

## A Lipsetian Theory of Institutional Change

Raouf Boucekkine  
Paolo G. Piacquadio  
Fabien Prieur

# A Lipsetian theory of institutional change

Raouf Boucekkine\*      Paolo G. Piacquadio<sup>†</sup>

Fabien Prieur<sup>‡</sup>

February 23, 2015

## Abstract

The paper addresses the role of education policies for institutional change. Our paradigmatic model consists of an autocratic elite and a mass of hand-to-mouth workers. The elite has full political and economic control. First, it anticipates and can avoid revolutionary threats through income redistribution. Second, it sets the education policy: a higher level of human capital results in a larger productivity of the national industry, but also in higher consumption aspirations of citizens (and thus more costly redistribution). Finally and in contrast to the recent literature on democratization games, the elite can stop the autocracy and initiate an institutional change. We show that perspective economic returns on education and resources play a crucial role: if sufficiently high, these may prompt high investment in education, human capital accumulation, and, eventually, an institutional change. Our theory of institutional change captures three essential dimensions of Lipset's view: the positive relationship between education and institutional change, the positive relationship between income and institutional change and, in a more stylized fashion, the negative relationship between inequality and institutional change.

**Keywords:** Democratization; human capital; Lipset's theory.

---

\*Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS, and senior member of the Institut Universitaire de France.

<sup>†</sup>Department of Economics, University of Oslo. Corresponding author. Email: [pgiacquadio@econ.uio.no](mailto:pgiacquadio@econ.uio.no).

<sup>‡</sup>INRA-LAMETA and University of Montpellier I. Email: [prieur@supagro.inra.fr](mailto:prieur@supagro.inra.fr).

# 1. Introduction

Whether education is a prerequisite for democratization or one of its side products is an important old question in both political sciences and economics. Perhaps the most influential contemporaneous work on this theme is due to Lipset (1960). According to Lipset, educated people are more prone to resort to negotiation and non-violent means to sort out conflicts, which is more compatible with democracy as the latter requires, on the top of everything, a significant civic engagement and participation.

A broader connected debate is the origin and evolution of institutions. A major contribution due to Acemoglu et al. (2001) puts forward the colonial origin of institutions. A central message conveyed by these authors and many others (see for example Rodrik et al., 2004) is the primacy of institutions: institutions cause development, and not the reverse.

The latter message goes at odds with the main Lipsetian statement, that's institutions are only an outcome of human capital accumulation and development. Glaeser et al. (2004) argue that the typical institutional indicators considered in the literature (including Acemoglu et al.) have a serious problem: they rather capture (volatile) outcomes than durable norms and rules as should institutional proxys do. For example, certain popular measures of constraints on governments are shown to be twice as volatile (in terms of average within country deviations) as Barro's measure of years of schooling. Glaeser et al. ultimately find that "human capital is a more basic source of growth than are institutions", fully in line with the Lipsetian view. A few years earlier, Barro reached similar conclusions (Barro, 1999). While the econometric debate is still undecided (see Acemoglu et al., 2005, or Castello-Climent, 2008, for example), it's fair to say that the Lipsetian theory of democratization is up to now a valid and powerful framework for the analysis of institutional transformations (see Przworski et al., 2000, for a general and applied view).

In contrast to the abundant empirical literature, there are only few attempts to capture the Lipsetian view in a theoretical framework. A well-known contribution to this line of research is Glaeser et al. (2007). The main idea conveyed in this paper is that education raises the benefits of civic engagement pretty much as social capital, therefore leading to a larger social and political involvement. Glaeser et al. (2007) go indeed a bit further as they also claim that education does not only favor the emergence of democracy, it also helps stabilizing democracy. More recently, other aspects of Lipset theory are being the object of theoretical developments. For example, Jung and Sunde (2014) have investigated the Lipset claim that democracy

is more likely in countries with more equal distributions of resources.

In this paper, we develop a model of institutional change that allows us to explore three essential features of Lipset’s theory. These are the link between education and democracy, the link between income and democracy, and the link between inequality and democracy.

We study the paradigmatic case of an autocratic elite with full political and economic power. In line with the recent literature on democratization games (see Acemoglu and Robinson, 2006), the elite anticipates the existence and extent of revolutionary threats and can act in such a way that revolutions are avoided. In particular, the elite is able to keep under control the incentives to revolt of the hand-to-mouth workers through an appropriate redistribution policy.

The threshold level of consumption below which workers would revolt is endogenous in the model and depends on the education level of the population. As citizens become more educated, they also become more aware of the political situation (see Zaller, 1992) and they tend to be more politically sophisticated (see Luskin, 1990, and Neuman, 1986). Moreover, as suggested by Campante and Chor (2012) to explain the Arab Spring events, they also have higher income expectations and require better working opportunities.<sup>1</sup> Thus, larger human capital has a positive effect on the revolt-threshold and increases the redistribution-cost of avoided revolutions.

Consequently, the education policy set by the elite has two opposite effects. On the one side, human capital accumulation has the above-described “awareness” cost, making workers more demanding. On the other side, human capital is a production factor in the economy, so education enhances labor productivity and triggers economic growth.

Finally, the elite is assumed to have control over the timing of institutional change. This contrasts with the recent literature on democratization games (cited above): the democratization process is not started by citizens via a costly revolution –since the revolution threat is fully internalized– but by the elite.

While our paradigmatic view cannot directly explain the recent Arab spring events, it may help understanding some aspects of certain non-violent democratization processes as those observed in South Asia. Yet, the inherent research questions in such a context are clearly nontrivial: Would an autocratic elite as ours –with absolute political and economic control– decide to leave the power on his own at

---

<sup>1</sup>Education is certainly more multifaceted than what economists generally assume. As argued by Chomsky (1996), education can be paradoxically used to obtain “ignorance,” with the aim of standardization and domination of populations. This view is supported empirically by Castello-Climent and Mukhopadhyay (2013). For the sake of tractability, we abstract from non-monotonic effects of education on peoples desire for freedom.

finite time? If yes, what determines the duration of dictatorships? Of course, the answer depends on the regime following the institutional change. We show that a particularly painful post-dictatorship period would not encourage the autocrat to leave his office. As a particular case, we do show that in the situation where the continuation payoff is zero, which corresponds to the well-known “end of the world” problem, the autocratic elite will optimally decide to stay forever in office and adjust his redistribution and education policies accordingly (in a way that will be explained later in the main text).

We pay a particular attention to a case which we shall label “Lipsetian”. In our theory, a Lipsetian elite anticipates that their well-being after the institutional change is an increasing function of socio-economic development, measured by the stock of human capital. As a consequence, it might decide to invest in education despite the higher cost of redistribution and will eventually give up political and economic power by inducing an institutional change. As we shall see, the characterization of optimal institutional changes –in the form of voluntary power dismissal– is a quite involved mathematical problem.

A few additional observations are in order here before getting to the analytical specifications. A key aspect of our model is the capacity of the autocratic elite to anticipate and control the revolutionary threat. While this is an unrealistically strong assumption, it enforces our results. Despite unanticipated revolutions –and the corresponding institutional changes– can never occur, the Lipsetian elite might decide to optimally lead the country to democratization. This implies that our theory underestimates the occurrence of institutional changes; however, by looking only at the incentives of the elite, it is also more likely to capture those institutional changes that lead to stable democracies, confirming, from a different perspective, the results by Glaeser et al. (2007).

Last but not least, the autocratic elite has access to a constant windfall of resources, which can be exported or supplied to the national industry. The presence of resources enriches the model with a further dimension. Our comparative statics with respect to windfall resources are consistent with Lipset’s theory, stating that higher incomes are favorable for democracy: in our model, a larger amount of resources endows the elite with the necessary wealth to take the path of worker’s education and guarantees a sufficient long-run payoff, despite this is associated with an institutional change. Beyond making democratization more likely, a larger amount of resources also leads to a shorter duration of the authoritarian regime.

The windfall of resources also allows us to compare our predictions with game-theoretic models of institutional change or of natural resources-induced conflicts.

Our conclusions are in line with those obtained within the former stream of literature where dictatorships are ended through revolutions and not voluntarily as in our model. In particular, the idea that the initial wealth of a country leads to earlier conflicts has also been put forward also by Boucek et al. (2014). This result seems at odds with the resource curse hypothesis, according to which resource abundance may undermine the democratization process. However, the resource curse hypothesis receives only a moderate empirical support: some economist argue that resource wealth strengthens autocratic regimes (see among others, Ross 2001, and Tsui, 2011); others find no evidence of a positive relationship between resource wealth and the stability of autocratic regimes (Alexeev and Conrad, 2009, and Haber and Menaldo, 2011). The new mechanism presented in this paper –through the investment in education and the development of citizens’ aspiration for democracy– can rationalize these contrasting opinions: resource windfalls cannot alone trigger the democratization process. This also explains why in certain countries with large endowment in natural resources, investment in education is particularly weak (see Gylfason, 2001). In our setting, even though a country has a given large level of windfall rents, human capital accumulation (and therefore democratization) is suboptimal at least in two cases: when the effectiveness of education investment is not large enough and/or when the share of wealth accruing to the elite after democratization is not large enough.

