

---

## WORKING PAPER SERIES

2019-ECO-01

### The rural exodus and the rise of Europe

**Thomas Baudin\***

IESEG School of Management and LEM-CNRS (UMR 9221)  
and IRES, Université catholique de Louvain,

**Robert Stelter**

Max Planck Institute for Demographic Research

IESEG School of Management  
Lille Catholic University  
3, rue de la Digue  
F-59000 Lille  
[www.ieseg.fr](http://www.ieseg.fr)  
Tel: 33(0)3 20 54 58 92

---

\* Corresponding author: *IESEG School of Management, Socle de la Grande Arche, 1 Parvis de La Défense - F-92044 Paris La Defense cedex, France,*  
Email address: [t.baudin@ieseg.fr](mailto:t.baudin@ieseg.fr)

# The rural exodus and the rise of Europe\*

Thomas Baudin<sup>†</sup> and Robert Stelter<sup>‡</sup>

January 15, 2019

## Abstract

To assess the importance of the rural exodus in fostering the transition from stagnation to growth, we propose a unified model of growth and internal migrations. Using an original set of Swedish data, we identify the deep parameters of our model. We show that internal migration conditions had to be favorable enough to authorize an exodus out of the countryside in order to fuel the industrial development of cities. We then compare the respective contribution of shocks on internal migration costs, infant mortality and inequalities in agricultural productivity to the economic take-off and the demographic transition that occurred in Sweden. Negative shocks on labor mobility generate larger delays in the take-off to growth compared to mortality shocks equivalent to the Black Death. Deepening inequalities of productivity in the agricultural sector, like it has been done by enclosure movements, contributes to accelerate urbanization at the cost of depressed economic growth.

*Keywords:* Demographic transition, Industrialization, Rural exodus, Mortality differentials, Fertility differentials.

*JEL Classification Numbers:* J11, J13, O41.

---

\*We would like to thank Annika Westberg for helping us with SHiPs data. We are deeply grateful for the computational infrastructure provided by Max Planck Computing and Data Facility and the assistance offered by Markus Rampp. The paper greatly benefited from discussions with Philippe Bocquier, Shankha Chakraborty, David de la Croix, Alexia Fürnkranz-Prskawetz, Paula Gobbi, Victor Hiller, Joël Machado, Andreas Irmen, Fabian Kindermann, Carl Mosk, Jim Oeppen, Holger Strulik, Uwe Sunde and Nico Voigtlaender. We also thank participants to seminars and conferences in Louvain-la-Neuve, Paris, Hamburg, Aix en Provence, Moscow and Brussels for their helpful comments. This project has benefited from the scientific and financial supports of ARC project 15/19-063 on "Family transformations" (French speaking community of Belgium) and Project MALYNES (ANR-18-CE26-0002). The traditional disclaimer applies.

<sup>†</sup>IÉSEG School of Management, LEM UMR 9221 and IRES, Université catholique de Louvain, E-Mail: t.baudin@ieseg.fr, corresponding author

<sup>‡</sup>Max Planck Institute for Demographic Research, E-Mail: stelter@demogr.mpg.de.

# 1 Introduction

Among historians and demographers, the rural exodus is reputed to have been one of the key forces behind the development and the urbanization of European economies in the 19<sup>th</sup> and 20<sup>th</sup> centuries (Saville, 2013; van der Woude *et al.*, 1995). Surprisingly, to the best of our knowledge, economists never questioned this qualitative result deeply. To fill that gap, we estimate the contribution of the rural exodus to the rise of European economies. To do so, we build and estimate the deep parameters of a unified model of growth and internal migration. Using an original dataset on Sweden, we show that the conditions of migration between the countryside and cities have strongly shaped the timing and the intensity of the transition to growth. Among these conditions, we highlight migration costs determining labor mobility, mortality differentials between areas and inequalities of productivity in the agricultural sector. We find that negative shocks on labor mobility generate larger delays in the take-off to growth compared to mortality crisis equivalent to the Black Death. We also show that the gradual enclosure movement that took place in Sweden has magnified inequalities in labor productivity in agriculture. It has expedited the rural exodus and urbanization but at the price of making cities and the whole country poorer, a situation evidenced in today's Sub-Saharan Africa and South-East Asia (see Fay and Opal, 1999 and Gollin *et al.*, 2016).

To determine if the rural exodus had a strong influence on the process of growth, we have first collected data on Sweden between 1760 and 1960. From this data collection, we identify a series of facts in three domains: demographics, production and education. First, compared to the countryside, cities were suffering higher child and infant mortality rates as well as lowered fertility. If the mortality differential vanished smoothly along the process of industrialization, fertility differentials were more persistent and started to vanish only around 1870. Urbanization, that we define as the increase in the proportion of people living in cities,<sup>1</sup> started to progress after 1850.<sup>2</sup> Second, the growth rates of GDP per capita became higher in cities than in the countryside as early as 1820, a date at which a clear break in industrial growth can be detected. Still, total industrial production first exceeded total agricultural production around 1900. Progressively, cities became leading economic areas with an average GDP per capita 7.5 times higher than in the countryside by 1960. Industrialization has been accompanied by a

---

<sup>1</sup>This definition implies that experiencing a rural exodus is neither a necessary nor a sufficient condition for urbanization, as the natural population growth rates of cities and countryside differ.

<sup>2</sup>In 1850, around 10% of the Swedish population were living in cities; but by 1960, this share had increased to 52.3%.

continuous decrease of the price of industrial goods relative to the price of agricultural goods. Third, an educational revolution took place after 1850 making the enrollment rate into primary education exploding until 1960.

To the best of our knowledge, our database is the first one to evidence these facts in a unique setting.<sup>3</sup> These facts highlight a key historical timeline: both urbanization and educational improvements seem to have resulted from the pre-existing differences in economic growth rates between the cities and the countryside. The theory we develop is in line with this empirical finding: the rural exodus which has allowed urban revolutions has been initiated by the technical progress of industries which has attracted peasants to cities. We nevertheless show that the conditions which have governed the migration of peasants to cities have been a key element of the economic success of Sweden. As discussed in Section 6, such a conclusion can be extended to other European countries as well as contemporary Sub-Saharan Africa.

To evidence how the conditions governing the rural exodus have influenced the economic take-off of Sweden, we estimate the deep parameters of an overlapping-generations model that explicitly distinguishes between cities and the countryside. We assume that modern industries are based in cities and that agrarian production takes place in the countryside.<sup>4</sup> Adults are endowed with a certain level of human capital and an intrinsic ability to work in the agricultural sector.<sup>5</sup> Additionally, mortality is area specific. Individuals can decide to stay in their area of birth or to move to the other area. But, they can choose their place of residence only once, at the beginning of adulthood. Return migration is excluded. Once individuals have settled down, they decide on the number of children they have, the educational level they want their surviving children to achieve, and the levels of agricultural and industrial goods they consume.

In early stages of our economy, the society is trapped in a Malthusian regime. It endoge-

---

<sup>3</sup>To the best of our knowledge, Sweden is, with Denmark, the only country that offers reliable local level data on fertility and mortality before the industrial revolution. We focus on Sweden as data are available on a longer time span. See Baudin and Stelter (2019) for additional information about the Danish case.

<sup>4</sup>Assuming that industries are based exclusively in cities is a rather strong assumption. Indeed, Ogilvie and Cerman (1996) document significant levels of proto-industrialization in Western Europe's countryside during and even before the industrial revolution. Nevertheless, here we define industrialization as the development of modern industries that rely on mass production, efficient work organization, a large labor force, and large amounts of human capital.

<sup>5</sup>By the intrinsic abilities to work in the agricultural sector, we mean the number of efficient units of agricultural labor characterizing each person. It is exogenous and fixed once for all for each dynasty. This intrinsic ability may be seen either as an agricultural specific human capital or a productivity factor related to soil quality.

nously starts to experience a global technological progress once the overall workforce has reached a certain threshold.<sup>6</sup> As a result of this technological progress, more and more parents become willing and able to educate their children. However, the improved returns of human capital accumulation is not sufficient to start the process of industrialization; living in cities, and thus leaving the countryside, has to become desirable, as well. This is the key role our model gives to the rural exodus: as natural movements of the population cannot sustain urbanization, generations of peasants have to migrate to urban areas.

What factors might prevent the rural exodus from occurring? One potential factor is the mortality penalty in cities. Mortality peaks in cities could significantly lower urbanization and industrialization rates. Indeed, urbanization and industrialization require a more intense rural exodus while this latter is less desirable. The economic cost of the exodus to cities may also be crucial. When migration cost are high, the prevailing technological progress may not allow for development simply because the economic opportunities offered by cities are not good enough to attract peasants to urban areas. Finally, living in a prosperous rural economy may reduce the incentive to move to cities. Thus, the richer and the more egalitarian the countryside is, the less intense the rural exodus and the need for industrialization are likely to be.<sup>7</sup> These are the three main mechanisms that may prevent industrialization in our set-up and these are the three main mechanisms we quantify in this paper.

We estimate the deep parameters of our model in order to match fertility in cities and the countryside as well as urbanization rates in Sweden between 1760 and 1960. Our model reproduces the main characteristics of these time series very well. We do not try to replicate the dynamics of economic moments, such as shifts in GDP per capita in rural and urban areas, in overall education, or in relative prices of industrial goods. However, our over-identification checks indicate that our model replicates these dynamics well.<sup>8</sup> The three facts documented in our paper are then reproduced qualitatively and quantitatively by our model. Such a result guarantees that our estimation strategy delivers reliable results and allows us to proceed to the evaluation of the importance of

---

<sup>6</sup>As modern industries are the only one using human capital for production and because they are located in cities, we could have assumed that endogenous technological progress arises when the size of urban populations reaches a certain threshold. This agglomeration effect is tested in Appendix D.9.

<sup>7</sup>Considering income differentials between cities and the countryside as a driving force of urbanization is in line with most of the literature on the topic. See for instance Gollin *et al.* (2016).

<sup>8</sup>In Section 4.4, we discuss the reasons why we do not try to match economic moments. The main reason pertains to issues with the definitions of GDP per capita at the local level and of school enrollment rates.

the conditions which have governed the exodus out of the countryside.

Our first experiment evaluates the effects of migration costs on the economic and demographic revolutions of Sweden. We first present evidence that limited migration costs were a necessary condition to allow the rural exodus and then the economic take-off, the fertility transition and the transition to mass education. Indeed, in a scenario in which moving from one area to the other becomes impossible, we see that Sweden stagnates in perpetual poverty. In our benchmark estimation, the cost of migration represents 20.4% of the average income per capita in cities by 1760. Multiplying this cost by five would have delayed the fertility transition in the countryside by a generation or 20 years while multiplying it by 20 would have delayed this transition by almost a century.<sup>9</sup> Higher migration costs are also accompanied by a dramatic increase in levels of economic inequality between cities and the countryside. Indeed, when the migration cost is 20 times higher, we predict that the differential in rural and urban gdp per capita would have been 10 times higher in 1960 than it is in the benchmark.

Our second experiment relates to mortality crisis. We simulate the dynamics of our economy in a case in which, rather than decreasing, infant mortality rates (IMR) were frozen between 1760 and 1960. Such a stagnation would have left the economic dynamics almost unchanged; while, compared to the benchmark, the number of births would have strongly increased, the number of surviving offspring would have decreased. This finding of a rather weak effect of mortality reduction is in line with the results of Galor (2005) and Doepke (2004)<sup>10</sup> and contradicts traditional demographic theories positing that the mortality decline was the main driver of the fertility decline and the associated economic take-off.

In our third experiment, we focus on inequalities of productivity in the agricultural sector. We test a scenario in which the distribution of intrinsic abilities to work in the agricultural sector becomes more unequal. This mean-preserving spread scenario echoes the enclosure movement, which occurred in Sweden along the beginning of the 19<sup>th</sup> century. At that time, pieces of land were redistributed and privatized. While some farmers were lucky and received very productive land, others received land of poor quality or no land at all. In the terms of our model, this movement leads to a

---

<sup>9</sup>As is discussed in Section 4.3, at first sight, multiplying the migration costs by five or 20 may seem excessive by the standards of 1760; nevertheless, one has to keep in mind that real gdp per capita has been multiplied by more than 22 between 1750 and 1960.

<sup>10</sup>Other papers have examined the impact of other types of mortality on the economic take-off. See, among others, the enlightening papers of Cervellati and Sunde (2011), Strulik and Werner (2016), Boucekkine *et al.* (2002), Strulik and Weisdorf (2014) and Voigtländer and Voth (2012)

mass exodus out of the countryside, and to a pauperization of the migrant population. Because of Malthusian checks, the fertility levels of both the urban and the rural populations are strongly depressed. The GDP per capita growth rate accelerates temporarily in both urban and rural areas, but this effect vanishes after one period. The greatly reduced population size postpones the onset of technological progress, and thus growth in income per capita. As a result, we find that the average GDP per capita in Sweden in 1960 would have been much lower in both industrial and agricultural areas compared to our benchmark estimations. It therefore appears that the agricultural revolution was a key determinant of the timing of the economic and demographic Swedish revolutions, but that it may not have been a necessary condition for the emergence of the industrial revolution; a theory that was defended by Clark (2002).

The model we propose contributes to the Unified Growth Theory (hereafter, UGT). UGT provides a consistent framework for explaining how the economies of Western Europe moved from Malthusian stagnation to modern growth, and how this economic take-off is related to the demographic transition. A strand of literature has recently emerged that suggests that certain conditions must be met for the general mechanisms of the UGT to operate. For instance, Squicciarini and Voigtländer (2014) show that a rising total factor productivity is not necessarily sufficient to induce human capital accumulation and the economic take-off. They highlight the role played by upper-tail human capital and scientific knowledge.<sup>11</sup> We complement this literature by emphasizing how the rural exodus had to be affordable for industrialization to occur.

Together with Adams (2017) and Greenwood and Seshadri (2002), we are among the few economists who investigate the role of urban-rural differential demography in the process of economic development. With English data, Adams (2017) calibrates the parameters of a growth model in which rural-urban differences are key.<sup>12</sup> He then predicts that a country will grow more rapidly if it has “*a lower initial urban share, a higher initial income growth rate, or a higher initial population growth rate*”, a prediction that he then tries to validate on cross-sectional data. Greenwood and Seshadri (2002) focus on the rural-urban fertility differential in the US. While they find that technological

---

<sup>11</sup>The Unified Growth Theory has produced a large number of other refinements. Galor (2011) provides an enlightening review of this literature.

<sup>12</sup>Other papers have used calibration techniques applied to the UGT. Lagerlöf (2006) presents a numerical exercise for the model of Galor and Weil (2000) by choosing parameters either from the literature or currently observed moments. Cervellati and Sunde (2005) as well as Strulik and Weisdorf (2014) present illustrative numerical simulations to highlight the respective role of longevity and child survival. Cervellati *et al.* (2013) use panel data for 59 countries since 1880 for a structural estimation in a unified growth model.

progress is essential for the reduction of fertility in rural and urban sectors, they do not study the precise effects of the exodus and of mortality from either a theoretical or a quantitative perspective.

Our paper is organized as follows. Section 2 presents the empirical regularities characterizing the Swedish economic take-off from the 18th to the 20th century. Section 3 proposes our unified model of growth and migrations while Section 4 estimates the deep parameters of this latter. We estimate the effect of shocks on migration costs, mortality rates and agricultural productivity in Section 5. Importantly, in Section 6, we show both how the facts we have highlighted for Sweden are symptomatic of most of Western European countries and how our results can be generalized to these countries.

## 2 Data and facts

In the 20<sup>th</sup> century, demographers and economists collected historical data throughout Europe at the national and the local levels. These data have been widely used by the UGT to reconstruct a global demo-economic history of Western Europe. We complement this approach by providing a dataset for Sweden that makes a clear distinction between rural and urban areas for the series of facts along three dimensions: demography, economics and education.<sup>13</sup>

**Demographic facts.** Historisk statistik för Sverige (Centralbyran, 1969) provides information on rural (“Landsbygd”) and urban (“Städer”) population sizes from 1800 to 1962. Information is available every five years from 1800 to 1855, and on an annual basis in the more recent population data (1856 to 1962). To compute the urban and rural populations for earlier periods, we use the SHiPs Demographic Data Base (SHiPs). SHiPs contains a rich set of demographic variables for the 1749–1859 period for up to around 3,000 parishes, and a classification of these parishes as having an urban or a rural status. This status is institutionally, rather than demographically defined.<sup>14</sup> As the proportion of the Swedish population covered by the SHiPs data varied greatly from one year to another, we decided to use data only for the years covering at least 60% of the Swedish population.

---

<sup>13</sup>Data comparability issues necessarily arise. In Appendix A.1, we describe the general procedures we used for data harmonization and smoothing.

<sup>14</sup>It prevents reclassification issues; for more on this topic, see Bocquier and Costa (2015) and our discussion in Appendix A.3.



Figure 1: OBSERVED (DOTS) AND SMOOTHED POPULATION DYNAMICS (SOLID LINES) IN SWEDEN

In Figure 1, we illustrate the dynamics of the Swedish population size as well as the evolution of the proportion of the population living in cities, thereby obtaining an approximation of urbanization. While the size of the population increased continuously during the whole period, an acceleration occurred around 1815. The share of the population in cities remained constant over the first 100 years of observations. In the middle of the 19<sup>th</sup> century, urbanization took off. From 1850 to 1960, the share of individuals living in urban areas increased by up to 50%.

Demographers like Woods (2003) have emphasized that cities experienced a mortality penalty relative to the countryside at the time of the industrial revolution. In a perfect framework, we would reconstruct the dynamics of mortality at all ages to get a complete picture of this phenomenon. But the restrictions on the availability and the quality of data do not allow us to build such a time series for a country like Sweden in the 18<sup>th</sup> and 19<sup>th</sup> centuries. Still, by using data from the SHiPs database and Historisk statistik för Sverige, we are able to examine the local and global dynamics of infant mortality rates (IMR), which are important determinants of the number of births within a family, as well as of the number of surviving children.<sup>15</sup>

<sup>15</sup>Brändström (1993) presented Swedish infant mortality rates at a provincial level, and found differences within provinces with similar urbanization rates.

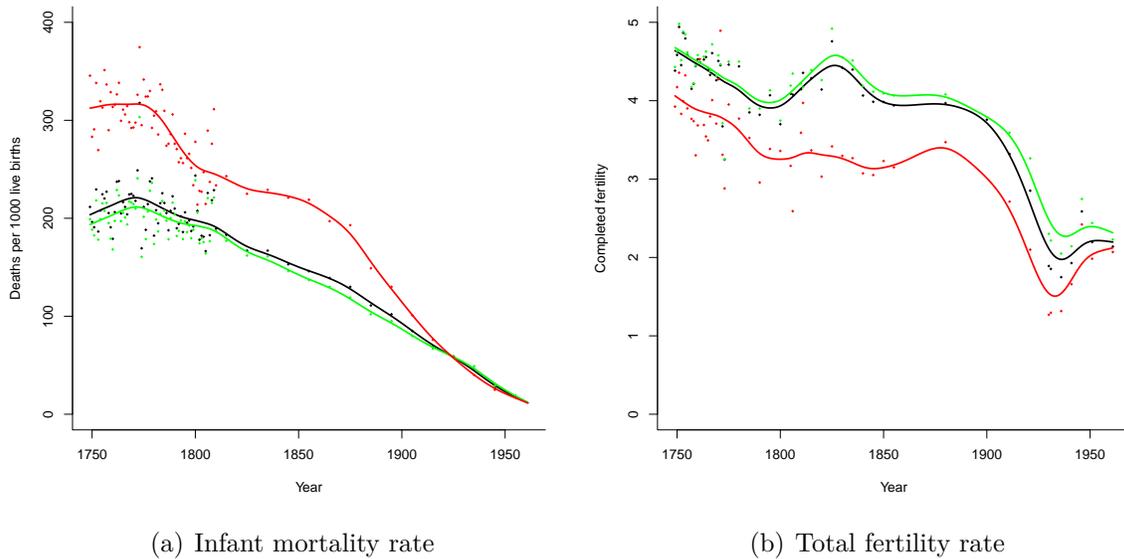


Figure 2: TRENDS (SOLID LINES) AND OBSERVED (DOTS) INFANT MORTALITY AND FERTILITY IN SWEDEN (BLACK), THE COUNTRYSIDE (GREEN) AND CITIES (RED).

Figure 2a indicates that in the middle of the 18<sup>th</sup> century, infant mortality was quite high in Sweden. In rural areas, only four out of five children on average survived until age one. The situation was even worse in cities, where the IMR was 10 percentage points higher. Around 1775, infant mortality started to diminish. The IMR trend in rural Sweden and across the country was quite linear over almost all of the next 200 years. The history of urban infant mortality was characterized by sharp declines followed by periods of relative stagnation. The first dramatic drop occurred at the end of the 18<sup>th</sup> century, and shrank the rural-urban mortality gap to five percentage points. Thereafter, however, the survival probability gains in rural areas exceeded those in cities. By the middle of the 19<sup>th</sup> century, the differential had expanded to 10 percentage points. After this point, there was a sustained trend toward convergence. By the early 20<sup>th</sup> century, the rural-urban IMR gap had vanished.

Looking at population age's pyramids (as those exposed in Figure 14 in Appendix A.2), we see that the respective age structures in cities and the countryside were very different. Using Crude Births Rate (CBR hereafter) to measure local fertility rates may then lead to strongly biased estimations as the number of women of fertile ages living in these areas greatly differ. For this reason, we use Coale's (general) index of fertility to estimate dynamics of the total fertility rate (TFR).<sup>16</sup> As for most of our variables, we merge data

<sup>16</sup>We discuss the method and the advantages of alternative fertility measures more deeply in Ap-

from the SHiPs Demographic Data Base for the 1749–1855 period, and from several statistical yearbooks for the 1910–1970 period. As data on the age structure in specific areas are not available for the 1855–1910 period, we have chosen to use the calculation of Mosk (1980) for the year 1880 to complement our data collection. Looking at Figure 2b, we can see that the total fertility rate for Sweden in the year 1749 was equal to 4.6. Fertility was slightly higher in rural (4.7) and lower in urban areas (4.1). Thus, rural fertility exceeded urban fertility by around 15% at the beginning of our observations. From 1750 to 1880, fertility in the countryside fluctuated without showing a clear reduction, which is a key characteristic of Malthusian regimes. In the same period, urban fertility experienced a first significant decline and then stagnated. From around 1880 to the end of our period of observation, fertility in both areas decreased, despite having reached a maximum differential of close to 60% around 1930. At that point, both fertility rates started to increase and converged around 1960.

As was stated in the introduction, fertility was lower and mortality was higher in cities than in the countryside (Figure 2b and 2a). Without a significant rural exodus, the share of the population living in the countryside should have increased. Yet, Figure 1b shows that the opposite was the case.

**Economic facts.** The Swedish Historical National Accounts provide information on production per sector in Sweden since 1560. Economic activities are divided into seven main sectors: agriculture, manufacturing and industry, building and construction, transport and communication, private services, public services, and services of dwellings; see Schön *et al.* (2012). We use the sector *agriculture* to approximate agricultural output in rural areas, and the sector *manufacturing and industry* to approximate the industrial production in cities. We do not include the production of any of the other sectors in either of these two sectors. In other words, our definition of industrial production is quite conservative.<sup>17</sup> Time series are available in current and in constant factor prices (Mill. SEK).

---

pendix A.2.

<sup>17</sup>Applying less restrictive assumptions to the industrial sector shifts its production upward and increases the income differential in GDP per capita. Alternative classifications are illustrated in Appendix A.4.

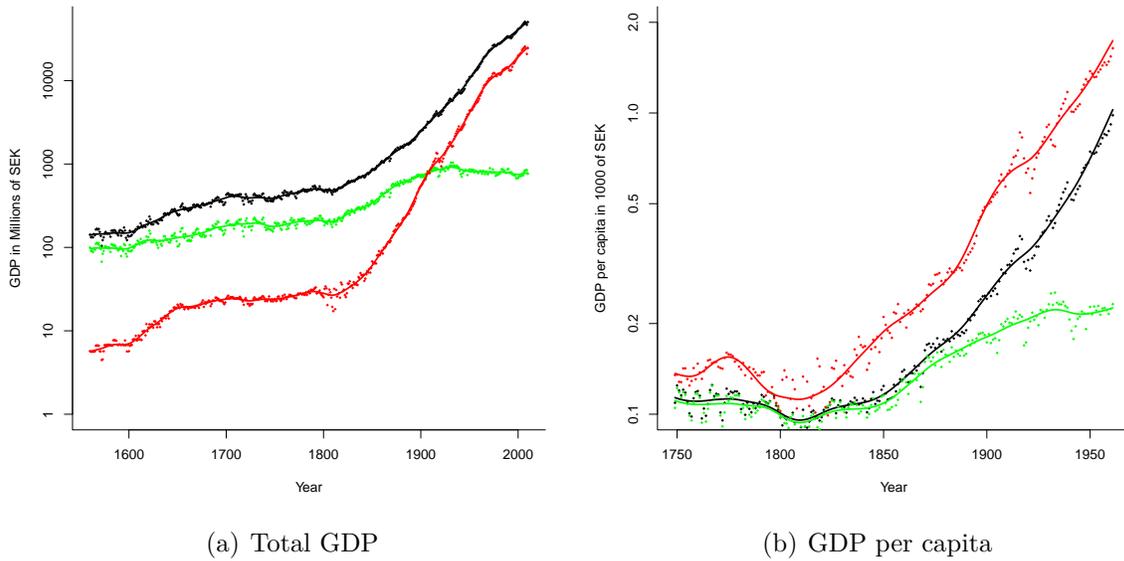


Figure 3: OBSERVATIONS (DOTS) AND TRENDS (SOLID LINES) IN SWEDISH (BLACK), INDUSTRIAL (RED), AND AGRICULTURAL (GREEN) GDP AND GDP PER CAPITA.

Figure 3 shows the evolution of Swedish GDP in constant prices at both the national and the local levels. As our data start very early, we are able to observe a period of stagnation even if as early as in 1610 aggregate production started to grow slightly; mainly driven by industrial production. At the end of the 16<sup>th</sup> century, agricultural production was around 14 times higher than industrial production. This difference decreased to a factor of eight up to the beginning of the industrial revolution. Around 1820, a sharp break occurred, which marks the beginning of the transition to the modern economic regime in Sweden. At that point, total industrial production exploded, growing at the annual constant rate of 3.49% from 1820 to 2010. Agricultural production also increased, but started to stagnate around 1930.

From population estimates, we have been able to reconstruct the evolution of GDP per capita in cities and the countryside. Figure 3b shows a great divergence between the two areas. Growth in gdp per capita of cities began as early as 1820, while the rural gdp per capita remained around its pre-crisis level until around 1850. Around 1850 both areas and overall Sweden experienced sustained growth – the beginning of the modern area. Nevertheless, the sub-country measure is based on the assumption that people in cities were working exclusively in industries, while people in rural areas were working exclusively in agriculture. This is a rather strong assumption, which neglects commuting and proto-industrialization. We discuss this issue and relax the assumption

in Section 4.4 and Appendix A.5.

