Centre for Studies in Economics and Finance

WORKING PAPER NO. 599

Human Capital Distribution and the Transition from Stagnation to Growth

Mario F. Carillo

January 2021

University of Naples Federico II
University of Salerno
Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance
DEPARTMENT OF ECONOMICS – UNIVERSITY OF NAPLES
80126 NAPLES - ITALY
Tel. and fax +39 081 675372 – e-mail: csef@unina.it
ISSN: 2240-9696

WORKING PAPER NO. 599

Human Capital Distribution and the Transition from Stagnation to Growth

Mario F. Carillo*

Abstract
This research argues that differences in the distribution of human capital across countries and their impact on the advancement and the adoption of technology contributed to the differential timing of the transition from the Malthusian stagnation to modern growth and the persistent differences in income per capita across the globe. Polarization in the distribution of human capital within an economy implied a trade-off between innovation and adoption of technologies that determined the transition from stagnation to growth. Despite the contribution of the upper tail of the human capital distribution to technological innovation, the absence of wide group of educated individuals among the working population delayed technology adoption and the transition from stagnation to growth.

Keywords: Economic Growth; Human Capital Distribution; Demographic Transition; Long-run Development.

JEL Classification: I24, J13, J24, O30, O40.

Acknowledgements: I thank Oded Galor for insightful comments. I also benefited from conversations with Veronique Gille, Lothar Grall, Tullio Jappelli, Mark Klemp, Anastasia Litina, Vincenzo Lombardo, and Giulia Santangelo. I also thank seminar participants at Brown University, UNISA, the Ascea Summer School in Economic Growth, the Strasbourg ACDD Conference, and the Capri Summer School in Economic Growth for helpful comments.

* Università di Napoli Federico II and CSEF. Email: mariofrancesco.carillo@unina.it.

Table of contents

1 Introduction

Throughout most of human existence, societies around the globe exhibited living standards at subsistence level and limited economic growth. In the last two centuries, several regions of the world witnessed a transition to an era of unprecedented growth in living standards. Differences in the timing of this transition from the so-called Malthusian stagnation to modern growth contributed significantly to the disparity in income per capita that we observe across societies today (Galor, 2011). Understanding the determinants of the timing of this transition is crucial to deepen our understating of the development process.

Several studies have emphasized the crucial role of human capital in explaining the transition from stagnation to growth. The joint evolution of technological progress and human capital has been shown to influence investment in the quality of children (Galor and Weil, 1999, 2000; Carillo, 2020), life expectancy (Cervellati and Sunde, 2005), and the adoption of growth-enhancing institutions (Doepke and Zilibotti, 2005). While the upper tail of the human capital distribution is a key determinant of the innovation process that led to the Industrial Revolution (Mokyr, 2002, 2005; Jacob, 2014), a large number of studies suggest that a sufficiently educated labor force was important for the adoption of innovations in the production process (Griliches, 1957; Nelson and Phelps, 1966; Benhabib and Spiegel, 2005b), and more generally for economic growth (Becker et al., 2011; Madsen and Murtin, 2017). While previous works mainly emphasize the effects of the upper tail of the human capital distribution and mass education in isolation, their joint evolution, and effects on the long-term trajectory of the economy are not fully understood.

This paper advances the hypothesis that differences in the distribution of human capital across societies and their impact on the advancement and adoption of technology contributed to the differential timing of the transition from the pre-industrial stagnation to modern growth and long-term development. Polarization in the distribution of human capital within an economy (i.e. fat tails at the two ends of the human capital distribution) had two implications for long-run growth. First, it implied a trade-off between technological innovation and adoption. Despite the contribution of the upper tail of the human capital distribution to technological innovation, the absence of a wide group of educated individuals among the working population delayed technological adoption and the transition from stagnation to growth. Second, it implied a trade-off between the timing of the transition and long-run growth. By harnessing adoption, low polarization also delayed the transition out of the Malthusian stagnation, during which improvements in living standards are mainly devoted to the number of children, fueling population growth for a longer period. The resulting large population at the time of the transition, and the associated large number of highly educated individuals fostered innovations, which could be adopted by the educated labor force, ultimately enhancing long-run economic growth and convergence to early-takeoff economies.

In the model presented in this paper, technology adoption has two main purposes. First, innovation can have a direct effect on production improvements only if it is adopted. An emblematic example is the precursor of the steam engine, developed by the Italian scholar Giambattista Della Porta (1535? - 1615), who used steam power to pump water already in 1606. The first commercially successful engine did not appear until around 1712 (Brown, 2002, pp. 60). It was invented by Thomas Newcomen and paved the way for the Industrial Revolution (Stuart, 1829). Yet, his first application was to pump water as well (Rolt, 1963). Thus, absent the appropriate conditions for technology adoption, even the most powerful innovation in history may not trigger a technological breakthrough.

Second, adoption serves as a mechanism to store, make accessible, and transmit knowledge, which ultimately can be improved over time. This channel was especially relevant in pre-industrial ages, when a common scientific language was not yet developed, generating difficulties in storing and accessing information. For instance, the first scientific journal — the *Royal Society of London* — appeared in 1662 (Bekar and Lipsey, 2004). Only at that time, innovations were stored and made accessible to the scientific community. Before then, they were transmitted across generations mainly when they were adopted and thus transformed from theoretical projects into material "gadgets" (Ashton, 1955). These could be stored and understood despite the lack of a common scientific language. They could also be improved over generations (Bekar and Lipsey, 2004) based on a process of "tinkering" (Mokyr, 1990) that paved the way for a basic mechanization experimentation based (Musson and Robinson, 1989). This indirect effect of adoption on technological advancement is primarily conducted by artisans skilled in tinkering and producing gadgets, which supplied a combination of intellectual and manual labor and thus composed the mass in the middle of the human capital distribution.

Therefore, low polarization in human capital and the associated wide share of the labor force able to understand and employ new ideas facilitated their adoption that, through the creation of gadgets materially representing these ideas could be improved over time, with feedback effects on the accumulation of knowledge and technological progress. The interplay of this feedback effect together with the direct effect on production implies that studying the distribution of human capital is a crucial element in the analysis of the transition from stagnation to growth and for our understanding the disparities in living standards across the globe today.

This paper mainly contributes to three strands of the literature. First, it may reconcile seemingly contrasting findings on persistence vis-à-vis reversals in economic performance across the globe. While a large literature has documented persistence in economic development and technology adoption over thousands of years (Comin et al., 2010; Putterman and Weil, 2010; Chanda et al., 2014), others have emphasized reversals in the process of development (Acemoglu et al., 2002). Factors such as geography and historical institutions are important elements to explain these findings. The present paper complements this literature by mapping these historical conditions to differences in the prevalence of an educated labor force and an educated elite. In turn, providing a novel mechanism of how these differences in historical factors may influence long-run growth.

Second, the paper bridges the long-run growth literature with the literature emphasizing technology adoption as a byproduct of specific skills and human capital levels. Several studies have hypothesized that human capital and technology are complementary (Griliches, 1957; Nelson and Phelps, 1966). This view induced the emergence of a body of empirical studies in cross-country settings (Benhabib and Spiegel, 2005a) as well as across localities within countries (Foster and Rosenzweig, 1996; Carillo, 2020) pointing at human capital as an essential precondition for technology adoption. By investigating the interlink between the distribution of human capital and technology adoption, this paper links this literature to the body of works studying the transition from stagnation to growth.

Third, this paper employs a tractable approach that introduces different skill levels in explaining the transition to modern growth a Unified Growth Theory framework. Thus, it extends the literature studying factors that influenced the transition from stagnation to growth, including human capital (Galor and Weil, 1999, 2000), life expectancy (Cervellati and Sunde, 2005), physical development (Croix and Licandro, 2013), and fertility choices (Strulik and Weisdorf, 2008).

The paper is organized as follows. Section 2 presents the theoretical model, which includes the human capital polarization and technological adoption. Section 3 explores the link between human capital polarization, the timing of the transition out of the Malthusian stagnation, and long-run growth. Section 4 concludes.

2 The Model

Consider an overlapping generation model that evolves over infinite discrete time. Every period t, a finite homogeneous good, Y_t, can be produced according to two alternatives regimes of production, defined respectively as *old regime* and *new regime*. To model in a tractable fashion the distribution of human capital in the population, I employ three levels of skills.[1] Factors of production are three sources of labor force reflecting three levels of human capital of workers.[2] The three sources of labor force are: manual labor, human capital intensive labor and a combination between the two.

