The rural exodus and the rise of Europe*

Thomas Baudin† and Robert Stelter‡

December 13, 2021

Abstract

We propose a unified model of growth and internal migration and identify its deep parameters using an original set of Swedish data. We show that internal migration conditions had to be favorable enough to induce an exodus out of the countryside in order to fuel the industrial development of cities and the demographic transition of the country. We then compare the respective effects of shocks to internal migration costs, to infant mortality and to the productivity of rural industry to the economic take-off and demographic transition that occurred in Sweden. Negative shocks to internal mobility generate larger delays in the take-off of growth than do mortality shocks equivalent to the bubonic plague. These delays are dramatic when rural industry is less productive in the early phase of industrialization. The economic and demographic dynamics of Sweden were very similar to those of the rest of Europe at the time of industrialization, which allows us to generalize our findings.

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

JEL Classification Numbers: J11, J13, O41.

---

*This paper has 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 the participants of seminars and conferences in Louvain-la-Neuve, Paris, Hamburg, Aix-en-Provence, Moscow, Rostock, Nanterre and Brussels for their helpful comments. We would like to thank Annika Westberg for helping us with the SHiPs data. We are deeply grateful for the computational infrastructure provided by the Max Planck Computing and Data Facility and the assistance offered by Markus Rampp. This project has benefited from the scientific and financial support of ARC project 15/19-063 on “Family transformations” (French-speaking community of Belgium) and Project MALYNES (ANR-18-CE26-0002). Robert Stelter acknowledges financial support from the Max Geldner foundation. The traditional disclaimer applies.

†IÉSEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM - Lille Economie Management, F-59000 Lille, France and IRES, Université catholique de Louvain, E-Mail: t.baudin@ieseg.fr, corresponding author

‡University of Basel, Faculty of Business and Economics and Max Planck Institute for Demographic Research, E-Mail: robert.stelter@unibas.ch.

1 Introduction

Among historians and demographers, the rural exodus is reputed to have been one of the key forces behind the development and urbanization of European economies in the nineteenth and twentieth centuries (Saville, 2013; van der Woude *et al.*, 1995). Surprisingly, to the best of our knowledge, economists have never deeply questioned this qualitative result. We fill that gap by estimating the contribution of the rural exodus to the rise of European economies. To do so, we build a unified model of growth and internal migration and estimate its deep parameters. Using an original dataset on Sweden, we show that the conditions for migration between the countryside and cities have strongly shaped the timing and intensity of the transition to growth. Among these conditions, we highlight migration costs, mortality differentials between areas and the initial productivity of rural industries. We obtain two main results. First, shocks that increase migration costs generate larger delays in the take-off to growth than do mortality crises equivalent to a bubonic plague or influenza pandemic. Second, the negative effects of high internal migration costs are significantly attenuated by higher initial levels of productivity in rural industries. The more productive the latter were, the less economies needed to rely on the rural exodus and urbanization to start industrializing.

We first collect data on Sweden from 1760 to 1960. From this data collection, we identify a series of facts in three domains: demographics, economics and education. First, compared to the countryside, cities were experiencing higher child and infant mortality rates as well as lower fertility rates. While the mortality differential vanished smoothly during the process of industrialization, fertility differentials were more persistent and started to vanish only in approximately 1870. Urbanization, which we define as an increase in the proportion of people living in cities, started to progress after 1850.[1] Second, if average GDP per capita also started to grow only around 1850, the growth rates of total production became higher in industries than in agriculture as early as 1820. At that date at which a clear break in industrial growth can be detected. Nevertheless, total industrial production first exceeded total agricultural production around 1900. Importantly, industrialization occurred in two phases: until the turn of the twentieth century, rural industries triggered this process, while after 1900, urban industries became its epicenter. Cities progressively became the leading economic areas. Indus-

[1] In 1850, approximately 10% of the Swedish population was living in cities; however, by 1960, this share had increased to 52.3%.

trialization was accompanied by a continuous decrease in the price of industrial goods relative to the price of agricultural goods. Third, an educational revolution took place after 1850, causing the enrollment rate in primary education to explode until 1960.

Our database is the first to demonstrate these facts in a unique setting. To the best of our knowledge, Sweden is —along with Denmark— the only country that offers reliable information for consistently computing fertility and mortality in cities and the countryside before the Industrial Revolution. We focus on Sweden, as data are available over a longer timespan; see Baudin and Stelter (2021) for additional information about Denmark. Our data highlight a key historical fact: both urbanization and educational improvements seem to have resulted from the pre-existing differences in economic growth rates between industry and agriculture. The theory we develop is in line with this empirical finding.

To demonstrate 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 agricultural production takes place exclusively in rural areas, while industrial production may occur in both cities and the countryside. What distinguishes these two types of industrial production is the technology they employ: urban industries are the only industries that produce with human capital. Each adult is endowed with a certain level of human capital and an intrinsic ability to work in the rural sectors.[2] There is perfect mobility of labor between the two rural sectors. Individuals can decide to stay in their area of birth or to pay a cost to move to the other area. However, they can choose their place of residence only once, at the beginning of adulthood. Return migration is excluded, and mortality is area specific. Once individuals have settled down, they decide on the number of children to have, the educational level they want their surviving children to achieve, and the levels of agricultural and industrial goods they consume.

In the early stages of our economy, the society is trapped in a Malthusian regime.[3] It endogenously starts to experience global technological progress once the overall workforce has reached a certain threshold. As a result, an increasing number of parents become willing and able to educate their children, attracting such people to urban ar-

[2]If both agriculture and rural industries use rural abilities, they do so under different technological regimes. Skills are not initially equally distributed among adults, such that our framework is a heterogeneous agents model.

[3]Interestingly enough, Edvinsson (2017) evidences the existence of a Malthusian regime in Sweden before 1800 while positive checks may have persisted even until 1920.

eas and magnifying the growth in productivity in urban industries. This is the key role our model gives to the rural exodus: as natural movements of the population cannot sustain urbanization, generations of rural workers have to migrate to urban areas to fuel the virtuous circle of urbanization and human capital-based industrialization. If they do not, the second phase of industrialization, based in urban areas, may not be sustainable, clearly and permanently reducing development prospects.

What factors might prevent or weaken the rural exodus? One potential factor is the mortality penalty in cities. Mortality peaks in cities could significantly lower the incentives to settle in urban areas. The economic cost of the exodus to cities may also be crucial. When migration costs 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 rural workers to urban areas. Finally, living in a prosperous rural economy may reduce the incentive to move to cities. Thus, the more productive rural industries are initially, the less intense the rural exodus is likely to be.[4] These are the three main mechanisms that may prevent the rural exodus from fueling human capital-based industrialization in our setup, and these are the three main mechanisms we quantify in this paper.

We calibrate the parameters of our model to match the fertility rates in cities and the countryside as well as urbanization rates and the share of rural workers involved in industry 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 other economic moments, such as shifts in GDP per capita, changes in overall education, or changes in the relative prices of industrial goods. However, our overidentification checks indicate that our model replicates these dynamics convincingly.[5] Such a result guarantees that our estimation strategy delivers reliable results and allows us to proceed to an evaluation of the importance of the conditions that governed the exodus out of the countryside.

In our first experiment, we focus on the importance of productivity in rural industries during the eve of the Industrial Revolution. The first phase of European industrialization is marked by the importance of these industries in the countryside, be they cottage industries or larger industries.[6] This seems especially true for Sweden, as discussed by

[4]Considering income differentials between cities and the countryside as a driving force for urbanization is in line with most of the literature on the topic. See, for instance, Gollin *et al.* (2016).

[5]In Section 4.5, we discuss the reasons why we do not try to match additional moments. The main reason pertains to issues with the definitions of GDP per capita and school enrollment rates.

[6]The development of rural industries during the eve of the Industrial Revolution has often been

Schön (2010). We simulate our economy under alternative initial productivities in 1760s' rural industry. When these industries are initially very productive, they are better able to sustain the first phase of industrialization and cause rural areas to be relatively richer. Nevertheless, they also make urban places poorer, as urban areas suffer from decreases in the relative price of industrial goods. Perhaps more surprisingly, higher initial productivity in rural industries also impoverishes the entire country in the long run. Indeed, the initial distortion in local incomes translates into depressed fertility rates in both urban and rural places, delaying the onset of technological progress. As a result, the higher the initial productivity in rural industries is, the lower the growth rates of GDP per capita are over the period 1820–1900 and the more time the economy needs to reach its long-run growth path and.

Our second experiment evaluates the effects of migration costs on the economic and demographic revolutions in Sweden. Alone, high migration costs, even infinite ones, would not have prevented the occurrence of the first (rural) phase of industrialization, and they would not have prevented the fertility transition from setting in. Nevertheless, they would have had detrimental effects in the long run, as the second phase of the Industrial Revolution would have been strongly affected. In our benchmark estimation, by 1760, the cost of migration represented 28.4% of the average potential income of industrial workers. Multiplying this cost by five would have magnified fertility differentials between urban and rural places during the fertility transition. Furthermore, the urbanization process would have been severely delayed, as by 1900, urbanization rates would not have exceeded their reference level in 1760. These two movements would have contributed to the reinforcement of the demographic pressure on rural areas, maintaining their populations at low levels of productivity and thus depressing their incomes. As a result, the higher the internal migration costs are, the stronger the economic inequalities between the two areas.

In a complementary analysis, we show that the detrimental effects of high migration costs are magnified when the initial productivity of rural industries is reduced. We show, for instance, that if their initial productivity had been divided by 10, rural industries would not have been able to sustain the first phase of industrialization, giving a much more crucial role to urban industries. Thus, in a scenario in which internal migration

qualified as protoindustrialization after Mendels (1972). Economic historians have had severe disagreements on the reality of this phenomenon, but they have all agreed on the fact that prior to the so-called Industrial Revolution, a non-negligible portion of industrial work was located in rural areas. See Ogilvie and Cerman (1996) for a complete discussion of the concept of protoindustrialization, as well as Section 6.

and urbanization are impossible (infinite migration costs), Sweden's development would have been totally different: no fertility transition would have occurred in rural areas by 1960, the rural population would have exploded at the price of severe pauperization, and rural-urban income inequalities would have climaxed. Our experiments therefore suggest that more than the initial productivity level of rural industries or migration costs alone, what truly mattered for the transition to modern economic growth was the interplay between these two elements. To the best of our knowledge, we are the first to demonstrate and quantify such a result.

Our last experiment relates to mortality crises. We simulate the dynamics of our economy in a case in which infant mortality rates (IMRs), rather than decreasing, were frozen between 1760 and 1960. Such stagnation would have left the economic dynamics almost unchanged, while compared to those in the benchmark scenario, the number of births in this scenario is greatly increased and the number of surviving offspring is decreased. Equivalent results are found when simulating a mortality shock equivalent to the bubonic plague for the year 1800. This rather weak effect of mortality variations is in line with the results of Galor (2005) and Doepke (2004)[7] and contradicts traditional demographic theories positing that the mortality decline was the main driver of the fertility decline and the associated economic take-off.

The model we propose contributes to the unified growth theory (hereafter, UGT) initiated by Galor and Weil (2000) and Galor and Moav (2002). 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. Recent literature has suggested that certain conditions must be met for the general mechanisms of the UGT to operate. For instance, Squicciarini and Voigtländer (2015) show that rising total factor productivity is not necessarily sufficient to induce human capital accumulation and an economic take-off. They highlight the role played by upper-tail human capital and scientific knowledge.[8] We complement this literature by emphasizing how the rural exodus had to be affordable for industrialization—at least its second phase—to occur.

Together with Adams (2021) and Greenwood and Seshadri (2002), we are among the few

[7]Some papers have documented a significant impact of other types of mortality, such as young adult mortality and longevity, on the economic take-off. See, among others, the enlightening papers of Tamura (2002), Cervellati and Sunde (2011), Strulik and Werner (2016), Boucekkine *et al.* (2002), Strulik and Weisdorf (2014) and Voigtländer and Voth (2012).

[8]A large number of other refinements to unified growth theory have been produced. Galor (2011) provides an enlightening review of this literature.

economists to investigate the role of urban-rural differential demography in the process of economic development. With English data, Adams (2021) calibrates the parameters of a growth model in which rural-urban differences are key.[9] He then predicts that a country will grow less rapidly if it has a higher initial urbanization rate, a prediction validated through cross-sectional data. Greenwood and Seshadri (2002) focus on the rural-urban fertility differential in the US. They find that technological progress was essential for the reduction in fertility in rural and urban sectors, and we complement their approach by studying the precise effects of the rural exodus on the emergence of technological progress and growth from both a theoretical and a quantitative perspective. Finally, our paper is also in line with Brunt and García-Peñalosa (2021), who provide evidence that urbanization caused technological progress throughout the Industrial Revolution, one of our core mechanisms.

Our paper is organized as follows. Section 2 presents the empirical regularities characterizing the Swedish economic take-off from the eighteenth to the twentieth century. Section 3 proposes our unified model of growth and migration, while Section 4 estimates the deep parameters of this model. We investigate the effects of shocks to productivities in rural industries, to migration costs and to mortality rates in Section 5. Importantly, in Section 6, we show both how the facts we have highlighted for Sweden are representative of most Western European countries and how our results can be generalized to these countries. Section 7 concludes.

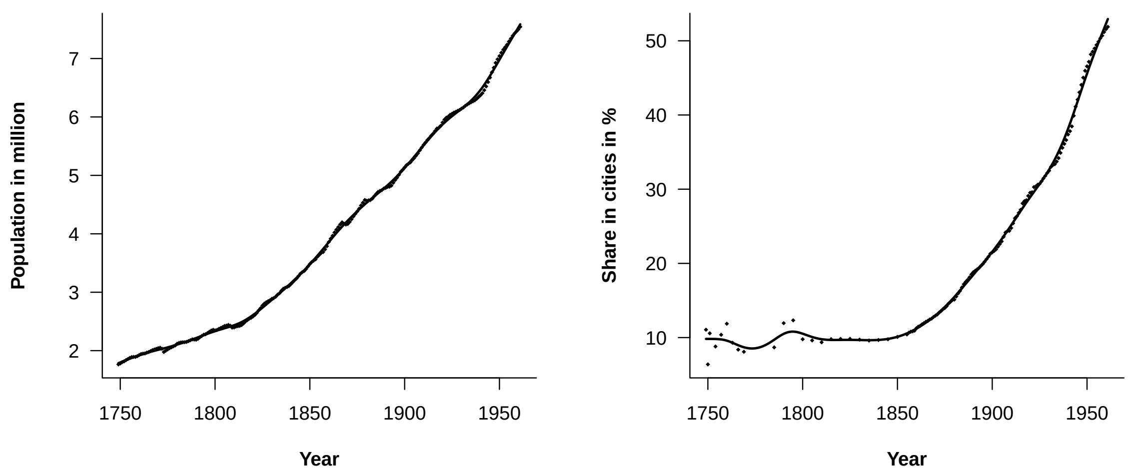
2 Data and facts

In the twentieth century, demographers and economists collected historical data throughout Europe at the national and 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 a series of facts along three dimensions: demographics, economics and education.[10]

[9]Other papers have applied calibration techniques to UGT. Lagerlöf (2006) presents a numerical exercise conducted with the model of Galor and Weil (2000) by choosing parameters either from the literature or currently observed moments. Cervellati and Sunde (2005) and Strulik and Weisdorf (2014) present illustrative numerical simulations that highlight the respective roles of longevity and child survival. Cervellati *et al.* (2013) use panel data for 59 countries beginning in 1880 to structurally estimate a unified growth model.

[10]Data comparability issues necessarily arise. In Appendix A.1, we describe the general procedures we used for data harmonization and smoothing. All statistics are computed using current Swedish

Demographic facts. We gather data from Historisk statstik för Sverige (Centralbyrån, 1969), Swedish Historical Population Statistics (SHiPs) and various Statistical Yearbooks. A key feature of these three sources is that they allow the use of a common and time-consistent definition of urban and rural places following Centralbyrån (1969). The definition of cities therefore follows legal principles and administrative terms rather than population numbers. Receiving a city charter granted certain rights, for instance, to trade and artisanship. These city privileges define the division of urban ("Städer") and rural ("Landsbygd") places.[11] By not relying on population size or density in our dataset, we limit the need for reclassification and its related biases; see Bocquier and Costa (2015) for a discussion.



(a) Population (b) Share of population in cities

Figure 1: Observed (dots) and smoothed (solid lines) population dynamics in Sweden.

We define the urbanization rate as the ratio of persons living in cities to the total population. Historisk statstik för Sverige (Centralbyrån, 1969) provides information on rural and urban population sizes from 1800 to 1962. Information is available every five years from 1800 to 1855 and on an annual basis beginning in 1856. To compute the urban and rural populations for earlier periods, we use the SHiPs dataset. SHiPs contains a rich set of demographic variables for the 1749–1859 period for approximately 3,000 parishes that we have been able to classify as urban or rural, again following

borders.

[11] Centralbyrån (1969) provides an overview of Swedish cities and their year of foundation.

Centralbyrån (1969).[12] As the proportion of the Swedish population covered by the SHiPs data varies greatly from one year to another, we decided to use only data from years that cover at least 60% of the Swedish population.

In Figure 1, we illustrate the dynamics in both the Swedish population size and its urbanization rate. 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 at approximately 10% over the first 100 years of the observation period.[13] In the middle of the nineteenth century, urbanization took off. From 1850 to 1960, the share of individuals living in urban areas increased to over 50%.

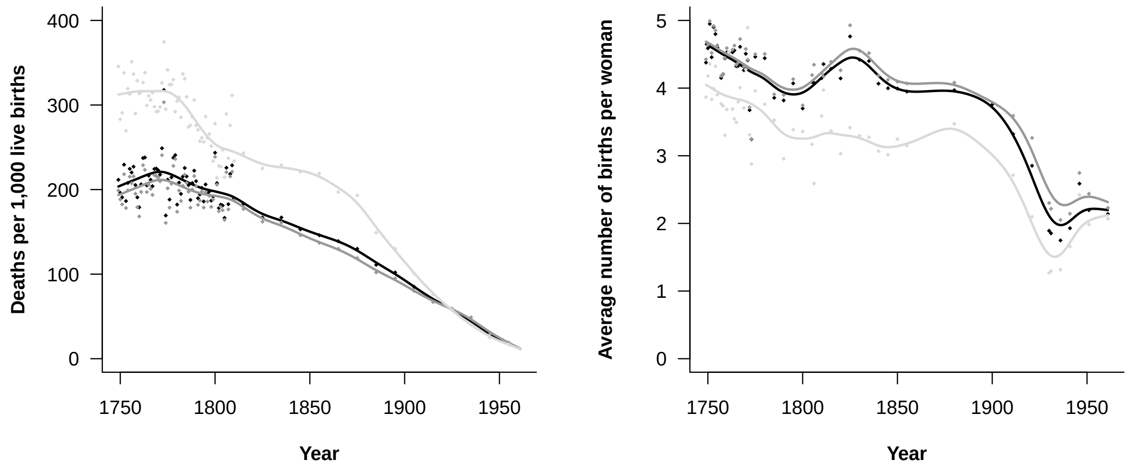
Demographers such as Woods (2003) have emphasized that cities experienced a mortality penalty relative to the countryside at the time of the Industrial Revolution. In a perfect framework, we would reconstruct the dynamics of mortality for all ages to obtain a complete picture of this phenomenon. However, restrictions on the availability and the quality of data do not allow us to build such a time series for a country such as Sweden in the eighteenth and nineteenth centuries. Nevertheless, by using data from the SHiPs database and Historisk statstik för Sverige, we are able to examine the local and global dynamics in infant mortality rates, which are important determinants of the number of births within a family, as well as of the number of surviving children.[14]

Figure 2a documents high infant mortality rates in the middle of the eighteenth century. 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 overall was quite linear for 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 eighteenth century and decreased the rural-urban mortality gap to five percentage points. Thereafter, however, the gains to survival probability in rural areas exceeded those in cities. By the middle of the nineteenth century, the differential had expanded to 10 percentage points. After this point, there was a sustained trend toward convergence. By the early twentieth century, the rural-urban IMR gap had vanished.

[12]This exercise has been possible thanks to the help of Umeå University and its retrieval tool, discussed in Appendix A.2.

[13]Schön (2010) also documents a share of 10% of the total population living in towns and cities for the first half of the nineteenth century.

[14]Brändström (1993) presents Swedish infant mortality rates at the provincial level and evidence on intra-rural and intra-urban differentials.



(a) Infant mortality rate (b) Total fertility rate

Figure 2: Smoothed (solid lines) and observed (dots) infant mortality and fertility in Sweden overall (black), in the countryside (dark gray) and in cities (light gray).

Looking at the population age pyramids (Figure 18 in Appendix A.3), we see that the respective age structures in cities and in the countryside were very different. Using the crude birth rate (CBR) to measure local fertility rates may therefore lead to biased estimations, as the number of women of fertile ages living in these areas differed greatly. For this reason, we compute the general fertility index developed by Coale (1969) and translate its values into total fertility rates (TFRs).[15] We merge data from SHiPs for the period 1749–1855 and from several statistical yearbooks for the period 1910–1960. As data on the age structure in specific areas are not available for the 1855–1910 period, we have chosen to use the calculation from 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 1749 was equal to 4.6. Fertility was slightly higher in rural areas (4.7) and lower in urban areas (4.1). Thus, rural fertility exceeded urban fertility by approximately 15% at the beginning of our observation period. From 1750 to 1850, fertility in the countryside fluctuated without clearly decreasing, which is a key characteristic of Malthusian regimes. In the same period, urban fertility experienced its first significant decline and then stagnated. From approximately 1880 to the end of our observation period, fertility in both areas decreased, despite having reached a

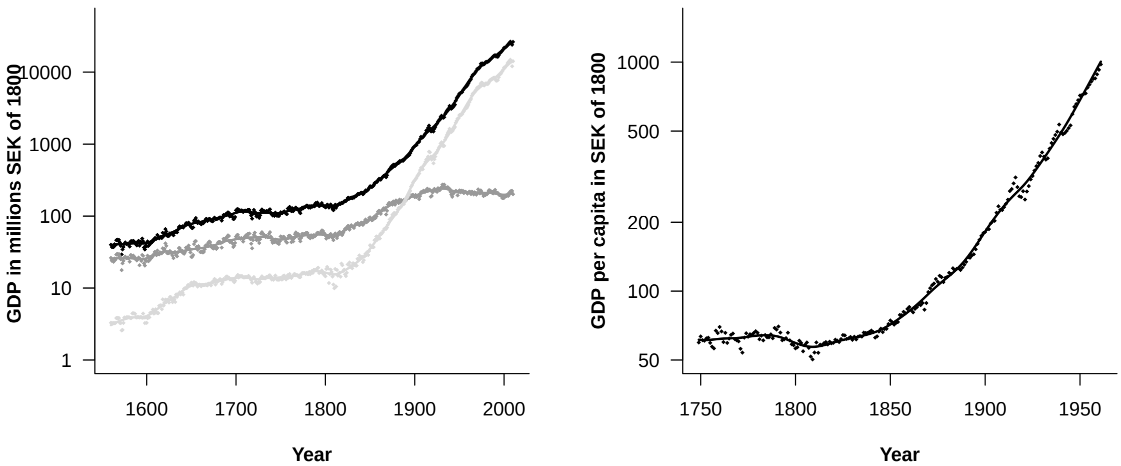
[15] We discuss the method and the advantages of using alternative fertility measures in Appendix A.3.

maximum differential of close to 60% around 1930. At that point, both fertility rates started to increase and converged around 1960.

As 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. However, Figure 1b shows that the opposite was the case.

Economic facts. The Swedish Historical National Accounts (SHNA) provide information on production per sector in Sweden since 1560. Economic activity is 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 and Krantz (2012). We use the sector *agriculture* to approximate agricultural output and the sector *manufacturing and industry* to approximate industrial production. We do not include production data from any of the other sectors. In other words, our definition of industrial production is quite conservative. Time series are available in constant factor prices since 1560 and in current prices since 1800 (mill. SEK).

Figure 3 shows the evolution of Swedish GDP in constant prices of the year 1800 nationally and by sector. As our data start very early, we are able to observe a period of stagnation even though aggregate production started to grow slightly as early as 1610, mainly driven by industrial production. At the end of the sixteenth century, agricultural production was approximately 6 times higher than industrial production. This difference decreased to a factor below 4 by 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 a constant annual rate of 3.5% from 1820 to 2010. Agricultural production also increased but started to stagnate in approximately 1930. As a result, total GDP started to grow sustainably in 1820. Nevertheless, as shown in Figure 1a, population was also rapidly growing at the same time such that total GDP growth translated into sustained growth in GDP per capita only around 1850 (see Figure 3b).



(a) GDP

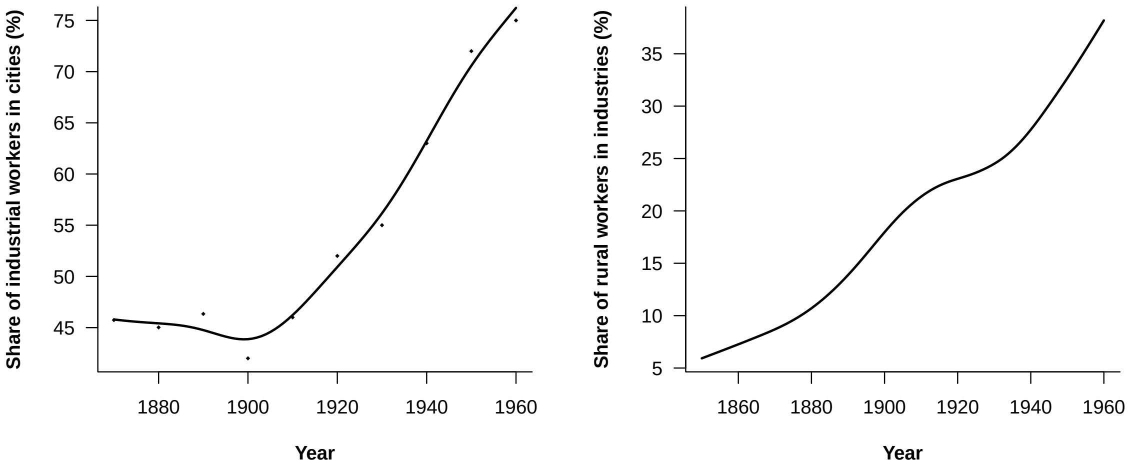
(b) GDP per capita

Figure 3: Observations of (dots) and smoothed data on (solid lines) total GDP and average GDP per capita at the national (black), industrial (light gray), and agricultural (dark gray) levels.

Combining data from SHNA, from Berger *et al.* (2012) and from Centralbyrån (1969) enables us to estimate the allocation of labor across areas and sectors.[16] Our first key finding is that among agricultural workers, more than 90% were located in the countryside throughout our observation period. Later, this motivates our theoretical assumption that agricultural production occurs exclusively in the countryside. Our second key observation appears in Figure 4, in which we document two main phases in Swedish industrialization. In the first phase, from the middle of the nineteenth to the early twentieth century, industrialization was mainly driven by the countryside. Indeed, in a context in which urbanization was already increasing significantly, the share of industrial employment located in cities decreased in favor of rural places (Figure 4A). This testifies to the stronger vitality of rural industrialization than of its urban counterpart until 1900. In the second phase, cities again became the center of industrial development. Despite the fact that industrial labor continued to crowd out agricultural labor in the countryside, the share of industrial workers located in cities increased by three quarters around 1960. This two-phase episode finds a remarkable echo in the rich literature on Swedish industrialization; see, for instance, Berger *et al.* (2012), Berger (2019), Söderberg (1984) and the rich set of contributions by Schön (2008, 2010, 2012).[17]

[16]See Appendix A.4 for details.