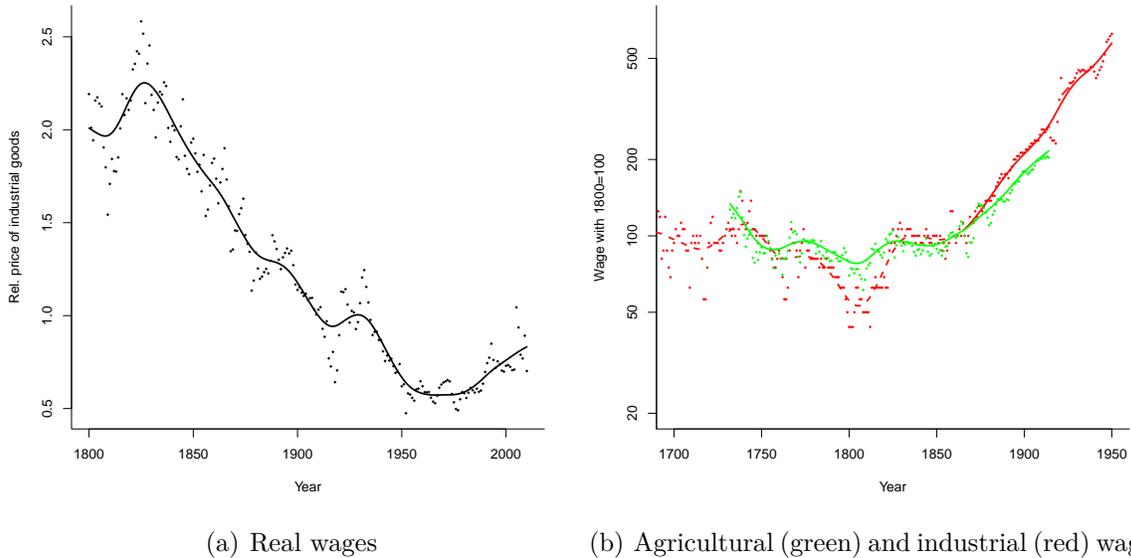


Figure 4: EVOLUTION OF OBSERVATIONS (DOTS) AND TRENDS (SOLID LINES) IN SWEDEN'S RELATIVE PRICES AND WAGES.

Industrialization was accompanied by a strong decrease in the prices of manufactured goods relative to the prices of agricultural goods. Still using the Swedish National Accounts, we divide GDP in current prices by GDP in constant prices in each area, which allows us to approximate the dynamics of prices in rural and industrial sectors since 1800. During the period of de-industrialization at the beginning of the 19<sup>th</sup> century, the relative prices of industrial goods first increase slightly, and then decline after 1825; see Figure 4a. At the same time wages normalized to 100 in 1860<sup>18</sup>, illustrated in Fig 4b, declined. Both sectors, industries and agriculture, underwent a period of low wages at the beginning of the 19<sup>th</sup> century, whereat the insection was higher in industrial than in farm wages. While the gdp per capita already started to grow permanently around 1820 in cities and 1850 in total Sweden, we evidence a delay in wages. Industrial wages took-off in the 1860s. Agricultural wages followed in the 1870s. We show similar facts for France and Great-Britain in Section 6.

**Educational facts** Like everywhere else in Western Europe, industrialization and economic modernization in Sweden were accompanied by an educational revolution.

---

<sup>18</sup>Male agricultural data originates from Jörberg (1972) and is deflated with the time series presented in Edvinsson and Söderberg (2010). Industrial wages from Söderberg (2010) present wages for unskilled laborer in Stockholm until 1864 (dashed plain red trend in 4b) and industrial wages afterwards (plain red trend in Fig. 4b).

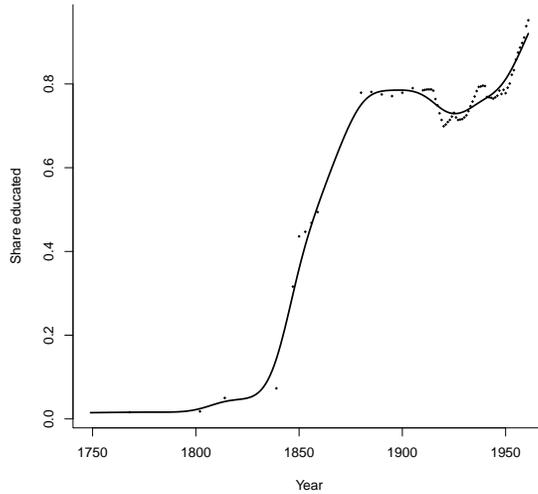


Figure 5: OBSERVATIONS (DOTS) AND TREND (SOLID LINES) IN SWEDISH EDUCATION.

This development has been highlighted by Galor (2005), among others. While this approach is not perfect, we have decided to approximate the dynamics of education using primary education enrollment rates, defined as percentage of the population aged 5–14 who were attending school. Enrollment rates in primary (and lower secondary) school have been extracted from de la Croix *et al.* (2008). Before the 19<sup>th</sup> century, enrollment in primary education was negligible, and remained low until 1842, when compulsory elementary education was introduced (Flora, 1983). At that point, primary education enrollment rates quickly rose to around 80%, and remained at this level until the end of World War II. A further increase in the post-war period is observable in Figure 5. By 1960, almost all children were enrolled in school.

Summarizing our stylized facts, we note that urbanization and educational improvements have been preceded by the emergence of differentiated economic growth rates between the cities and the countryside as well as by the start of the demographic transition in cities around 1775, a time when mortality started to decrease.

### 3 The model

We assume a closed economy in discrete time ( $t = 0, 1, 2, \dots, \infty$ ) that is composed of overlapping generations. Individuals live for two periods, childhood and parenthood;<sup>19</sup> and they can live either in the countryside  $R$  or in cities  $U$ . For simplicity, we consider asexual reproduction, and disregard the importance of local marriage markets; see Voigtländer and Voth (2013) for a discussion. The number of people living in area  $\mathbb{A}$  at date  $t$  is denoted  $N_t^{\mathbb{A}}$ . We denote  $\mathcal{I}_t^{\mathbb{A}}$  as the set containing the people living in area  $\mathbb{A}$ , who are indexed by  $i = \{1, 2, \dots, N_t^{\mathbb{A}} - 1, N_t^{\mathbb{A}}\}$ .

#### 3.1 Production sectors

In terms of production technologies, we assume that the only type of activity occurring in the countryside is agriculture, while the only type of activity occurring in cities is manufactured production.<sup>20</sup> In rural areas, firms produce the agricultural output  $Y_t^R$  with efficient labor and land following a technology with constant returns to scale:

$$Y_t^R = A_{R,t} L_{R,t}^{1-\theta} X^\theta. \quad (1)$$

$A_{R,t}$  denotes the agricultural total factor productivity, and  $\theta$  is the partial output elasticity of land. Agricultural labor is defined as  $L_{R,t} \equiv \sum_{\mathcal{I}^R} a^i I_t^i$  where  $I_t^i$  denotes the labor force participation of person  $i$ , while  $a^i$  is her intrinsic ability to work in the agricultural sector.  $X$  is the effective quantity of land. Agricultural efficiency units of labor are compensated by  $w_t^R$ . Given that land is assumed to be a public good, there are no property rights (Galor and Weil, 2000); furthermore, competition is perfect and so wage in the agricultural sector equals production per capita:

$$w_t^R = A_{R,t} L_{R,t}^{-\theta} X^\theta = \frac{Y_t^R}{L_{R,t}}.$$

Manufactured goods are produced in urban areas using efficient labor  $H_t$ :

$$Y_t^U = A_{U,t} H_t^\gamma, \quad (2)$$

---

<sup>19</sup>By doing so, we disregard potentially important phenomena like savings or intergenerational co-residence. On this latter point, see the recent contribution of Pensieroso and Sommacal (2017).

<sup>20</sup>We discuss this assumption and the related critics of Clark *et al.* (2012) in Appendix A.5.

with  $A_{U,t}$  as total factor productivity in the industrial sector,  $\gamma$  as partial output elasticity of efficient labor, and  $Y_t^U$  as manufactured output. The quantity of efficient industrial labor is equal to the total amount of human capital within the industrial workforce. We denote  $h_t^i$  the level of human capital of a person  $i$  born in  $t - 1$  such that  $H_t \equiv \sum_{\mathcal{I}^U} h_t^i I_t^i$ . As for agriculture, we assume that in the industrial sector competition is perfect and workers are self-employed. We get that  $w_t^U = \frac{p_t Y_t^U}{H_t}$ , with  $p_t$  denoting the relative price of manufactured goods in terms of the agricultural goods, these latter being our numeraire.

We assume the existence of general knowledge. We denote the stock of this knowledge  $A_t$ , it endogenously starts to increase when the size of the population exceeds a critical value  $\bar{N}$ :

$$\frac{A_{t+1} - A_t}{A_t} = \begin{cases} 0 & \text{if } N_t < \bar{N} \\ \Gamma > 0 & \text{if } N_t \geq \bar{N}. \end{cases} \quad (3)$$

This assumption is close to that of Galor and Weil (1999), which was formalized in Galor and Weil (2000).<sup>21</sup> Furthermore, we assume that the respective sector specific technological situation is as follows:

$$A_{R,t} = B_R A_t \quad \text{and} \quad A_{U,t} = B_U A_t H_t^{1-\gamma}. \quad (4)$$

For all  $\mathbb{A}$ ,  $B_{\mathbb{A}} \in \mathbb{R}^+$  is a scale parameter allowing for initial technological differences between industries and agriculture. Together with Equation (2), Equation (4) is at the core of our theory: benefiting from an (ex-post)  $AK$  technology, human capital is productive in cities only. This means that urbanization may be a pre-requirement for industrialization.

Technological progress is related to population density and human capital in a complex way. First, we assume that technological progress in modern industries is magnified by the accumulation of human capital in cities. A great amount of literature evidences that cities and their industries required a specific and agglomerated accumulation of human capital to develop, and that they were producing industrial knowledge in return.<sup>22</sup> Second, because the size of the total population has to reach a critical threshold,

---

<sup>21</sup>An alternative would be to assume that technological progress appears when the size of urban populations reaches a threshold. This would mean that population growth in cities was a pre-requirement for the take-off, which is not the case in our model. We nevertheless test this alternative version in Appendix D.9; neither the quality of the fit nor our main predictions are changed significantly.

<sup>22</sup>See for instance the enlightening contributions of Simon and Nardinelli (1996, 2002) in the case of England and US. Magnifying effects of agglomeration (through peer effects and imitation) on the

household decisions in terms of fertility and education are the endogenous drivers of the appearance of technological progress: in the Malthusian world, families accumulating human capital are also those having the highest fertility rates. Be them rural or urban, these families fuel the society on its way to the technological revolution.

### 3.2 Individuals

We assume the existence of a subsistence level for income, denoted  $\bar{c} > 0$ , below which people only consume agricultural goods  $c_t$ .<sup>23</sup> This is a state of extreme poverty in which having kids and consuming industrial goods is impossible. When people's income levels increase sufficiently, they can consume industrial goods  $d_t$ , give birth to  $n_t$  children, educate them  $e_t$  and migrate. Migration arises at the beginning of adult life and entails a fix cost  $\kappa$ : If the person moves from one area to another,  $\kappa$  equals  $\bar{\kappa} > 0$ , and is otherwise zero. For the sake of simplicity, we assume that children are born after potential migration, and are costly in terms of time: giving birth to a child requires a part  $\xi$  of the parental time unit. Furthermore, with exogenous probability  $q_t^{\mathbb{A}}$ , the child survives to age one. The child then needs additional time from her parents, which is denoted  $\zeta^{\mathbb{A}}$ , and is area-specific. When having  $n_t$  children and living in area  $\mathbb{A}$ , parental labor supply is  $I_t \leq 1 - (\xi + \zeta^{\mathbb{A}} q_t^{\mathbb{A}}) n_t$ . Finally, education entails industrial good costs, such that each unit of education costs the parents  $\beta p_t$ , and their budget constraints can be written as:

$$c_t + p_t d_t + \beta q_t^{\mathbb{A}} p_t n_t e_t = \left(1 - [\xi + \zeta^{\mathbb{A}} q_t^{\mathbb{A}}] n_t\right) \omega_t^{\mathbb{A}} - \kappa. \quad (5)$$

Because raising children takes time, the number of surviving children a mother can have is limited to  $n_M \equiv \frac{1}{\xi + \zeta^{\mathbb{A}} q_t^{\mathbb{A}}}$ .  $\omega_t^{\mathbb{A}}$  is the potential labor income of a parent living in area  $\mathbb{A}$  at date  $t$ . Potential income is only determined by labor income, such that:

$$\begin{aligned} \text{if } \mathbb{A} = R \text{ then } \omega_t^R &= a w_t^R \\ \text{if } \mathbb{A} = U \text{ then } \omega_t^U &= w_t^U h_t, \end{aligned}$$

with  $w^{\mathbb{A}}$  denoting the wage per efficient unit of labor in sector  $\mathbb{A}$ . Similar to de la Croix accumulation of human capital in cities compared to rural areas are theorized for instance by Glaeser (1999). The specific use of human capital in industries is also a central assumption of papers like Tamura (2002).

<sup>23</sup>For reasons of clarity, we hereafter exclude the superscript for the individuals  $i$ .

and Doepke (2003), human capital accumulates according to:

$$h_{t+1} = (v + e_t)^\phi h_t^{1-\phi}, \quad (6)$$

where  $v > 0$  ensures that in the absence of educational investment, human capital does not converge to zero. Each dynasty is endowed with a specific initial level of human capital  $h_0 > 0$  and a specific efficiency of agricultural activities  $a$ .  $h_0$  and  $a$  differ from one dynasty to another, and are drawn from two, a priori distinct, exponential distributions. There is an intrafamily transmission of human capital, whereby  $1 - \phi$  denotes the elasticity with which parents' human capital enters children's human capital accumulation.

Regardless of whether they live in urban or rural areas, parents are characterized by similar preferences and face the same technology to accumulate human capital. The utility of an agent living in area  $\mathbb{A}$  is denoted  $W_t^{\mathbb{A}}$ . When  $c_t \leq \bar{c}$ ,  $W_t^{\mathbb{A}} \rightarrow -\infty$  by definition, while when  $c_t > \bar{c}$ :

$$W_t^{\mathbb{A}} = \alpha \ln(c_t - \bar{c}) + (1 - \alpha) \ln(d_t + \varepsilon) + \rho \left[ \ln(n_t q_t^{\mathbb{A}}) + \ln h_{t+1} \right]. \quad (7)$$

As parents are characterized as having “warm glow preferences,” we assume that they do not value the future well-being of their children, but that they do care about future human capital  $h_{t+1}$ . As  $\varepsilon$  is a positive scalar, equations 6 and 7 imply that education and industrial goods are superior goods.

## Parental behavior

We assume that young adults have to decide between migrating or staying where they were born before having children, educating them, consuming and working. The individual decision problem is then a two stage decision problem that we solve backward: (i) future parents calculate their optimal behavior in each location given the prevailing economic conditions; (ii) they then decide where to live by comparing their expected indirect utilities.

In Table 1, we list the six potential regimes adults can live in. Regimes are obtained by maximizing the parental utility function (7) with respect to budget constraints (5), maximal fertility, and minimal consumption constraints. In regime R1, by definition, being poor leads to childlessness and very low consumption levels.<sup>24</sup> In regime R2,

---

<sup>24</sup>For a discussion about childlessness, see Baudin *et al.* (2015b).

Table 1: REGIMES IN WHICH ADULTS CAN LIVE

Regime	$c_t$	$n_t$	$d_t$	$e_t$
R1	$\omega_t^\Delta$	0	0	0
R2	$\frac{\alpha\omega_t^\Delta}{\alpha+\rho} + \frac{\rho}{\alpha+\rho}\bar{c}$	$\frac{\rho(\omega_t^\Delta - \bar{c})}{(\alpha+\rho)(\xi + \zeta^\Delta q_t^\Delta)\omega_t^\Delta}$	0	0
R3	$\frac{\alpha(\omega_t^\Delta - \kappa)}{\alpha+\rho} + \frac{\rho}{\alpha+\rho}\bar{c}$	$\frac{\rho(\omega_t^\Delta - \kappa - \bar{c})}{(\alpha+\rho)(\xi + \zeta^\Delta q_t^\Delta)\omega_t^\Delta}$	0	0
R4	$\frac{\alpha\tilde{\omega}_t^\Delta}{1+\rho} + \tilde{c}$	$\frac{\rho(\tilde{\omega}_t^\Delta - \bar{c})}{(1+\rho)(\xi + \zeta^\Delta q_t^\Delta)\omega_t^\Delta}$	$\frac{1-\alpha}{1+\rho} \frac{(\tilde{\omega}_t^\Delta - \bar{c})}{p_t} - \varepsilon$	0
R5	$\frac{\alpha(\omega_t^\Delta - \kappa)}{\alpha+\rho} + \frac{\rho}{\alpha+\rho}\bar{c}$	$\frac{\rho(1-\phi)(\omega_t^\Delta - \kappa - \bar{c})}{(\alpha+\rho)\pi_t}$	0	$\frac{\phi\pi_t - (1-\phi)\beta p_t q_t^\Delta v}{(1-\phi)\beta p_t q_t^\Delta}$
R6	$\frac{\alpha\tilde{\omega}_t^\Delta}{1+\rho} + \tilde{c}$	$\frac{\rho(1-\phi)(\tilde{\omega}_t^\Delta - \bar{c})}{(1+\rho)\pi_t}$	$\frac{1-\alpha}{1+\rho} \frac{(\tilde{\omega}_t^\Delta - \bar{c})}{p_t} - \varepsilon$	$\frac{\phi\pi_t - (1-\phi)\beta p_t q_t^\Delta v}{(1-\phi)\beta p_t q_t^\Delta}$

with  $\tilde{\omega}_t^\Delta = \omega_t^\Delta - \kappa + \varepsilon p_t$ ,  $\tilde{c} = \frac{1-\alpha+\rho}{1+\rho}\bar{c}$  and  $\pi_t = (\xi + \zeta^\Delta q_t^\Delta)\omega_t^\Delta - \beta p_t q_t^\Delta v$

having children is possible and is optimal, but as in Regime R1, the number of children is limited by  $\bar{c}$ , and it is impossible for individuals to move from one area to another, to consume industrial goods, and to educate their children. Moving from one area to another becomes possible in regime R3. R1 to R3 are Malthusian regimes. For a given set of parameters, Regimes R4 and R5 cannot co-exist while both of them describe situations where the parental income is too low to afford all kinds of consumptions. When the cost of education  $\beta$  is low relatively to the Stone-Geary element  $\varepsilon$ , then R5 exists and describes a situation in which parents can educate their children, but they cannot consume industrial goods. In the opposite situation where  $\beta$  is relatively high, parents can afford industrial goods but cannot educate their children. Finally, regime R6 is a state of higher wealth in which there are no binding constraints, except for budget.

Given the system of prices  $\{p_t, w_t^U, w_t^R\}$  and mortality rates  $\{q_t^U, q_t^R\}$ , the regime an individual lives in depends on her income  $\omega_t$  and migration costs. In Appendix B, we characterize this behavior for any configuration of our set of parameters; we also learn that the conditions prevailing on the market for industrial goods will determine the timing of the take-off to modern economic growth and to urbanization, as they drive the rise of the national demand for industrial products, and then for human capital. Obviously, these conditions are deeply rooted by the parameters of our model. In Appendix B.2, we highlight the non-monotonicity that characterizes the relationship between fertility and income.

## Should I stay or should I go?

In deciding where she wants to live, a person born in area  $\mathbb{A}$  has to compare the value of living in a city to the value of living in the countryside. She will decide to live in a city if:

$$V[U|\mathbb{A}, h_t, \omega_t^U, q_t^U] \geq V[R|\mathbb{A}, h_t, \omega_t^R, q_t^R]. \quad (8)$$

For extreme values of  $h_t$ , the choice between living in the countryside and living in a city is very easy to determine. First, if the potential income in the destination area,  $\omega_t^{\mathbb{A}}$ , is very low (because either the wage in area  $\mathbb{A}$  or individual productivity is very low), the person is not able to move, and is therefore forced to stay in the area where she was born. This is the case when  $\omega_t^{\mathbb{A}} < \bar{c} + \bar{\kappa}$ .

We can also show that there is a threshold  $\check{h}_t$ , such that when  $\forall h_t \geq \check{h}_t$ , a person will prefer to live in a city. Indeed, at the individual level,  $\omega_t^R = aw_t^R$  does not depend on  $h_t$ . Hence, an increase in  $h_t$  will change the person's allocation of her total investment in children between quality and quantity, but it will not alter her potential income. In other words, a higher  $h_t$  does not change the set of choices an individual living in the countryside has. By contrast, the set of choices a person living in a city has enlarges when  $h_t$  increases, which implies that the value of living in a city increases monotonously with  $h_t$ , but that this is not the case for the value of living in the countryside. The threshold  $\check{h}_t$  then follows. But if this is the case, what prevents the existence of an equilibrium without farmers? The answer to this question is both the decreasing marginal productivity of agricultural labor and the need to consume agricultural goods.

Importantly, the value of living in one place decreases with the infant mortality rate that prevails in this place. For this reason, higher infant mortality rates in cities may hinder a rural exodus. The returns to human capital will have to be greater to attract new workers if infant mortality is higher.

### 3.3 Market equilibria

We have four markets: the labor market in cities, the labor market in the countryside, the industrial goods market, and the agricultural goods market. Thanks to Walras'

law, we know that markets are clearing if the three following conditions are met:

$$w_t^U = A_t^U p_t \quad (9)$$

$$\sum_{i \in \mathcal{I}_R} a^i I_t^i = X \left( \frac{A_t^R}{w_t^R} \right)^{\frac{1}{\theta}} \quad (10)$$

$$\sum_{i \in \mathcal{I}_U} d_t^i + \sum_{i \in \mathcal{I}_R} d_t^i = A_t^U \sum_{i \in \mathcal{I}_U} h_t^i I_t^i \quad (11)$$

Eq. (10) ensures equilibria in the urban labor market. Because of the  $AK$  production technology, wages do not depend on labor supply. Eq. (11) ensures equilibria on the rural labor market. Demand is a decreasing function of wages, which are equal to the average product, since land is a public good. Labor supply depends on the number of children of the people living in the countryside. The more children these people have, the less they participate in the labor market. Finally, condition (11) describes the equilibria on the industrial goods market. The demand comes from both cities and the countryside, while the supply is ensured by cities only. As for the agricultural production in the countryside, the higher the fertility of workers in cities, the lower the total production of industrial goods.

The resource constraints of the economy writes as follows:

$$Y_t = p_t Y_t^U + Y_t^R = \sum_{i \in \mathcal{I}_U} (c_t^i + p_t d_t^i + \beta p_t q_t^U e_t^i n_t^i) + \sum_{i \in \{\bar{\mathcal{I}}_U \cup \bar{\mathcal{I}}_R\}} \bar{\kappa} + \sum_{i \in \mathcal{I}_R} (c_t^i + p_t d_t^i + \beta p_t q_t^R e_t^i n_t^i)$$

where  $\bar{\mathcal{I}}_U \subseteq \mathcal{I}_U$  contains the newcomers to the cities who had to pay  $\bar{\kappa}$ , while  $\bar{\mathcal{I}}_R \subseteq \mathcal{I}_R$  contains the newcomers to the countryside who also had to pay  $\bar{\kappa}$ .

## 4 Identification

To determine how well our theory fits the data, we use the simulated method of moments (hereafter, SMM). We first set some parameters a priori, and then determine the value of the remaining parameters by minimizing the squared distance between empirical and simulated demographic moments. Economic moments are then used for overidentification checks. Before discussing how we proceed for each type of parameter, we first explain how we infer total fertility rates from the model's predictions, and

second how human capital and agricultural abilities are initially distributed among the 5,000 dynasties that will generate our simulated moments.

For consistency we use again the two step procedure of Section 2 to approximate the Total Fertility Rate  $n_t^{\mathbb{A}}$  in area  $\mathbb{A}$  at date  $t$ . We first aggregate individual fertility decisions to compute a general Coale Index (Coale, 1969) taking into account that our population is composed of single persons only. We denote the Coale Index prevailing in area  $\mathbb{A}$  at date  $t$  as  $\hat{I}_t^{\mathbb{A}}$ . Then, we approximate the TFR by using the TFR of the Hutterites denoted  $\mathcal{N}^H$  such as  $n_t^{\mathbb{A}} = \hat{I}_t^{\mathbb{A}} * \mathcal{N}^H$ . By doing so, we allow for comparison between our model predictions and the data we have collected in a directly interpretable way.<sup>25</sup>

For each of our 5,000 dynasties, we draw an initial human capital  $h_0^i$  from an exponential distribution, with  $\lambda$  as the average. To ensure that the distribution is fulfilled ex post, we determine the initial value of each household by the inverse of the distribution such that  $h_0^i = -\frac{1}{\lambda} \ln\left(1 - \frac{i}{N_0+1}\right)$ . Turning to agricultural abilities  $a_0^i$ , we assume for the sake of convenience that its distribution is a symmetrical permutation of human capital. In other words, the dynasty with the highest human capital has the lowest agricultural abilities  $h_0^i = h^{\text{MAX}}$ , then  $a_0^i = a^{\text{MIN}} = h^{\text{MIN}}$ , and vice versa. The median dynasty in terms of human capital is also the median dynasty in terms of agricultural abilities. Finally, we assume that the adults with the highest initial human capital initially are located in cities, and that the adults with the highest agricultural abilities are located in the countryside.

## 4.1 A priori fixed parameters

We use the data for IMRs we have collected to build Figure 2a and pick values every 20 years. We then apply the following formula,  $q_t^{\mathbb{A}} = 1 - \text{IMR}_t^{\mathbb{A}}$ , to obtain survival probabilities presented in Table 2.<sup>26</sup>

Table 2: INFANTS' SURVIVAL PROBABILITY

Year	1760	1780	1800	1820	1840	1860	1880	1900	1920	1940	1960
$q^R$	0.797	0.794	0.807	0.828	0.847	0.866	0.889	0.913	0.937	0.961	0.987
$q^U$	0.684	0.691	0.746	0.765	0.775	0.791	0.830	0.886	0.933	0.967	0.988

<sup>25</sup>For more details see Appendix A.2.

<sup>26</sup>We describe the general procedure in more detail in Appendix A.1.

We initially allocate the adults across areas according to the observation in 1760, such that  $\delta_0 = 9.6\%$ .  $1 - \theta$  denotes the elasticity of the agricultural production function with respect to labor. In the literature, values of  $1 - \theta$  vary between 0.6 in Adamopoulos (2008) and 0.71 in Desmet and Parente (2012). We fix the elasticity of labor ( $1 - \theta$ ) to 0.65, and hence  $\theta = 0.35$ . Additionally, we assume that a share  $\xi = 0.04$  of each period is required to give birth. This assumption corresponds to a little less than 10 months for each pregnancy. Implicitly, this fixes the maximum number of pregnancies to 25. Finally, we normalize the size of land  $X$  and the initial TFP  $A_0$  to one.

## 4.2 Moments to be matched and minimum distance

A set of 15 parameters are still to be estimated. We have chosen this set of parameters to minimize the squared distance between the 33 observed and predicted moments:

$$f(p) = [d - s(p)] W [d - s(p)]'. \quad (12)$$

Here,  $d$  denotes the vector of empirical moments and  $s(p)$  denotes the vector of simulated moments depending on  $p$ , the vector of parameters to be estimated. For simplicity, we assume a weighing matrix  $W$  with  $\frac{1}{d^2}$  on the diagonal and zeroes elsewhere.

Table 3 presents the 33 moments to be matched. Taking the IMRs as given, we aim to match the values of Total Fertility Rate in the cities and in the countryside, as well as the distribution of the population between urban and rural areas. Hence, we use only the demographic dynamics to fix the parameters while we use the economic data for overidentification checks.