Manual labor, L_t, reflects the amount of labor supplied by individuals characterized by the lowest level of human capital. Human capital-intensive labor, H_t, reflects the amount of labor supplied by highly educated labor force, that is relatively more likely to innovate and thus are called *innovators*. A third source of labor force, which represents the middle of the

[1] The use of a discrete distribution is in line with the indivisibility of skills and degrees.

[2] Considering also land as a constant factor of production with the absence of property rights (Galor and Moav, 2002), the results would not change qualitatively. Considering capital as a factor of production would complicate the model to the point of intractability.

human capital distribution, is given by a combination between manual labor and intellectual labor, M_t, which is supplied by individuals that are enough educated to understand and adopt new innovations, thus they are called *adopters*.

2.1 Production

Production may take place according to two alternative production regimes, defined as *old regime* and *new regime*.

2.1.1 The Old Regime

The *old regime* of production is such that only manual labor and highly educated labor are employed.[3]

$$Y_t^o = A_t^o H_t^\alpha L_t^{1-\alpha} = A_t^o L_t h_t^\alpha \tag{1}$$

where h_t is the proportion of highly educated over manual labor force, given by

$$h_t \equiv \frac{H_t}{L_t} \tag{2}$$

In other words, in early stages of technological development, new inventions, which were mainly produced by high-skilled labor force, can be directly adopted by the manual labor force. The idea is that when the level of knowledge is low, new innovations are not too complex to adopt and thus understanding of the practical functioning is possible even for low-skilled labor. In particular, even though their theoretical foundation could be unknown to the masses, the simplicity of early technology is reflected in simplicity of its adoption, implying that there is no necessity of specific forms of education or skills to adopt them in production.

Examples of technologies of this sort include ley farming: an agricultural system where land is alternately seeded for grain and left fallow. During the fallow period the soil is filled with roots of grasses and other plants. While sophisticated chemical skills were needed in order to invent

[3]Following Aghion and Howitt (1992), innovators may be considered as monopolists of an intermediate sector of innovations. Despite the complexity of such alternative approach, the result would be identical.

ley farming, which was introduced in one of the most advanced forms at the beginning of the seventeenth century (Stapledon et al., 1948), it could be easily understood and adopted by farmers. They would simply need to know which grasses to cultivate during the fallow period to ultimately increase agricultural yields.

Highly educated individuals are modeled as directly contributing to production through the creation of innovations. Moreover, as discussed in Section 2.3, they also indirectly contribute to knowledge accumulation. Thus, in early stages of technological development, the creation of innovations required human capital, but their adoption did not. However, manual laborers' limited understanding of the new techniques may have harnessed the improvement of these newly adopted techniques. This additional element will be discussed in Section 2.3.

2.1.2 The New Regime

Production in the *new regime* takes place according to a production function that includes manual labor, innovators' labor, and adopters' labor force.

$$Y_t^n = A_t^n H_t^\beta M_t^\phi L_t^{1-\beta-\phi} = A_t^n L_t h_t^\beta m_t^\phi \tag{3}$$

where m_t is the proportion of adopters over manual labor force,

$$m_t \equiv \frac{M_t}{L_t} \; ; \; h_t \equiv \frac{H_t}{L_t} \tag{4}$$

The rationale is that, when technology is more complex, innovations are mainly adopted by individuals sufficiently educated to understand and employ them in the production process. This is the case of most advanced technical innovations, whose comprehension and adoption typically requires a basic level of skills. Artisans and craftsmen, for instance, had specific training aimed at employing gadgets. This sort of labor force, supplying a combination of manual and intellectual labor, composed the mass in the middle of the human capital distribution.

2.1.3 Factor Prices

Markets are perfectly competitive, the inverse demands for factors of production depend on the regime employed. The inverse demand for highly skilled labor, given (1) and (3),

$$w_t^h = \begin{cases} \alpha A_t^o h_t^{\alpha-1} & \text{if } Y_t^o > 0 \\ \beta A_t^n h_t^{\beta-1} m_t^{\phi} & \text{if } Y_t^n > 0 \end{cases} \tag{5}$$

where w_t^h is the wage of innovators. The inverse demand for manual labor, given (1) and (3), is

$$w_t^l = \begin{cases} (1-\alpha) A_t^o h_t^{\alpha} & \text{if } Y_t^o > 0 \\ (1-\beta-\phi) A_t^n h_t^{\beta} m_t^{\phi} & \text{if } Y_t^n > 0 \end{cases} \tag{6}$$

where w_t^l is the wage of unskilled labor. The inverse demand for adopters' labor, given (3), is

$$w_t^m = \phi A_t^n h_t^{\beta} m_t^{\phi-1} \quad \text{if } Y_t^n > 0 \tag{7}$$

where w_t^m is the wage of adopters, that will be employed only in new regime. Moreover, given (5) and (6), the wage ratio of the innovators over manual labor is

$$\frac{w_t^h}{w_t^l} = \begin{cases} \frac{\alpha}{1-\alpha}\frac{1}{h_t} & \equiv \omega\left(h_t^o\right) & \text{if } Y_t^o > 0 \\ \left(\frac{\beta}{1-\beta-\phi}\right)\frac{1}{h_t} & \equiv \omega\left(h_t^n\right) & \text{if } Y_t^n > 0 \end{cases} \tag{8}$$

Given (7) and (6) the wage ratio of the adopters over manual labor is given by

$$\frac{w_t^m}{w_t^l} = \left(\frac{\phi}{1-\beta-\phi}\right)\frac{1}{m_t} \equiv \omega^m\left(m_t\right) \text{ if } Y_t^n > 0 \tag{9}$$

From the properties of the production functions, it follows that wage ratios are characterized by the following properties:

$\omega^{'}(j_t) < 0$, $\lim_{j_t \to 0} \omega^j(j_t) \to \infty$, $\lim_{j_t \to \infty} \omega^j(j_t) \to 0$ with $j = m, h$ and $\forall j_t \in \left[0, \infty\right)$.

2.2 The Individual Choice

Consider an economy in which individuals live for two periods of time: childhood and parenthood. During the first period of life they consume a fraction of parental endowment that consists in one unit of time. All decisions are made in the adult period of life. Parents are endowed with one unit of time as manual labor, l, adopters labor, m, or innovators labor, h, depending on the level of education they received during childhood. Such endowment is allocated between children rearing and consumption.

2.2.1 Preferences and Budget Constraints

Preferences are defined over parental consumption and the potential aggregate income of their children (Galor and Mountford, 2008). Parents i, where $i = l, m, h$, choose the number of children $n^{i,j}$ for each level of education j, with $j = l, m, h$, and parental utility from each child depends on the wage she gets on the market. In other terms, parents get their own utility according to the utility function

$$u_t^i = (1-\gamma)lnc_t^i + \gamma ln\big(w_{t+1}^l n_t^{i,l} + w_{t+1}^m n_t^{i,m} + w_{t+1}^h n_t^{i,h}\big) \tag{10}$$

where c_t^i is parental consumption at time t, $n_t^{i,j}$ is the number of children of type j reared by parent i at time t.[4]

The budget constraint is given by

$$c_t^i + w_t^i\big(n_t^{i,l}\tau^l + n_t^{i,m}\tau^m + n_t^{i,h}\tau^h\big) \leq w_t^i \tag{11}$$

2.2.2 Optimization

$$\{c_t^i, n_t^{i,l}, n_t^{i,m}, n_t^{i,h}\} = argmax\Big[(1-\gamma)lnc_t^i + \gamma ln\big(w_{t+1}^l n_t^{i,l} + w_{t+1}^m n_t^{i,m} + w_{t+1}^h n_t^{i,h}\big)\Big] \tag{12}$$

subject to

$$c_t^i + w_t^i\big(n_t^{i,l}\tau^l + n_t^{i,m}\tau^m + n_t^{i,h}\tau^h\big) \leq w_t^i \tag{13}$$

$$c_t^i \geq \tilde{c} \tag{14}$$

where $\tau^l < \tau^m < \tau^h$. In particular τ^j is the cost of having a child of type j with $j = l, m, h$, therefore the higher the level of human capital of the offspring the higher the cost of producing a child with that particular level of education[5].

[4]Notice that, since mortality is not explicitly modeled, n_t can be interpreted as the number of surviving children.