[17]In his contribution to the EH.net Encyclopedia, Schön (2008) argues that "*During the first half*



(a) Location of industrial workers

(b) Share of rural workers in industry

Figure 4: Allocation of workers across industries and locations

As the SHNA provide the number of workers per sector and GDP from 1850 onward, we can compute the dynamics in Sweden's GDP per worker (Fig. 5a). In both sectors, we observe an almost constant positive growth rate. With an average annual growth rate of 2.1 %, industrial GDP per worker clearly outperformed that in the agricultural sector (1.4%).[18] Industrialization was accompanied by a strong decrease in the prices of manufactured goods relative to the prices of agricultural goods (Fig. 5b). Still using the SHNA and beginning in 1800, we divide GDP in current prices (Y_t^s) by GDP in constant prices ($\bar{Y}_t^s$) in the sectors (s) agriculture ($s = A$) and industry ($s = I$):

$$p_t^A = \frac{Y_t^A}{\bar{Y}_t^A} \quad \text{and} \quad p_t^I = \frac{Y_t^I}{\bar{Y}_t^I}$$

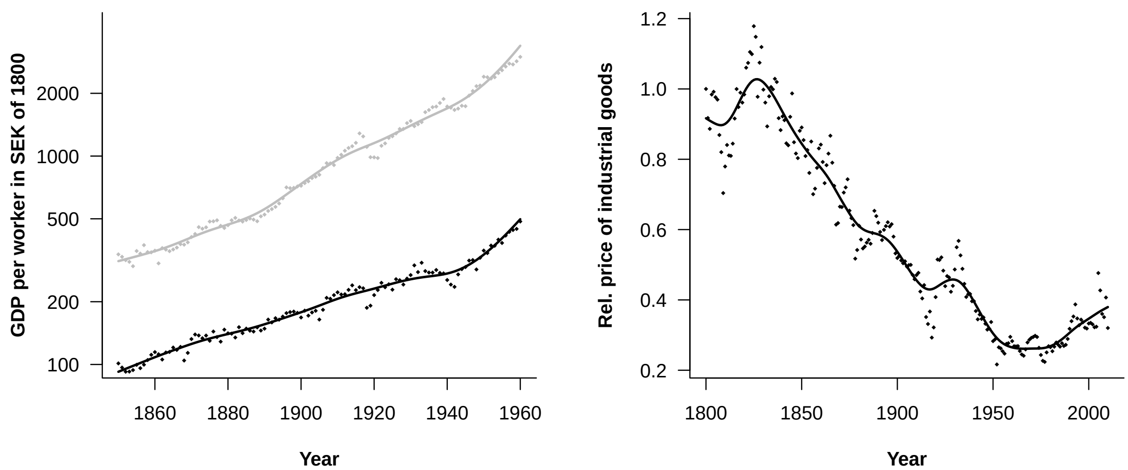
Then, the dynamics of the relative price of industry in terms of agricultural goods follow:

$$p_t = \frac{p_t^I}{p_t^A}.$$

of the 19th century (...) many non-agrarian activities such as the iron industry, the saw mill industry and many crafts as well as domestic, religious and military services were performed in rural areas." This argument is further developed in key contributions such as Schön (2012) and Schön (2010).

[18]Even if our theory does not offer any predictions regarding wages per sector, the latter may be used to evaluate whether our presentation of the timing of growth is plausible. In Figure 20 of Appendix A.5, we present wages by sector, normalized to 100 in 1800, and we can again identify the fact that industry started to outperform agriculture around 1820, at least in terms of growth rates.

During the period of stagnation at the beginning of the nineteenth century, the relative prices of industrial goods first slightly increased and then declined after 1825.



(a) GDP per worker in agriculture (black) and industry (gray).

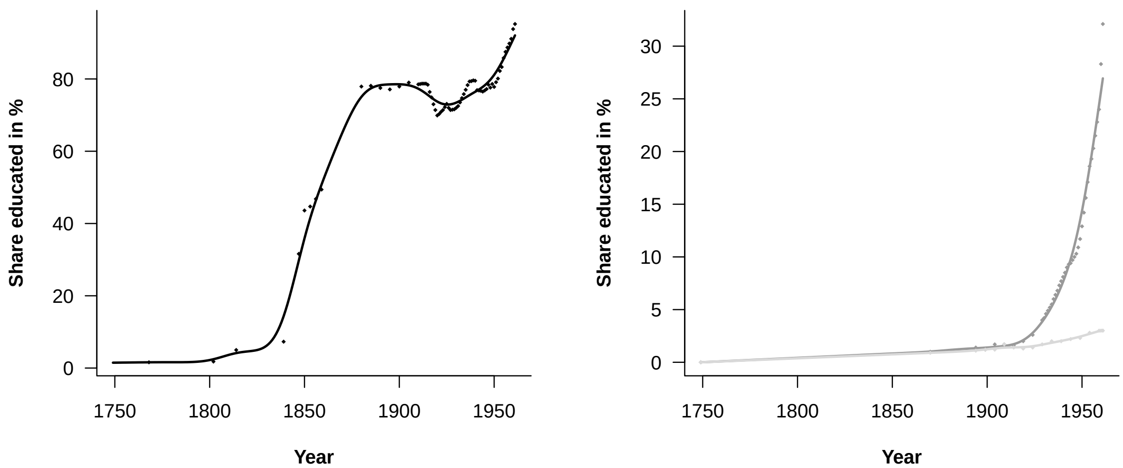
(b) Relative prices

Figure 5: Evolution of observations (dots) and smoothed observations (solid lines) of Sweden's GDP per worker and its relative prices for industrial goods.

Educational facts As was the case everywhere else in Western Europe, industrialization and economic modernization were accompanied by an educational revolution. This development has been highlighted by Galor (2005), among others. We extract enrollment rates for *primary and lower secondary*, *integrated secondary* and *tertiary* education from de la Croix *et al.* (2008). Figure 6a displays the primary and lower secondary education enrollment rates, defined as the percentage of the population aged 5–14 who were attending school. Before the nineteenth century, enrollment was negligible and remained low until 1842, when compulsory elementary education was introduced (Flora, 1983). At that point, primary (and lower secondary) education enrollment rates quickly rose to approximately 80% and remained at this level until the end of World War II.[19] A further increase during the postwar period is observable in Figure 6. By 1960, almost all children were enrolled in primary or lower secondary education.

[19]It has been widely documented that although school became compulsory in 1842 in Sweden, enrollment rates stayed well below 100% during the nineteenth century ; see Flora (1983); Benavot and Riddle (1988) or Soysal and Strang (1989), among many others.

Figure 6b shows the enrollment rates in integrated secondary and tertiary education.[20] The take-off in integrated secondary and tertiary education was delayed by almost a century compared to that of primary and secondary education. Afterwards, the integrated secondary enrollment rates accelerated and already exceeded one third around 1960. By contrast, the gains in tertiary education remained low and increased only slowly. Until 1960, less than 5% of those aged 20–24 were enrolled in the highest level of education.



(a) Primary and lower secondary education

(b) Integrated secondary education (dark gray) and tertiary education (light gray).

Figure 6: Observations (dots) and smoothed observations (solid lines) of Swedish educational enrollment.

In summary, our data distinguish three main periods: before 1850, between 1850 and 1900, and 1900 onwards. Until 1850, Sweden was experiencing a stable regime of economic stagnation in which agriculture was the dominant activity while GDP per capita was constant. Urbanization rates were stable, as the rural exodus compensated for the higher level of mortality and lower fertility in cities. Despite this overall stagnation, fertility rates first fell in urban areas, and mortality started to decrease everywhere slightly before 1800. Industrialization and significant growth rates appeared around 1850, but industrialization occurred in two phases. In the first phase, between 1850 and 1900, it occurred mainly in the countryside, while after 1900, cities took the lead. It

[20]Enrollment rates in integrated secondary and tertiary education indicate the number of pupils in the age groups 15–19 and 20–24, respectively. For details, see de la Croix *et al.* (2008).

was at that time that fertility rates exhibited their most dramatic decrease in history in both areas. Overall, the years 1840–1850 correspond to extraordinary transformations, as GDP per worker exploded hand-in-hand with a sharp increase in the urbanization and education rates.

3 The model

We assume a closed economy in discrete time ($t = 0, 1, 2, ..., \infty$). Its population is composed of overlapping generations. Individuals of each generation live for two periods, childhood and adulthood.[21] They can live either in the countryside R or in cities U, and they choose where to live at the beginning of their adult life once for all. This decision is made before any other decision. Each person has the freedom to settle wherever he or she wants, but migration is costly. The number of adults living in area $j = \{R, U\}$ at date t is denoted N_t^j; it is endogenous and is determined at the equilibrium of the economy.

3.1 Production sectors

We assume that agricultural production can take place in the countryside only, while industry can be located both in cities and in the countryside but not with the same technological regime. Rural industries produce goods that are fully homogeneous with the industrial goods produced in cities; the price of each unit of industrial good at date t is denoted p_t. Because we normalize the price of agricultural goods to 1, p_t also corresponds to the relative price between industrial and agricultural goods.

3.1.1 Production structures in the countryside

Adults living in rural areas can work either in the agricultural sector or in rural industry. We assume that each worker is endowed with a units of efficient rural labor that she can use in agriculture or in rural industry. a reflects the worker's intrinsic ability to work in rural production sectors and is transmitted from parents to children without decay.[22]

[21] In setting up our model in this way, we disregard potentially important phenomena such as savings or intergenerational co-residence. On this latter point, see the recent contribution of Pensieroso and Sommacal (2019).

[22] This simplifying assumption implies that all members of the same dynasty share the same rural abilities.

In the spirit of Galor and Weil (2000) but also Lagerlöf (2006) and Strulik and Weisdorf (2014), we assume that production in these sectors is ensured by self-employed workers earning their average production.

Agricultural production sector $\mathcal{A}$. Mimicking Galor and Weil (2000), agricultural production $Y_t^{\mathcal{A}}$ is ensured through a constant returns to scale technology using efficient labor $L_t^{\mathcal{A}}$ and land X:

$$Y_t^{\mathcal{A}} = (L_t^{\mathcal{A}})^{1-\theta}(A_t^{\mathcal{A}}X)^{\theta}. \tag{1}$$

θ is the partial output elasticity of land, and $A_t^{\mathcal{A}}$ reflects the state of the technology, which evolves endogenously. The quantity of efficient labor is determined by the adults living in the countryside and working in the agricultural sector. As shown in Section 3.2, productive efforts at the individual level strongly depend on fertility decisions. We assume that there are no property rights to the land such that the return to land is null and workers receive their average production. We define $w_t^{\mathcal{A}}$ as the average production per efficient unit of labor in the agricultural sector at time t such that:

$$w_t^{\mathcal{A}} \equiv \frac{Y_t^{\mathcal{A}}}{L_t^{\mathcal{A}}} = \left(\frac{A_t^{\mathcal{A}}X}{L_t^{\mathcal{A}}}\right)^{\theta}. \tag{2}$$

Rural industrial production sector $\mathcal{I}$. We allow industrial production to be located in the countryside.[23] Rural industries also use efficient labor in quantity $L_t^{\mathcal{I}}$ and are endowed with a technology $A_t^{\mathcal{I}}$:

$$Y_t^{\mathcal{I}} = A_t^{\mathcal{I}}\left(L_t^{\mathcal{I}}\right)^{1-\psi}. \tag{3}$$

The absence of stocks of capital and land implies that workers receive their average production per efficient unit of work $w_t^{\mathcal{I}}$, as in agriculture:

$$w_t^{\mathcal{I}} \equiv p_t\frac{Y_t^{\mathcal{I}}}{L_t^{\mathcal{I}}} = p_t A_t^{\mathcal{I}}\left(L_t^{\mathcal{I}}\right)^{-\psi}. \tag{4}$$

We assume perfect mobility of labor between agriculture and rural industries. It implies that the following non-arbitrage condition ensuring the equality of average productivities between the two sectors has to be verified at the equilibrium of the economy at

[23] This assumption is in line with the history of Sweden but even more generally of Western European countries. See, among others, the key contributions of Lampe and Sharp (2015), Bairoch (1997), Berger *et al.* (2012), Schön (2010).

any date t:

$$L_t^{\mathcal{A}} = \left(L_t^{\mathcal{I}}\right)^{\frac{\psi}{\theta}} \left[\frac{A_t^{\mathcal{A}} X}{(p_t A_t^{\mathcal{I}})^{\frac{1}{\theta}}}\right] \equiv \boldsymbol{v}\left(L_t^{\mathcal{I}}\right) \tag{5}$$

This condition implies that, for a given p_t, any productivity shock in favor of the agricultural sector (an exogenous increase in $A_t^{\mathcal{A}}$) would provoke a reallocation of the labor force from rural industry to agriculture and vice versa. A negative shock to the price of industrial goods would also reallocate labor from rural industry to agriculture.

3.1.2 Production structures in cities

Reasonably enough, we assume that cities can only produce industrial goods. The applied technology in urban industry differs from that in rural industry.[24] Indeed, urban industries produce industrial goods using human capital H_t:

$$Y_t^{\mathcal{M}} = A_t^{\mathcal{M}} H_t^{\gamma}, \tag{6}$$

with $A_t^{\mathcal{M}}$ representing total factor productivity in the urban industrial sector, γ representing the partial output elasticity of efficient labor, and $Y_t^{\mathcal{M}}$ representing urban manufactured output. The available quantity of efficient industrial labor is equal to the total amount of human capital within the urban industrial workforce. As human capital is fully owned by workers, we again find that each unit of efficient labor is compensated with its average productivity, implying that:

$$w_t^{\mathcal{M}} = \frac{p_t Y_t^{\mathcal{M}}}{H_t} = p_t A_t^{\mathcal{M}} H_t^{\gamma-1} \tag{7}$$

Therefore, what distinguishes the technology used in rural and urban industries is both the type of skills each industry uses and its level of technology. On this last point, we later assume that the accumulation of human capital is a source of positive externalities in urban industry, such that, along the lines of Romer (1986), $A_t^{\mathcal{M}}$ is an increasing function of H_t.

[24]Such a hypothesis is supported by Squicciarini and Voigtländer (2015) and Lees and Lees (2007), among many others.

3.1.3 Technological progress

The triplet $\{A_t^{\mathcal{A}}, A_t^{\mathcal{I}}, A_t^{\mathcal{M}}\}$ characterizes the state of technology at time t as well as the relative productivities across sectors. 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 adult population exceeds a critical value $\overline{N}$:

$$\frac{A_{t+1} - A_t}{A_t} = \begin{cases} 0 & \text{if} \quad N_t < \overline{N} \\ g > 0 & \text{if} \quad N_t \geq \overline{N}. \end{cases} \tag{8}$$

This assumption is close to that of Galor and Weil (1999), which was formalized in Galor and Weil (2000). It does not necessarily mean that the country-level population density had to be large enough to allow a country to innovate. It rather implies that the density had to be large enough to capture the technological progress coming from the rest of the world. This is consistent with the view that the technological revolution of the eighteenth and nineteenth centuries was a northwestern European phenomenon rather than a more localized phenomenon; see, for instance, Bairoch (1997). The respective sector-specific technological situations are as follows:

$$A_t^{\mathcal{A}} = B^{\mathcal{A}} A_t \quad , \quad A_t^{\mathcal{I}} = B^{\mathcal{I}} A_t \quad \text{and} \quad A_t^{\mathcal{M}} = B^{\mathcal{M}} A_t H_t^{1-\gamma}. \tag{9}$$

For all x, $B^x \in \mathbb{R}^+$ is a scale parameter allowing for initial differences in technology between sectors. Equations (6) and (9) imply that the accumulation of human capital is a source of technological progress for urban industries. Urban industries then benefit from an AK-type technology, as written in the ex-post production function:

$$Y_t^{\mathcal{M}} = B^{\mathcal{M}} A_t H_t. \tag{10}$$

Average productivity becomes independent of the quantity of human capital used in the sector such that $w_t^{\mathcal{M}} = B^{\mathcal{M}} A_t$.

In the framework we develop, technological progress is related to population density and human capital in a complex way. First, from Equation (8) and what is standard in UGT, family decisions regarding fertility and education in the Malthusian world fuel society on its way to the technological revolution. Net fertility at the family level has to be high enough to cause the whole population to reach a certain density in order to spread knowledge. Second, technological progress in urban industry is magnified by the accumulation of human capital in cities and so by the trade-off between the quality and quantity of children. A large number of studies show that cities and their industries required a specific and agglomerated accumulation of human capital to develop and that they produced industrial knowledge in return.[25] This mechanism gives a central role to migration decisions and urbanization.

3.2 Individuals

3.2.1 Population structure

At date $t = 0$, the initial population is composed of $N_0 \in \mathbb{N}$ individuals who, by having children, generate N_0 dynasties. These settlers are indexed by $i = \{1, 2, 3, ..., N_0\}$ and give birth to $n_0^i \in \mathbb{R}$ children.[26] Among these n_0^i children, only $q_0^i n_0^i$ survive to adulthood, as q_0^i denotes the child survival rate experienced by settler i at period 0. This survival rate depends on the area in which the settler lives, as explained in the next subsection. At date $t = 1$, these $q_0^i n_0^i$ survivors themselves decide on the number of offspring to whom they give birth.

At date t, the dynasty s designates all the individuals descending from settler $i = s$ throughout history. Therefore, the current size of any dynasty i at date $t+1$ corresponds to the number of living adult offspring of settler i and is denoted ℓ_{t+1}^i such that:

$$\ell_{t+1}^i = \prod_{b=0}^{t} q_b^i n_b^i \tag{11}$$

The total size of the adult population at date $t+1$ is denoted N_{t+1} and corresponds to:

$$N_{t+1} = \sum_{i=1}^{N_0} \ell_{t+1}^i \tag{12}$$

Each dynasty i differs from the others in its initial endowments in terms of rural skills (a^i) and in terms of human capital (h_0^i). Rural skills are given once for all to each

[25] See, for instance, the enlightening contributions of Simon and Nardinelli (1996, 2002) and Brunt and García-Peñalosa (2021) for the cases of England and the US. The magnifying effects of agglomeration (through peer effects and imitation) on the accumulation of human capital in cities compared to that in rural areas is theorized, for instance, by Glaeser (1999). The specific use of human capital in industry is also a central assumption in papers such as Tamura (2002).

[26] In line with the literature from Becker (1960) to Galor (2011), fertility decisions are represented with continuous numbers.

dynasty and transmitted without decay from parents to children, and human capital is accumulated at the level of individuals as described in the next section.

3.2.2 Maximization problem

In this subsection, we describe the budget and technological constraints faced by the representative agent of dynasty i at date t, as well as her utility function. Preferences between and within dynasties are identical. To simplify the notation and without loss of generality, we suppress the superscript i from our notation in what follows.

The utility of the representative agent of a given dynasty is denoted W_t. We assume the existence of a subsistence level of income, denoted $\bar{c} > 0$, below which people consume only agricultural goods c_t. This is a situation of extreme poverty that does not allow individuals to make choices and causes them to suffer infinitely negative utility. One can think of this hypothesis as echoing the possibility of famines, which have plagued the history of Western Europe; see, for instance, Alfani and Gráda (2018). When adults' income levels (ω_t) are sufficiently high, they can consume industrial goods d_t, give birth to n_t children, educate them to level e_t and migrate, with $\{c_t, d_t, n_t, e_t\} \in \mathbb{R}^{4+}$.

Accordingly, we assume that when $\omega_t \leq \bar{c}$, $W_t \to -\infty$ by definition and $c_t = \omega_t$, $n_t = e_t = d_t = 0$. In contrast, when $\omega_t > \bar{c}$, W_t is defined as follows:

$$W_t = \alpha \ln(c_t - \bar{c}) + (1-\alpha)\ln(d_t + \varepsilon) + \rho\left[\ln\left(n_t q_t^j\right) + \ln h_{t+1}\right]. \tag{13}$$

As parents have "warm glow preferences", we assume that they do not value the future well-being of their children but that they do care about their future human capital h_{t+1}. This human capital translates into higher potential income and therefore higher utility if the child decides to live in a city in the future.[27] As ε is a positive scalar, Equations (13) and (17) imply that education and industrial goods are superior goods.

Migration between urban and rural areas arises at the beginning of adult life and entails a fixed cost κp_t corresponding to κ units of industrial goods. Let's denote $\mathbb{1}_\kappa$ an indicator function that takes value one if the person moves from one area to another and zero otherwise. For the sake of simplicity, we assume that children are born after

[27] Importantly enough, we do not assume differences in preferences between urban and rural persons because migration decisions are endogenous; it would allow people to decide on their own preferences. Nevertheless, assuming, for instance, that urban persons value more the human capital of their children than rural persons may help to capture the dynamics of the tradeoff between quality and quantity over time in a better way. We present this alternative framework in Appendix D.2.

potential migration and are costly in terms of time: giving birth to a child requires a portion ξ of the parental time unit. With exogenous probability q_t^j, the child survives to age one. A surviving child needs additional time from her parent, which is denoted ζ^j and is area-specific. When having n_t children and living in area j, parents work for a portion, $1-\left(\xi+\zeta^j q_t^j\right)n_t$, of their time. Finally, education entails an industrial goods cost, such that each unit of education costs the parents βp_t. Parental budget constraints can be written as:

$$c_t + p_t d_t + \beta q_t^j p_t n_t e_t = \left(1-\left[\xi+\zeta^j q_t^j\right]n_t\right)\omega_t^j - \kappa p_t \mathbb{1}_\kappa. \tag{14}$$

Because raising children takes time and given q_t^j, the maximum number of births an individual can have is limited to $n_M^j \equiv \frac{1}{\xi+\zeta^j q_t^j}$. ω_t^j is the potential labor income of a parent living in area j at date t, such that:

$$\text{if} \quad j=R \quad then \quad \omega_t^R = a w_t^{\mathcal{A}} = a w_t^{\mathcal{I}}, \tag{15}$$

$$\text{if} \quad j=U \quad then \quad \omega_t^U = w_t^{\mathcal{M}} h_t, \tag{16}$$

The non-arbitrage condition (Equation 5), implies that in the countryside, no matter how individuals allocate their labor between industry and agriculture, their income remains the same. Furthermore, as explained in Subsection 3.1.3, because an AK-type technology prevails in urban industry, $w_t^{\mathcal{M}}$ is independent of human capital. Hence, potential labor income in cities is linear in the level of the human capital of the agent. Similar to de la Croix and Doepke (2003), human capital accumulates according to:

$$h_{t+1} = (\upsilon + e_t)^{\phi} h_t^{1-\phi}, \tag{17}$$

where $\upsilon > 0$ 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 an exponential distribution. There is an intrafamily transmission of human capital, whereby $1-\phi$ denotes the elasticity with which parental human capital enters into the children's human capital accumulation.

Individual behavior

Because young adults have to decide between migrating or staying where they were born before making other decisions, their maximization problem is solved backwards: (i) future parents calculate their optimal behavior in each location given the prevailing economic conditions; (ii) they decide where to live by comparing their indirect utilities in the two locations.

In Table 1, we list the six potential regimes adults can live in. Regimes are obtained by maximizing the parental utility function (13) with respect to their budget constraints (14), maximal fertility constraint, and minimal consumption constraints. Superscript j indicates the location of the person either in a city or in the countryside, and the value of $\mathbb{1}_\kappa$ depends on whether the person has migrated or not.

Table 1: Regimes in which adults can live

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

with $\tilde{\omega}_t^j=\omega_t^j-\kappa p_t\mathbb{1}_\kappa+\varepsilon p_t$, $\tilde{c}=\frac{1-\alpha+\rho}{1+\rho}\bar{c}$ and $\pi_t=\left(\xi+\zeta^j q_t^j\right)\omega_t^j-\beta p_t q_t^j\upsilon$

In regime R1, being poor by definition leads to childlessness and very low consumption levels.[28] In regime R2, having children is possible and optimal, but it is not optimal for individuals to move from one area to another, to consume industrial goods, or to educate their children. Moving from one area to another becomes potentially optimal in regime

[28]For a discussion of childlessness, see Baudin *et al.* (2015).

R3. R1 to R3 are Malthusian regimes in the sense that any increase in labor income translates into higher fertility. For a given set of parameters, Regimes R4 and R5 cannot coexist, while both regimes describe situations in which parental income is too low to optimally afford all types of consumption. When the cost of education β is low relative to the Stone-Geary element ε, then R5 exists and describes a situation in which parents educate their children but in which it is not optimal to consume industrial goods. In the opposite situation in which β is relatively high, parents consume industrial goods, but they do not educate their children: regime R4. Finally, regime R6 is a state of higher wealth in which there are no binding constraints except for the budget constraint.

Given the system of relative price p_t, potential incomes $\{\omega_t^U, \omega_t^R\}$ and mortality rates $\{q_t^U, q_t^R\}$, we can determine the regime in which each person lives conditional on the area she decides to live in. In Appendix B.1, we characterize individual behavior for any configuration of our set of parameters.

The conditions prevailing on the market for industrial goods determine the timing of the take-off of modern economic growth and of urbanization, as they drive the rise of the national demand for industrial products and then for human capital. Obviously, these conditions are deeply rooted in 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 j has to compare the value of living in a city with the value of living in the countryside. To express agents' location decisions, we introduce indirect utility functions $v^*(p_t, \omega_t^j(x_t), q_t^j)$ where x_t denotes the area specific abilities of an individual. More formally, $v^*(p_t, \omega_t^j(x_t), q_t^j) = \max_{c_t, d_t, n_t, e_t} W_t$ (s.t.) Eq. (14), (15) and (16), $n_t < n_M^j$, $c_t > \bar{c}$ and usual non-negativity constraints. The explicit expression of v^* directly follows from Table 1 and Appendix B.1.

From this, we can assert that an individual will decide to establish in a city under the following condition:

$$v^*(p_t, \omega_t^U(h_t), q_t^U) \geq v^*(p_t, \omega_t^R(a), q_t^R) \tag{18}$$

Let's remember that the place of birth plays an important role in this comparison of indirect utilities as in case of migration from one place to another, agents have to pay κp_t. For extreme values of the endowment set $\{a, h_t\}$, the choice between living in the countryside and living in a city is very easy to determine. As long as migration costs can be covered, any person characterized by $\frac{a}{h_t} \to 0$ should live in an urban area, while any person characterized by $\frac{a}{h_t} \to +\infty$ should live in a rural area. For intermediate values of the ratio, the prevailing child mortality rates in both areas play a significant role in the migration decision.

We can also show that there is a threshold $\check{h}_t$ such that $\forall\ h_t \geq\ \check{h}_t$, a person will prefer to live in a city. Indeed, at the individual level, ω_t^R does not depend on h_t. Hence, when living in the countryside, an increase in h_t modifies the optimal allocation between the quality and quantity of children, but it does not enlarge the set of potential choices. By contrast, the set of choices of an urban person enlarges when h_t increases, which implies that the value of living in a city increases monotonically with h_t. The threshold $\check{h}_t$ then follows. However, if this is the case, what prevents the existence of an equilibrium without agricultural production? The answer to this question is both the decreasing average 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 must be greater in order to attract new workers if infant mortality is higher.

To describe how the economy evolves along a general dynamic equilibrium in the next section, we introduce the sets S_t^j with $j = \{U, R\}$ containing the adults living in cities and in the countryside at date t. The composition of these sets is endogenously determined by the model and evolves over time. Each individual can belong to one and only one of these sets, and the allocation of people is driven by inequality (18). The absence of heterogeneity within generations of the same dynasty implies that all living members of a given generation of a given dynasty live in the same area. The composition of the sets is key for determining the quantity of efficient labor available in each area. Taking into account the fertility decisions of individuals, we can write:

$$L_t^{\mathcal{A}} + L_t^{\mathcal{I}} = \sum_{i \in S_t^R} \ell_t^i a^i \left(1 - \left[\xi + \zeta^R q_t^R\right] n_t^i\right) \quad (19)$$

$$H_t = \sum_{i \in S_t^U} \ell_t^i h_t^i \left(1 - \left[\xi + \zeta^U q_t^U\right] n_t^i\right) \quad (20)$$

Each of the sets S_t^j also corresponds to specific distributions of rural and urban skills. The initial shape of these distributions is assumed to be exponential but evolves endogenously over time. In Appendix C.5, we show how these associated distributions evolve along our calibrated general equilibrium.

