following maximization problem:

$$V_t^r(\lambda_t) = \max_{\lambda_t} \left\{ \tau y_t - qb\tau y_t \lambda_t + \beta \mathbb{E}_t \left[V_{t+1}^r(\lambda_t) \right] \right\} \quad \text{s.t. (14)}$$

Taking a logarithmic transformation yields:

$$\begin{aligned}
\ln V_t^r(\lambda_t) &= \max_{\lambda_t} \Big\{ \ln\tau + \ln y_t + \ln(1-qb\lambda_t) + \beta\mathbb{E}_t\big[\ln V_{t+1}^r(\lambda_t)\big] \Big\} \quad \text{s.t. (14)} \\
&\approx \max_{\lambda_t} \Big\{ \ln\tau + \ln y_t - qb\lambda_t + \beta\mathbb{E}_t\big[\ln V_{t+1}^r(\lambda_t)\big] \Big\} \quad \text{s.t. (14)} \\
&= \max_{\lambda_t} \Big\{ \ln\tau + 2\mu_t + \Delta_t^2 - qb\lambda_t + \beta\mathbb{E}_t\big[\ln V_{t+1}^r(\lambda_t)\big] \Big\} \quad \text{s.t. (14)}
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{E}_t\big[\ln V_{t+1}^r(\lambda_t)\big] &= \ln q + \ln\tau + m_{t+1} \\
&= \ln q + \ln\tau + 2\mu_t + \Delta_t^2 + 2(\Omega-\delta) - 2\Omega(1-\lambda_t)\eta_t - (1-\lambda_t)\eta_t
\end{aligned}$$

Since the program is linear in λ_t, it is easy to note that

$$\frac{\partial \ln V_t^r}{\partial \lambda_t} = \beta(1+2\Omega)\eta_t - qb$$

that is negative if and only if condition (26) is satisfied, that is if the cost of the public investment, relative to tax revenues, is high enough.

We now show that there exists a cutpoint-cost of the public good $\bar{b}$ such that for every $b > \bar{b}$ the ruler strictly prefer to going-for-broke instead of playing a strategy $\lambda_t \geq \lambda_t^*$. We first show that, at time $T2$, the value function of going-for-broke is strictly greater than the value taken by being congruent enough (we show it for $\lambda_t = \lambda_t^*$; a fortiori it hold for any $\lambda_t > \lambda_t^*$). Let us call $\mathcal{V}_t^r \equiv \ln V_t^r(\lambda_t^*)$.

$$\mathcal{V}_t^r < T_t$$

$$\begin{aligned}
\ln\tau + 2\mu_t + \Delta_t^2 > (1+\beta)\ln\tau + 2(1+\beta)\mu_t + (1+\beta)\Delta_t^2 - qb + b(1-q)\frac{1}{\eta_t^m} + \\
+ \beta\ln q + 2\beta(\Omega-\delta) - \beta(1+2\Omega)(1-q)\exp\left(\frac{\Delta_t^2}{2}\right)
\end{aligned}$$

Solving by b yields $\bar{b}$:

$$\bar{b}(\delta) \equiv \frac{\beta\big[\ln\tau q + m_t + 2\beta(\Omega-\delta) - \beta(1+2\Omega)(1-q)\exp\big(\frac{\Delta_t^2}{2}\big)\big]}{q-(1-q)\frac{1}{\eta_t^m}}.$$

Note that $\bar{b}(\delta)$ is a decreasing function of the depreciation rate of the human capital of the producers so as for high level of δ Go-for-Broke is more likely to be the optimal strategy in the stage game. In particular, there exists a threshold δ^* such that for any $\delta < \delta^*$ Going-for-Broke is not an admissible strategy for the ruler. To show that we need to solve the following inequality:

$$\bar{b}(\delta) > 1,$$

that holds for

$$\delta < \Omega - \frac{1}{2\beta}\left[q - (1-q)\frac{1}{\eta_t^m} - \beta \ln \tau q + \beta m_t - \beta(1+2\Omega)(1-q)\exp\left(\frac{\Delta_t^2}{2}\right)\right]$$
$$\equiv \delta^*$$