[5]Alternatively, one may argue that, despite the lower level of human capital intrinsic in

The optimal level of consumption is given by,

$$c_t^i = \begin{cases} \tilde{c} & \text{if } (1-\gamma)w_t^i < \tilde{c} \\ (1-\gamma) & \text{if } (1-\gamma)w_t^i \geq \tilde{c} \end{cases} \tag{15}$$

The amount of time invested in child rearing is given by,

$$n_t^{i,l}\tau^l + n_t^{i,m}\tau^m + n_t^{i,h}\tau^h = \begin{cases} \frac{w_t^i - \tilde{c}}{w_t^i} & \text{if } (1-\gamma)w_t^i < \tilde{c} \\ \gamma & \text{if } (1-\gamma)w_t^i \geq \tilde{c} \end{cases} \tag{16}$$

During the old regime of production,

$$\begin{array}{ll} n_t^{i,h} = 0 & \text{if } w_t^h/w_t^l < \tau^h/\tau^l \\ n_t^{i,h} > 0 \text{ and } n^{i,l} > 0 & \text{only if } w_t^h/w_t^l = \tau^h/\tau^l \\ n_t^{i,l} = 0 & \text{if } w_t^h/w_t^l > \tau^h/\tau^l \end{array} \tag{17}$$

which means that if the wages ratio equals the cost ratio all the types available in the old regime will exist.

Equivalently, during the new regime of production,

$$\begin{array}{rlll} n_t^{i,h} = 0 & \textit{if } w_t^h/w_t^l < \tau^h/\tau^l & \textit{or } w_t^h/w_t^m < \tau^h/\tau^m \\ n_t^{i,m} = 0 & \textit{if } w_t^m/w_t^l < \tau^m/\tau^l & \textit{or } w_t^m/w_t^h < \tau^m/\tau^h \\ n_t^{i,l} = 0 & \textit{if } w_t^h/w_t^l < \tau^h/\tau^l & \textit{or } w_t^h/w_t^m < \tau^h/\tau^m \\ (n_t^{i,h}, n_t^{i,m}, n_t^{i,l}) \gg 0 & \textit{only if } w_t^h/w_t^l = \tau^h/\tau^l & \textit{and } w_t^m/w_t^l = \tau^m/\tau^l \end{array} \tag{18}$$

Lemma 1 *Consider the old regime of production. There exists a unique ratio of innovators to manual labor ratio, $(h^o)^*$ such that*

$$\frac{w_t^{o,h}}{w_t^{o,l}} = \omega((h^o)^*) = \frac{\tau^h}{\tau^l} \tag{19}$$

artisans' skills with respect to philosophers or mathematicians, the scarcity of certain skills, such as the carpenter or armorer ones, entails difficulties in acquiring them with the consequence of higher costs. However, most artisans' skills were acquired through job training or, in more advanced stages of urbanization, through apprenticeship under the supervision of masters. Both these approaches of acquiring this source of human capital are characterized by a higher degree of economies of scale with respect to the acquisition of high level human capital, in turn implying a lower cost relative to other forms of human capital.

where,

$$
\begin{aligned}
n_t^{i,l} &= 0 \quad \text{if } h_t < (h_t^o)^* \\
n_t^{i,h} &= 0 \quad \text{if } h_t > (h_t^o)^*
\end{aligned}
\tag{20}
$$

Proof. The uniqueness of $(h_t^o)^*$ follows from the properties of $\omega((h_t^o)^*)$. The remaining part is a corollary of (17). □

Hence, during the old regime, if $h_t < (h_t^o)^*$ the relative reward for having uneducated offspring is low with respect to the relative cost and thus there are no incentives to raise them, implying an increase in h_t. Whereas, if $h_t > (h_t^o)^*$ there are no incentives to raise high-human-capital offspring, implying a decrease in h_t up to the equilibrium proportion,$(h_t^o)^*$.

Corollary 1 *If the old regime of production is employed then $h_t = (h_t^o)^*$, that is,*

$$h_t = (h_t^o)^* \quad \text{if } Y_t^o > 0 \tag{21}$$

and therefore wages for innovators are

$$w_t^h = \alpha A_t^o \left[(h_t^o)^*\right]^{\alpha-1} \quad \text{if } Y_t^o > 0 \tag{22}$$

wages for manual labor are

$$w_t^l = (1-\alpha) A_t^o \left[(h_t^o)^*\right]^{\alpha} \quad \text{if } Y_t^o > 0 \tag{23}$$

and thus

$$(h_t^o)^* = \left(\frac{\alpha}{1-\alpha}\right)\frac{\tau^l}{\tau^h} \tag{24}$$

where the latter comes from (8), given Lemma 1.

Importantly, notice from (24) that during the old regime the optimal proportion of innovators over manual labor force is constant over time, that is,

$$(h_t^o)^* = (h^o)^* \ \forall t \tag{25}$$

Lemma 2 *Consider the new regime. There exists a unique innovators to manual labor ratio, $(h^n)^*$, and a unique adopters to manual labor ratio, m^*, such that*

$$\frac{w_t^{n,h}}{w_t^{n,l}} = \omega((h_t^n)^*) = \frac{\tau^h}{\tau^l} \tag{26}$$

$$\frac{w_t^m}{w_t^{n,l}} = \omega(m_t^*) = \frac{\tau^m}{\tau^l} \tag{27}$$

$$\begin{aligned} n_t^{i,h} &= 0 \quad \text{if } h_t > (h^n)^* \quad \text{or } m_t < m_t^* \\ n_t^{i,m} &= 0 \quad \text{if } m_t > m_t^* \quad \text{or } h_t < (h^n)^* \\ n_t^{i,l} &= 0 \quad \text{if } h_t < (h^n)^* \quad \text{or } m_t < m_t^* \end{aligned} \tag{28}$$

Proof. The uniqueness of $(h_t^n)^*$ and m_t^* follows from the properties of $\omega((h_t^n)^*)$ and $\omega(m_t^*)$ respectively. The remaining part is a corollary of (18). □

Hence, during the new regime, if $h_t > (h_t^n)^*$ there are no incentives to raise highly-educated children, implying a reduction in h_t. However, also in the case in which $m_t < m_t^*$ there are no incentives to raise neither highly educated children nor uneducated children because the relative reward of raising offspring with an intermediate level of education is higher, therefore resources will move in this direction, increasing m_t.

In other words, if the proportion of one of the three source of labor force is lower than optimal, the potential relative wage of that child is higher than the relative cost, inducing parents to invest their resources in rearing offspring with that particular level of education, ultimately increasing their relative proportion and reducing their relative wage until equations (26) and (27) are satisfied.

Corollary 2 *If the new regime of production is employed then $h_t = (h_t^n)^*$ and $m_t = m_t^*$, that is,*

$$h_t = (h_t^n)^* \text{ and } m_t = m_t^* \quad \text{if } Y_t^n > 0 \tag{29}$$

and therefore wages for innovators are

$$w_t^h = \beta A_t^n \left[(h^n)^*\right]^{\beta-1} \left[m^*\right]^{\phi} \quad \text{if } Y_t^n > 0 \tag{30}$$

wages for adopters are

$$w_t^m = \phi A_t^n \left[(h^n)^*\right]^{\beta} \left[m^*\right]^{\phi-1} \quad \text{if } Y_t^n > 0 \tag{31}$$

wages for manual labor are

$$w_t^l = (1-\beta-\phi) A_t^n \left[(h^n)^*\right]^{\beta} \left[m^*\right]^{\phi} \quad \text{if } Y_t^n > 0 \tag{32}$$

and thus

$$(h_t^n)^* = \left(\frac{\beta}{1-\beta-\phi}\right)\frac{\tau^l}{\tau^h} \tag{33}$$

$$m_t^* = \left(\frac{\phi}{1-\beta-\phi}\right)\frac{\tau^l}{\tau^m} \tag{34}$$

where (33) *and* (34) *come from* (8) *and* (9) *, given Lemma 2.* Notice from (33) and (34) that during the new regime the optimal ratios of innovators to manual labor and adopters to manual labor are constant over time, that is,

$$(h_t^n)^* = (h^n)^* \text{ and } m_t^* = m^*; \ \forall t \tag{35}$$

Furthermore from (19) and (26),

$$\omega((h^o)^*) = \omega((h^n)^*) \tag{36}$$

that implies

$$\left(\frac{\alpha}{1-\alpha}\right)(h^n)^* = \left(\frac{\beta}{1-\beta-\phi}\right)(h^o)^* \tag{37}$$

Notice that, plausibly assuming that production in the old regime is manual-labor-intensive ($(1-\alpha) < (1-\beta-\phi)$), the prevalence of innovators with respect to manual labor force is higher in the new regime, that is,

$$(h^n)^* > (h^o)^* \tag{38}$$

2.3 Technological Progress

Suppose that, during the old regime, technology formation between time t and $t+1$, depends on the number of innovators in the economy at time t:

$$\frac{A^o_{t+1} - A^o_t}{A^o_t} = \Omega(H_t) \tag{39}$$

where A^o_0 is historically given and the innovation function $\Omega(H_t)$ is an increasing and concave function $\Omega' > 0$ and $\Omega'' < 0$ and $\Omega \in (0, \infty)$.