3.3 General dynamic equilibrium

At period t, given a pre-determined stock of the adult population N_t, technological levels $\{A_t^{\mathcal{A}}, A_t^{\mathcal{I}}, A_t^{\mathcal{M}}\}$, the sets S_{t-1}^U and S_{t-1}^R, the level of human capital h_t^i of each current generation of each dynasty, as well as the size of each dynasty ℓ_t^i, the temporary general equilibrium of our economy is described by (i) the sets S_t^U and S_t^R, which reallocate dynasty $i = \{1, 2, 3, ..., N_0\}$ between U and R according to Equation (18), (ii) the associated sets of individual decisions $\{c_t^{i*}, d_t^{i*}, n_t^{i*}, e_t^{i*}\}$ for each member of dynasty $i = \{1, 2, .., N_0\}$ that maximize (13) with respect to (14) as well as the maximal fertility and minimal consumption constraints, (iii) the aggregate quantities $\{Y_t^{\mathcal{A}}, Y_t^{\mathcal{I}}, Y_t^{\mathcal{M}}, H_t, L_t^{\mathcal{A}}, L_t^{\mathcal{I}}\}$ defined in Equations (1), (3), (10), (19) and (20), as well as (iv) the relative price p_t^*, which ensures the satisfaction of the non-arbitrage condition (5) as well as equilibrium in the agricultural goods market (Equation 21) and the industrial goods market (Equation 22):

$$Y_t^{\mathcal{A}} = \sum_{i \in S_t^U} \ell_t^i c_t^{i*} + \sum_{i \in S_t^R} \ell_t^i c_t^{i*} \tag{21}$$

$$Y_t^{\mathcal{I}} + Y_t^{\mathcal{M}} = \sum_{i \in S_t^U} \ell_t^i (d_t^{i*} + \beta q_t^U e_t^{i*} n_t^{i*} + \kappa \mathbb{1}_\kappa^i) + \sum_{i \in \mathcal{S}_t^R} \ell_t^i (d_t^{i*} + \beta q_t^R e_t^{i*} n_t^{i*} + \kappa \mathbb{1}_\kappa^i). \tag{22}$$

We assume that at each date t, our temporary equilibrium exists and is unique. Recall that each current generation of a dynasty i experiences a specific child survival rate q_t^i that equals either the urban survival rate q_t^U if the dynasty's adult generation settles in cities or the rural survival rate q_t^R if the dynasty settles in the countryside. The dynamics of our general equilibrium are then characterized by the evolution of the state variables $h_t^i, \ell_t^i, N_t, A_t$ such that $\forall\ i = \{1,2,3,...,N_0\}$:

$$h_{t+1}^{i*} = \left[h_t^{i*}\right]^{1-\phi}\left[e_t^{i*}+\nu\right]^{\phi} \tag{23}$$

$$\ell_{t+1}^{i*} = q_t^i n_t^{i*}\ell_t^{i*} \tag{24}$$

$$N_{t+1}^{*} = \sum_{i=1}^{N_0}\ell_t^{i*}q_t^i n_t^{i*} \tag{25}$$

$$A_{t+1}^{*} = \begin{cases} A_t^* \ \text{ if } \ N_t^* < \bar{N} \\ (1+g)A_t^* \ \text{ if } \ N_t^* \geq \bar{N} \end{cases} \tag{26}$$

3.4 From stagnation to growth

Regarding the dynamics of our economy, some intuition can be provided. In order to better understand the main mechanisms at play in the long run, we write the expressions of average production per capita in agricultural and urban industrial sectors, disregarding dynasties' indices for sake of simplicity. Denoting g_{x_t} the growth rate of variable x at time t and using log approximations, we obtain:[29]

$$y_t^{\mathcal{M}} = \frac{Y_t^{\mathcal{M}}}{L_t^{\mathcal{M}}} = A_t h_t^U \quad \Rightarrow \quad g_{y_t^{\mathcal{M}}} = g_{A_t} + g_{h_t^U}$$

$$y_t^{\mathcal{A}} = \frac{Y_t^{\mathcal{A}}}{L_t^{\mathcal{A}}} = \left(\frac{A_t^{\mathcal{A}}X}{L_t^{\mathcal{A}}}\right)^{\theta} \quad \Rightarrow \quad g_{y_t^{\mathcal{A}}} = \theta\left[g_{A_t} - g_{L_t^{\mathcal{A}}}\right]$$

Under plausible initial conditions, there exists a set of sufficient conditions such that a balanced growth path exists in the long run. Let's assume that initially, $N_t < \bar{N}$ such that no exogenous technological progress exists, while the entire population lives in Regimes 2 to 4 whatever their place of living. The economy is trapped in a temporary Malthusian stagnation where fertility and income are positively linked while GDP per capita is stable. If fertility rates are above the replacement level on average, population acts as an underlying factor pushing the economy to $N_t > \bar{N}$. In that situation, a first engine of growth appears: technological progress. As income per capita starts to increase in both areas, families switch progressively from Malthusian types regimes to interior regimes (R5 and 6). At the beginning of this transition, the economy is in a post-Malthusian regime where both population and average GDP per capita increase hand in hand (Galor and Weil (2000)). As regimes 5 and 6 become more and more

[29] The perfect mobility between rural sectors ensures that $g_{y_t^{\mathcal{I}}} = g_{y_t^{\mathcal{A}}}$.

prevalent, households substitute quality to quantity of children and a second engine of growth appears in urban areas: the accumulation of human capital. The economy enters progressively the modern growth regime where growth of GDP per capita is no more positively linked to growth in population size. At that stage, economic growth in cities is potentially explosive as, stepwise, all urban persons would enter Regime 6. Neglecting inter-dynastic inequalities as a matter of simplicity, we would get that in this situation:

$$\frac{h_{t+1}^U}{h_t^U} = \left[\frac{\phi\left([\xi + \zeta^U q_t^U] A_t^{\mathcal{M}} h_t^U - \beta p q_t^U \nu\right)}{(1-\phi)\beta p q_t^U h_t^U}\right]^{\phi}$$

Along a growth path where h_t growth indefinitely, we obtain that the growth rate of human capital per capita is:

$$g_{h_t} \to \left(\frac{\phi[\xi + \zeta^U q_t^U] A_t^{\mathcal{M}}}{(1-\phi)\beta p q_t^U}\right)^{\phi} - 1 \tag{27}$$

In the long run, the growth rate of human capital in cities is increasing with the technological level $A_t^{\mathcal{M}}$ which is itself growing during the transition. Nevertheless, under the acceleration of growth, the fertility transition is on its way and fertility recedes in both areas. Explosive growth in cities will come to an end at the condition that the average fertility in the economy becomes lower than one, making population size decreasing below $\bar{N}$ in finite time. Simple computations show that on a balanced growth path where human capital grows at a positive rate and where $\zeta^U > \zeta^R$, rural fertility is higher than urban fertility.[30] It implies a simple sufficient condition for the existence of a non-explosive growth path:

$$\lim_{\omega_t \to +\infty} n_t^R = \frac{\rho(1-\phi)}{(1+\rho)[\xi + \zeta^R q_t^R]} < 1$$

Under such a condition, N_t is decreasing along the BGP.[31] Such a property is rather aligned with recent empirical evidence, as we document in the last paragraph of Section 4.5. Once all households are in Regime 6 and provided that relative price p_t is constant along the BGP, the existence of the agricultural sector is ensured only if $g_{y_t^{\mathcal{A}}} = g_{y_t^{\mathcal{M}}}$,[32] what is possible if the active population in agriculture decrease at the

[30]The condition $\zeta^U > \zeta^R$ is satisfied in our calibration-simulation exercise, see Section 4.4.

[31]Population size could be stabilized at a level below $\bar{N}$ along the BGP relying on additional mechanisms like international immigration or public policies increasing net fertility to one.

[32]See Appendix B.3 for a proof.

following rate: $g_{L_t^A} = -\frac{g_{h_t^U}}{\theta}$. This de-growth of the rural labor force is nourished by low fertility rates (general decrease of the population) and a constant exodus from countryside to cities. In Appendix B.3, we show that in a simplified version of the model where inequalities vanish along the BGP and everyone lives in Regime 6, there exists a rate of rural exodus ensuring the stability of urbanization rates and equality of GDP per capita growth in both areas along the BGP. This long-run situation is the one our economy converges to when we calibrate and simulate, see Section 4.4.

Finally, the convergence to a BGP may not be reached in two situations. A first possibility is that no initial growth in the technological level appears because fertility rates in the Malthusian regimes are lower than or equal to one. In that situation, the economy is trapped permanently into Malthusian stagnation. A second possibility is that average fertility in the long run is above one such that output growth is explosive.

4 Calibration

To determine how well our theory fits the data, we use the simulated method of moments (hereafter, SMM) and proceed in three steps: first, we define the initial situation of our economy. Second, we set certain parameters a priori, and third, we determine the value of the remaining parameters by minimizing the squared distance between the empirical and simulated moments for fertility, urbanization and the allocation of rural labor. The remaining economic and educational moments are used as checks for overidentification.

4.1 Initial situation

Our initial population consists of $N_0 = 5,000$ adults. Each of these adults founds his or her own dynasty. In line with the observations from 1760, we initially allocate 9.6% of these dynasties to cities. For each of the 5,000 dynasties, we draw an initial level of human capital h_0^i from an exponential distribution, with λ as the average. To ensure that the distribution is fulfilled ex post, we determine the initial value of each dynasty 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 ability a^i, we assume that its distribution is a symmetrical permutation of human capital. In other words, the dynasty with the highest human capital has the lowest agricultural ability. The median dynasty in terms of human capital is also the median dynasty in terms of agricultural ability. This simplifying assumption ensures economic and demographic heterogeneity in rural and urban areas. The heterogeneity is important to avoid that all individuals within an area live in the same regime and switch regimes at the same time. It enables smoother and more realistic dynamics.[33]

Finally, we assume that the adults with the highest initial human capital are initially located in cities and that the adults with the highest agricultural ability are located in the countryside. These two assumptions are grounded in the idea that the moments we observe in 1760 are the fruit of rational decisions made in the past: those with a comparative advantage in agriculture or rural industry remained in the countryside. Dynasties with the highest human capital have a comparative advantage in the activities of modern industry and the highest incentives to locate in cities.

4.2 A priori fixed parameters

Infant survival probability is exogenous and approximated by the observations of IMRs in Figure 2a. We pick values every 20 years and apply the formula $q_t^j = 1 - \text{IMR}_t^j$ to obtain the survival probabilities presented in Table 2.[34]

Table 2: Infant survival probabilities

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

$1-\theta$ denotes the elasticity of the agricultural production function with respect to labor. In the literature, the value of $1-\theta$ varies 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 slightly less than 10 months for each pregnancy. Implicitly, this fixes the maximum number of pregnancies to 25. Finally, we normalize the amount of land X and initial TFP A_0 to one.

[33] We illustrate this argument in Appendix D.1, where we propose an alternative calibration and simulation exercise with $a^i = \bar{a}$ for all dynasties. Even if satisfactory, the fit of the predicted moments reduces significantly compared to our benchmark framework. In particular, the educational transition is less smooth. All rural parents start to educate their children at the exact same moment; the educational revolution is "rugged".

[34] We describe the general procedure in more detail in Appendix A.1.

4.3 Moments to be matched and minimum distance

A set of 17 parameters remains to be estimated. We chose this set of parameters to minimize the squared distance between 44 observed and predicted moments:

$$f(p) = [d - s(p)]\, W\, [d - s(p)]' . \tag{28}$$

Here, d denotes the vector of empirical moments, and $s(p)$ denotes the vector of simulated moments that depend 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 zeros elsewhere.

To obtain our empirical moments, we first interpolated and smoothed our data using the Hodrick-Prescott filter (Hodrick and Prescott, 1997) with a smoothing parameter equal to 1200. We then selected observations from the smoothed time series for each 20 years from 1760 to 1960.

As we only have information on the allocation of labor in rural areas between 1850 and 1960 (Fig. 4b), we extrapolated the time series by assuming that the average annual growth rate of the share of industrial labor in the countryside was 1% per year. Table 3 presents the 44 moments to be matched. Taking the IMRs as given, we aim to match the value of the total fertility rate in cities and in the countryside, the proportion of rural labor allocated to industry b_t, and the urbanization rate v_t.

Table 3: Moments to be matched by the model

	1760	1780	1800	1820	1840	1860	1880	1900	1920	1940	1960
n_t^R	4.499	4.184	4.009	4.487	4.303	4.062	4.046	3.792	3.126	2.278	2.326
n_t^U	3.881	3.619	3.253	3.310	3.156	3.229	3.393	3.011	2.064	1.673	2.116
v_t	0.096	0.089	0.105	0.097	0.097	0.112	0.155	0.216	0.289	0.381	0.523
b_t	0.024	0.030	0.036	0.044	0.054	0.073	0.101	0.180	0.231	0.277	0.382

To allow for a comparison of our model predictions and the empirical moments, we have to approximate the total fertility rate $\hat{n}_t^j$ in area j at date t from our model's predictions of fertility. For consistency, we apply the two-step procedure introduced in Section 2. We first aggregate individual fertility decisions and then normalize them based on the fertility of the Hutterites[35] (Coale, 1969):

$$\hat{I}_t^j = \frac{\sum_{i \in S_t^j} n_t^i \ell_t^i}{N_t^j \frac{5}{2} \sum_o n_o^H}$$

where n_t^i denotes the number of births per person in dynasty i of size ℓ_t^j. n_o^H represents the age-specific fertility rates of the Hutterites. To take into account the fact that we model a unisex population, we only consider half of the adult population. In a second step, we approximate the TFR in area j at date t by multiplying Coale's index $\hat{I}_t^j$ by the fertility of the Hutterites, denoted $\mathcal{N}^H$, such that $\hat{n}_t^j = \hat{I}_t^j * \mathcal{N}^H$.

To estimate our deep parameters, we follow the general strategy in Baudin *et al.* (2015). We first apply the genetic optimization algorithm PIKAIA (Charbonneau, 2002) to search for 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). In contrast to Baudin *et al.* (2015), we model fertility as continuous, and in each parametric situation explored by PIKAIA and UOBYQA, we need to compute the general equilibrium dynamics of our economy as defined in Section 3.3. To do so, we first use the simulated annealing algorithm SIMANN (Goffe *et al.*, 1994) to determine at least the region of the global optimum. If SIMANN does not find the equilibrium, we 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, the whole exercise is very demanding in terms of computational power. To address this challenge, we adjusted PIKAIA to allow for dynamic parallelization and applied a hybrid MPI-OpenMP programming strategy within Fortran 90.[36]

We do not try to replicate any further economic moments for three reasons. First, by limiting the number of empirical moments required for identification, we reduce our computational time 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, sectors and goods. This puts considerable pressure on our computational resources. Second, we would like to preserve some time series for overidentification checks. Third, for some moments such as education and GDP

[35]Hutterites are German Anabaptists reputed to have fertility rates close to maximal reproduction capacities.

[36]See Baudin and Stelter (2021) for more details.

per capita, the mapping between the empirical and theoretical moments is not direct enough to include the latter in the minimization algorithm (see Section 4.5).

4.4 Results

Our calibration strategy leads to the 17 estimated parameters presented in Table 4. If individuals want to leave their area of origin, they have to pay 0.946 units of industrial goods. By 1760, migration costs $\kappa * p_t$ represent 28.4% of the average potential income of workers in modern industries. 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, and especially during episodes of economic stress, "*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 were 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 each of these productive systems necessarily entailed costs that for the sake of simplicity, we assume to be independent of the type of migration.

In the specific case of Sweden, where the land topology is not always favorable and distances between (large) cities are not small, the low quality of the transportation infrastructure was a clear barrier to information acquisition. Berger and Enflo (2017) explain 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 the absolute cost of migration to decline because of the shrinking relative price between 1760 and 1960 by almost 90%.[37] In addition, the relative cost of migration decreased

[37] In Appendix C.4, we show that, in our simulations, migrations were quantitatively much less important before 1800 than after, and even almost absent for some years. This result has to be interpreted carefully: in a context where migration costs represent almost 30% of the average potential income and given that families reallocated themselves in 1760 already, migration was not a viable option during some periods for the vast majority of the population. In reality, sizable internal migration existed before the industrialization in Sweden, especially short distance migration as explained by Dribe (2003) as well as Andersson (2018). One could imagine an extension of our model where migration costs would be heterogeneous. It would allow for short distance migration even in pre-industrial times.

continuously due to improvements in the average potential income of industrial workers and represented only 0.01% of the latter in 1960.

Table 4: PARAMETERS FROM THE CALIBRATION EXERCISE

Parameter	Sym.	Value	Parameter	Sym.	Value
Preference for c_t	α	0.250	Cost of education	β	0.101
Preference for children	ρ	0.428	Costs of moving	$\bar{\kappa}$	0.946
Minimum consumption	$\bar{c}$	0.037	Initial TFP in agric.	$(B^{\mathcal{A}}X)^{\theta}$	6.797
Stone-Geary element for d_t	ε	9.925	Initial TFP in rural ind.	$B^{\mathcal{I}}$	10.93
Child rearing cost in R	ζ^R	0.259	Initial TFP in urban ind.	$B^{\mathcal{M}}$	0.677
Child rearing cost in U	ζ^U	0.345	Elasticity in rural ind.	$1-\psi$	0.707
Basic human capital	υ	2.615	Growth rate of A	g	0.692
Elasticity of HC	ϕ	0.045	Critical population size	$\bar{N}$	15,028
Average HC in 1760	λ	0.605			

The initial total factor productivity in rural industries of 10.93 is approximately 16 times higher than the total factor productivity in the industrial sector of cities (0.677). The elasticity of rural industrial production with respect to labor input $(1-\psi)$ is 0.707 and, hence, slightly above that in the agricultural sector (0.650). Initial TFP in the agricultural sector is 6.797. The critical population size for technological progress to appear is $\bar{N} = 15,028$, which corresponds to an adult population approximately 3 times higher than that in 1760. An average annual growth rate of the population of 1.9% would enable global technological progress to start as early as 1820, this latter equals 2.7% per year.

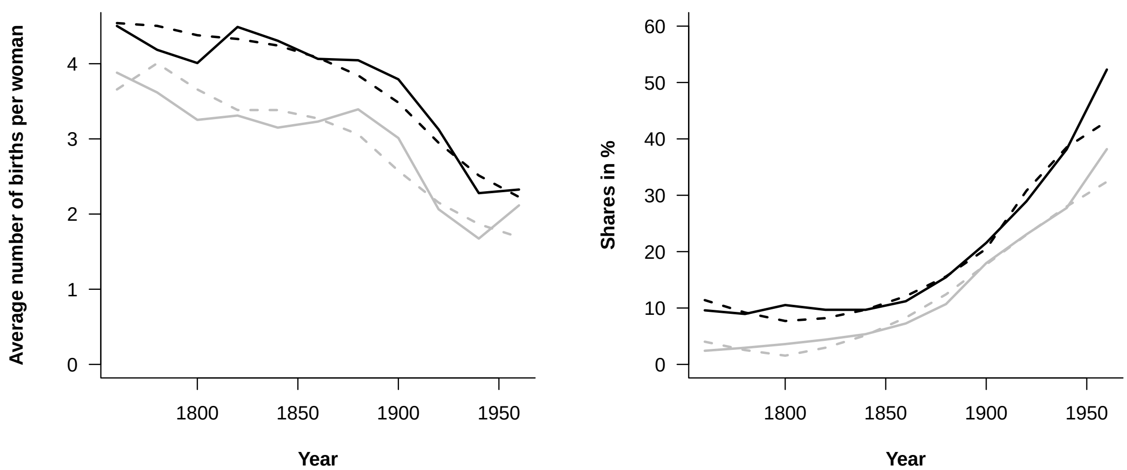
The minimum consumption of agricultural goods ($\bar{c}$) is 0.037 and represents 5.1% of predicted average potential income in 1760. For the poorest dynasty, this amounts to approximately $\frac{3}{8}$ of potential income and to more than 60% of labor income.[38] In our benchmark simulation, no dynasty ends-up in Regime 1 and disappears, see Figure 9b. For the set of parameters chosen by our algorithm, the existence of Regime 1 is then inconsequential.[39] The weight of agricultural goods within the utility function (α) is equal to 0.25, which is below the weight of educated children ($\rho = 0.43$). Each

Though more realistic, such an extension would severely complexify our theory without changing its main predictions and its overall quantitative fit.

[38]It corresponds to approximately 11.7% of the average available income. This value is plausible, as Gary (2018) estimates that on average, in 1760, fewer than 50 days of casual work were required to earn a subsistence basket.

[39]Notice that in alternative parameter sets, poverty may lead to the disappearance of some dynasties.

birth amounts to 4.0% in the time constraint. When interpreting the time required to raise a surviving child, we have to keep in mind that we assume a unisex population. Therefore, a surviving child requires an additional 13% of each parent's available time in the countryside and 17% in the cities. In the absence of infant mortality, this would lead to a maximal TFR of approximately six children in the countryside and five children in cities.



(a) Urban (gray) and rural (black) fertility

(b) Share of rural workers in industry (gray) and share of population in cities (black)

Figure 7: Observed (solid) and simulated (dashed) demographic dynamics and allocation of labor in Sweden, 1760–1960.

This parameter set enables our model to reproduce the characteristics of the Swedish fertility transition in a convincing way. We predict that from 1760 to 1960, the average number of births per woman dropped from 3.7 to 1.7 in cities and from 4.5 to 2.2 in rural areas (see Figure 7a). On the whole, we reproduce almost perfectly the magnitude of the fertility decline in both areas.[40] Our model is able to explain 92.7% of the variation in the TFR in the countryside and 84.5% of the variation in cities.[41]

[40]Our model predicts an increase in urban fertility between the first two periods, a result that is not in line with the data. This is because the dynasties are not perfectly allocated across areas in the initial state and then they reallocate themselves. We find that in 1760, the initial period in the simulation, some dynasties relocated from the countryside to a city (see Figure 31a and 31b in Appendix C.4). Approximately 5% of young adults in rural areas left their place of origin and moved to a city, where they represented 30% 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 κ caused them to live in a Malthusian regime.

[41]These values correspond to the R^2 of an ordinary least squares regression in which the observed

Regarding the urbanization rates and the share of rural labor dedicated to rural industry, the fit of our predicted values is almost perfect, except during the last period. As illustrated in Figure 7b, we tend to underestimate the most recent shares. Thus, we predict an increase of only 31.7 percentage points compared to the actual increase of 42.7 percentage points in the share of individuals in cities. Nevertheless, our model reproduces 96.3% of the variation in urbanization rates. The average distance between our empirical and simulated moments regarding urbanization rates is equal to 12.7% on average. In Appendix C.4, we present our estimation of the intensity of the rural exodus in Sweden. We estimate that it mainly prevented de-urbanization through 1840 and then became the driver of urbanization. Our predictions on the share of labor in the countryside that was active in the industrial sector explains 97.4% of the variation and on average only deviates by 3.1% from the observations.[42]

4.5 Overidentification checks

The high quality of our fit 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 five main series for which we have both data and predictions: GDP per person of working age, the average annual growth rates for GDP per worker in each area, the relative prices, the local allocation of industrial workers and enrollment rates into education.

The predicted GDP per worker in constant relative prices, $\hat{y}_t^{\text{real}}$, is computed according to:

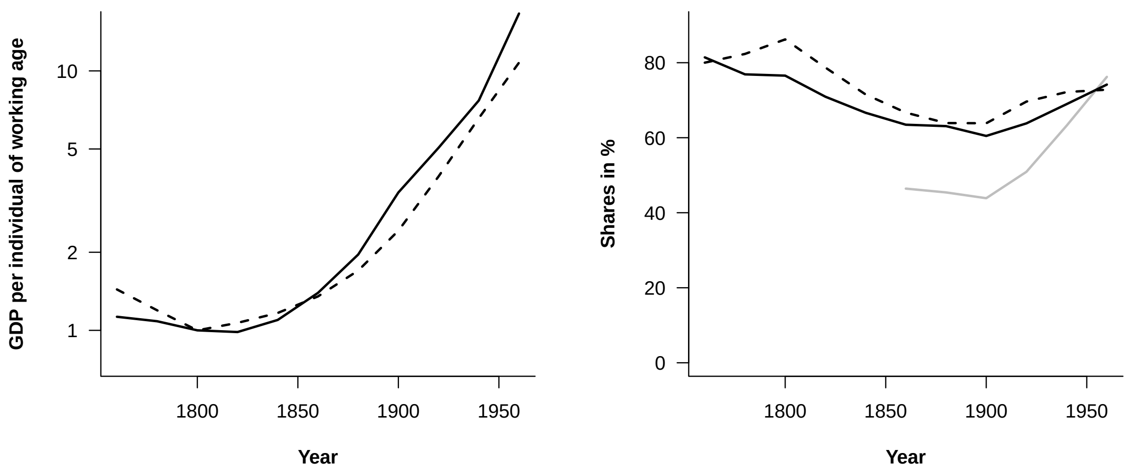
$$\hat{y}_t^{\text{real}} = \frac{Y_t^{\mathcal{A}} + p_{1800}\left(Y_t^{\mathcal{I}} + Y_t^{\mathcal{M}}\right)}{N_t}$$

To compare these predictions with the observed GDP per worker, we only account for the sectors *agriculture* and *manufacturing and industries* in the latter. The population in working age is approximated by multiplying the observed share of individuals of working age (15–64) from Centralbyrån (1999) by the observed population size.

TFR is the dependent variable and the predicted TFR 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 predicted moments in percentage points. In the case of fertility, the distance equals 10.5% in cities and 6.1% in the countryside. See Appendix C.1 for a further discussion.

[42]Even if the fit of the predicted moments speaks in favor of our estimation strategy, it does not guarantee that our set of estimated parameters is located in a well-behaved region of the parameter space nor that each parameter is well identified. We show in Appendix C.3 that our estimation suffers from neither of these two issues.

Figure 8a illustrates that we predicted the evolution of Swedish GDP per individual of working age normalized to one in 1800 as convincingly as we did the demographic data. Our model explains 98.3% of the observed variation, with an average deviation between the predicted and observed moments of 18.9%.[43] In addition, the available data from SHNA allow us to compare our predictions regarding the average annual growth rates for GDP per worker by sector with the observations from between 1860 and 1960. With averages of 2.3%, our predicted values in the industrial sector slightly exceed observed values of 2.1%. In the agricultural sector, we predict an average annual growth rate of GDP per worker of only 0.93%, while the observed value of 1.4% is higher.



(a) GDP per individual of working age

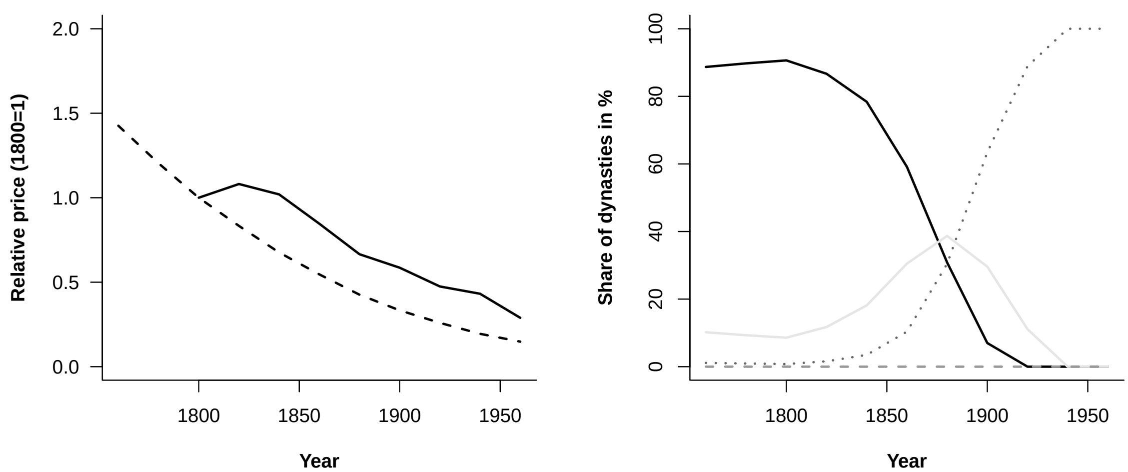
(b) Share of industrial workers in cities

Figure 8: Simulated (dashed) and observed (solid) total GDP per adult of working age and share of industrial workers located in urban areas: raw data (gray), adjusted data (solid black), and model predictions (dashed black).

The Industrial Revolution (in general and in Sweden in particular) occurred in two phases that we document through the U-shaped evolution of the fraction of industrial workers located in cities (Figure 4a): in the first phase of industrialization, rural industry was the driving force and attracted an increasing share of industrial workers; in the second phase, cities took over the lead (in the twentieth century). While our model predicts the allocation of industrial workers, the mapping between the model and the

[43] Using the GDP per capita or taking into account all sectors instead of being limited to industry and agriculture also leads to a convincing reproduction of the empirical moments; see Appendix C.1.

data is not direct. According to Centralbyrån (1969), only half of the urban population was active in industry during the Industrial Revolution; see Table 6 in Appendix A.4. The other half was engaged in other sectors, such as services and administration. In our model, for the sake of tractability, we assume that every urban adult is engaged in industry. For this reason, in Figure 8b, our predicted share of industrial workers located in cities (dashed black line) is much higher than the share computed in the data section (gray line). Nevertheless, from a qualitative point of view, we reproduce a U-shaped evolution consistent with the historical facts. For a more reasonable comparison between the model's predictions and the data, we can adjust the data by assuming that the whole urban population works in industry.[44] We then compare this modified share of industrial workers to the predictions of the model. The simulated moments almost perfectly fit these empirical moments with an average distance of only 7.1 percent. Overall, it is reassuring that our model reproduces this U-shaped pattern with its minimum at the beginning of the twentieth century.



