**REM WORKING PAPER SERIES**

# Education spread, technology, and population density

**José Pedro Pontes**

**REM Working Paper 0383-2025**

July 2025

**REM – Research in Economics and Mathematics**
Rua Miguel Lúpi 20,
1249-078 Lisboa,
Portugal

ISSN 2184-108X

**REM – Research in Economics and Mathematics**

Rua Miguel Lupi, 20
1249-078 LISBOA
Portugal

Telephone: +351 - 213 925 912
E-mail: rem@iseg.ulisboa.pt

https://rem.rc.iseg.ulisboa.pt/

https://twitter.com/ResearchRem

https://www.linkedin.com/company/researchrem/

https://www.facebook.com/researchrem/

# Education spread, technology, and population density

by

José Pedro Pontes[1]

**Date**: July 2025

**Abstract:** We model the expansion of (higher) education in an economy composed by regions that only differ in population density. The schooling process takes place sequentially across regions in descending order of demographic density and it implies a substitution of modern industrial technologies for traditional land-based ones.

Under the crucial assumption that young people may travel to school within the region where they live, but not across regions, the model explains why both the literacy rate and per capita income increase, albeit at a decreasing rate. Furthermore, is allows us to understand why the average students' commuting distance tends to rise despite the geographical decentralization of the educational system.

**Keywords:** Education Spread, Population Density, Spatial Monopolistic Competition, Wage Premium of Education, Modern versus Traditional Technology.

**JEL Classification:** O18, R11, I20.

---

José Pedro Pontes
Email pspontes@iseg.ulisboa.pt
[1]Universidade de Lisboa, ISEG Lisbon School of Economics & Management, ISEG Research in Economics and Management, Rua do Quelhas, 1200-781, Lisbon, Portugal.

# 1. Introduction

The role played by population density in facilitating the expansion of literacy has been acknowledged for a long time. If we follow Salop (1979), we can feature a spatial economy where households and their offspring are uniformly distributed. In this context, private schools decide first to enter freely, then they choose locations and eventually they set competitive tuition fees. If population rises, the costs of providing education decrease as Duranton and Puga (2004) noticed.

In the short run with a fixed number of schools, the fixed cost of each school – in terms of professors, buildings, libraries and labs - is “shared” by a larger number of students, so that the cost of providing education per student consistently falls.

In the long run, a larger cohort of young people enables more schools to break even, thereby reducing the average distance between the student’s residence and the closest school and hence the average commuting cost. Such an enhancement in the “matching” of students to schools indeed matters, as Spiess and Wrohlich (2010) and Frenette (2006) deemed that concerning higher education the student’s travel cost to the nearest university is an important hindrance to enrolment.

Since in many countries secondary education is now compulsory, we deal here with the spread of higher education, which corresponds to ISCED 5-8. In this line of reasoning, we can observe that population density in Portugal explains about 90 percent of total variation in higher education schooling rates across NUTS2 regions (see Pontes, 2024).

While population density seems inversely related with the average production and transport costs of providing education, we observe that the schooling activity tends to spread out of densely populated urban areas to more sparsely inhabited rural territories. Since this evolution increase average education costs, this paper discusses the policy options that arise from such a loss in efficiency related with the teaching activity.

At this stage of reasoning, we notice that literacy spread is closely associated with technical progress. In fact, Galor and Weil (2000) in line with Schultz (1964) argued that the rise in aggregate productivity derived from the adoption of new production techniques stimulates school attainment since young people traditionally educated within the family are likely to be less able to deal with new activities and challenges in their professional activity.

As in Salop (1979), the provision of education by increasing returns schools to young people who must travel to the nearest school faces the competition of an *outside good*, which is produced under constant returns to scale and therefore it is available everywhere in the economy without need for transportation.

We assume that school attainment enables a young person to switch from a “traditional” technology to a “modern” one. Following Tamura (2002), while the former method of production has constant returns in relation to both labour and land, the latter technology exhibits constant returns to labour only. Hence an increasing employment density has an opposite effect on labour productivity under the two kinds of technology. In fact, it strictly decreases output per worker under the “traditional” method of production as labour productivity is positively associated with the amount of land that is available to each worker. By contrast, a rise in employment density does not change the output per worker under the “modern” technology.

A rise in population density may *per se* lead to a transition from the “traditional” to the “modern” technology. As Boserup (1965) says,

> *“…the intensification of agriculture may compel cultivators and agricultural labourers to work harder and more regularly. This can produce changes in work habits which help to raise overall productivity. (p. 105)*

Nonetheless, the reduction in the parcel of land available to an individual worker might likely not decrease his output only if it is offset by an enhanced training derived from schooling. Furthermore, a rising population enables more schools to break even by allowing that more users share fixed teaching assets.[2]

According to Boserup (1965), an economy may be composed by two kinds of territory. Firstly, there are urban, densely populated regions, where production takes place under a “modern” technology, which is carried out by high-wage workers with school attainment. Secondly, the economy also contains sparsely populated areas, where low-wage, non-educated workers carry out production under a “traditional” technology. If interregional migrations are negligible, such a structural difference might remain invariant for a long period.

In this paper, we model education spread by considering that the economy is made up by a set of regions with diverse population density. We presuppose that travel costs across regions are prohibitive, so that workers may not move either during training or professional work while they can move within the region where they live.

[2] This sharing mechanism concerns not only schools but also other fixed infrastructure such as roads. Glover and Simon (1975) showed that population density fosters the construction of roads especially those that require large fixed outlays such as paved roads.

We further assume that in each time the Government attempts to formally educate people living in a different region and it selects the regions in a descending order of demographic density, i.e. it first achieves an educational investment in the densest region, then in the secondly densest area and so on.