That is, technological progress during early stages of development depends on the population of innovators in the economy. Note that $\Omega(0) > 0$, that is, in the absence of innovators, there would still be technological advancements.

During the new regime of production, in addition to the effect of innovators, the presence of a share of the labor force that can adopt innovations and improve them over time is an additional source of technological accumulation

$$\frac{A^n_{t+1} - A^n_t}{A^n_t} = \Omega(H_t)\ \big(1 + \lambda(M_t/N_t)\big) \tag{40}$$

where A^n_0 is historically given and the adoption rate $\lambda' > 0$, $\lambda'' < 0$ and $\lambda \in (0, \infty)$ with $\lambda(0) > 0$. The adoption rate λ depends on the fraction of adopters in the economy, $\frac{M_t}{N_t}$, the higher the fraction of adopters the closer the economy is to perfect adoption. Note that technological progress is faster in the new regime.[6]

During the old regime, despite the fact that the new regime of production is not operative, knowledge advancement permits the *potential* productivity of the new regime to grow over time. That is, when the new regime is not efficient, adopters are not employed in the production process, and thus not rewarded on the market, however there is a latent technology advancement due to those workers that, throughout a process of tinkering and learning by doing in laboratories rather than on the job, acquire those skills that are necessary to the process of adoption and make the new regime more efficient.

[6] This is ensured by the assumption that even if the share of adopters is null, the adoption rate is assumed to be positive (i.e. $\lambda(0) > 0$).

2.4 Viability of production regimes

The two regimes are available at each point in time, thus each agent chooses the preferred regime depending on the reward he can get. In other terms,

$$\text{Every agent } i \text{ chooses the regime } j \text{ if } w_t^{j,i} \geq w_t^{-j,i} \forall i = l, m, h; \forall t \tag{41}$$

Lemma 3 *At each point in time, only one regime of production is operative.*

Proof. It comes from condition (41) given Lemma 1 and Lemma 2. □

Considering that each agent chooses the preferred regime depending on the reward he can get, lemmas 1 and 2 imply that, during the old and the new regime respectively, the proportions of factors of production in the economy are constant over time (see (24), (33) and (34)). Thus, factors of production are not free to adjust up to an equilibrium wage that permits the coexistence of the two regimes. Conversely, in each regime wages are given (but not constant: they depend on the level of technology A_t^j, *for* $j = old, new$) and at each point in time agents compare such wages determining the operative regime.

Lemma 4 *The new regime is economically viable if*[7]

$$w_{t+1}^{n,l} \geq w_{t+1}^{o,l} \tag{42}$$

where $w_{t+1}^{n,l}$ *is the wage that uneducated children at time* t *will get at time* $t+1$ *in regime* $j = old, new$

Proof. It comes from condition (41) given (19), (26) and (27). □

Since the wage ratios are constant over time, if the new regime is economically viable for the manual laborers it will be economically viable for all agents in the economy. During the old regime adopters are not rewarded, therefore parents will choose to invest in that particular level of education only when the new regime will be operative. That is, only when children

[7]It is assumed that, in the case in which wages for a specific source of labor force are equal in both regimes, the new regime is preferred. However, as will be clear in the following, this equality can persist only for one period of time.

that are educated at high level (innovators) or at low level (manual labor) will get a higher wage in that regime. Therefore, the difference between the wage that manual workers can earn in the new regime and the one available in the old regime represents a threshold rule for the transition to the new regime of production.

2.5 The Time Path of the Economy

2.5.1 Technological Progress

The productivity parameters are restricted so that the new regime is not economically viable in period 0, that is,

$$\frac{A_0^o}{A_0^n} > \frac{1-\beta-\phi}{1-\alpha} \frac{[(h^n)^*]^\beta [m^*]^\phi}{[(h^o)^*]^\alpha} \tag{43}$$

Lemma 5 It exists a time t^* such that the new regime is viable, that is,

$$\exists t^* | \forall t \geq t^*, w_t^{n,l} \geq w_t^{o,l} \tag{44}$$

Proof. It comes from lemma 4, (32) *and* (23)*, given* (39) *and* (40) □

Since the productivity of the new regime grows faster[8] and given that the unique source of time variation of wages is due to total productivity growth, there exists a point in time, t^*, in which the new regime yields a level of wages higher than the old one, leading to the transition to the new regime. This also means that, regardless of where the economy is positioned in the h_t, m_t plane (see Figure 1) at some point it will experience the transition from stagnation to growth.

Lemma 6 It exists a time t^c such that the Malthusian constraint is no longer binding, that is,

$$\exists t^c | \forall t \geq t^c, w_t^{o,l} \geq \tilde{c}/(1-\gamma) \tag{45}$$

[8]Although the new regime of production is not employed, knowledge advancements imply improvements in the potential technology. Innovations stimulate productivity of the old regime as well as knowledge advancements that employ adopters' skills through, for instance, learning by doing rather than market returns. See, for instance, Galor and Mountford (2008).

Proof. It comes from equation (23) *, given* (39) □

Equilibrium wages increase of time due to technology advancement, therefore necessarily exists a point in time at which the subsistence constraint is no longer binding.

The new regime of production implies a level of knowledge sufficiently advanced to consolidate an educated labor force, which can be achieved when the subsistence constraint is no longer binding. Thus, it is assumed that $t^* = t^c$, which is equivalent to assume that the wage level w_{t^*} such that $w_{t^*} \equiv w_t^{n,l} = w_t^{o,l}$ is such that $w_{t^*} = \tilde{c}/(1-\gamma)$. The assumption implies that the Malthusian constraint will be binding during the old regime and the escape from stagnation is associated with the transition to the new regime of production. While unnecessary, this assumption simplifies the analysis and is broadly consistent with historical patterns.

2.5.2 The Timing of the Transition

Given the equilibrium quantities h^*; m^* and the threshold G_t, it possible to solve for the time at which an economy will experience the transition to the new regime, t^*. Where G_t is such that (42) is satisfied with equality, that is,

$$G_t = \left\{G(h_t, m_t)|w_t^{n,l} - w_t^{o,l} = 0\right\} \ \forall t \leq t^* \tag{46}$$

where $h_t \equiv h_t^n$.

Lemma 7 *The threshold G_t, before the transition, is a function of the proportions of factors of production in the new regime and parameters of the model, that is,*

$$G_t = w_t^{n,l} - w_t^{o,l} = G\big(h_t^n, m_t; A_t^o, A_t^n, \zeta\big) = 0 \ \forall t \leq t^* \tag{47}$$

where $\zeta = \zeta(\tau^l, \tau^h; \alpha, \beta, \phi)$

Proof. It comes from (23), (24) *and* (82) *given Lemma 5 noting that only the old regime is operative.* □

Thus, before the transition takes place, only the old regime is operative (see condition (43)). The threshold G_t, $\forall t < t^*$, represents the sets of points