(a) Observed (solid) and simulated (dashed) evolution of the relative industrial price

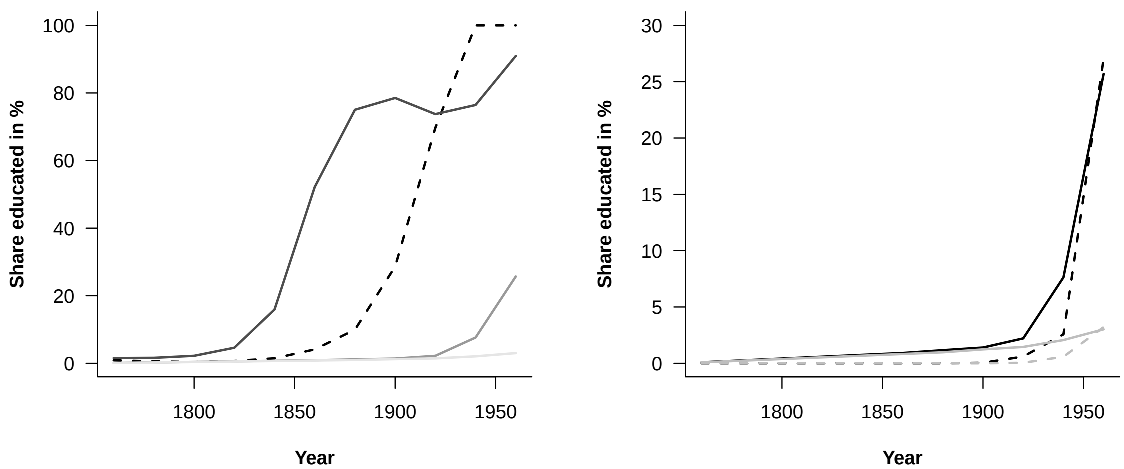
(b) Share of dynasties in regime 1–3 (solid black), regime 4 (solid light gray), regime 5 (dashed dashed medium gray) and regime 6 (dotted dark gray)

Figure 9: Evolution of the relative industrial prices and the share of dynasties by regimes of Tab. 1.

Our next counterfactual check concerns the dynamics of the prices of industrial goods

[44] More precisely, using Centralbyrån (1969) and Table 3, we compute the ratio between the number of urban workers and the number of urban workers plus rural industrial workers.

relative to the prices of agricultural goods. As Figure 9 shows, the data we have collected for this check span the 1800–1960 period. For illustrative purposes, we normalized the predicted and estimated relative prices in 1800 to one. While the model does not predict the stagnation in relative prices that occurred at the beginning of the nineteenth century, it captures most of the historical decline in relative prices that followed during the Industrial Revolution. Our simulated relative prices explain 97.6% of the variation in actual relative prices.



(a) Predicted share educated (black) and observed enrollment rates into primary (dark gray), integrated secondary (medium gray) and tertiary (light gray) education.

(b) Integrated secondary (black) and tertiary (gray) education with educational thresholds $\bar{e}^S$ and $\bar{e}^T$, respectively.

Figure 10: Observed enrollment rates into education (solid) and simulated shares of educated agents (dashed).

Regarding education, we first have to determine how the continuous educational investments in the model can predict enrollment rates. One simple idea is to define educational thresholds that mimic educational levels. Figure 10a depicts the predicted share of children with any educational investments compared to the primary (and lower secondary), integrated secondary and tertiary education enrollment rates. While we observe a clear delay in the onset of a noteworthy share of ever-educated children, improvements are as fast as in the observed primary education rates. As early as 1940, all children received at least some educational investments in our simulated world.

In Figure 10b, we introduce thresholds that are required to at least mimic integrated secondary and tertiary education. Minimizing the distance between the observed and predicted evolution, we fix the thresholds at $\bar{e}^S = 8.402$ and $\bar{e}^T = 15.115$ to approximate integrated secondary and tertiary education, respectively. These values correspond to 41.5% of the average labor income required to provide a secondary education to a child in 1760. Through 1960, these costs declined to 1%. We observe a similar decline—from 74.4% to 1.8%—in the case of tertiary education. From a qualitative perspective, we still observe the delays and periods of catching up. However, from a quantitative point of view, we are able to explain 96.5% of the variation in the case of secondary education and even 68.2% of the variation in tertiary education.[45]

By keeping infant mortality rates constant at their 1960 level, we can run the model forward into more recent years and the future. Although we strongly advise caution in interpreting these forecasts due to the model's limitations and the constant IMRs, we can at least report our predicted moments along the balanced growth path.[46] The predicted growth rates of GDP per worker converge to approximately 1.4% per year, while the average TFR converges to 1.64 and the urbanization rate is approximately two-thirds of the population. Interestingly, these predictions are close to those of international organizations for the coming years. For instance, Statistics-Sweden (2018) predicts a TFR of 1.88 in 2070; OECD (2014) predicts that real GDP per capita will grow at a rate of 1.99% per year between 2040 and 2060.[47] Finally, urbanization rates are expected to reach 93.2% of the population according to United-Nations (2018). This last rate seems especially high compared to our projections, but one should keep in mind that the estimation we provide is constrained by the city charter-based definition of cities used in Sweden in the past, constituting a lower bound on what is in reality based on a standard demographic definition of urban areas.

[45]For simplicity, we propose a constant threshold to "mimic" educational levels. In line with Galor and Weil (2000), we could also argue that the acceleration in technological progress required increasingly higher levels of education to maintain the relative educational level. Such a dynamic threshold would reduce the temporary underestimation in secondary and tertiary education enrollment rates.

[46]See details in Appendix C.2.

[47]Fertility rates below the replacement level are observed among many industrialized economies, like Japan and Germany having respectively 1.36 and 1.54 births per woman in 2019 according to the World Bank. In particular, several European countries experience what demographers have designated as lowest-low fertility levels since the 1990s (Kohler *et al.*, 2002; Billari and Kohler, 2004; Sobotka, 2004).

5 Historical experiments

Three factors are crucial in our theory: the initial productivity in rural industry, migration costs and mortality factors. In this section, we discuss the role of each using counterfactual experiments.

5.1 Productivity in rural industries

We evaluate the extent to which the initial productivity in rural industry has influenced the take-off of modern economic growth and its associated urbanization and fertility transitions. We simulate the dynamics of our economy for two main scenarios: a decrease in the initial aggregate productivity in rural industry ($A_0^{\mathcal{I}}$) and an increase therein. More precisely, we assume alternative values of $A_0^{\mathcal{I}}$ such that $A_0^{\mathcal{I}'} = xA_0^{\mathcal{I}}$ with $x = \{\frac{1}{10}, \frac{1}{5}, \frac{1}{2}, 1, 2, 5, 10\}$. The results of these experiments are presented in Figures 11 and 12.

5.1.1 Lower initial productivity

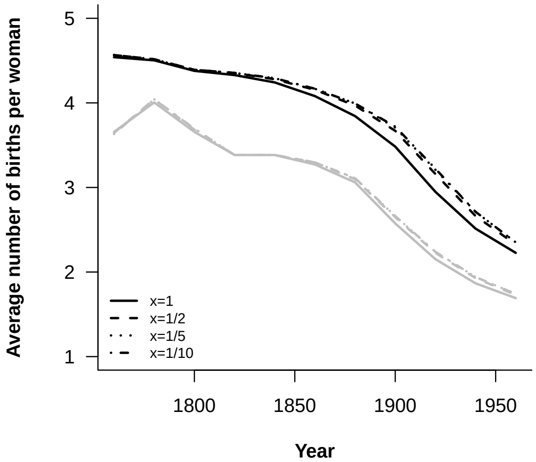
One can expect that reduced productivity in rural industry would push populations to leave rural areas while reallocating rural workers from industry to agriculture. These two movements appear clearly in Figures 11c and 11e after 1850. In the extreme scenario in which the initial productivity in rural industry is divided by 10, urbanization rates in 1960 are increased by 10.9 percentage points compared to those in the benchmark scenario. This difference is 8.1 percentage points when initial productivity is divided by only 2. Within rural areas, the reallocation of workers between agriculture and industry that we observe in the benchmark scenario, is already severely limited when initial productivity is divided by 2 and becomes almost nonexistent once productivity is divided by 5 (Figure 11c).

Fertility in the countryside increases (Figure 11a) through two complementary mechanisms. First, due to the perfect mobility between rural sectors, a lower industrial productivity reduces potential income both in industry and in agriculture, which increases fertility among people in the interior regimes. This effect is reversed for people in Malthusian-type regimes (those with a lower ratio a/h), which could then counterbalance the rise in fertility. Nevertheless, a second effect takes place that leads to a stronger rural exodus by pushing the relatively less productive agents out of rural areas.

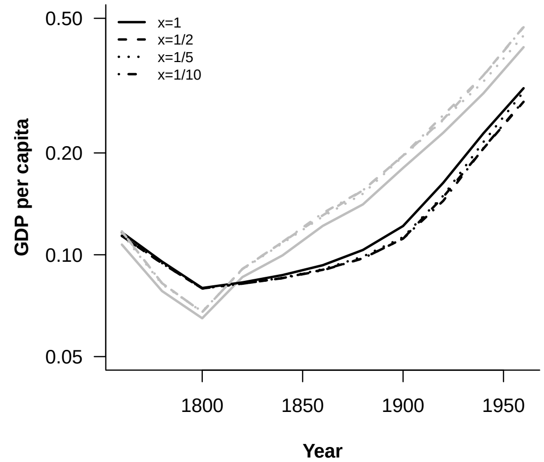
This strongly limits the reduction in fertility due to the reinforcement of the Malthusian checks on poorer populations. In this scenario compared to the benchmark, there is a surplus of rural migrants in cities. If the fertility of these migrants is depressed by the payment of migration costs, they will also be characterized by relatively lower levels of human capital on average. For that reason, after one period of settlement, they enjoy a fertility maximizing potential income and thus have on average more children than the rest of the urban population.[48] Accompanying this increased fertility, educational investments are reduced, thereby delaying the educational revolution, a movement amplified by the increase in the relative price of industrial goods provoked by the lower average productivity in industry; see panels d and f of Figure 11.

The initial productivity in rural industry has no effect on the long-run growth rate of the economy, whether in rural or urban places. Nevertheless, compared to that in the benchmark scenario, the accelerated rural exodus in this scenario permanently increases urban income. Industrialization is almost completely triggered by urban industries benefiting from general technological progress and the accumulation of human capital, and the latter is stronger than in the benchmark case (Figure 11b). Importantly, the negative effect of the shock to agricultural income is more than compensated for by the positive effect on industrial income. When initial productivity is divided by 2, the average income per capita in 1960 (in current relative prices) increases by 6.2% compared to that in the benchmark scenario, and this change increases to 8.8% when productivity is divided by 5.

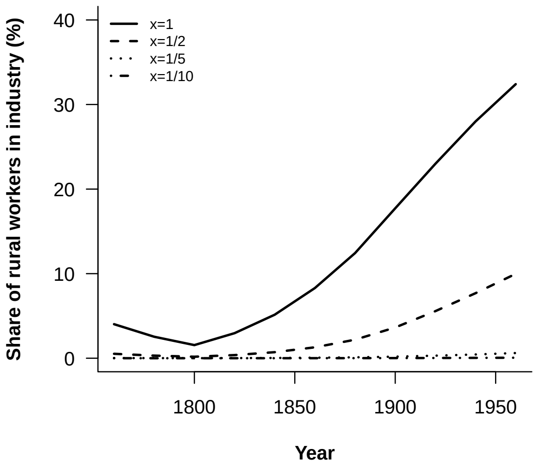
[48]In Figure 21 of Appendix B.2, these persons live in the neighborhood of $\bar{\omega}^j$.



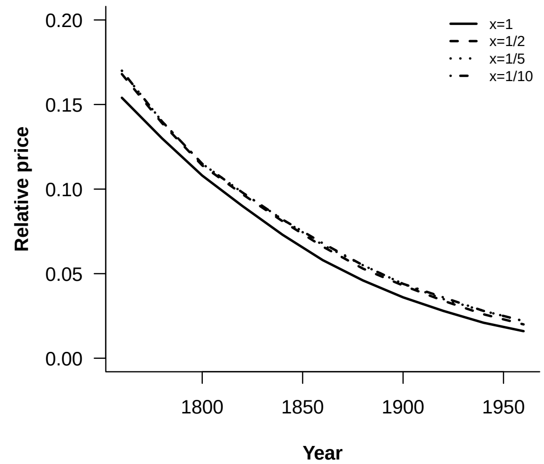
(a) Urban (gray) and rural (black) fertility



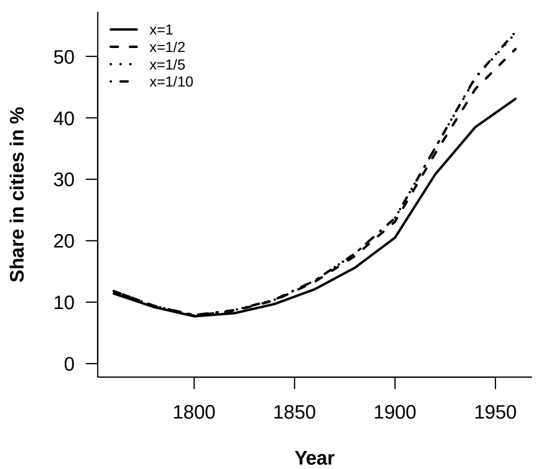
(b) GDP in current relative prices per capita in cities (gray) and in the countryside (black)



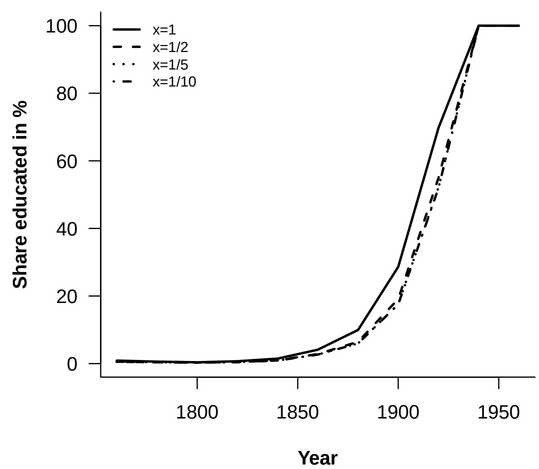
(c) Share of industrial workers in rural areas



(d) Relative price of industrial goods



(e) Urbanization



(f) Education

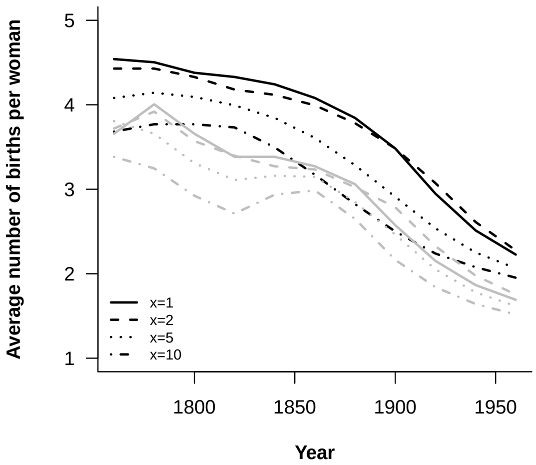
Figure 11: Swedish demographic and economic dynamics for reduced initial productivity in rural industry.

5.1.2 Higher initial productivity

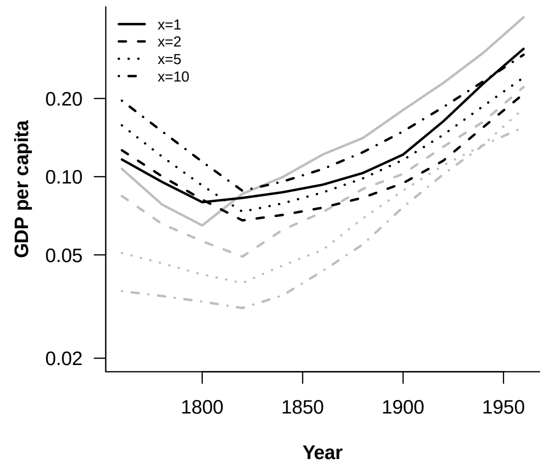
Compared to the benchmark, in this scenario, workers are reallocated from urban industry and agriculture to rural industry (Figure 12c). This reallocation slows down and clearly delays the urbanization process (Figure 12e). When the initial productivity is multiplied by 2, urbanization rates in 1960 are already reduced by more than 14.5 percentage points. More generally, the stronger the initial productivity in rural industry, the later the onset of the urban transition.[49] Within the countryside, the reallocation of workers from agriculture to industry occurs in two steps. At the time of the shock, compared to that in the benchmark scenario, the proportion of industrial workers increases substantially. However, then, in the absence of general technological progress (which appears only in 1820 in all scenarios), the reduction in the relative price of industrial goods tends to reallocate workers from rural industry to agriculture (see Equation 5). However, once technological progress sets in, this progress more than compensates for the decrease in relative prices, and the share of industrial workers in rural areas starts to increase again, reaching between 40% and 60% depending on the magnitude of the shock.

Figure 12b shows how a favorable initial condition in rural industry tends to reduce income in the long run in all areas. In the initial phase, in rural places, industrial and agricultural income are higher due to the higher $A_0^{\mathcal{I}}$, which translates to depressed fertility rates. In cities, income levels are clearly reduced through a general equilibrium effect: the vitality of rural industries increases the supply of industrial goods, clearly decreasing their relative price (Figure 12d). Lowered industrial prices translate into reduced potential income in cities as, in cities, the value of the produced units of good decreases. This negative income shock translates into depressed fertility rates (Figure 12a), as it pushes a large share of the urban population into an impoverished Malthusian state. As a result, population growth is slower, and technological progress appears one period later. If long-run growth rates are not altered, this later take-off substantially reduces income levels in the long run in both places. When initial productivity is multiplied by two, the average income per capita is divided by 1.49 in rural places and 1.85 in urban places by 1960. Note that these income losses in the long run are limited by the earlier educational revolution due to the initial increase in educational investment. Overall, in the transition period from 1820 to 1900, the higher

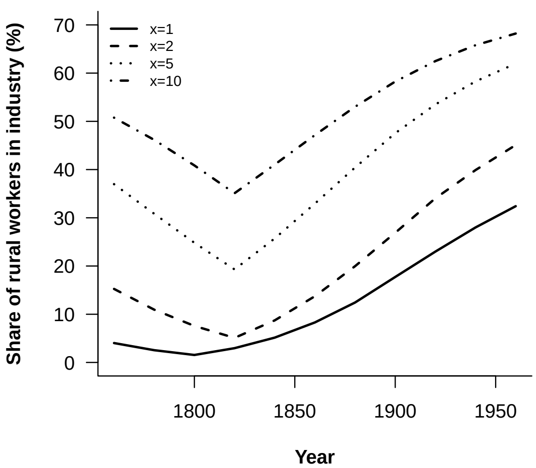
[49] Note that in the initial period, a higher $A_0^{\mathcal{I}}$ provokes an urban exodus; this effect is quite sizable when $A_0^{\mathcal{I}}$ is multiplied by 10.



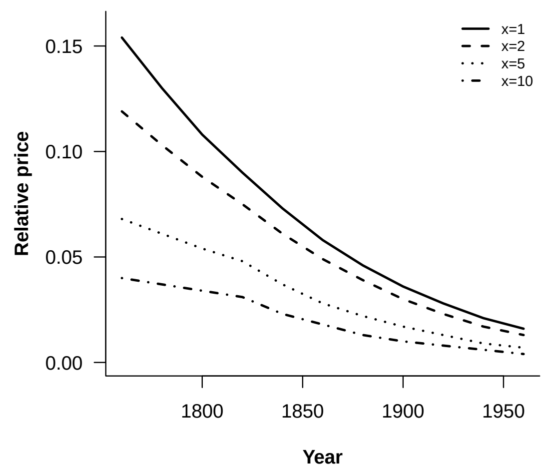
(a) Urban (gray) and rural (black) fertility



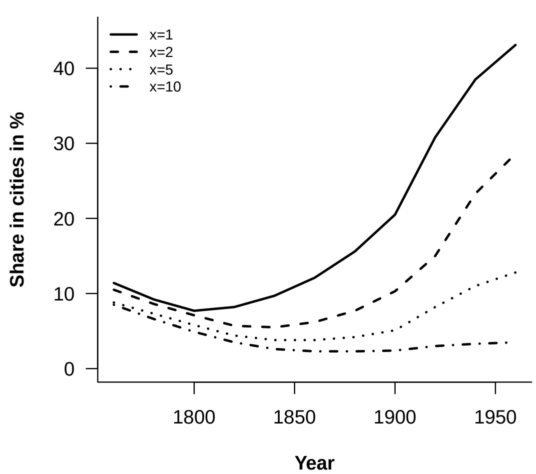
(b) GDP in current relative prices per capita in cities (gray) and in the countryside (black)



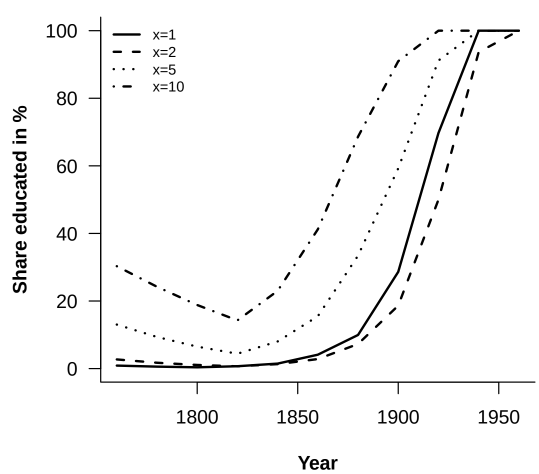
(c) Share of industrial workers in rural areas



(d) Relative price of industrial goods



(e) Urbanization



(f) Education

Figure 12: Swedish demographic and economic dynamics for increased initial productivity in rural industry.

the initial productivity in rural industry is, the lower the growth rates of income per capita. Because of this, despite being earlier, fertility transitions tend to last longer such that fertility levels in 1960 tend to be similar to those in the benchmark situation or to even surpass them.

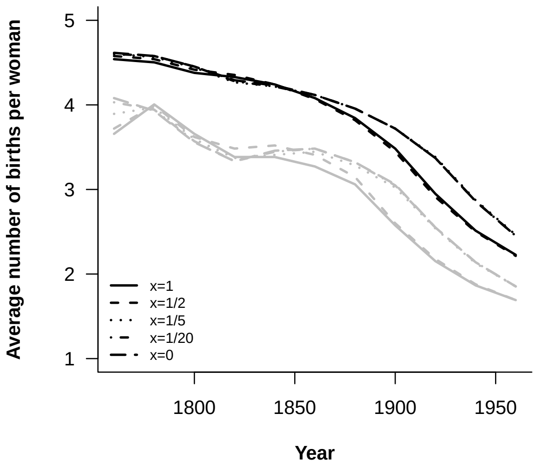
This first set of historical experiments suggests that the initial productivity of rural industries has been a key driver of urbanization and its associated transition towards growth. While beneficial to each sector in the first phase of industrialization, a better initial condition in rural industry tends to delay the second phase of the industrialization process and then to permanently impoverish the economy.

5.2 Migration costs

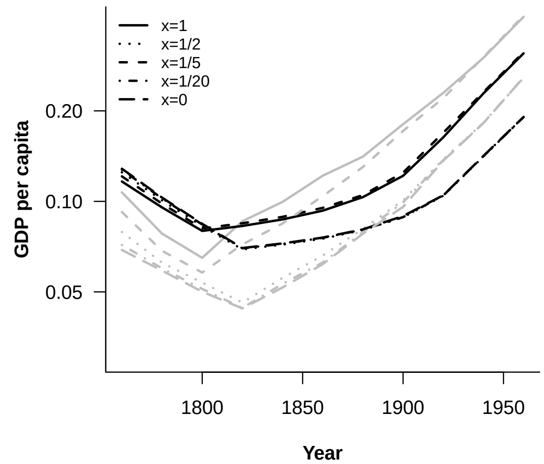
As we explained in Section 4.4, migration costs are deeply rooted in the transportation infrastructure, in topological conditions, and in information asymmetries. Recent papers by de la Croix *et al.* (2017) and Ogilvie (2014) have also documented how institutions such as guilds facilitated or hindered internal migration in Europe before industrialization. In this section, we estimate the impact of restrictions on internal migration via variation in migration costs on the economic and demographic dynamics of Sweden. More precisely, we simulate our dynamic general equilibrium model for alternative values of κ such that $\kappa' = x\kappa$ with $x = \{0, \frac{1}{20}, \frac{1}{5}, \frac{1}{2}, 1, 2, 5, 20, +\infty\}$. As for productivity shocks, we separate our analyses into a decrease and an increase in migration costs.

5.2.1 Lower migration costs

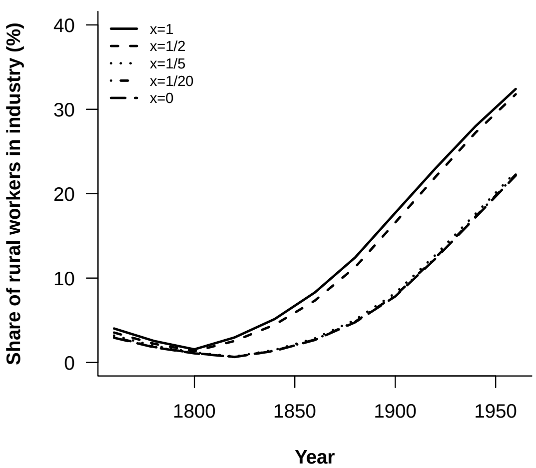
Compared to the benchmark scenario, in this scenario, decreasing migration costs provokes an initial reallocation of people from rural to urban places (see Figure 13e) and so from more to less fertile places (Figure 13a). The greater the decrease in migration costs, the larger this reallocation. As a result, average fertility rates are reduced to a point at which the population size surpasses the required threshold for TFP growth one period later. This delay in the economic take-off implies a permanent negative income shock, as average productivities are reduced in all sectors compared to those in the benchmark situation; this is reinforced by the induced substantial delays in the educational revolution (see Figure 13f).



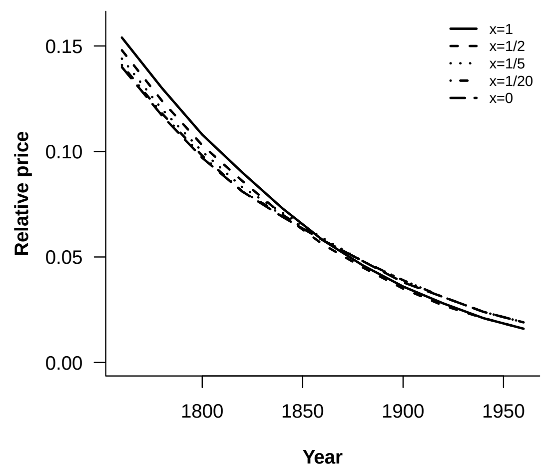
(a) Urban (gray) and rural (black) fertility



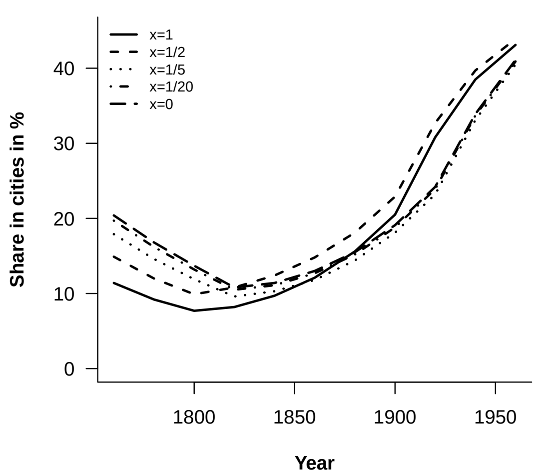
(b) GDP in current relative prices per capita in cities (gray) and countryside (black)



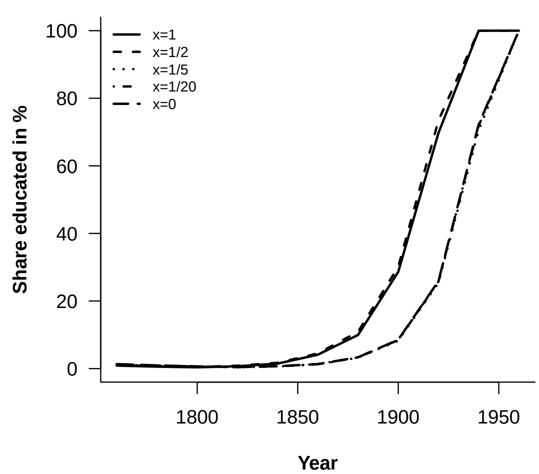
(c) Share of industrial workers in rural areas



(d) Relative price of industrial goods



(e) Urbanization



(f) Education

Figure 13: Swedish demographic and economic dynamics under reduced migration costs.

