On the emergence of public education in land-rich economies

ARTICLE *in* JOURNAL OF DEVELOPMENT ECONOMICS · FEBRUARY 2008

Impact Factor: 2.13 · DOI: 10.1016/j.jdeveco.2007.12.004 · Source: RePEc

Fernando Tohmé
Universidad Nacional del Sur

Daniel Heymann
University of Buenos Aires

Carlos Dabús
National Scientific and Technical Research C…

On the Emergence of Public Education in Land-Rich Economies

Sebastian Galiani
Washington University in St. Louis

Daniel Heymann
CEPAL

Carlos Dabús and Fernando Tohmé *
Universidad Nacional del Sur - CONICET

December 3, 2007

Abstract

We analyze the emergence of large-scale education systems by modeling the incentives that the economic elite could have (collectively) to accept taxation destined to finance the education of credit-constrained workers. Contrary to previous work, in our model this incentive does not arise from a complementarity between physical and human capital in manufacturing. Instead, we emphasize the demand for human-capital-intensive services by high-income groups. Our model seems capable to account for salient features of the development of Latin America in the 19th century, where, in particular, land-rich countries such as Argentina established an extensive public education system and developed a sophisticated service sector before starting significant manufacturing activities.

*We are grateful to an anonymous referee and to P. Somaini for many helpful comments and suggestions. We also thank the comments of M. Garibotti, P. Gerchunoff, J. Fanelli, F. Weinschelbaum and seminar participants at DEGIT conference of Growth and Development, Universidad de las Americas and Universidad de San Andres. Corresponding Author: Sebastián Galiani, Washington University in St. Louis, MO, US. E-mail: galiani@wustl.edu. A previous version of this paper circulated under the title: "Land-Rich Economies, Education and Economic Development".

1 Introduction

In a world with imperfect capital markets, low-income workers are constrained in their private investment in education. Thus, the nature (and, more starkly, the presence or absence) of a public school system critically determines the extent and the evolution of human capital accumulation. Historically, different societies have performed quite differently in the establishment and the growth of schooling. The United States and Canada had developed a substantial school system already in colonial times; by 1850, every northern state of the US had enacted a law strongly encouraging or requiring localities to establish "free schools", open to all children and supported by general taxes. The rest of the hemisphere trailed far behind those two countries in education and literacy. The most progressive Latin American countries, such as Argentina and Uruguay, lagged more than fifty years behind the U.S. and Canada in providing massive primary schooling and attaining high levels of literacy. Most of Latin America was unable to achieve these standards until well into the twentieth century, if then (Mariscal and Sokoloff, 2000).

Social decisions on education are influenced by broad political factors, but they also respond to economic considerations. Reciprocally, changes in the levels of schooling and literacy affect the social structure and may contribute to alter the political institutions that determine the educational institutions themselves (see, among others, Glaeser et al. (2004), which present evidence supporting the view that higher education leads to more democratic political systems). The question naturallly arises as to why did some countries invest heavily in the education of broad segments of the population while others lagged behind.

Galor and Moav (2006) -GM from now on- provide an interesting explanation: capitalists, as a group, may have incentives to invest in the education of the labor force because the productivity of physical capital in manufacturing production increases with the input of human skills. That is, capitalists can gain from tax-financing the emergence of a public education system in order to raise the return on their assets by increasing the supply of a complementary factor, human capital. This argument seems relevant to North America (and to Western Europe), but it would have difficulties in explaining the Latin America experience.

Galor et al. (2005) extendss the analysis in GM by assuming that, although human skills contribute to increase the productivity of industrial capital, they provide no benefits for landlords as such. Then, if landlords have veto power over policies, they would block or delay the growth of public education (see also Bourguignon and Verdier (2000) for a complementary argument). This hypothesis can account for the delay of most Latin American countries, but it still does not rationalize the intermediate cases of the Southern Cone countries and Costa Rica, which started to develop important schooling systems in the second part of the 19th century, with a polity strongly influenced by landholders.

In this paper we present a model to study the emergence of public education in land-abundant economies, where the interests of landlords essentially dictate policies, and which need not engage in the production of manufactured goods, since their demand for these may be wholly satisfied by imports. Worker skills are assumed to make no contribution to production in this agricultural sector. Thus, the existence of more educated workers would generate no advantage for the elite in terms of the returns on their productive assets, as it does in the argument of GM. However, we assume that the skill-intensity of output and consumption baskets increases with income levels, specifically because the production of some services requires the input of educated workers.

The argument is based on three central elements: First, individual preferences over consumption goods imply changes in the composition of individual spending as income grows, embodied in Engel curves. Second, the production of sophisticated services (which are non-tradeable goods) is intensive in human capital. Third, investment in education by individual households is constrained by lack of access to credit (see, for example, Banerjee and Newman, 1991; Galor and Zeira, 1993, and Benabou, 1996). We also assume that the quantity and quality of labor are imperfect substitutes. This implies that the number of high-income agents may have strong effects on how many individuals are subsidized to accumulate human capital.

Thus, the size of the elite, as the group who demands goods particularly intensive in human capital, may have strong effects on the number of educated workers. This would rationalize a link between historical conditions, especially with regard to the distribution of land, and social choices regarding the scope and the financing of the education system. Widespread education would emerge earlier in agricultural-based economies when land is highly productive and its property sufficiently distributed as to create a demand for a sizeable number of educated workers. The proposition corresponds with the case made by Engermann and Sokoloff (2000), who indicated that the greater degree of inequality in Latin America, as compared to North America, played an important role in explaining the different behaviors regarding the establishment of educational institutions (see also Mariscal and Sokoloff, 2000 and Easterly, 2006).

The experience of Argentina provides an illustration of the argument. In the second half of the 19th century, Argentina became increasingly integrated in the world economy as a large producer and exporter of agricultural goods, and an importer of manufactures. At the same time, the composition of rural production changed significantly, as agriculture expanded over cattle rising activities, a shift that favored less extensive forms of production (see, among others, Adelman, 1994). While the distribution of land and incomes was still more unequal than in North America, where wheat production was mainly based on family farms, it was less concentrated than in other Latin American economies. The expansion of agricultural activities allowed a very substantial growth of the urban population, especially in the city of Buenos Aires. Apart from its administrative functions as the capital of the country, the city developed an increasingly large and sophisticated service sector. At the same time, the country experienced what is widely considered one of the key processes in its history: the emergence of the system of public education, associated with the emblematic figure of Domingo Faustino Sarmiento.

However, the progress of education was not immediate, but rather went along with the development of the economy and, actually with a growing demand for skills (see Martínez Paz, 2003). It was only in 1875 that the Province of Buenos Aires passed a comprehensive law on public instruction, while the corresponding national instrument was introduced in 1884. The large regional differences in rates of scholarization and literacy indicate that, directly or indirectly, spending on education depended very much on the economic configuration of the localities. In 1895, in the city of Buenos Aires, almost 60% of the children of ages 6-14 attended school, doubling the national average; while the rate of illiteracy was only 20% in the city, against 57% in the country as a whole (and nearly 80% in poor jurisdictions far from the central agricultural regions, like La Rioja or Neuquén). Furthermore, the type of education provided by the Argentine government seemed to correspond more to the economic incentives perceived at the time than to the vision of the founders of the system. For example, Sarmiento believed that the educational system had to form individuals to work in the agricultural and industrial sectors but, instead, it tended to qualify individuals for work in services (see Tedesco, 1993). These features are rationalized through the simple analytical construction that we develop in the paper.