(h_t, m_t), where $h_t^n \equiv h_t$, such that the new regime is viable, therefore, it can be represented on a h_t, m_t plane (see Figure 1). The new regime is viable when the equilibrium proportions (h^*, m^*) satisfies the threshold rule G_t. The timing of the transition can be measured considering the time elapsed between period 0 and the period in which the optimal proportions of factors of production belong to the threshold, $t^*|(h^*, m^*) \in G_t^*$. Figure 1 depicts the movement of the threshold, G_t, until the transition is experienced (i.e. $t \leq t^*$). Therefore, the timing of the transition, t^*, is a function of the distance from the equilibrium point (h^*, m^*) and the curve $G_{t=0}$. More specifically, time is given by the ratio between distance and speed, thus the timing of the transition from stagnation to growth is given by,

$$t^* = \frac{d^*}{s_{t^*}} \tag{48}$$

where, d^* is the minimum distance between (h^*, m^*) and $G_{t=0}$, s_{t^*} is the speed of convergence to the new regime[9]. Therefore, noticing (24), (33) and (34) the time of the transition can be expressed as follows,

$$t^* = t(\tau^l, \tau^m, \tau^h; \xi) \tag{49}$$

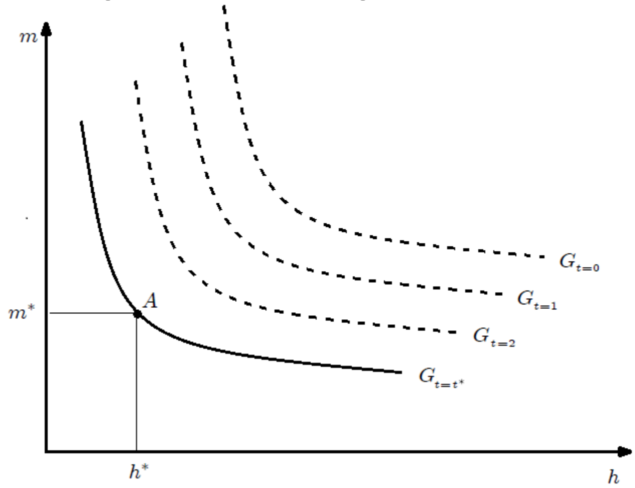
where $\xi \equiv \xi(\alpha, \beta, \phi, A_0^o, A_0^n, \lambda(0))$. The costs of raising offspring, τ^j, with $j = l, m, h$, determine polarization in distribution of human capital in the economy. Therefore, analyzing the effect of variations in such parameters on t^*, it is possible to investigate the effect of polarization in the human capital distribution on the timing of the transition from stagnation to growth.

3 Comparative Statics

Environmental and institutional differences may explain variations in the costs of raising children, influencing the human capital distribution and the long- evolution of economic performance. In the following, I explore the effect of the changes in the parameters and their effects on the transition to modern growth and on the growth rate of the economy.

[9]See Appendix for the specification of d^* and s_{t^*}

Figure 1: The Timing of the Transition



Notes: At t^* Economy A will experience the transition from the old to the new regime.

3.1 Timing of the Transition

Each point in the h, m plane depicted in Figure 2 represents a potential distribution of human capital once the transition occurs. An economy, depending on the cost parameters, will be located on one point that represents the optimal distribution of human capital. The distance from such point to the threshold $G_{t=0}$ is proportional to the timing of the transition from stagnation to growth. Variations in the costs of raising children with different levels of education — τ^l, τ^m and τ^h — are associated with changes in the distribution of human capital and the transition timing, t^*, which explains the link between polarization and the transition from stagnation to growth.

As depicted in Figure 2, an increase in the cost of raising highly educated children, τ^h, reduces the prevalence of innovators implying a shift from point A to point B which, in turn, corresponds to a larger distance AA' with respect to BB'. Thus economy B would need more time to achieve the transition with respect to economy A.[10] Low polarization in

[10]The effect of the speed function due to variations in τ^j, with $j = l, m, h$, exacerbates the effect of the change in the parameters on the distance function. Therefore, it is sufficient to analyze the effect on the distance to understand the direction of the overall effect on the timing on the transition.

the distribution of human capital implied a relatively low level of innovation in the economy, ultimately delaying the transition from stagnation to growth. Similarly, high cost of raising adopters, τ^m, lowers the share of adopters and shifts the economy from A to C, delaying the transition (i.e. $CC' > AA'$). Finally, high cost of raising uneducated offspring, τ^l, entails a larger prevalence of innovators and adopters, which accelerate the process of technological progress implying an early take-off.

Lemma 8

$$\frac{\partial t^*}{\partial \tau^l} < 0; \frac{\partial t^*}{\partial \tau^m} > 0; \frac{\partial t^*}{\partial \tau^h} > 0 \tag{50}$$

Proof. See Appendix. □

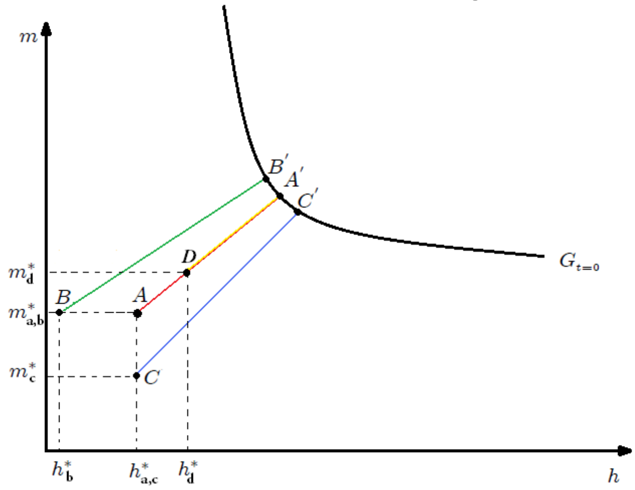
This cross-country comparison is linked to the assumption that economies considered in the analysis are isolated from a technological point of view. This condition is in line with the imperfect transmission of technology across societies in early stages of development. Thus, ni line with the advanced theory, technologies diffused mostly when innovations were adopted in the production process entailing the creation of gadgets that could be transmitted both across space and time.

The model predicts that, in early stages of development, polarization in the distribution of human capital implied a trade-off between innovation and adoption of innovation, determining the timing of the transition from stagnation to growth. The following section explores the effect of polarization on economic growth in the aftermath of the escape from the Malthusian stagnation.

3.2 Long-Run Growth

During the Malthusian stagnation, improvements in output are eroded by proportional increases in population. The model shows that increases in parental income induce larger fertility rates, and thus limited improvement in living standards. As a result, growth in output per capita is null during the old regime when the Malthusian constraint is binding. Then, at the time of the exodus from the Malthusian trap, the transition to a modern growth regime leads to a significant growth in output per capita,

Figure 2: Polarization and the Timing of the Transition



Notes: *Case 1* $[\tau^l_d > \tau^l_a]$ - In country D the level of human capital is higher with respect to A; *Case 2* $[\tau^h_b > \tau^h_a]$ - In country B polarization is lower with respect to A ; *Case 3* $[\tau^m_c > \tau^m_a]$ - In country C polarization is higher with respect to A.

which is generated by the interlink between human capital and technological progress.

The growth rate of output in the two regimes is given by[11],

Lemma 9

$$a)\ g^o_y = 0 \ \forall \ t < t^* \tag{51}$$

$$b)\ g^n_y = \Omega(H_t)\ \big(1 + \lambda(M_t/N_t)\big)\ \forall \ t \geq t^* \tag{52}$$

Proof. *See Appendix* □

Lemma 10

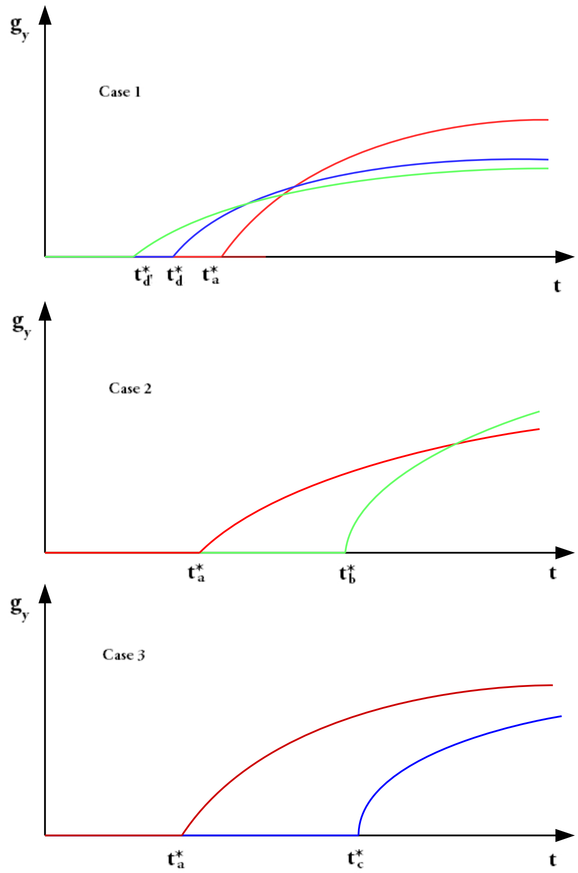
$$g^n_y = \Omega(n^{t-t^*})\chi\ \left(1 + \lambda(M_t/N_t)\right) \quad \forall \ t \geq t^* \tag{53}$$

Proof. *See Appendix* □

Equation (53) provides the link between polarization and the growth rate of output per capita. The results are illustrated in Figure 3. In the

[11]Where $g^j_y = \frac{Y^j_{t+1}/N^j_{t+1} - Y^j_t/N^j_t}{Y^j_t/N^j_t}$ and $j = old, new$.