B.3 Proof of Proposition 5

Assume $\ln e_t^i \sim \mathcal{N}(\mu_t, \Delta_t^2)$, such that $\ln(e_t/e_t^m) = \Delta_t^2/2$. Given condition (26), the optimal congruence rate is

$$\lambda_t^* = 1 - \frac{1-q}{q}\frac{1}{\eta_t^m}.$$

Substituting it into the average future income m_{t+1} and the ruler's rent V_t^r yields equations (32) and (33), respectively.

Claim 1: *Income is hill-shaped with respect to inequality.*

Proof: Straight differentiation shows that income is increasing w.r.t. inequality iff:

$$\frac{\partial m_{t+1}}{\partial \Delta_t^2} \geq 0 \Longleftrightarrow 1 \geq \left(\frac{1}{2} + \Omega\right)\left(\frac{1-q}{q}\right)\exp\left(\frac{\Delta_t^2}{2}\right)$$
$$\Longleftrightarrow \Delta_t^2 \leq 2\ln(q) - 2\ln(1-q) + 2\ln(2) - 2\ln(1+2\Omega) \equiv \bar{\Delta}^c$$

Claim 2: *Politicians rents are hill-shaped with respect to inequality.*

Proof:

Straight differentiation shows that politicians rents are increasing w.r.t. inequality iff:

$$\frac{\partial \ln V_t^r}{\partial \Delta_t^2} \geq 0 \Longleftrightarrow \frac{1+\beta}{\beta} \geq \frac{1}{2}(1+2\Omega)\frac{1-q}{q}\exp\left(\frac{\Delta_t^2}{2}\right)$$
$$\Longleftrightarrow \Delta_t^2 \leq 2\ln 2 + 2\ln(1+\beta) - 2\ln\beta + 2\ln q - 2\ln(1-q) - 2\ln(1+2\Omega) \equiv \bar{\Delta}^r$$

Claim 3: *Citizens bliss point is smaller than ruler's.*

Proof: We need to demonstrate that $\bar{\Delta}^c < \bar{\Delta}^r$. It comes out from the definitions in Claim 1 and 2:

$$\bar{\Delta}^c < \bar{\Delta}^r \Longleftrightarrow \ln(1+\beta) > \ln\beta$$

that is always true.

B.4 Proof of Proposition 6

A stable steady state is a point $(e_\infty, \lambda_\infty)$ with the curve $\chi(\lambda)$ cuts the curve $\Lambda(e)$ from above. An unstable steady state corresponds in each case to an intersection from below.

The dynamical system (31) reduces to a one-dimensional recursion: $e^i_{t+1} = \chi(e^i_t, \Lambda(e^i_t))$. It has the following features:

(i)

$$\chi(0) = 0$$

(ii)

$$\chi'(e^i_t) = \frac{(1-\delta)(1-\Omega)}{\left[1 - \Omega\left(1 - \frac{e^i_t}{e^m_t}\frac{1-q}{q}\right)\right]^2} \geq 0 \quad \text{by SOC}$$

(iii)

$$\chi''(e^i_t) = -\frac{(1-\delta)(1-\Omega)}{\left[1 - \Omega\left(1 - \frac{e^i_t}{e^m_t}\frac{1-q}{q}\right)\right]^3}\frac{\Omega}{e^m_t}\frac{1-q}{q} \leq 0 \quad \text{by SOC}$$