The reduction in κ limits the importance of rural industry in the first phase of the industrialization process. As seen in Figure 13c, the proportion of rural workers involved in industry remains stable throughout the nineteenth century, while urbanization rates are already higher than or equal to the benchmark. Interestingly, after the initial reallocation of workers, agricultural income remains higher than industrial income for a relatively longer period of time. This explains why rural migration to cities is almost absent from 1760 to 1860, leaving natural movements of the population to reduce the urbanization rate. Because of this, after 1850–1870, the urbanization rate increased again because of the rural exodus but did not fully catch up to its benchmark value by 1960.

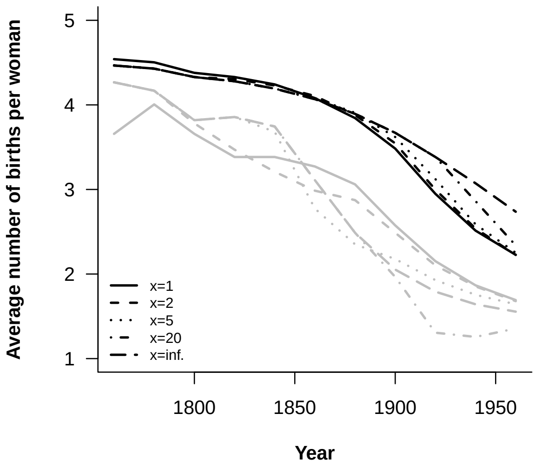
5.2.2 Increase in migration costs

In the extreme scenario in which $\kappa' \to +\infty$, no migration between the two areas occurs (Figure 14, long dashed lines). Driven by natural population movements, urbanization rates converge to arbitrarily small values, while rural industry becomes the epicenter of industrialization until 1960. This is verified in Figure 14c, where by 1960, more than 50% of the rural population works in rural industry. Globally speaking, compared to the benchmark, in this scenario, workers are reallocated from urban industries to rural sectors, while within the countryside, the share of industrial workers grows.

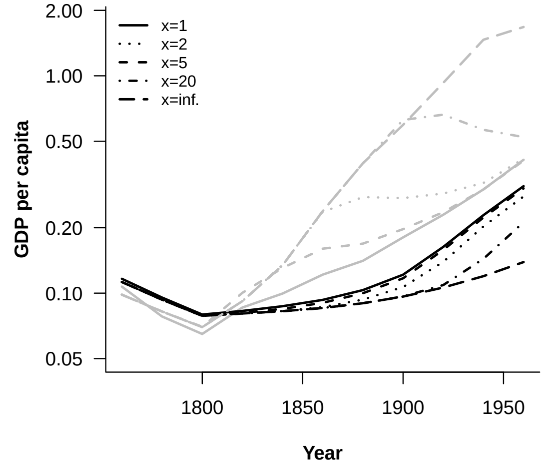
Due to the vitality of the rural industrial sector, until 1870, GDP per capita in rural areas in this scenario is quite comparable to that in the benchmark (Figure 14b). Nevertheless, after 1870, the absence of a rural exodus coupled with higher fertility rates increased population pressure in rural places, translating into low levels of productivity and therefore low GDP per capita. By 1960, GDP per capita would have been divided by 2.2 in the absence of internal migration. On the other hand, cities would not have welcomed new workers from rural areas, those workers who are characterized by lower levels of human capital than the residents. As a result, average urban income driven by technological progress and high average levels of investment in education would have exploded to become 4.1 times higher than in the benchmark. This explosion in rural-urban income inequalities would be fueled by the persistence of high fertility differentials between the two areas (Figure 14a). Indeed, higher differentials in potential income would have translated into higher fertility differentials and therefore higher workforce participation differentials. We estimate that in the absence of the rural exodus, the fertility differential between the countryside and cities would have reached 76% by 1960, while it was only 32.6% in the benchmark scenario.

In less extreme scenarios in which migration costs are multiplied by 2 up to 20, we observe comparable movements. The higher the increase in migration costs, the later the urbanization rate increases, but these rates do ultimately increase. When migration costs are multiplied by two, we do not observe a substantial change in the urbanization dynamics, but once these costs are multiplied by 5, urbanization rates surpass their 1760 level only after 1900: the onset of the urban transition is delayed by 50 years (Figure 14e). As a result, compared to the benchmark, in this scenario, even if the fertility transition is still delayed in the countryside and completed more rapidly in cities, these movements are less dramatic than those in the no-movement scenario. Indeed, compared to that no-movement scenario, pauperized migrants ultimately arrive in cities and thus increase fertility. In terms of the long run, in all the scenarios, the economy reaches a path of sustained economic growth, but the larger the increase in migration costs is, the higher the transitory economic loss, permanently depressing long-run levels of GDP per capita.

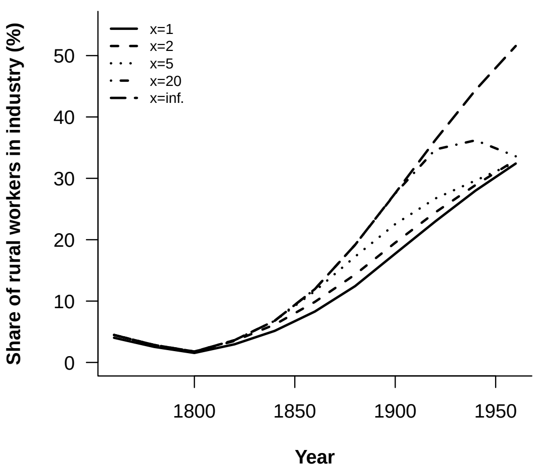
Overall, this analysis suggests that migration costs were an important determinant of the economic and demographic dynamics in Sweden. Indeed, the greater these costs are, the later the fertility, educational and economic transitions to modern regimes. Nevertheless, the existence of rural industry somehow mediates this effect, as rural industry makes urbanization unnecessary during the first phase of industrialization. That being said, by permanently altering fertility and the rural exodus, higher migration costs deepen the economic inequalities between cities and the countryside and permanently impoverish the country. In the next subsection, we dig deeper into this mechanism.



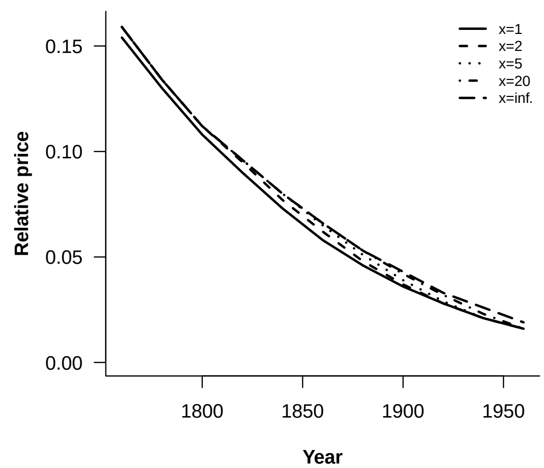
(a) Urban (gray) and rural (black) fertility



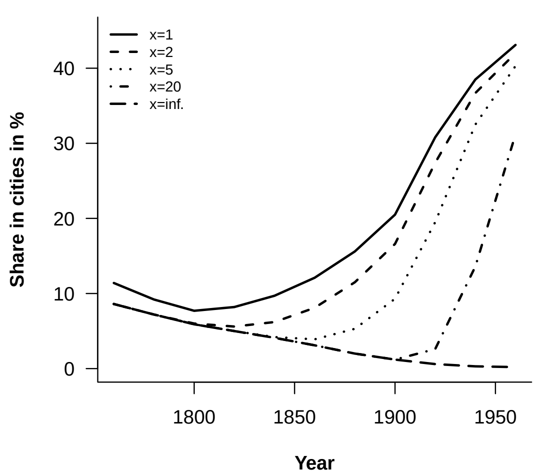
(b) GDP in current relative prices per capita in cities (gray) and countryside (black)



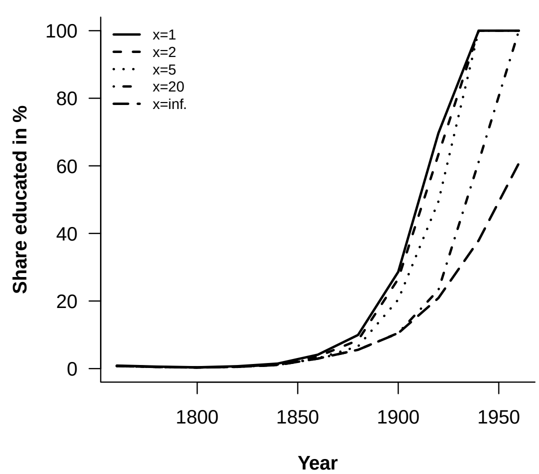
(c) Share of rural workers in rural areas



(d) Relative price of industrial goods



(e) Urbanization



(f) Education

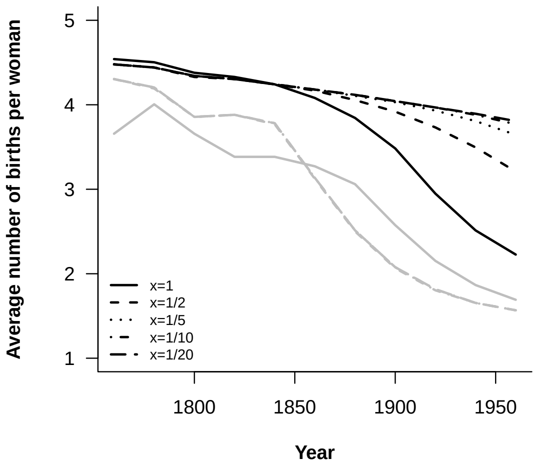
Figure 14: Swedish demographic and economic dynamics under increased migration costs.

5.2.3 Interplay between rural industries and migration costs

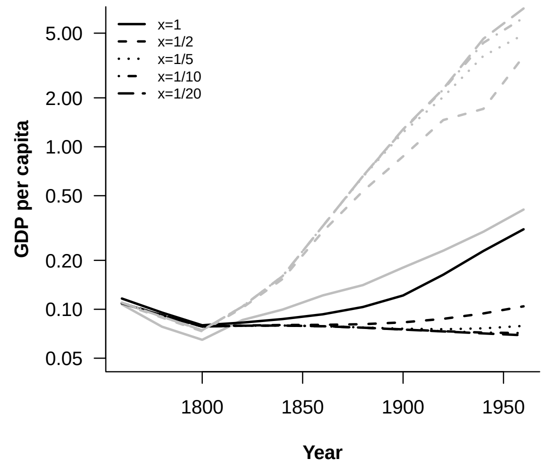
In this section, we establish that high initial levels of productivity in rural industry protected economies against the side effects of unfavorable migratory conditions. To do so, we simulate our economy under infinite migration costs and analyze whether reducing $A_0^{\mathcal{I}}$ delays the transition to growth and to modern demographic behavior. We test the following scenarios with $\kappa \to +\infty$: $A_0^{\mathcal{I}'} = xA_0^{\mathcal{I}}$ with $x = \{\frac{1}{20}, \frac{1}{10}, \frac{1}{5}, \frac{1}{2}, 1\}$.

In Figure 15a, we show that the lower $A_0^{\mathcal{I}'}$ is, the greater the delay in the onset of the fertility transition in rural areas, which is provoked by the absence of the rural exodus. When rural industry is no longer able to sustain the first phase of industrialization, the fertility transition in rural areas tends to vanish, while the fertility dynamics in cities are almost unchanged. In the scenario in which $A_0^{\mathcal{I}}$ is divided by 10, no fertility transition would have occurred in the countryside until 1960, and the rural TFR would have remained above 4 children per woman. Recall that in the benchmark scenario, the reallocation of agricultural workers to rural and urban industry could reduce the negative effects of population pressure in the agricultural sector. In the present scenario, these two mechanisms (industrialization in the countryside and the rural exodus) no longer exist. Coupled with the explosion in the size of the rural population, this leads to the pauperization of rural people.

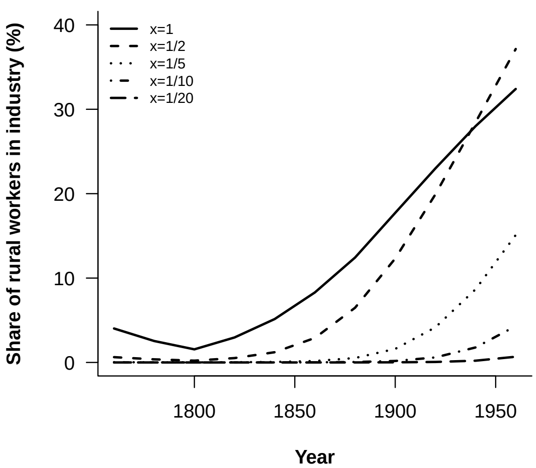
The fertility dynamics in cities would not be fundamentally changed, which holds in all our scenarios in which $\kappa \to +\infty$; furthermore, we see no newcomers from the countryside. The absence of a fertility transition in the countryside would have led to imbalances in the natural movements of the population. De-urbanization would have been more intense; a small elite would have populated the cities and enjoyed incomes that were 10.9 times higher than those in the benchmark scenario. In the previous section, increasing the migration costs to infinity would have led to an already substantial differential in terms of average GDP per capita between cities and the countryside, as the average urban GDP would have been 12.1 times higher than its rural counterpart. However, now, when initial productivity per capita is divided by 10, this differential amounts to 87.5. Rural-urban inequalities are largely driven by the combination of migratory and rural production conditions.



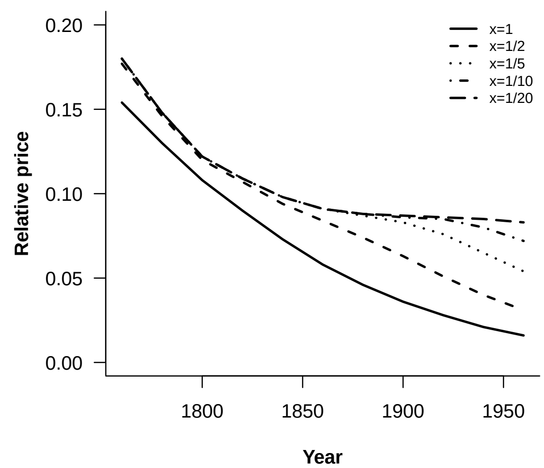
(a) Urban (gray) and rural (black) fertility



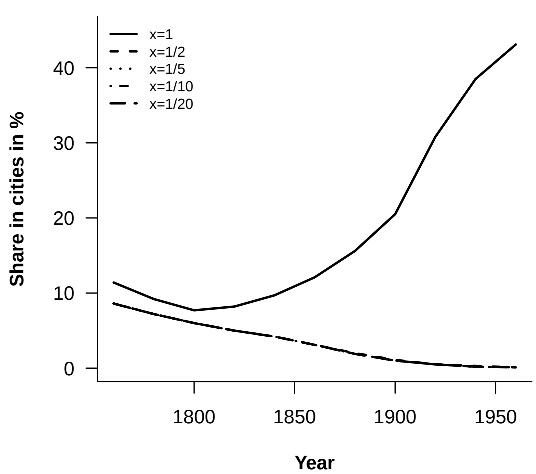
(b) GDP in current relative prices per capita in cities (gray) and countryside



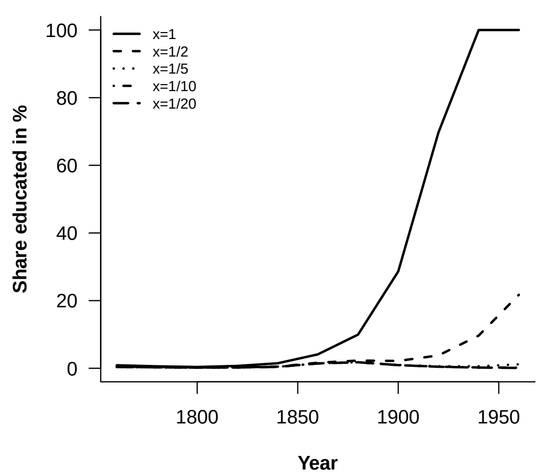
(c) Share of rural workers in rural areas



(d) Relative price of industrial goods



(e) Urbanization



(f) Education

Figure 15: Effect of the absence of rural exodus under reduced productivities in rural industries.

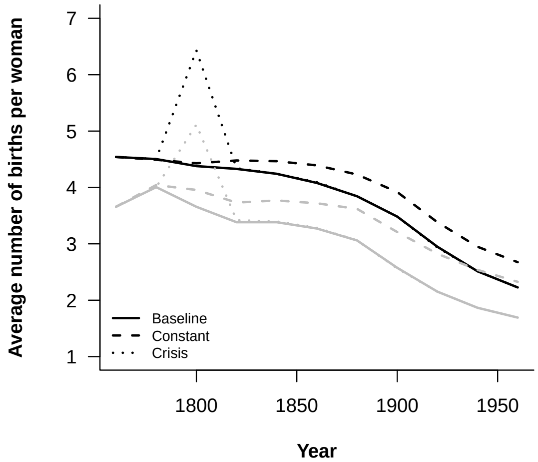
Overall, our results show that the vitality of rural industry during the eve of the Industrial Revolution protected Sweden from high relative migration costs and contributed to the containment of income inequalities between urban and rural places. Rural industrialization significantly reduced the need for high urbanization rates and a large rural exodus to fuel industrialization.

5.3 Mortality crises

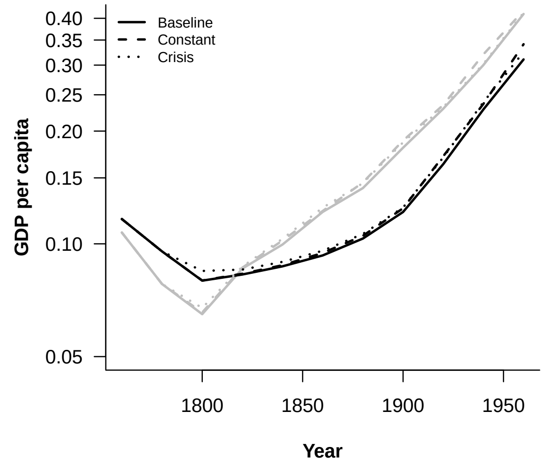
In line with papers such as Lagerlöf (2003), Weisdorf (2004), Strulik and Weisdorf (2014), and Bar and Leukhina (2010), we discuss the impact of IMRs on the demographic and economic dynamics of Sweden. Our main objective here is to evaluate the role that mortality reductions play in the take-off of modern economic growth and the fertility transition in our theory. For this purpose, we analyze (*i*) what the dynamics of Sweden would have been if IMRs had stagnated instead of declining and (*ii*) the effect of a mortality shock as important as the shocks provoked by the bubonic plague or by influenza pandemics.

Holding mortality rates constant at their 1760 level in both areas has a strong positive effect on the average number of births, especially in cities, as Figure 16 shows. 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 37.5% in cities and 20.1% in the countryside in 1960. Nevertheless, the number of surviving births would have declined by 4.8% and 3.1%. As a result, compared to that in the benchmark, the overall population size in this scenario would have been reduced by 13.1% after two centuries of mortality stagnation.[50] We find only a very limited impact of mortality stagnation on both urbanization and education rates; see Figures 16e and 16f. Positive Malthusian checks that reduce the net number of births are also not strong enough to postpone the onset of technological progress. They tend to foster the GDP per capita growth rates in both areas, even if this impact is quite contained.

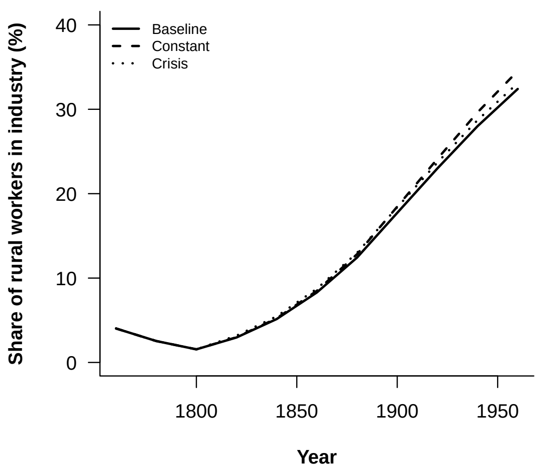
[50] In Table 9 in Appendix C.6, we provide the series of net changes in the number of births in each area.



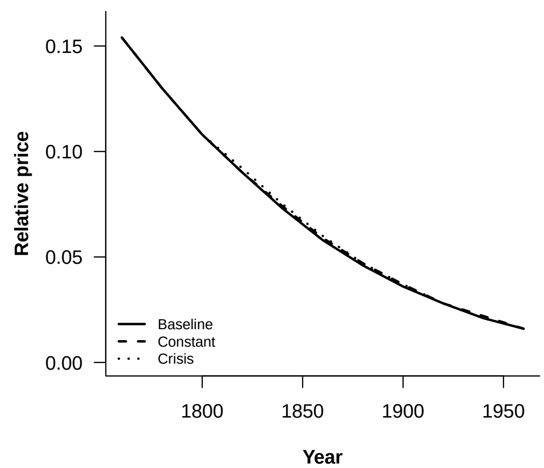
(a) Urban (gray) and rural (black) fertility



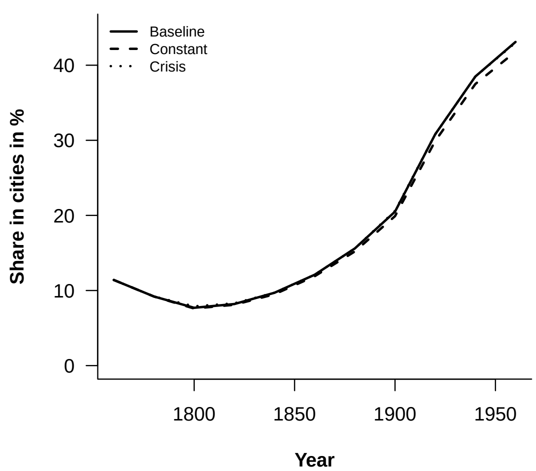
(b) GDP in current relative prices per capita in cities (gray) and countryside (black)



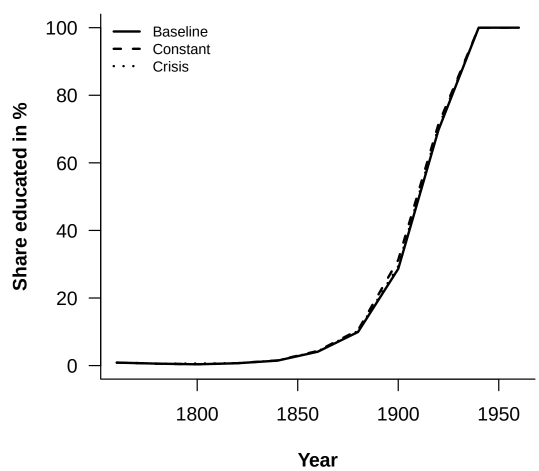
(c) Share of industrial workers in rural areas



(d) Relative price of industrial goods



(e) Urbanization



(f) Education

Figure 16: The impact of infant mortality on the Swedish demographic and economic evolutions.

We then simulate the impact of a mortality crisis by assuming that IMRs increased to 50% in 1800 (dotted lines in Figure 16). Such a rate should be considered a kind of upper limit for a mortality crisis, such as the influenza pandemics in 1831–33 and 1847–48 or the famine in 1772–73. The bubonic plague appeared 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 parishes with IMRs of up to 50% for farmers' children in terms of the five-year average in 1830-34, we use this upper bound for our simulation. At the time of the shock, the crude number of births increases substantially in both areas, but this number returns to its benchmark value as early as the next period. Interestingly, as highlighted in Appendix C.6, the mortality crisis leads to a decrease in net fertility. This lowers net fertility and increases GDP per capita through two mechanisms: it increases the labor supply and reduces the dependency ratio. As in the case of mortality stagnation, the impact on population growth is not sufficiently strong to delay the onset of global technological progress.

Overall, the effect of the IMR is quite restricted compared to that of the variation in κ or $A_0^{\mathcal{I}}$. 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 (2021) discusses issues related to the choice of the utility function. The impact of mortality might also increase if young adult mortality (Tamura, 2006; Tamura and Simon, 2017) or longevity (Boucekkine *et al.*, 2002; de la Croix, 2017) are taken into account.

6 Discussion

In the tradition of UGT, we present a fairly general model without adjusting the framework to account for country specificities and quantify the underlying mechanisms using the history of Sweden. Neglecting specific Swedish policies or institutional changes may raise some concerns about the validity of our predictions for Sweden, while limiting the quantitative analysis to Swedish data may cast doubt on the validity of our predictions outside Sweden. In the next paragraphs, we first show that Sweden is representative of other European countries along key economic and demographic dimensions during the observation period. Then, we discuss why key Swedish historical particularities do not undermine the generality of our conclusions.

The empirical Swedish demographic facts, which we have reproduced to calibrate our model, are representative of Western Europe between 1700 and 1950. First, the urban death penalty has been widely documented at the European level; see, for instance, Knodel (1974), Bairoch (1997) or Woods (2003). Second, rural populations were more fertile than their urban counterparts, as has been shown for many other countries, including France (Bocquier and Brée, 2018), Belgium (Bocquier and Costa, 2015), Japan (Mosk, 1980) and Denmark (Baudin and Stelter, 2021). In all these countries, the number of births per woman in rural areas exceeded that in urban areas. Third, despite these unfavorable natural population movements, urbanization nevertheless occurred in all these economies. Since neither international immigration nor emigration were able to sustain urbanization in European countries during the time we study, urbanization was only possible because of a massive and sustained rural exodus.

In one way or another, the two phases of industrialization we observe in the Swedish evolution are detectable in many other European economies. After Mendels (1972), the early phase of industrialization has often been associated with the concept of *protoindustrialization*. If economic historians have never agreed on a unified theory of protoindustrialization, they have also never disagreed on the importance of rural industry during the eve of the Industrial Revolution (Ogilvie and Cerman, 1996). Industrial activity increased in the countryside, which was less regulated and institutionalized than cities. The key industries varied across countries and did not coincide with the image of modern industries. Production was organized around simple tasks, such as nail production or weaving, and enabled peasant farmers and semiproletarianized workers to earn additional income (Houston and Snell, 1984). These still-agrarian economies were shaped by the role of domestic production and stepwise regional specialization. The British textile industry is a prominent example. In Denmark, the milk and butter industries were important contributors (Lampe and Sharp, 2019). In Sweden, early industrialization was based on "*a combination of iron core deposits and access to charcoal*" (Schön, 2010) as well as the textile industry and lasted until circa 1890.

The transition to the second phase of industrialization shifted the character of industrial production towards the image of modern industry. With mechanization, specialization and the movement out of homes, cities became the natural ecosystem for the Industrial Revolution; see Bertinelli and Black (2004) and Bairoch (1997) for their enlightening contributions. This transition appears to have been a global phenomenon across industrialized European economies and made urbanization a necessary driver of industrialization. Nevertheless, its timing and intensity differed between leaders such as Britain and Belgium and laggards such as France and Sweden. During the second part of the eighteenth century, when the Industrial Revolution began in Britain, Sweden was economically less advanced than the leading countries. It experienced a breakthrough in modern industry at the turn of the twentieth century (Schön, 2010), when cities became the epicenter of industrialization (Berger *et al.*, 2012). Progressive specialization encouraged the growth of the transport sector and the rise of exchange and trade via markets. Sweden became a small open economy and, hence, sensitive to price shocks in both directions. Not surprisingly, industrial prices relative to agricultural prices decreased throughout the European Industrial Revolution; see, for instance, Knick (1998) for textile prices in England and Mokyr (1977) for a more general discussion.