Figure 3: The Long Run Effect of Polarization



Notes: *Case 1* $[\tau^l_{d'} > \tau^l_d > \tau^l_a]$ - In country D the level of human capital is higher with respect to Country A; *Case 2* $[\tau^h_b > \tau^h_a]$ - In country B polarization is lower with respect to Country A ; *Case 3* $[\tau^m_c > \tau^m_a]$ - In country C polarization is higher with respect to Country A.

first case, economy D has a higher cost of raising uneducated offspring with respect to A, $\tau_d^l > \tau_a^l$, therefore in D more educated children will be raised since their relative cost is lower. In turn, entailing a higher level of human capital in the economy and an earlier transition (as shown in Figure 3). However, increasing τ^l the benefits of the early transition are eroded by a small population. In turn, inducing a lower long-run growth in output per capita (Figure 3, case 1).

The second case is the effect of an increase in the cost raising highly educated children, τ^h, which implies lower polarization in B than in A (Figure 2) and a delayed transition, which is compensated by faster long-run growth (Figure 3, case 2). Low polarization (high τ^h) by delaying the timing of the transition also implies that the Malthusian mechanism lasts longer, in turn inducing a higher level of population once the transition is experienced. Subsequently, the larger number of innovators supported by a larger fraction of the population able to adopt such innovation, ultimately implied higher growth after the take-off from the Malthusian stagnation. Thus the model predicts a reversal of fortunes: the initial disadvantage in the timing of the transition induced by low polarization is compensated by an advantage in long-run growth rate after the transition occurs.

The last case is the case in which the cost of raising offspring enough educated to contribute to technology formation through the adoption of innovations in the production process, τ^m, is higher in economy C with respect to A (as shown in Figure 2). Economy C is disadvantaged both in low prevalence of adopters and in terms of population size, and therefore it will experience a late transition and a slow growth (Figure 3, case 3).

4 Conclusions

This paper examines the effect of the distribution of human capital on the advancement and adoption of technology and their influence on the timing of the transition from stagnation to growth and the long-run evolution of the economy. While highly educated individuals enhanced innovation, the absence of an educated labor force hindered technology adoption, delaying the transition from stagnation to growth. Yet, a wide group of educated individuals in the labor force induced faster economic growth in the modern period, compensating for the delayed transition due to the limited presence of highly educated individuals.

The advanced theory offers a number of testable predictions of the link between human capital distribution and long-run growth, which can be linked to the parameters of the models. The cost of raising children can be mapped to environmental factors that generate differences in the cost of producing food. The parameters indicating the cost of educating children at different levels can be mapped to institutional factors influencing the cost of acquiring education for the masses and for the elites.

While the econometric identification of these effects is beyond the scope of this paper and is left for future research, the present study offers new insights on the effect of the distribution of human capital and the long-term development process. In turn, indicating a promising avenue for future empirical investigation of the links between the distribution of human capital and the long-term evolution of living standards.

References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly journal of economics*, 117(4):1231–1294.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica: Journal of the Econometric Society*, pages 323–351.

Ashton, T. S. (1955). *An economic history of England: the 18th century*, volume 3. Taylor & Francis.

Becker, S. O., Hornung, E., and Woessmann, L. (2011). Education and catch-up in the industrial revolution. *American Economic Journal: Macroeconomics*, 3(3):92–126.

Bekar, C. and Lipsey, R. (2004). Science, institutions and the industrial revolution. *Journal of european economic history*, 33(3):709–753.

Benhabib, J. and Spiegel, M. M. (2005a). Chapter 13 human capital and technology diffusion. volume 1 of *Handbook of Economic Growth*, pages 935 – 966. Elsevier.

Benhabib, J. and Spiegel, M. M. (2005b). Human capital and technology diffusion. *Handbook of economic growth*, 1:935–966.

Brown, R. (2002). *Society and economy in modern Britain 1700-1850*. Routledge.

Carillo, M. F. (2020). Agricultural Policy and Long-Run Development: Evidence from Mussolini's Battle for Grain. *The Economic Journal*. ueaa060.

Cervellati, M. and Sunde, U. (2005). Human capital formation, life expectancy, and the process of development. *American Economic Review*, 95(5):1653–1672.

Chanda, A., Cook, C. J., and Putterman, L. (2014). Persistence of fortune: Accounting for population movements, there was no post-columbian reversal. *American Economic Journal: Macroeconomics*, 6(3):1–28.

Comin, D., Easterly, W., and Gong, E. (2010). Was the wealth of nations determined in 1000 bc? *American economic journal. Macroeconomics*, 2(3):65–97.

Croix, D. d. l. and Licandro, O. (2013). The child is father of the man: Implications for the demographic transition. *The Economic Journal*, 123(567):236–261.

Doepke, M. and Zilibotti, F. (2005). The macroeconomics of child labor regulation. *American Economic Review*, pages 1492–1524.

Foster, A. D. and Rosenzweig, M. R. (1996). Technical change and human-capital returns and investments: Evidence from the green revolution. *The American Economic Review*, 86(4):931–953.

Galor, O. (2011). *Unified growth theory*. Princeton University Press.

Galor, O. and Moav, O. (2002). Natural selection and the origin of economic growth. *The Quarterly Journal of Economics*, 117(4):1133–1191.

Galor, O. and Mountford, A. (2008). Trading population for productivity: theory and evidence. *The Review of Economic Studies*, 75(4):1143–1179.

Galor, O. and Weil, D. N. (1999). From malthusian stagnation to modern growth. *The American Economic Review*, 89(2):150–154.

Galor, O. and Weil, D. N. (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American economic review*, pages 806–828.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica, Journal of the Econometric Society*, pages 501–522.

Jacob, M. C. (2014). *The first knowledge economy: Human capital and the European economy, 1750–1850*. Cambridge University Press.

Madsen, J. B. and Murtin, F. (2017). British economic growth since 1270: the role of education. *Journal of Economic Growth*, 22(3):229–272.

Mokyr, J. (1990). *The lever of riches: Technological creativity and economic progress*. Oxford University Press.

Mokyr, J. (2002). *The gifts of Athena: Historical origins of the knowledge economy*. Princeton University Press.

Mokyr, J. (2005). Long-term economic growth and the history of technology. In *Handbook of economic growth*, volume 1, pages 1113–1180. Elsevier.

Musson, A. E. and Robinson, E. (1989). *Science and technology in the industrial revolution*, volume 3. Taylor & Francis.

Nelson, R. R. and Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *The American Economic Review*, 56(1/2):pp. 69–75.

Putterman, L. and Weil, D. N. (2010). Post-1500 Population Flows and The Long-Run Determinants of Economic Growth and Inequality*. *The Quarterly Journal of Economics*, 125(4):1627–1682.

Rolt, L. T. C. (1963). *Thomas Newcomen: the prehistory of the steam engine*. David and Charles.

Stapledon, R. G., Davies, W., et al. (1948). Ley farming. *Ley Farming*, (Edn 2 (revised)).

Strulik, H. and Weisdorf, J. (2008). Population, food, and knowledge: a simple unified growth theory. *Journal of Economic growth*, 13(3):195.

Stuart, R. (1829). *A descriptive history of the steam engine*. Number 24502. printed for Knight and Lacey.