The remainder of the paper is organized as follows. In Section 2, we present the model of the economy. In Section 3, we setup the elite maximization problem. In Section 4 and 5, we first study all the possible solutions, we then compare them to determine the optimal choices, and finally study the implications of the model. We conclude in Section 6. All the proofs are gathered in the appendix.

## 2. The model

Time is continuous, let  $t \in [0, \infty)$ . For notational simplicity, the time index is omitted when there is no risk of confusion. In each period, the elite manages a constant windfall of resources  $R > 0$ . Resources have two possible uses: a part of it can be exported on the international primary good market –let the exported quantity be denoted by  $X$ – and for the remaining part, it can be supplied internally to the manufacturing sector –let this quantity be denoted by  $Q \leq R - X$ . The international price of the resources is constant over time and is denoted  $p^x > 0$ .

National firms operate in a competitive environment and produce a homogeneous commodity  $Y$  using two inputs: resources  $Q$  and human capital  $H$ . With a Cobb-

Douglas specification  $Y = F(Q, H) = A Q^\alpha H^{1-\alpha}$ , resource price and wage (per unit of human capital) are:

$$p = \alpha \frac{Y}{Q} \quad (1)$$

and

$$w = (1 - \alpha) \frac{Y}{H}. \quad (2)$$

The resource rent of the elite, obtained from export  $X$  and national supply  $Q$ , is allocated between the elite's consumption  $C$ , transfers to workers  $\Theta$ , and education investment  $E$ :

$$p^x X + pQ \geq C + \Theta + E. \quad (3)$$

Investment in the education sector increases human capital according to the following accumulation function:

$$\dot{H} = h(E, H) = hE - \delta H \quad (4)$$

where  $h > 0$  measures the effectiveness of the education investment and  $\delta > 0$  is the depreciation rate of human capital.

In each period, a unitary mass of workers –the poor citizens– inelastically supply their human capital  $H$  to national firms and earn an equilibrium wage  $w$ , determined by (2). Their income is completed by the transfers  $\Theta$  and is entirely consumed in each period.

If workers find their consumption not large enough, they can decide to revolt against the elite. The threshold consumption for revolution is the sum of a subsistence consumption –let  $s > 0$  denote such level– and an awareness component, which depends on their human capital level. As outlined in the introduction, the idea is that as workers get more and more educated, they become more aware of the political situation in the country, they become more politically sophisticated, and require better working opportunities. As a result, they would not revolt only if their consumption was sufficiently large. Let the “awareness” component of consumption be linear in human capital, with awareness parameter (the multiplier of human capital)  $\phi > 0$ . Then, workers decision to revolt at each  $t$  is:

$$\begin{cases} \text{revolt} & \text{if } wH + \Theta < s + \phi H \\ \text{not revolt} & \text{otherwise} \end{cases} \quad (5)$$

Let  $T \in \mathbb{R}_+ \cup \{\infty\}$  be the time at which the autocratic regime ends (permanent dictatorship holds when  $T = \infty$ ). A sharing rule describes how the wealth of the

economy is distributed to agents at  $T$ ; in particular, since wealth is an increasing function of human capital, we shall assume that the elite expects a factor  $\pi > 0$  of the amount of human capital at  $T$ . As explained in the introduction, such an assumption may also be interpreted as a Lipsetian trait: human capital is tightly connected with negotiation and absence of violence in Lipset's theory; thus, a continuation payoff increasing with human capital is a natural assumption. Elites highly regard the possibility of leaving the power without being exposed to full expropriation and political violence. The larger parameter  $\pi$ , the larger this impact of human capital on the post-autocratic regime is valued by the elites.

The intertemporal well-being of the elite is given by:

$$U^e = \int_0^T e^{-\rho t} u(C) dt + e^{-\rho T} \pi H_T \quad (6)$$

where the instantaneous utility function is  $u(C) \equiv \frac{C^{1-\gamma}}{1-\gamma}$  with  $\gamma \in (0, 1)$  and  $\rho > 0$  is the discount rate.<sup>2</sup>

The elite is particularly powerful. It is able to control the consumption/income of workers, and thus their willingness to revolt, in three different ways: *(i)* directly, by setting the transfer  $\Theta$ ; *(ii)* indirectly, by deciding how many resources to supply to the national industry  $Q$ ; and *(iii)* dynamically, by investing more or less in education  $E$  and thus setting their level of human capital. Furthermore, the elite controls the political transition process and chooses the timing  $T$  (possibly infinite) for the institutional change.

Before moving to the analysis, we shall assume that the resource windfall is sufficiently large so that it is possible to sustain a dictatorial regime when human capital is zero.

**Assumption 1.**  $p^x R > s$ : *the value of resources is larger than the subsistence consumption of the workers and gives the elite some freedom in how to allocate such wealth.*

### 3. The elite maximization problem

The elite seeks to maximize utility (6), subject to the budget constraint (3), equilibrium prices (1) and (2), the revolution decisions of workers (5), and the dynamics of

---

<sup>2</sup>The assumption that the elasticity of consumption be positive ensures that the utility is positive for any value of consumption. This is needed for the continuation payoff at the time of institutional change, i.e.  $\pi H_T \geq 0$ , to be intrapersonal comparable. When instead  $\gamma > 1$ , utility levels are strictly negative and an immediate institutional change (independently of the human capital level) is always optimal.

human capital (4). The decision variables are the use of resources  $Q$ , own consumption  $C$ , transfers  $\Theta$ , and education  $E$ . Substituting  $\Theta$  from the non-revolt condition of workers (5), the optimization problem of the elite can be written as an optimal stopping problem, where  $T$  is the time until which the constraint is met. Formally:

$$\begin{aligned} \max_{\{Q,E,T\}} \quad & \int_0^T e^{-\rho t} u(p^x (R - Q) + AQ^\alpha H^{1-\alpha} - E - s - \phi H) dt + e^{-\rho T} \pi H(T) \\ \text{s.t.} \quad & \dot{H} = hE - \delta H \\ & H(T) = H_T \text{ is free} \\ & E \geq 0 \\ & \text{with } H(0) = H_0 \text{ given} \end{aligned}$$

Maximizing the criterion with respect to national resource supply  $Q$ , requires that national prices equalize international ones, i.e.  $p = p^x$ , and sets the optimal ratio between resources and human capital as follows:<sup>3</sup>

$$\frac{Q}{H} = \left( \frac{\alpha A}{p^x} \right)^{\frac{1}{1-\alpha}} \quad (7)$$

Let  $\lambda$  be the costate variable associated with human capital. Its equilibrium dynamics is:

$$\dot{\lambda} = (\delta + \rho) \lambda - C^{-\gamma} \left( (1 - \alpha) A \left( \frac{Q}{H} \right)^\alpha - \phi \right) \quad (8)$$

At the equilibrium, optimal investment in education is such that marginal utility of consumption equals the marginal cost of investing in education. Let  $\mu$  be the Lagrange multiplier associated to the positivity constraint on education, then the optimal consumption of the elite is:

$$C = (\lambda h + \mu)^{-\frac{1}{\gamma}} \quad (9)$$

The slackness conditions on education investment  $E$  require:

$$\mu \geq 0 \quad \text{and} \quad \mu E = 0 \quad (10)$$

The optimal time  $T < \infty$  for violating the no-revolt condition is such that the

---

<sup>3</sup>The optimal resource supply of the elite to the national industry would determine a price wedge between international and internal resource prices in case of costly redistributive transfers. In this case, the elite would find it more profitable to redistribute income to workers by oversupplying resources and, indirectly, determining a wage increase. While this extension is potentially relevant for an empirical assessment, the results discussed are not affected.

current value of the Lagrangian be equal to the value of the salvage function; i.e.:

$$u(C(T)) + \lambda [hE(T) - \delta H(T)] = \rho\pi H(T) \quad (11)$$

Finally, the transversality condition requires that:

$$\lambda(T) = \frac{\partial S(H)}{\partial H} \Big|_{H=H(T)} = \pi \quad (12)$$

Convexity of the problem with respect to optimal education investment  $E$  and internally supplied resources  $Q$  guarantees that the second order conditions are always satisfied. The second order conditions for the optimal stopping problem are not necessarily met and will be discussed in more details in the following.

We solve the model in 3 steps: we first study the dynamics of the system when education investments are strictly positive; we then study the case of either zero investments in education or alternating periods of positive and zero investments; and finally we characterize optimality by combining the results.