This framework is analogous the one put forward by Boucekkine, De La Croix and Peeters (2007), which describes the evolution of schooling in a region where population steadily declines. In each time, the Government faces a new generation of young people, and it sets up a network of schools dedicated to them. When the members of the cohort have completed education, these schools are scrapped, and a new set of schools are built to meet the needs of the following generation. Each person must be trained and assigned to work within his generation.

This conceptual setting allows us to rationalize five main stylized facts concerning the link of education spread with economic development. First, the higher education system grows steadily, albeit at a decreasing rate. Second, the expansion of college education increases aggregate productivity, despite this effect tends to fade away as schooling rates rise. Third, the "wage premium" of university education declines with the rise of tertiary literacy rates. Fourth, while expanding the higher education system appears to decentralize both in terms of college locations and the geographical origin of students. Lastly, even though the college network indeed decentralizes, the evolution of average students' commuting distances does not follow a simple pattern. While it seems to have decreased in a first stage as new regional universities were created, it likely increased afterwards as the growth in schooling rates was eventually biased to low-density regions with large average travel distances.

# 2. Main stylized facts concerning the spread of higher education and economic development

We try to provide a reasonable explanation to a collection of main trends of the growth and geographical spread of higher education (ISCED 5-8).

**The first trend consists in the fast growth in tertiary educational attainment albeit at a decreasing rate**.

For the US, Gordon (2018) argues that the annual growth rate in the share of people with a complete four-year college degree was about 3.7 percent in the period 1915-1980 and it fell to 1.3 percent between 1980 and 2016. For Europe, Pontes and Buhse (2019) found a strong $\beta-$ convergence speed of about $-4.0$ of the share of people aged between 25 and 34 endowed with a complete higher education degree. For Portugal, Pontes (2025) finds that the average annual growth rates in the share of people older than 15 with complete tertiary education evolved from 6.4 percent in the period 1981-2001 to 4.8 percent during 2001-2021, a change that reflects the comparatively low initial schooling rates in the university.

**The second trend consists in a positive initial impact of higher education spread on labour productivity and real per capita GDP, which gradually slows down and eventually disappears**.

The increase in labour productivity in the US was reduced by a factor of three along successive time periods. In fact, the average annual growth rate decreased from 2.8 percent during 1920-1970 to less than1.8 in 1970-2006 and less than 1 percent in the time interval 2006-2016. For Europe, Pontes and Buhse (2019) computed a $\beta-$ convergence speed of about $-2.3$ for real GDP per capita, which is much lower than the $-4.0$ estimated speed for college attainment levels. Using PORDATA-INE data, we can compute the average annual growth rates of real per capita GDP in periods 1981- 2001 and 2001-2021 as 3.0 and 0.4 percent, respectively.

**The third trend consists in the decline in the "wage premium" of higher education as measured by the average difference between the wages earned by a graduate and by a worker with secondary education only.**

The wage premium of higher education in the US increased fast between 1980, it continued to rise slowly during 2000-2010 and then remained constant during the economic recovery in 2010-2015. In addition, graduate underemployment, which is defined as a graduate performing a job that does not require college education of 25, increased from 36 to 46 percent of the set of 25 years old graduates between 2000 and the period 2009/2011. At the OECD level, the wage premium of higher education diminished 3 percent between 2006 and 2016, while for Portugal the decrease was much more severe reaching about 23 percent.

**The fourth main trend was the decentralization of the higher education system, both in terms of the spatial distribution of universities and the geographical origin of students.**

For 27 European countries, Pontes and Buhse (2019) estimated a high $-4.0$ $\beta-$ convergence speed of the share of tertiary-educated people during 2004-2018 thus showing that the higher education system strongly decentralized during that period. For Portugal, the average annual growth rates of the share of college educated people between 2001 and 2021 were 3.0 percent for the major metropolitan areas, which host the oldest universities (i.e., Lisbon and Oporto metropolitan areas and the NUT3 region of Coimbra) and about 7.7 percent for the remaining and less populated regions. Such a spread out of higher education was supported by the creation since 1974 of twelve regional universities and fifteen polytechnic institutions, which converged to university standards.

**The fifth main trend consists in the fact that average students' commuting distances eventually increased even though the college network became considerably more decentralized.**

According to Pontes (2025), the average student's travel distance in Portugal decreased between 1981 and 2001 and then increased between 2001 and 2021, when it attained a higher level than in the initial year 1981. Two main factors account for the increase in average commuting in the more recent period. On the one hand, despite the number of higher education institutions increased between 1981 and 2001, it seems to have slightly diminished in 2001-2021, a change that concerns private colleges only. On the other hand, as we noticed above the growth rate of enrolled young people was considerably higher in territories with low college density, so that those students faced larger commuting distances.

# 3. A model of education spread

## 3.1. Main assumptions

We consider an economy composed by $T+1$ regions. Each region $t=0,1,2,\ldots,T$ consists in a circle of unit length and has a population $L_t$. Regions are indexed in descending order of population density, so that we have,

$$L_0 > L_1 > L_2 > \ldots > L_{T-1} > L_T \tag{1}$$

We assume that in time $t=0,1,2,\ldots,T$, private schools freely enter a region, then select locations and finally decide a tuition fee. Hence, formal education is set up sequentially across regions by descending order of demographic density. In time 0, schools install in region $0$, then a new set of schools settles in region 1, and so on.

Formal education implies both school fixed costs $F$ and commuting cost with $\tau$ standing for the student's travel cost per unit of distance. We assume that schooling is economically feasible in every region, a condition that is expressed by the fact that population density $L_t$ is bounded from below by unit education costs.

The infimum of $L_t$ combines two constraints. Firstly, we must have $L_t > 1$, i.e., the number of potential students should exceed the minimal number of schools in the region. Secondly, population density should be high enough in relation to the costs of education, which is expressed by $L_t > \frac{25}{16}(\tau F)$. Hence, the infimum of $L_t$ is expressed by the inequality,

$$L_T > \max\left\{1, \frac{25}{16}(\tau F)\right\} \tag{2}$$

We presuppose that interregional travel costs faced by both students and workers are prohibitive, so that a young person might neither study nor work in a region other than the one where he lives.