The rest of the paper is organized as follows. The next section describes the setup of the model. In section 3 we analyze the emergence of public education in land-rich economies which have not gone through a previous stage of industrialization. In section 4 we present the comparative dynamics of the model. Conclusions are then presented in section 5.

2 Setup of the model

The economy can potentially produce, and consume, two commodities: an agricultural tradable good and services. The scenarios that we consider would correspond to countries which have agricultural comparative advantages, and import manufactured goods. However, in this version of the model, we leave the consumption, as well as the potential production, of industrial goods out of the analysis. We also disregard capital flows[1].

The agricultural good can be produced with a technology which uses only land and capital (this is homogeneous with the agricultural output). The landlords in this sector form an elite with the power to determine the level and the allocation of government spending given the taxation technology. In this economy, members of the elite group are the exclusive owners of land and capital. For the sake of the argument, we consider that there is another activity that produces the agricultural good using only labor, and which effectively operates

[1] Potential extensions that include an industrial sector are discussed in Galiani et al. (2005).

as a reservoir of unskilled workers. Production in the land-intensive agriculture sector is only constrained by the available stocks of land and capital. In the labor-only sector, skills are supposed to add to the productivity of workers, according to a pre-determined function (which we will later represent as linear, so that earnings will be proportional to the level of skills). This assumption establishes the reservation wage of workers of different skills, whose alternative occupation is in the service sector[2].

The accumulation of agricultural capital is the driving impulse of change in the system. At low levels of income, only agricultural goods are consumed. There is no demand for services. The service sector, where production depends on the level of skills of the workers employed, is therefore inactive. However, at some point, the rising wealth of landlords may induce structural change in the economy due to a diversification of consumption into services.The emergence of activities that use skilled labor and that produce goods demanded by the elite can generate an incentive for that group to support and to finance the education of a segment of workers.

We consider an overlapping generations economy, with a setup similar to that in GM. Agents live for two periods, and there is no population growth (i.e., individuals have a single offspring). At the beginning of the dynamics there are two kinds of dynasties, landowners and workers, who differ only in their factor endowments. Individuals in the first group own land, which is not traded in equilibrium, and some physical capital; they do not supply labor, either skilled or unskilled. Workers are endowed with a basic set of labor capacities, which can be increased by acquiring education. However, young agents of worker origin have no access to credit to self-cover the costs of acquiring skills against the future increase in earnings.

In every period, the young agent of each dynasty receives a non-negative (but not necessarily strictly positive) bequest from the old agent of his lineage. Those bequests are potentially taxable and, in the model, such taxes will fund the spending on public education when and if this emerges. The land owned by an individual landlord is automatically transferred to his offspring, but this transfer is not included in the definition of bequests. Young agents use their after-tax bequests to accumulate assets. In the case that we analyze, only landlords save, and the resources are used to acquire physical capital for agricultural production. Old agents carry out productive activities, they consume, and decide whether and how much to transfer to their offspring as bequests.Young agents do not consume (or, equivalently, their consumption is included in that of their parent). We will assume that the level of income of all workers is such that they do not leave bequests, and that their consumption consists only of agricultural goods.

[2]The identification with an agricultural activity of the sector that acts as a source of unskilled labor is of course a simplification corresponding to this two-good economy setup. It is meant as a very basic way to specify the opportunity costs of the workers who will receive education. In the equilibrium generated by the model, all skilled workers will be employed in the production of services.

2.1 Technology and Production

Agricultural goods are denoted A and services N. The factors of production are: labor, L, land, T (in fixed supply) and physical capital, X (which is homogeneous with the agricultural good). The bequests transferred by the previous generation provide young agents with resources which are applied to capital formation or, if applicable, to pay taxes that finance spending in education.

2.1.1 The Agricultural Sector

The production of the market-oriented agricultural sector is described by a constant-return function $F(\mathbf{T}, \mathbf{X})$, where $\mathbf{T}$ and $\mathbf{X}$ are, respectively, the total surface of land and the aggregate capital used in that sector. W suppose, for simplicity, that the distribution of land and capital is uniform. If the number of production units (and landlords) in the economy is m, the output of a farm is then given by:

$$y^A \ = \ \frac{1}{m}F(\mathbf{T}, \mathbf{X}) \ = \ F(\frac{\mathbf{T}}{m}, \frac{\mathbf{X}}{m}) \ = \ f(X^A)$$

where $X^A = \frac{\mathbf{X}}{m}$ is the capital stock of the average production unit, which has been carried over from the previous period. Since the total endowment of land is fixed, we treat $\frac{\mathbf{T}}{m}$ as a constant. Capital fully depreciates in the period. The per-landlord production function $f(\cdot)$ has standard properties.

The good A can alternatively be produced with a technology using only labor, at constant returns. The productivity of an unskilled worker in this 'subsistence' activity is denoted by $\widetilde{w}$; educated workers with a skill level indicated by h generate a volume of output given by w^h; in what follows we will specify $w^h = \widetilde{w}h$. We normalize to unity the productive capacity (labor units) of an unskilled individual, i.e. $h = 1$. Given the employment opportunities in this sector, the market wages of workers with skills h in alternative activities will be bounded by: $w(h) \geq w^h$.

2.1.2 Education

Education increases the productivity of workers by generating skills that raise their working capacity. The inputs of the education activity are assumed to consist solely of good A. The skills of an individual in period $t + 1$ are a function of the resources spent on the agent's education (e) in period t:

$$h_{t+1} = h(e_t)$$

with $h(0) = 1$, $h'(0) = \infty$, $h'(\infty) = 0$, $h' > 0$ and $h'' < 0$. Equivalently, the costs of providing a worker with a level of skills h_{t+1}would be given by a convex function: $e_t = e(h_{t+1})$.

2.1.3 The Service Sector

The second consumable good, N, is interpreted as an urban, relatively sophisticated service, the production of which requires only the use of labor[3]. An individual who demands services hires workers competitively, and consumes the output they produce, which depends on the number of persons who participate in the supply of the services, as well as on their skills:

$$y^N(n, \overline{h})) = \psi(n)\overline{h}$$

where y^N is the volume of services per individual consumer, produced by n workers who have an average level of skills $\overline{h}$. We assume that both n and the level of skills h are continuous variables; then, $n = \int n(h)dh$, where $n(h)$ is the number of workers of skill h employed in the production of services. The mean level of skills used by an individual who demands services is then given by $\overline{h} = (\int n(h)hdh)/n$. If the wage associated with a level of skills h is $w(h)$, the value of services consumed by an individual is given by $\int w(h)n(h)dh$.

An important assumption is that the quantity of services consumed grows without bounds with the level of skills of the workers who participate in their production, but that it only increases up to a point with the number of such workers. The intuition behind this specification is simply that, for a wide range of services (from medical attention to entertainment, say), when a certain level of the quantity of suppliers is reached, additional workers make little difference for the utility of the resulting consumption, but this is enhanced with an increase in the skills applied to production. The function ψ would then be strictly concave. Although this condition is sufficient to define the qualitative features of our argument, we make the stronger assumption that the marginal contribution of additional workers goes to zero at a finite number $\widetilde{n}$ of service suppliers. Thus, ψ is increasing in n up to a maximum $\widetilde{n}$. For simplicity, we consider the following function:

$$\psi(n) \;=\; \min[n, \widetilde{n}]$$

2.2 Preferences and consumption demands

Individuals, within as well as across generations, are identical in their preferences and innate abilities. Preferences are defined over consumption-bequest bundles, $c = \langle c_A, c_N, b^o \rangle$, $c \in \mathcal{R}^3_+$, where c_j is the quantity consumed of good j ($j = A$ or N), while b^o is the bequest, measured in terms of agricultural goods, left to

[3] The assumption here is that, having left aside the potential demand for human capital in an industrial sector, the production of services would be the only source of a demand for skills. This is clearly a simplification, which disregards other important activities which require the input of skilled workers, such as the public administration and education itself. However, those demands for skills would be likely to appear when the economy has reached a stage of sophistication and consumption diversification as to induce a significant "final demand" for educated labor such as the one emphasized in the model.

an offspring (we reserve the notation b for the bequest received by the agent when young). Individual welfare is assumed to vary directly with the amount of resources left as bequest, independently of their use by the offspring. This implies, in particular, that savings depend only on the income of the adults who are deciding on bequests, and not on the expected return on assets for the recipients of the next generation[4].