## 4. Optimality candidates

### 4.1. Education-driven institutional change

Define the following two coefficients:  $\Omega$  is the instantaneous return on human capital and  $\chi$  is the instantaneous return on education investment. These are hereafter defined:

$$\begin{aligned} \Omega &\equiv \frac{1-\alpha}{\alpha} p^x \left( \frac{\alpha A}{p^x} \right)^{\frac{1}{1-\alpha}} - \phi \\ \chi &\equiv h\Omega - \delta. \end{aligned}$$

Straightforward manipulation of the necessary optimality conditions yields the general solution, valid for any solution with strictly positive education  $E > 0$ :

$$\begin{cases} C(t) = C_0 e^{\frac{(\chi-\rho)t}{\gamma}} \\ H(t) = \left( H_0 + \frac{h(p^x R - s)}{\chi} - \frac{\gamma h C_0}{\rho - \chi(1-\gamma)} \right) e^{\chi t} + \frac{\gamma h C_0}{\rho - \chi(1-\gamma)} e^{\frac{(\chi-\rho)t}{\gamma}} - \frac{h(p^x R - s)}{\chi} \\ \lambda(t) = \frac{C_0^{-\gamma}}{h} e^{(\rho-\chi)t} \end{cases} \quad (13)$$

Then, we can establish that:

**Proposition 1.** (i) *There is no solution combining permanent dictatorship and positive education.*

(ii) *There may be a solution combining institutional change and positive education.*

(a) This solution is characterized by an increasing stock of human capital, an institutional change in finite time  $T = T(H_0, R, \pi)$ , and a corresponding end-point stock of human capital

$$H_T = \frac{h}{\rho - \chi} \left( \frac{\gamma(h\pi)^{-\frac{1}{\gamma}}}{1 - \gamma} + p^x R - s \right) > 0. \quad (14)$$

(b) Necessary conditions for the existence of such solution for all  $H_0 \in [0, H_T]$  are:

$$\begin{cases} \chi < \rho \\ p^x R - s > (h\pi)^{-\frac{1}{\gamma}}. \end{cases} \quad (15)$$

(c) Sufficient conditions for the existence of a unique solution are (15) and  $\chi > \underline{\chi}$ , with  $\underline{\chi} \in (0, \rho)$  the unique solution of

$$e^{\frac{\gamma \underline{\chi}}{\rho - \underline{\chi}}} = \frac{\rho[(1 - \gamma)(\rho - \chi(1 - \gamma)) + \gamma^2 \underline{\chi}]}{(1 - \gamma)(\rho - \chi)^2}. \quad (16)$$

The condition  $\chi < \rho$  sets an upper bound on the instantaneous returns to human capital and is a necessary condition for the existence of a solution with education-driven institutional change. Under the opposite condition  $\chi \geq \rho$ , autocracy is too growth-friendly. The elites can invest in education and, due to the high returns, this stimulates growth of output (and citizens' consumption) while being compatible with the respect of the no-revolution constraint. However, the solution with permanent dictatorship and positive education is not relevant because it would either violate the resource constraint, or imply resource imports to become infinite (with  $X$  tending to  $-\infty$ ). The second necessary existence condition in (15) states that resource windfalls net of the intrinsic subsistence consumption level should be larger than the level of consumption the elite just enjoys at the date dictatorship ceases,  $C^1(T)$ . This condition ensures that the elite has enough resources at date  $T$  to not only reach the optimal consumption level  $C^1(T)$  (as it is determined by the optimal stopping condition (12)) but also maintain the education level compatible with an increasing path for human capital (as it is dictated by the second order optimality condition). Finally, a sufficient condition for existence requires that the returns to human capital be higher than a threshold  $\underline{\chi}$ , defined by (16). High enough returns to human capital logically guarantees that it is worthwhile for the elite to engage in the path of education and sustained capital accumulation.

Under  $\chi < \rho$ , the time path of consumption is decreasing whereas the stock of human capital is increasing. The intuition runs as follows. For the elite to find it optimal to democratize it should be able to accumulate a sufficient amount of human

capital, which will directly affect the wealth it will hold in the post-dictatorship regime, and will also guarantee that it can enjoy its wealth in a peaceful environment. Thus investment in human capital should be favored over consumption. Moreover, by investing a lot in human capital, the elite fosters the development of citizens' claims for a freer system through the increasing awareness mechanism. In order to delay the political regime change the elite has no other option but to transfer a lot of resources to the citizens. This also comes at the expense of its own consumption (growth).

Note that under the conditions of Proposition 1, solutions that combine positive education and a revolution in finite time exist for any  $H_0 \leq H_T$ . In other words, the stated conditions guarantee the existence of a solution with education-driven institutional change independently of the initial endowment in human capital. This is a reasonable feature of our model: it would otherwise be difficult to explain why some countries are doomed to dictatorial regimes exclusively based on their initial stock of human capital and would also raise the issue of identifying this initial period (of the development process). Importantly, this doesn't mean that the initial stock of human capital is irrelevant to our analysis. As far as the optimality analysis is concerned, one expects that this variable will be crucial to determine which one of the optimality candidates yields the optimum.

Proposition 1 also gives a first insight for the predictions of our model concerning the Lipsetian links between human capital and democracy, and between resources and democracy. An obvious implication of the proposition is the incompatibility between permanent dictatorship and education (i), which is in accordance with Lipset's theory. The reason of this incompatibility follows from the no-revolt constraint: it's impossible to satisfy this constraint with a positive level of education investment as the consumption aspirations associated with positive education will end up exceeding the redistribution capacities of the elite. In addition, it is worth emphasizing some other interesting features of the first optimality candidates. They are summarized in the next two corollaries.

**Corollary 1.** *The solution with education-driven institutional change is impossible if, ceteris paribus,*

- (i) *the windfall of resources,  $R$ , is not large enough;*
- (ii) *the international price of resources,  $p^x$ , is not large enough;*
- (iii) *the effectiveness of the education investment,  $h$ , is not large enough; or*
- (iv) *the share of wealth accruing to the elite after the institutional change,  $\pi$ , is not large enough.*

These properties are in line with Lipset's theory in two essential aspects.<sup>4</sup> First of all, the model predicts that a large amount of resources (or of their export price) is a precondition for the emergence of a non-dictatorial regime through human capital accumulation. As mentioned in the introduction, this goes into the same direction as most of the theories developed following Acemoglu and Robinson (2006), in which revolutions are launched by citizens and not planned by the governing elites.

However, the resource wealth of a country (measured by  $p^x R$ ) is not sufficient to guarantee that a solution with investment in education and institutional change exists. Two further factors are necessary conditions for the emergence of a democratization path: a sufficient return on the investment in education and a sufficient reward for the elite at the time of institutional change. Importantly, the interaction between the resource wealth and these factors is likely to be responsible for the mixed support for the natural resource curse hypothesis (see Ross, 2001, Hodler, 2006, Alexeev and Conrad, 2009, Tsui, 2011, and Haber and Menaldo, 2011) and is in line with the empirical studies pointing at the mis-management of education in several oil-exporting countries (see Gylfason, 2001).

A more Lipsetian reading of the corollary would simply observe that there is no institutional-change driven by human capital accumulation if the dictators do not value human capital enough, given the resource wealth at their disposal.

Beyond the existence of a solution with institutional change, it is also useful to stress how the time-to-democratization is affected by the parameters of the model (see the comparative statics exercise at the end of Appendix A.1).

**Corollary 2.** *The optimal time for institutional change,  $T = T(H_0, R, \pi)$ , is:*

- (i) decreasing in the initial endowment in human capital,  $H_0$ ;*
- (ii) decreasing in the constant flow of resource windfalls,  $R$ ; and*
- (iii) decreasing in the share of wealth accruing to the elite after the institutional change,  $\pi$ .*

The first two features strengthen the correlation between wealth and democratization discussed before. The larger is the initial stock of human capital (another possible measure of human wealth) or the windfall of resources, the quicker are the Lipsetian elites in driving the country into an institutional change. While the larger windfall is also associated to a larger level of human capital at the time of institutional change, such an effect is absent for the initial human capital level.

---

<sup>4</sup>Note that they are a direct consequence of the second necessary condition in (15). Also note that the parameters  $R$  and  $\pi$  do not show up in the sufficient condition  $\chi > \underline{\chi}$  since they don't enter into the expressions of  $\Omega$ . In contrast, the parameters  $h$  and  $p^x$  enter this condition through  $\Omega$ .