There are two main reasons why we do not try to replicate economic moments in addition to demographic moments. First, by minimizing the number of empirical moments to be reproduced, we reduce our computational times and guarantee the convergence of our minimization algorithms. Indeed, we are trying to replicate dynamic data in the context of a general equilibrium model with heterogeneous agents and multiple markets. This puts a lot of pressure on computational resources. Second, education and GDP per capita have to be redefined to allow for comparisons of data and theory. Allowing our code to choose the best definition is extremely costly in terms of computational time, while imposing a single arbitrary definition greatly limits the ability of the model to fit the data.<sup>27</sup> Section 4.4 shows that our model performs quite well for education,

---

<sup>27</sup>A simple example comes from the definition of primary school enrollment rates. See the discussion

GDP, and the relative prices.

Table 3: MOMENTS TO BE MATCHED BY THE MODEL

	1760	1780	1800	1820	1840	1860	1880	1900	1920	1940	1960
$n^R$	4.499	4.184	4.009	4.487	4.303	4.062	4.046	3.792	3.126	2.278	2.326
$n^U$	3.881	3.619	3.253	3.310	3.156	3.229	3.393	3.011	2.064	1.673	2.116
$\delta$	0.096	0.089	0.105	0.096	0.088	0.110	0.155	0.216	0.289	0.381	0.523

To obtain our empirical moments, we first interpolated and smoothed our data using the Hodrick-Prescott filter (Hodrick and Prescott, 1997). We then have selected observations of the smoothed time series for each 20 years from 1760 to 1960, as documented in Table 3.

To estimate our deep parameters, we follow the general strategy of Baudin *et al.* (2015b). We first apply the genetic optimization algorithm PIKAIA (Charbonneau, 2002) to search the region in the parameter space where the global maximum is located. Taking this information as representing the initial values, we identify the precise parameters by UOBYQA, an algorithm that applies Powell’s unconstrained optimization method (Powell, 2002). Contrary to Baudin *et al.* (2015b), in each parametric situation, we need to compute the general equilibrium dynamics of our economy. For each parametric condition explored by PIKAIA and UOBYQA, we need to equalize supply and demand on markets in order to find equilibrium prices and individual behaviors at the equilibrium. To do so, we first use the simulated annealing algorithm SIMANN (Goffe *et al.*, 1994) to determine the region of the global optimum, and then continue with FMINSEARCH; an algorithm based on Powell’s method and proposed by Fehr and Kindermann (2018). As a numerical algorithm is used inside a numerical algorithm, this numerical exercise is very demanding in terms of computational power; we then adjusted PIKAIA in order to allow for a dynamic parallelization, and apply a hybrid MPI-OpenMP programming strategy within Fortran 90, which we describe more deeply at the beginning of Appendix D.1.

### 4.3 Results

The strategy described above leads to the values presented in Table 4 for our 15 estimated parameters. If individuals want to leave their area of origin, they have to pay

---

on education in Appendix D.3.

1.89 units of agricultural goods. This migration cost has to be compared to the average income of an industrial worker, as in general people move from the countryside to the city. By 1760,  $\kappa$  represents 20.4% of the average income of industrial workers. At first glance, this cost seems to be high, but it is in line with the findings of Dribe (2003), who argues that in Sweden during this period, “*high costs of migration, (...) and lack of information on conditions in other places might have kept (...) migration from being a viable option in many cases.*” Indeed, in the specific case of pre-industrial European countries, asymmetries of information coupled with the low quality of the transportation infrastructure are key determinants of migration costs. More than two centuries ago, cities and the countryside were separate economic environments that followed different rules and different models of labor organization (see, for instance, Bairoch, 1997 and Mokyr, 2010). Adapting to these productive systems necessarily entailed costs that, for the sake of simplicity, we assume to be fixed and independent of the type of migration.

In the specific case of Sweden, where the distances between large cities are not small and the land topology is not always favorable, the low quality of the transportation infrastructure was a clear barrier to information acquisition. Berger and Enflo (2017) explained this situation in their enlightening study of the history of rail and transportation in Sweden. Interestingly, they found that the costs of transportation were very high and the transportation infrastructure (which was based on water) was highly inefficient before 1820-1850, but that these costs decreased sharply and rapidly following the construction of the first railroads in the country. In line with this description, we predict that the relative costs of migration decreased continuously, and represented only 5.4% of the average income of industrial workers by 1960.

The initial total factor productivity in the agricultural sector is 250.25, and is therefore more than 670 times higher than the total factor productivity in industries (0.37). This result is mainly driven by the normalization of the size of land to one. The critical population size to make technological progress appearing is equal to  $\bar{N} = 12,445$ , which means that the global technological progress starts as early as in 1820, and equals 1.7% per year.

The subsistence level ( $\bar{c}$ ) represents 7.1% of the predicted average income in the country by the year 1760, which given the inequalities that prevail in both areas, may represent a much larger share of individual income among poor households. The weight of agricultural goods within the utility function ( $\alpha$ ) is equal to 0.5, which is more than twice the weight of educated children ( $\rho = 0.23$ ). Each birth requires 4.0% of available

Table 4: PARAMETERS IN THE SWEDISH CALIBRATION EXERCISE

Parameter	Sym.	Value	Parameter	Sym.	Value
Costs of moving	$\bar{\kappa}$	1.888	Average HK in 1760	$\lambda$	1.773
Critical pop. size	$\bar{N}$	12445	Basic human capital	$\nu$	1.230
Growth rate of A	$g$	0.412	Preference for children	$\rho$	0.233
Preference for $c_t$	$\alpha$	0.501	Initial TFP in R	$B_0^R$	250.25
Cost of education	$\beta$	0.044	Initial TFP in U	$B_0^U$	0.372
Stone-Geary element for d	$\varepsilon$	1.549	Elasticity of HK	$\phi$	0.239
Child rearing cost in R	$\zeta_R$	0.110	Minimum consumption	$\bar{c}$	0.544
Child rearing cost in U	$\zeta_U$	0.145			

parental time. If the child survives, an additional 11% of this time is required in the countryside and an additional 14.5% of this time is required in a city to bring up the child. In the absence of infant mortality, this would lead to a maximal fertility of around seven children in the countryside and five children cities.

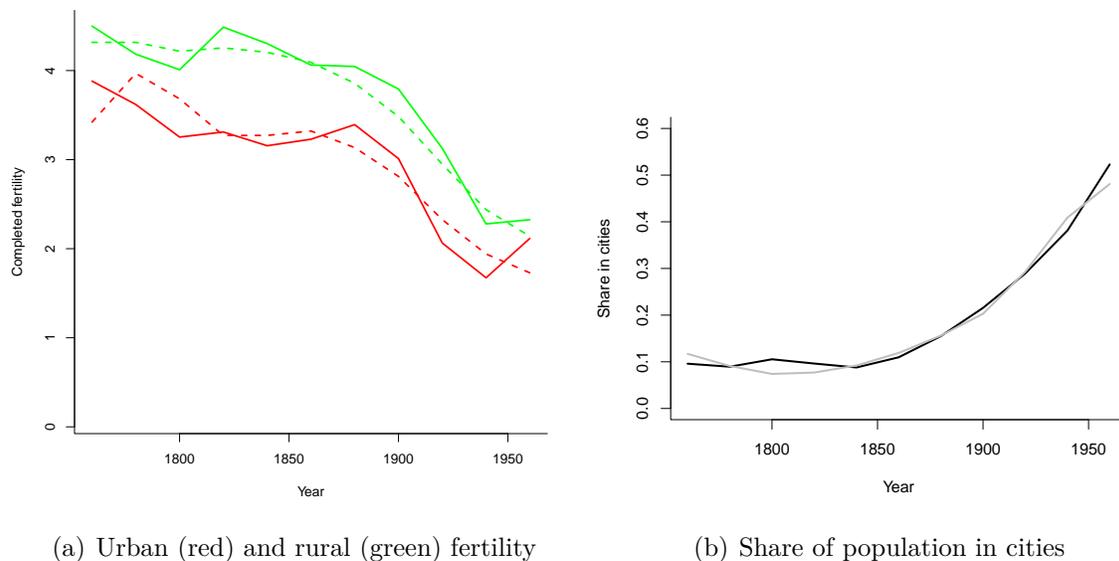


Figure 6: OBSERVED (SOLID) AND SIMULATED (DASHED) DEMOGRAPHIC DYNAMICS IN SWEDEN, 1760–1960

The parameter set enables our model to reproduce the characteristics of the Swedish demographic dynamics between 1760 and 1960 in a convincing way. Figure 6 shows these results in more detail. From 1760 to 1960, we predict that the average number of

births per women dropped from 3.4 to 1.7 in cities and from 4.3 to 2.1 in rural areas. On the whole, we predict almost perfectly the magnitude of fertility decrease in both areas.<sup>28</sup> Our model is able to explain 95.1% of the variation in fertility in rural areas and 82.3% of the variation in cities.<sup>29</sup> Regarding urbanization rates, we predict an increase of 36.4 percentage points, compared to the actual increase of 42.7 percentage points. Still, our model reproduces 97.9% of the variation in urbanization rates. The average distance between our empirical and simulated moments regarding urbanization rates is equal to 13.6% on average. In Appendix D.4, we deliver our estimation of the intensity of the rural exodus in Sweden. We estimate that it mainly prevented de-urbanization up to 1840 and then became the driver of urbanization.

#### 4.4 Overidentification checks

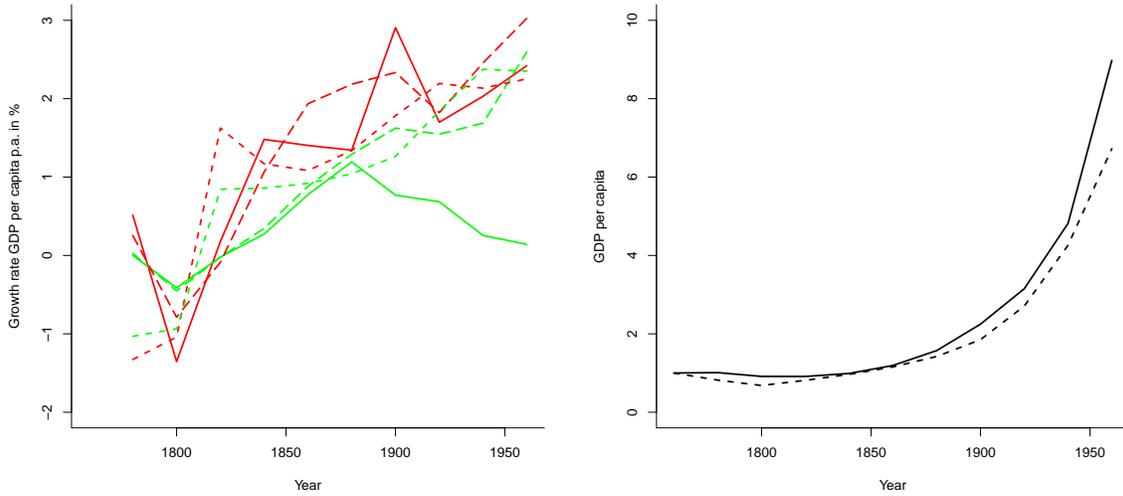
The good quality of our fit for demographic dynamics does not guarantee that the model delivers reliable results. Indeed, it may perform badly in predicting the economic and educational moments we do not try to match. There are four main series for which we have both data and predictions: average annual growth rates of GDP per capita in each area, relative prices, wages and enrollment rates into primary education.

The average distance between the predicted and the empirical growth rates of GDP per capita is somehow larger than it is for the demographic data. The distance equals 0.9 percentage points for the agricultural sector and 0.6 percentage points for the industrial sector. As is shown in Figure 7a, except for the initial period, our predicted industrial growth rates are quite realistic. Regarding agriculture, we tend to over-predict the growth rate of GDP per capita after 1880. This tendency toward over-prediction becomes stronger over time, and is mainly due to the fact that with the increase in urbanization after 1880, commuting habits emerged. Thus, estimating GDP per capita

---

<sup>28</sup>Our model predicts an increase in fertility between the two first periods; a result that is not in line with the data. This is due to the fact that the dynasties are not perfectly allocated across areas in the initial situation, and then reallocate themselves. We find that in 1760, the initial period in the simulation, some dynasties reallocated from the countryside to cities (see Figure 22 in Appendix D.4). 5.1% of young adults in rural areas left their place of origin and moved to cities, where they represented 50% of the young adults. These relatively poor dynasties were associated with a significant reduction in average fertility in cities. They had low levels of human capital, and paying  $\bar{\kappa}$  made them living in the Malthusian regime.

<sup>29</sup>These values correspond to the  $R^2$  of an ordinary least squares regression in which the observed total fertility rate is the dependent variable, and the predicted total fertility rate is the independent variable. Another way to evaluate the quality of the model's fit is to compute the average distance between the observed and the predicted moments in percentage points. In the case of fertility, the distance equals 10.8% in cities and 5.4% in the countryside. See Appendix D.2 for more discussion.



(a) Average annual growth rate of GDP per capita in agriculture (green) and industry (red)

(b) Total GDP per capita

Figure 7: OBSERVED (SOLID), ADJUSTED OBSERVED (LONG DASHED) AND SIMULATED (DASHED) GDP PER CAPITA AND ITS GROWTH RATES

in the rural sector by dividing total agricultural production by the size of the population living in rural areas becomes less justified. Thanks to the data of Centralbyran (1969) that we present in Table 7 of Appendix A.5, we are able to tackle this issue. The long dashed lines of Figure 7a display the evolution of the growth rate of GDP per capita after commuting and proto-industrialization are integrated. It is notable that the overestimation issue disappears and the difference between the predicted and the observed growth rate drops to less than 0.5 percentage points.<sup>30</sup>

As illustrated in Figure 7b, we reproduce quite well the dynamics of overall Swedish GDP per capita. Our model explains 99.1% of the observed variation, although the average deviation between predicted and observed moments is 15.8%.<sup>31</sup>

Our next counterfactual check concerns the dynamics of the prices of industrial goods relative to the prices of agricultural goods. As Figure 8b shows, the data we have collected for this check span the 1800–1960 period, or a shorter period of time than the data used for the other moments. To avoid scaling issues, we have normalized the

<sup>30</sup>We discuss this issue in more details in Appendix A.5.

<sup>31</sup>It is important to recall here that the value that we call the average GDP per capita in Sweden in 1960 is nothing more than the total production of the agricultural and industrial sectors translated into agricultural prices, and divided by the size of the Swedish population. This value is therefore a strongly underestimated version of the real GDP per capita of Sweden at that time.

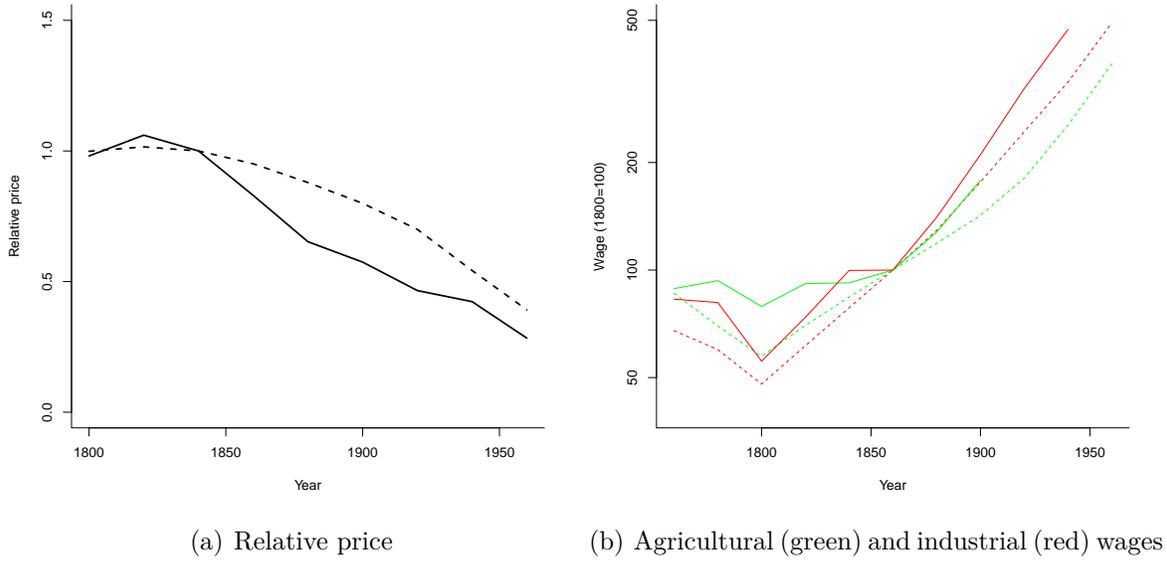


Figure 8: OBSERVED (SOLID) AND SIMULATED (DASHED) RELATIVE PRICES AND WAGES

predicted relative prices in 1800 by the observed prices. It is reassuring to see that our model captures most of the historical adjustments of relative prices. We are even able to reproduce the initial stagnation that preceded the industrial revolution. Our simulated relative prices explain 88.8% of the variation, even though the average distance between our prediction and the data equals 29.1%.

The last check regarding the economic dimension of facts concerns the dynamics of wages. Between 1760 and 1940, we can compare the estimated and predicted industrial wages. Our predictions are able to explain 98.7% of the variation. The time span in the estimated trends is shorter for the agricultural wages and only allows to compare estimated and predicted values between 1760 and 1900. With an average deviation of -14.4% the fit is also convincing. Finally, the predictions of the model reproduce the most important qualitative findings on the dynamics of normalized wages: the u-shaped pattern with their minimum in 1800 and the initially higher agricultural wages and the reverse order after 1860.

Regarding education, we first have to determine how the model might predict primary education enrollment rates, as what we predict for each individual is her investment in the education of her children, but not her children's primary school enrollment. We consider that a child is enrolled in primary school if she is receiving at least  $\bar{e} = 0.513$  units of education. We interpret this value as follows: by 1760, the average family

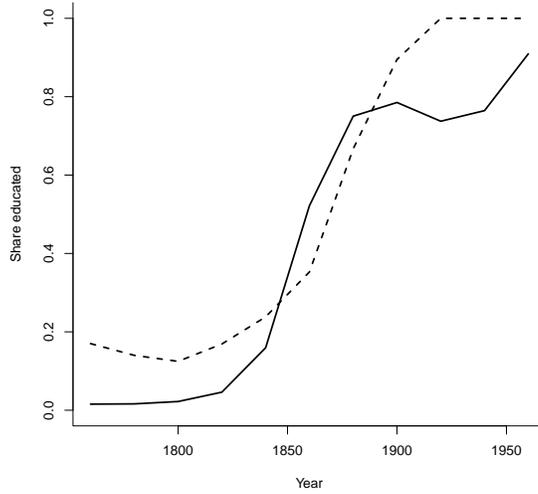


Figure 9: OBSERVED (SOLID) AND SIMULATED (DASHED) ENROLLMENT RATES INTO PRIMARY EDUCATION

would have to spend 2.74% of their income to allow one of the children to complete primary school, and 4.5%<sup>32</sup> to allow all of the children to finish primary school. As Sweden became richer while fertility and the price per unit of education declined, these amounts became smaller, declining to 0.17% and 0.16%, respectively, by 1960.

Given this last hypothesis, we are able to explain 89.6% of the variation in education. Figure 9 shows how we reproduced the S-shaped pattern of educational dynamics, despite our tendency to overestimate enrollment rates in the beginning and at the end of our period of observation.<sup>33</sup>

## 5 Historical experiments

In this section, we discuss the importance of three phenomena we believe are crucial in our theory: migration conditions, mortality factors, and the allocation of agricultural skills.

<sup>32</sup>This value corresponds to  $\frac{\beta * p_{1760} * \bar{e}}{\bar{y}}$  where  $\bar{y}_{1760}$  denotes the average GDP per workforce in Sweden in 1760.

<sup>33</sup>For transparency purposes, we show what our predicted dynamics would be under alternative assumptions about the value of  $\bar{e}$  in Figure 21 of Appendix D.3.

## 5.1 Restrictions on internal migration

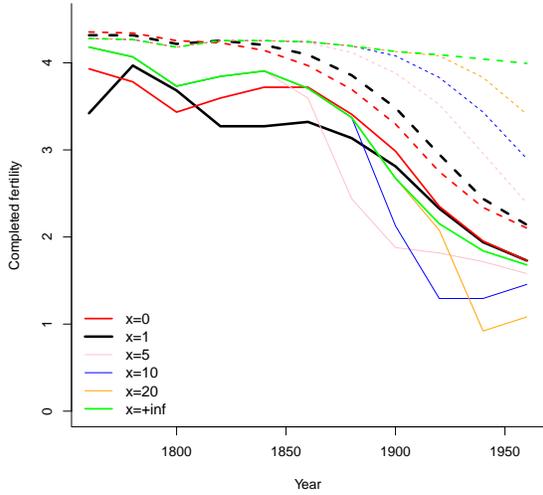
As we explained in Section 4.3, migration costs are deeply rooted in transportation infrastructure, topological conditions, and information asymmetries. Recent papers by De la Croix *et al.* (2017) and Ogilvie (2014) have also documented how institutions like guilds facilitated or hindered internal migration in Europe even before industrialization. In this section, we estimate the impact of restrictions on internal migration via variations in migration costs on the economic and demographic dynamics of Sweden. More precisely, we simulate our dynamic general equilibrium model for alternative values of  $\bar{\kappa}$ , such that  $\bar{\kappa}' = x\bar{\kappa}$  with  $x = \{0, \frac{1}{20}, \frac{1}{10}, \frac{1}{5}, 1, 5, 10, 20, +\infty\}$ . The results of these experiments are presented in Figure 10.

We begin with an extreme case:  $\bar{\kappa}' = 0$ , or total freedom of movement. Any person wanting to leave her area of origin is able to do so. As is shown in Figure 10 (solid red line), the overall impact of this change in  $\bar{\kappa}$  remains limited even if fertility becomes much higher in cities. It increases because cities welcome more newcomers from the countryside than in the benchmark scenario, and these newcomers do not have to pay the migration cost  $\bar{\kappa}$ , which increases their fertility. Hence, in the periods in which the rural exodus is weak, fertility in cities does not increase. This is mainly the case at the end of the 18<sup>th</sup> century. By contrast, fertility in the countryside declines slightly due to general equilibrium effects: as the intensity of the exodus becomes stronger, the labor supply in the countryside becomes lower, which increases the wage rate per unit of efficient labor. As a result, the fertility of the people not living in the Malthusian regime decreases, while the fertility of the people trapped in this regime increases. The first effect dominates the second. Finally, urbanization would have accelerated and the educational revolution, as well. When analyzing the less extreme cases in which  $\bar{\kappa}$  is divided by either five, 10, or 20, we find the same results as for  $\bar{\kappa}' = 0$ , but with smaller magnitudes, see Figure 23 in Appendix D.5.

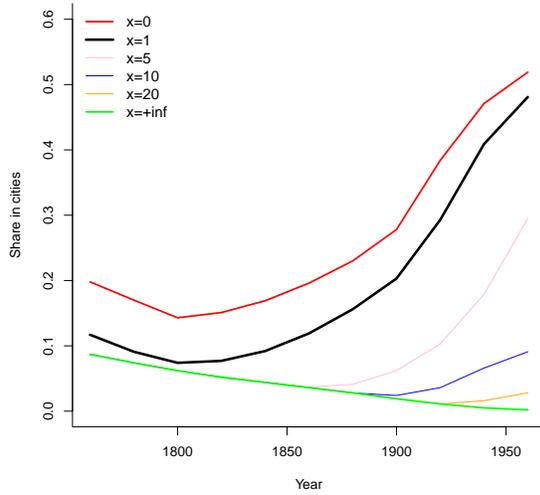
In our second extreme scenario,  $\bar{\kappa}' \rightarrow +\infty$  such that no migration between the two areas is possible (Figure 10, green lines). A first important result is that no fertility transition would have occurred until 1960 in the countryside, as the TFR would have been only 6.7% lower than in 1760.<sup>34</sup> Supplemented by the reduction of mortality rates, the absence of rural exodus would have led to an explosion in the size of the rural population. By 1960, this population would have had a GDP per capita that was 4.1 times lower than that of the benchmark scenario. From 1760 to 1880, fertility in cities

---

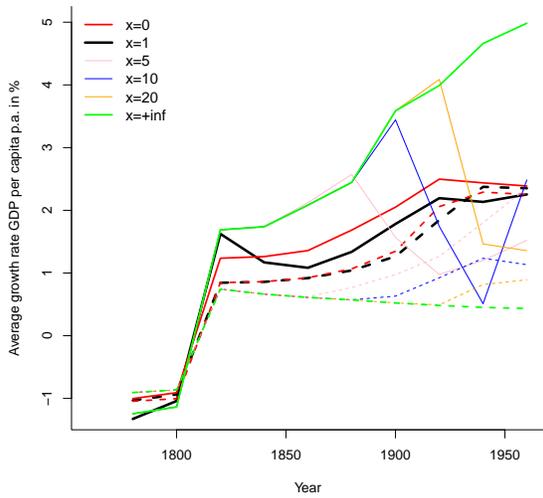
<sup>34</sup>This does not mean that the fertility transition would not have occurred at all as we can identify an inverse U-shaped dynamics of fertility in the countryside when  $\bar{\kappa}' \rightarrow +\infty$ .



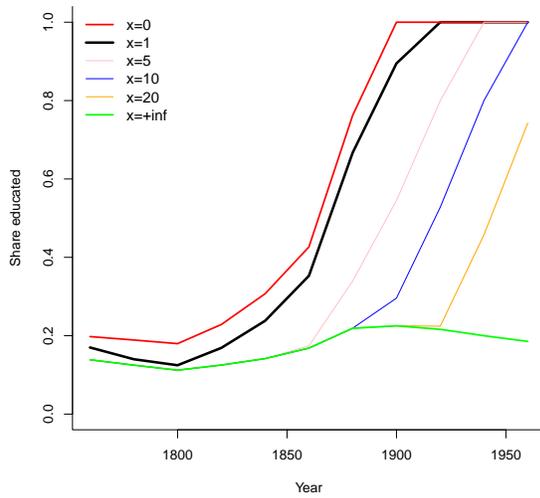
(a) Urban (solid) and rural (dashed) fertility



(b) Urbanization



(c) Industrial (solid) and agricultural (dashed) growth rates of GDP per capita



(d) Education

Figure 10: SWEDISH DEMOGRAPHIC AND ECONOMIC DYNAMICS UNDER ALTERNATIVE MIGRATION SCENARIOS IN 1760–1960

would have been higher than in the benchmark. This is due to the absence of newcomers from the countryside, as most newcomers would have been living in a Malthusian regime with a low fertility rate. Nevertheless, it is important to note that the fertility transition would have occurred in cities, and it would have been more intense. Indeed, compared to the Benchmark, the accumulation of human capital would not have suffered from the presence of the pauperized newcomers from the countryside. It would have led to higher growth rates of GDP per capita, and thus to greater adjustments in fertility. Not surprisingly, imbalances in the natural movements of the population would have resulted in a de-urbanization (see Figure 10b), with a small elite living in cities and enjoying incomes that were 10.9 times higher than in the benchmark scenario.

Obviously, a scenario without any migration between the two areas is far from realistic, but it gives us a good idea of the forces at play when the costs of migration are multiplied by five, 10, and 20. Multiplying migration costs by five, 10, and 20 may seem excessive. However, the average real GDP per capita in Sweden has multiplied by slightly more than 22 between 1760 and 1960. This implies that we only test scenarios in which the relative costs of migration would have receded or, at worst, stagnated in real terms. Multiplying the costs of migration by five would have delayed the fertility transition in the countryside by 20 years, while multiplying it by 20 would have delayed it by almost a century.