This sequential process of education development is analogous to the gradual change of the educational system in a region with a declining population that was described by Boucekkine, De La Croix, and Peeters (2007). In such a framework, a new generation of young people joins college in time $t$. A dedicated network is built in this time, and it is attended by this generation of students only. Then, these colleges are scrapped, and a new set of schools are set up to host the next generation of students. Each young person must complete education in a college specifically built for his generation.

Given the analogy of the multiregional model with the evolution of schooling in a region with a declining demography and for simplicity, we will model the situation in continuous time $t \in [0, T]$. Consequently, we deal with an economy with a continuum of circular regions where population density $L(t)$ strictly decreases with time/region $t$ while being bounded from below by inequality (2)

## 3.2. The school network in region/time $t$

We consider the circular region $t$, which is inhabited by $L(t)$ households. Each household is assumed to have one offspring. For simplicity we assume that the upbringing of a young person and its engagement in college and professional occur in the same time $t$. We presuppose that households and their offspring are uniformly distributed in space. We drop the index $t$ in this section for ease of exposition.

The equilibrium network of schools is modelled according to Salop (1979) spatial monopolistic competition. According to this framework $n$ schools enter the region and they select symmetric locations, a pattern that reflects the symmetry in the spatial distribution of young people. Figure 1 plots the equilibrium locations of $n = 4$ schools.

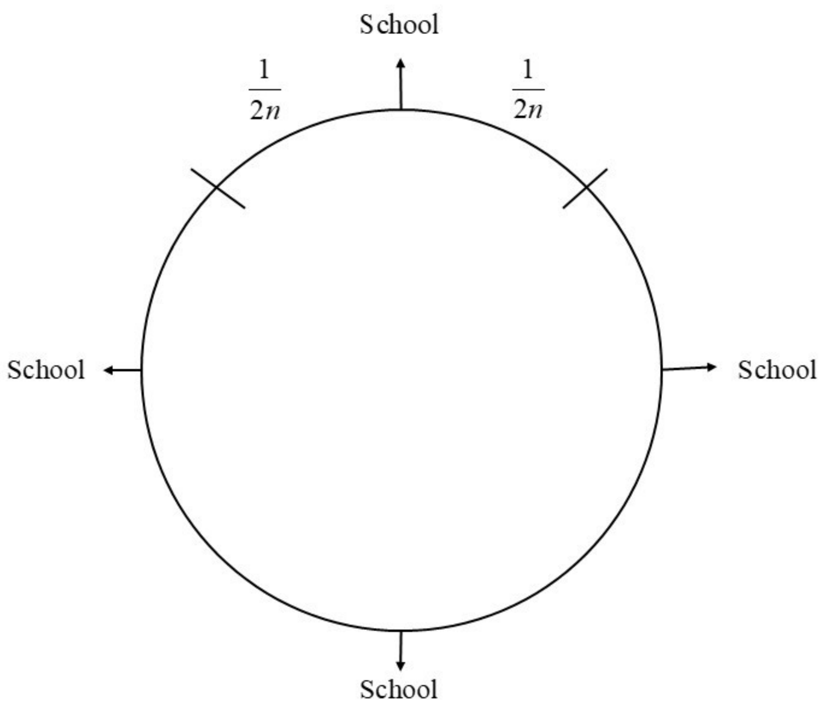


Figure 1: Symmetric network of $n = 4$ schools

We assume that each school bears a fixed cost $F$ and – for simplicity – has zero variable costs. To enrol in college, a young person pays a tuition fee $p$ and bears an additional travel cost $\tau\tilde{d}$, where $\tilde{d}$ is the distance to the nearest university and $\tau$ is the commuting cost per unit of distance. Hence, the total cost of education borne by a youngster is,

$$p + \tau\tilde{d} \tag{3}$$

Clearly a young person will decide to join college if the wage premium of education, i.e., the difference between $w^S$, the wage of skilled labour which he earns from graduating, and the wage $w^U$ of unskilled, uneducated labour is sufficient to offset the education cost (3). This condition may be written as,

$$\begin{aligned} w^S - \left(p + \tau\tilde{d}\right) \geq w^U \quad \text{or} \\ \Delta \equiv w^S - w^U \geq p + \tau\tilde{d} \end{aligned} \tag{4}$$

where $\Delta \equiv w^S - w^U$ stands for the wage premium of education.

We now follow Salop (1979) to derive the equilibrium number $n^*$ of symmetrically located schools and uniform tuition fee $p^*$. We now assume that when schools quote the tuition $p^*$ all youngsters prefer to join college rather engaging in the labour market immediately. This is equivalent to stating that the demand addressed to a school is not constrained by the prevailing wage premium $\Delta$. Figure 2 illustrates this situation.