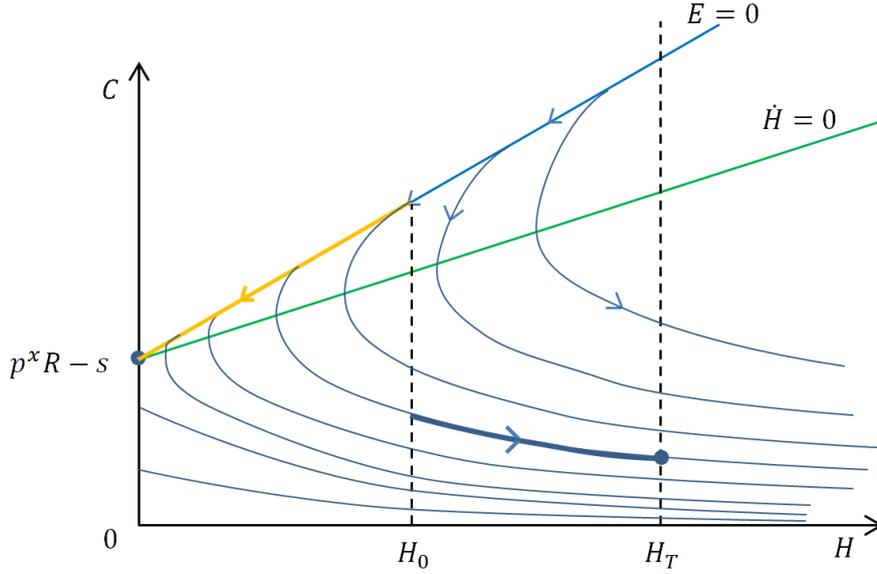


Figure 1: Phase diagram.

Finally, the optimal time-to-democratization is decreasing in  $\pi$ . The elite compensates a less favorable sharing rule by increasing the human capital of the country at the institutional change,  $H_T$ . This requires a longer period of investment in education.

To end up this discussion, it is important to measure the elite's payoff associated to the solution with education-driven institutional change. Let such optimality candidate be referred to as regime 1; then the present value (for the elite) of following this regime is given by (hereafter, the optimal time for institutional change is expressed in terms of  $H_0$  only):

$$V^1(H_0) = e^{-\rho T(H_0)} \left( \frac{\gamma(h\pi)^{-\frac{(1-\gamma)}{\gamma}}}{(1-\gamma)(\rho - \chi(1-\gamma))} (e^{\frac{(\rho - \chi(1-\gamma)T(H_0))}{\gamma}} - 1) + \pi H_T \right).$$

The typical dynamics corresponding to this first possible solution is depicted in Figure 1.

In the next section the other optimality candidates are briefly reviewed.

## 4.2. No-education and permanent dictatorship

The general solution corresponding to no investments in education, i.e.  $E = 0$ , is given by (the superscript 2 refers to the regime with no education):

$$\begin{aligned} H^2(t) &= H e^{-\delta t} \\ C^2(t) &= p^x R - s + \Omega H^2(t) \\ \lambda^2(t) &= e^{(\rho+\delta)t} \left( L - \int \Omega C^2(u)^{-\gamma} e^{-(\rho+\delta)u} du \right) \end{aligned}$$

with  $H, L$  two constants to be determined. It can easily be shown that:

**Proposition 2.** *(i) There always exists a solution combining permanent dictatorship with no investment in education. This solution is characterized by a decreasing flow of consumption and a decreasing stock of human capital. Consumption asymptotically converges toward  $C^2(\infty) = p^x R - s$ , while the stock of human capital vanishes.*

*(ii) There may be a solution alternating periods of investment in education with periods of no investment, but these never provide a candidate for optimality.*

Solutions with no education exist for any level of the stock of human capital; whereas, solutions featuring a regime change from positive to zero education can be disregarded because they are always dominated by the former type. It means that we are left with two optimality candidates with distinct features. Put differently, the trade-off faced by the elite is crystal-clear. Either it chooses to rely on resource windfalls and not to invest in education in order to labor force uneducated and docile. But this requires to sacrifice education-driven economic growth. Or, the elite engages in a policy of sustained investment in education, which promotes the accumulation of human capital but also goes hand in hand with the development of people claims for democracy. As expected, variables like the returns to education (and human capital), the initial stock of human capital, the discount rate but also the share of wealth accruing to the elite after a revolution will play a central role in explaining what is the elite's best option.

For the remainder of the analysis, it is useful to retrieve the value function corresponding to the optimality candidate with no education. The present value, which is given by:

$$V^2(H_0) = \int_0^\infty \frac{1}{1-\gamma} (p^x R - s + \Omega H_0 e^{-\delta t})^{1-\gamma} e^{-\rho t} dt.$$

Before determining the optimal choice of the elite, we compare regime 1 and 2 in terms of their implications for the link between inequalities and institutional change.

### 4.3. Inequality implications of a Lipsetian elite

So far, we have addressed the links income-institutional change and human capital-institutional change. In these respect, we have shown that the predictions of the model are consistent with Lipset's theory. It remains to study the link inequalities-institutional change.

As workers are a homogeneous mass of individuals, the only way to appraise inequalities in a direct and elementary way is by tracking the consumption of the elites vs the consumption of workers. Although this is not completely in the spirit of Lipset's theory concerning this aspect (see Jung and Sunde, 2013, for a tighter connection), this exercise turns out to be worthwhile. Recall that the workers' income is entirely devoted to consumption. At any solution, we have  $C_W^i(t) = s + \phi H^i(t)$  for  $i = 1, 2$ . Index the consumption of the rich elite by  $R$ . Let  $I^i(t) \equiv \frac{C_P^i(t)}{C_R^i(t)}$  be the index of inequalities at solution  $i = 1, 2$ . Then, we can establish the following result.

**Proposition 3.** *At the solution with education-driven institutional change, inequalities continuously shrink. At the solution with permanent dictatorship and no investment in education, the opposite result holds if:*

$$R > \frac{(1 - \alpha)s}{\phi\alpha} \left( \frac{\alpha A}{p^x} \right)^{\frac{1}{1-\alpha}}. \quad (17)$$

A couple of comments are in order here. First, we observe that along the transition process to non-dictatorship, inequalities do decrease. It is as if in order to prepare the ground for a democratic regime, the elite has to progressively reduce the income (consumption) gap between the two groups. Intuitively, since the elite invests in human capital along this path, growth is stimulated. But the positive growth effect is dominated by the negative effect due to increasing awareness and the elite has no option but to sacrifice part of her consumption to satisfy the no-revolt constraint and delay the date of leaving office. Second and not surprisingly, permanent dictatorship implies a widening of inequalities if resource windfalls are high, the awareness cost is large, the international resource price is high, and the level of subsistence consumption is low. Under these conditions, the dictator is able to fill the revolution constraint at least cost. By not investing in human capital, the people are maintained under control while the elite becomes richer and richer relative to the workers.

The next section investigates the optimality of the above-identified solutions.

## 5. Optimal solution

The optimality analysis boils down to a study of the relative performance of the solution with education-driven institutional change vs. the solution with permanent dictatorship and no-education. To conduct this analysis, we proceed to a comparison of the present values associated with our two optimality candidates. The results can be summarized as follows:

**Proposition 4.** *Let  $H_0 \in [0, H_T]$ . The following cases can arise:*

*(i) The solution with permanent dictatorship and no-education is optimal for all  $H_0$  iff  $V^2(H_T) > V^1(H_T)$ ;*

*(ii) The solution with education-driven institutional change is optimal for all  $H_0$  iff  $V^1(0) > V^2(0)$ ;*

*(iii) Otherwise, a human capital poverty trap arises. There exists  $\bar{H} \in [0, H_T]$  such that the solution with education-driven institutional change is optimal iff  $H_0 \geq \bar{H}$ .*

Both the no-education regime with persistent dictatorship and the education regime with democratization can arise. Depending on the parameters, it might be possible that: *(i)* the first alternative is chosen independently of the initial stock of human capital; *(ii)* the second alternative is chosen independently of the initial stock of human capital; and *(iii)* the regime choice depends on the initial human capital stock, a low stock is associated to no-education investment and infinite horizon dictatorship while a large stock is associated to education investment and democratization in finite time.

This result sheds light on the relationship between education, development, and democratization. First, education is necessary for both development and democratization: it is the engine of economic growth and, by increasing the workers awareness, it is also responsible for the institutional change.<sup>5</sup> Second, education investments might be optimal for the ruling elite, despite it might lead to more democratic institutions, as their political power gets substituted by economic returns. Third, the existence of a poverty trap is particularly interesting for it teaches that development aid leading to “small” increases in human capital might not be sufficient for a regime switch and thus fails to have permanent effects on development and institutions of the recipient country.

---

<sup>5</sup>An effective education is however not a sufficient condition for institutional change: when the return on education are sufficiently high, the additional wealth in the hands of the elites could be sufficient to compensate for the larger awareness of the workers. In this case, a developing dictatorship could be sustained, although no equilibrium paths exist.

The next result further emphasizes the conditions under which the elite finds the democratization path optimal. Even though it proves difficult to express the condition  $V^2(H_T) > V^1(H_T)$  in terms of the parameters, we can establish the following sufficient condition for optimality.