Like all Western European economies, Sweden experienced major institutional changes regarding education and property rights. Primary education became compulsory in 1842 with the "Folkskola," marking the onset of the transition to mass education (Flora, 1983). The great variety in the timing of these reforms across Europe may suggest that the institutional dimension of the rise in primary education is not key for the general economic take-off.[51] In contrast, the recent literature has highlighted the importance of the accumulation of upper-tail human capital during the process of industrialization; see the contributions of Squicciarini and Voigtländer (2015), Mokyr (2005) and De la Croix *et al.* (2020). Our quantitative exercise captures one key feature of this literature, namely, the fact that in the early phase of industrialization, human capital accumulated among the economic elite. Later, it passed to the rest of the population; see Appendix C.5.

As in Britain, enclosure acts redistributed cultivated land and enabled growth in the agricultural sector. The Swedish enclosure movement started around 1749. Another wave of stepwise enclosures occurred between 1803 and 1807. Finally, villages without an enclosure in 1827 implemented one after the new act in 1827. As documented by Rozental (1956), the Swedish enclosure movement was part of a larger agrarian reform favoring enclosure in France, Germany, Spain, Italy and the Netherlands.

That said, it would be inexact to claim that Sweden did not have any unique features. While the steam engine had become increasingly important in European industrial

[51] Prussia introduced compulsory education with the famous school edict in 1717, while its industrialization was rather late (Van Horn Melton, 2003). In contrast, in England (and Wales), as the leader in industrialization, primary education was not compulsory before 1870 when compulsory primary education was established with the Elementary Education Act (Flora, 1983).

production, it played a minor role in Sweden's relatively small industrial sector, which relied more heavily on water power. The stagnation of Stockholm is another unique phenomenon. From 1750 to 1850, Stockholm experienced a century of stagnation that cannot be related to any specific economic or demographic crisis nor to revolutionary or war episodes. This century of deindustrialization was marked by a shrinking fraction of active Swedish industrial workers in Stockholm's factories. Their share among all industrial workers declined from 50% to 12% between 1760 and 1846. At the same time, many of the textile industries closed, and real wages declined or stagnated. Limitations to capital accumulation facilitated egalitarian tendencies. The distribution of real estate became more equal. From a demographic perspective, mortality was high and exceeded births. Only the rural exodus that relieved the population pressure in surrounding areas prevented the clear shrinking of the urban population. Finally, with the rather late but successful adaption to modern capital- and skill-intensive industries such as mechanical engineering, Stockholm recovered from its stagnation circa 1850 (Söderberg *et al.*, 2003).

Despite these particularities, we can safely conclude that Sweden was part of the wider "European world economy." Swedish dynamics can be placed between the pioneering British dynamics and those of most continental European countries, which were hit hard by the Napoleonic wars in approximately 1800.

7 Conclusion

Based on unified growth theory, we present a dynamic general equilibrium model with two areas, two sectors of production and an explicit migration decision. We are the first to use a full-fledged simulated method of moments in a unified theory to evaluate the role that internal migration conditions played in the economic and demographic development of Western Europe. In the specific case of Sweden, we show that a key determinant of the take-off of sustained economic growth and the associated fertility transition was the balance between the initial ability of rural industry to sustain the first phase of industrialization and the prevailing migration costs during the eve of the Industrial Revolution.

Taken alone, variation in the initial productivity of rural industry has nontrivial implications. When rural industry is simulated to be highly productive in the middle of the eighteenth century, rural areas are prosperous, and the first phase of the industrialization process is quite salient. Nevertheless, rural prosperity tends to depress fertility in the entire country as well as to depress urbanization rates. These two movements tend to delay and weaken the second phase of the industrialization process, which is based on urban and human capital-intensive industries. Even if the long-run growth rate of the economy is not reduced, the convergence to that rate becomes increasingly longer and slower, which impoverishes the country.

Based on our estimations, increases in migration costs tend to delay the fertility transition but never prevent it from occurring. Nevertheless, maintaining high demographic pressure on the rural production sectors tends to impoverish the countryside and maintain a rich but vanishing elite in cities. However, migration costs, even when they are infinite, cannot prevent the start of both the fertility transition and the modern growth regime. Indeed, the Swedish rural industry was sufficiently productive to allow for some industrialization and accumulation of wealth.

In scenarios in which rural industrial productivity is reduced, rural industry can no longer sustain the first phase of industrialization, which causes the rural exodus and urbanization to become necessary conditions for industrialization, an industrialization based on human capital accumulation. In this situation, if internal migration costs become prohibitive, urban industry cannot sustain this process, and the country becomes trapped in secular poverty. It is in this sense that we claim that what mattered for the long-run development of early Western European countries was the balance between the productivity of rural industry in the middle of the eighteenth century and internal migration costs. It is in this sense that we claim that the conditions surrounding the rural exodus have been a key determinant of the dynamics of European economies.

While our quantitative exercise focuses on Sweden, we show that the Swedish case is representative of other Western European economies during the same time period. Most of the demographic, economic and educational facts that we document for Sweden hold across European countries, such as the UK, France and Denmark. Thus, we are convinced that our main conclusions about the role of migration conditions, infant mortality and the productivity in rural industry are generalizable to the European context. Nevertheless, country specifics, such as the pioneering role of Britain in the industrialization of Europe or the revolution and Napoleonic wars in the French case, likely altered the timing and magnitude 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 to 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 effects of this policy in the specific context of the one-child policy would help to understand how, today, restrictions to internal mobility influence economic growth.

References

ADAMOPOULOS, T. (2008). Land inequality and the transition to modern growth. *Review of Economic Dynamics*, **11** (2), 257–282.

ADAMS, J. J. (2021). Urbanization, long-run growth, and the demographic transition. *Journal of Demographic Economics*, **Forthcoming**.

ALFANI, G. and GRÁDA, C. Ó. (2018). The timing and causes of famines in Europe. *Nature Sustainability*, **1** (6), 283–288.

ANDERSSON, M. (2018). *Migration i 1600-talets Sverige: Älvsborgs lösen 1613–1618.* Ph.D. thesis, Diss. Universus Academic Press.

BAIROCH, P. (1997). *Victoires et déboires: histoire économique et sociale du monde du XVIe 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. and GOBBI, P. E. (2015). Fertility and childlessness in the United States. *The American Economic Review*, **105** (6), 1852–1882.

— and STELTER, R. (2021). A dynamic general equilibrium approach to migration in economic history. *High Performance Computing and Data Science in the Max Planck Society*, pp. 60–61.

BECKER, G. S. (1960). An economic analysis of fertility. In *Demographic and economic change in developed countries*, Columbia University Press, pp. 209–240.

BENAVOT, A. and RIDDLE, P. (1988). The expansion of primary education, 1870-1940: Trends and Issues. *Sociology of Education*, pp. 191–210.

BERGER, T. (2019). Railroads and rural industrialization: Evidence from a historical policy experiment. *Explorations in Economic History*, **74** (101277).

— 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.

—, — and HENNING, M. (2012). Geographical location and urbanisation of the Swedish manufacturing industry, 1900–1960: evidence from a new database. *Scandinavian Economic History Review*, **60** (3), 290–308.

BERTINELLI, L. and BLACK, D. (2004). Urbanization and growth. *Journal of Urban Economics*, **56** (1), 80–96.

BILLARI, F. and KOHLER, H.-P. (2004). Patterns of low and lowest-low fertility in Europe. *Population studies*, **58** (2), 161–176.

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.

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.

Brunt, L. and García-Peñalosa, C. (2021). Urbanisation and the onset of modern economic growth., aix-Marseilles School of Economics - Working Papers - WP 2021 - Nr 01.

Centralbyrån, K. S. (1914). Statistisk årsbok för Sverige 1914. *Stockholm (and subsequent editions)*.

Centralbyrån, S. (1969). Historisk statistik for Sverige. *Del I. Befolkning. Andra upplagan 1720–1967*.

Centralbyrån, S. (1999). Population development in Sweden in a 250-year perspective. *Historisk statistik för Sverige*.

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.

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.

de la Croix, D. (2017). Did Longer Lives Buy Economic Growth? From Malthus to Lucas and Ben-Porath. In M. Cervellati and U. Sunde (eds.), *Demographic Change and Long-Run Development*, MIT Press, pp. 69–89.

De la Croix, D., Docquier, F., Fabre, A. and Stelter, R. (2020). The Academic Market and the Rise of Universities in Medieval and Early Modern Europe (1000-1800).

de la Croix, D. and Doepke, M. (2003). Inequality and growth: why differential fertility matters. *The American Economic Review*, **93** (4), 1091–1113.

—, — 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.

—, Lindh, T. and Malmberg, B. (2008). Swedish economic growth and education since 1800. *Canadian Journal of Economics*, **41** (1), 166–185.

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. (2017). The response of vital rates to harvest fluctuations in pre-industrial sweden. *Cliometrica*, **11** (2), 245–268.

— 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.

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

— 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.

Gary, K. (2018). *Work, Wages and Income: Remuneration and Labor Patterrns in Sweden 1500-1850*. Ph.D. thesis, Lund University.

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.

Houston, R. and Snell, K. D. (1984). Proto-industrialization? Cottage industry, social change, and industrial revolution. *The Historical Journal*, **27** (2), 473–492.

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.

Kohler, H.-P., Billari, F. C. and Ortega, J. A. (2002). The emergence of lowest-low fertility in Europe during the 1990s. *Population and Development Review*, **28** (4), 641–680.

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.

Lampe, M. and Sharp, P. (2015). Just add milk: a productivity analysis of the revolutionary changes in nineteenth-century Danish dairying. *The Economic History Review*, **68** (4), 1132–1153.

— and — (2019). *A land of milk and butter: how elites created the modern Danish dairy industry*. University of Chicago Press.

Lees, A. and Lees, L. H. (2007). *Cities and the making of Modern Europe, 1750-1914*, vol. 39. Cambridge University Press.

Mendels, F. F. (1972). Proto-industrialization: the first phase of the industrialization process. *The Journal of Economic History*, pp. 241–261.

Mokyr, J. (1977). Demand vs. supply in the industrial revolution. *The Journal of Economic History*, **37** (4), 981–1008.

— (2005). Long-term economic growth and the history of technology. In *Handbook of economic growth*, vol. 1, Elsevier, pp. 1113–1180.

— (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.

OECD (2014). *Long-term baseline projection - OECD Economic Outlook: Statistics and Projections*. Tech. Rep. 95, Organisation for Economic Cooperation and Development, https://doi.org/10.1787/data-00690-en.

Ogilvie, S. (2014). The economics of guilds. *The Journal of Economic Perspectives*, **28** (4), 169–192.

— and Cerman, M. (1996). *European proto-industrialization: an introductory handbook*. Cambridge University Press.

Pensieroso, L. and Sommacal, A. (2019). Agriculture to Industry: the End of Intergenerational Coresidence. *Review of Economic Dynamics*, **34**, 87–102.

Powell, M. J. (2002). UOBYQA: unconstrained optimization by quadratic approximation. *Mathematical Programming*, **92** (3), 555–582.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy*, **94** (5), 1002–1037.

Saville, J. (2013). *Rural depopulation in England and Wales, 1851-1951*, vol. 9. Routledge.

Schön, L. (2008). Sweden–Economic Growth and Structural Change, 1800-2000. *EH. net Encyclopedia.*

Schön, L. (2010). *Sweden's road to modernity: an economic history.* SNS förlag Stockholm.

Schön, L. (2012). *An economic history of modern Sweden.* Routledge.

Schön, L. and Krantz, O. (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.

Sobotka, T. (2004). Is lowest-low fertility in Europe explained by the postponement of childbearing? *Population and Development Review*, **30** (2), 195–220.

Söderberg, J. (1984). A long-term perspective on regional economic development in Sweden, Ca. 1550–1914. *Scandinavian Economic History Review*, **32** (1), 1–16.

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.

Söderberg, J., Jonsson, U. and Persson, C. (2003). *A stagnating metropolis: the economy and demography of Stockholm, 1750–1850*, vol. 13. Cambridge University Press.

Soysal, Y. N. and Strang, D. (1989). Construction of the first mass education systems in nineteenth-century Europe. *Sociology of Education*, pp. 277–288.

Squicciarini, M. P. and Voigtländer, N. (2015). Human capital and industrialization: Evidence from the age of enlightenment. *The Quarterly Journal of Economics*, **130** (4), 1825–1883.

Statistics-Sweden (2018). *Sveriges framtida befolkning 2018–2070.* Demografiska rapporter 2018:1, Statistics Sweden.

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.

Swedish Historical Population Statistics (). The demographic data base on Swedish historical population statistics. CEDAR, Umeå University.

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

— (2006). Human capital and economic development. *Journal of Development Economics*, **79** (1), 26–72.

— and Simon, C. (2017). Secular fertility declines, baby booms, and economic growth: International evidence. *Macroeconomic Dynamics*, **21** (7), 1601–1672.

United-Nations (2018). *World Urbanization Prospects: The 2018 Revision*. Online edition, United Nations, Department of Economic and Social Affairs - Population Division.

van der Woude, A., Hayami, A. and De Vries, J. e. (1995). *Urbanization in history: a process of dynamic interactions*. Oxford University Press.

Van Horn Melton, J. (2003). *Absolutism and the eighteenth-century origins of compulsory schooling in Prussia and Austria*. Cambridge 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.

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 Data

A.1 General procedure

To develop a dataset that enabled us to quantify our theory, we gathered data from different sources. Focusing on data spanning more than two centuries presents some challenges. Definitions may vary across alternative sources, while in the same data source, such as official statistics, definitions may have changed over time. As explained in the main text, one of the main features of our data collection is that the definitions of urban and rural places are consistent across all our data sources. We document and discuss this fact in A.2. Another potential issue comes from Sweden's border changes throughout history. We limit ourselves to the modern territory of Sweden. Despite these elements of caution, alternative sources often provide varying numbers for the same year. To ensure that we are using the most reliable value in each case, we selected the values that are either identified as less biased in the literature, originate from official statistics, or reduce inconsistencies. We finally combined the sources to construct time series while trying to minimize breaks due to conflicting definitions.

Even if rich, our time series are incomplete because several observations are missing. To fill these gaps, we applied linear interpolation. As we are not interested in short-term fluctuations, this approach is defensible. In the next step, we smoothed the data, and as the time series no longer included missing values, we applied the Hodrick-Prescott filter (Hodrick and Prescott, 1997) with a smoothing parameter of 1200. Finally, we picked values every 20 years, which we then tried to match or used for overidentification checks.

A.2 Definition of rural and urban areas

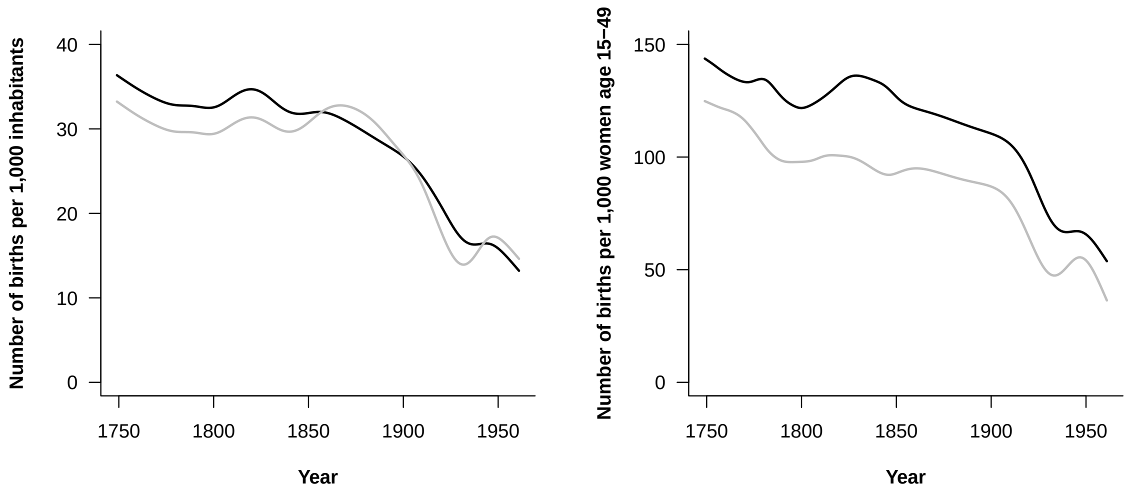
While our data on urban and rural populations originate from different sources, all of these sources base their definition of cities and the countryside on Centralbyrån (1969). This division follows legal principles and administrative terms and, hence, is not necessarily linked to population size or density. Municipal charters enabled trade and artisanship and, hence, were intended to stimulate economic development in the selected areas. This idea is close to the definition of a city used in our model: namely, places with modern technologies that support economic growth.

As explained in Section 2, data from SHiPs are based on records at the parish level. We have been able to associate each parish with its urban or rural status thanks to a retrieval tool offered by Umeå University (available upon request). Despite its numerous advantages, this tool does not allow us to reclassify parishes from urban to rural areas or vice versa over our observation period. That is, during our observation period, each parish in the SHiPs data has one and only one urban/rural status even if an official reclassification occurred over time. This limitation may seem major at first glance but it is not at all. Indeed, from 1749 to 1855, only four geographic units switched from rural to urban: Östersund, Skellefteå, Haparanda and Borgholm. Östersund has been a city since 1786 and is classified as urban in our dataset; Skellefteå became an urban place in 1845 and is classified as rural in our dataset. Haparanda also became a city in 1842, but unlike Skellefteå, it is classified as urban. However, SHiPs provides data on this parish beginning only in 1851, when it was already urban. Finally, Borgholm (1816) is not included in our data. We can therefore reasonably assert that our dataset does not suffer from any significant bias due to the absence of reclassifications.

A.3 Age structures and differential fertility in Sweden

Most commonly, crude birth rates (CBRs) are used to study the demographic transition; see, e.g., Dyson (2011) and Bocquier and Costa (2015). Figure 17a presents the trend in Dyson's CBRs for Sweden between 1750 and 1960. These estimated fertility differentials between cities and the countryside are likely to be biased, as the age distributions of women in rural and urban areas are potentially different. We illustrate these varying age structures in Fig. 18, in which we present age pyramids for Sweden's cities and countryside in 1910 (Centralbyrån, 1914). The proportion of women of fertile ages is significantly higher in cities (29.1%) than in the countryside (22.6%).[52]

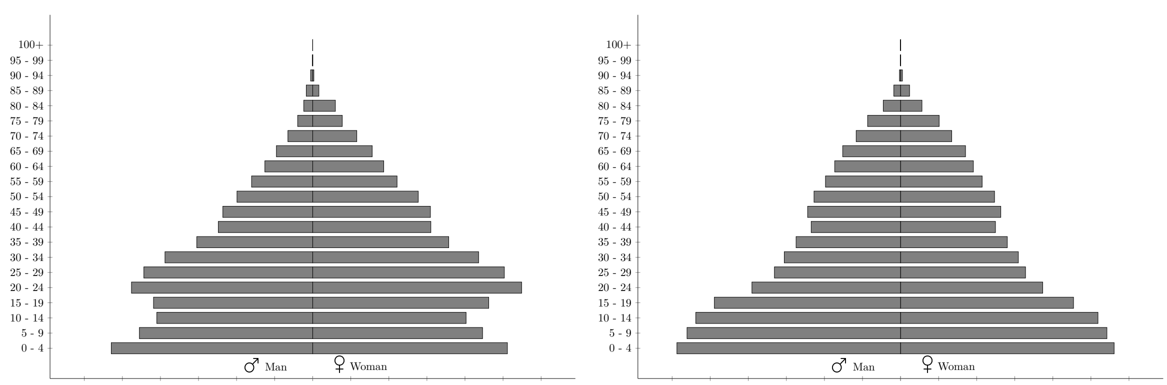
[52] Note that this difference deserves our particular attention when examining the effects of the rural exodus, as it was largely induced by the exodus. Figure 18 documents the flow of women of fertile ages from the countryside into cities, which could lead to an overestimation of fertility in cities relative to that in the countryside.



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

Figure 17: Swedish fertility in cities (gray) and in the countryside (black) according to different measures, 1760–1960.

Relying on the so-called generalized fertility rate (GFR) is one way to correct for biases due to dissimilar age structures. The GFR divides the number of births by the person-years lived among all women of fertile ages (15–49). As illustrated in Fig. 17b, merging data from SHiPs Swedish Historical Population Statistics and from statistical yearbooks, these corrections significantly change the ordering and dynamics of urban and rural fertility.



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

Figure 18: Age structure of Sweden's population by sex and area in 1910, according to the Statistical Yearbook, Sweden, 1914.

Though better than the CBR, the GFR cannot account for differences within the distribution of women of fertile ages. Since the exodus contributed to the varying age structures, it is important to rely on a fertility indicator that controls for differences in fertile ages. An indicator based on age-specific fertility rates such as the TFR is the optimal solution. However, data limitations do not allow us to use this first-best solution directly, and so we proceed in two steps. First, we estimate the general fertility index from Coale (1969). 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 with 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^j as the value of Coale's index for area j at time t:

$$I_t^j = \frac{B_t^j}{\sum_{o=1}^{T} x_{o,t}^j n_o^H}$$

where B_t^j denotes the number of births in area j at time t, while $x_{o,t}^j$ is the number of individuals in area j 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 as shown in the table below:[53]

Table 5: Age-specific fertility rates of Hutterites

o	15–19	20–24	25–29	30–34	35–39	40–44	45–49
n_o^H	0.300	0.550	0.502	0.447	0.406	0.222	0.061

Source: Coale (1969).

In a second step, we use Coale's index to approximate the TFR simply by multiplying the index by the TFR of the Hutterites. At the country level, we can compare our estimations with direct estimations of Sweden's TFR from the Human Fertility Database beginning in 1861. In fact, our estimation approximates the official dynamics very well.

[53] According to Wetherell (2001), these values are an underestimation of Hutterite fertility due to ad hoc assumptions made by Coale (1969). Re-estimating age-specific Hutterite fertility rates leads to the following age-specific fertility rates: 0.623, 0.549, 0.502, 0.447, 0.406, 0.236, and 0.063, with much higher fertility among women aged 15–19. Higher age-specific fertility rates should be used in cases of high marriage rates at very young ages. However, since 1750, more than 5% of the Swedish female population was never married at these ages, and as the impact on the estimated general fertility index is limited, the original estimation of Coale is preferable in the Swedish case. Furthermore, using the original estimation allows us to use estimations from the Princeton fertility project when necessary.

From 1861 to 1960, both the timing and magnitude of the fertility transition fit well. Only the fertility level in the 1860s is slightly underestimated.

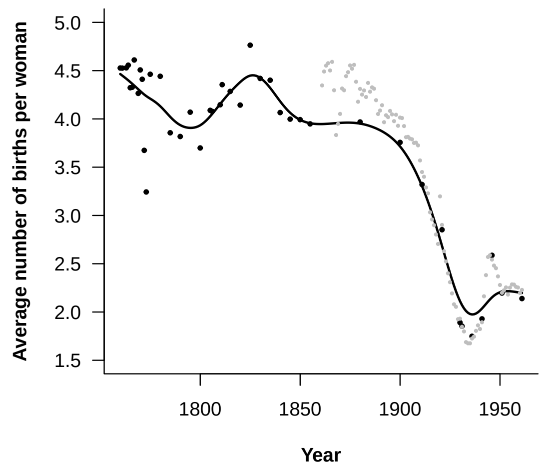


Figure 19: Comparison of Swedish fertility estimations: Authors estimated TFRs and their trend (black), TFR from the Human fertility database (gray).

A.4 Share of workers by location

Berger *et al.* (2012), Berger (2019) as well as Söderberg (1984) document a key feature of Swedish industrialization: rural areas triggered the first phase of industrialization, while urban areas took the lead thereafter. There are multiple reasons for such a situation; among them, patterns of regional specialization, natural resource availability and railroad development are key. While our theory assumes that agricultural activities are located exclusively in the countryside, it predicts the distribution of industrial workers between urban and rural places. To evaluate the quality of that prediction and the importance of rural industrialization for the Swedish economic take-off, we need to build a time series related to this fact.

We take advantage of two main data sources. The first is, once again, Centralbyrån (1969). Those data are displayed in Table 6, which presents the distribution of persons living from each kind of activity in each type of area. Retired persons are classified as working in their original industry, domestic servants in the industry of their employer and dependents in the industry of their head of family. We use these data as an approximation of the distribution of workers between sectors from 1870 to 1900.

Table 6: Share of individuals in each sector by area

	Total		Rural		Urban	
Year	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: Centralbyrån (1969), Table 23.

Our second source of information is Berger *et al.* (2012), who explicitly quantify the distribution of industrial workers rather than persons across cities and the countryside from 1900 to 1960. These data are decennial, as shown in Table 7, from which we use the line "Nonrural areas." Using one line or the other has its pros and cons. Using the largest 50 cities guarantees that we only include areas benefiting from city charters but not all of them, which then logically leads to an underestimation of the share of industrial workers located in urban areas, as shown in the table. On the other hand, using the concept of nonrural areas does not guarantee that all urban places used here received city charters, but it guarantees that most of the places that benefitted from these charters are included. We opt for the second choice while noting that the correlation between the two measures is 0.7.

Table 7: Urban share of manufacturing employment, 1900–1960.

Urban definition	1900	1910	1920	1930	1940	1950	1960
50 largest cities	0.39	0.41	0.46	0.49	0.51	0.53	0.51
Nonrural areas	0.42	0.46	0.52	0.55	0.63	0.72	0.75

Source: Berger *et al.* (2012), Table 5.

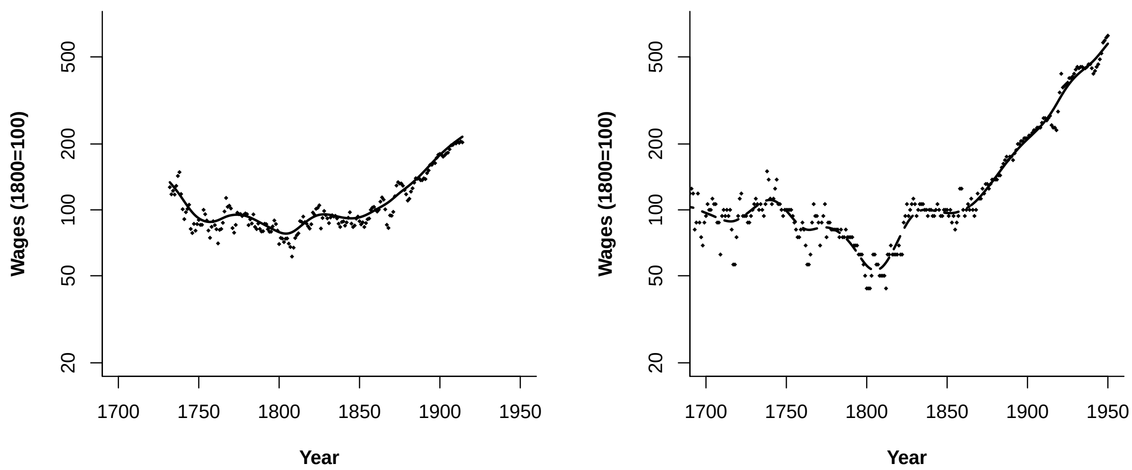
From these two data sources, we reconstruct a smoothed time series for the distribution of industrial workers between cities and the countryside from 1870 to 1960; we define $\mathcal{S}_t$ as the share of industrial workers located in the countryside. Based on this, we can compute the distribution of rural workers between agriculture and industry from 1850 to 1960. To do so, we first compute the number of industrial workers located in the countryside by multiplying $\mathcal{S}_t$ by the total number of industrial workers in Sweden (denoted M_t) as provided by the SHNA. Then, assuming that all agricultural workers (size J_t) are located in rural areas, we compute the share of workers allocated to industry within rural areas as $\frac{\mathcal{S}_t M_t}{\mathcal{S}_t M_t + J_t}$. We then obtain Figure 4.

In our calibration exercise, we further assume that from 1760 to 1850, the proportion of industrial workers located in the countryside increased by 1% per year, a reasonable assumption.