Turning to cities, we see that the effect on fertility is rather dramatic. When  $\bar{\kappa}'$  increases, fertility is initially higher than in the benchmark scenario because, as explained above, many children of farmers who were supposed to move to cities and depress the average fertility rate, are not there anymore. But once the rural exodus takes place (around the middle of the 19<sup>th</sup> century, when  $\bar{\kappa}$  is multiplied by five), fertility is strongly reduced precisely for the reason mentioned previously: as the newcomers are poorer than in the benchmark scenario, their fertility is considerably lower. Higher migration costs are associated with a more severe pauperization of the low-skilled in cities, and thus with heavily depressed fertility in this population. For this reason, the higher the  $\kappa'$ , the later and the more intense the fertility transition in cities and the stronger and more persistent the fertility differential between cities and the countryside. This last result echoes remarkably well these of Daudin *et al.* (2018) who show, in the case of French departments, that the railway transportation revolution reducing migration costs has contributed to the reduction of fertility differentials between urban and rural areas.

Not surprisingly, increases of  $\bar{\kappa}$  greatly limit urbanization, and delay the educational revolution. The growth rate of GDP per capita responds according to the previous

mechanisms: it is initially higher in cities as the pauperized newcomers have not yet arrived, and then it becomes lower under the effect of magnified pauperization. In the countryside, growth is systematically lowered: as fewer people are able to move to cities, average productivity is reduced.

This numerical exercise clearly shows that the rural exodus does affect more than just urbanization. While a missing rural exodus expedites the demographic transition in urban areas, the transition simply disappears in rural areas. The rural exodus also clearly supports the transition to mass education and a more egalitarian distribution of welfare, as its absence would limit the economic take-off to a small elite in cities. Thus, the rural exodus acts above all as a mechanism that prevents inequality levels from exploding by serving as a vector of economic opportunities for the “talented mass” from the countryside.

## 5.2 Mortality crisis

In line with papers like Lagerlöf (2003), Weisdorf (2004), Strulik and Weisdorf (2014), and Bar and Leukhina (2010), we discuss the impact of IMRs on the demographic and economic dynamics of Sweden. Our main objective here is to evaluate the role that our theory gives to mortality reductions in the take-off to modern economic growth and the fertility transition. For this purpose, we analyze what the dynamics of Sweden would have been if IMRs had stagnated instead of receding.

Keeping mortality rates constant in both areas at their 1760 level has a strong positive effect on the average number of birth, especially in cities, as Figure 20 shows in Appendix C.1. This is mainly due to a replacement effect, as the TFR counts all deliveries. We predict that the stagnation of IMRs would have increased the TFR by 30.9% in cities and 15.7% in the countryside in 1960. Nevertheless, the number of surviving births would have declined by 12.6% and 6.6%, respectively. As a result, the overall population size would have been reduced by 30% after two centuries of mortality stagnation.<sup>35</sup> We find only a very limited impact of mortality stagnation on both urbanization and education rates; see Figures 20a and 20b. The positive Malthusian checks reducing the net number of births are also not strong enough to postpone the onset of technological progress. They tend to foster growth rates of GDP per capita in both areas, even if this impact is quite contained.

---

<sup>35</sup>In Tables 8 and 9 in Appendix C.2, we provide the series of net changes in the number of births in each area.

In Appendix C, we try to determinate whether a major mortality crisis such as the Black Death could have prevented the take-off to modernity. Nevertheless, even in these extreme cases, mortality crisis are unable to alter the agenda of the economic and demographic Swedish revolutions.

### 5.3 Agricultural productivity shocks

As documented by Bairoch (1997), the English agricultural revolution led to improvements in the productivity of a few agrarian workers through innovations and the effects of the enclosure acts, but it resulted in lower productivity for the majority of agrarian workers. In Sweden, the agricultural revolution started with the *Storskifte Enclosure Act* of 1749, which redistributed land to farmers. As Svensson (2006) showed, this institutional movement was accompanied by the radical *Enskiftet* of 1803-07, which consolidated landholdings. Such reforms led to increased levels of inequality between farmers, as some received highly productive pieces of land, while others were not as lucky. It is important to note, however, that these redistributive reforms might have not modified the aggregate productivity of agricultural labor.

In this section, we simulate the impact of a shock on agricultural productivity that would preserve the mean productivity of the population, but would increase its variance (mean-preserving spread). Such a shock echoes the effect of the agricultural revolution. We mainly want to determine whether those higher levels of inequality in the countryside may have accelerated the transition to modern economic and demographic regimes, as the push theory of the rural exodus assumes.<sup>36</sup> From a technical point of view, we simply implement the following transformation of the distribution of agricultural abilities  $a_i$ :

$$a_i' = \frac{a_i^n \sum_{i=1}^{\bar{N}} a_i}{\sum_{i=1}^{\bar{N}} a_i^n}, \quad \text{with } n = \{2, 4, 6\}$$

where  $\bar{N}$  denotes the initial adult population size in our model. The higher  $n$ , the more unequal the distribution of skills.

Regarding fertility, *in the initial period*, worsened economic conditions push a large number of people out of the countryside, causing urbanization rates to increase sharply. The people who left the countryside are typically those who have experienced productivity

---

<sup>36</sup>Scenarios with a more equal distribution of agricultural skills are displayed in Appendix D.6.

losses but who are productive enough in the industrial sector to cover their migration costs. Because these individuals have to cover their migration costs and their human capital is relatively low,<sup>37</sup> they remain poor even after moving to the city, which limits their fertility. Thus, Malthusian positive checks are at play.

When we look at the countryside, we see that three kinds of agents remain after the shock. The first group is made up of people with high agricultural abilities, and who thus benefited from the productivity shock. They live in the interior regime, and as their wages have increased, their fertility has decreased. The second group consists of those individuals who have suffered from the shock and cannot afford to pay migration costs. They live in the Malthusian regime, as their wages and fertility have been sharply reduced. Thus, most of these people are too poor to have children. They live in Regime I of Table 1 and their dynasty will disappear at the next period.<sup>38</sup> The third group consists of those who suffered limited productivity losses and live in the interior regime. Their situation did not deteriorate enough to cause them to leave the countryside and move to a city. They tend to increase their fertility in response to the decline in their wages. As the two negative effects are stronger than the positive effect, fertility in the countryside decreases strongly.

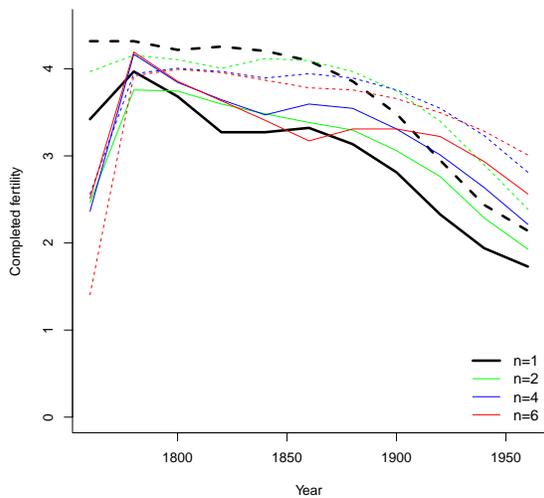
*In the subsequent periods*, another positive Malthusian check is at play: as the poorest rural dynasties have been childless, the fertility levels of the survivors become increasingly close to the benchmark values. When we look at fertility in cities, we see that fertility levels even exceed the benchmark values. This is mainly due to the fact that after the big push that occurred in the previous period, almost no rural exodus takes place; see Figure 28 in Appendix D.8. As no one in the urban population has to pay  $\bar{\kappa}$ , fertility is higher than in the benchmark scenario. The rural exodus remains persistently low or even absent during the next century, which leads to higher fertility levels. In addition, the fertility transition in both areas is postponed and dampened due to lower average incomes in both areas compared to the benchmark. In particular, starting in the middle of the 19<sup>th</sup> century, the weakened quality-quantity trade-off hinders a rapid fertility transition.

Regarding GDP per capita, it increases strongly in both areas *during the first period*. This effect comes from endemic childlessness, depressed fertility, and the mass exodus to cities, which cause rural populations to decrease in size and marginal labor productivity

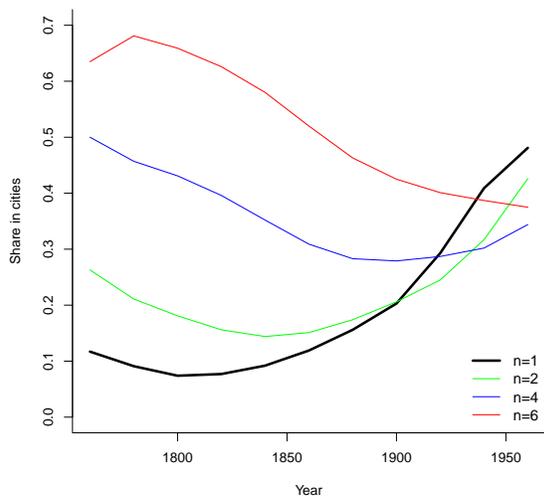
---

<sup>37</sup>In the benchmark scenario, these dynasties decided to stay in the countryside, which is a sign of a relatively small stock of human capital.

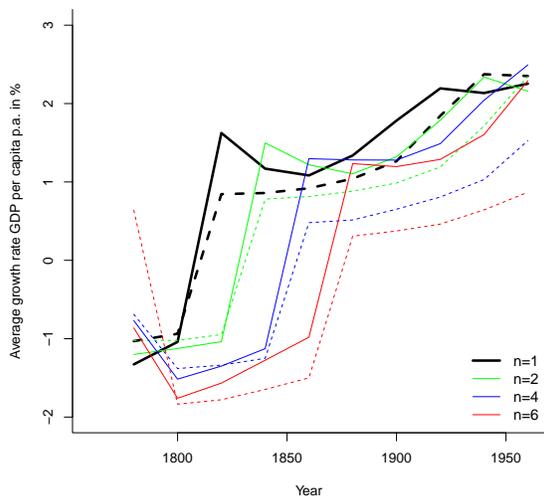
<sup>38</sup>This fact is documented extensively by Baudin *et al.* (2015a) for developing countries.



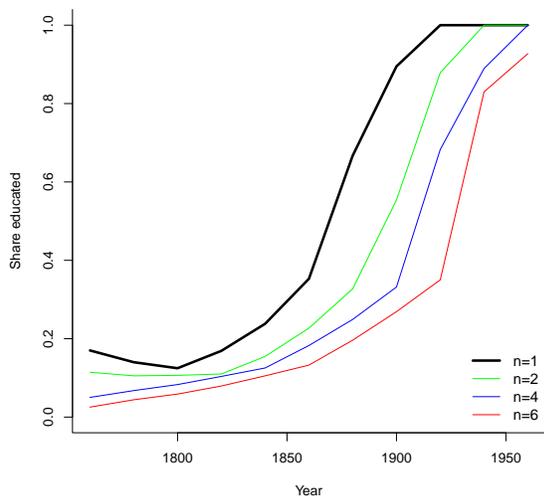
(a) Fertility: cities (solid), countryside (dashed)



(b) Urbanization



(c) GDP per capita in industry (solid) and in agriculture (dashed)



(d) Education

Figure 11: SWEDISH DEMOGRAPHIC AND ECONOMIC DYNAMICS WITH INCREASING INEQUALITIES IN AGRICULTURAL ABILITIES IN 1760–1960

to increase; a pattern documented by Voigtländer and Voth (2012). In cities, the very large population of low-skilled migrants causes the average level of GDP per capita to decrease. As the size of the adult population participating to the workforce increases while average fertility declines, the dependency ratio shrinks, thereby increasing average GDP per capita. As the second effect dominates the first effect, the initial level of GDP per capita rises above the benchmark.

However, this initial increase is followed by a long period of stagnation or even of recession in the case in which  $n$  is equal to four or six. This is a classical Malthusian effect: as the size of the population has been strongly reduced, the higher income of people translates into higher fertility rates and thus into higher dependency ratios, which in turn contribute to the persistence of low or even negative growth rates of GDP per capita. The reduced population size (below  $\bar{N}$ ) traps the economy into the Malthusian regime up to 1880, when  $n = 6$ . If in all of our scenarios the growth rate of GDP per capita finishes by reaching that of our benchmark in 1960, the effects in terms of income levels remain strong. When  $n = 2$ , the level of GDP per capita in 1960 is reduced by 34.9% in rural areas and 37.4% in cities. The lower income levels postpone the transition to mass education.

Interestingly enough, in our scenario enclosures turned out to be key determinants of the intensity and the timing of the rural exodus. Nevertheless, the more severe those movements were, the later the onset of the economic take-off, and the poorer the country became in the long run. Stated a bit differently, this means that the economic and demographic pauperization of migrant rural populations was a brake on, rather than an engine of, the industrial revolution.

It is also interesting to notice that the persistent economic recession caused by the increase of inequalities is accompanied by high urbanization rates. In the case where  $n = 6$ , urbanization rates maintain above 60% while gdp per capita recedes over almost a century. In this situation, urbanization of the economy has not been accompanied by income growth and industrialization.

## 6 Discussion

In tradition with the UGT, we present a fairly general model with two areas and sectors to emphasize the role of the rural exodus for the arisen of Sweden as representative for Western Europe. Without adjusting the general framework to country specifics,

we quantify its mechanisms for the history of Sweden. Neglecting specific policies or institutional changes may raise some concerns about the validity of our predictions for Sweden, while the limitation to Swedish data may question the validity of our predictions outside Sweden. In this section, we show in which dimensions Sweden is emblematic of other European countries at the same period. We also discuss the role of some important institutional and policy changes in Sweden. We then explore some additional key historical facts regarding Europe as well as more contemporary data in today's developing countries.

## 6.1 Was Sweden a typical European country?

If the demographic empirical facts, we have reproduced to calibrate our model, come from Swedish data exclusively, they are symptomatic for Western Europe between 1700 and 1950. First, the higher urban than rural mortality has been widely documented at the European level, see for instance Knodel (1974), Bairoch (1997) or Woods (2003). Regarding fertility, Mosk (1980) in the case of Sweden and Japan and more recently Bocquier and Costa (2015) in the case of Belgium and Sweden and Bocquier and Brée (2018) in the case of France have evidenced that rural populations were more fertile than their urban counterparts. The number of births per woman in rural areas exceeded those in urban areas. All these papers have established that despite these unfavorable natural population movements, urbanization nevertheless occurred. Since neither international immigration nor emigration were able to sustain urbanization in European countries at that time, urbanization was only possible because of a massive and sustained rural exodus. Finally, in their enlightening contributions, Bertinelli and Black (2004) and Bairoch (1997) have shown how cities were the natural ecosystem of the industrial revolution. They made a case for urbanization as a necessary condition for industrialization. Not surprisingly, industrial prices relative to agricultural prices decreased all along the European industrial revolution, see for instance Knick (1998) for textile prices in England and Mokyr (1977) for a more general discussion.

This said, it would be inexact to claim that Sweden was not characterized by some unique features. As documented by Schön (2010), during the second part of the 18<sup>th</sup> century, when the Industrial Revolution began in Britain, Sweden was economically less advanced compared to Great-Britain or Belgium for instance. While the steam engine was increasingly important in the European industrial production, it played a minor role in Sweden's relatively small industrial sector that relied more on water power.

Like many other European countries, the Industrial Revolution of Sweden took place in two phases. Before 1850, the industrial development was based on “a combination of iron core deposits and access to charcoal” (Schön, 2010) as well as the textile industry. In this first phase of the Industrialization, the still agrarian economy was shaped by the role of domestic production and a step-wise regional specialization. During the second phase of industrialization, this specialization has encouraged the transport sector and the arisen of exchange and trade via markets. Sweden became a small open economy and, hence, sensitive to price shocks in both directions. Like in other continental European economies the consequences of the Napoleonic wars have been large in the short run as it slowed down industrial expansion.

Beside the arisen of markets another important institutional change started around in 1749. As in Britain, the Enclosure Act redistributed cultivated land and enabled growth in the agricultural sector. Another wave of step-wise enclosures occurred between 1803 and 1807. Finally, villages without a Enclosure until 1827 did so after the new Act in 1827. As documented by Rozental (1956), the Swedish enclosure movement is part of a larger agrarian reform favoring enclosure in France, Germany, Spain, Italy and Holland.

The education reform of 1842 corresponds to the end of the first phase of the industrial revolution. With the construction of railroads in the 1850s, Sweden entered in sustained economic growth like the forerunners of industrialization. Between 1870 and 1910 the average growth rate of gdp per capita was even higher than in other Western European Countries such as Germany, Denmark, the Netherlands or the UK (Berger, 2016). Beside growth in TFP, an increasing capital stock was also a driving force behind growth at that time, a reality that our model cannot reproduce

We conclude that Sweden was part of the wider “European world economy”. The Swedish dynamics can be placed between the British pioneer and most continental European countries which were hit much more by the Napoleonic wars around 1800. Like all other economies Sweden dynamic is featured by its own climate, geographic and cultural situation, but also by the general European context. Dynamics of price altered exports and imports of the small open economies and Sweden was directly and indirectly involved in political conflicts. As its context can be regarded to be no more or less representative than most other European countries we rely our quantitative exercise on its data thanks to its exceptional data availability.

## 6.2 Upper-tailed human capital, wages versus gdp per capita

There are at least two facts at the European level that we did not emphasize in the case of Sweden so far. The first one concerns the accumulation of upper-tailed human capital. In their influential paper, Squicciarini and Voigtländer (2014) evidence how the accumulation of upper-tailed human capital has been a prerequisite for the accumulation of human capital and growth at the scale of the entire population. If we do not tackle this issue explicitly, our model delivers interesting predictions regarding this phenomenon. We have computed the Gini index characterizing the distribution of human capital in our artificial population. We can distinguish three periods. In the first period, before the industrial revolution from 1760 to 1840, a Gini coefficient of 0.174 indicates strong inequalities in the distribution of human capital.<sup>39</sup> This latter is accumulated only among the rich dynasties. Rich dynasties are either located in cities where they are already well endowed in terms of human capital or in the countryside, meaning that the agrarian elite also accumulates human capital. A second phase appears, right after the “Folkskola” reform of 1842 making primary school compulsory for four years; at that time, industrialization and the rural exodus accelerated. Between 1860 and 1900 the average Gini Index was 0.044. Hence, inequalities in the distribution of human capital decrease dramatically what corresponds to the phase where the mass get access to education and human capital. With an average Gini of 0.063 the period between 1920 and 1960 corresponds to higher inequalities but these latter remain well beyond those existing before the Industrial Revolution.

Interestingly enough, our simulation exercises make a strong prediction about the findings of Squicciarini and Voigtländer (2014) on the role of upper-tail human capital. A negative shock on labor mobility would have prevented the accumulation of upper-tailed human capital to translate into the accumulation of human capital for the entire population. Indeed, we have evidenced that when the migration cost becomes very high, human capital accumulates in small cities where an educational elite lives while the majority of the population lives in the countryside without being educated. Such a movement is associated with the explosion of income inequalities between cities and countryside.

The second fact we want to discuss emerges mainly from the a series of key contributions by Clark (Clark, 2002, 2007). Clark has documented the evolution of wages in Great

---

<sup>39</sup>To compute these Gini coefficients, we have first computed a Gini Index for each period of simulation and we have then averaged these values under the assumption that each generation has the same weight.

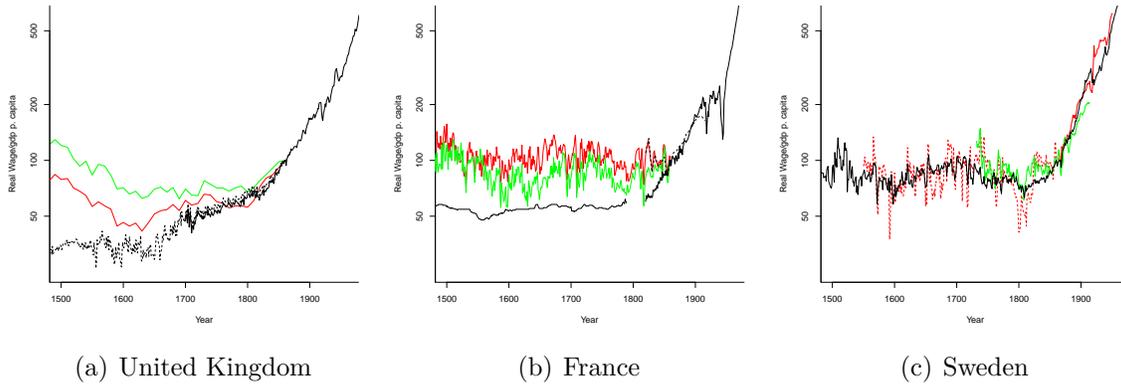


Figure 12: DYNAMICS OF WAGES AND GDP PER CAPITA IN SWEDEN, GREAT-BRITAIN AND FRANCE

Fig. 12a: wages of craftsmen (plain red) and farmers (plain green) with data from Clark (2002). Fig. 12b: with wages from building laborers (red) and agricultural laborers (green) from Ridolfi (2017) and on average (plain black) from Lévy-Leboyer and Bourguignon (1990). Fig. 12c: wages for unskilled laborers until 1864 in Stockholm (dashed red) and industrial workers in Sweden since 1865 (plain red) from Söderberg (2010). Male agricultural wages from Jörberg (1972) deflated with Edvinsson and Söderberg (2010). Gdp per capita in black (dashed black for England) from Bolt *et al.* (2018).

Britain before and during the Industrial Revolution at the sectoral level. As shown in Figure 12a, the average salary of Craftsmen and Farmers embraced a long run U-shaped dynamics. After being divided by almost 2 between 1500 and 1650, wages started to increase smoothly (Craftsmen) or stagnated (Farmers) until 1800 and then exploded. As highlighted in Figure 12b, such a long run pattern of wages cannot be found in France. Until the end of the 18<sup>th</sup> century wages of both agricultural and building laborer varied in a limited way. Around 1855 average real wages in France entered a period of sustained growth. Sweden interlinks elements of the British and French case. In Figure 12c, we plot wages for unskilled labor in Stockholm until 1864 supplemented by the industrial wages as time series on industrial wages from Söderberg (2010) and male agricultural wages from Jörberg (1972). The latter has been deflated according to Edvinsson and Söderberg (2010). The onset of sustained growth is close to the French one in the middle of the 19<sup>th</sup> century. However, like in the British case, the normalized agricultural wages tend to be higher than the industrial ones in the period of stagnation.

From the data displayed in Figure 12, one key fact emerges: the correlation between sectoral wages and gdp per capita embraces a U-shape pattern through time in all the three countries. In Table 6.2, we display the evolution of this correlation between three main periods: 1500-1650, 1651-1800 and 1801-1960. If after 1800, France, Sweden and

Country	Correlation with	1500-1650	1651-1800	1801-1960
England	Craftsmen’s Wages	0.517	0.137	0.713
England	Farmers’ Wages	0.506	0.016	0.781
France	Building workers’ Wages	0.205	-0.178	
France	Farm laborers’ Wages	-0.001	-0.028	
France	Average Wages			0.893
Sweden	Farm Wages		0.669	0.974
Sweden	Unskilled Laborers’ Wages	0.657	0.408	0.683
Sweden	Industrial Workers’ Wages			0.951

Table 5: CORRELATION BETWEEN AVERAGE GDP PER CAPITA AND VARIOUS SECTORAL WAGES IN FRANCE, ENGLAND AND SWEDEN BETWEEN 1500 AND 1960.

Data from Clark (2002), Ridolfi (2017), Lévy-Leboyer and Bourguignon (1990), Söderberg (2010), Jörberg (1972), Edvinsson and Söderberg (2010) and Bolt *et al.* (2018). England until 1860, French average wages 1820–1915, French farm and building workers’ wages until 1789, unskilled laborers’ wages in Sweden from 1801–1864, Swedish industrial workers’ wages 1865–1950, Swedish farm Wages 1732–1800 and 1801–1914.

England are characterized by a strong positive linear correlation between real wages and gdp per capita, situations are more contrasted in earlier times. England has been marked by a low correlation between the end of the Thirty Years War (1648) and the beginning of the period of sustained growth. In the same period, the correlations are even slightly negative in France. If we also evidence the same U-shape pattern in Sweden, one difference emerges: the correlation remains higher than in the two other countries. To the best of our knowledge, these variations of the correlation between wages and gdp per capita through time and space have never been studied systematically. The absence of physical capital from our model may prevent us to delivering such an investigation but this latter remains far beyond the scope of this paper.

### 6.3 Urbanization without growth

Finally, we would like to highlight predictions of our model regarding the interplay of urbanization and industrialization. The specified model and the baseline scenario are in line with the traditional view of “industrial cities”. By contrast, in the counter-factual exercise related to shocks increasing inequalities of productivity in the agricultural sector, urbanization arises without economic growth and industrialization. Such a result has also been evidenced in modern developing economies by Gollin *et al.* (2016) and Fay and Opal (1999). With the emergence of “consumption cities” in resource-exporting

economies, urbanization does not require an industrialization. Gollin *et al.* (2016) argue that the love for variety in consumption attracts people from the countryside to move into cities . Our explanation for the absence of the association between urbanization and growth is more aligned with Fay and Opal (1999) as it is driven by inequalities in the countryside as well as urban-rural income gap drive the rural exodus.

## 7 Conclusion

Based on the unified growth theory, we present a dynamic general equilibrium model with two areas, two goods and an explicit migration decision. Applying the simulated method of moments, we are the first to use a full-fledged calibration-simulation strategy in a unified theory to evaluate the role the rural exodus played in the economic and demographic development of Western Europe.

In the specific cases of Sweden, we show that favorable conditions of exodus were necessary for both the fertility transition and the economic take-off. We estimate that multiplying the cost of migration by five or even 20 would not have been sufficient to prevent the take-off to modern growth, but that it would have delayed it and weakened its intensity. Higher migration costs would have also magnified the inequalities between cities and the countryside, which were initiated by the Industrial Revolution. While GDP per capita in cities would have increased, farmers would have severely impoverished. The pattern of the fertility transition would have unfolded differently in cities than in the countryside. While it would have been delayed in both areas, it would have been weakened in rural areas and intensified in urban areas. We also show that making the exodus even more costly would have simply put a stop to development; at some point, the countryside would have pauperized, while a vanishing elite would have haunted dying cities.

Like for the rural exodus, we find that the agricultural revolution crucially altered the economic and demographic dynamics. Increasing levels of inequality in agricultural productivity also delayed the economic-take off and weakened the fertility transition, while a more equal distribution of skills had rather limited effects. More importantly, it seems that pauperizing the rural population and pushing them out of the countryside through enclosures may not have helped industrial cities to escape the Malthusian regime. On the contrary, it would have delayed and weakened this movement. It may even lead to situations where urbanization becomes high while economic growth is

absent if not negative, an empirical regularity met in today's Sub-Saharan Africa and South-East Asia. Finally, the effects of infant mortality crises and IMRs more generally are no more than modest.

While our quantitative investigation studies the Swedish economy, we show that Sweden was a typical example for a Western European economy at that time. Most demographic, economic and educational facts we document for Sweden hold across European countries, like the UK, France or Denmark, see Baudin and Stelter (2019) for the latter. Thus, we are convinced that the main conclusion on the role of migration conditions, infant mortality and the agricultural revolution are generalizable in the European context. Still, country specifics, such as the pioneering role of Britain in the Industrialization of Europe or the Revolution and the Napoleonic War in the French case, likely altered timings and magnitudes of the presented mechanisms.

From a methodological point of view, our investigation of the rural exodus illustrates how advanced structural estimation methods and counterfactual experiments in dynamic general equilibrium settings may help in evaluating the accuracy of refinements of the UGT. Finally, it is quite interesting to note that the freedom to move between cities and the countryside is not guaranteed everywhere today. For instance, as documented by Chun-Chung and Henderson (2006), China still imposes restrictions on migration from the countryside to cities. Estimating the cost of this policy in the specific context of the one-child policy and its recent relaxation is on our research agenda.