**Proposition 5.** *The solution with education-driven institutional change is optimal for all  $H_0 \in [0, H_T]$  if:*

$$p^x R - s > e^{\frac{\gamma}{1+\gamma}} (h\pi)^{-\frac{1}{\gamma}}. \quad (18)$$

This sufficient condition for optimality is stronger than the second necessary condition for existence of such solution in (15). The next corollary mirrors Corollary 1 and confirms the previous intuition about which factors are crucial for the decision of a Lipsetian elite to educate the population and drive the country out of autocracy.

**Corollary 3.** *The solution with education-driven institutional change is optimal if, ceteris paribus,*

- (i) *the windfall of resources,  $R$ , is sufficiently large;*
- (ii) *the international price of resources,  $p^x$ , is sufficiently large; or*
- (iii) *the share of wealth accruing to the elite after the institutional change,  $\pi$ , is sufficiently large.*

The effect of a windfall of resources  $R$  and the share of wealth  $\pi$  were already discussed in the previous sections; as the intuition is unchanged, we shall skip it here. We shall instead stress the way the international price of resources  $p^x$  leads to optimality and discuss the absence of the effectiveness of the education investment  $h$ , despite its presence in the sufficient condition (18). When the international price of resources increases, two counterbalancing effects arise: on the one side, the available wealth of the elite is increased; on the other side, supplying resources to the national industry becomes relatively less appealing, the returns on human capital  $\Omega$  are reduced and, consequently, so are the returns on education investment  $\chi$ . When  $p^x$  is large, a further implication holds: the condition for uniqueness of the solution with education-driven institutional change will be violated and possible multiple education choices can arise.

Finally, a large effectiveness of education  $h$  might not ensure the optimality of the democratization regime with education. Formally, this counter intuitive result is due to the first necessary condition (15), stating that the return on education  $\chi$  is bounded from above. As mentioned, when this condition is violated the returns on education are non-realistically high and would allow the autocratic elite can invest in education, pay-out the increasing costs of revolt-avoiding redistribution, and increase the own consumption indefinitely.

Proposition 5 and Corollary 3 illustrate that institutional change initiated by the Lipsetian elite is a matter of having the right conditions. A large stock of resources might not trigger education policies and democratization if the education sector doesn't ensure good enough economic returns to the elite. A permanent positive shock to international resource prices might give the elite the wealth needed to invest in education and human capital accumulation, but this opportunity will not be taken if the wealth prospects at the time of institutional change are not sufficiently compelling.

## 6. Conclusion

In this paper we have developed a Lipsetian theory of institutional change. The key mechanism at play is the feedback effect of education policies on the awareness of workers, measuring their understanding of the political system, their political sophistication, and, more in general, their reluctance to accept a dictatorship. Human capital makes national industry more productive and is the engine of economic growth, but has a political cost in terms of the larger services/transfers that workers require to be refrained from revolting.

The main result of the paper is to show that two possible regimes can arise, depending on the relative magnitude of the education incentives. The first case is that of countries where the ruling elite supports a permanent dictatorship characterized by low education, low growth, low level of worker's life conditions. The second case is that of countries where the elite favors long-term economic interests by investing in education and human capital accumulation, achieve a high growth path and improving life conditions; the cost of this development is an unavoidable removal of dictatorship at finite time.

Our theory is compatible with Lipset in the three essential dimensions: the positive link between human capital and institutional change, the positive link between income and institutional change and, in a more stylized fashion, the negative link between inequality and institutional change.

## References

- [1] Acemoglu, D. & J. Robinson, 2001. A Theory of Political Transitions, *American Economic Review*, vol. 91(4), pages 938-963.
- [2] Acemoglu, D. & S. Johnson & J. Robinson & P. Yared, 2005. From Education to

- Democracy?, *American Economic Review, Papers and Proceedings*, vol. 95(2), pages 44-49.
- [3] Acemoglu, D. & J. Robinson, 2006. *Economic Origins of Dictatorship and Democracy*, Cambridge University Press.
- [4] Alexeev, M. & R. Conrad. 2009. The Elusive Curse of Oil. *Review of Economics and Statistics*, vol. 91(3), pages 586-598.
- [5] Amit, R. 1986. Petroleum reservoir exploitation: Switching from primary to secondary recovery. *Operations Research*, vol 34, pages 534-549.
- [6] Barro, R., 1999. Determinants of Democracy, *Journal of Political Economy*, vol. 107, pages 158-183.
- [7] Boucekine, R. & F. Prieur & K. Puzon, 2014. On Political Regime Changes in Arab Countries, AMSE Working Papers 1401, Aix-Marseille School of Economics.
- [8] Campante, F. & D. Chor, 2012. Why Was the Arab World Poised for Revolution? Schooling, Economic Opportunities, and the Arab Spring, *Journal of Economic Perspectives*, vol. 26(2), pages 167-88.
- [9] Castello-Climent, A. & A. Mukhopadhyay, 2013. Mass education or a minority well educated elite in the process of growth: The case of India, *Journal of Development Economics*, vol. 105(C), pages 303-320.
- [10] Castello-Climent, A. 2008. On the distribution of education and democracy, *Journal of Development Economics*, vol. 87(2), pages 179-190.
- [11] Chomsky, N., 1996. *Class Warfare: Interviews with David Barsamian*. Common Courage Press Monroe.
- [12] Collier, P. & A. Hoeffler, 2004. Greed and grievance in civil war, *Oxford Economic Papers*, vol. 56(4), pages 563-595.
- [13] Glaeser, E. & G. Ponzetto & A. Shleifer, 2007. Why does democracy need education?, *Journal of Economic Growth*, vol. 12(2), pages 77-99.
- [14] Glaeser, E. & R. La Porta & F. Lopez-de-Silanes & A. Shleifer, 2004. Do Institutions Cause Growth?, *Journal of Economic Growth*, vol. 9(3), pages 271-303.
- [15] Gylfason, T., 2001. Natural resources, education, and economic development, *European Economic Review*, vol. 45(4-6), pages 847-859.

- [16] Haber, S. & V. Menaldo. 2011. Do Natural Resources Fuel Authoritarianism? A Reappraisal of the Resource Curse. *American Political Science Review*, vol. 105(1), pages 1-26.
- [17] Hodler, R. 2006. The curse of natural resources in fractionalized countries, *European Economic Review*, vol. 50(6), pages 1367-1386.
- [18] Jung, F. & U. Sunde, 2014. Income, inequality, and the stability of democracy: Another look at the Lipset hypothesis, *European Journal of Political Economy*, vol. 35(C), pages 52-74.
- [19] Lipset, S., 1960. *Political Man: The Social Bases of Modern Politics*. New York: Doubleday.
- [20] Luskin, R., 1990. Explaining Political Sophistication, *Political Behavior*, vol. 12 (4), pages 331-361.
- [21] Neumann, R., 1986. *The Paradox of Mass Politics: Knowledge and Opinion in the American Electorate*. Harvard University Press.
- [22] Przeworski, A. & M. Alvarez & J. Cheibub & F. Limongi, 2000. *Democracy and Development: Political Institutions and Well-being in the World*. Cambridge: Cambridge University Press.
- [23] Rodrik, D. & A. Subramanian & F. Trebbi, 2004. Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development, *Journal of Economic Growth*, vol. 9(2), pages 131-165.
- [24] Ross, M. 2001. Does Oil Hinder Democracy? *World Politics*, vol. 53, pages 325-361.
- [25] Tsui, K. 2011. More Oil, Less Democracy: Evidence from the Worldwide Crude Oil Discoveries. *Economic Journal*, vol. 121, pages 89-115.
- [26] van der Ploeg, F. & D. Rohner, 2012. War and natural resource exploitation, *European Economic Review*, vol. 56(8), pages 1714-1729.
- [27] Zaller, J., 1992. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge university press.

## A. Proofs

### A.1. Proof of Proposition 1

#### A.1.1. Item (i)

When regime 1 is permanent, from the transversality condition we must have:

$$C_0 = \frac{\rho - \chi(1 - \gamma)}{\gamma h} \left( H_0 + \frac{h(p^x R - s)}{\chi} \right)$$

So existence requires  $\rho > \chi(1 - \gamma)$ . More generally, the ordering between  $\chi$  and  $\rho$  is crucial to discuss the nature of any potential solution. First assume that  $\rho > \chi$ , then from the expression of the stock of human capital

$$H(t) = \frac{\gamma h C_0}{\rho - \chi(1 - \gamma)} e^{\frac{(\chi - \rho)t}{\gamma}} - \frac{h(p^x R - s)}{\chi},$$

we obtain that  $H_\infty = -\frac{h(p^x R - s)}{\chi} < 0$ , which is impossible. Thus, in this case, necessarily the system reaches the frontier  $E = 0$  in finite time. Next, consider the alternative,  $\chi(1 - \gamma) < \rho < \chi$  ( $\rho = \chi$  is a knife-edge situation). In this case, both consumption and human capital are varying at the constant rate  $(\chi - \rho)/\gamma$ . Hereafter we disregard this case because either it is not compatible with the resource constraint or, if resource imports are allowed, it implies that  $X \rightarrow -\infty$ , which is not a reasonable feature of the model.