Appendices

A The Speed Function

From equation (47), the threshold can be written, without loss of generality, as,

$$G_t \equiv \big[h_t\big]^{\beta}\big[m_t\big]^{\phi}\frac{(1-\beta-\phi)}{(1-\alpha)}\big((h^o)^*\big)^{-\alpha} - \Big(\frac{A_t^o}{A_t^n}\Big) = 0 \tag{54}$$

Therefore the speed of convergence to the new regime is calculated as the movement the points $(h_t, m_t) \in G_t$ have to do each period in order to compensate the time variation of the productivity ratio, $\Big(\frac{A_t^o}{A_t^n}\Big)$. Therefore, speed at $t+1$ is given by

$$\begin{aligned}
s_{t+1} &= \Big\{\big[h_{t+1}\big]^{\beta}\big[m_{t+1}\big]^{\phi} - \big[h_t\big]^{\beta}\big[m_t\big]^{\phi}\Big\}\frac{(1-\beta-\phi)}{(1-\alpha)} && (55)\\
&= \Bigg\{\frac{\big[h_{t+1}\big]^{\beta}\big[m_{t+1}\big]^{\phi} - \big[h_t\big]^{\beta}\big[m_t\big]^{\phi}}{\big[h_t\big]^{\beta}\big[m_t\big]^{\phi}}\Bigg\}\big[h_t\big]^{\beta}\big[m_t\big]^{\phi}\frac{(1-\beta-\phi)}{(1-\alpha)} && (56)\\
&\approx ln\Bigg(\frac{\big[h_{t+1}\big]^{\beta}\big[m_{t+1}\big]^{\phi}}{\big[h_t\big]^{\beta}\big[m_t\big]^{\phi}}\Bigg)\big[h_t\big]^{\beta}\big[m_t\big]^{\phi}\frac{(1-\beta-\phi)}{(1-\alpha)} && (57)\\
&= ln\Big(\frac{A_{t+1}^o}{A_t^o}\frac{A_t^n}{A_{t+1}^n}\Big)\frac{A_t^o}{A_t^n}\big((h^o)^*\big)^{\alpha} && (58)\\
&= \Big[ln\Big(\frac{A_{t+1}^o}{A_t^o}\Big) - ln\Big(\frac{A_{t+1}^n}{A_t^n}\Big)\Big]\frac{A_t^o}{A_t^n}\big((h^o)^*\big)^{\alpha} && (59)\\
&\approx \|\Big[-\Omega(H_t)\lambda(0)\Big]\frac{A_t^o}{A_t^n}\big((h^o)^*\big)^{\alpha}\| && (60)\\
& && (61)
\end{aligned}$$

where it is understood that speed cannot be negative, thus the absolute value is considered.

B The Distance Function

The distance d^* is the square of the minimum distance between the point (h^*, m^*) and the curve $G_{t=0}$[12]. The (h^*, m^*) point is given by equations (33)

[12]It is considered the distance squared for simplicity of calculations. The results are not affected by this simplification.

and (34); the function $G_{t=0}$ is given by the threshold function in period 0, that is

$$G_0 \equiv \left[h_t\right]^{\beta}\left[m_t\right]^{\phi}\frac{(1-\beta-\phi)}{(1-\alpha)}\left((h^o)^*\right)^{-\alpha}-\left(\frac{A_0^o}{A_0^n}\right)=0 \tag{62}$$

In order to find the minimum distance between a point and a curve, first it is necessary to find a point $(\tilde{h},\tilde{m}) \in G_{t=0}$ that minimize the distance function, that is,

$$\{\tilde{h},\tilde{m}\} = argmin\Big\{\left(\tilde{h}-h^*\right)^2+\left(\tilde{m}-m^*\right)^2\Big\} \tag{63}$$

taking into account that, since $(\tilde{h},\tilde{m}) \in G_{t=0}=0$, then $\tilde{m}=\tilde{m}(\tilde{h})$. Thus, substituting $(\tilde{h},\tilde{m})$ into the generic distance function[13] squared the distance function, d^* is obtained, that is,

$$d^*=\left\{\left[\tilde{h}(h^*,m^*,(h^o)^*)-h^*\right]^2+\left[\tilde{m}\left(\tilde{h}(h^*,m^*,(h^o)^*),(h^o)^*\right)-m^*\right]^2\right\}$$

Notice that from the distance minimization it is sufficient to find $\tilde{h}$ to uniquely determine $\tilde{m}$. Whereas it is not possible to find an explicit solution for $\tilde{h}$, the sign of variations in the cost parameters are derived through the implicit function theorem.

C Comparative Statics on the Transition Timing

Given that

$$t^*=\frac{d^*}{s_{t^*}}$$

the comparative statics exercise is made on the speed function, s_{t^*}, and on the distance function, d^*.

Given equation for the speed derived above, it is straightforward to

[13]The generic distance function, d, is given by

$$d=\sqrt{\left\{\left[\tilde{h}-h^*\right]^2+\left[\tilde{m}-m^*\right]^2\right\}}$$

derive that

$$\partial s_{t^*}/\partial \tau^h < 0 \forall t \leq t^* \tag{64}$$

and

$$\partial s_{t^*}/\partial \tau^l > 0 \forall t \leq t^* \tag{65}$$

The comparative statics exercise on d^* implies,

$$\frac{\partial d^*}{\partial \tau^m} > 0; \ \frac{\partial d^*}{\partial \tau^h} > 0; \ \frac{\partial d^*}{\partial \tau^l} < 0 \tag{66}$$

In order to make the comparative statics exercises on the distance function, d^*, it can be useful to consider the quantities $\tilde{m}$ and $\tilde{h}$. Where, from (62)

$$\tilde{m} = \Big[\frac{\big((h^o)^*\big)^\alpha \frac{1-\alpha}{1-\beta-\phi} A_0^o}{\tilde{h} A_0^n}\Big]^{1/\phi} \tag{67}$$

from the first order conditions of the minimization problem given by (63), taking into account (67), thus $\tilde{h}$ is implicitly defined by function $\tilde{K}$, where,

$$\tilde{K} \equiv (h - h^*) - \frac{1}{\phi}\Big[\Big(\frac{\big((h^o)^*\big)^\alpha \frac{1-\alpha}{1-\beta-\phi} A_0^o/A_0^n}{\tilde{h}^\beta}\Big) - m^*\Big]\Big(\frac{\big((h^o)^*\big)^\alpha \frac{1-\alpha}{1-\beta-\phi} A_0^o/A_0^n}{\tilde{h}^\beta}\Big)\frac{\beta}{\tilde{h}} = 0 \tag{68}$$

The comparative statics is done for each of the parameters τ^j with $j = l, m, h$, where derivatives of $\tilde{h}$ are implicitly derived from $\tilde{K}$. The effect on d^* of a variation in τ^m is given by,

$$\frac{\partial d^*}{\partial \tau^m} = 2\big[\tilde{h} - h^*\big]\Big[\frac{\partial \tilde{h}}{\partial m^*}\frac{\partial m^*}{\partial \tau^m}\Big] + 2[\tilde{m} - m^*]\Big[\Big(\frac{\partial \tilde{m}}{\partial \tilde{h}}\frac{\partial \tilde{h}}{\partial m^*} - 1\Big)\frac{\partial m^*}{\partial \tau^m}\Big] > 0 \tag{69}$$

D Population

The number of individuals in the manual labor force

$$N_t^l = \begin{cases} (1-\alpha)N_t & \forall t < t^c \\ \frac{(1-\beta-\phi)/\tau^l}{(1-\beta-\phi)/\tau^l + \beta/\tau^h + \phi/\tau^m} N_t & \forall t \geq t^* \end{cases} \tag{70}$$

The number of Adopters

$$N_t^m = \frac{(\phi)/\tau^m}{(1-\beta-\phi)/\tau^l + \beta/\tau^h + \phi/\tau^m} N_t \quad \forall t \geq t^* \tag{71}$$

The number of Innovators

$$N_t^h = \begin{cases} \alpha N_t & \forall t < t^c \\ \frac{(\beta)/\tau^h}{(1-\beta-\phi)/\tau^l+\beta/\tau^h+\phi/\tau^m} N_t & \forall t \geq t^* \end{cases} \tag{72}$$

Proof. It comes from the optimization, given the equilibrium quantities $(L_t^*, M_t^*, H_t^*) = (N_t^l l_t^{l,*}, N_t^m l_t^{m,*}, N_t^h l_t^{h,*})$, where l_t^j is the amount of working time of individual $j = l, m, h$ □

D.1 Population Dynamics

$$N_{t+1} = \begin{cases} N_t\left(\frac{1}{(1-\alpha)\tau^l+\alpha\tau^h}\right)\left[1 - \frac{\tilde{c}}{A_t^o}\left(\frac{1+h^{o,*}}{((h^o)^*)^\alpha}\right)\right] & \forall t < t^* \\ \\ N_t\big((1-\beta-\phi)/\tau^l + \beta/\tau^h + \phi/\tau^m\big)\gamma & \forall t \geq t^* \end{cases} \tag{73}$$