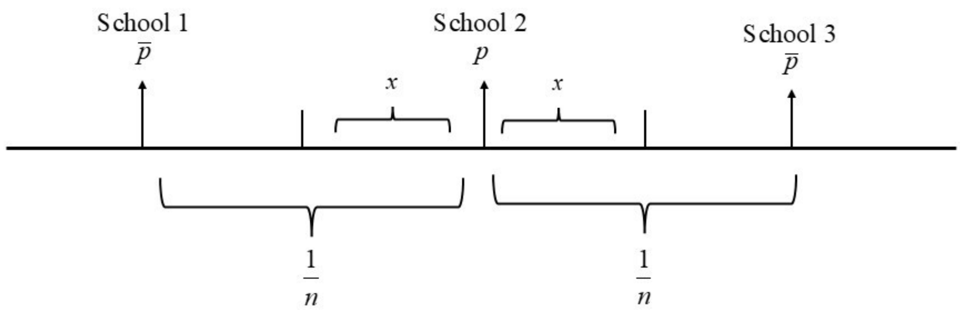


Figure 2: Competition in tuition fees by schools

Let us assume that the representative school 2 in Figure 2 charges tuition $p$, while its neighbours at distance $\frac{1}{n}$ quote $\overline{p} \neq p$. Let $x$ be the distance of school 2 the border of its market area. Then, $x$ is determined by the equality of the full prices of education provided by two neighbouring schools,

$$p + \tau x = \overline{p} + \tau\left(\frac{1}{n} - x\right) \tag{5}$$

By solving for $x$ we obtain,

$$x = \frac{\overline{p} - p}{2\tau} + \frac{1}{2n} \tag{6}$$

Hence, the profit earned by the representative school 2 as a function of $p$ and $n$ is,

$$\pi(p,n) = pL(2x) - F \tag{7}$$

which becomes through substitution of (6) in (7),

$$\pi(p,n) = pL\left(\frac{\overline{p} - p}{\tau} + \frac{1}{n}\right) - F \tag{8}$$

Given the number of schools $n$ and the fact that $\pi$ is a strictly concave function of $p$, each firm computes the uniform equilibrium tuition $p^*$ through the condition $\frac{\partial \pi}{\partial p} = 0$ to obtain,

$$p^* = \frac{\tau}{n} \tag{9}$$

Since $n$ schools are uniformly spaced in equilibrium and they charge the same tuition $p^*$, the distance between the school and its most distant student is $x = \frac{1}{2n}$. Consequently, the profit earned by each school in equilibrium is,

$$\pi(p^*,n) = p^* L\left(\frac{1}{n}\right) - F \tag{10}$$

Through substitution of (9) in (10) the school profit becomes,

$$\pi(p^*,n) = \frac{L\tau}{n^2} - F \tag{11}$$

With free entry, the school profit is reduced to zero. Condition

$$\pi\left(p^*, n\right) = \frac{L\tau}{n^2} - F = 0$$

determines the equilibrium number $n^*$ of symmetric schools.

$$n^* = \sqrt{\frac{L\tau}{F}} \tag{12}$$

By substituting $n^*$ from (12) in (9), we obtain the uniform fob mill price $p^*$.

$$p^* = \sqrt{\frac{\tau F}{L}} \tag{13}$$

Even though education is supplied by monopolistically competitive firms, we presuppose that the government can control the equilibrium number of schools by setting the cost parameters $\tau$ and $F$.

On the one hand, the government can decrease the travel cost $\tau$ b subsidizing student's accommodation, thereby reducing the number $n^*$ of colleges in equilibrium and raising the average student's commuting distance $\frac{1}{4n^*}$.

On the other hand, the public authority can subsidize a share of college fixed cost thereby increasing the density of universities. By contrast, it might increase fixed costs by constraining colleges to provide equipment such as libraries, laboratories, sports infrastructure and so on. In either case, this will have an impact on the density $n^*$ of higher education institutions.

While being necessary, condition (4) is not sufficient for a young person living at distance $\tilde{d}$ from the nearest university to enrol in tertiary education. We assume in addition that a student earns a high wage $w^S$ only if *all youngsters in the region* also decide to complete college. This coordination requirement greatly simplifies our analysis, and we might view it as reflecting local positive externalities in the formation of human capital. As Diamond (1982) argued, such externalities follow from the fact that a graduate only finds a suitable job if his specialized skill is matched to other complementary specialists within a work organization or the local labour market.

Hence in our model differences in schooling rates necessarily derive from asymmetric population density across regions and might not arise within the same homogeneous region.

If $n^*$ schools enter and locate symmetrically, the maximal distance between a student's residence and the closest school is just $\frac{1}{2n^*}$. Consequently, since all schools quote the uniform tuition $p^*$, a necessary and sufficient condition for inequality (4) is just,

$$\Delta \equiv w^S - w^U = p^* + \tau\left(\frac{1}{2n^*}\right) \tag{14}$$

By substituting $p^*$ and $n^*$ from (13) and (12) in (14), we obtain the condition,

$$\Delta \equiv w^S - w^U = \sqrt{\frac{\tau F}{L}} + \frac{\tau}{2\sqrt{\frac{L\tau}{F}}}$$

which can be simplified to,

$$\Delta \equiv w^S - w^U = \left(\frac{3}{2}\right)\sqrt{\frac{\tau F}{L}} \tag{15}$$

It remains to model the wages of educated and uneducated workers. We presuppose that both categories of workers engage in productive activities characterized by constant returns to scale. Graduates produce the composite consumer good by means of an "industrial" modern technology that exhibits constant returns to scale in relation to labour according to the production function,

$$Y_{ind} = A_{ind} L \tag{16}$$

where $Y_{ind}$ is total output, $A_{ind}$ is a total factor productivity or TFP, and $L$ is total labour input. Consequently, the output per worker is,

$$y_{ind} \equiv \frac{Y_{ind}}{L} = A_{ind} \tag{17}$$

which is a constant.

By contrast, uneducated workers are assumed to produce the composite good by an "agricultural" or "housing" traditional technology that shows constant returns to scale in relation to both labour *and* land according to the Cobb-Douglas production function,

$$Y_{ag} = A_{ag} L^{\alpha} S^{(1-\alpha)} \tag{18}$$

where $Y_{ag}$ is total output of land-based activity, $A_{ag}$ is TFP of this kind of technology and $S$ is the total input of land.