The preferences defined below capture the intuition that consumers prefer more diversified consumption bundles as their income increases, but must surpass certain threshold levels of satisfaction of "basic" wants before including new goods in their purchases. We partition the consumption space into three subsets (which may be thought of as "stages"), according to which income thresholds have been reached. Within each stage, preferences are described by a (stage-specific) Stone-Geary function. Preferences and the associated demand curves have implicit an ordering of the goods which are part of the consumption basket at different levels of income. There are three stages: i) consumption of agricultural goods (A), only; ii) consumption of A and positive bequests; iii) consumption of A and services (N) as well as positive bequests[5].

The preferences are concretely specified as follows:

$$u(c) = \begin{cases} c_A & \text{if } c \in \mathcal{C}^1 \\ (\min(c_A, \widetilde{c}_{A2}) - \widetilde{c}_{A1})^{1-\beta}(\min(b^o, \widetilde{b}_2^o))^{\beta} + k_1 & \text{if } c \in \mathcal{C}^2 \\ \left[(c_A - \widetilde{c}_{A2})^{1-\alpha} c_N^{\alpha}\right]^{1-\beta} (b^o - \widetilde{b}_2^o)^{\beta} + k_2 & \text{if } c \in \mathcal{C}^3 \end{cases}$$

$$\begin{aligned} \mathcal{C}^1 &= \left\{c \in R_+^3 : c_A \leq \widetilde{c}_{A1}\right\} \\ \mathcal{C}^2 &= \left\{c \in R_+^3 : \left(\widetilde{c}_{A1} < c_A \leq \widetilde{c}_{A2} \text{ or } b^o \leq \widetilde{b}_2^o\right)\right\} \\ \mathcal{C}^3 &= \{c \in R_+^3 : \widetilde{c}_{A2} < c_A \text{ and } \widetilde{b}_2^o < b^o\} \end{aligned}$$

$$\begin{aligned} k_1 &= u(\widetilde{c}_{A1}, 0, 0);\ k_2 = u(\widetilde{c}_{A2}, 0, \widetilde{b}_2^o); \\ \widetilde{b}_2^o &= \frac{\beta}{1-\beta}(\widetilde{c}_{A2} - \widetilde{c}_{A1}) \end{aligned}$$

Here, c_A, b^o denote, respectively, the consumption of agricultural goods of the old agent, and the bequest left to the offspring; c_N is the volume of services

[4]This form of bequest motive (i.e., the "joy of giving") is common in the recent literature on income distribution and growth. The assumption that the rate of return is irrelevant in the decision to leave bequests does not affect the qualitative results emphasized in the paper, although it may have strong implications in other contexts.

[5]In a setup including manufacturing goods, the preferences of the agents would be specified so as to include an intermediate step between cases ii) and iii), where the agent demands manufactured goods together with agricultural products and positive bequests, before entering the stage where skill-intensive services are consumed (see Galiani et al. (2005)).

consumed, i.e. $c_N = \overline{h}\min(n,\widetilde{n})$, and therefore depends on the number of workers employed by the individual (n), and on their average level of skills ($\overline{h}$).

This utility function generates consumption-bequest choices with the following features (see Lemmas 1, 2, 3 and 4 in the Appendix):

- If the resources of the old individual, denoted by i in terms of the agricultural good, are smaller than a first threshold: $i < \widetilde{i}_1 = \widetilde{c}_{A1}$, all those resources are allocated to the consumption of good A. The agent neither leaves bequests nor consumes services.

- When income is in an intermediate region: $\widetilde{i}_1 < i \leq \widetilde{i}_2 = \widetilde{c}_{A2} + \widetilde{b}^o_2$, the agent consumes A and leaves positive bequests, following a rule with constant marginal propensities to consume:

  $$c_A = \widetilde{c}_{A1} + (1-\beta)(i - \widetilde{i}_1)$$

  $$b^o = \beta(i - \widetilde{i}_1)$$

  The share of agricultural consumption in spending falls with income, as implied in traditional Engel curves.

- When $i > \widetilde{i}_2$, the individual allocates spending to services. Let $w(h)$ be the market wages of the workers of skill h, assuming that such workers exist in the economy. Given the assumptions on the opportunity costs of educated workers, $w(h)$ must be such that $w(h) \geq \widetilde{w}h$, where $\widetilde{w}$ is the productivity of unskilled workers ($h = 1$) in the labor-only sector. Suppose that there is a set **S** of skill levels for which the market wages satisfy $w(h) = \widetilde{w}h$, and that h^S is the highest level of skills belonging to that set. Then, if $\alpha(1-\beta)(i - \widetilde{i}_2) \leq \widetilde{n}\ \widetilde{w}h^S$, the agent will only demand labor belonging to the set S. The average level of skills of the workers contributing to the production of services would then be given by: $\alpha(1-\beta)(i - \widetilde{i}_2) = (n\overline{h})\widetilde{w}$, and the agent would be indifferent to the composition of the workforce providing services as long as that condition applies. The intuition behind this criterion is that the set **S** defines the levels of skills available in the market with minimal wages per unit labor capacity. As long as the consumer has not reached the upper bound on the number of service suppliers (so that he employs $n < \widetilde{n}$ workers), it is advantageous to hire only workers of this type, since that minimizes the cost of purchasing a given total amount of working capacity. The decision rule can be simply implemented by an individual agent if workers with skills h^S are hired in number $n = \alpha(1-\beta)(i - \widetilde{i}_2)/\widetilde{w}h^S$. Thus, in this region, as income increases, individual consumption effectively varies by raising the number of workers within the same skill group.

- With high enough incomes ($\alpha(1-\beta)(i - \widetilde{i}_2) > \widetilde{n}\widetilde{w}h^S$), the individual agent would potentially contemplate hiring workers with skills higher than

h^S at some wage levels above the reservation earnings ($w(h) > \widetilde{w}h$). If such workers were unavailable in the market, then the agent would simply employ the saturation number of workers of skill h^S, and allocate optimally the remaining income between consumption of A and bequests:

$$c_A = \widetilde{c}_{A2} + \frac{(1-\alpha)(1-\beta)}{\beta + (1-\alpha)(1-\beta)}(i - \widetilde{i}_2 - \widetilde{n}\widetilde{w}h^S)$$

$$b^o = \widetilde{b}^o_2 + \frac{\beta}{\beta + (1-\alpha)(1-\beta)}(i - \widetilde{i}_2 - \widetilde{n}\widetilde{w}h^S)$$

This case will be important in the political economy equilibrium discussed in the next section. We note here that the consumer has no excess demand for labor of type h^S, since he already employs the saturation number of those workers (i.e., the demand curve is truncated at $n = \widetilde{n}$).