(iv)

$$\chi'(e^i_t) \geq 1 \Longleftrightarrow \begin{cases} e^i_t \in \left[0, \frac{q}{1-q}\frac{e^m_t}{\Omega}(1-\Omega)^{1/2}\left((1-\delta)^{1/2} - (1-\Omega)^{1/2}\right)\right] & \text{iff} \quad \Omega \geq \delta \\ e^i_t \in \left[\frac{q}{1-q}\frac{e^m_t}{\Omega}(1-\Omega)^{1/2}\left((1-\delta)^{1/2} - (1-\Omega)^{1/2}\right), 0\right] & \text{elsewhere} \end{cases}$$

(v)

$$e^i_{t+1} = e^i_t = e_\infty$$

in four fixed points:

$$e^{(1)}_\infty = 0$$

$$e^{(2)}_\infty = \begin{cases} \left(1 - \frac{\delta}{\Omega}\right)\left(\frac{q}{1-q}\right)e^m_\infty \geq 0 & \text{iff} \quad \Omega \geq \delta \\ \left(1 - \frac{\delta}{\Omega}\right)\left(\frac{q}{1-q}\right)e^m_\infty \leq 0 & \text{elsewhere} \end{cases}$$

A fixed point e_∞ is stable if and only if

$$\left.\frac{\mathrm{d}\chi(e)}{\mathrm{d}e}\right|_{e=e_\infty} < 1$$

Computation tells us that $e^{(2)}_\infty$ is the only stable fixed point if $\Omega(q) \geq \delta$ (or if $q \geq \bar{q}(\beta, \delta, \tau) \equiv \frac{1}{2}\frac{\delta}{\beta(1-\tau)}$), whereas $e^{(1)}_\infty = 0$ is unstable. In the latter case ($\Omega(q) \leq \delta$), $e^{(1)}_\infty = 0$ is the unique stable fixed point.

In the former case, $e_\infty^{(2)}$ depends on the trajectory of the median agent of the distribution, that is then pivotal. In time 1, from (31), is evident that $e_1^i \geq e_0^i$ if the citizen i is naive enough, i.e. if

$$e_0^i \leq \Big(1 - \frac{\delta}{\Omega}\Big)\Big(\frac{q}{1-q}\Big)e_0^m$$

More sophisticated agents instead will be driven by the political process to decrease the investment in human capital. This reasoning implies a convergence process according to which there exists a catching-up period $t^* > 0$ such that $e_{t^*}^i = e_{t^*}^m$ and $F(e_{t^*}^i)$ is degenerate for every $t \in [t^*, \infty)$. The dynamics of the system turns into a degenerate ergodic process described by the following linear law of motion:

$$e_t = \left[\frac{1-\delta}{1 - \Omega\big(1 - \frac{1-q}{q}\big)}\right]^t e_0^m \qquad \forall t \in [t^*, \infty)$$

Let us call $D = \frac{1-\delta}{1-\Omega\left(1-\frac{1-q}{q}\right)}$. It follows that e_∞ diverges for every $D > 1$, i.e. for every

$$q > \frac{2\beta(1-\tau)+\delta}{4\beta(1-\tau)} \equiv \tilde{q}(\beta, \delta, \tau)$$

whereas it converges to zero for every $D < 1$. Finally it converges to e_0^m for every $D = 1$. Note that $\tilde{q} \geq \bar{q}$ meaning that $D > 1$ when $\Omega > \delta$.

B.5 Proof of Proposition 7

The individual optimal tax rate is obtaining by solving the MPBE by backward induction. Each citizen i, endowed with e_t^i in time t, anticipates which is the optimal congruence rate of the ruler, λ_t^*, and accordingly solve maximization (36).

FOC requires $\mathcal{V}'(\tau) = 0$. Since the maximization problem is quite complex we constrain parameters to get analytical results to the following values: $\beta = 1$ and $\delta = 0$. We also initially set $q = 0.25$ to consider a pessimistic scenario, i.e. one in which politicians is rationally expected to cheat, but later we allow q to vary. We also define $\epsilon^i = e_t^i / e_t^m$ to be the individual relative level of political sophistication in each period t (to simplify notations we drop the subscript t).