A.5 Wages

Figure 20 displays the dynamics in Swedish wages by sector, normalized to 100 in 1800. The agricultural wages in Fig. 20a are for males only. These data originate from Jörberg (1972) and are deflated by the time series presented in Edvinsson and Söderberg (2010). Industrial wages from Söderberg (2010) are the wages for unskilled workers in Stockholm through 1864 (dashed line in 20b) and industrial wages thereafter (solid black line in Figure 20b). At the beginning of the nineteenth century, both sectors—industry and agriculture—experienced a period of low wages. The decline in industrial wages was even stronger than in agriculture wages. In the 1820s, industrial wages started to increase and permanently took off in the 1860s. Agricultural wages followed in the 1870s.[54]

[54] For a comprehensive discussion of Swedish wages, see Gary (2018).



(a) Agricultural wages (b) Industrial wages

Figure 20: Evolution of observations (dots) and smoothed observations (solid lines) of Swedish wages.

B Theory

B.1 Regimes

To characterize the optimal decisions of adults living in area j for any configuration of our parameter set, we first define some critical values for wages and the relative price of industrial goods. $\bar{\omega}_t$ is the potential income above which a parent living in area j who had paid $\kappa \mathbb{1}_\kappa$ decides to educate her children:

$$\bar{\omega}_t \equiv \frac{\upsilon p_t \beta q_t^j}{\phi(\xi + \zeta^j q_t^j)}.$$

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

$$\hat{\omega}_t \equiv \left[\frac{\alpha+\rho}{1-\alpha}\varepsilon + \kappa \mathbb{1}_\kappa\right] p_t + \bar{c}.$$

Finally, $p^*(\kappa \mathbb{1}_\kappa)$ denotes the limit price below which poor parents who can afford $\bar{c}$ but cannot migrate could nevertheless educate their children, and $\tilde{p}(\kappa \mathbb{1}_\kappa)$ is the limit price above which $\bar{\omega}_t > \hat{\omega}_t$:

$$p^*(\kappa\mathbb{1}_\kappa) \equiv \frac{\phi(\xi+\zeta^j q_t^j)\bar{c}}{\upsilon\beta q_t^j - \phi(\xi+\zeta^j q_t^j)\kappa\mathbb{1}_\kappa} \quad , \quad \tilde{p}(\kappa\mathbb{1}_\kappa) \equiv \frac{\phi(\xi+\zeta^j q_t^j)\bar{c}}{\upsilon\beta q_t^j - \phi(\xi+\zeta^j q_t^j)\left[\kappa\mathbb{1}_\kappa + \frac{(\alpha+\rho)\varepsilon}{1-\alpha}\right]}.$$

Proposition 1

There exist $\{\bar{\bar{\upsilon}}_t(\kappa\mathbb{1}_\kappa), \bar{\upsilon}_t(\kappa\mathbb{1}_\kappa)\} \in \mathbb{R}^{2+}$***, defined as***

$$\bar{\upsilon}_t(\kappa\mathbb{1}_\kappa) = \frac{[(\alpha+\rho)\varepsilon + (1-\alpha)\kappa\mathbb{1}_\kappa]\,\phi\left(\xi+\zeta^j q_t^j\right)}{(1-\alpha)\,\beta q_t^j} \quad , \quad \bar{\bar{\upsilon}}_t(\kappa\mathbb{1}_\kappa) = \frac{\phi\left(\xi+\zeta^j q_t^j\right)\kappa\mathbb{1}_\kappa}{\beta q_t^j}$$

such that:

- ***If*** $\upsilon \leq \bar{\bar{\upsilon}}_t(\kappa\mathbb{1}_\kappa)$***, then*** $\tilde{p}(\kappa\mathbb{1}_\kappa) < 0$ ***and*** $p^*(\kappa\mathbb{1}_\kappa) \leq 0$***, and so the optimal decisions for people living in area*** j ***are:***

Not moving person ($\mathbb{1}_\kappa = 0$):			*Moving person* ($\mathbb{1}_\kappa = 1$):		
if $\omega_t^j \leq \bar{c}$	*then*	*R1*	*if* $\omega_t^j \in [\bar{c}+p_t\kappa, \hat{\omega}_t]$	*then*	*R5*
if $\omega_t^j \in]\bar{c}, \hat{\omega}_t]$	*then*	*R5*	*if* $\omega_t^j > \hat{\omega}_t$	*then*	*R6*
if $\omega_t^j \geq \hat{\omega}_t$	*then*	*R6*			

- ***If*** $\upsilon \in]\bar{\bar{\upsilon}}_t(\kappa\mathbb{1}_\kappa), \bar{\upsilon}_t(\kappa\mathbb{1}_\kappa)]$***, then*** $\tilde{p}(\kappa\mathbb{1}_\kappa) \leq 0 < p^*(\kappa\mathbb{1}_\kappa)$ ***such that:***

  when $p_t \leq p^*(\kappa\mathbb{1}_\kappa)$***, the optimal decisions for people living in area*** j ***are:***

Not moving person ($\mathbb{1}_\kappa = 0$):			*Moving person* ($\mathbb{1}_\kappa = 1$):		
if $\omega_t^j \leq \bar{c}$	*then*	*R1*	*if* $\omega_t^j \in [\bar{c}+p_t\kappa, \hat{\omega}_t]$	*then*	*R5*
if $\omega_t^j \in]\bar{c}, \hat{\omega}_t]$	*then*	*R5*	*if* $\omega_t^j > \hat{\omega}_t$	*then*	*R6*
if $\omega_t^j \geq \hat{\omega}_t$	*then*	*R6*			

  while if $p_t > p^*(\kappa\mathbb{1}_\kappa)$***, then:***

Not moving person ($\mathbb{1}_\kappa = 0$):

$$\begin{array}{lll} \textit{if } \omega_t^j \leq \bar{c} & \textit{then} & R1 \\ \textit{if } \omega_t^j \in]\bar{c}, \bar{\omega}_t] & \textit{then} & R2 \\ \textit{if } \omega_t^j \in]\bar{\omega}_t, \hat{\omega}_t] & \textit{then} & R5 \\ \textit{if } \omega_t^j \geq \hat{\omega}_t & \textit{then} & R6 \end{array}$$

Moving person ($\mathbb{1}_\kappa = 1$):

$$\begin{array}{lll} \textit{if } \omega_t^j \in [\bar{c}+p_t\kappa, \bar{\omega}_t] & \textit{then} & R3 \\ \textit{if } \omega_t^j \in [\bar{\omega}_t, \hat{\omega}_t] & \textit{then} & R5 \\ \textit{if } \omega_t^j > \hat{\omega}_t & \textit{then} & R6 \end{array}$$

- ***If*** $\upsilon > \bar{\upsilon}_t(\kappa\mathbb{1}_\kappa)$***, then*** $0 < p^*(\kappa\mathbb{1}_\kappa) < \tilde{p}(\kappa\mathbb{1}_\kappa)$ ***such that three situations may arise.***

  First, if $p_t \leq p^*(\kappa\mathbb{1}_\kappa)$***, then:***

  Not moving person ($\mathbb{1}_\kappa = 0$):

  $$\begin{array}{lll} \textit{if } \omega_t^j \leq \bar{c} & \textit{then} & R1 \\ \textit{if } \omega_t^j \in]\bar{c}, \hat{\omega}_t] & \textit{then} & R5 \\ \textit{if } \omega_t^j \geq \hat{\omega}_t & \textit{then} & R6 \end{array}$$

  Moving person ($\mathbb{1}_\kappa = 1$):

  $$\begin{array}{lll} \textit{if } \omega_t^j \in [\bar{c}+p_t\kappa, \hat{\omega}_t] & \textit{then} & R5 \\ \textit{if } \omega_t^j > \hat{\omega}_t & \textit{then} & R6 \end{array}$$

  For intermediate prices $p_t \in]p^*(\kappa\mathbb{1}_\kappa), \tilde{p}(\kappa\mathbb{1}_\kappa)]$***, we obtain:***

  Not moving person ($\mathbb{1}_\kappa = 0$):

  $$\begin{array}{lll} \textit{if } \omega_t^j \leq \bar{c} & \textit{then} & R1 \\ \textit{if } \omega_t^j \in]\bar{c}, \bar{\omega}_t] & \textit{then} & R2 \\ \textit{if } \omega_t^j \in]\bar{\omega}_t, \hat{\omega}_t] & \textit{then} & R5 \\ \textit{if } \omega_t^j \geq \hat{\omega}_t & \textit{then} & R6 \end{array}$$

  Moving person ($\mathbb{1}_\kappa = 1$):

  $$\begin{array}{lll} \textit{if } \omega_t^j \in [\bar{c}+p_t\kappa, \bar{\omega}_t] & \textit{then} & R3 \\ \textit{if } \omega_t^j \in [\bar{\omega}_t, \hat{\omega}_t] & \textit{then} & R5 \\ \textit{if } \omega_t^j > \hat{\omega}_t & \textit{then} & R6 \end{array}$$

  Finally, if p_t ***is greater than*** $\tilde{p}(\kappa\mathbb{1}_\kappa)$***, then:***

  Not moving person ($\mathbb{1}_\kappa = 0$):

  $$\begin{array}{lll} \textit{if } \omega_t^j \leq \bar{c} & \textit{then} & R1 \\ \textit{if } \omega_t^j \in]\bar{c}, \hat{\omega}_t] & \textit{then} & R2 \\ \textit{if } \omega_t^j \in]\hat{\omega}_t, \bar{\omega}_t] & \textit{then} & R4 \\ \textit{if } \omega_t^j > \bar{\omega}_t & \textit{then} & R6 \end{array}$$

  Moving person ($\mathbb{1}_\kappa = 1$):

  $$\begin{array}{lll} \textit{if } \omega_t^j \in [\bar{c}+p_t\kappa, \hat{\omega}_t] & \textit{then} & R3 \\ \textit{if } \omega_t^j \in [\hat{\omega}_t, \bar{\omega}_t] & \textit{then} & R4 \\ \textit{if } \omega_t^j > \bar{\omega}_t & \textit{then} & R6 \end{array}$$

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

Proof.

To prove Proposition 1, we first present and prove four lemmas. To simplify the notation, the time indices have been suppressed.

Lemma 1 $n_k < n_M \; \forall k = \{1,2,...,6\}$ *where* k *indexes the regimes described in Table 1.*

Proof. An inspection of each n_k yields the following results:

- $n_1 = 0 < n_M$

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

- n_3 is continuous with respect to ω, $\forall \omega^j \in \mathbb{R}^+$. Furthermore, $\frac{dn_3}{d\omega^j} = \frac{\bar{c}+\kappa \mathbb{1}_\kappa p}{(\alpha+\rho)(\xi+\zeta^j q^j)\omega^j} > 0$, $n_3(\bar{c}+\kappa \mathbb{1}_\kappa p) = 0$ and $\lim_{\omega^j \to +\infty} n_3(\omega^j) = \frac{\rho}{\alpha+\rho} n_M < n_M$. We can therefore conclude that $n_3 < n_M$, $\forall \omega^j$.

- n_4 is continuous with respect to ω, $\forall \omega^j \in \mathbb{R}^+$. Furthermore, n_4 is a monotonic functionof ω^j, as the sign of $\frac{dn_4}{d\omega^j}$ does not depend on ω^j:

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

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

- Simple computations yield the conclusion that:

$$n_5(\omega^j) \le n_3(\omega^j) < n_M \;,\; \forall \omega^j \ge \bar{\omega}$$
$$n_6(\omega^j) \le n_4(\omega^j) < n_M \;,\; \forall \omega^j \ge \bar{\omega}$$

■

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

- $\{c_1, n_1, d_1, e_1\}$ *is, by assumption, the solution to the individual maximization program* $\forall \omega^j \le \bar{c}$.

- $\{c_2, n_2, d_2, e_2\}$ *maximizes 13 subject to 14,* $n_k < n_M$, $c \ge \bar{c}$, $n_k > 0$ *and* $d_k = e_k = 0$ *when* $\mathbb{1}_\kappa = 0$.

- $\{c_3, n_3, d_3, e_3\}$ *maximizes* 13 *subject to* 14, $n_k < n_M$, $c \geq \bar{c}$, $n_k > 0$ *and* $d_k = e_k = 0$ *when* $\mathbb{1}_\kappa = 1$.

- $\{c_4, n_4, d_4, e_4\}$ *maximizes* 13 *subject to* 14, $n_k < n_M$, $c \geq \bar{c}$, $n_k > 0$ *and* $d_k > 0$ *and* $e_k = 0$ $\forall \kappa \geq 0$.

- $\{c_5, n_5, d_5, e_5\}$ *maximizes* 13 *subject to* 14, $n_k < n_M$, $c \geq \bar{c}$, $n_k > 0$ *and* $d_k = 0$ *and* $e_k > 0$ $\forall \kappa \geq 0$.

- $\{c_6, n_6, d_6, e_6\}$ *maximizes* 13 *subject to* 14, $n_k < n_M$, $c \geq \bar{c}$, $n_k > 0$ *and* $d_k > 0$ *and* $e_k > 0$ $\forall \kappa \geq 0$.

Proof. All the results are obtained through traditional analysis of the maximization problems described in each subcase.

■

Lemma 3 *When* $\upsilon \leq \bar{\bar{\upsilon}}(\kappa \mathbb{1}_\kappa)$, *then* $\tilde{p}(\kappa \mathbb{1}_\kappa) < 0$ *and* $p^*(\kappa \mathbb{1}_\kappa) \leq 0$. *When* $\upsilon \in]\bar{\bar{\upsilon}}(\kappa \mathbb{1}_\kappa), \bar{\upsilon}(\kappa \mathbb{1}_\kappa)]$, *then* $\tilde{p}(\kappa \mathbb{1}_\kappa) \leq 0 < p^*(\kappa \mathbb{1}_\kappa)$, *while when* $\upsilon > \bar{\upsilon}(\kappa \mathbb{1}_\kappa)$, *then* $0 < p^*(\kappa \mathbb{1}_\kappa) < \tilde{p}(\kappa \mathbb{1}_\kappa)$.

Proof. We first show that $p^*(\kappa \mathbb{1}_\kappa) \leq 0$ when $\upsilon \leq \bar{\bar{\upsilon}}(\kappa \mathbb{1}_\kappa)$, while $\tilde{p}(\kappa \mathbb{1}_\kappa) \leq 0$ when $\upsilon \leq \bar{\upsilon}(\kappa \mathbb{1}_\kappa)$:

$$
\begin{aligned}
p^*(\kappa \mathbb{1}_\kappa) \leq 0 \quad &\Leftrightarrow \quad \frac{\phi(\xi + \zeta^j q^j)\bar{c}}{\upsilon \beta q^j - \phi(\xi + \zeta^j q^j)\kappa \mathbb{1}_\kappa} \leq 0 \\
&\Leftrightarrow \quad \upsilon \leq \frac{\phi\left(\xi + \zeta^j q^j\right)\kappa \mathbb{1}_\kappa}{\beta q^j} = \bar{\bar{\upsilon}}(\kappa \mathbb{1}_\kappa) \\
\tilde{p}(\kappa \mathbb{1}_\kappa) \leq 0 \quad &\Leftrightarrow \quad \frac{\phi(\xi + \zeta^j q^j)\bar{c}}{\upsilon \beta q^j - \phi(\xi + \zeta^j q^j)\left[\kappa \mathbb{1}_\kappa + \frac{(\alpha+\rho)\varepsilon}{1-\alpha}\right]} \leq 0 \\
&\Leftrightarrow \quad \upsilon \leq \frac{[(\alpha+\rho)\varepsilon + (1-\alpha)\kappa \mathbb{1}_\kappa]\,\phi\left(\xi + \zeta^j q^j\right)}{(1-\alpha)\,\beta q^j} = \bar{\upsilon}(\kappa \mathbb{1}_\kappa)
\end{aligned}
$$

Then, we address the comparison between $\tilde{p}(\kappa \mathbb{1}_\kappa)$ and $p^*(\kappa \mathbb{1}_\kappa)$:

$$
\tilde{p}(\kappa \mathbb{1}_\kappa) > p^*(\kappa \mathbb{1}_\kappa) \Leftrightarrow \frac{\phi(\xi + \zeta^j q^j)\bar{c}}{\upsilon \beta q^j - \phi(\xi + \zeta^j q^j)\left[\kappa \mathbb{1}_\kappa + \frac{(\alpha+\rho)\varepsilon}{1-\alpha}\right]} > \frac{\phi(\xi + \zeta^j q^j)\bar{c}}{\upsilon \beta q^j - \phi(\xi + \zeta^j q^j)\kappa \mathbb{1}_\kappa} \tag{29}
$$

If $\upsilon \in]\bar{\bar{\upsilon}}(\kappa \mathbb{1}_\kappa), \bar{\upsilon}(\kappa \mathbb{1}_\kappa)]$, the denominator on the left-hand side of inequality (29) is negative such that inequality (29) is never satisfied. This means that $\tilde{p}(\kappa \mathbb{1}_\kappa) \leq 0 < p^*(\kappa \mathbb{1}_\kappa)$.

By contrast, if $\upsilon > \bar{\upsilon}(\kappa \mathbb{1}_\kappa)$, inequality (29) simplifies to $\frac{\alpha+\rho}{1-\alpha}\varepsilon\phi(\xi+\zeta^j q^j) > 0$, which is always satisfied, meaning that if $\upsilon > \bar{\upsilon}(\kappa \mathbb{1}_\kappa)$, $\tilde{p}(\kappa \mathbb{1}_\kappa) > p^*(\kappa \mathbb{1}_\kappa) > 0$.

■

Lemma 4

- *If* $\upsilon \leq \bar{\bar{\upsilon}}(\kappa \mathbb{1}_\kappa)$, *then* $\tilde{p}(\kappa \mathbb{1}_\kappa) < 0$ *and* $p^*(\kappa \mathbb{1}_\kappa) \leq 0$ *such that* $\forall\ p > 0$, $\bar{\omega} < \bar{c} + p\kappa \mathbb{1}_\kappa \leq \hat{\omega}$.
- *If* $\upsilon \in]\bar{\bar{\upsilon}}(\kappa \mathbb{1}_\kappa), \bar{\upsilon}(\kappa \mathbb{1}_\kappa)]$, *then* $\tilde{p}(\kappa \mathbb{1}_\kappa) \leq 0 < p^*(\kappa \mathbb{1}_\kappa)$ *such that:*
  - *if* $p \leq p^*(\kappa \mathbb{1}_\kappa)$, *then* $\bar{\omega} < \bar{c} + p\kappa \mathbb{1}_\kappa \leq \hat{\omega}$
  - *if* $p > p^*(\kappa \mathbb{1}_\kappa)$, *then* $\bar{c} + p\kappa \mathbb{1}_\kappa < \bar{\omega} < \hat{\omega}$
- *If* $\upsilon > \bar{\upsilon}(\kappa \mathbb{1}_\kappa)$, *then* $0 < p^*(\kappa \mathbb{1}_\kappa) < \tilde{p}(\kappa \mathbb{1}_\kappa)$ *such that:*
  - *if* $p \leq \tilde{p}(\kappa \mathbb{1}_\kappa)$, *then* $\bar{c} + p\kappa \mathbb{1}_\kappa < \bar{\omega} \leq \hat{\omega}$
  - *if* $p > \tilde{p}(\kappa \mathbb{1}_\kappa)$, *then* $\bar{c} + p\kappa \mathbb{1}_\kappa < \hat{\omega} < \bar{\omega}$

The proof of this lemma directly follows from Lemma 3 and the comparison between $\bar{w}$ and $\bar{c} + p\kappa \mathbb{1}_\kappa$, as well as the comparison between $\bar{w}$ and $\hat{w}$:

$$\bar{\omega} < \bar{c} + p\kappa \mathbb{1}_\kappa \quad \Leftrightarrow \quad p < \frac{\phi(\xi+\zeta^j q^j)(\bar{c}+\kappa \mathbb{1}_\kappa)}{\upsilon\beta q^j} = p^*(\kappa \mathbb{1}_\kappa)$$

$$\bar{\omega} < \hat{\omega} \quad \Leftrightarrow \quad p < \frac{\bar{c}+\kappa \mathbb{1}_\kappa}{\frac{\upsilon\beta q^j}{\phi(\xi+\zeta^j q^j)} - \frac{\alpha+\rho}{1-\alpha}\varepsilon} = \tilde{p}(\kappa \mathbb{1}_\kappa).$$

■

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

B.2 Non-monotonicity of fertility behavior

We study the specific situation of a person who does not move when υ is relatively high, while the equilibrium price prevailing on the market for industrial goods takes on intermediate values.[55] In this situation, we find that:

$$n_t = \begin{cases} 0 & \text{if } \omega_t^j \leq \bar{c} \\ \frac{\rho\left(\omega_t^j - \bar{c}\right)}{(\alpha+\rho)\left(\xi+\zeta^j q_t^j\right)\omega_t^j} & \text{if } \bar{c} < \omega_t^j \leq \bar{\omega}_t \\ \frac{\rho(1-\phi)\left(\omega_t^j - \bar{c}\right)}{(\alpha+\rho)\left[\left(\xi+\zeta^j q_t^j\right)\omega_t^j - \beta p_t q^j \upsilon\right]} & \text{if } \bar{\omega}_t < \omega_t^j \leq \hat{\omega}_t \\ \frac{\rho(1-\phi)\left(\tilde{\omega}_t^j - \bar{c}\right)}{(1+\rho)\left[\left(\xi+\zeta^j q_t^j\right)\omega_t^j - \beta p_t q_t^j \upsilon\right]} & \text{if } \omega_t^j > \hat{\omega}_t \end{cases} \quad ; \quad \frac{\partial n_t}{\partial \omega_t^j} \begin{cases} = 0 & \text{if } \omega_t^j \leq \bar{c} \\ > 0 & \text{if } \bar{c} < \omega_t^j \leq \bar{\omega}_t \\ \leq 0 & \text{if } \bar{\omega}_t < \omega_t^j \leq \hat{\omega}_t \\ \leq 0 & \text{if } \omega_t^j > \hat{\omega}_t \end{cases}$$

The relationship between optimal fertility and wages is presented in Figure 21. For our example 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 capacity. Any increase in ω_t^j will allow this person to have more children. When her income exceeds $\bar{\omega}_t^j$, she educates her children, and a trade-off between the quality and the quantity of children takes place. The higher ω_t^j is, 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 ω_t^j becomes higher than $\hat{\omega}^j$, investing in the education of children can be done without having to give up the consumption of industrial goods.

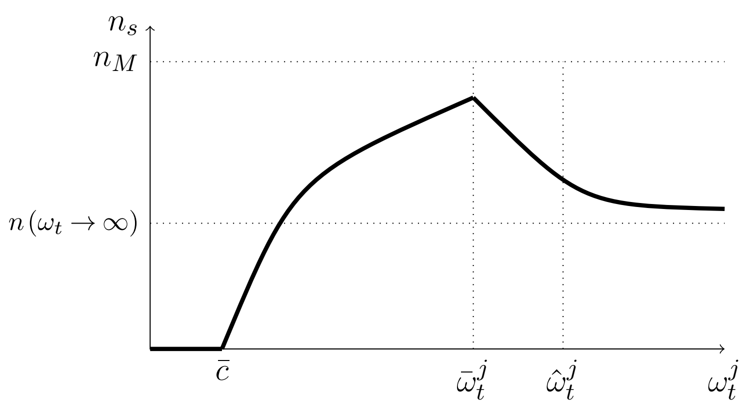


Figure 21: Fertility and income of parents who do not move.

[55] Using the notation from Appendix B.1, we study a case in which $\upsilon > \bar{\upsilon}(\kappa \mathbb{1}_\kappa)$ and $p_t \in]p^*(\kappa \mathbb{1}_\kappa), \tilde{p}(\kappa \mathbb{1}_\kappa)]$.

B.3 Long-run properties

In this subsection, we first explain why GDP per capita in urban industries and agriculture (and rural industries by extension) should grow at the same rate on a Balanced Growth Path (BGP hereafter). Then, we show that in a simplified version of the model, we can characterize a rate of rural exodus compatible with this equality and the stability of urbanization in the long run.

Let's assume that on the BGP, everyone lives in Regime 6. Indirect utilities depend positively on income per capita in each area; as there is no area specific differences in preferences, the functional form linking indirect utilities to income is the same whatever the place of living. Let's recall that an adult will decide to live in the countryside if:

$$v^*(p_t,\omega_t^R(a),q_t^R) > v^*(p_t,\omega_t^U(h_t),q_t^U)$$

$$\Leftrightarrow \alpha \ln\left[\frac{\alpha}{1+\rho}(\omega_t^R+\varepsilon p)\right]+(1-\alpha)\left[\frac{1-\alpha}{1+\rho}(\frac{\omega_t^R+\varepsilon p}{p})\right]+\rho\ln(\omega_t^R+\varepsilon p)-\rho(1-\phi)$$
$$+\ln\left[(\xi+\zeta^R q_t^R)\omega_t^R-\beta q_t^R\nu\right]\geq \alpha\ln\left[\frac{\alpha}{1+\rho}(\omega_t^U+\varepsilon p)\right]+(1-\alpha)\left[\frac{1-\alpha}{1+\rho}(\frac{\omega_t^U+\varepsilon p}{p})\right]$$
$$+\rho\ln(\omega_t^U+\varepsilon p)-\rho(1-\phi)\ln\left[(\xi+\zeta^U q_t^U)\omega_t^U-\beta q_t^U\nu\right]$$

In order to ensure that at least some members of the last dynasty populating the countryside agree to stay in this area, the above inequality needs to be verified what is done in the long run only if ω_t^R and ω_t^U grow at the same rate. From this, we get that $g_{y^{\mathcal{A}}}=g_{y^{\mathcal{M}}}$ on a BGP provided that the relative price is constant.