Proof. It comes from the optimization, given the equilibrium quantities $(N_t^j)$with $j = l, m, h$, given $\sum_i n_t^{i,j} = N_{t+1}^j$ $\forall i \in (1, N_t)$ where i is the number of parents at time t □

E Lemma 9

$$a)\ g_y^o = 0\ \forall\ t < t^* \tag{74}$$

Proof. Lemma 9 a) comes from feasibility condition that is such that

$$Y_t^o = \tilde{c}N_t^o\ \forall t^* < t \tag{75}$$

Where N_t given derived by equations (70), (71) and (72) considering that $t < t^*$ □

$$b)\ g_y^n = \Omega(H_t)\ \big(1 + \lambda(M_t/N_t)\big)\ \forall\ t \geq t^* \tag{76}$$

Proof. Lemma 9 b) comes from the fact that from equation (77), L_t is a constant fraction of N_t. Where N_t is derived by (the sum of) equations (70), (71) and (72) □

F Aggregate Labor Allocation

$$L_t = \begin{cases} \frac{\tilde{c}N_t}{A_t^o}\left[\frac{\alpha}{1-\alpha}\frac{\tau^l}{\tau^h}\right]^{-\alpha} & \forall t < t^* \\ \frac{(1-\gamma)(1-\beta-\phi)/\tau^l}{(1-\beta-\phi)/\tau^l+\beta/\tau^h+\phi/\tau^m}N_t & \forall t \geq t^* \end{cases} \tag{77}$$

$$H_t = \begin{cases} \frac{\tilde{c}N_t}{A_t^o}\left[\frac{\alpha}{1-\alpha}\frac{\tau^l}{\tau^h}\right]^{1-\alpha} & \forall t < t^* \\ \frac{(1-\gamma)(\beta)/\tau^h}{(1-\beta-\phi)/\tau^l+\beta/\tau^h+\phi/\tau^m}N_t & \forall t \geq t^* \end{cases} \tag{78}$$

$$M_t = \begin{cases} 0 & \forall t < t^* \\ \frac{(1-\gamma)(\phi)/\tau^m}{(1-\beta-\phi)/\tau^l+\beta/\tau^h+\phi/\tau^m}N_t & \forall t \geq t^* \end{cases} \tag{79}$$

G Equilibrium wages

$$w_t^h = \begin{cases} \alpha A_t^o\big[(h^o)^*\big]^{\alpha-1} & \text{if } Y_t^o > 0 \\ \beta A_t^n\big[(h^n)^*\big]^{\beta-1}\big[m^*\big]^{\phi} & \text{if } Y_t^n > 0 \end{cases} \tag{80}$$

$$w_t^m = \phi A_t^n\big[(h^n)^*\big]^{\beta}\big[m^*\big]^{\phi-1} \quad \text{if } Y_t^n > 0 \tag{81}$$

$$w_t^l = \begin{cases} (1-\alpha)A_t^o\big[(h^o)^*\big]^{\alpha} & \text{if } Y_t^o > 0 \\ (1-\beta-\phi)A_t^n\big[(h^n)^*\big]^{\beta}\big[m^*\big]^{\phi} & \text{if } Y_t^n > 0 \end{cases} \tag{82}$$

H The Dynamical System

In order to find H_{t^*} it is necessary to solve the following dynamical system.

From (73) and (78)

$$H_{t+1} = H_t\frac{A_t^o}{A_{t+1}^o}\Big(\frac{1}{(1-\alpha)\tau^l+\alpha\tau^h}\Big)\Big[1-\frac{\tilde{c}}{A_t^o}\Big(\frac{1+(h^o)^*)}{((h^o)^*)^{\alpha}}\Big)\Big] \ \forall t < t^* \tag{83}$$

where $((h^o)^*)$ is given by (24).

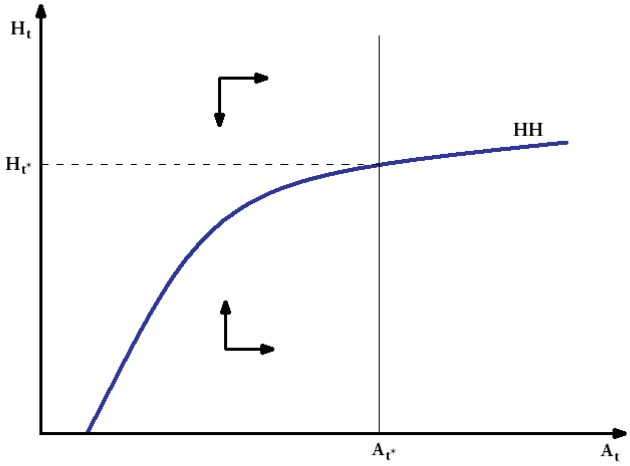
From (39)

$$A_{t+1}^o = A_t^o\big[1+\Omega(H_t)\big] \tag{84}$$

In other terms the dynamical system is given by

$$H_{t+1} = f(A_{t+1}^o, A_t^o)H_t \tag{85}$$

Figure 4: The Dynamical System



Notes: H_t^* is the solution to the dynamical system with $A_t^o = A_{t^*}$.

and

$$A_{t+1}^o = f(A_t^o)H_t \tag{86}$$

Therefore, the AA locus is defined as

$$AA = \{(H_t, A_t^o)|\Delta A_t^o = 0\} \tag{87}$$

where $\Delta A_t^o \equiv A_{t+1}^o - A_t^o = A_t^o[1 + \Omega(H_t)] - A_t^o = 0$ if $\Omega(H_t) = 0$ which is not feasible since $\Omega(0) > 0$ by construction. The intuition is that in this economy technology does not have a steady state in levels because there will always be technological advancements.

The HH locus is given by

$$HH = \{(H_t, A_t^o)|\Delta H_t = 0\} \tag{88}$$

where $\Delta H_t = H_{t+1} - H_t = 0$. Such condition, taking into account (86), is satisfied by

$$H_t = \Omega^{-1}\Big(\big(\frac{1}{(1-\alpha)\tau^l + \alpha\tau^h}\big)\big[1 - \frac{\tilde{c}}{A_t^o}\big(\frac{1+(h^o)^*)}{((h^o)^*)^\alpha}\big)\big] - 1\Big)\ \forall t \leq t^* \tag{89}$$

Let define, H_t as,

$$H_t \equiv \Omega^{-1}\Big(\chi(A^o_t)\Big)\ \forall t \leq t^* \tag{90}$$

The (85) implies that $\frac{\partial A_t}{\partial H} > 0$ and $\frac{\partial^2 A_t}{\partial H_t^2} < 0$ as represented by the blue line in Figure 4.

Finally, $H_{t^*} = \Big\{H_t | A^o_t = A_{t^*}\Big\}$ where $A_{t^*} = \Big\{A^o_t | w^{o,l}_t = \tilde{c}/(1-\gamma)\Big\}$ where $w^{o,l}_t$ is the equilibrium wage in the old regime, given by (23). That is

$$A_{t^*} = \frac{\tilde{c}}{(1-\gamma)(1-\alpha)\big((h^o)^*\big)^{\alpha}} \tag{91}$$

I Lemma 10

$$g^n_y = \Omega(n^{t-t^*})\chi\ \lambda(M_t/N_t) \quad \forall\ t \geq t^* \tag{92}$$

Proof. It comes from equation (76), given that from the dynamics of population, $\lambda(M_t/N_t)$ is constant over, considering that H_t is given by

$$H_t = n^{t-t^*} H_{t^*}\ \forall t > t^*$$

. Where n is given by

$$n = \Big(\frac{1-\beta-\phi}{\tau^l} + \frac{\beta}{\tau^h} + \frac{\phi}{\tau^m}\Big)\gamma \tag{93}$$

Given that H_{t^*} is the solution of the dynamical system, which is given by

$$H_{t^*} = \Omega^{-1}(\chi(A_{t^*}))$$

, where $\chi(A_{t^*})$ is given by

$$\chi(A_{t^*}) = \Big(\frac{1}{(1-\alpha)\tau^l + \alpha\tau^h}\Big)(1 - (1 + \big((h^o)^*\big))(1-\alpha)(1-\gamma)) \tag{94}$$

Where in the latter I make use of equation (91). □