- In this region of high incomes, if there is a supply of workers with $h > h^S$, such that $w(h) > \widetilde{w}h$, the individual agent will never demand workers with skills levels such that for these values the wage curve $w(h)$ is locally concave. The consumer will instead pick a combination of skills that minimizes the mean wage for a given average level of skills. If, on the other hand, the wage curve is locally strictly convex, just one level of skills will be chosen for consumption. In case the supply of labor with high skills ($h > h^S$) is described by a continuous, convex wage function, the agent will employ the saturation level of $\widetilde{n}$ workers, and their common level of skills h will satisfy: $\alpha(1-\beta)(i-\widetilde{i}_2) = \widetilde{n}w(h)(\epsilon_{w(h)}(1-\alpha(1-\beta))+\alpha(1-\beta))$, where $\epsilon_{w(h)}$ is the elasticity of the wage curve at h. The expression that multiplies $\widetilde{n}w(h)$ is larger than unity if and only if $\epsilon_{w(h)} > 1$, which is in fact the case for $h > h^S$ such that $w(\cdot)$ is convex at h. Therefore, the value of spending in services falls short of the level $\alpha(1-\beta)(i - \widetilde{i}_2)$. This is because once the critical number of workers $\widetilde{n}$ is reached, any increases in income will require an upgrade in skills, which in turn amounts to a more than proportional increase in wages (due to the convexity of $w(h)$).

The individual demands just described are predicated on a market wage function, which in turn depend on the supply of workers with various degrees of qualification. In the following section we discuss how a government that represents the group of (homogeneous) landlords will choose the skill level of the workers.

3 The emergence of public education

In our setup, educated workers will only be demanded when the income of the representative landlord is high enough as to induce the consumption of services. Thus, before studying the emergence of public education, we consider the transition from an agricultural economy to a society that diversifies its consumption.

3.1 Diversified consumption

In the initial state, we assume, the amount of available resources is such that only good A is produced and consumed. However, the economy is not in a "growth trap", so that the income of landlords is above the threshold that induces positive bequests ($i > \widetilde{i}_1$). Here, unskilled workers are employed in the subsistence sector, consuming $\widetilde{w}$. Landlords invest in physical capital all the bequest they receive, and choose their consumption/bequest bundle in the second period. The demand functions discussed in the previous section induce the following dynamics:

$$X_t = \beta(f(X_{t-1}) - \widetilde{i}_1).$$

Suppose that the agricultural sector is sufficiently productive and landlords sufficiently frugal so that at the capital stock X_0 available to the first generation, $\beta(f(X_0) - \widetilde{i}_1) > X_0$, implying that capital accumulation proceeds from that point. Then, the difference equation admits two steady states, with the lower one unstable and the higher, $\widehat{X}_1^h$, locally stable. If $f(\widehat{X}_1^h) > \widehat{i}_2 = f(\widetilde{X}_2)$, the economy will reach the stage where landlords start consuming services and continue to accumulate capital until reaching a steady state in that region. Assuming that this condition holds (see Proposition 1 in the Appendix), we can now study the emergence of public education.

3.2 Public education decided and financed by the elite

The consumption of services generates the potential of a demand for skills. The quantity of services consumed by an individual can vary by changing two variables: the number of workers hired, which has been assumed not to exceed the level $\widetilde{n}$, and the level of skills of the workers that provide the services. Since in the relevant range of incomes only landlords (a population of m agents) will consume the output of the service sector, employment in this activity will not exceed $\widetilde{n}m$. We assume that the population of unskilled workers in the "subsistence" agricultural sector is larger than $\widetilde{n}m$. Consequently, non-educated workers (with an ability indicated by $h = 1$) receive the reservation wage $\widetilde{w}$. Skilled workers, with $h > 1$, will earn at least their income in the subsistence sector, $h\widetilde{w}$. Young unskilled workers would benefit from receiving education. However, these individuals are credit-constrained, and individual landlords do not have incentives to finance the education of young workers who, by assumption, can freely choose their employment and cannot commit to the repayment of "education loans". As a consequence, some kind of collective action might improve the welfare of the elite. Thus, young landlords who anticipate their future demand for services, acting as a group, might accept to finance the education of some workers by way of a tax on the bequests left to them.

Assume that the young landlords are represented by a government, capable of enforcing taxes and operating a public education system. The government runs a balanced budget, and spends exclusively on education. Taxes are raised as lump-sum transfers from the bequests received by the young landlords. Therefore, these agents bear the cost of the education system.We also assume that the authority can limit the size of the set of individuals who receive education (in practice, this may be done by varying the geographical coverage of the education system, or by determining conditions of schooling such that some groups have preferential access).

The planner optimizes the welfare of the representative landlord, subject to a resource constraint that incorporates the tradeoff between capital accumulation and spending in education, and internalizes the effects of its decisions on the labor costs faced by the consumer of services. Then, if b is the bequest inherited by the young landlord, n is the total number of workers who will receive skills, E the aggregate spending in education per-landlord (or, equivalently, the level of taxes), $\overline{w}$ the average wage of workers in the service sector in the market equilibrium that will result given the planner's choices, the problem is:

$$\max \left[(c_A - \widetilde{c}_{A2})^{1-\alpha} c_N^{\alpha}\right]^{(1-\beta)} (b^o - \widetilde{b}_2^o)^{\beta} + k_2$$

subject to:

$$c_N = \overline{h}\min(n, \widetilde{n})$$

$$f\left[b - E\right] - n\overline{w} = c_A + b^o$$

If $n(h)$ workers receive education that gives them skills h, and $w(h)$ is their market wage, then $n = \int n(h)dh$, $\overline{h} = (\int n(h)hdh)/n$, and $\overline{w} = (\int n(h)w(h)dh)/n$.

The government that represents the elite clearly has incentives to avoid providing rents to educated workers. Consequently, the planner's optimum would imply that $w(h) = \widetilde{w}h$ for all the skill levels h present in the market. If the demand for services is high enough, so that it is not optimal to employ only unskilled workers, the government will tax its constituents in order to start a public education system.Since the cost of education is given by a convex function, if the planner wishes to endow workers with a certain average level of skills ($\overline{h}$), it would want to do it at minimal cost, that is, by providing a uniform level of education $e(\overline{h})$. Then, the total spending in education would be given by $E = n\overline{h}$.

When the demand for services arises, it will first be satisfied by increasing the number of non-educated workers who are drawn from the "subsistence" sector and employed in the service sector. This goes on until each landlord is hiring $\widetilde{n}$ such workers. Once income is sufficient to motivate the landlord group to favor spending in the generation of skills, the education system will cover the "saturation level" given by $\widetilde{n}m$.Thus, education would start with a widespread, but not universal, coverage at an elementary level. For higher incomes of the elite, the growth of education would raise the skill level of the educated population, but without increasing the number of skilled workers. This establishes a difference with GM, where the elite would choose to educate all the available labor force.

Here, the economy retains *dual* features, since, absent a productive sector that would absorb the supply of unskilled workers in the labor-only activity, the individuals who stay in this reservoir would earn incomes fixed at $\widetilde{w}$ even as capital accumulation raises the revenues of landlords, and their increasing demand for services gets reflected in higher levels of skills and earnings for the workers who get access to the education system.