Consequently, output per worker in the traditional technology is,

$$y_{ag} = \frac{Y_{ag}}{L} = A_{ag}\left(\frac{L}{S}\right)^{(\alpha-1)} \tag{19}$$

Hence, output per worker in the traditional technology decreases with population density $\frac{L}{S}$. An alternative way of expressing $y_{ag}$ is,

$$y_{ag} = A_{ag}\left(\frac{S}{L}\right)^{(1-\alpha)} \tag{20}$$

Output per worker under the traditional technology depends positively on the amount of land that the worker may use as an input.

Since the TFP terms $A_{ind}$ and $A_{ag}$ are exogenous, we may assume that they are both equal to 1. Moreover, the total input of land $S$ is the unit measure o distance, hence equal to 1. With these simplifications, aggregate production functions (16) and (18) become,

$$Y_{ind} = L \tag{21}$$

and,

$$Y_{ag} = L^{\alpha} \tag{22}$$

In turn, the wages of educated and uneducated workers can be written as,

$$\begin{aligned} w^S &= y_{ind} = 1 \\ w^U &= L^{(\alpha-1)} \end{aligned} \tag{23}$$

From (23), the wage premium of higher education is,

$$\Delta \equiv w^S - w^U = 1 - L^{(\alpha-1)} \tag{24}$$

Since by (2) we have $L > 1$, $\Delta$ is positive and increases strictly with population density $L$.

## 3.3 The spread of formal education across regions

We now consider the continuum of circular regions with population density $L(t)$, which is a strictly decreasing function. In time $t \in [0,T]$, higher education is established in region $t$. The schooling process takes place in descending order of population density, so that schools enter firstly in region 0, then in region 1 and so on.

By applying (24) to (15), we can define a function of time $g(t)$ with parameters $\tau$ and $F$ as the difference between the wage premium of education $\Delta$ and the maximal full price of higher education.

$$g(t) \equiv \left[1 - L(t)^{(\alpha-1)}\right] - \left(\frac{3}{2}\right)\sqrt{\frac{\tau F}{L(t)}} \tag{25}$$

Then, a necessary and sufficient condition of universal schooling attainment in region $t$ is simply derived.

We demonstrate the following Proposition.

**Proposition 1:** If $L(0)$ is sufficiently high and $L(T)$ is higher than but sufficiently close to $\max\left\{1, \frac{25}{16}(\tau F)\right\}$, then there is a threshold time/region $t^* \in [0,T]$ such that we have,

$$\begin{aligned} &g(t) > 0 \text{ if } 0 \le t < t^* \\ &g(t) = 0 \text{ if } t = t^* \\ &g(t) < 0 \text{ if } t > t^* \end{aligned} \tag{26}$$

**Proof**: From (25) it follows that $g(t)$ is a continuous function over the domain $[0,T]$.The first derivative of $g(t)$ is,

$$g'(t) = L'(t)\left[\frac{3}{4}(\tau F)^{\frac{1}{2}} L(t)^{-\frac{3}{2}} + (1-\alpha)L(t)^{(\alpha-2)}\right] < 0 \tag{27}$$

Hence, $g(t)$ is a strictly decreasing function.

We compute the endpoints of $g(t)$. Firstly, we notice that,

$$g(0) = 1 - L(0)^{(\alpha-1)} - \frac{3}{2}\sqrt{\frac{\tau F}{L(0)}} \tag{28}$$

Hence, $g(0) > 0$ for an arbitrarily high value of $L(0)$. Then the other endpoint is,

$$g(T) = 1 - L(T)^{(\alpha-1)} - \frac{3}{2}\sqrt{\frac{\tau F}{L(T)}} \tag{29}$$

From (2), $\max\left\{1, \frac{25}{16}(\tau F)\right\}$ is an infimum of $L_T$. Assume first that $\tau F \leq \frac{16}{25}$, so that we should have $L_T > 1$. If $L_T = 1$, then (29) becomes,

$$g(1) = -\frac{3}{2}\sqrt{\tau F} < 0 \tag{30}$$

Therefore, if $L(T) > 1$ but arbitrarily close to $1$, then we can say that $g(T)$ is negative by continuity.

Alternatively, assume that $\tau F \geq \frac{16}{25}$, so that we should have $L_T > \frac{26}{15}(\tau F)$. If $L(T) = \frac{25}{16}(\tau F)$, then $g(T)$ becomes,

$$g\left(\frac{25}{16}\tau F\right) = -\left[\frac{1}{5} + \left(\frac{25}{16}\tau F\right)^{(\alpha-1)}\right] < 0 \tag{31}$$

Therefore, if $L(T) > \frac{25}{16}(\tau F)$ but arbitrarily close to $\frac{25}{16}\tau F$, then we can say that $g(T)$ is negative by continuity.

We can thus state that there is a time/region threshold $t^*$ whose properties are expressed by (26). We plot function $g(t)$ in Figure 3.