#### A.1.2. Item (ii)

From the second optimality condition 12, for the stopping time  $T$ , we have

$$C_0 = (h\pi)^{-\frac{1}{\gamma}} e^{-\frac{(\chi - \rho)T}{\gamma}}.$$

Substituting the expression in (13), we obtain:

$$\begin{aligned} C(t) &= (h\pi)^{-\frac{1}{\gamma}} e^{\frac{1}{\gamma}(\chi - \rho)(t - T)} \\ H(t) &= \varphi(T) e^{\chi t} + \frac{\gamma h (h\pi)^{-\frac{1}{\gamma}} e^{\frac{(\chi - \rho)(t - T)}{\gamma}}}{\rho - \chi(1 - \gamma)} - \frac{h(p^x R - s)}{\chi} \end{aligned}$$

where  $\varphi(T) \equiv \left( H_0 + \frac{h(p^x R - s)}{\chi} - \frac{\gamma h (h\pi)^{-\frac{1}{\gamma}} e^{-\frac{(\chi - \rho)T}{\gamma}}}{\rho - \chi(1 - \gamma)} \right)$ . The second optimal stopping condition (11) can be rewritten as:

$$\frac{\gamma (h\pi)^{-\frac{1 - \gamma}{\gamma}}}{1 - \gamma} + h\pi(p^x R - s) + \chi\pi H(T) = \rho\pi H(T) \quad (19)$$

where the LHS is the marginal benefit of waiting (achieving utility from consumption  $C(T)$  and the advantage from increasing human capital  $\pi\dot{H}(T)$ ), while the RHS is the marginal cost of waiting (delaying the scrap value). Since  $\frac{\gamma(h\pi)^{-\frac{1-\gamma}{\gamma}}}{1-\gamma} + h\pi(p^x R - s) > 0$ ,  $\chi \geq \rho$  implies that LHS > RHS for each level of human capital.

When  $\chi < \rho$ , the RHS increases faster than the LHS: marginal benefit is first larger and then smaller than marginal cost. Then, we can define the optimal endpoint  $H(T) = H_T$

$$H_T = \frac{h}{\rho - \chi} \left( \frac{\gamma(h\pi)^{-\frac{1}{\gamma}}}{1 - \gamma} + p^x R - s \right) > 0. \quad (20)$$

The second order condition (SOC) for the optimal stopping problem is satisfied iff  $(\chi - \rho)\dot{H}(T) < 0$ , which requires  $\dot{H}(T) > 0$ .

From the continuity of the state variable, we have:

$$H_T = \varphi(T) e^{\chi T} + \frac{\gamma h (h\pi)^{-\frac{1}{\gamma}}}{\rho - \chi(1 - \gamma)} - \frac{h(p^x R - s)}{\chi} \quad (21)$$

Rearranging we obtain:

$$\varphi(T) e^{\chi T} = \frac{\rho h}{\rho - \chi} \left( \frac{p^x R - s}{\chi} + \frac{\gamma^2 (h\pi)^{-\frac{1}{\gamma}}}{(1 - \gamma)(\rho - \chi(1 - \gamma))} \right) \quad (22)$$

Let  $F(T)$  be the LHS and  $G > 0$  the RHS. By the monotonicity of the path  $\{H(t)\}_{t=0}^T$  (see the next item) and the SOC,  $H_0 \leq H_T$  needs to hold. Hereafter, let's pay attention to an interior solution, i.e.  $H_0 < H_T$ . This is equivalent to  $F(0) < G$ . Moreover,  $F(\infty) = -\infty$ . The sign of the derivative of  $F(T)$ :

$$F'(T) = e^{\chi T} \left( \chi H_0 + h(p^x R - s) - h(h\pi)^{-\frac{1}{\gamma}} e^{-\frac{(\chi - \rho)T}{\gamma}} \right).$$

For the existence of an optimal interior  $T$  for democratization, it's first necessary that  $F'(0) > 0$ , which equivalent to:

$$H_0 > \underline{H}_0 = \frac{h(h\pi)^{-\frac{1}{\gamma}} - h(p^x R - s)}{\chi}, \quad (23)$$

and for the interval  $(\underline{H}_0, H_T)$  to be non-empty, we must impose

$$p^x R - s + \frac{(h\pi)^{-\frac{1}{\gamma}}(\chi - \rho(1 - \gamma))}{\rho(1 - \gamma)} > 0. \quad (24)$$

This is equivalent to the SOC.

One way to tackle (and simplify) the existence issue is to make sure that the optimality candidate exists for any initial level of the stock of human capital  $H_0 \in [0, H_T]$ . Otherwise, we would have to discuss and justify the relevance of some particular intervals for existence. Thus, hereafter we scrutinize the conditions under which such a candidate exists for  $H_0 = 0$ . Put  $H_0 = 0$  in all the expressions above. Then,  $F'(0) > 0$  simplifies to

$$p^x R - s > (h\pi)^{-\frac{1}{\gamma}}, \quad (25)$$

and under this condition the SOC is satisfied.

Next we define the value that maximizes  $F(T)$ , say  $\tilde{T}$ , as:

$$\tilde{T} = \frac{\gamma}{\rho - \chi} \ln \left( \frac{p^x R - s}{(h\pi)^{-\frac{1}{\gamma}}} \right), \quad (26)$$

and for the existence of (at most two)  $T^* > 0$  that solve(s) (22), it must hold that  $F(\tilde{T}) > G$ . This condition can be restated as a condition on the parameters of the technology and preferences. Indeed,  $F(\tilde{T}) > G \Leftrightarrow$

$$\frac{(\rho - \chi)^2}{\rho(\rho - \chi(1 - \gamma))} e^{\frac{\gamma\chi}{\rho - \chi}} \left[ \frac{p^x R - s}{(h\pi)^{-\frac{1}{\gamma}}} \right]^2 > \frac{p^x R - s}{(h\pi)^{-\frac{1}{\gamma}}} + \frac{\gamma^2 \chi}{(1 - \gamma)(\rho - \chi(1 - \gamma))},$$

which is a simple polynomial of degree 2 in  $\frac{p^x R - s}{(h\pi)^{-\frac{1}{\gamma}}}$ . Now, under (25), this ratio is larger than 1. So, a sufficient condition for  $F(\tilde{T}) > G$  is:

$$e^{\frac{\gamma\chi}{\rho - \chi}} > \frac{\rho[(1 - \gamma)(\rho - \chi(1 - \gamma)) + \gamma^2 \chi]}{(1 - \gamma)(\rho - \chi)^2}, \quad (27)$$

where both terms of the inequality above are greater than 1. A quick inspection of the properties of the LHS and RHS of (27), seen as functions of parameter  $\chi$ , reveals that there exists a unique threshold  $\underline{\chi} \in (0, \rho)$  such that (27) holds iff  $\underline{\chi} < \chi$ .

The reasoning above is also valid for any  $H_0 > 0$ . Actually, the existence of a solution for the particular value  $H_0 = 0$  implies that such a solution exists for any  $H_0 > 0$ . In general, the optimal stopping time can be expressed as a function of  $H_0$ :  $T^* = T(H_0)$ , with, by differentiating (22)

$$\frac{\partial T}{\partial H_0} = - \frac{1}{\chi\varphi(T(H_0)) + \varphi'(T(H_0))}.$$

Given that we want  $\frac{\partial T}{\partial H_0} < 0$  (uniqueness of the optimal trajectory), only the solution corresponding the increasing part of  $F(T)$  is relevant, i.e. one has  $F'(T) > 0$ , which is equivalent to:

$$p^x R - s > e^{\frac{(\rho-\chi)T}{\gamma}} (h\pi)^{-\frac{1}{\gamma}}. \quad (28)$$

Comparative statics on  $T^*$  that solves (22), given that this equation can be rewritten as  $F(T, R, \pi) = G(R, \pi)$ . From the implicit function theorem, we have:

$$\begin{aligned} \frac{\partial T}{\partial R} &= \frac{\frac{\partial G}{\partial R} - \frac{\partial F}{\partial R}}{\frac{\partial F}{\partial T}} \\ \frac{\partial T}{\partial \pi} &= \frac{\frac{\partial G}{\partial \pi} - \frac{\partial F}{\partial \pi}}{\frac{\partial F}{\partial T}} \end{aligned}$$

given that  $\frac{\partial G}{\partial R} - \frac{\partial F}{\partial R} = \frac{hp^x(\rho-1)}{\chi} < 0$  and  $\frac{\partial G}{\partial \pi} - \frac{\partial F}{\partial \pi} = -\frac{h^2(h\pi)^{-\frac{1}{\gamma}}}{\rho-\chi(1-\gamma)} \left( \frac{\rho\gamma}{(\rho-\chi)(1-\gamma)} + e^{\frac{(\rho-\chi)T}{\gamma}} \right) < 0$ . Now given that we chose the solution that satisfies  $\frac{\partial F}{\partial T} > 0$  (because we want  $T'(H_0) < 0$  for all  $H_0 \geq 0$ ), we can conclude that  $T^*$  is decreasing w.r.t both  $R$  and  $\pi$ .