Another feature of the political economy equilibrium is that the planner equates the marginal rate of substitution between services and the basket of consumption of goods-bequests with the full marginal cost of labor capacity, which includes, in addition to the wage $\widetilde{w}$, the marginal cost of producing skills in terms of foregone output of good A. At the the level of skills h^* that the planner chooses to supply, in the market equilibrium, landlords will demand $\widetilde{n}$ workers with those skills, while spending less than the amount $\alpha(1-\beta)(i-\widetilde{i}_2)$ at a wage $\widetilde{w}h$.

Since the number of workers employed has reached saturation, there is no excess of demand for labor with those skills, and the wage is a market equilibrium. At that point, when landlords of this young generation reach old age, the income generated by their after-tax bequests could support the demand of service workers with skills $h > h^*$ (and wages $\widetilde{w}h > \widetilde{w}h^*$). However, the authority chooses not to offer that higher- intensity education, since the costs in terms of foregone market income for the landlord group would exceed the benefits in terms of higher quality of services consumed.

Summarizing, we have that the goverment of the elite chooses taxes and spending in education in the following way (see Proposition 2 in the Appendix):

- If $\alpha(1-\beta)(i-\widetilde{i}_2) \leq \widetilde{n}\widetilde{w}$, then spending in education is zero, and no worker receives skills $n(h) = 0$ for all $h > 1$.

- If $\alpha(1-\beta)(i-\widetilde{i}_2) \geq \widetilde{n}\widetilde{w}$, then $\widetilde{n}$ workers per landlord will receive education. All these individuals will have a level of skills h^* satisfying the condition:

  $$MRS_{N,Ab^0} = \frac{\alpha(1-\beta)}{1-\alpha(1-\beta)} \frac{f(b-\widetilde{n}e(h^*)) - \widetilde{i}_2 - \widetilde{n}\widetilde{w}h^*}{\widetilde{n}h^*} = \frac{(f'(b-\widetilde{n}e(h^*))}{h'(e(h^*))} + \widetilde{w} = MC_{N,i}$$

  where f' and h' are, respectively, the marginal productivities of capital in agriculture and of education in the production of skills. The left hand side of the equation measures the marginal rate of substitution between services (proportional to labor skills) and the bundle formed by good A and bequests. The right hand side indicates the full marginal cost of labor units, given by the wage and by the output foregone from raising education expenditures, and taxes, to increase average skills at the margin.

Alternatively, the condition for an optimum can be written:

$$\alpha(1-\beta)(f(b-\widetilde{n}e(h^*))-\widetilde{i}_2) \;=\; \widetilde{n}\widetilde{w}h^*(\alpha(1-\beta)+(1-\alpha(1-\beta))(1+\frac{f'(b-\widetilde{n}e(h^*))}{\widetilde{w}h'(e(h^*))})$$

so that $\widetilde{n}\widetilde{w}h^* < \alpha(1-\beta)(f(b-\widetilde{n}e(h^*))-\widetilde{i}_2)$.

- The chosen level of education and skills rises with the value of inherited bequests:
$$\frac{\partial e}{\partial b} > 0.$$

If the elite is subdivided, in the sense that the education system is destined to satisfy the demands of groups with different incomes and demands for services, there can be a distribution of people who receive different levels of education; the size of the members of each level would be a function of the size of the set of landlords that demands services requiring those skills. In any case, if all landlords require the services of $\widetilde{n}$ workers, the size of the group of educated workers would be proportionate to the size of that elite. We explore this issue further in the next section.

4 Comparative Dynamics

The distribution of land ownership influences the timing and the scope of the introduction of education. Consider two economies that start with equal aggregate levels of resources (land, capital and population), and have the same productivity in the labor-only sector, but differ in the dispersion of the ownership of physical assets. The economy with more concentrated holdings will have initially, and will maintain over time, higher levels of capital (and land) per landlord. This implies that, if both economies eventually enter the stage with diversified consumption and a demand for educated labor on the part of the elite, the emergence of a group of skilled workers would take place sooner where individual landlords are wealthier. However, since the elite is small, education would reach a comparatively small number of individuals, with high skills, proportionate to the (large) incomes of the landlord group. Conversely, an economy with a more diluted distribution of land would generate an educational system at a later date, but this would provide instrucyio to larger segments of the population.

This argument can be stated as follows:

Theorem 1 *Let $\epsilon_1(\mathbf{T}, \mathbf{X}_0)$ and $\epsilon_2(\mathbf{T}, \mathbf{X}_0)$ be two economies with the same aggregate endowments of land and initial capital stock, and the same production function. The number of landlords is, respectively, m_1 and m_2, with $m_1 < m_2$. Suppose that both economies will eventually reach the stage C^3 of diversified consumption of the landlord group. Then:*

- *If both economies enter C^3 at dates t_1 and t_2, respectively, $t_1 < t_2$.*
- *When education emerges in economy $\in_2$, the number of educated workers will be larger than in $\in_1$, while their level of skills will be lower.*

Proof: *Given the assumption that both $\in_1$and $\in_2$ enter C^3 at finite dates $t_1, t_2 > 0$, at the initial date $t = 0$ both must be at stage C^2, where the representative landlord in each economy consumes good A and leaves bequests to his offspring. While economy j remains in this stage, the average landlord has an income:*

$$i^j_{t+1} = f^j(X^j_t) = f(X^j_t, \frac{T}{m_j})$$

where T is the aggregate stock of land (equal in both economies), m_j the number of landlords in $\in_j$, and X^j_t is the capital stock owned by the young representative landlord at time t, resulting from the bequest left by the previous generation:

$$X^j_t = \beta(i^j_t - \widetilde{i}_1) = \beta(f^j(X^j_{t-1}) - \widetilde{i}_1)$$

Thus, $i^j_{t+1}(X^j_{t+1})$ is increasing in i^j_t (X^j_t). Since $m_1 < m_2$, $f^1(X) > f^2(X)$ and the initial capital stock per landlord, $X^1_0 > X^2_0$. Therefore, $i^1_0 > i^2_0$. Then, it must be that $i^1_t > i^2_t$ for all $t < \min(t_1, t_2)$. Since the timing of the transitions to stage C^3 are given by the conditions: $i^j_{t_j} = \widetilde{i}_2$, it follows that $t_1 < t_2$, as claimed. Moreover, it has been shown that once the authority representing the group of landlords starts an education system, this will reach $\widetilde{n}m_j$ young workers in economy $\in_j$. Therefore, when education emerges in $\in_2$, it will have a wider coverage than in $\in_1$. Since spending in education per worker is an increasing function of the pre-tax income of the average landlord, the level of skills will be higher in economy $\in_1$ $\square$

5 Conclusion

Incentives and the institutions that mold them clearly matter for development (see, among others, North, 1981; Landes, 1998; and Acemoglu et al., 2005). However, institutions themselves can vary according to political and economic configurations, that is, they are endogenously determined. Human capital accumulation is a noticeable example of this interaction between institutional and economic factors.

In this paper we study the emergence of public education in land-abundant economies that, for standard reasons of comparative advantage do not engage in the production of manufactured goods, and where the interests of landlords dictate policies.We abstract here from arguments emphasizing the role of social conflict in inducing the extension of education and promoting political change. Instead, we emphasize that, in certain circumstances, the economic interests of elites themselves may operate as a source of change.

In that regard, the argument is analogous with that in Galor and Moav (2006). However, in that paper, capitalists, as a group, may have incentives to invest in the education of the labor force because the productivity of physical capital in manufacturing production increases with the input of human skills. By constrast, in the economies we are considering, that channel would not operate. The mechanism that we study derives from the demand for human-capital-intensive services by high-income groups. Analytically, a demand for education, voluntarily financed by a landholding elite, can be generated in a setup with heterogeneous goods, where consumption preferences are non-homothetic and the demand for skill-intensive commodities emerges at comparatively high levels of income.