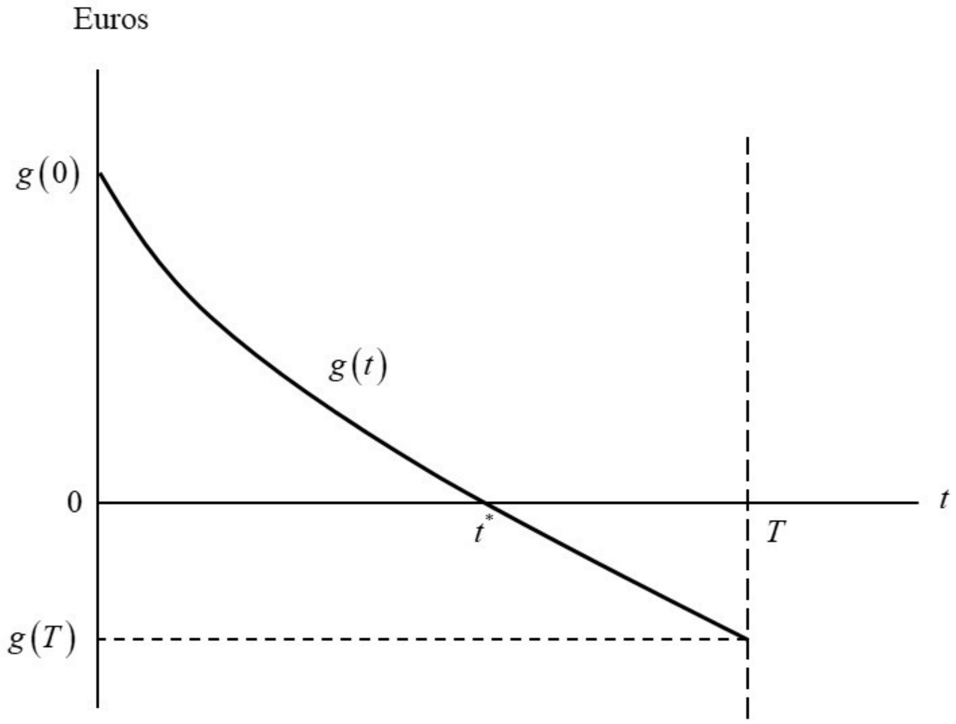


Figure 3: Set of regions with formal education

**QED**

# 4. Explaining the main empirical trends of higher education spread

We now use the model to account for the main stylized facts of the expansion of the university network, which were described in Section 2. In what follows, $t^*$ stands for the time/region threshold whose existence and uniqueness was proved in the latter section.

The following results are demonstrated ahead.

**Proposition 2: The higher education system tends to grow, albeit at a decreasing rate.**

In time $0 \le t \le t^*$, the country schooling rate is,

$$h(t) = \frac{\int_0^t L(r)\,dr}{\int_0^T L(r)\,dr} \tag{32}$$

For $0 < t < t^*$, the first and second derivatives of $h(t)$ are,

$$h'(t) = \frac{L(t)}{\int_0^T L(r)\,dr} > 0 \tag{33}$$

and,

$$h''(t) = \frac{L'(t)}{\int_0^T L(r)\,dr} < 0 \tag{34}$$

These are also the right-hand first and second derivatives of $h(t)$ for $t = 0$.

Moreover, for $t^* < t \le T$ we have,

$$h(t) = \frac{\int_0^{t^*} L(r)\,dr}{\int_0^T L(r)\,dr} \tag{35}$$

which is a constant in relation to $t$. **QED**

**Lemma 1:** Let $t^*$ be the time/region threshold defined in Proposition 1. Then, for $t \le t^*$ the total cost of education in time/region $t$ is,

$$C(t) = \frac{5}{4}\sqrt{L(t)\tau F} \tag{36}$$

**Proof:** Given the symmetry assumptions depicted in Figure 1, the average distance between a student's residence and the nearest school is,

$$\frac{1}{4n(t)}$$

Hence, the average commuting cost is,

$$\frac{\tau}{4n(t)}$$

Clearly, travel costs borne by all students are,

$$\frac{\tau L(t)}{4n(t)} \tag{37}$$

Through substitution of $n(t)$ from (12) in (37), we obtain,

$$\frac{1}{4}\sqrt{L(t)\tau F} \tag{38}$$

In addition, total fixed costs of education are,

$$n(t)F \tag{39}$$

Through substitution of $n(t)$ from (12) in (39), we obtain total fixed costs of education as,

$$\sqrt{L(t)\tau F} \tag{40}$$

By summing (38) and (40) we obtain $C(t)$.**QED**

**Proposition 3:** Let $t^*$ be the time/region threshold defined in Proposition 1. Then per capita income $y(t)$ is positive and increases in time albeit at a diminishing rate.

**Proof:** For $0 \le t \le t^*$, total output from (21) and (22) is,

$$Q(t) = \int_0^t L(r)\,dr + \int_t^T L(r)^\alpha\,dr \tag{41}$$

From Lemma 1, the accumulated cost of education in time $t$ is,

$$C(t) = \int_0^t \left(\frac{5}{4}\sqrt{L(r)\tau F}\right)dr \tag{42}$$

Total income in time $t$ is,

$$Y(t) = Q(t) - C(t) = \left[\int_0^t L(r)\,dr + \int_t^T L(r)^\alpha\,dr\right] - \int_0^t \left(\frac{5}{4}\sqrt{L(r)\tau F}\right)dr \tag{43}$$

Conditions (1) and (2) are sufficient to ensure that $Y(t)$ is positive.

Total population is $\int_0^T L(r)\,dr$, which does not depend on time/region index $t$. Hence, the sign of the first and second derivatives of per capita income have the same sign as the corresponding derivatives of $Y(t)$.

The first derivative of $Y(t)$ is,

$$Y'(t) = L(t) - L(t)^\alpha - \frac{5}{4}\sqrt{L(t)\tau F} \tag{44}$$

By factoring out $L(t)$ from (44), we obtain,

$$Y'(t) = L(t)\left[1 - L(t)^{(\alpha-1)} - \frac{5}{4}\sqrt{\frac{\tau F}{L(t)}}\right] \tag{45}$$

Clearly, the sign of $Y'(t)$ is the same as the sign of the expression in brackets in (45) .

We recall from the proof of Proposition 1 that function $g(t)$ defined by (25) has two properties namely,

$$g'(t) < 0 \text{ and } g(t^*) = 0$$

Hence, for $t \leq t^*$ we have the inequality,

$$g(t) = 1 - L(t)^{(\alpha-1)} - \frac{3}{2}\sqrt{\frac{\tau F}{L(t)}} \geq 0 \tag{46}$$

Clearly, inequality (46) is a sufficient condition of the positivity of the expression in brackets in the right hand side of (45). Indeed, we have,

$$1 - L(t)^{(\alpha-1)} - \frac{5}{4}\sqrt{\frac{\tau F}{L(t)}} > 0 \tag{47}$$

Consequently, we can say that $Y'(t) > 0$.