In this scenario the only plausible solution is given by

$$\tau^*(\epsilon^i | q = 0.25) = 2 + \frac{3}{8}(1 - 3\epsilon^i) + \\ + \frac{1}{2}\frac{\sqrt{\frac{9}{16}(1-3\epsilon^i)^4 + \frac{3}{8}(1-3\epsilon^i)^3 + 17(1-3\epsilon^i)^2 - 20(1-3\epsilon^i) - 16}}{1 - 3\epsilon^i}$$

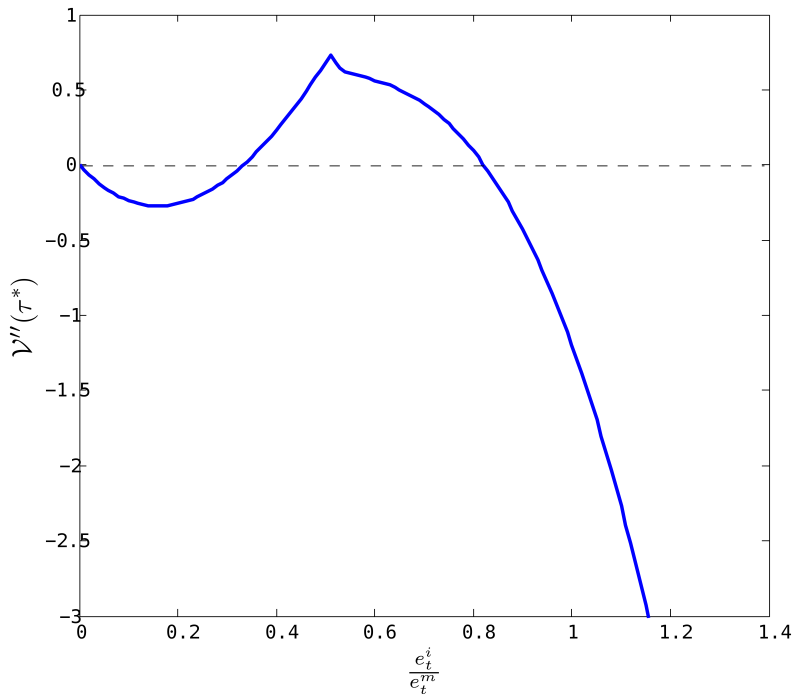


Figure 17: Second order condition of the taxation maximization problem.

depicted in Fig. 11(a).

SOC tells us that $\tau^*(\epsilon^i|q = 0.25)$ is a maximum for any $\tau^*(\epsilon^i|q = 0.25)$ satisfying the following inequality:

$$\mathcal{V}''(\tau^*(\epsilon^i|q = 0.25)) < 0$$

$$\mathcal{V}'' = (1 - 3\epsilon^i)\Big[\frac{1}{2}(1 - 3\epsilon^i)\big(\tau^* - \frac{3}{8}\big) - 1\Big] < 0.$$

The SOC problem is illustrated in Fig. 17. As one may see $\tau^*(\epsilon^i|q = 0.25)$ is not a maximum for $\epsilon^i \in (1/3, 9/10)$ for neither solutions. Then we proceed by computing corner solutions. We limit to show it for $\epsilon^i = 1/2$, but same results apply for any $\epsilon^i \in (1/3, 9/10)$. It is straightforward to see that

$$\mathcal{V}(\tau = 1|\epsilon^i = 0.5, q = 0.25) = 0,$$

whereas

$$\mathcal{V}(\tau = 0|\epsilon^i = 0.5, q = 0.25) = \frac{9}{10}(e_t^i)^2 > 0.$$

Since $\mathcal{V}(\tau|\epsilon^i = 0.5, q = 0.25)$ is a monotonically decreasing function and

$$\mathcal{V}(\tau = 0|\epsilon^i = 0.5, q = 0.25) > \mathcal{V}(\tau = 1|\epsilon^i = 0.5, q = 0.25),$$

we conclude that $\tau = 0$ is a corner solution for $\epsilon^i = 0.5$.