Assuming as we do that the quantity and quality of labor are not perfect substitutes, the number of high-income agents may have strong effects on how many individuals are subsidized to accumulate human capital. Hence, the diffusion of education would depend on the size of the elite and, indirectly, on the degree of concentration of land ownership.

Our model seems capable to account for salient features of the development of Latin America in the 19th century, where, in particular, land-rich countries such as Argentina established an extensive public education system and developed a sophisticated service sector before starting significant manufacturing activities and with policies strongly influenced by a landholding elite.

6 References

Acemoglu, D.; S. Johnson and J. Robinson (2005): "Institutions as the Fundamental Cause of Long-Run Growth", forthcoming in Aghion, P. and Durlauf, S. (eds.): **Handbook of economic growth**, NorthHolland, Amsterdam.

Adelman, J. (1994): **Frontier Development: land, labor, and capital on the wheatlands of Argentina and Canada,** Oxford University Press, New York.

Banerjee, A. and A. Newman (1991): "Risk-Bearing and the Theory of Income Distribution", *Review of Economic Studies* 58: 211-235.

Benabou, R. (1996): "Equity and Efficiency in Human Capital Investment: The Local Connection", *Review of Economic Studies* 63: 237-264.

Bourguignon, F. and T. Verdier (2000): "Is Financial Openness Bad for Education? A Political Economy Perspective on Development", *European Economic*

Review 44: 891-903.

Easterly, W. (2006): "Inequality does Cuase Underdevelopment: Insights from a New Instrument", forthcoming in *Journal of Development Economics.*

Engerman, S. and K. Sokoloff (2000): "Institutions, Factor Endowments, and Paths of Development in the New World", *Journal of Economic Perspectives* 14: 217-232.

Galiani, S., D. Heymann, C. Dabús and F. Thomé (2005): "Land-Rich Economies, Education and Economic Development", Universidad de San Andrés, Mimeo.

Galor, O. and O. Moav (2006): "Das Human Kapital", *Review of Economic Studies* 73, 85-117.

Galor, O.; O. Moav and D. Vollrath (2005): "Land Inequality and the Emergence of Human Capital Promoting Institutions", Brown University Economics Working Paper 05-03.

Galor, O. and J. Zeira (1993): "Income Distribution and Macroeconomics", *Review of Economic Studies* 60: 35-52.

Glaeser, E., R. La Porta, F. Lopez-de-Silanes and A. Shleifer (2004): "Do Institutions Cause Growth?", *Journal of Economic Growth* 9 (3), 271-303.

Landes, D.S. (1998): **The Wealth and Poverty of Nations: Why Some are so Rich and Some so Poor**, Norton, New York.

Mariscal, E. and K. Sokoloff (2000): "Schooling, Suffrage, and Inequality in the Americas, 1800-1945", in S. Habber (ed.) **Political Institutions and Economic Growth in Latin America**, Hoover.

Martínez Paz, F. (2003): "Enseñanza Primaria, Secundaria y Universitaria (1862-1914)", in **Nueva Historia de la Nación Argentina, Vol. 6**, Academia Nacional de la Historia, Planet.

North, D. (1981): **Growth and Structural Change**, Norton, New York.

Tedesco, J. C. (1993): **Educación y Sociedad en la Argentina (1880-1945),** Ediciones Solar, Buenos Aires.

7 Appendix

Lemma 1 *The consumption-bequest choices of an agent are as follows:*

- *If* $i \leq \widetilde{c}_{A1}$ *then* $c \in \mathcal{C}^1$*:* $c_A = i$*. Let* $\widetilde{i}_1 = \widetilde{c}_{A1}$*.*
- *If* $\widetilde{i}_1 < i \leq \widetilde{i}_2 = \widetilde{c}_{A2} + \widetilde{b}_2^o$ *then* $c \in \mathcal{C}^2$ *and demands are:*

$$c_A = (1-\beta)(i - \widetilde{i}_1) + \widetilde{c}_{A1}$$

$$b^o = \beta(i - \widetilde{i}_1)$$

- *If* $i > \widetilde{i}_2$ *then* $c \in \mathcal{C}^3$*. Let* $w(h)$ *be the wage of a worker of level of skill* h *(*$w(1) = \widetilde{w}$*). Denote by* $\mathbf{S}$ *the set of skills available in the market with the minimum ratio* $\frac{w(h)}{h}$*, indicating the wage per unit of skill:*

$$\mathbf{S} = \{h : \frac{w(h)}{h} \leq \frac{w(\hat{h})}{\hat{h}},\ \textit{for any}\ \ \hat{h}\}$$

  and for every $h \in \mathbf{S}$*,* $w(h) = \widetilde{w}h$*.*

- *Assume* $c \in C^3$*, and let* h^S *be the maximum level of skills in* $\mathbf{S}$*. Then, if the income of the agent satisfies* $\alpha(1-\beta)(i - \widetilde{i}_2) < \widetilde{n}w(h^S)$*, the individual will only hire workers belonging to set* $\mathbf{S}$*, and will choose a quantity and skill composition of those workers with the total labor capacity corresponding to the level of income of the consumer. That is, the level of skills of suppliers of services will be* $h^* = \int_{\mathbf{S}} n(h)hdh \in \mathbf{S}$ *and their number* $n^* = \int_{\mathbf{S}} n(h)dh$ *(where* $n(h)$ *is a density function describing the distribution of skills) such that* $\alpha(1-\beta)(i - \widetilde{i}_2) = n^*\widetilde{w}h^*$

  Bequests, and the consumption of agricultural goods, would be given by:

$$b^o = \beta(i - \widetilde{i}_2) + \widetilde{b}_2^o$$

$$c_A = \widetilde{c}_{A2} + (1-\beta)(1-\alpha)(i - \widetilde{i}_2)$$

$$c_N = \frac{(1-\beta)\alpha(i - \widetilde{i}_2)}{\widetilde{w}}$$

Proof*: It is immediate from the maximization of* $u(c)$ *subject to the budget constraints at each stage we have:*

- *In $\mathcal{C}^1$, the only variable of choice verifies* $c_A = i$.
- *In $\mathcal{C}^2$, the two variables of decision, c_A and b^o are such that* $c_A = (1-\beta)(i-\widetilde{i}_1) + \widetilde{c}_{A1}$ *and* $b^o = \beta(i - \widetilde{i}_1)$.
- *In $\mathcal{C}^3$, the variables are c_A, b^o and c_N, which are* $c_A = \widetilde{c}_{A2} + (1-\beta)(1-\alpha)(i-\widetilde{i}_2)$, $b^o = \beta(i-\widetilde{i}_2) + \widetilde{b}^o_2$ *and* $c_N = \frac{(1-\beta)\alpha(i-\widetilde{i}_2)}{p_N}$.