From (44), we may compute the second derivative of $Y(t)$.

$$Y''(t) = L'(t) - \alpha L(t) L'(t) - \frac{5}{8} L'(t)\sqrt{\frac{\tau F}{L(t)}} \tag{48}$$

By factoring out $L'(t)$ from (48), we obtain,

$$Y''(t) = L'(t)\left[1 - \alpha L(t)^{(\alpha-1)} - \frac{5}{8}\sqrt{\frac{\tau F}{L(t)}}\right] \tag{49}$$

Clearly, the sign of $Y''(t)$ is opposite to the sign of the expression in brackets in (49) . Moreover, inequality (47) is a sufficient condition of,

$$1 - \alpha L(t)^{(\alpha-1)} - \frac{5}{8}\sqrt{\frac{\tau F}{L(t)}} > 0 \tag{50}$$

Hence, we may argue that $Y''(t) < 0$.

We deal now with the case where $t > t^*$. Total income now is,

$$Y(t) = \left[\int_0^{t^*} L(r)dr + \int_{t^*}^{T} L(r)^{\alpha} dr\right] - \int_0^{t^*}\left(\tfrac{5}{4}\sqrt{L(r)\tau F}\right)dr \tag{51}$$

which differs from (43) in that the current time/region $t$ is replaced by the fixed threshold $t^*$ defined by Proposition 1 as the limit of the integrals of integration.

Consequently, $Y(t)$ is now a constant function of $t$, so that its first and second derivatives are zero. **QED**

**Proposition 4**: The wage premium of education steadily declines.

**Proof:** We recall from (24) that the wage premium of education in time $t$ is,

$$\Delta(t) \equiv w^S - w^U(t) = 1 - L(t)^{(\alpha-1)} \tag{52}$$

Its first derivative is,

$$\Delta'(t) = (1-\alpha) L(t)^{(\alpha-2)} L'(t) < 0 \tag{53}$$

**QED**

We state now a Proposition that needs not to be demonstrated.

**Proposition 5:** The higher education system tends to decentralize both in terms of the geographical origin of students and the spatial distribution of universities.

Furthermore, we can prove the following Proposition.

**Proposition 6:** Even though the college network decentralizes in time, the average commuting cost borne by students eventually increases.

**Proof:** From Figure 1, the average distance travelled by students in time/region $t \le t^*$ is $\frac{1}{4n(t)}$, where $n(t)$ is the number or density of colleges in region $t$. Hence, we may assess the evolution of commuting distance as the inverse time path of $n(t)$, which from (12) is given by,

$$n(t) = \sqrt{\frac{L(t)\tau}{F}} \tag{54}$$

Clearly $n(t)$ is strictly decreasing as,

$$n'(t) = \frac{L'(t)}{2}\sqrt{\frac{\tau}{FL(t)}} < 0 \tag{55}$$

Let $\tilde{n}(t)$ the average density of schools considering all regions up to time $t \le t^*$, i.e.,

$$\tilde{n}(t) = \frac{1}{t}\int_0^t n(r)\,dr \tag{56}$$

The first derivative of $\tilde{n}(t)$ is,

$$\tilde{n}'(t) = \frac{n(t)t - \int_0^t n(r)\,dr}{t^2} \tag{57}$$

By using the Mean Value Theorem of integrals, we know that there is a $c \in (0,t)$ that satisfies,

$$n(c)t = \int_0^t n(r)\,dr \tag{58}$$

Hence, the first derivative of $\tilde{n}(t)$ becomes,

$$\tilde{n}'(t) = \frac{1}{t}\left[n(t) - n(c)\right] < 0 \tag{59}$$

The derivative (59) is negative because $n(t) < n(c)$. Indeed, we have $c < t$ and $n(r)$ is strictly decreasing according to (55). **QED**

# 5. Discussion

While Propositions 1 to 5 explain reasonably main empirical trends of higher education spread and economic development, Proposition 6 does not fit exactly with the evidence.

Empirical evidence for Portugal gathered by Pontes (2025) shows that the average college students' commuting distance first decreased between 1981 and 2001 before it eventually increased between 2001 and 2021, when it eventually became higher than the initial 1981 level. Proposition 6 does not account for this initial temporary decline in average commuting distance, a gap which we believe to follow from our extreme assumption that travel costs across regions are *infinite*.

Commuting costs across regions might be considerably higher than internal travel costs, but they are finite. This difference does not matter much when we consider sparsely populated regions because – as we show in Proposition 4 - in those territories the wage premium of education $\Delta(t)$ is low. Consequently, in this instance the wage premium of education likely fails to cover the high interregional travel cost.

By contrast, for high density regions and during the early stages of the schooling process, wage premia of education are quite high, so that students living in low density areas might find advantageous to travel to core regions to attend college despite the high commuting costs. Our assumption of infinite interregional commuting costs rules out this commuting pattern with the consequence that we probably underestimate the average students' commuting distance in the early stages of the schooling process.

# 6. Policy recommendations

We now try to design a policy aimed at expanding tertiary education while having a definite impact on the level of concentration of the university system. Our advice is condensed in the following Proposition.

**Proposition 7:** The government can raise the schooling rate $h(t^*)$ by diminishing either the fixed cost $F$ , or the unit travel cost $\tau$ .If the policy aims to centralize the college network, then it should reduce $\tau$ . By contrast, if the public authority purports to decentralize the higher education system, then the appropriate policy is to decrease $F$ .