### A.1.3. Monotonicity of trajectories

The value function at any  $H^1(t_i) = H_i$  taken on the optimal path is given by:

$$V^1(H_i) = e^{-\rho\theta(H_i)} \left( \frac{\gamma(h\pi)^{-\frac{(1-\gamma)}{\gamma}}}{(1-\gamma)(\rho-\chi(1-\gamma))} \left[ e^{\frac{(\rho-\chi(1-\gamma))\theta(H_i)}{\gamma}} - 1 \right] + \pi H_T \right)$$

with  $\theta(H_i) = T(H_i) - t_i$ , the optimal time-to-go before stopping, which doesn't depend on  $t_i$ .

If there exists an optimal trajectory of type 1 from some  $H_0$  with  $H(t)$  non monotone, it must be true that  $H$  is decreasing first, then increasing. This implies that there exists  $(t_1, t_2)$ , with  $t_1 < t_2$ , such that:  $H^1(t_1) = H^1(t_2)$ . Thus, we have  $\theta(H_1) = \theta(H_2)$ : The time that elapses between  $t_1$  and  $T(H_1)$  must be the same as the one between  $t_2$  and  $T(H_2)$ . This yields a contradiction because the optimal trajectory is uniquely defined ( $H_T$  is invariant) and (initial) consumptions at  $t_1$  and  $t_2$  necessarily differ.

Finally note that the solution with positive education and a revolution in finite time yields the following present value to the elites:

$$V^1(H_0) = e^{-\rho T(H_0)} \left( \frac{\gamma(h\pi)^{-\frac{(1-\gamma)}{\gamma}}}{(1-\gamma)(\rho-\chi(1-\gamma))} \left( e^{\frac{(\rho-\chi(1-\gamma)T(H_0))}{\gamma}} - 1 \right) + \pi H_T \right) \quad (29)$$

This completes the proof of Proposition 1.

## A.2. Proof of Proposition 2

### A.2.1. Item (i)

If regime 2 is permanent, then from the transversality condition  $L = 0$ , and the solution reduces to (using the superscript 2):

$$\begin{aligned} H^2(t) &= H_0 e^{-\delta t} \\ C^2(t) &= p^x R - s + \Omega H^2(t) \\ \lambda^2(t) &= -e^{(\rho+\delta)t} \int_0^t \Omega C(u)^{-\gamma} e^{-(\rho+\delta)u} du (< 0). \end{aligned}$$

The value function is given by:

$$V^2(H_0) = \int_0^\infty \frac{1}{1-\gamma} (p^x R - s + \Omega H_0 e^{-\delta t})^{1-\gamma} e^{-\rho t} dt. \quad (30)$$

### A.2.2. Item (ii)

Consider a trajectory  $(\{H^1\}, \{C^1\})$  that reaches the locus  $E = 0$  at date  $t_1$  for some stock  $H^1(t_1) = \tilde{H}$  and consumption  $C^1(t_1) = \tilde{C}$ . From the dynamical system, both  $H^1$  and  $C^1$  are all decreasing w.r.t. time. The approach is to consider a solution with permanent  $E = 0$  as a limit case of the solution with a regime change from  $E > 0$  to  $E = 0$ . Let's work with the general solution obtained by combining regimes 1 and 2. For the time being, let  $t_1$  be given. Recall that the general solution in each regime is:

$$\begin{aligned} C^1(t) &= C_0 e^{\frac{(\chi-\rho)t}{\gamma}} \\ H^1(t) &= \left( H_0 + \frac{h(p^x R - s)}{\chi} - \frac{\gamma h C_0}{\rho - \chi(1-\gamma)} \right) e^{\chi t} + \frac{\gamma h C_0}{\rho - \chi(1-\gamma)} e^{\frac{(\chi-\rho)t}{\gamma}} - \frac{h(p^x R - s)}{\chi} \end{aligned}$$

and,

$$\begin{aligned} H^2(t) &= \tilde{H} e^{-\delta(t-t_1)} \\ C^2(t) &= p^x R - s + \Omega H^2(t) \end{aligned}$$

From the continuity of consumption at  $t_1$ , we obtain:  $C_0 = (p^x R - s + \Omega \tilde{H}) e^{-\frac{(\chi-\rho)}{\gamma} t_1}$  and  $C^1(t) = (p^x R - s + \Omega \tilde{H}) e^{\frac{(\chi-\rho)(t-t_1)}{\gamma}}$

From the continuity of the state variable at  $t_1$ ,  $\tilde{H}$  can be expressed as a function of  $t_1$ :  $\tilde{H} = \zeta(t_1)$  with

$$\zeta(t_1) = \frac{(\rho - \chi(1 - \gamma)) \left[ \left( H_0 + \frac{h(p^x R - s)}{\chi} - \frac{\gamma h(p^x R - s)}{\rho - \chi(1 - \gamma)} e^{-\frac{\chi - \rho}{\gamma} t_1} \right) e^{\chi t_1} - \frac{h(\rho - \chi)(p^x R - s)}{\chi(\rho - \chi(1 - \gamma))} \right]}{\rho - \chi(1 - \gamma) + \gamma \Omega h \left( e^{\frac{(\rho - \chi(1 - \gamma))}{\gamma} t_1} - 1 \right)}$$

So the value corresponding to this trajectory can be written as:

$$V(t_1) = \frac{1}{1-\gamma} \left[ \int_0^{t_1} (p^x R - s + \Omega\zeta(t_1))^{1-\gamma} e^{\frac{(1-\gamma)(\chi-\rho)}{\gamma}(t-t_1)} e^{-\rho t} dt + \int_{t_1}^{\infty} (p^x R - s + \Omega\zeta(t_1)e^{-\delta(t-t_1)})^{1-\gamma} e^{-\rho t} dt \right]$$

Taking the derivative w.r.t  $t_1$  yields:

$$\begin{aligned} \frac{\partial V}{\partial t_1} &= \frac{1}{1-\gamma} (p^x R - s + \Omega\zeta(t_1))^{1-\gamma} e^{-\rho t_1} + \\ &+ \frac{1}{1-\gamma} \int_0^{t_1} e^{-\frac{(\rho-\chi(1-\gamma))t}{\gamma}} e^{\frac{(1-\gamma)(\rho-\chi)}{\gamma}t_1} (p^x R - s + \Omega\zeta(t_1))^{-\gamma} [(1-\gamma)\Omega\zeta'(t_1) + \\ &+ \frac{(1-\gamma)(\rho-\chi)}{\gamma} (p^x R - s + \Omega\zeta(t_1))] - \\ &+ \frac{1}{1-\gamma} (p^x R - s + \Omega\zeta(t_1))^{1-\gamma} e^{-\rho t_1} + \\ &+ \int_{t_1}^{\infty} \Omega(\zeta'(t_1) + \delta\zeta(t_1))e^{-\delta(t-t_1)} (p^x R - s + \Omega\zeta(t_1)e^{-\delta(t-t_1)})^{-\gamma} e^{-\rho t} dt \end{aligned}$$

Taking the limit when  $t_1 \rightarrow 0$ , we obtain:

$$\lim_{t_1 \rightarrow 0} \frac{\partial V}{\partial t_1} = \int_0^{\infty} \Omega(\zeta'(0) + \delta\zeta(0)) (p^x R - s + \Omega\zeta(t_1)e^{-\delta t})^{-\gamma} e^{-(\delta+\rho)t} dt,$$

and it's clear that the sign of the limit is determined by the sign of the expression  $\zeta'(0) + \delta\zeta(0)$ . Direct computations yield:  $\zeta(0) = H_0$  and the derivative of  $\zeta(t_1)$  evaluated at  $t_1 = 0$  is given by:  $\zeta'(0) = -\delta H_0$ . Thus, we obtain  $\zeta'(0) + \delta\zeta(0) = 0$ . From multi-stage optimal control theory (interpreting the change from  $E > 0$  to  $E = 0$  as a regime switching problem), we know that a necessary condition for an immediate switch  $t_1 = 0$  is  $\lim_{t_1 \rightarrow 0} \frac{\partial V}{\partial t_1} \leq 0$  (see Amit [5], Theorem 1). Thus trajectories of the 1-2 type are always dominated by the ones associated with permanent  $E = 0$ .

The last eventuality is a candidate with a regime change from 2 to 1. Suppose the economy starts in regime  $E = 0$  and enters the region with positive education at  $t_1 < \infty$ . Then, there are two options. Either the economy stays in region 1 till the revolution. This would implies the crossing of the locus  $\dot{H} = 0$  in finite time. But this is excluded by the argument provided in Appendix A.1 because the trajectory  $\{H^1(t)\}$  must be monotonous. Or, the economy stays for a while in regime 1 before going back in regime 2. But this is not optimal if one refers to the reasoning developed just above. The economy prefers to directly settle on the locus  $E = 0$  rather than to start in regime 1 before a switch, in finite time, to regime 2.