## References

- ADAMOPOULOS, T. (2008). Land inequality and the transition to modern growth. *Review of Economic Dynamics*, **11** (2), 257–282.
- ADAMS, J. J. (2017). Urbanization, long-run growth, and the demographic transition, mimeo.
- BAIROCH, P. (1997). *Victoires et déboires: histoire économique et sociale du monde du XVIIe siècle à nos jours*, vol. I, II and III. Gallimard.
- BAR, M. and LEUKHINA, O. (2010). Demographic transition and industrial revolution: A macroeconomic investigation. *Review of Economic Dynamics*, **13** (2), 424–451.
- BAUDIN, T., DE LA CROIX, D., GOBBI, P. *et al.* (2015a). Development policies when accounting for the extensive margin of fertility. *Discussion Papers (IRES-Institut de Recherches Economiques et Sociales)*, **2015003**.
- , — and GOBBI, P. E. (2015b). Fertility and childlessness in the united states. *The American Economic Review*, **105** (6), 1852–1882.
- and STELTER, R. (2019). Danish urbanization and the demographic transition, mimeo.

- BERGER, T. (2016). Engines of growth: essays in swedish economic history.
- and ENFLO, K. (2017). Locomotives of local growth: The short-and long-term impact of railroads in sweden. *Journal of Urban Economics*, **98**, 124–138.
- BERTINELLI, L. and BLACK, D. (2004). Urbanization and growth. *Journal of Urban Economics*, **56** (1), 80–96.
- BOCQUIER, P. and BRÉE, S. (2018). A regional perspective on the economic determinants of urban transition in 19th-century france. *Demographic Research*, **38**, 1535–1576.
- and COSTA, R. (2015). Which transition comes first? urban and demographic transitions in belgium and sweden. *Demographic Research*, **33**, 1297–1332.
- BOLT, J., INKLAAR, R., DE JONG, H. and VAN ZANDEN, J. (2018). Rebasings ‘maddison’: new income comparisons and the shape of long-run economic development, gGDC Research Memorandum 174.
- BOUCEKKINE, R., DE LA CROIX, D. and LICANDRO, O. (2002). Vintage human capital, demographic trends, and endogenous growth. *Journal of Economic Theory*, **104** (2), 340–375.
- BRÄNDSTRÖM, A. (1993). Infant mortality in sweden—1750-1950—past and present research into its decline. In C. Carlo A and P. P. Viazzo (eds.), *The decline of infant mortality in Europe? 1850–1950: four national case studies*, pp. 19–34.
- BRÄNDSTRÖM, A. and SUNDIN, J. (1981). Infant mortality in a changing society: The effects of child care in a swedish parish 1820-1894.
- CENTRALBYRÅN, K. S. (1914). Statistisk årsbok för Sverige 1914. *Stockholm (and subsequent editions)*.
- CENTRALBYRAN, S. (1969). Historisk statistik för sverige. *Del I. Befolkning. Andra upplagan 1720–1967*.
- CERVELLATI, M., MURTIN, F. and SUNDE, U. (2013). Unified growth empirics: 1880-2000.
- and SUNDE, U. (2005). Human capital formation, life expectancy, and the process of development. *The American Economic Review*, pp. 1653–1672.
- and — (2011). Life expectancy and economic growth: the role of the demographic transition. *Journal of Economic Growth*, **16** (2), 99–133.
- CHARBONNEAU, P. (2002). An introduction to genetic algorithms for numerical optimization. *NCAR Tech. Note TN-450+ IA*, 74pp.
- CHUN-CHUNG, A. and HENDERSON, V. (2006). How migration restrictions limit agglomeration and productivity in china. *Journal of Development Economics*, **80** (2), 350–388.
- CLARK, G. (2002). Land rental values and the agrarian economy: England and wales, 1500–1914. *European Review of Economic History*, **6** (3), 281–308.
- (2007). The long march of history: Farm wages, population, and economic growth, england 1209–1869. *The Economic History Review*, **60** (1), 97–135.
- , CUMMINS, J. and SMITH, B. (2012). Malthus, wages, and preindustrial growth. *The Journal of Economic History*, **72** (02), 364–392.

- COALE, A. (1969). The decline of fertility in europe from the french revolution to world war ii. In C. L. J. Behrman, S.J. and R. Freedman (eds.), *Fertility and Family Planning: A World View.*, University of Michigan Press, pp. 3–24.
- DAUDIN, G., FRANCK, R. and RAPOPORT, H. (2018). Can internal migration foster the convergence in regional fertility rates? evidence from 19th century france. *The Economic Journal*, **Forthcoming**.
- DE LA CROIX, D. and DOEPKE, M. (2003). Inequality and growth: why differential fertility matters. *The American Economic Review*, **93** (4), 1091–1113.
- DE LA CROIX, D., DOEPKE, M. and MOKYR, J. (2017). Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy. *The Quarterly Journal of Economics*, **133** (1), 1–70.
- DE LA CROIX, D., LINDH, T. and MALMBERG, B. (2008). Swedish economic growth and education since 1800. *Canadian Journal of Economics*, **41** (1), 166–185.
- DEMOGRAPHIC DATA BASE (). The demographic data base on Swedish historical population statistics. CEDAR, Umeå University.
- DESMET, K. and PARENTE, S. L. (2012). The evolution of markets and the revolution of industry: a unified theory of growth. *Journal of Economic Growth*, **17** (3), 205–234.
- DOEPKE, M. (2004). Accounting for fertility decline during the transition to growth. *Journal of Economic growth*, **9** (3), 347–383.
- DRIBE, M. (2003). Migration of rural families in 19th century southern sweden. a longitudinal analysis of local migration patterns. *The History of the Family*, **8** (2), 247–265.
- DYSON, T. (2011). The role of the demographic transition in the process of urbanization. *Population and Development Review*, **37** (s1), 34–54.
- EDVINSSON, R. and SÖDERBERG, J. (2010). The evolution of swedish consumer prices 1290-2008. In R. Edvinsson, T. Jacobson and D. Waldenström (eds.), *Exchange rates, prices, and wages, 1277-2008*, Ekerlids Förlag & Sveriges Riksbank, pp. 412–452.
- FAY, M. and OPAL, C. (1999). Urbanization without growth: a not-so-uncommon phenomenon. *The World Bank*.
- FEHR, H. and KINDERMANN, F. (2018). *Introduction to Computational Economics using Fortran*. Oxford University Press.
- FLORA, P. (1983). *State, Economy and Society in Western Europe 1815-1975: A Data Handbook in Two Volumes. Vol. I: the Growth of Mass Democracies and Welfare States*. Campus Verlag.
- GALOR, O. (2005). The demographic transition and the emergence of sustained economic growth. *Journal of the European Economic Association*, **3** (2-3), 494–504.
- (2011). *Unified growth theory*. Princeton University Press.
- and WEIL, D. N. (1999). From malthusian stagnation to modern growth. *The American Economic Review*, **89** (2), 150–154.
- and — (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *The American Economic Review*, pp. 806–828.
- GLAESER, E. L. (1999). Learning in cities. *Journal of urban Economics*, **46** (2), 254–277.

- GOFFE, W. L., FERRIER, G. D. and ROGERS, J. (1994). Global optimization of statistical functions with simulated annealing. *Journal of Econometrics*, **60** (1), 65–99.
- GOLLIN, D., JEDWAB, R. and VOLLRATH, D. (2016). Urbanization with and without industrialization. *Journal of Economic Growth*, **21** (1), 35–70.
- GREENWOOD, J. and SESHADRI, A. (2002). The us demographic transition. *AEA Papers and Proceedings*, **92** (2), 153–159.
- HODRICK, R. J. and PRESCOTT, E. C. (1997). Postwar us business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, **29** (1), 1–16.
- HOLMBERG, M. (2016). The ghost of pandemics past: revisiting two centuries of influenza in sweden. *Medical Humanities*, pp. 141–147.
- JÖRBERG, L. (1972). *A history of prices in Sweden 1732-1914. 1. Sources, methods, tables*. Gleerup.
- KNICK, H. C. (1998). Cotton textile prices and the industrial revolution. *The Economic History Review*, **51** (1), 49–83.
- KNODEL, A. J. (1974). *The decline of fertility in Germany, 1871-1939, Office of Population Research and others*, vol. 2. Princeton University Press.
- LAGERLÖF, N.-P. (2003). From malthus to modern growth: Can epidemics explain the three regimes?\*. *International Economic Review*, **44** (2), 755–777.
- LAGERLÖF, N.-P. (2006). The galor–weil model revisited: A quantitative exercise. *Review of Economic Dynamics*, **9** (1), 116–142.
- LÉVY-LEBOYER, M. and BOURGUIGNON, F. (1990). *The French economy in the nineteenth century: an essay in econometric analysis*. Cambridge University Press.
- METCALFE, T. S. and CHARBONNEAU, P. (2003). Stellar structure modeling using a parallel genetic algorithm for objective global optimization. *Journal of Computational Physics*, **185** (1), 176–193.
- MOKYR, J. (1977). Demand vs. supply in the industrial revolution. *The Journal of Economic History*, **37** (4), 981–1008.
- (2010). Understanding growth in europe, 1700-1870: Theory and evidence. *in Stephen Broadberry and Kevin H. O’Rourke, eds., The Cambridge Economic History of Modern Europe – Vol I, Cambridge: Cambridge University Press*, pp. 7–42.
- MOSK, C. (1980). Rural-urban fertility differences and the fertility transition. *Population Studies*, **34** (1), 77–90.
- OGILVIE, S. (2014). The economics of guilds. *The Journal of Economic Perspectives*, **28** (4), 169–192.
- and CERMÁN, M. (1996). *European proto-industrialization: an introductory handbook*. Cambridge University Press.
- PENSIEROSO, L. and SOMMACAL, A. (2017). Agriculture to industry: the end of intergenerational coresidence. *IREES Discussion Paper 2017-7*.
- POWELL, M. J. (2002). Uobyqa: unconstrained optimization by quadratic approximation. *Mathematical Programming*, **92** (3), 555–582.

- RIDOLFI, L. (2017). *L'histoire immobile? Six centuries of real wages in France from Louis IX to Napoleon III: 1250-1860*. Tech. rep., LEM Working Paper Series.
- ROZENTAL, A. A. (1956). The enclosure movement in france. *American Journal of Economics and Sociology*, **16** (1), 55–71.
- SAVILLE, J. (2013). *Rural depopulation in England and Wales, 1851-1951*, vol. 9. Routledge.
- SCHÖN, L. (2010). *Sweden's road to modernity: an economic history*. SNS förlag Stockholm.
- SCHÖN, L., KRANTZ, O. *et al.* (2012). Swedish historical national accounts 1560-2010. *Lund Papers in Economic History*, **123**.
- SIMON, C. and NARDINELLI, C. (1996). The talk of the town: Human capital, information, and the growth of english cities, 1861 to 1961. *Explorations in Economic History*, **33** (3), 384–413.
- and — (2002). Human capital and the rise of american cities, 1900–1990. *Regional Science and Urban Economics*, **32** (1), 59–96.
- SÖDERBERG, J. (2010). Long-term trends in real wages of labourers. In R. Edvinsson, T. Jacobson and D. Waldenström (eds.), *Exchange rates, prices, and wages, 1277-2008*, Ekerlids Förlag & Sveriges Riksbank, pp. 453–478.
- SQUICCIARINI, M. P. and VOIGTLÄNDER, N. (2014). Human capital and industrialization: Theory and evidence from 18th century encyclopédie subscriptions. *UCLA Anderson, manuscript*.
- STRULIK, H. and WEISDORF, J. (2014). How child costs and survival shaped the industrial revolution and the demographic transition. *Macroeconomic Dynamics*, **18** (01), 114–144.
- and WERNER, K. (2016). 50 is the new 30—long-run trends of schooling and retirement explained by human aging. *Journal of Economic Growth*, **21** (2), 165–187.
- SUNDIN, J. and WILLNER, S. (2007). *Social change and health in Sweden: 250 years of politics and practice*. Swedish National Institute of Public Health.
- SVENSSON, P. (2006). Peasants and entrepreneurship in the nineteenth-century agricultural transformation of sweden. *Social Science History*, **30** (3), 387–429.
- TAMURA, R. (2002). Human capital and the switch from agriculture to industry. *Journal of economic Dynamics and Control*, **27** (2), 207–242.
- VAN DER WOUDE, A., HAYAMI, A. and DE VRIES, J. E. (1995). *Urbanization in history: a process of dynamic interactions*. Oxford University Press.
- VOIGTLÄNDER, N. and VOTH, H.-J. (2012). The three horsemen of riches: Plague, war, and urbanization in early modern europe. *Review of Economic Studies*, **80** (2), 774–811.
- and VOTH, H.-J. (2013). How the west “invented” fertility restriction. *The American Economic Review*, **103** (6), 2227–64.
- WEISDORF, J. L. (2004). From stagnation to growth: Revisiting three historical regimes. *Journal of Population Economics*, **17** (3), 455–472.
- WETHERELL, C. (2001). Another look at coale's indices of fertility,  $i_f$  and  $i_g$ . *Social Science History*, **25** (4), 589–608.
- WOODS, R. (2003). Urban-rural mortality differentials: An unresolved debate. *Population and Development Review*, **29** (1), 29–46.

## A Swedish data

### A.1 General procedure

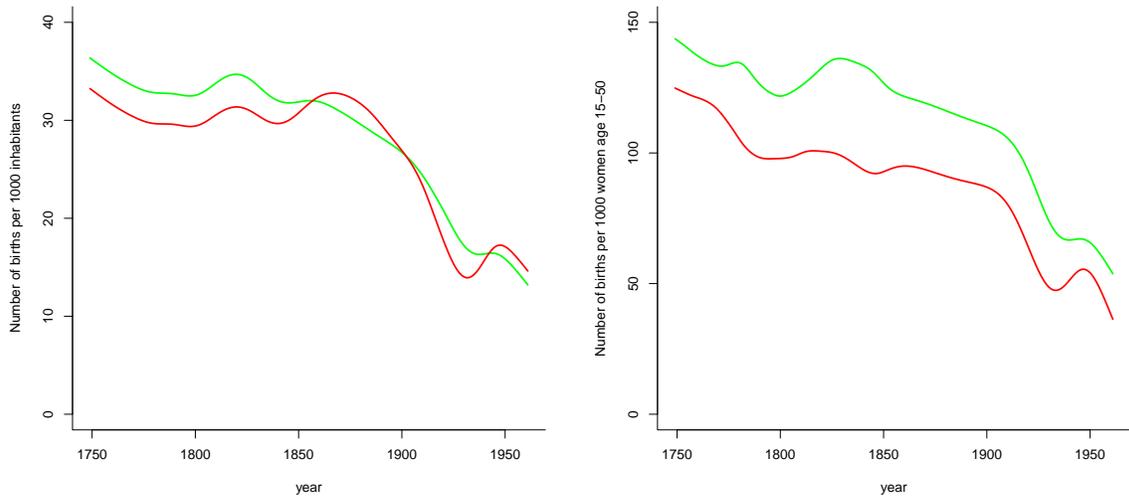
To develop a dataset that enables us to check our theory, we gathered data from different sources, like original census data, books, and available databases. Focusing on data over more than two centuries presents us with two main challenges. First, the definitions vary as they come from different sources. Additionally, even those definitions that come from official statistics have changed over time. This issue is of particular interest with respect to the separation of urban and rural areas. We document and discuss applied definitions in A.3. Second, countries' borders changed. We combine the sources to construct time series, while trying to minimize breaks due to conflicting definitions. Different sources often provide varying numbers in the same year. To ensure we are using the most reliable value in each case, we select the values that are either identified as less biased in the literature, originate from official statistics, or reduce inconsistencies.

While the Swedish data are rich, our time series are incomplete, because several observations are missing. In order to close these gaps, we apply linear interpolation. As we are not interested in short-term fluctuations, this approach is still defensible. We then smooth the data, and as the time series no longer include missing values, we apply the Hodrick-Prescott filter (Hodrick and Prescott, 1997). Finally, we pick values every 20 years, which we then try to match or use for overidentification checks.

### A.2 Age structures and differential fertility in Sweden

In general, Crude Birth Rates (CBR) are used to study the demographic transition; see, e.g., Dyson (2011) or Bocquier and Costa (2015). Figure 13a presents the trend of Dyson's CBRs for Sweden between 1750 and 1960. These estimated fertility differentials between cities and countryside are likely to be biased, as the age distributions of women between rural and urban areas are potentially different. Indeed, if we only consider women of fertile ages (15–49) and measure fertility using the so-called Generalized Fertility Rate (GFR), we see a clear fertility differential in the second part of the 19<sup>th</sup> century: fertility in the countryside was much higher than in the cities; see Figure 13b.

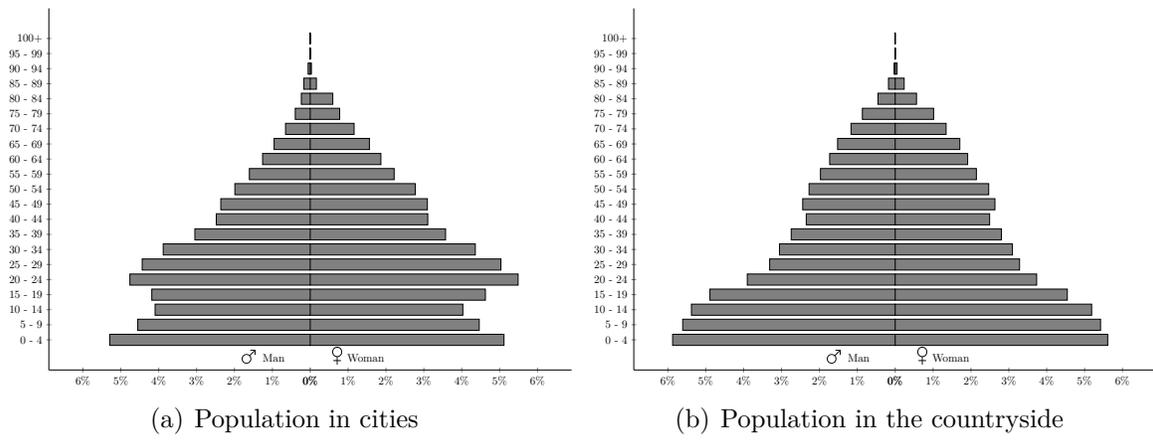
The varying findings for the CBRs and the GFRs are indeed induced by the differences in age structures illustrated in Figure 14 using the example of Sweden 1910Centralbyrån



(a) Crude Birth Rate (b) Generalized Fertility Rate

Figure 13: SWEDISH FERTILITY IN CITIES (RED) AND COUNTRYSIDE (GREEN) ACCORDING TO DIFFERENT MEASURES, 1760–1960

(1914). The share of fertile women in the total population of cities (29.1%) significantly exceeds the share in the countryside (22.6%). This difference deserves our particular attention when examining the effects of the rural exodus, as it was largely induced by the exodus. Figure 14 somehow documents the flow of women of fertile ages from the countryside into the cities, which could lead to an overestimation of fertility in cities relative to in the countryside.



(a) Population in cities (b) Population in the countryside

Figure 14: AGE STRUCTURES OF SWEDEN’S POPULATION BY SEX AND TYPE OF AREA IN 1910, ACCORDING TO THE STATISTICAL YEARBOOK SWEDEN 1914.

Since the exodus contributed to varying age structures, it is important to rely on a

fertility indicator that controls for the age structure. Ideally, we would use an indicator that applies age specific fertility rates such as the TFR. However, since data limitation do not allow this first-best solution, we proceed in two steps: we firstly estimate Coale’s index and secondly, use its values to approximate the TFR. While the GFR only excludes men and women not of fertile ages, Coale’s index of (general) fertility additionally provides a way to control for differences within the fertile age range. Coale (1969) proposed an index that normalizes the number of births by the number of births that would have occurred if the studied population had experienced the age-specific fertility of the Hutterites, which is reputed to be the closest to biological limits. This index then belongs to  $(0, 1)$  and is positively correlated to the fertility of the population under study. This is an efficient way to normalize the number of births by the prevailing age structure. We denote  $I_t^{\mathbb{A}}$  as the value of Coale’s index for area  $\mathbb{A}$  at time  $t$ .<sup>40</sup> It is calculated as follows:

$$I_t^{\mathbb{A}} = \frac{B_t^{\mathbb{A}}}{\sum_{o=1}^T \varepsilon_{o,t}^{\mathbb{A}} n_o^H}$$

where  $B_t^{\mathbb{A}}$  denotes the number of births in area  $\mathbb{A}$  at time  $t$ , while  $\varepsilon_{o,t}^{\mathbb{A}}$  is the number of individuals in area  $\mathbb{A}$  who belong to age group  $o$  at time  $t$ . Finally,  $n_o^H$  denotes the Hutterite age-specific fertility level. According to Coale (1969) these age specific fertility rates are:

Table 6: AGE-SPECIFIC FERTILITY RATES OF HUTTERITES

$o$	15–19	20–24	25–29	30–34	35–39	40–44	45–49
$n_o^H$	0.3	0.55	0.502	0.447	0.406	0.222	0.061
$\tilde{n}_o^H$	0.623	0.549	0.502	0.447	0.406	0.236	0.063

Source: Coale (1969).

According to Wetherell (2001) these values are an underestimations of Hutterites fertility due to ad hoc assumptions done by Coale (1969). Re-estimating Hutterites’ age specific fertility rates leads to  $\tilde{n}_o^H$  with a much higher fertility among women age 15–19. Still, the impact on Coale’s index is limited as illustrated in Fig. 15a. The higher age specific fertility rates should be used in case of high marriage rates in very young ages. However, since 1750 more than 5% of the Swedish female population were never married in these ages, such that the original estimation of Coale is preferable in the

<sup>40</sup>  $I_t^{\mathbb{A}}$  corresponds to the Coale’s index of general fertility.

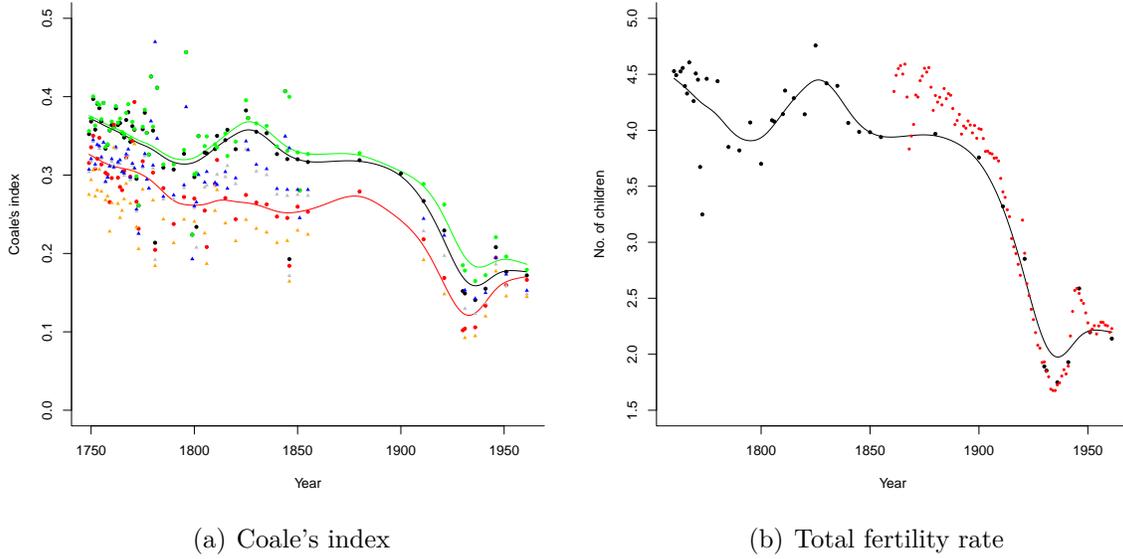


Figure 15: FIT OF SWEDISH FERTILITY ESTIMATIONS

Fig. 15a Coale's index estimated with the original values in rural (green), urban (red) and total Sweden (black) as well as the index computed with the re-estimated values from Wetherell (2001) in rural (blue), urban (orange) and total Sweden (gray). The black dots and the solid black line in Fig. 15b mark the estimated TFRs and their trend. Red dots add the TFR from the Human fertility database.

Swedish case. Furthermore, it allows us to use estimations of the Princeton fertility project.

In a second step, we can use the estimation of Coale's index to approximate the TFR simply by multiplying the index by the TFR of the Hutterites. On the country level, we can compare our estimations with direct estimations of the TFR since 1861. In fact, our approximation does a very good job. As illustrated in Figure 15b estimations from Coale's index reproduce the trend of Sweden's TFR from Human fertility database.

We apply the same strategy for the fertility predicted by the model and approximate Coale's index in a first step. Following the main idea presented in Coale (1969), we relate fertility in the model to the maximum fertility of the Hutterites:  $\hat{I}_t^{\mathbb{A}} = \frac{\sum_{i,t} n_{i,t}^{\mathbb{A}}}{N_t^{\mathbb{A}} \frac{5}{2} \sum_o n_o^H}$ . To compare our single sex population with the data, we divide the total population in area  $\mathbb{A}$  by two, as we are considering a single-sex population. Furthermore, we multiply the age-specific fertility rates of Hutterites by five to account for the five-year age groups. In a second step, we multiply Coale's index by 12.44 to approximate the TFR which also corresponds to the cohort fertility in the model.

### A.3 Definition of rural and urban areas

Information on urban and rural populations originate from different sources. The data we used for the first part of our study period, from 1749 to 1799, originate from SHiPs, Demographic Data Base and use institutional criteria. SHiPs provides us with population estimates by parish. We then identify rural and urban parishes using a database retrieval tool for urban-rural parishes that is also offered by Umeå University. The institutional criteria in SHiPs defines urban areas in the early stages based on municipal rights. These rights were intended to concentrate trade and artisanship in certain locations, and thus to support economic activities and facilitate control over the economy. This idea is close to the definition of a city used in our model: namely, that places with modern technologies are able to support economic growth. Additionally, the classification of a place as urban or rural does not change if the population increases or shrinks.

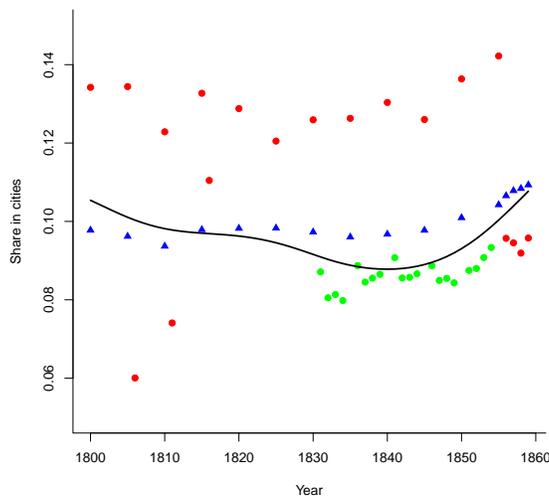


Figure 16: SWEDISH URBANIZATION RATES 1800–1859 FROM SHiPs (ESTIMATIONS CONSIDERED (GREEN) AND NOT CONSIDERED (RED)), OFFICIAL STATISTICS (BLUE) AND THE APPLIED TREND (BLACK)

For the period from 1800 to 1960, we use official statistics and their demographic criteria. An agglomeration requires more than 10,000 inhabitants to be a city. As both datasets are available between 1800 and 1859, we can compare them, which we do in Figure 16. It should be noted that in the first half of the 19<sup>th</sup> century, populations in urban areas are overrepresented by SHiPs’ data relative to the official data.<sup>41</sup>

<sup>41</sup>This finding is in line with our intuition: not all cities with municipal rights have more than

## A.4 Definition of industrial production

Different assumptions on the classification of agricultural and industrial production are possible within the seven categories of the Swedish Historical National Accounts. Figure 17a illustrates that the use of the categories “agriculture” for agriculture and “manufacturing” for industries is based on restrictive assumptions in the latter case. By also taking into account “building and construction” and “transportation and communication,” industrial production increases. Subtracting agricultural from total GDP obviously defines an upper limit for industrial production. In this case, industrial production would have exceeded agricultural output for the whole 1760-1960 study period.