**Proof:** We compute the partial derivatives of $g(.)$ from (25) in relation to $F$ and $\tau$ .

$$\frac{\partial g(.)}{\partial F} = -\frac{3}{4}\sqrt{\frac{\tau}{L(t)F}} < 0 \tag{60}$$

and,

$$\frac{\partial g(.)}{\partial \tau} = -\frac{3}{4}\sqrt{\frac{F}{L(t)\tau}} < 0 \tag{61}$$

Let us assume that the Government purports to increase the schooling rate $h(t^*)$ , which was defined in (32) with $t = t^*$ . Reducing either $F$ or $\tau$ shifts the $g(t)$ curve upwards thereby raising the time/region threshold from $t_1^*$ to $t_2^*$ in Figure 4.

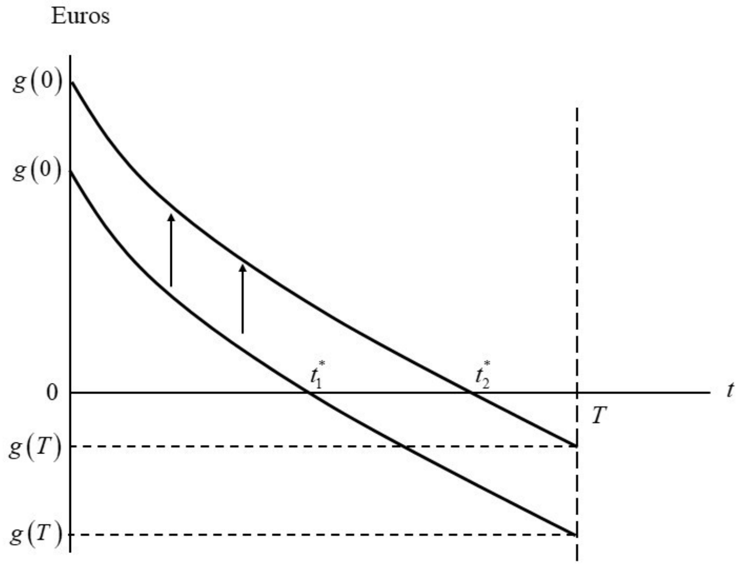


Figure 4: Increasing the schooling rate by diminishing either fixed or transport costs

We recall from (59) with $t = t^*$ that the average density of schools in the economy strictly decreases with threshold $t^*$. Hence by shifting $t^*$ to the right in Figure 4 a decrease in either $F$ or $\tau$ exerts an *indirect negative* effect on the average density of schools $\tilde{n}\left(t^*\right)$.

However, in addition to this *negative indirect effect,* reducing $\tau$ has also a *direct negative effect* on the school density in each region $n(t)$ which is clear from (54). Consequently, diminishing travel costs has an *unambiguous negative effect* on the average school density in the economy.

By contrast, from (54) reducing $F$ has a *direct positive effect* on the school density in each region $n(t)$. Consequently, despite reducing $F$ helps the development of formal education, its impact on the concentration level of the college system is unclear.

A similar demonstration can be made for the situation where the Government purports to expand literacy by decentralizing the college network. In this case, reducing the fixed cost is the appropriate policy tool. **QED**

Since the quality of a college is correlated with the size of fixed costs, we may state the following corollary.

**Corollary:** If a country expands literacy by decentralizing the school network, then this will be achieved at the price of decreasing the quality of education. By contrast, if the expansion of higher education results from an enhancement of students' commuting, then it is compatible with the maintenance of average quality.

# 7. Concluding remarks

We try to rationalize the process of higher education development in a multiregional economy. The model accounts for the fast growth in the college system, albeit at a decreasing rate and it further explains why per capita income also experiments a similar evolution. In addition, it also provides a reasonable explanation for the apparent paradox that the geographical decentralization of universities seems to eventually lead to a rise in the average students' commuting distance.

Our approach to the schooling process was based on the strong assumption that travel costs across regions are prohibitive, so that young people are constrained to attend college in the same region where they live, or not to attend at all. While this assumption seems to be not too strict during the later stages of education spread, it appears unrealistic in the earlier stages when the wage premium of education is high. To relax this assumption, both at the empirical and the theoretical level is a challenge that is left for future research.

# References


Boserup, E. (1965). The conditions of agricultural growth. The economics of agrarian change under population pressure. London, George Allen & Unwin.

Boucekkine, R., De la Croix, D., & Peeters, D. (2007). Early literacy achievements, population density, and the transition to modern growth. *Journal of the European Economic Association*, *5*(1), 183-226.

Diamond, P. A. (1982). Aggregate demand management in search equilibrium. *Journal of political Economy*, *90*(5), 881-894.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics* (Vol. 4, pp. 2063-2117). Elsevier.

Frenette, M. (2006). Too far to go on? Distance to school and university participation. *Education Economics*, *14*(1), 31-58.

Galor, O., & Weil, D. N. (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American economic review*, *90*(4), 806-828.

Glover, D. R., & Simon, J. L. (1975). The effect of population density on infrastructure: the case of road building. *Economic Development and Cultural Change*, *23*(3), 453-468.

Gordon, R. J. (2018). Declining American economic growth despite ongoing innovation. *Explorations in Economic History*, *69*(C), 1-12.

Pontes, J. P. (2024). Asymmetries in education spread: “Equilibrium” versus “Social Optimum”.

Pontes, J. P. (2025). Education spread and economic development: Portugal 1981–2021. *The Annals of Regional Science*, *74*(2), 1-19.

Pontes, J. P., & Buhse, A. P. (2019). The development of higher education in Europe as a 'coordination game'.

Salop, S. C. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, 141-156.

Schultz, T. W. (1964). Transforming traditional agriculture. New Haven: Yale University Press.

Spiess, C. K., & Wrohlich, K. (2010). Does distance determine who attends a university in Germany? *Economics of Education Review*, *29*(3), 470-479.

Tamura, R. (2002). Human capital and the switch from agriculture to industry. *Journal of economic Dynamics and Control*, *27*(2), 207-242.