This completes the proof of Proposition 2.

### A.3. Proof of Proposition 3

At solution 1 (democratization), the index of inequalities is:

$$I^1(t) = \frac{(s\chi - \phi h(p^x R - s))}{\chi} (h\pi)^{\frac{1}{\gamma}} e^{\frac{(\rho-\chi)(t-T)}{\gamma}} + \phi\varphi(T) (h\pi)^{\frac{1}{\gamma}} e^{\frac{(\rho-\chi(1-\gamma))t}{\gamma}} e^{\frac{(\chi-\rho)T}{\gamma}} + \frac{\phi\gamma h}{\rho - \chi(1-\gamma)}$$

Take the derivative w.r.t time:

$$\dot{I}^1(t) = \frac{(h\pi)^{\frac{1}{\gamma}}}{\gamma} e^{\frac{(\rho-\chi)(t-T)}{\gamma}} [(\rho - \chi)(s\chi - \phi h(p^x R - s)) + \phi(\rho - \chi(1 - \gamma))\varphi(T)e^{\chi T}].$$

The sign of the derivative is given by the sign of the term, denoted  $\Psi$ , between squared brackets. Evaluating this coefficient at  $t = 0$ , we obtain

$$\Psi = (\rho - \chi)s + \phi(\rho - \chi(1 - \gamma))H_0 + \gamma\phi h \left[ p^x R - s - (h\pi)^{-\frac{1}{\gamma}} e^{\frac{(\rho-\chi)T}{\gamma}} \right],$$

which is positive according to (28). Thus  $\dot{I}^1(t) > 0$  for all  $t \in [0, T]$ .

At solution 2 (permanent dictatorship), the index is simply given by:

$$I^2(t) = \frac{s + \phi H^2(t)}{p^x R - s + \Omega H^2(t)},$$

with derivative

$$\dot{I}^2(t) = \frac{p^x \dot{H}}{(p^x R - s + \Omega H^2(t))^2} \left[ \phi R - \frac{1 - \alpha}{\alpha} \left( \frac{\alpha A}{p^x} \right)^{\frac{1}{1-\alpha}} s \right].$$

Thus,

$$R > \frac{(1 - \alpha)s}{\phi\alpha} \left( \frac{\alpha A}{p^x} \right)^{\frac{1}{1-\alpha}}$$

is sufficient to conclude that  $\dot{I}^2(t) < 0$  for all  $t \in [0, \infty)$ .

### A.4. Proof of Proposition 4

The proof of the first item of Proposition 4 works by contradiction and relies on a time consistency requirement for optimal trajectories.

Let's assume that  $V^1(H_T) < V^2(H_T)$ . Regarding the ordering between the value functions at the other boundary, two case are possible.

**Case 1.** Let  $V^1(0) < V^2(0)$ . If the curves  $V^1(H_0)$  and  $V^2(H_0)$  intersect then the number of intersections must be even. For instance, consider that there are two intersections  $\hat{H}$  and  $\tilde{H}$ , with  $0 < \tilde{H} < \hat{H} < H_T$ . By construction, we have  $V^1(H_0) >$

$V^2(H_0)$  for all  $H_0 \in (\tilde{H}, \hat{H})$ ,  $V^1(H_0) < V^2(H_0)$  for all  $H_0 \in [0, \tilde{H}) \cup (\hat{H}, H_T]$ ,  $V^1(\tilde{H}) = V^2(\tilde{H})$  and  $V^1(\hat{H}) = V^2(\hat{H})$ .

At  $H_0 = \tilde{H}$ , there exist two optima, i.e. the elite is indifferent between following path 1 (with positive education) or path 2 (no education). If the economy settles on path 1, then from what has been established in Appendix A.1, human capital increases. But by construction again, for any  $H$  varying in  $(\hat{H}, H_T]$ ,  $V^1(H) < V^2(H)$ : The elite would prefer to follow path 2 rather than path 1, which implies that the solution considered is not time consistent. This yields a contradiction. If the elite chooses path 2, then from Appendix A.2, human capital decreases monotonically. But  $V^1(H) > V^2(H)$  for all  $H \in (\tilde{H}, \hat{H})$ . There is a non degenerate period of time during which the elite would prefer in fact being in regime 1. Based on the time consistency requirement, we obtain a contradiction.

**Case 2.** Let  $V^1(0) > V^2(0)$ . This implies that the number of intersections between  $V^1(H_0)$  and  $V^2(H_0)$  (if any) is odd. Let's work with a unique intersection at  $\check{H}$  that is such that  $V^1(H_0) > V^2(H_0)$  for all  $H_0 \in [0, \check{H})$ ,  $V^1(H_0) < V^2(H_0)$  for all  $H_0 \in (\check{H}, H_T]$  and  $V^1(\check{H}) = V^2(\check{H})$ .

At  $H_0 = \check{H}$ , there is multiplicity of optima. Either, the elite may adopt the regime with positive education and human capital increases. Again by construction, we have  $V^1(H) < V^2(H)$  for all  $H \in (\check{H}, H_T]$ , which yields a contradiction. Or, it may choose not to invest in education, in which case  $H$  decreases and a contradiction arises too.

Proofs of the remaining items are left to the reader since they exactly follow the same line. In particular the reasoning of the second item is symmetric when one works with  $V^1(0) > V^2(0)$ . As for the third item, assuming that  $V^1(0) \leq V^2(0)$  and  $V^1(H_T) \geq V^2(H_T)$  (with one strict inequality), it's easy to show that there exists a unique intersection between the two value functions, for a critical initial stock of human capital  $\bar{H}$  such that  $V^1(H_0) \geq V^2(H_0) \Leftrightarrow H_0 \geq \bar{H}$ .

## A.5. Proof of Proposition 5

For the sake of exposition, let  $T(0)$  be denoted by  $T$ . Then,  $V^1(0) > V^2(0)$  if and only if:

$$e^{-\rho T} \left( \frac{\gamma(h\pi)^{-\frac{(1-\gamma)}{\gamma}}}{(1-\gamma)(\rho - \chi(1-\gamma))} \left[ e^{\frac{(\rho - \chi(1-\gamma))T}{\gamma}} - 1 \right] + \pi H_T \right) > \frac{(p^x R - s)^{1-\gamma}}{\rho(1-\gamma)};$$

which, by the definition of  $H_T$  in (20) and by (22), yields:

$$\frac{(p^x R - s)^{1-\gamma}}{\rho(1-\gamma)} < \frac{\gamma(h\pi)^{-\frac{(1-\gamma)}{\gamma}}}{\rho(1-\gamma)} e^{\frac{(\rho-\chi)(1-\gamma)T}{\gamma}} + \frac{h\pi(p^x R - s)}{\rho} e^{(\chi-\rho)T}. \quad (31)$$

Denote the RHS of (31) by  $J(T)$ . It follows that  $J(0) = \frac{h\pi}{\rho} \left( \frac{\gamma}{1-\gamma} (h\pi)^{-\frac{1}{\gamma}} + p^x R - s \right) > 0$ ,  $\lim_{T \rightarrow \infty} J(T) = \infty$  and:

$$J'(T) = \frac{h\pi(\rho - \chi)}{\rho} e^{(\chi-\rho)T} \left[ (h\pi)^{-\frac{1}{\gamma}} e^{\frac{(\rho-\chi)T}{\gamma}} - (p^x R - s) \right].$$

We observe that  $J'(T) \leq 0 \Leftrightarrow T \leq \tilde{T}$ , where  $\tilde{T}$  has been defined in (26). Thus, imposing  $J(\tilde{T}) > \frac{(p^x R - s)^{1-\gamma}}{\rho(1-\gamma)}$ , which is equivalent to:

$$p^x R - s > e^{\frac{\gamma}{1+\gamma}} (h\pi)^{-\frac{1}{\gamma}}, \quad (32)$$

is sufficient to conclude that  $V^1(0) > V^2(0)$ .

Moreover, whatever the regime, the value functions are strictly increasing in  $H_0$ , it's clear that a sufficient condition for having  $V^2(H_0) > V^1(H_0)$  for all  $H_0 \in [0, H_T]$  is  $V^2(0) \geq V^1(H_T)$ , which is equivalent to:

$$\frac{(p^x R - s)^{1-\gamma}}{\rho(1-\gamma)} \geq \pi H_T \Leftrightarrow \frac{(p^x R - s)^{1-\gamma}}{\rho(1-\gamma)} \geq \frac{\gamma(h\pi)^{-\frac{(1-\gamma)}{\gamma}}}{(\rho - \chi)(1-\gamma)} + \frac{\pi h(p^x R - s)}{\rho - \chi}. \quad (33)$$