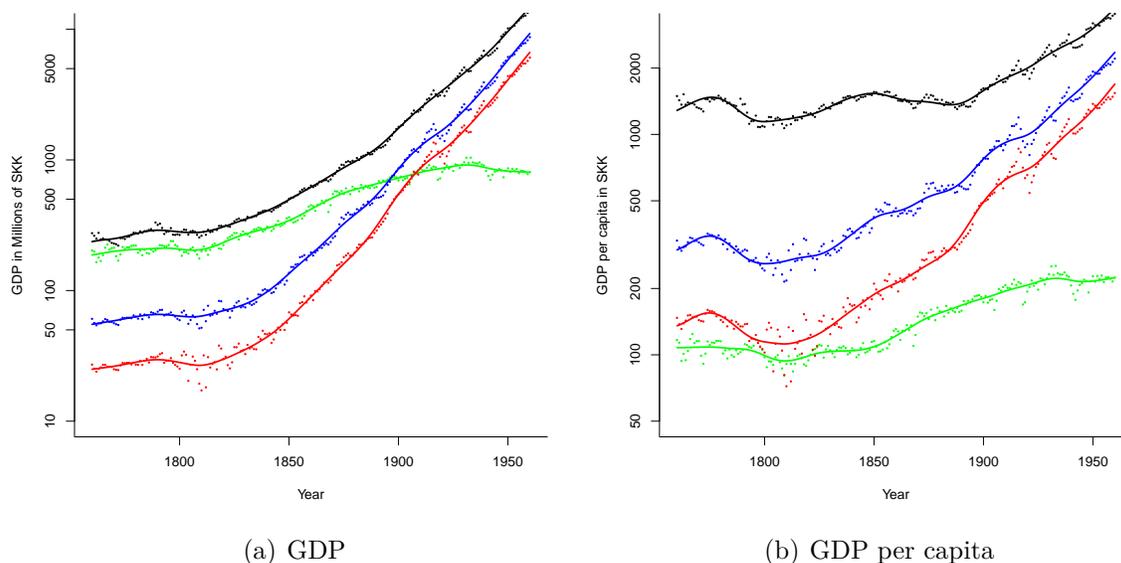


Figure 17: SWEDEN’S AGRICULTURAL PRODUCTION (GREEN) COMPARED TO ASSUMPTIONS ABOUT INDUSTRIAL PRODUCTION: MANUFACTURING (RED), MANUFACTURING, BUILDING AND CONSTRUCTION AND TRANSPORTATION AND COMMUNICATION (BLUE) AND GDP REDUCED BY AGRICULTURE (BLACK)

Our assumption that industries are located in cities and that agricultural activities take place in the countryside, is the basis for the GDP per capita presented in Figure 17b. Even the conservative assumption leads us to conclude that income was higher in the industrial sector. If we had have applied less restrictive assumptions, the initial income

---

10,000 inhabitants. However, as the number of places with more than 10,000 inhabitants increased over time, this distinction should have vanished or changed. Several places without municipal rights might exceed the threshold, such that the demographic criteria lead to higher values. Biases in the database due to missing parishes are an alternative explanation for the lower official values. This argument is supported by the changing levels in Figure 16.

would have been up to a 12 times higher in cities than in the countryside, compared to 1.25 times higher in the restrictive case. Since Clark’s critique that not all industries were located in cities and not all farming was done in rural areas, might be even more relevant for the other sectors, we rely on our conservative assumption, and present a critical discussion in the next appendix.

## A.5 GDP per capita, GDP per employee and commuting

Our benchmark measure of GDP per capita in each area assumes that all people living in the countryside work in agriculture, while all people living in cities work in industries. Stated differently, we do not consider commuting behaviors or proto-industrialization. Clark *et al.* (2012) argued that “*urbanization rates in preindustrial England provide a very poor guide to the share of the population employed in farming.*” This may also be true for Sweden. Centralbyran (1969) provided data points on total GDP and the number of employees per sector from 1850 to 1950; see Table 7. The measure allows us to compare the dynamics of GDP per capita using our benchmark definition to GDP per employee in each sector of production, and to GDP per capita when commuting is taken into consideration. As illustrated by Figure 18a, GDP per employee is higher than GDP per capita, which makes sense given that the denominator of the GDP per capita includes both non-active individuals and those employed in other sectors of production.

Table 7: SHARE OF INDIVIDUALS WORKING IN EACH SECTOR BY AREA

Year	Total		Rural		Urban	
	Agric.	Indust.	Agric.	Indust.	Agric.	Indust.
1870	72.4	14.6	82.4	9.1	5.2	51.7
1880	67.9	17.4	79.2	11.2	4.6	51.7
1890	62.1	21.7	75.5	14.3	4.7	53.4
1900	55.1	27.8	69.1	20.4	3.7	54.7
1910	48.8	32.0	63.7	25.1	3.6	53.0
1920	44.0	35.0	60.8	28.2	3.8	51.3
1930	39.4	35.7	56.6	29.6	3.6	48.6
1940	34.1	38.2	52.3	31.8	3.6	48.9
1950	24.6	42.7	43.5	36.9	3.5	49.2

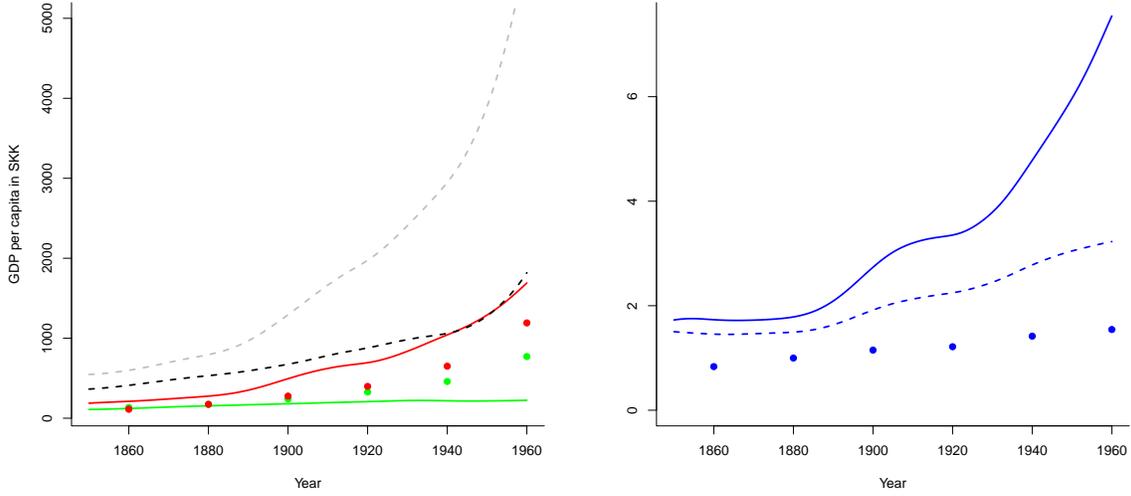
Source: Centralbyran (1969).

An attentive reading of Table 7 suggests that in 1870, a kind of (proto)-industrialization was taking place in rural areas. However, the increasing share of rural residents working in industry is also related to improvements in transportation that enabled them to commute. Still, the share of rural residents working in the agricultural sector was still very high, whereas levels of agricultural employment were negligible in the urban areas during the whole period. These two findings validate our assumption that living in the countryside is a requirement for working in agriculture. We also note that the share of individuals working in the industrial sector was constant over almost the whole period, and represented half of the population in cities. The rest of the urban population was working in sectors like transportation and communication and public services.

To take into account the increasing share of employees in industries living in rural areas, we compute new values of GDP per capita in each area as follows:

$$y_R = s_{R,A} \frac{Y_A}{s_{R,A}N_R + s_{U,A}N_U} + s_{R,I} \frac{Y_U}{s_{R,I}N_R + s_{U,I}N_U},$$

where  $s_{M,J}$  denotes the proportion of persons from area  $M$  working in sector  $J$  while  $N_M$  denotes the population size of area  $M$ . This transformation first involves recomputing the average GDP per capita in each sector by dividing the total production by the number of workers in the sector. Then, when we compute the average GDP per capita in rural areas, we simply multiply the shares of people living in rural areas and working in the agricultural sector by the average GDP of this sector and add the share of people from rural areas working in industries multiplied by the average GDP per capita in this sector. The results of these transformations appear as dots in Figure 18. The main result of this exercise is that when we consider the effect of commuting on local wealth, the period of stagnation that we document for the countryside after 1930 seems to disappear. This finding is in line with the predictions of our theory.



(a) GDP per capita or employee

(b) Ratio between GDP in both sectors/areas

Figure 18: GDP PER CAPITA (SOLID LINES) AND MODIFIED GDP PER CAPITA (DOTS) IN RURAL (GREEN) AND URBAN (RED) AREAS VERSUS GDP PER EMPLOYEE (DASHED LINE) IN AGRICULTURE (BLACK) AND INDUSTRY (GRAY)

## B Theory

### B.1 Regimes

In order to characterize the optimal decisions of adults living in area  $\mathbb{A}$  for any configuration of our parameters set, we first define some critical values on wages and the relative price of industrial goods:  $\bar{\omega}_t$  is the potential income above which a parent living in area  $\mathbb{A}$  and who had paid  $\kappa$ , decides to educate her children:

$$\bar{\omega}_t \equiv \frac{vp_t\beta q_t^{\mathbb{A}}}{\phi(\xi + \zeta^{\mathbb{A}}q_t^{\mathbb{A}})}.$$

The potential income above which she decides to consume industrial goods is:

$$\hat{\omega}_t \equiv \frac{\alpha + \rho}{1 - \alpha} \varepsilon p_t + \kappa + \bar{c}.$$

Finally,  $p^*(\kappa)$  denotes the limit price below which poor parents who can afford  $\bar{c}$  but cannot migrate, could nevertheless educate their children, and  $\tilde{p}$  is the limit price above

which  $\bar{\omega}_t > \hat{\omega}_t$ :

$$p^*(\kappa) \equiv \frac{(\bar{c} + \kappa)\phi(\xi + \zeta^{\mathbb{A}}q_t^{\mathbb{A}})}{v\beta q_t^{\mathbb{A}}} , \quad \tilde{p}(\kappa) \equiv \frac{\bar{c} + \kappa}{\frac{v\beta q_t^{\mathbb{A}}}{\phi(\xi + \zeta^{\mathbb{A}}q_t^{\mathbb{A}})} - \frac{\alpha + \rho}{1 - \alpha}\varepsilon}.$$

**Proposition 1**

*There exist  $\bar{v}_t > 0$ , defined as*

$$\bar{v}_t = \frac{(\alpha + \rho)\varepsilon\phi(\xi + \zeta^{\mathbb{A}}q_t^{\mathbb{A}})}{(1 - \alpha)\beta q_t^{\mathbb{A}}},$$

*such that:*

- **If  $v \leq \bar{v}_t$ , then  $\tilde{p}(\kappa) \leq 0 < p^*(\kappa)$ , such that:**

*when  $p_t \leq p^*(\kappa)$ , optimal decisions of people living in area  $\mathbb{A}$  are:*

*Not moving person ( $\kappa = 0$ ):*

- if  $\omega_t^{\mathbb{A}} \leq \bar{c}$  then R1*
- if  $\omega_t^{\mathbb{A}} \in ]\bar{c}, \hat{\omega}_t]$  then R5*
- if  $\omega_t^{\mathbb{A}} \geq \hat{\omega}_t$  then R6*

*Moving person ( $\kappa = \bar{\kappa}$ ):*

- if  $\omega_t^{\mathbb{A}} \in [\bar{c} + \bar{\kappa}, \hat{\omega}_t]$  then R5*
- if  $\omega_t^{\mathbb{A}} > \hat{\omega}_t$  then R6*

*while, when  $p_t > p^*(\kappa)$ , then:*

*Not moving person ( $\kappa = 0$ ):*

- if  $\omega_t^{\mathbb{A}} \leq \bar{c}$  then R1*
- if  $\omega_t^{\mathbb{A}} \in ]\bar{c}, \bar{\omega}_t]$  then R2*
- if  $\omega_t^{\mathbb{A}} \in ]\bar{\omega}_t, \hat{\omega}_t]$  then R5*
- if  $\omega_t^{\mathbb{A}} \geq \hat{\omega}_t$  then R6*

*Moving person ( $\kappa = \bar{\kappa}$ ):*

- if  $\omega_t^{\mathbb{A}} \in [\bar{c} + \bar{\kappa}, \bar{\omega}_t]$  then R3*
- if  $\omega_t^{\mathbb{A}} \in [\bar{\omega}_t, \hat{\omega}_t]$  then R5*
- if  $\omega_t^{\mathbb{A}} > \hat{\omega}_t$  then R6*

- **If  $v > \bar{v}_t$ , then:  $0 < p^*(\kappa) < \tilde{p}(\kappa)$ , such that three situations may arise.**

**First, if  $p_t \leq p^*(\kappa)$ , then:**

Not moving person ( $\kappa = 0$ ):

if  $\omega_t^A \leq \bar{c}$  then R1  
if  $\omega_t^A \in ]\bar{c}, \hat{\omega}_t]$  then R5  
if  $\omega_t^A \geq \hat{\omega}_t$  then R6

Moving person ( $\kappa = \bar{\kappa}$ ):

if  $\omega_t^A \in [\bar{c} + \bar{\kappa}, \hat{\omega}_t]$  then R5  
if  $\omega_t^A > \hat{\omega}_t$  then R6

**For intermediate prices,  $p_t \in ]p^*(\kappa), \tilde{p}(\kappa)]$ , we get that:**

Not moving person ( $\kappa = 0$ ):

if  $\omega_t^A \leq \bar{c}$  then R1  
if  $\omega_t^A \in ]\bar{c}, \bar{\omega}_t]$  then R2  
if  $\omega_t^A \in ]\bar{\omega}_t, \hat{\omega}_t]$  then R5  
if  $\omega_t^A \geq \hat{\omega}_t$  then R6

Moving person ( $\kappa = \bar{\kappa}$ ):

if  $\omega_t^A \in [\bar{c} + \bar{\kappa}, \bar{\omega}_t]$  then R3  
if  $\omega_t^A \in [\bar{\omega}_t, \hat{\omega}_t]$  then R5  
if  $\omega_t^A > \hat{\omega}_t$  then R6

**Finally, if  $p_t$  becomes greater than  $\tilde{p}(\kappa)$ , then:**

Not moving person ( $\kappa = 0$ ):

if  $\omega_t^A \leq \bar{c}$  then R1  
if  $\omega_t^A \in ]\bar{c}, \hat{\omega}_t]$  then R2  
if  $\omega_t^A \in ]\hat{\omega}_t, \bar{\omega}_t]$  then R4  
if  $\omega_t^A > \bar{\omega}_t$  then R6

Moving person ( $\kappa = \bar{\kappa}$ ):

if  $\omega_t^A \in [\bar{c} + \bar{\kappa}, \hat{\omega}_t]$  then R3  
if  $\omega_t^A \in [\hat{\omega}_t, \bar{\omega}_t]$  then R4  
if  $\omega_t^A > \bar{\omega}_t$  then R6

where R1–R6 correspond to the regimes described in Table 1.

**Proof.**

To prove Proposition 1, we first expose and prove four lemmas. To alleviate the notations, the time indices have been suppressed.

**Lemma 1**  $n_i < n_M \forall i = \{1, 2, \dots, 6\}$  where  $i$  labels the regimes described in Table 1.

**Proof.** The inspection of each  $n_i$  yields the following results:

- $n_1 = 0 < n_M$

•  $n_2$  is continuous with respect to  $\omega$ ,  $\forall \omega^{\mathbb{A}} \in \mathbb{R}^+$ . Furthermore,  $\frac{dn_2}{d\omega^{\mathbb{A}}} = \frac{c}{(\alpha+\rho)(\xi+\zeta^{\mathbb{A}}q^{\mathbb{A}})\omega^{\mathbb{A}}} > 0$ ,  $n_2(\bar{c}) = 0$  and  $\lim_{\omega^{\mathbb{A}} \rightarrow +\infty} n_2(\omega^{\mathbb{A}}) = \frac{\rho}{\alpha+\rho}n_M < n_M$ . We can then conclude that  $n_2 < n_M$ ,  $\forall \omega^{\mathbb{A}}$ .

•  $n_3$  is continuous with respect to  $\omega$ ,  $\forall \omega^{\mathbb{A}} \in \mathbb{R}^+$ . Furthermore,  $\frac{dn_3}{d\omega^{\mathbb{A}}} = \frac{c+\kappa}{(\alpha+\rho)(\xi+\zeta^{\mathbb{A}}q^{\mathbb{A}})\omega^{\mathbb{A}}} > 0$ ,  $n_3(\bar{c} + \kappa) = 0$  and  $\lim_{\omega^{\mathbb{A}} \rightarrow +\infty} n_3(\omega^{\mathbb{A}}) = \frac{\rho}{\alpha+\rho}n_M < n_M$ . We can then conclude that  $n_3 < n_M$ ,  $\forall \omega^{\mathbb{A}}$ .

•  $n_4$  is continuous with respect to  $\omega$ ,  $\forall \omega^{\mathbb{A}} \in \mathbb{R}^+$ . Furthermore, it is a monotonous function of  $\omega^{\mathbb{A}}$ , as the sign of  $\frac{dn_4}{d\omega^{\mathbb{A}}}$  does not depend on  $\omega^{\mathbb{A}}$ :

$$\frac{dn_4}{d\omega^{\mathbb{A}}} = \frac{\rho(\kappa + \bar{c} - \varepsilon p)}{(1 + \rho)(\xi + \zeta^{\mathbb{A}}q^{\mathbb{A}})\omega^{\mathbb{A}}}$$

Rapid calculations show that  $\lim_{\omega^{\mathbb{A}} \rightarrow +\infty} n_4(\omega^{\mathbb{A}}) = \frac{\rho}{1+\rho}n_M < n_M$  and  $n_4(\hat{\omega}) < n_M$ . We can then conclude that  $n_4 < n_M$ ,  $\forall \omega^{\mathbb{A}}$ .

• Simple computations yield the conclusion that:

$$\begin{aligned} n_5(\omega^{\mathbb{A}}) &\leq n_3(\omega^{\mathbb{A}}) < n_M, \quad \forall \omega^{\mathbb{A}} \geq \bar{\omega} \\ n_6(\omega^{\mathbb{A}}) &\leq n_4(\omega^{\mathbb{A}}) < n_M, \quad \forall \omega^{\mathbb{A}} \geq \bar{\omega} \end{aligned}$$

■

**Lemma 2** *Using results from Lemma 1 and Table 1, it follows that:*

- $\{c_1, n_1, d_1, e_1\}$  is, by assumption, the solution of the individual maximization program  $\forall \omega^{\mathbb{A}} \leq \bar{c}$ .
- $\{c_2, n_2, d_2, e_2\}$  maximizes 7 subject to 5,  $n_i < n_M$ ,  $c \geq \bar{c}$  and  $n_i = d_i = e_i = 0$  when  $\kappa = 0$ .
- $\{c_3, n_3, d_3, e_3\}$  maximizes 7 subject to 5,  $n_i < n_M$ ,  $c \geq \bar{c}$ ,  $n_i > 0$  and  $d_i = e_i = 0$  when  $\kappa > 0$ .
- $\{c_4, n_4, d_4, e_4\}$  maximizes 7 subject to 5,  $n_i < n_M$ ,  $c \geq \bar{c}$ ,  $n_i > 0$  and  $d_i > 0$  and  $e_i = 0$   $\forall \kappa \geq 0$ .
- $\{c_5, n_5, d_5, e_5\}$  maximizes 7 subject to 5,  $n_i < n_M$ ,  $c \geq \bar{c}$ ,  $n_i > 0$  and  $d_i = 0$  and  $e_i > 0$   $\forall \kappa \geq 0$ .

- $\{c_6, n_6, d_6, e_6\}$  maximizes  $\gamma$  subject to 5,  $n_i < n_M$ ,  $c \geq \bar{c}$ ,  $n_i > 0$  and  $d_i > 0$  and  $e_i > 0$   $\forall \kappa \geq 0$ .

**Proof.** All the results are obtained through the traditional analysis of the maximization problems described in each sub-case.

■

**Lemma 3** When  $v \leq \bar{v}$ ,  $\tilde{p}(\kappa) \leq 0 < p^*(\kappa)$  while when  $v > \bar{v}$ ,  $0 < p^*(\kappa) < \tilde{p}(\kappa)$

**Proof.** We first show that  $\tilde{p}(\kappa) \leq 0$  when  $v \leq \bar{v}$ :

$$\begin{aligned} \tilde{p}(\kappa) \leq 0 &\Leftrightarrow \frac{\bar{c} + \kappa}{\frac{v\beta q^{\mathbb{A}}}{\phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}})} - \frac{\alpha + \rho}{1 - \alpha} \varepsilon} \leq 0 \\ &\Leftrightarrow v \leq \frac{(\alpha + \rho) \varepsilon \phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}})}{(1 - \alpha) \beta q^{\mathbb{A}}} = \bar{v} \end{aligned}$$

Then we deal with the comparison between  $\tilde{p}(\kappa)$  and  $p^*(\kappa)$ :

$$\tilde{p}(\kappa) > p^*(\kappa) \Leftrightarrow \frac{\bar{c} + \kappa}{\frac{v\beta q^{\mathbb{A}}}{\phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}})} - \frac{\alpha + \rho}{1 - \alpha} \varepsilon} > \frac{(\bar{c} + \kappa) \phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}})}{v\beta q^{\mathbb{A}}} \quad (13)$$

If  $v \leq \bar{v}$ , the denominator of the left-hand side of Inequality (13) is negative such that Inequality (13) is never satisfied. This means that  $\tilde{p}(\kappa) \leq 0 < p^*(\kappa)$ . By contrast, if  $v > \bar{v}$ , Inequality (13) summarizes to  $\frac{\alpha + \rho}{1 - \alpha} \varepsilon \phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}}) > 0$  which is always satisfied, meaning that if  $v > \bar{v}$ ,  $\tilde{p}(\kappa) > p^*(\kappa) > 0$ .

■

**Lemma 4**

- If  $v \leq \bar{v}$ , then  $\tilde{p}(\kappa) \leq 0 < p^*(\kappa)$  such that:
  - if  $p \leq p^*(\kappa)$ , then  $\bar{\omega} < \bar{c} + \kappa \leq \hat{\omega}$
  - if  $p > p^*(\kappa)$ , then  $\bar{c} + \kappa < \bar{\omega} < \hat{\omega}$
- If  $v > \bar{v}$ , then  $0 < p^*(\kappa) < \tilde{p}(\kappa)$  such that:
  - if  $p \leq \tilde{p}(\kappa)$ , then  $\bar{c} + \kappa < \bar{\omega} \leq \hat{\omega}$
  - if  $p > \tilde{p}(\kappa)$ , then  $\bar{c} + \kappa < \hat{\omega} < \bar{\omega}$

The proof of this lemma directly flows from Lemma 3 and the comparison between  $\bar{w}$  and zero, as well as the comparison between  $\bar{w}$  and  $\hat{w}$ :

$$\begin{aligned}\bar{w} < \bar{c} + \kappa &\Leftrightarrow p < \frac{\phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}})(\bar{c} + \kappa)}{v\beta q^{\mathbb{A}}} = p^*(\kappa) \\ \bar{w} < \tilde{w} &\Leftrightarrow p < \frac{\bar{c} + \kappa}{\frac{v\beta q^{\mathbb{A}}}{\phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}})} - \frac{\alpha + \rho}{1 - \alpha} \varepsilon} = \tilde{p}(\kappa).\end{aligned}$$

■

Inspecting Lemmas 1 to 4 leads us to conclude that Proposition 1 describes well the solution to the maximization problem exposed in Subsection 3.2.

## B.2 Non-monotonicity of fertility behavior

We study the specific situation of a non-moving person in a case in which  $v$  is rather high, while the equilibrium price prevailing on the market for industrial goods has intermediate values.<sup>42</sup> In this situation, we find that:

$$n_t = \begin{cases} 0 & \text{if } \omega_t^{\mathbb{A}} \leq \bar{c} \\ \frac{\rho(\omega_t^{\mathbb{A}} - \bar{c})}{(\alpha + \rho)(\xi + \zeta^{\mathbb{A}} q_t^{\mathbb{A}})\omega_t^{\mathbb{A}}} & \text{if } \bar{c} < \omega_t^{\mathbb{A}} \leq \bar{\omega}_t \\ \frac{\rho(1 - \phi)(\omega_t^{\mathbb{A}} - \kappa - \bar{c})}{(\alpha + \rho)[\phi(\xi + \zeta^{\mathbb{A}} q_t^{\mathbb{A}})\omega_t^{\mathbb{A}} - \beta p_t q_t^{\mathbb{A}} v]} & \text{if } \bar{\omega}_t < \omega_t^{\mathbb{A}} \leq \hat{\omega}_t \\ \frac{\rho(1 - \phi)(\hat{\omega}_t - \bar{c})}{(1 + \rho)[\phi(\xi + \zeta^{\mathbb{A}} q_t^{\mathbb{A}})\omega_t^{\mathbb{A}} - \beta p_t q_t^{\mathbb{A}} v]} & \text{if } \omega_t^{\mathbb{A}} > \hat{\omega}_t \end{cases}; \quad \frac{\partial n_t}{\partial \omega_t^{\mathbb{R}}} \begin{cases} = 0 & \text{if } \omega_t^{\mathbb{A}} \leq \bar{c} \\ > 0 & \text{if } \bar{c} < \omega_t^{\mathbb{A}} \leq \bar{\omega}_t \\ \leq 0 & \text{if } \bar{\omega}_t < \omega_t^{\mathbb{A}} \leq \hat{\omega}_t \\ \leq 0 & \text{if } \omega_t^{\mathbb{A}} > \hat{\omega}_t \end{cases}$$

The relationship between optimal fertility and wages is represented in Figure 19. For our exemplary person who decided not to migrate, having an income lower than  $\bar{c}$  means that she is unable to give birth. If her wage is higher, she can have children in a Malthusian context in which fertility is limited by earning capacities. Any increase in  $\omega_t^{\mathbb{A}}$  will allow this person to have more children. When her income exceeds  $\bar{\omega}_t^{\mathbb{A}}$ , she can afford to educate children, and a trade-off between the quality and the quantity of children takes place. The higher the  $\omega_t^{\mathbb{A}}$ , the higher the cost of quantity relative to the cost of quality. Having higher wages incentivizes parents to have fewer children who are better educated. Once  $\omega_t^{\mathbb{A}}$  becomes higher than  $\hat{\omega}_t^{\mathbb{A}}$ , investing in the education of children can be done without having to give up to the consumption of industrial goods.

<sup>42</sup>Using the notations of Appendix B.1, we study a case in which  $v > \bar{v}_t$  and  $p_t \in ]p^*(\kappa), \tilde{p}(\kappa)[$ .