On the other hand, by the definition of $\mathbf{S}$, *since for* $h = 1$, $\frac{w(1)}{1} = \widetilde{w}$, *and from* $w(h) \geq \widetilde{w}h$, *we have that for every* $h \in \mathbf{S}$, $w(h) = \widetilde{w}h$. *From this, it follows that since the value of a single individual with skills h is* $p_N h = \widetilde{w}h$ *then* $p_N = \widetilde{w}$. *Therefore, the demand of services can be satisfied by a combination of skills drawn from* $\mathbf{S}$ *in any number less or equal than* $\widetilde{n}$. $\square$

Lemma 2 *If i is such that* $\alpha(1-\beta)(i-\widetilde{i}_2) > \widetilde{n}\widetilde{w}h^S$ *but in the labor market only* $h \leq h^S$ *levels of skills are available, the demands and bequests become:*

$$c_A = \widetilde{c}_{A2} + \frac{(1-\alpha)(1-\beta)}{\beta + (1-\alpha)(1-\beta)}(i - \widetilde{i}_2 - \widetilde{n}\widetilde{w}h^S)$$

$$b^o = \widetilde{b}^o_2 + \frac{\beta}{\beta + (1-\alpha)(1-\beta)}(i - \widetilde{i}_2 - \widetilde{n}\widetilde{w}h^S)$$

$$c_N = \widetilde{n}h^S$$

Proof: *Immediate from the maximization of $u(c)$ subject to the budget constraint and the additional condition* $c_N = \widetilde{n}h^S$. $\square$

Lemma 3 *Given a value of expenditure in services, denoted by $p_N c_N$, and defining the average level of wages and skills of the workers that participate in the production of those services:* $\int w(h)g(h)dh = \overline{w}$, $\int hg(h)dh = \overline{h}$:

- *If the function $w(h)$ is strictly convex, then an agent demands only one quality of labor, with skills* $h = \overline{h}$
- *If the function $w(h)$ is convex, then whenever the agent demands a mix of skills, he is indifferent with hiring only one quality of labor, with skills* $h = \overline{h}$.
- *If $w(h)$ is strictly concave on a closed set of skills* $\mathbf{H}$, *the agents will demand zero workers with skills in the interior of* $\mathbf{H}$.

Proof: *The maximal production of services for a given cost solves:*

$$\max\ \psi(n)\int hg(h)dh \quad (*)$$

$$s.t.\ \ p_N c_N = n\int w(h)g(h)dh$$

$$and \int g(h)dh \ = \ 1$$

where the distribution $g(h) = \frac{n(h)}{n}$ *where* n *is the total number of workers and* $n(h)$ *those of skill level* h.

The dual problem of cost minimization is,

$$\min\ n\int w(h)g(h)dh \quad (**)$$

$$s.t.\ \bar{y} \ = \ \psi(n)\int hg(h)dh$$

$$and \int g(h)dh \ = \ 1$$

If $w(\cdot)$ *is assumed to be strictly convex,* $\int w(h)g(h)dh \ > \ w(\int hg(h)dh)$ *for every skill distribution* $g(h)$. *Cost minimization implies that the optimal demand is concentrated on a single skill level* $\overline{h}$, *so that* $g(h)$ *is non-zero only at* $\overline{h}$. *Problem* $(*)$ *then reduces to:*

$$\max \psi(n)\overline{h},\ s.t.\ p_N c_N = nw(\overline{h})$$

If $w(h)$ *is convex (but not strictly),* $\int w(h)g(h)dh \ \geq \ w(\int hg(h)dh)$. *Local cost minimization implies now that the demand is concentrated on a convex interval* $\mathbf{H}$ *of skills levels on which* $g(h)$ *is non-zero. Problem* $(*)$ *then reduces to:*

$$\max \psi(\bar{n}), \bar{h} s.t.\ p_N c_N \ = \ nw(\bar{h})$$

where $\bar{n} = \int_{\mathbf{H}} n(h)dh$ *and* $\bar{h} = \int_{\mathbf{H}} hn(h)dh$. *Since* $\mathbf{H}$ *is assumed convex (recall that* h *is a continuous variable), there exists a single level* $h \in \mathbf{H}$ *such that* $h = \bar{h}$, *i.e. the agent is indifferent between choosing* h *or the combination of skill levels that yields* $\bar{h}$.

Finally, if $w(h)$ *is strictly concave for some values of* h, *then* $\int w(h)g(h)dh \ < \ w(\int hg(h)dh)$ *over a closed set* $\mathbf{H}$. *Local cost minimization implies that the demand will be concentrated on either* $\sup_{\mathbf{H}}$ *or* $\inf_{\mathbf{H}}$, *the extreme values in* $\mathbf{H}$. *Therefore, the agent will demand zero workers in the interior of* $\mathbf{H}$. $\square$

Lemma 4 *Suppose $\alpha(1-\beta)(i-\widetilde{i}_2) > \widetilde{n}\widetilde{w}h^S$ and that workers with skills $h \geq h^S$ can be hired at wages given by a continuous (also at h^S), convex function $w(h)$, the consumer employs $\widetilde{n}$ workers with a single level of skills h^* given by:*

$$\alpha(1-\beta)(i-\widetilde{i}_2) \;=\; \widetilde{n}w(h^*)(\epsilon_{w(h^*)}(1-\alpha(1-\beta))+\alpha(1-\beta))$$

where $\epsilon_{w(h^*)}$ is the elasticity of the wage function at h^*. Therefore, spending in services is $\widetilde{n}w(h^*) < \alpha(1-\beta)(i-\widetilde{i}_2)$.

Proof: *By Lemma 3 we have that at equilibrium levels of skills, i.e. in which the demand is non-zero, $w(h)$ is convex. Furthermore, demand will be either concentrated on a single level h or there exists such a level indifferent to a mix of demanded skills. In either case, we can assume that $c_N = \widetilde{n}h$. The maximization of $u(c)$ subject to this condition yields the relation between income and demand of services. The last claim follows from the fact that $\epsilon_{w(h^*)} > 1$ because $h^* > h^S$ and $w(\cdot)$ is convex at h^*. Therefore $\epsilon_{w(h)}(1-\alpha(1-\beta))+\alpha(1-\beta) > 1$.* □

Proposition 1 *Consider $\widehat{X}_1^h$ and $\widehat{X}_1^l$, the highest and lowest steady states of $X_t = \beta(f(X_{t-1})-\widetilde{i}_1)$in $\mathcal{C}^2$. Then, if $f(\widehat{X}_1^h) > \widetilde{i}_2 \;=\; f(\widetilde{X}_2) > f(X_0) > f(\widehat{X}_1^l)$, where the economy will reach stage 3 (landlords consume agricultural goods, bequests and services in $\mathcal{C}^3$)*

Proof: *First of all, notice that $f(0) < \widetilde{i}_1$ (otherwise no landlord would leave bequests) and $\beta(f(X)-\widetilde{i}_1)$ is a concave function. Therefore it has two steady states, namely $\widehat{X}_1^h$ and $\widehat{X}_1^l$ with $\widehat{X}_1^l < \widehat{X}_1^h$. Now, if $f(\widehat{X}_1^h) > f(X_0) > f(\widehat{X}_1^l)$ capital will be increasing in time while $\widehat{X}_1^h > X_0$. If $f(\widetilde{X}_2) > f(X_0)$ the economy will start in stage 2. Finally, If the capital required for entering stage 3, $\widetilde{X}_2$, is less than $\widehat{X}_1^h$, the economy does not get trapped in stage 2 and will eventually reach stage 3.* □

Proposition 2 *If the bequests b received by the present generation of young landlords is such that $(1-\beta)\alpha(f(b)-\widetilde{i}_2) \leq \widetilde{n}\widetilde{w}$, there is no public education. If $(1-\beta)\alpha(f(b)-\widetilde{i}_2) > \widetilde{n}\widetilde{w}$ the education system will provide a single level of education e to all the group of individuals who receive training, $\widetilde{n}$ workers will be educated and e will verify the following condition (with f' the marginal product of capital in agriculture and h' the marginal product of education in terms of skills, respectively):*

$$\frac{\alpha(1-\beta)}{1-\alpha(1-\beta)}\frac{f(b-\widetilde{n}e(h^*))-\widetilde{i}_2-\widetilde{n}\widetilde{w}h^*}{\widetilde{n}h^*} = \frac{(f'(b-\widetilde{n}e(h^*))+\widetilde{w}h'(e(h^*)))}{h'(e(h^*))}$$

Furthermore, the equilibrium level of education will be increasing with the landlords' income, i.e. $\frac{de}{db} > 0$.