A more optimistic scenario, i.e. one in which the ruler is expected to be congruent, is one with $q = 0.75$. In this scenario the only plausible solution is given by

$$\tau^*(\epsilon^i|q = 0.75) = \frac{2}{3} + \frac{3}{8}(1 - \frac{1}{3}\epsilon^i) + \\ + \frac{1}{2}\frac{\sqrt{\frac{9}{16}(1-\frac{1}{3}\epsilon^i)^4 + \frac{9}{2}(1-\frac{1}{3}\epsilon^i)^3 + \frac{25}{9}(1-\frac{1}{3}\epsilon^i)^2 - \frac{20}{3}(1-\frac{1}{3}\epsilon^i) - \frac{16}{9}}}{1-\frac{1}{3}\epsilon^i}$$

depicted in Fig. 11(b). SOC also shows that $\mathcal{V}''(\tau^*(\epsilon^i|q = 0.75)) < 0$ stands for any ϵ^i.

Finally one may see that for $q = 1$ each citizen, regardless of his own level of sophistication, knows that the politician in office is a perfect agent always playing $\lambda_t^* = 1$. In such scenario

$$\tau^*(q = 1) = \frac{7}{2} \pm \frac{1}{2}\sqrt{29} > 1.$$

References

[1] Abramowitz A., Lanoue D.J. and Ramesh S. (1988), *Economic conditions, causal attributions, and political evaluations in the 1984 presidential election*, Journal of Politics, Vol. 50, No. 4: 848–63.

[2] Acemoglu D., Johnson S., and Robinson J.A. (2001), *The colonial origins of comparative development: an empirical investigation*, AER, Vol. 91: 1369-1401.

[3] Acemoglu D., Johnson S., Robinson J.A. and Yared P. (2005), *From education to democracy?*, AER, Vol. 95, No. 2: 44-49.

[4] Acemoglu D., Egorov G., and Sonin K. (2013), *A Political Theory of Populism*, QJE, Vol. 128, No. 2: 771-805.

[5] Almond G.A. and Verba S. (1989), *The Civic Culture: Political Attitudes and Democracy in Five Nations*, SAGE Publications.

[6] Ashworth S. (2012), *Electoral accountability: recent theoretical and empirical work*, ARPS No. 15: 183-201

[7] Ashworth S., de Mesquita E.B. and Friedenberg A. (2013) *Accountability Traps*, Working Paper.

[8] Austen-Smith D., Banks J. (1988), *Elections, Coalitions, and Legislative Outcomes*, APSR, Vol. 82, No. 2: 405-422.

[9] Banks, J. (1990), *A model of electoral competition with incomplete information*, JET, Vol. 50, No. 2: 309-325.

[10] Banks, J., Sundaram R. (1993), *Adverse Selection and Moral Hazard in a Repeated Elections Model*, in Political Economy: Institutions, Information, Competition, and Representation, eds. W. Barnett, M.J. Hinich and N. Schofield. New York, Cambridge University Press.

[11] Barro R.J. (1973), *The Control of Politicians: An Economic Model*, PC, Vol. 14: 19-42

[12] Barro R. J. (1991), *Economic Growth in a Cross Section of Countries*, QJE, Vol. 106: 407-433.

[13] Barro R.J. (1999), *Determinants of democracy*, JPE, Vol. 107: 158-183.

[14] Barro R.J and Sala-I-Martin X. (1992), *Public Finance in Models of Economic Growth*, RES, Vol. 59, No. 4: 645-661.

[15] Benabou R. (1996), *Inequality and growth*, NBER Macroeconomics Annual 1996, Vol. 11.