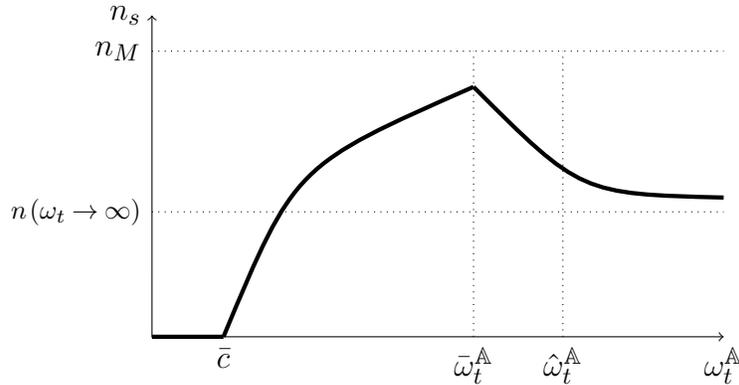


Figure 19: FERTILITY AND INCOME OF NON-MOVING PARENTS

## C Mortality crisis and net fertility in Sweden

In this exercise, we try to determine whether a major mortality crisis such as the Black Death could have prevented the take-off to modernity. This crisis might have altered both areas or only cities.

### C.1 Mortality crisis

We first simulate the impact of a mortality crisis by assuming that IMRs increased to 50% in 1800 (blue lines of Figure 20). Such a rate should be considered as a kind of upper limit for a mortality crisis caused by pandemics, such as the influenza pandemics in 1831–33 and 1847–48, or the famine in 1772–73.<sup>43</sup> At the time of the shock, the crude number of births increases very strongly in both areas, but they return to their benchmark value as early as in the next period. Interestingly enough, as highlighted in the next subsection, the mortality crisis leads to a decrease in net fertility. This lower net fertility then reinforces GDP growth through two mechanisms: it increases the labor supply and it reduces the dependency ratio. As in the case of mortality stagnation described in Subsection 5.2, the impact on population growth is not sufficiently strong to delay the onset of the global technological progress.

<sup>43</sup>While Sweden was barely affected by two influenza pandemics in the first half of the 19<sup>th</sup> century, the Black Death arrived for the last time in 1710–13, when Sweden lost 30–40% of its total population (Sundin and Willner, 2007). Since Brändström and Sundin (1981) found for the five-year averages in 1830–34 parishes with IMRs up to 50% for farmers' children, we use this upper bound for our simulation.

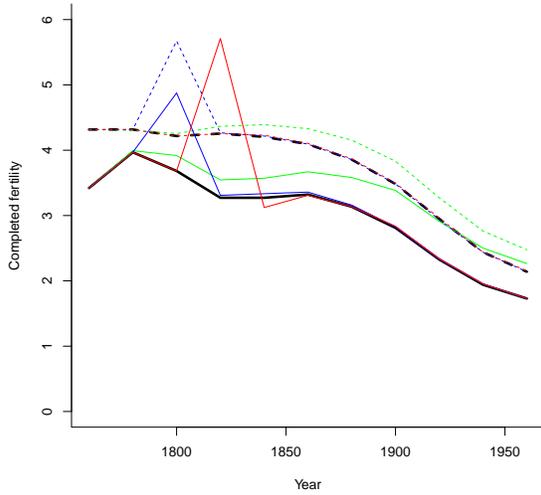
Regarding education, the temporary decrease in the net number of children translates into increased investments, and thus to a larger total stock of human capital in the economy. The mortality crisis implies a small but permanent increase in the GDP per capita compared to the benchmark scenario, while absolute GDP growth is not greatly altered (see Figure 25, Appendix D.7). For the same reason, reduced growth rates can be observed in the countryside.

The consequences of an asymmetric mortality shock that only hits cities due, for instance, to the higher population densities and bad hygienic conditions in urban areas, are quite similar. As documented by (Holmberg, 2016), Stockholm in 1833 was characterized by dramatic infant mortality rates, much higher than the remaining parts of the country; we have then chosen an IMR of  $\frac{2}{3}$  as our new upper bound here. With only one-third of children surviving in 1820, the peak in Coale's index is stronger than in the previous exercise, which indicates a strong replacement fertility. Nevertheless, due to the lower net fertility in cities, the share of the population living in cities is slightly lower than in the benchmark scenario, and the positive impact on education vanishes. Finally, compared to the symmetric mortality shocks, we observe the same overshooting followed by an undershooting in urban growth rates of GDP per capita. The dynamics of the countryside do not change significantly, since rural areas are not directly hit by the mortality crisis. The general equilibrium effects are limited in this scenario.

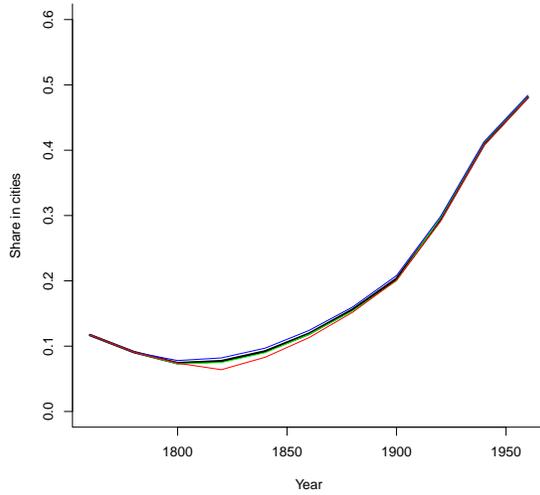
Overall, the effects of the IMRs are quite restricted compared to the variations of  $\bar{k}$ . The limited impact of infant (and child) mortality on the long-term economic and demographic dynamics in Western Europe has been shown by Galor (2011) in the case of France. We tend to confirm this result using advanced estimation techniques. However, our results may also be driven by the fact that we do not consider uncertainty about the number of surviving children, and that our utility function is log-linear. Adams (2017) discusses issues related to the choice of the utility function.

## C.2 Net fertility

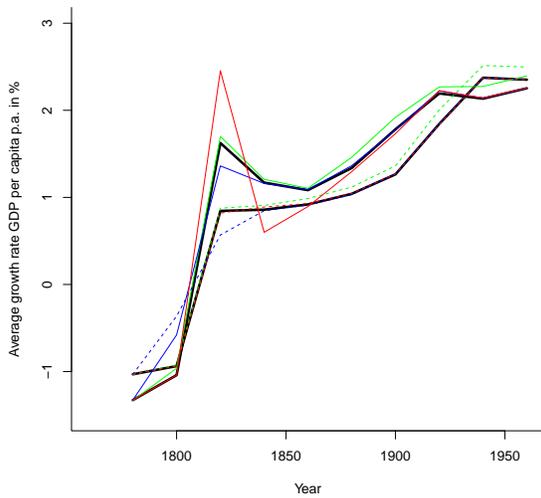
Coale's index relates the number of births of our population to the maximum fertility of the Hutterites, and thus our approximated TFR only measures the number of deliveries, not the number of surviving children. Since the latter information is also of interest – particularly for historical experiments on infant mortality – we modify the way we



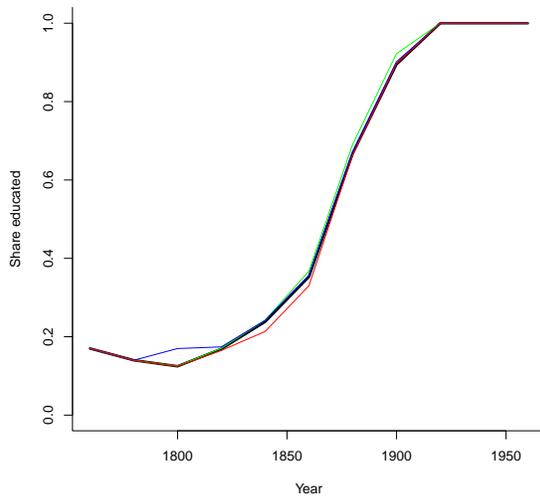
(a) Fertility in cities (solid) and countryside (dashed)



(b) Urbanization



(c) GDP per capita in industries (solid) and the agriculture (dashed)



(d) Education

Figure 20: Swedish demographic and economic dynamics in the benchmark scenario (black), with constant IMRs (green), a countrywide shock in IMRs in 1800 (blue), and a shock in cities' IMRs in 1820 (red), 1760–1960

Table 8: GROSS AND NET FERTILITY IN THE BENCHMARK AND NO IMPROVEMENTS IN THE IMR SCENARIO.

	Benchmark Estimation				No improvements in IMRs			
	Rural		Urban		Rural		Urban	
	TFR <sub>gr</sub>	TFR <sub>net</sub>	TFR <sub>gr</sub>	TFR <sub>net</sub>	TFR <sub>gr</sub>	TFR <sub>net</sub>	TFR <sub>gr</sub>	TFR <sub>net</sub>
1760	4.317	3.446	3.421	2.339	4.317	3.446	3.421	2.339
1780	4.317	3.421	3.968	2.749	4.304	3.433	3.993	2.737
1800	4.217	3.397	3.682	2.749	4.254	3.384	3.919	2.688
1820	4.254	3.521	3.273	2.500	4.366	3.483	3.545	2.426
1840	4.205	3.570	3.273	2.500	4.391	3.496	3.570	2.438
1860	3.856	3.545	3.321	2.625	4.329	3.446	3.670	2.513
1880	3.856	3.433	3.135	2.600	4.155	3.309	3.583	2.451
1900	3.483	3.185	2.811	2.500	3.832	3.048	3.384	2.314
1920	2.948	2.762	2.326	2.177	3.272	2.612	2.911	1.990
1940	2.438	2.339	1.941	1.878	2.762	2.202	2.500	1.704
1960	2.140	2.115	1.729	1.704	2.476	1.978	2.264	1.555

approximate the number of surviving children ( $TFR_{net}$ ) for Coale's index according to:

$$TFR_{net,t}^A = 12.44 * q_t^A \frac{\sum_{I^A} n_{i,t}^A}{N_t^A \frac{5}{2} \sum_o n_o^H}.$$

Using this approach, we can also take into account IMRs, and can find out whether the number of surviving children increases or decreases compared to the benchmark if we shock infant mortality. In Tables 8 and 9, we show how the number of surviving children ( $TFR_{net}$ ) might drop albeit the number of births ( $TFR_{gr}$ ) exceeds the benchmark level.

Table 9: GROSS AND NET FERTILITY IN SCENARIOS WITH MORTALITY SHOCKS.

	Countrywide crisis in IMRs in 1800				IMR crisis in cities in 1820			
	Rural		Urban		Rural		Urban	
	TFR <sub>gr</sub>	TFR <sub>net</sub>	TFR <sub>gr</sub>	TFR <sub>net</sub>	TFR <sub>gr</sub>	TFR <sub>net</sub>	TFR <sub>gr</sub>	TFR <sub>net</sub>
1760	4.217	3.446	3.421	2.339	4.217	3.446	3.421	2.339
1780	4.217	3.421	3.968	2.749	4.217	3.421	3.968	2.749
1800	5.673	2.836	4.876	2.438	4.218	3.409	3.682	2.749
1820	4.267	3.545	3.309	2.525	4.254	3.521	5.710	1.144
1840	4.218	3.583	3.334	2.575	4.230	3.583	3.122	2.413
1860	4.093	3.545	3.359	2.662	4.105	3.558	3.309	2.612
1880	3.856	3.433	3.160	2.625	3.869	3.446	3.135	2.600
1900	3.483	3.185	2.824	2.500	3.496	3.197	2.836	2.513
1920	2.948	2.749	2.326	2.177	2.961	2.774	2.339	2.177
1940	2.438	2.339	1.953	1.878	2.438	2.351	1.953	1.891
1960	2.140	2.115	1.729	1.704	2.152	2.115	1.729	1.717

## D Calibration and historical experiments

### D.1 Details on the estimation technic

Our hybrid parallelization strategy employs MPI (Message Passing Interface) to distribute parameter sets across multiple nodes of a compute cluster, and uses OpenMP threads to parallelize the computations of the equilibrium dynamics over the available CPU cores of each node. To this end, we have implemented a variant of the MPIKAIA algorithm (Metcalfé and Charbonneau, 2003), which closely follows their approach of dynamically scheduling parameter sets to different MPI processes (in order to cope with large variations in the computing times for the individual parameter sets), but without the need to designate a special master process (in order to avoid wasting computational resources in our hybrid approach). This numerical exercise is still very demanding in terms of computational power: it uses roughly 1,000 cores of a compute cluster, and takes around 15 hours.

## D.2 Statistics on the model's fit

In this appendix, we summarize in some statistics the fit of our model. We rely on two indicators: (i) linear regressions, where we regress the predicted  $Y_t^s$  on the observed moments  $Y_t^o$ :  $Y_t^o = \alpha_0 + \alpha_1 Y_t^s + \epsilon_t$ ; and (ii) the average difference in percentages or percentage points between estimated and predicted trends.<sup>44</sup>

As Table 10 shows, the model predicts the estimated trends very well for the demographic series of fertility and urbanization. While the fit is slightly worse for the economic moments we do not try to match, it is still convincing. This shows the quality of the predictions made by our model.

Table 10: EVALUATION OF MODEL'S FIT

Variable	$\alpha_0$	$\alpha_1$	$R^2$	%-deviation
Coale's index in U	0.02353	0.89678***	0.8225	10.80040775
Coale's index in R	0.01252	0.97842***	0.9511	5.395292368
Share in U	0.00059	1.01424***	0.9789	13.56167578
rel. price	-0.55090**	0.34096***	0.8875	29.12639981
Education	-0.04860	0.91418***	0.8975	411.2753615
GDP p.c.	-0.24274***	1.31475***	0.9905	19.98157651
Industrial wages	-0.19655	1.40979***	0.9872	-17.3487
Agricultural wages	0.09971	1.059***	0.8565	-14.435
Growth rate GDP p.c. in U	0.46210	0.7138***	0.5619	0.612813964
Growth rate GDP p.c. in U corr.	0.59850	0.7332***	0.5658	0.479772498
Growth rate GDP p.c. in R	0.1939	0.1807	0.195	0.928896329
Growth rate GDP p.c. in R corr.	0.279	0.7069***	0.7456	0.736642729

Interestingly enough, our model performs poorly regarding the fit of the growth rate of GDP per capita when commuting is not taken into account while it performs quite well when commuting is taken into account.

## D.3 Alternative educational thresholds

Figure 21 displays the observed primary education enrollment rates (solid line) using alternative critical values  $\bar{e}' = v\bar{e}$ , where  $v$  characterizes the multiple of the benchmark

<sup>44</sup>The latter is used for all growth rates, while the remaining values in column 4 of Table 10 are deviations in percentages.

value  $\bar{e}$ . Increasing the critical value of education depresses the initial education and postpones the onset of the measured educational revolution.

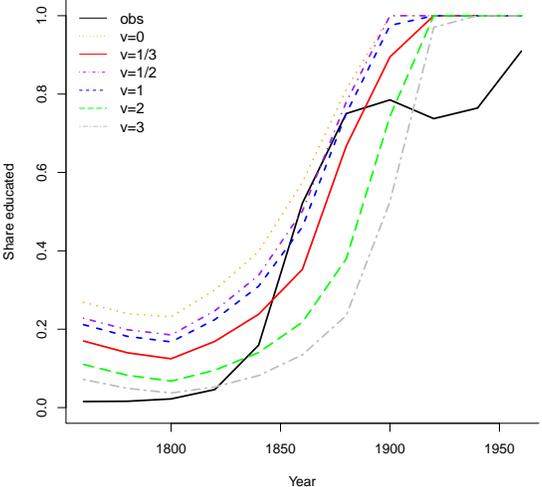
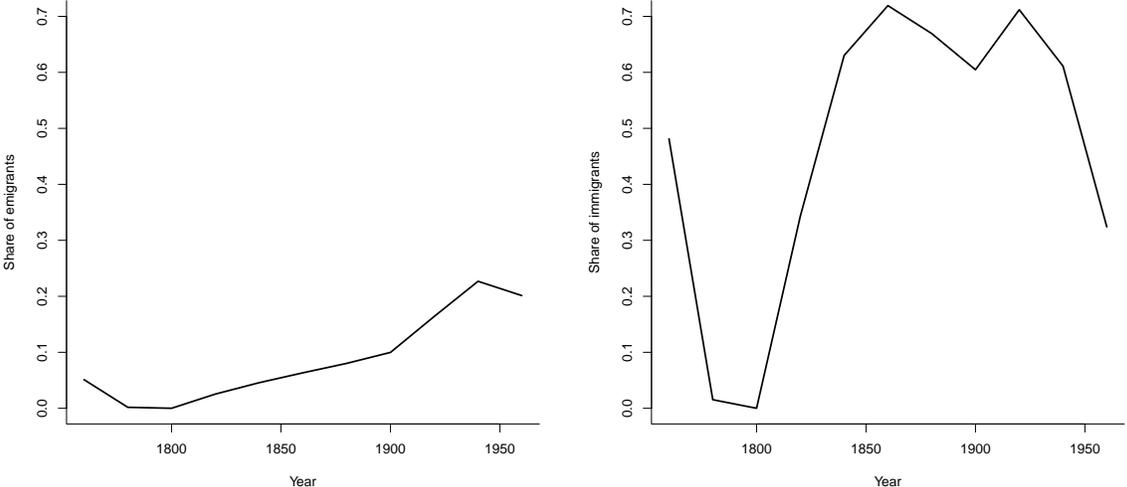


Figure 21: ESTIMATED (BLACK) VERSUS SIMULATED (COLORED) PRIMARY EDUCATION ENROLLMENT RATES FOR VARYING  $\bar{e}'$

### D.4 Estimation of the intensity of the rural exodus



(a) Share of emigrants from the countryside

(b) Share of immigrants in cities

Figure 22: SHARE OF YOUNG ADULTS MOVING RELATED TO THE PLACE OF BIRTH AND RESIDENCE

In Figure 22, we highlight here our estimation of the intensity of the rural exodus along the Swedish take-off. Except for the initial reallocation, the exodus is negligible until the beginning of the 19<sup>th</sup> century. Afterwards, our model predicts a rapid increase in internal migration, a finding that is in line with the efficiency gains in transportation at this time (Berger and Enflo, 2017). Up to 1840, the exodus mainly prevented de-urbanization. Thereafter, it provided the manufacturing sector with the required labor force it needed to enable the acceleration in growth and the industrial revolution. Until 1940, more than 40% of all young adults living in cities were born in the countryside.

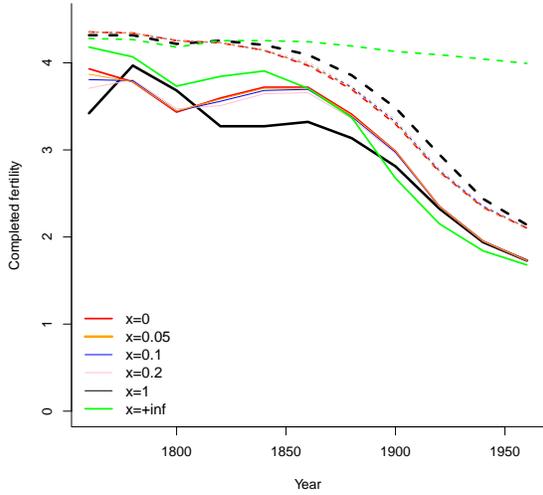
## D.5 Decrease in migration costs

Figure 23 illustrates the rather limited effects of shrinking the costs of migration on fertility, urbanization, education, and growth in GDP per capita. Not surprisingly, this reduction increases urbanization rates compared to the Benchmark scenario and increases fertility in cities as the migrant populations pauperize less. The effects on growth and education are rather limited what suggests that our estimated  $\bar{k}$  is low.

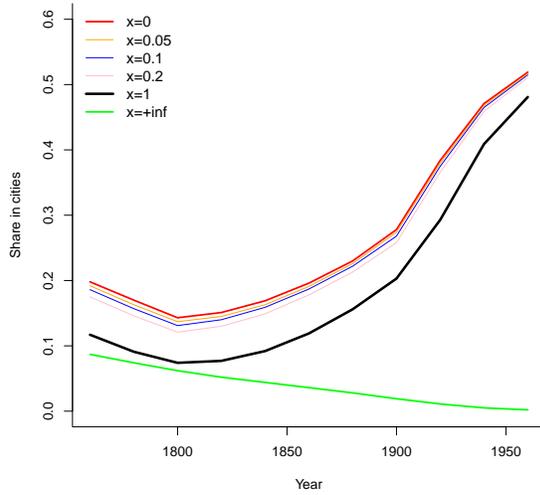
## D.6 More equal distribution of agricultural skills

While we discussed situations with a less equal distribution of agricultural skills in the main paper, this appendix briefly introduces a situation with  $n = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{6}\}$ . As Figure 24 shows, the impact is rather limited compared to the effect of higher levels of inequality.

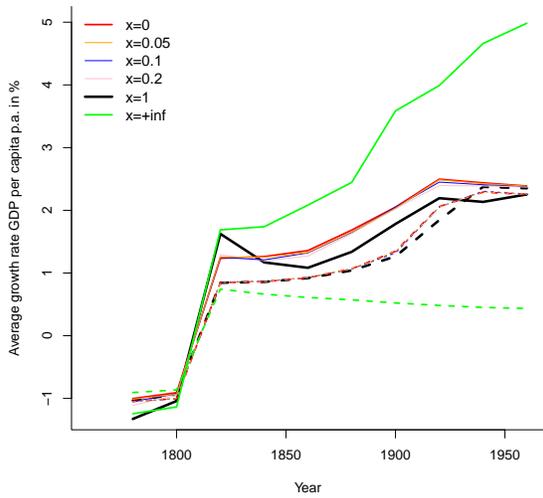
Unlike in the higher inequality scenario, fewer dynasties are willing to leave the countryside at the very beginning. As fertility in rural areas increases slightly, we observe less urbanization. Urban fertility both exceeds and is below the benchmark level for certain periods. Nevertheless, in general, the effects of a more equal distribution of agricultural skills on fertility, as on average growth rates of GDP per capita and the educational revolution, are limited. The educational revolution proceeds even faster than in the benchmark scenario.



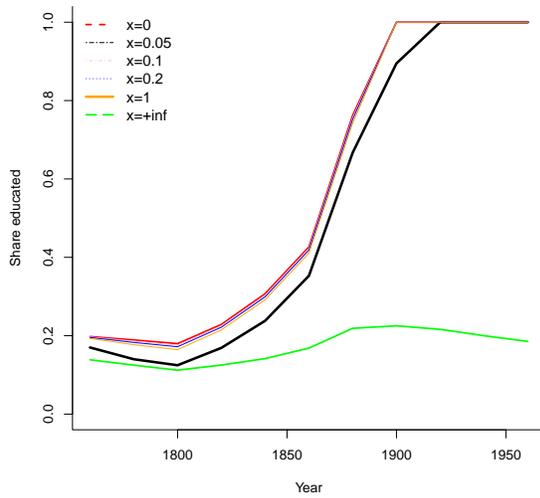
(a) Urban (solid) and rural (dashed) fertility



(b) Urbanization

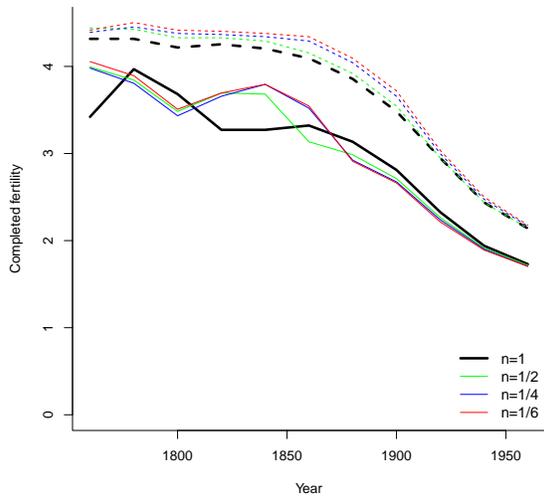


(c) Industrial (solid) and agricultural (dashed) growth rates of GDP per capita

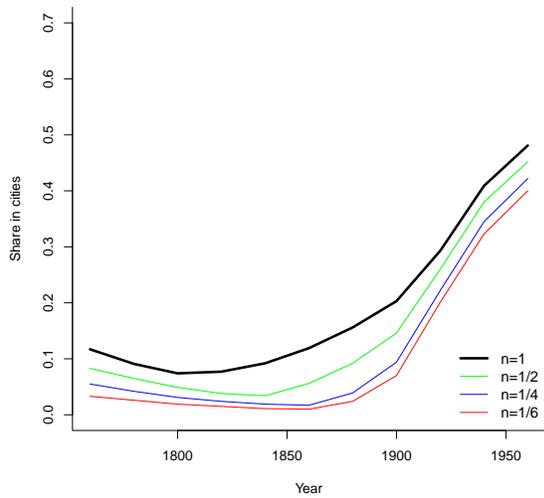


(d) Education

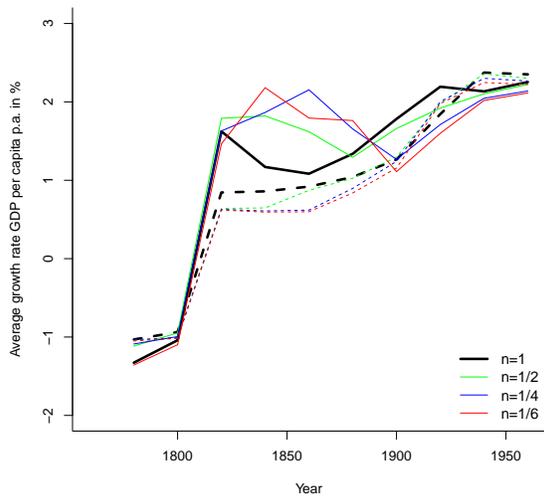
Figure 23: SWEDISH DEMOGRAPHIC AND ECONOMIC DYNAMICS WITH LOW MIGRATION COST SCENARIOS IN 1760–1960



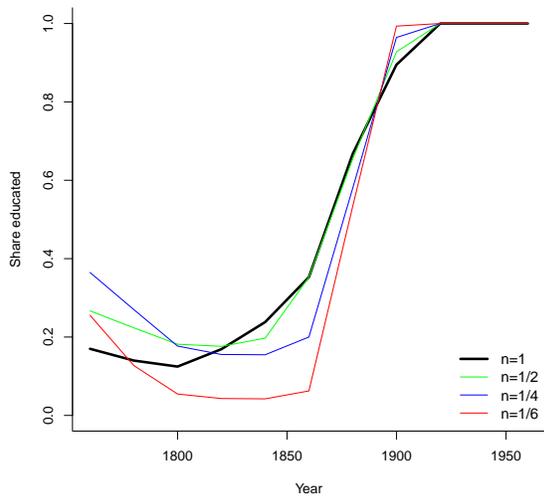
(a) Fertility in cities (solid) and countryside (dashed)



(b) Urbanization



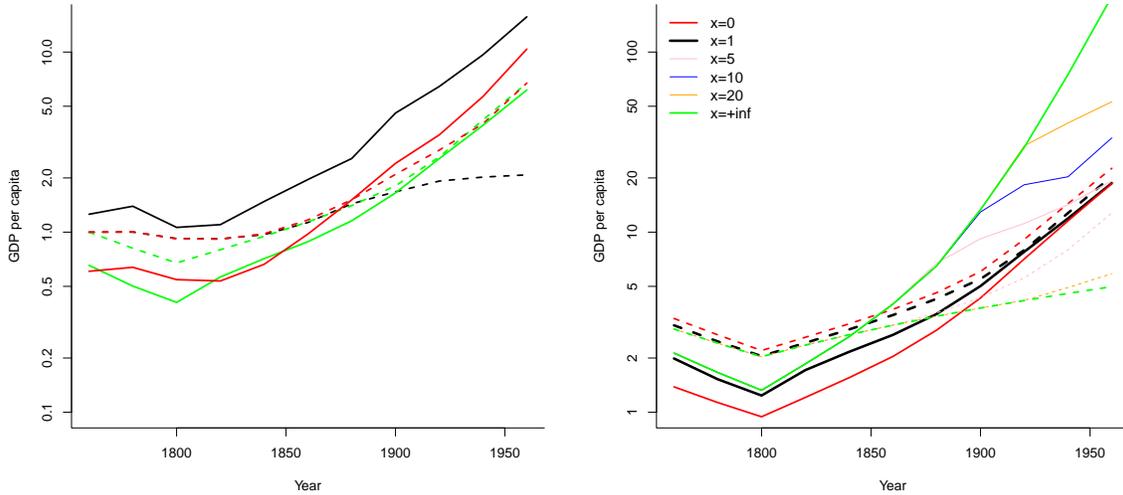
(c) GDP per capita in industries (solid) and agriculture (dashed)