Proof: *If* $(1-\beta)\alpha(f(b)-\widetilde{i}_2) \leq \widetilde{n}\widetilde{w}$ *landlords will hire only unskilled workers, i.e. with* $h = 1$*, at wage* $\widetilde{w}$*. They will gain nothing from starting public education, since their demands of services will already be fulfilled.*

If $(1-\beta)\alpha(f(b)-\widetilde{i}_2) > \widetilde{n}\widetilde{w}$ *it means that workers with abilities* $h > 1$ *can be hired. Consider now that the planner wants to select a level of education* e *that maximizes the income of the representative agent while ensuring that* $c_N = \frac{(1-\beta)\alpha(f(b-ne)-\widetilde{i}_2)}{\widetilde{w}}$*. This level is unique, since the demanded* h*, as has been argued in Lemma 3, is unique and* $h(e)$ *is strictly concave.*

This problem is equivalent to maximizing $b - ne$ *subject to* $(1-\beta)\alpha(f(b-ne)-\widetilde{i}_2) = n\widetilde{w}h(e)$*. The first order conditions with respect to* n *and* e *yield that*

$$n\widetilde{w}(h'(e)e \ - \ h(e)) \ \geq \ 0$$

i.e., the optimal level of education tax e^* *is such that the elasticity of* h *at* e^* *is* $\epsilon_{h(e^*)} \geq 1$*. An interior solution implies* $\epsilon_{h(e^*)} = 1$*, but by the concavity of* $h(e)$ *this condition may be fulfilled by rather high values of* e*. Therefore the solution may be accomplished by any* e *such that* $\epsilon_{h(e)} \geq 1$*, but since the Lagrangian function of the planner's problem* ($L(n,e) = b - ne - (1-\beta)\alpha(f(b-ne)-\widetilde{i}_2) - n\widetilde{w}h(e)$) *is strictly decreasing in* e*, the solution is at the lowest feasible level* e^**.*

In turn, differentiating $(1-\beta)\alpha(f(b-ne)-\widetilde{i}_2) = n\widetilde{w}h(e)$ *with respect to* e *yields:*

$$\frac{dn}{de} \ = \ -\frac{(1-\beta)\alpha f'(b-ne) + \widetilde{w}h'(e)}{(1-\beta)\alpha f'(b-ne)e + \widetilde{w}h(e)}.$$

It can be easily seen that for low values of e, $|\frac{dn}{de}| >> 1$*. That is, for the low* e^**, the optimal solution is to hire* $\widetilde{n}$ *workers such that* $(1-\beta)\alpha(f(b-\widetilde{n}e^*)-\widetilde{i}_2) = \widetilde{n}\widetilde{w}h(e^*)$*.*

Due to the decreasing returns of spending in education in terms of producing skills, cost minimization by the government implies that it will provide a uniform level of skills to all $\widetilde{n}$ *individuals who receive education.*

Once the planner has chosen the amount of individuals to be educated, he is able to optimize the utility of the representative landlord, taking into account his budget constraint as modified by the taxes on bequests which finance education. The utility function in stage C^3 *can be transformed by two monotone transformations, namely subtracting* k_2 *and taking logarithms:*

$$(1-\beta)(1-\alpha)\ln(c_A - \widetilde{c}_{A2}) + (1-\beta)\alpha \ln c_N + \beta \ln(b^o - \widetilde{b}^o_2)$$

Since now the per-landlord spending in education is $\widetilde{n}e$, the planner maximizes this function subject to the condition:

$$f(b-\widetilde{n}e) = c_A + b^o + \widetilde{n}w(h(e))$$

where $w(h(e))$ is the wage that the individuals with skills $h(e)$ will receive in the market when of working age. Notice that b is the bequest received from the previous generation, while b^o is the one that the current generation will leave for its descendents.

Due to the competition among workers, it will be the case that wages will equal the opportunity cost given by the returns to labor in the subsistence sector, so that $w(h(e)) = \widetilde{w}h(e)$.

Then, the problem can be reformulated as maximizing with respect to c_A, b^o and e

$$(1-\beta)(1-\alpha)\ln(c_A - \widetilde{c}_{A2}) + (1-\beta)\alpha \ln \widetilde{n}h(e) + \beta \ln(b^o - \widetilde{b}^o_2)$$

Subject to

$$f(b-\widetilde{n}e) = c_A + b^o + \widetilde{n}\widetilde{w}h(e)$$

The first order conditions can be summarized as:

$$\frac{(1-\alpha(1-\beta))(f'(b-\widetilde{n}e)+\widetilde{w}h'(e))}{\Big(f(b-\widetilde{n}e)-\widetilde{i}_2-\widetilde{n}\widetilde{w}h(e)\Big)} = \frac{\alpha(1-\beta)h'(e)}{\widetilde{n}h(e)}$$

Differentiating this expression, we have:

$$\widetilde{n}(1-\alpha(1-\beta))((f''(b-\widetilde{n}e)db-\widetilde{n}f''(b-\widetilde{n}e)de+\widetilde{w}h''(e)de)h(e)+(f'(b-\widetilde{n}e)+\widetilde{w}h'(e))h'(e)de) =$$
$$= \alpha(1-\beta)(h''(e)(f(b-\widetilde{n}e)-\widetilde{i}_2-\widetilde{n}\widetilde{w}h(e))de+h'(e)(f'(b-\widetilde{n}e)db-\widetilde{n}f'(b-\widetilde{n}e)de-\widetilde{n}\widetilde{w}h'(e)de)).$$

And regrouping terms:

$$\alpha(1-\beta)h'(e)f'(b-\widetilde{n}e)-n(1-\alpha(1-\beta)f''(b-\widetilde{n}e) \;=\; \frac{de}{db}((\widetilde{n}(1-\alpha(1-\beta))\widetilde{w}h(e)-$$
$$-\alpha(1-\beta)(f(b-\widetilde{n}e)-\widetilde{i}_2-\widetilde{n}\widetilde{w}h(e)))h''(e)+de(\widetilde{n}(1-\alpha(1-\beta)))(-\widetilde{n}f''(b-\widetilde{n}e)+$$
$$+(f'(b-\widetilde{n}e)+\widetilde{w}h'(e))h'(e))+\alpha(1-\beta)h'(e)(\widetilde{n}f'(b-\widetilde{n}e)+\widetilde{n}\widetilde{w}h'(e)))$$

Because of the properties of f and h, the left hand expression is positive. In turn, the expression that multiplies $\frac{de}{db}$ is positive if:

$$\widetilde{n}\widetilde{w}h(e) - \alpha(1-\beta)(f(b-\widetilde{n}e)-\widetilde{i}_2) \leq 0.$$

which, as said above, obtains from the initial decision of the planner on both the number and level of education of the workers to be trained. Therefore:

$$\frac{de}{db} > 0. \quad \square$$