[16] Benabou R. (2000), *Unequal societies: income distribution and the social contract*, AER, Vol. 90, No. 1: 96-129.

[17] Benabou R. (2002), *Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?*, Econometrica, Vol. 70, No. 2: 481-517

[18] Benabou R. (2013), *Groupthink: collective delusions in organizations and markets*, RES, Vol. 80: 429-462.

[19] Benabou R., Tirole, J. (2002), *Self confidence and personal motivation*, QJE, Vol. 117, No. 3: 871-915.

[20] Benhabib J., Schmitt-Groh S. and Uribe M. (2001), *Monetary Policy and Multiple Equilibria*, AER, Vol. 91, No. 1: 167-186.

[21] Besley T., Case, A. (1995), *Incumbent Behavior: Vote-Seeking, Tax-Setting, and Yardstick Competition*, AER Vol. 85, No. 1: 25-45.

[22] Besley T. (2006), *Principled Agents? The Political Economy of Good Government*, Oxford University Press.

[23] Bidner C., Francois P., and Trebbi F., (2014), *A Theory of Minimalist Democracy*, NBER Working Papers No. 20552, National Bureau of Economic Research, Inc.

[24] Bobba M., Coviello D. (2007), *Weak instruments and weak identification, in estimating the effects of education, on democracy*, EL, Vol. 96, No. 3: 301-306.

[25] Botero J., Ponce A., and Shleifer A. (2013), *Education, Complaints, and Accountability*, JLE, Vol. 56, No. 4: 959–996.

[26] Bourguignon F., Verdier T. (2000), *Oligarchy, democracy, inequality and growth*, JDE, Vol. 62: 285-313.

[27] Brender A. and Drazen A. (2005), *Political budget cycles in new versus established democracies*, JME, Vol. 52, No. 7: 1271-1295.

[28] Callender S. and Wilkie S. (2007), *Lies, damned lies, and political campaigns*, GEB, Vol. 60: 262-286.

[29] Castelló-Climent A. (2008), *On the distribution of education and democracy*, JDE, Vol. 87: 179-190.

[30] Chappell H.W. and Keech W.R. (1985), *A new view of political accountability for economic performance*, APSR: 10–27.

[31] Coate, S., Morris, S. (1995), *On the Form of Transfers in Special Interests*, JPE, Vol. 103, No. 6: 1210-35.

[32] Delli Carpini M.X. and Keeter S. (1997), *What Americans Know about Politics and Why It Matters*, Yale University Press.

[33] Durlauf S.N. (1996), *A Theory of Persistent Income Inequality*, JEG, Vol. 1: 75-93.

[34] Federico C.M. and Sidanius J. (2002), *Racism, ideology, and affirmative action revisited: the antecedents and consequences of 'principled objections' to affirmative action*, Journal of Personality and Social Psychology, Vol. 82, No. 4: 488.

[35] Fearon, J. (1999), *Electoral Accountability and the Control of Politicians: Selecting Good Candidates versus Sanctioning Poor Performance*, in Democracy, Accountability, and Representation, eds. A. Przeworski, S.C. Stokes and B. Manin, New York, Cambridge University Press.

[36] Ferejohn J. (1986), *Incumbent Performance and Electoral Control*, PC, Vol. 50, No. 1/3, Carnegie Papers on Political Economy, Volume 6: 5-25.

[37] Galor O. and Zeira J. (1993), *Income Distribution and Macroeconomics*, RES, Vol. 60, No. 1: 35-52.

[38] Glaeser E.L, La Porta R., Lopez-de-Silanes F., Shleifer A. (2004), *Do institutions cause growth?*, JEG, Vol. 9, 271-303.

[39] Glaeser E.L, Ponzetto G.A.M., and Shleifer A. (2007), *Why Does Democracy Need Education?*, JEG, Vol. 12, No. 2: 77–99.

[40] Glomm, G. and Ravikumar, B. (1992), *Public versus private investment in human capital endogenous growth and income inequality*, JPE, Vol. 100(4) :818-34.