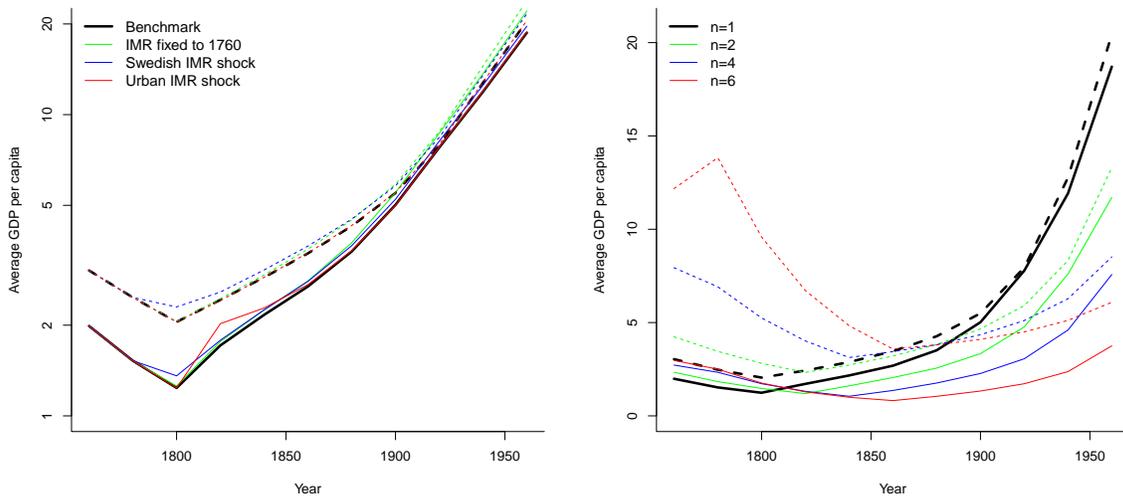
(d) Education

Figure 24: SWEDISH DEMOGRAPHIC AND ECONOMIC DYNAMICS WITH DECREASING INEQUALITIES IN AGRICULTURAL ABILITIES IN 1760–1960

## D.7 Dynamics of GDP per capita



(a) Estimated (black), corrected estimated (red) and predicted (green) average GDP p.c. (b) Average GDP p.c. for varying migration costs



(c) Average GDP p.c. for varying IMR scenarios (d) Average GDP p.c. for increasing inequalities in agricultural productivity

Figure 25: AVERAGE GDP PER CAPITA IN COUNTRYSIDE (DASHED) AND CITIES (SOLID LINES)

Figures 25 describe the dynamics of the average GDP per capita in rural and urban areas in our benchmark scenario (Figure 25a) as well as in our three experiments. In Figure 25a, we compare the observed GDP per capita from our benchmark definition

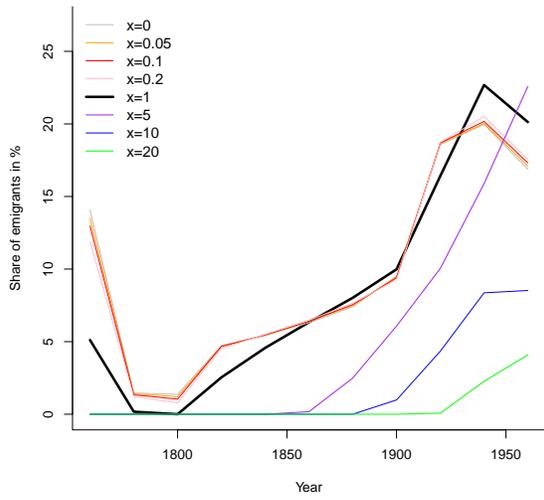
to both its alternative version that takes commuting into account (see Appendix A.4 for an explanation) and our predicted values. While the observed and the predicted values diverge toward the end of our time span, the difference is much smaller once we correct for commuting. In particular, our predictions reproduce very well the corrected estimations in the countryside.

The strong impact of the rural exodus on the average GDP per capita in both cities and the countryside is illustrated in Figure 25b. Increasing the costs boosts income in cities and depresses it in the countryside. The situation is reversed if migration is made less expensive. The average GDP per capita is slightly higher in rural and lower in urban areas. Shocks in infant mortality only slightly affect the average income level. The higher mortality is compensated for by small gains in average GDP per capita, as illustrated in Figure 25c. A less equal distribution of agricultural skills initially increases average GDP per capita in industry and agriculture. Due to the higher share of migrants, the income level drops below the benchmark first in the cities and later on in the countryside as well. A catch-up is observable only toward the end of our time series.

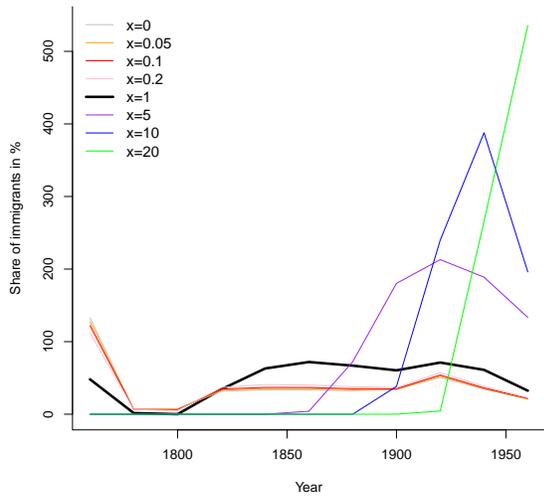
## D.8 Internal migrations

This appendix illustrates the evolution of the intensity of the rural exodus in our benchmark scenario and the three kinds of historical experiments. Figures 26a–28a compare the number of adults who leave the countryside with the number of adults who stay in the countryside. The share of immigrants in Figures 26b–28b adds the newcomers related to the young adults already living in cities.

In particular, changing migration costs and the distribution of agricultural skills strongly alter internal migration. Reducing migration costs increases the exodus at the beginning, while higher costs lead to a postponement of internal migration. As soon as income increases sufficiently to pay  $\bar{\kappa}'$ , we observe a catch-up effect, and the exodus exceeds the initial level; see Figure 26. The findings are similar for the distribution of skills in Figure 28. Greater inequalities lead to an intensification of the exodus at the beginning, while a more equal distribution leads to more internal migration later on. The effects of infant mortality crises are limited overall. Only large reductions in the population of cities after a shock that exclusively hits the urban areas would pull in so many newcomers that we would observe a noteworthy impact on the exodus.

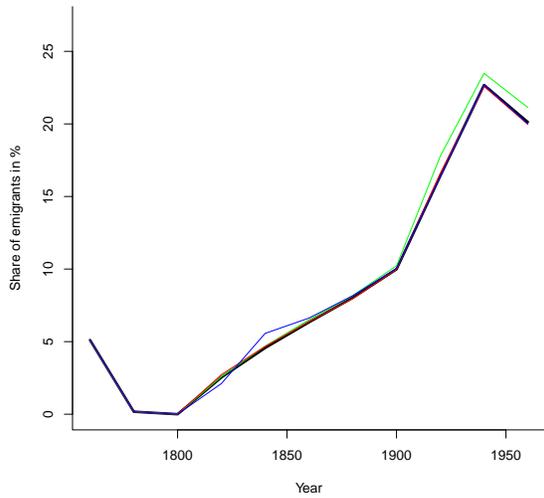


(a) Emigrants

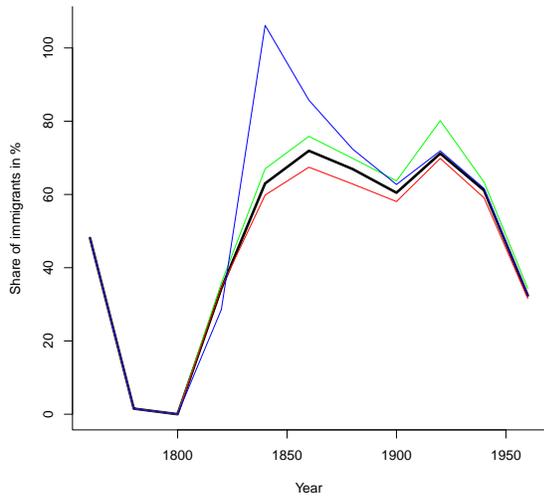


(b) Immigrants

Figure 26: CHANGING MIGRATION COSTS AND INTERNAL SWEDISH MIGRATION



(a) Emigrants



(b) Immigrants

Figure 27: INTERNAL SWEDISH MIGRATION IN THE BENCHMARK SCENARIO (BLACK), WITH CONSTANT 1760 IMR (GREEN), COUNTRYWIDE IMR SHOCK IN 1800 (RED) AND IN SWEDISH CITIES IN 1820 (BLUE)

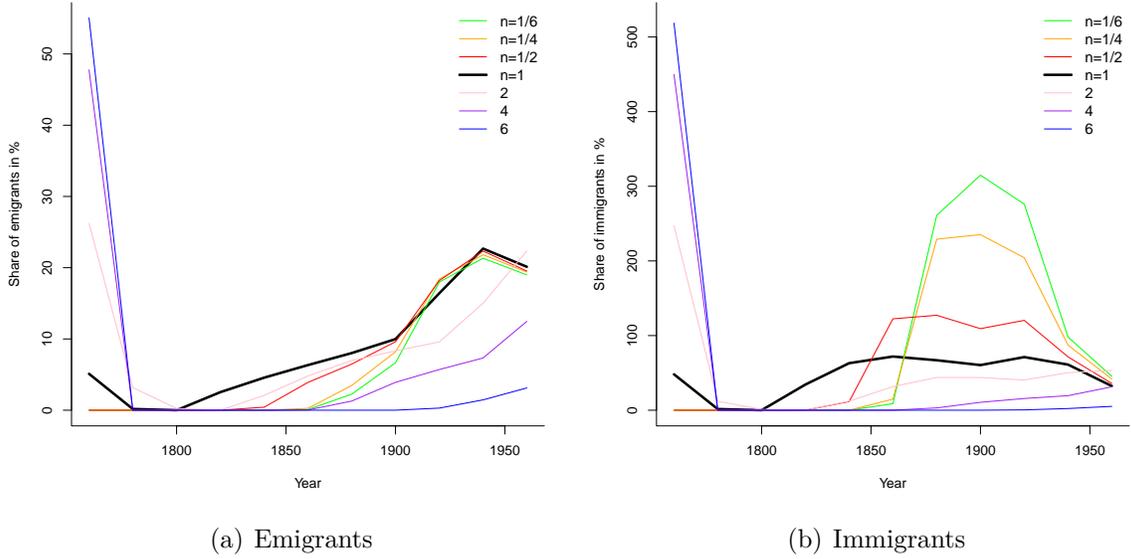


Figure 28: INEQUALITY IN AGRICULTURAL PRODUCTIVITY SKILLS AND INTERNAL SWEDISH MIGRATION

## D.9 Agglomeration effects

Here, we test a model where technological progress starts once the size of urban population, instead of the total population size, reaches  $\bar{N}$ . This new model is identical to the one presented in the main text except that equation (3) is replaced by:

$$\frac{A_{t+1} - A_t}{A_t} = \begin{cases} 0 & \text{if } N_{U,t} < \bar{N} \\ \Gamma > 0 & \text{if } N_{U,t} \geq \bar{N}. \end{cases} \quad (14)$$

We re-estimate our parameters for Sweden, they appear in Table 11. It is reassuring to see that, compared to our benchmark model, most of our parameters get close values. The main difference between the two estimations come from  $\bar{N}$  which is naturally lower in the alternative model as it corresponds to the urban population and not the total population. In fact, because  $\bar{N}$  is below the initial urban population size,  $A_t$  starts to grow as soon as 1760. As a result, incomes grow earlier; and so, the value of some parameters has been adjusted compared to the benchmark estimation, in order to continue replicating fertility and urbanization rates. For instance, the cost of education as well as  $\varepsilon$  have increased in order to maintain fertility high enough.

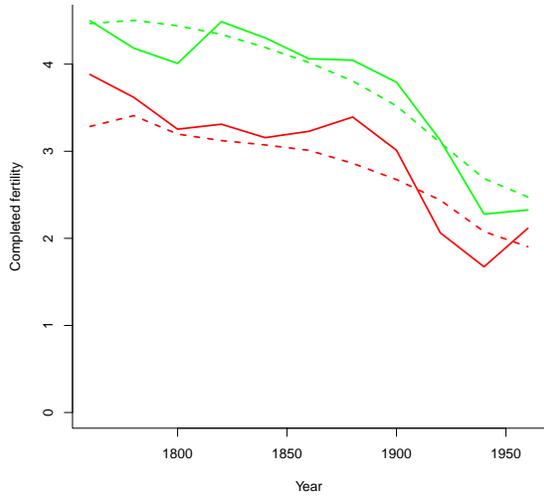
Table 11: PARAMETERS WITH AGGLOMERATION EFFECTS FOR SWEDEN

Parameter	Sym.	Value	Parameter	Sym.	Value
Costs of moving	$\bar{\kappa}$	1.967	Average HK in 1760	$\lambda$	1.985
Critical pop. size	$\bar{N}$	160	Basic human capital	$\nu$	0.299
Growth rate of A	$g$	0.420	Preference for children	$\rho$	0.248
Preference for $c_t$	$\alpha$	0.616	Initial TFP in R	$B_0^R$	251
Cost of education	$\beta$	0.122	Initial TFP in U	$B_0^U$	0.799
Stone-Geary element for d	$\varepsilon$	2.018	Elasticity of HK	$\phi$	0.210
Child rearing cost in R	$\zeta_R$	0.095	Minimum consumption	$\bar{c}$	0.028
Child rearing cost in U	$\zeta_U$	0.139			

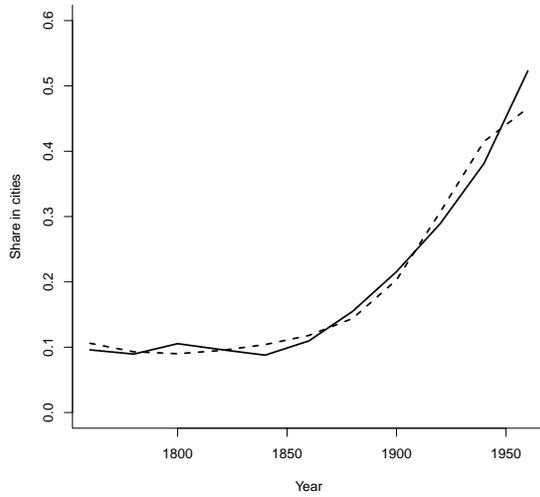
This new estimation set delivers a fit which is very close to the one of our Benchmark model, see Figure 29. We are still able to reproduce the dynamics of fertility and urbanization very well. The same is true for prices while we are not able to reproduce the initial negative growth rates of gdp per capita, contrary to our benchmark model. From this, we conclude that introducing agglomeration effects in the production of knowledge does not improve the explanatory power of our model. It could nevertheless deliver a radically different message regarding the respective effects of migration costs, mortality and agricultural productivity changes. For instance, now, early mortality peaks in cities could keep population size below critical levels such that the take-off to modern economic growth would be delayed if not canceled.

The effect of changes in migration conditions, mortality factors, and the allocation of agricultural skills, using our new estimation, appears in Figures 30 to 32. From a qualitative point of view, our results from Section 5 are kept unchanged. From a quantitative point of view, they are less severe.

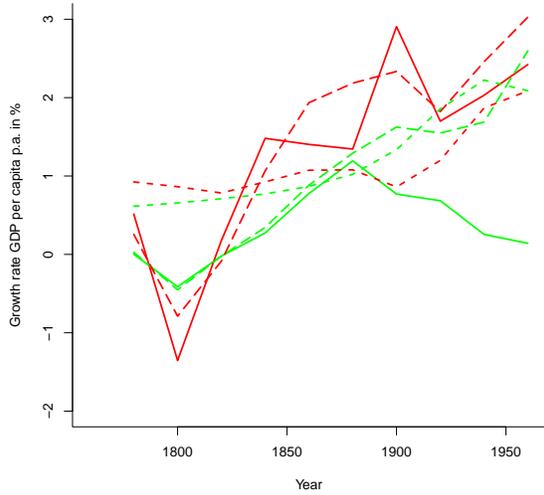
When migration costs increase, the fertility transition as well as the urban revolution appear later and they are then more rapid. The higher the costs of migration, the more inequalities between cities and the countryside progress. Regarding mortality crisis, their effect remain very contained. This result mainly comes from the fact that the initial population size in cities is high enough to guarantee the existence of the take-off to modern economic growth. Finally, an increase of inequalities in the distribution of agricultural abilities continues to provoke a pauperizing exodus from countryside to cities, but this latter is less important compared to our benchmark estimation. This



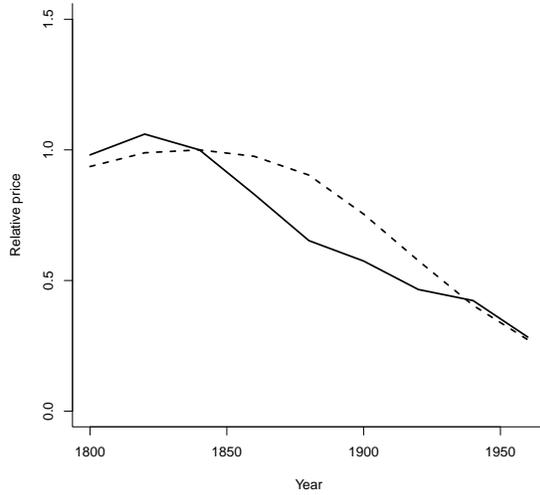
(a) Fertility in cities (solid) and countryside (dashed)



(b) Urbanization

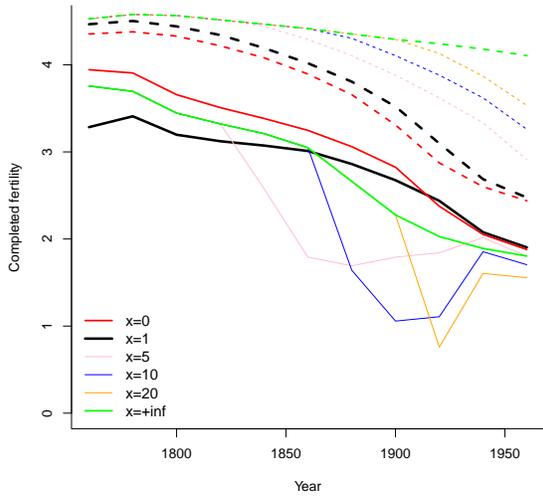


(c) GDP per capita in industries (solid) and the agriculture (dashed)

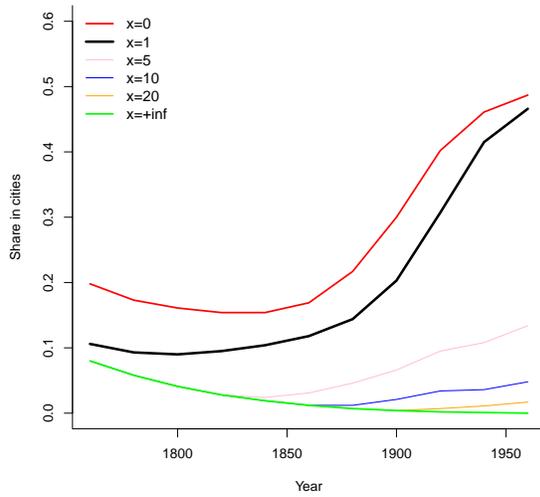


(d) Relative prices

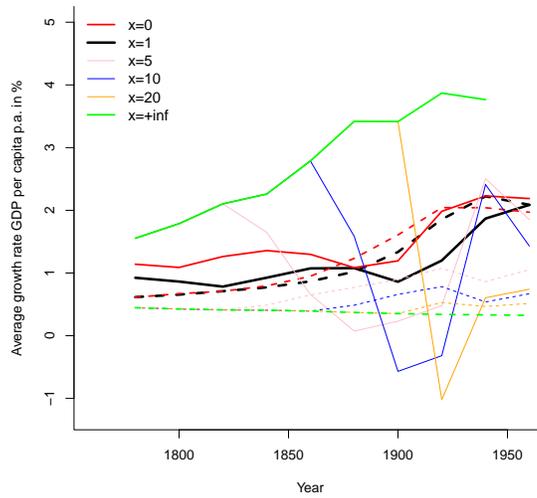
Figure 29: Model's fit and overidentification checks for Prices and Gdp per capita's growth rates



(a) Urban (solid) and rural (dashed) fertility

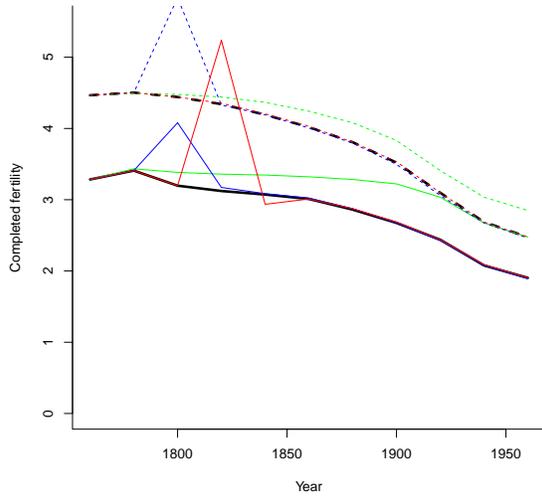


(b) Urbanization

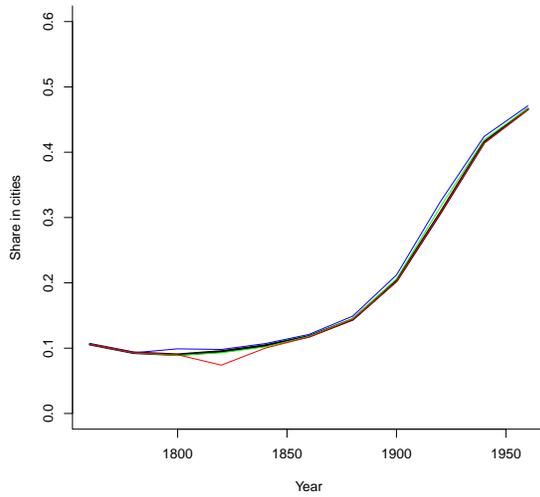


(c) Industrial (solid) and agricultural (dashed) growth rates of GDP per capita

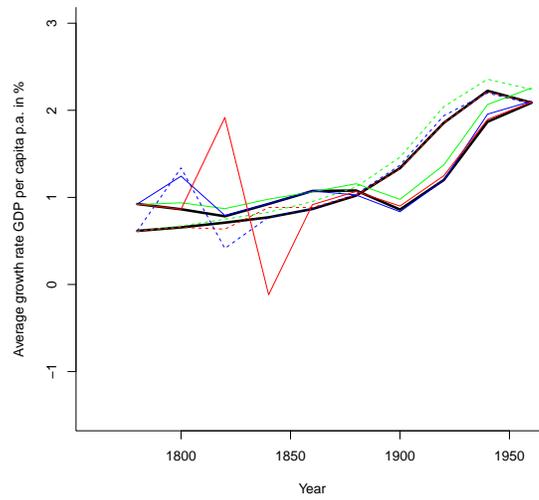
Figure 30: SWEDISH DEMOGRAPHIC AND ECONOMIC DYNAMICS UNDER ALTERNATIVE MIGRATION SCENARIOS IN 1760–1960



(a) Fertility in cities (solid) and countryside (dashed)

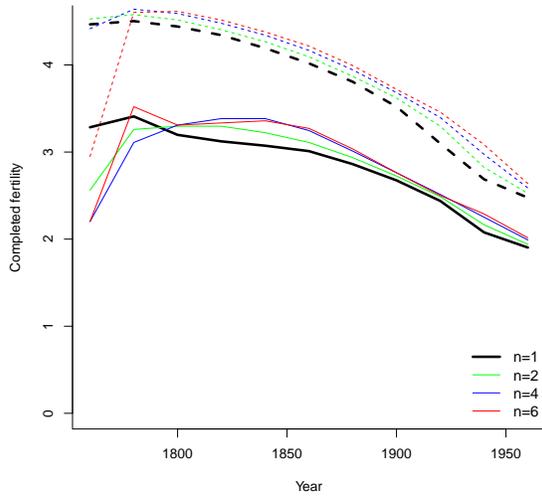


(b) Urbanization

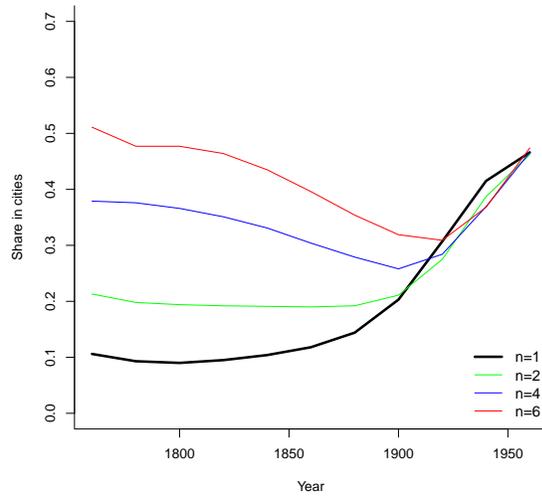


(c) GDP per capita in industries (solid) and the agriculture (dashed)

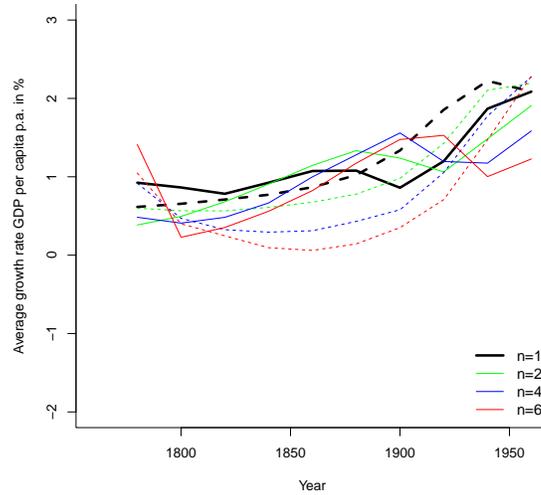
Figure 31: Swedish demographic and economic dynamics in the benchmark scenario (black), with constant IMRs (green), a countrywide shock in IMRs in 1800 (blue), and a shock in cities' IMRs in 1820 (red), 1760–1960



(a) Fertility in cities (solid) and countryside (dashed)



(b) Urbanization



(c) GDP per capita in industry (solid) and in agriculture (dashed)

Figure 32: SWEDISH DEMOGRAPHIC AND ECONOMIC DYNAMICS WITH INCREASING INEQUALITIES IN AGRICULTURAL ABILITIES IN 1760–1960

difference mainly comes from the fact that economic growth appears earlier here, what allow to alleviate pauperization. Regarding fertility, we continue to observe an initial reduction of the number of births while regarding urbanization, the U-Shape pattern is still there; nevertheless, those two movements are less marked. The take-off to modern economic growth remains delayed by the severity of the inequality shock.