[41] Gomez B.T. and Wilson J.M. (2001), *Political sophistication and economic voting in the American electorate: A theory of heterogeneous attribution*, AJPS: 899–914.

[42] Gradstein M. and Justman M. (1997), *Democratic choice of an education system: implications for growth and income distribution*, JEG, Vol. 2, No. 2: 169–183.

[43] Grossman G.M. and Helpman E. (2001), *Special interest politics*, The MIT press, Cambridge, Massachusetts.

[44] Guiso L., Sapienza P. and Zingales L. (2003), *People's opium? Religion and economic attitudes*, JME, Vol. 50, No. 1: 225-282.

[45] Harrington, J.E. (1993), *Economic Policy, Economic Performance, and Elections*, AER, Vol. 83, No. 1: 27-42.

[46] Helliwell J.F. and Putnam R.D. (2007), *Education and Social Capital*, Eastern Economic Journal, Vol. 33, No. 1: 1–19.

[47] Jaggers, K. and Marshall, M. (2003), *Polity IV project*, Center for International Development and Conflict Management, University of Maryland.

[48] Kartik N. and McAfee R.P. (2007), *Signaling Character in Electoral Competition*, AER, Vol. 97, No. 3: 852-870.

[49] Kaufmann D., Kraay A. and Mastruzzi M. (2010), *The Worldwide Governance Indicators: Methodology and Analytical Issues*, World Bank Policy Research Working Paper No. 5430.

[50] Köszegi B. (2014), *Behavioral Contract Theory*, JEL, Vol. 52, No. 4: 1075–1118.

[51] Lipset S.M. (1959), *Some social requisites of democracy: economic development and political legitimacy*, APSR, Vol. 53, 69-105.

[52] Lucas, R J. (1988), *On the mechanics of economic development*, JME, Vol. 22(1): 3-42.

[53] Luskin, R C. (1987), *Measuring Political Sophistication*, AJPS, Vol. 31 No. 4: 856–899.

[54] McGraw K.M. (2000), *Contributions of the cognitive approach to political psychology*, Political Psychology, Vol. 21, No. 4: 805–832.

[55] Meltzer A.H. and Richard S.F. (1981), *A Rational Theory of the Size of Government*, JPE, Vol. 89, No. 5: 914-927.

[56] Przeworski A., Alvarez M., Cheibub J.A., Limongi F. (2000), *Democracy and development: political institutions and material well-being in the world, 1950-1990*, Cambridge University Press, NY.

[57] Persson T. and Tabellini G.E. (2000), *Political economics: explaining economic policy*, MIT Press, Cambridge, Massachusetts.

[58] Persson T. and Tabellini G.E. (2002), *Do constitutions cause large governments? Quasi-experimental evidence*, EER, Vol. 46, No. 4-5: 908-918.

[59] Rogoff, K. (1990), *Equilibrium Political Budget Cycles*, AER, Vol. 80, No. 1: 21–36.

[60] Rogoff K. and Sibert A. (1988), *Elections and macroeconomic policy cycles*, RES, Vol. 55, No. 1: 1–16.

[61] Saint-Paul G. (1994), *Unemployment, wage rigidity, and the returns to education*, EER, Vol. 38, No. 3: 535–543.

[62] Shi M. and Svensson J. (2006), *Political budget cycles: Do they differ across countries and why?*, JPubE, Vol. 90, No. 8-9: 1367–1389.

[63] Sniderman P., Brody R.A. and Tetlock P.E. (1991), *Reasoning and Choice: Explorations in Political Psychology*, Cambridge: Cambridge University Press.

[64] Verba S. and Nie N.H. (1972), *Participation in America: Political Democracy and Social Equality*, Chicago: University of Chicago Press.

[65] Zaller J. (1992), *The nature and origins of mass opinion*, Cambridge university press.