Now, we want to show that in a simplified version of the model where (i) there is no inequalities, (ii) we disregard the rural industrial sector, (iii) $q^R n^R=\frac{q^R\rho(1-\phi)}{(1+\rho)[\xi+\zeta^R q^R]}<1$, iv) $\zeta^R<\zeta^U$ and (v) migration from rural to urban places is the only form of migration in the long run,[56] we can characterize a constant rate of rural exodus ensuring the stability of urbanization rates and the equality between growth rates in agriculture and urban industries. To do so, let's define the rate of rural exodus as δ_t. It corresponds to the proportion of children born in the countryside in period $t-1$ who decide to live in cities once adult in period t. In order to alleviate notations, let's introduce $\mathfrak{n}_t^j\equiv q_t^j n_t^j$ as the net fertility rate prevailing in area j. Let's also express the urbanization rate at date t, u_t such as:

[56]Parametric conditions (iii) to (v) are satisfied in our calibration and simulations.

$$u_t = \frac{\left(1+\mathfrak{n}_t^U\right) N_t^U}{(1+\mathfrak{n}_t^U)N_t^U + (1+\mathfrak{n}_t^R)N_t^R},$$

where N_t^U and N_t^R define respectively the total number of adults contained in the sets S_t^U and S_t^R; in simpler terms, N_t^U and N_t^R are the number of adults living in urban and rural areas. On a BGP, urbanization rates have to be constant; $u_{t+1} = u_t = u^*$ is satisfied at the following condition:

$$\frac{N_{t+1}^U}{N_t^U} = \frac{(1+\mathfrak{n}^U)N_{t+1}^U + (1+\mathfrak{n}^R)N_{t+1}^R}{(1+\mathfrak{n}^U)N_t^U + (1+\mathfrak{n}^R)N_t^R} \tag{30}$$

Because migration from rural to urban areas is the only type of migration prevailing along the BGP, we get that:

$$\begin{aligned} N_{t+1}^U &= \mathfrak{n}_t^U N_t^U + \delta_t \mathfrak{n}_t^R N_t^R \\ N_{t+1}^R &= \mathfrak{n}_t^R N_t^R (1-\delta_t) \end{aligned}$$

Using these expressions and expressing everything in function of $\frac{N_t^R}{N_t^U}$, we can re-express Equation 30 such as:

$$\mathfrak{n}^U + \delta_t \mathfrak{n}^R \frac{N_t^R}{N_t^U} = \frac{(1+\mathfrak{n}^U)n^U + (1+\mathfrak{n}^U)\delta_t \mathfrak{n}^R \frac{N_t^R}{N_t^U} + (1+\mathfrak{n}^R)\mathfrak{n}^R(1-\delta_t)\frac{N_t^R}{N_t^U}}{(1+\mathfrak{n}^U) + (1+\mathfrak{n}^R)\frac{N_t^R}{N_t^U}} \tag{31}$$

Simple computations based on the definition of u_t implies that:

$$\frac{N_t^R}{N_t^U} = \frac{1+\mathfrak{n}_t^U}{1+\mathfrak{n}_t^R}\frac{1-u_t}{u_t}$$

Using this, we can re-express Equation 31 along the BGP as:

$$\mathfrak{n}^U + \delta_t \mathfrak{n}^R \frac{1+\mathfrak{n}^U}{1+\mathfrak{n}^R}\frac{1-u}{u} = \mathfrak{n}^U u + \mathfrak{n}^R(1-u) + \frac{\delta_t \mathfrak{n}^R(\mathfrak{n}^U - \mathfrak{n}^R)(1-u)}{1+\mathfrak{n}^R} \tag{32}$$

This equation simplifies to:

$$u = \frac{\delta_t \mathfrak{n}^R(1+\mathfrak{n}^U)}{(\mathfrak{n}^R - \mathfrak{n}^U)(1+\mathfrak{n}^R(1-\delta_t))} \tag{33}$$

Knowing that $n^R > n^U$, this equation defines u as a function of δ_t and as this equation holds only when u_t is constant, it implies that δ_t has to be constant too. It remains now to determine whether we can find a constant value of δ_t ensuring the equality of growth rates in agricultural and industrial sectors. To do so, one has to remark that $g_{y^{\mathcal{A}}} = g_{y^{\mathcal{M}}}$ if and only if $-\theta g_{L^{\mathcal{A}}} = g_{h^U}$. In this simplified version of the model, the labor force in agriculture is defined as:

$$L_t^{\mathcal{A}} = \left[1 - (\xi + \zeta^R q^R) n^R\right] N_t^R$$

Because q^R and n^R are constant in the long run, we get that $\frac{L_{t+1}^{\mathcal{A}}}{L_t^{\mathcal{A}}} = \frac{N_{t+1}^R}{N_t^R}$. By definition, $N_{t+1}^R = (1-\delta_t)\mathfrak{n}_t^R N_t^R$ what leads to:

$$\frac{L_{t+1}^{\mathcal{A}}}{L_t^{\mathcal{A}}} = 1 + g_{L_t^{\mathcal{A}}} = (1-\delta_t)\mathfrak{n}^R$$

Because $N_t < \bar{N}$ in the long run, the necessary equality between agricultural and industrial growth implies that $\delta = 1 - \frac{1 - \frac{g_h^U}{\theta}}{\mathfrak{n}^R} \equiv \delta^*$. Following Equations 27 and 33, δ^* defines $u^* = u(\delta^*)$ along the BGP on which $g_{y^{\mathcal{A}}} = g_{y^{\mathcal{M}}} = g_{h^U} = \left(\frac{\phi(\zeta+\zeta^U q^U)A^*}{(1-\phi)\beta p q^U}\right)^{\phi} - 1$.[57]

C Calibration and historical experiments

C.1 Model fit

In this appendix, we summarize the fit of our model through some statistics. We rely on two indicators: (*i*) the R^2 from linear regressions in which 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.[58]

[57] If relative price p would not be constant along the BGP, this equality would become $g_{y^{\mathcal{A}}} = g_{y^{\mathcal{M}}} = g_{h^U} + g_p = \left(\frac{\phi(\zeta+\zeta^U q^U)A^*}{(1-\phi)\beta p q^U}\right)^{\phi} + g_p - 1$, where g_p would denote the growth rate of prices. A simple inspection of this equation indicates that growth rates of $g_{y^{\mathcal{A}}}$ and $g_{y^{\mathcal{M}}}$ can be constant only if $g_p = 0$ as otherwise, $\left(\frac{\phi(\zeta+\zeta^U q^U)A^*}{(1-\phi)\beta p q^U}\right)^{\phi}$ would not be constant.

[58] The latter is used for all growth rates, while the remaining values in Column 4 of Table 8 are deviations in percentages.

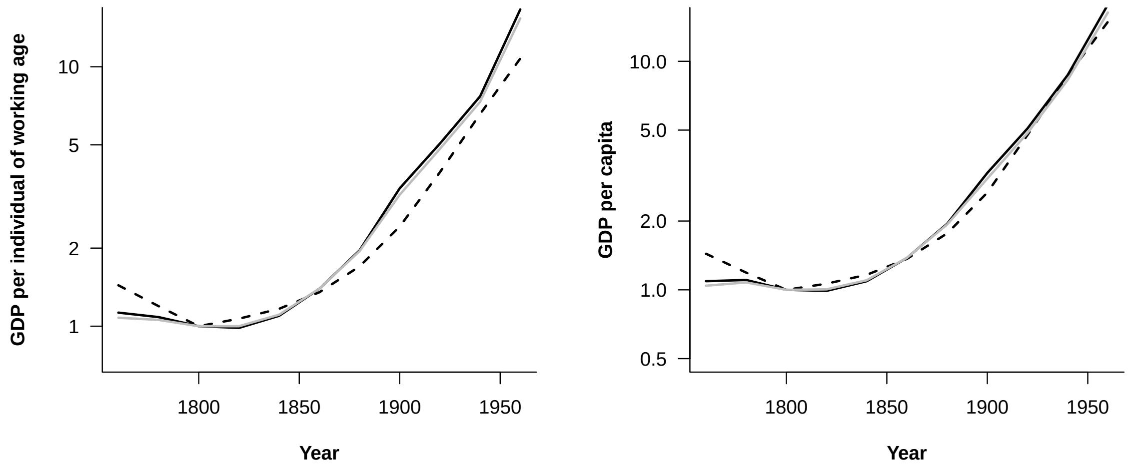
Table 8: Evaluation of model's fit

Variable	α_0	α_1	R^2	%-deviation
TFR in U	0.04142	0.82651***	0.8454	10.5412
TFR in R	0.02003	0.93412***	0.9267	6.0799
Share in U	-0.01113	1.10384***	0.9625	12.6544
Share rural workers in $\mathcal{I}$	-0.06858	1.07261***	0.9742	3.0520
rel. price	0.17690**	1.16442***	0.9763	41.1185
primary Education	0.23136*	0.68984**	0.5662	66.9882
secondary Education ($\bar{e}^S$)	0.012075*	0.919609***	0.9647	89.9950
tertiary Education ($\bar{e}^T$)	0.007819**	0.746687**	0.6817	92.7263
Share ind. workers in U	-1.4236*	2.8851*	0.7976	34.6646
Modified share ind. workers in U	0.1020	0.8098***	0.7844	7.0899
Growth rate agr. GDP per worker	0.268	1.262	0.2706	0.9293
Growth rate ind. GDP per worker	-1.3465	1.5275.	0.7618	0.3583
GDP[1] per person working age	-0.78094*	1.53468***	0.983	18.9446
GDP[2] per person working age	-0.62303*	1.41640***	0.9856	17.8128
GDP[1] per capita	-0.30939	1.17077***	0.9928	13.0816
GDP[2] per capita	-0.18767	1.08114***	0.9945	13.3854

[1] Only the industrial and agricultural sector, [2] All sectors included in the observations.

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$,. $p < 0.1$

As Table 8 shows, the model predicts the estimated trends very well for the time series we try to match: fertility, urbanization and the allocation of labor in the countryside between agriculture and industry. While the fit is less good for the moments we do not try to match, it is still convincing. For total GDP per person and in constant relative prices, normalized by its values in 1760, the model reproduces the observed values almost perfectly. The goodness of this fit depends neither on whether we include only industry and agriculture in GDP or all sectors nor whether GDP is divided by the total population or the working-age population (see Figure 22). The good fit of the overidentification checks shows the quality of the predictions made by our model.



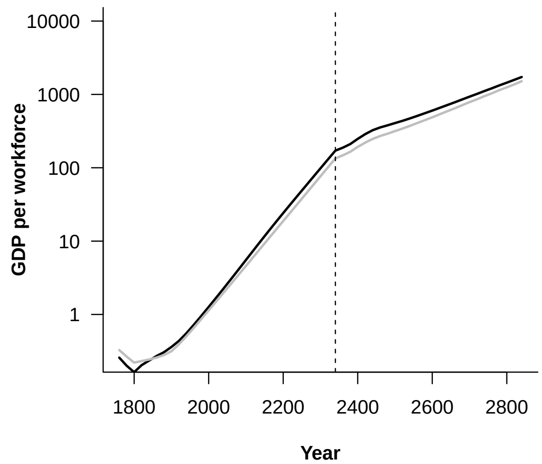
(a) GDP per working-age individual

(b) GDP per capita

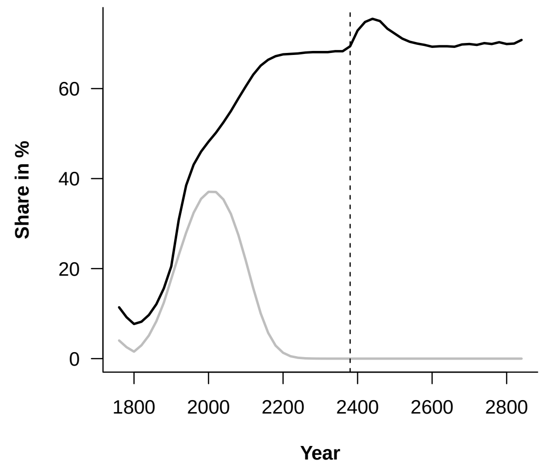
Figure 22: Predicted values for GDP (dashed black) versus observations with all observed sectors (gray) and with agriculture and industry only (solid black).

C.2 Long-run predictions of the model

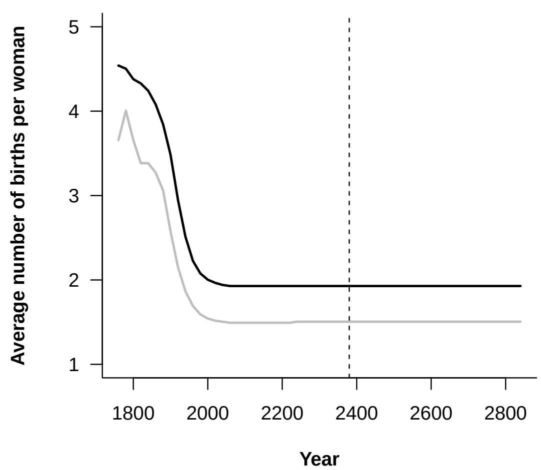
In this section, to investigate the long-run behavior of our model, we do not end our numerical exercise at the end of our observation window. Instead, we hold infant mortality constant at its final level in 1960 and run the model until the year 2840. Figure 23 illustrates the evolution of the GDP per worker in both sectors (Panel A), the share of individuals in cities and the fraction of rural workers active in industry (Panel B), the urban and rural TFR (Panel C) and the share of educated children (Panel D) from 1760 to 2840.



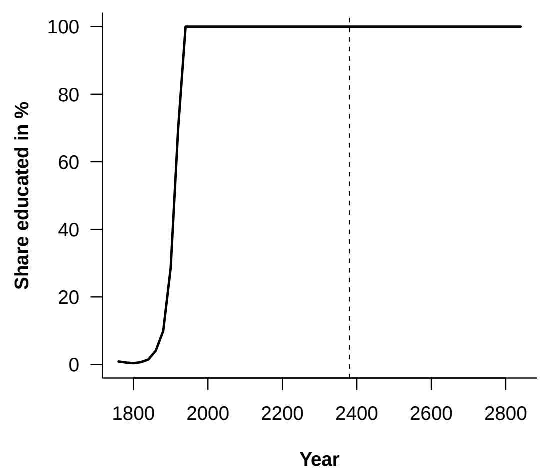
(a) GDP per worker in agriculture (black) and industries (gray)



(b) Share of individuals in cities (black) and fraction of rural workers in industries (gray)



(c) TFR in cities (gray) and the countryside (black)



(d) Share of educated offspring

Figure 23: Very long-run dynamics of the model.

Up to the year 2340 industrial and agricultural GDP per worker converge to the same growth rate of approximately 1.4% per annum, see Figure 23a. In 2360, population size becomes smaller than the threshold $\bar{N}$ and the TFP growth engine stops. At that date, we observe a kink in the dynamics of GDP per workforce. Limited to human capital accumulation as a growth engine, the new growth rate in both sectors converges to 0.5% per annum.

The share of rural workers active in industry reaches a maximum of almost 40% in the first part of the twenty-first century and converges asymptotically to zero afterwards.[59] In contrast, urbanization continues to increase and converges to a level slightly above two-thirds around the year 2200. As soon as the engine of TFP growth comes to a standstill, urbanization temporarily increases up to a peak of around 75% and moves back to around two-thirds afterwards.

In our original observation window, all children were already educated and, hence, the quantity-quality trade-off was in play. While the level of education further increases, this share remains equal to one in Figure 23d. With the increase in income, the fertility of an increasing number of dynasties converges to the area-specific steady-state levels of approximately 1.93 children per woman in the countryside and 1.50 children per woman in cities (see also Figure 21).

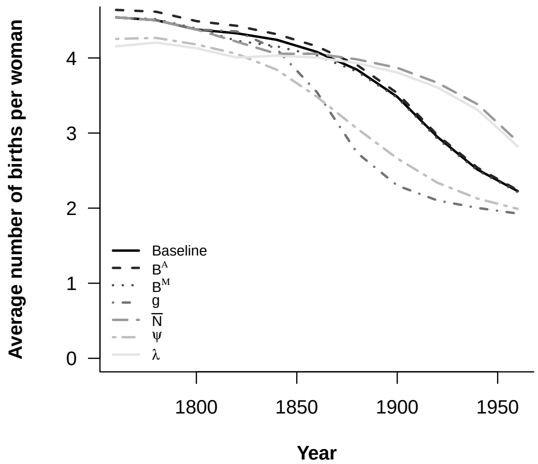
C.3 Sensitivity analysis

Overidentification checks and the fit of the matched moments speak in favor of our estimation strategy. In this appendix, we discuss whether our set of estimated parameters is located in a well-behaved region of the parameter space and whether each parameter is well identified.

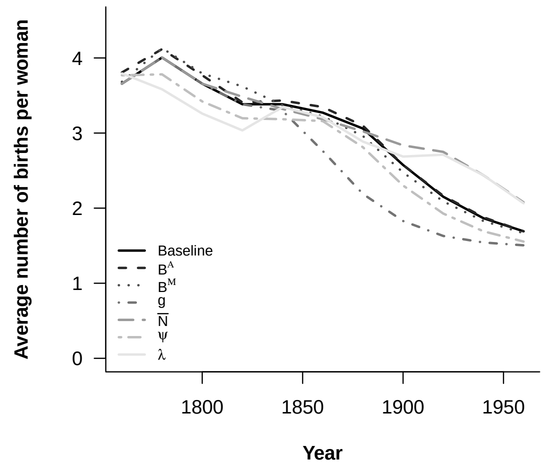
Modifying poorly identified parameters may either not change the predictions of our model at all or may alter the predictions in a way that is not aligned with the theoretical model. To investigate whether some of our parameters suffer from poor identification, we simulate the dynamics of the economy while keeping all parameters equal to their benchmark value from Section 4.4 except for the parameter of interest, which is doubled.[60] As we are interested in identification issues, we highlight only the main changes induced by a simulated shock to each parameter. Nevertheless, we need to keep in mind that our theoretical framework is a general equilibrium model. Hence, changing any parameter will alter most predicted moments through general equilibrium effects at least.

[59] This long-run decrease is due to the absence of modern industry or a service sector in rural areas in our model. This simplification was key to deriving tractable results even though we acknowledge the importance of the service sector in modern Sweden. This dimension is left to future research.

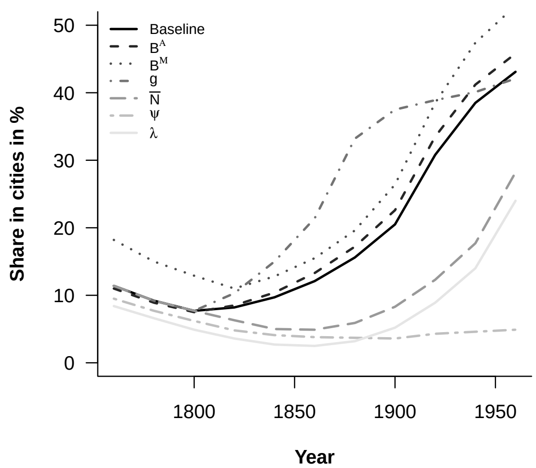
[60] To satisfy the condition $\psi \in (0, 1)$, we multiply its value by 1.25. κ and $A_0^{\mathcal{I}}$ are not discussed here, as they are discussed extensively in Section 5.



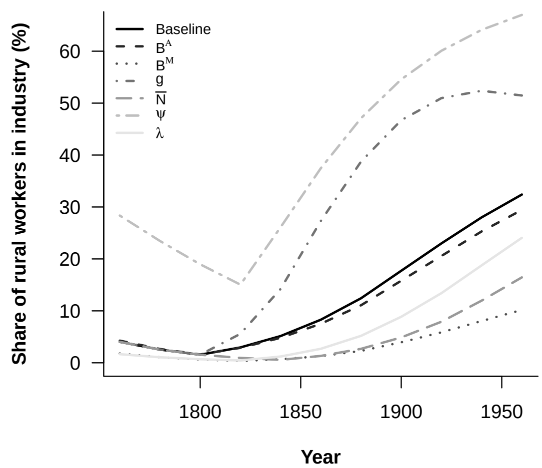
(a) Rural fertility



(b) Urban fertility



(c) Urbanization

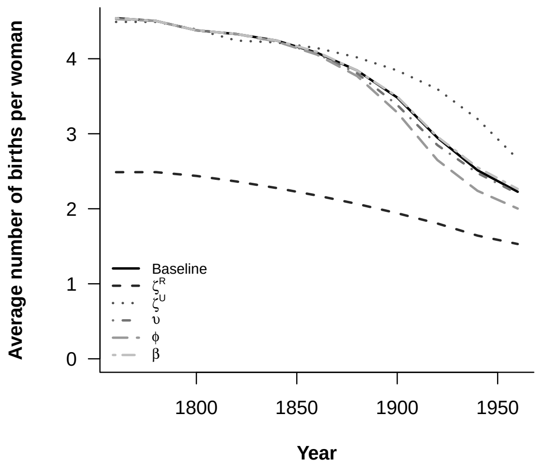


(d) Share of rural workers in industry

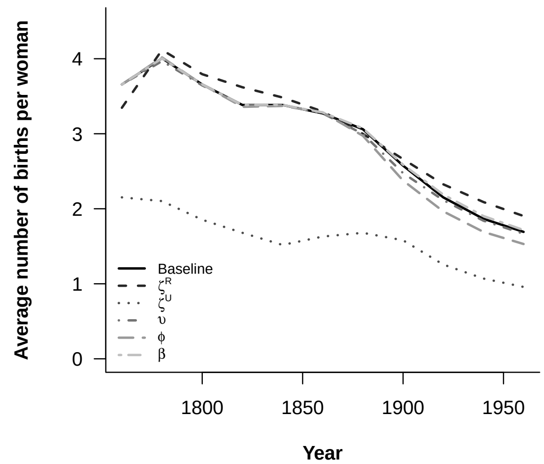
Figure 24: Sensitivity analysis on technological parameters. The values of the parameters are doubled compared to the baseline estimation. The parameter ψ is multiplied by 1.25.

The elasticity of labor input in rural industry ψ determines the allocation of labor between the industrial and agricultural sectors: the allocation as the share of industrial workers within the rural population (Fig. 24d) and as the share of individuals active in cities (Fig. 24c). Initial average ability λ determines earning capacity and, hence, also the allocation of labor as well as fertility levels. The available effective natural resources

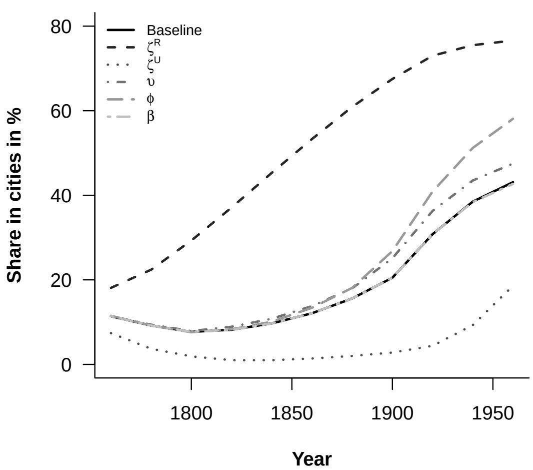
$B_0^{\mathcal{A}}X$ and the total factor productivity in modern urban industry $B_0^{\mathcal{M}}$ affect the initial income differential between cities and the countryside. The subsequent dynamics of this differential identifies the distribution of the population across areas. The income differential also identifies the initial total fertility rate in cities and in the countryside.



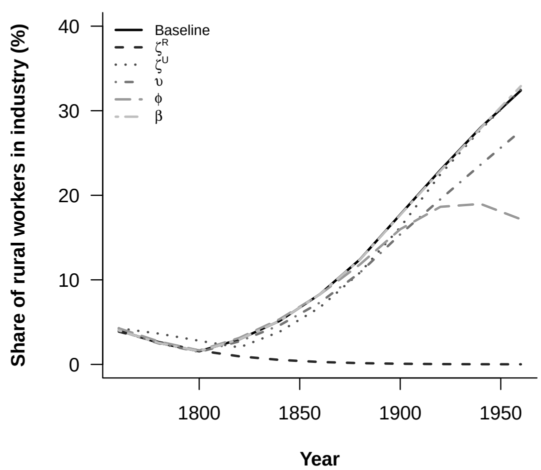
(a) Rural fertility



(b) Urban fertility



(c) Urbanization



(d) Share of rural workers in industry

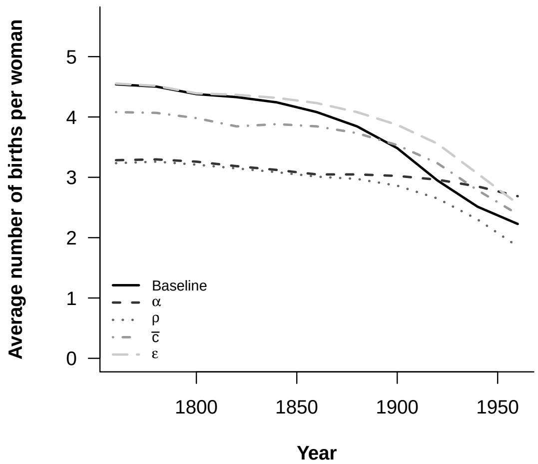
Figure 25: Sensitivity analysis of parameters for the quantity and quality of children. The values of the parameters are doubled compared to those from the baseline estimation.

$\bar{N}$ and g allow us to adjust for the population dynamics after the take-off of sustained economic growth (see Figure 24). The timing of the take-off mainly identifies $\bar{N}$, while the intensity of urbanization and of industrial labor participation in rural areas determine g.

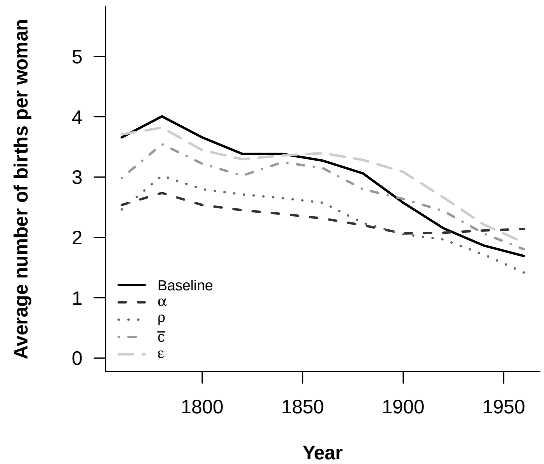
As shown in Figure 25, the fertility differential between the cities and the countryside mainly identifies the time costs of bringing up surviving children. ζ^U has a major impact on urban fertility, and ζ^R has a major impact on the number of births in rural areas. The missing children in the countryside increase the urbanization rate significantly. An increase in the cost of education β first affects the predictions for education. Hence, the cost of education has only an indirect effect on fertility during the second phase of the transition, when it allows us to adjust the TFR through the quality-quantity trade-off.

The value of the basic level of human capital υ is mainly identified by the transitional dynamics in the allocation of the population. Indeed, the higher υ is, the more the population is naturally adapted to production in the urban industrial sector. This makes the rural exodus easier at the cost of decreasing the share of rural workers engaged in industry in the countryside. Finally, the elasticity of human capital to education, ϕ, exerts a negative effect on fertility in both areas, but it is also identified thanks to the allocation of labor during the last decades.

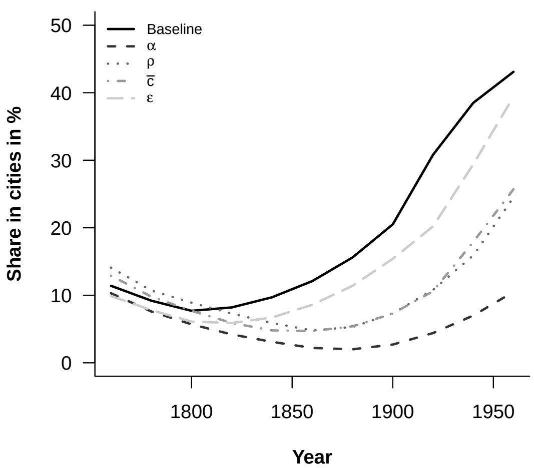
Turning to the preference parameters, Figure 26 illustrates two findings. First, multiplying α or dividing ρ by two has a strong and negative effect on initial fertility rates. Additionally, we observe a postponement of and reduction in urbanization as well as in the share of rural industrial workers. Not surprisingly, when the weight of agricultural consumption α relative to educated children or industrial goods increases, incentives to engage in the industrial sector decline. Similarly, an increase in the weight of consumption relative to educated children due to a lower ρ reduces incentives to acquire human capital and so to migrate to a city. Second, the timing of the fertility transition, urbanization and rural industrialization determine $\bar{c}$ and ε. The higher their values, the later are the transitions. One could worry about the fact that the roles of $\bar{c}$ and ε are the same, and similarly for α and ρ. Nevertheless, as soon as we divide the parameters by two instead of multiplying them, the apparent symmetries disappear (see Figures 27–30).



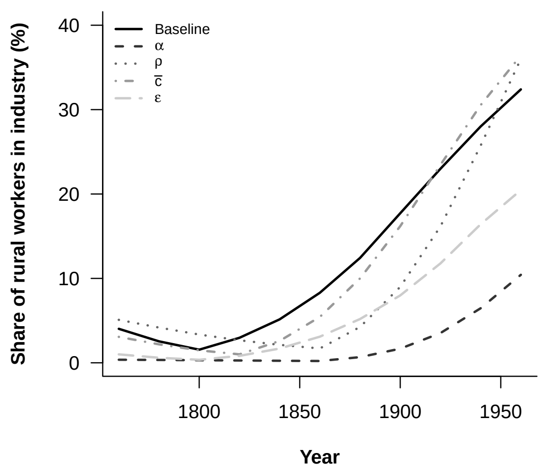
(a) Rural fertility



(b) Urban fertility



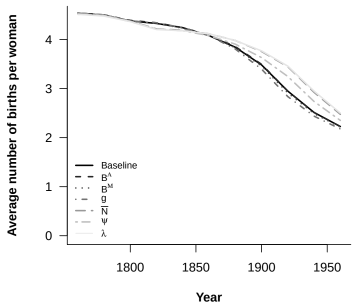
(c) Urbanization



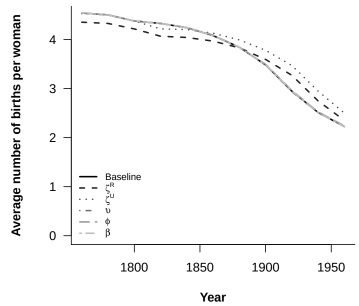
(d) Share of rural workers in industry

Figure 26: SENSITIVITY ANALYSIS OF PREFERENCE PARAMETERS. THE VALUES OF THE PARAMETERS α, ε AND $\bar{c}$ ARE DOUBLED COMPARED TO THOSE FROM THE BASELINE ESTIMATION. THE PARAMETER ρ IS DIVIDED BY 2.

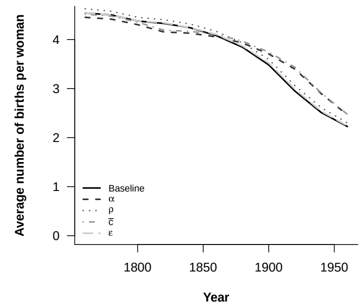
In Figures 27–30, we show how a variation of 5% in each parameter changes our predictions. It allows us to determine if our parameter set is located in a well-behaved region of our parameter space; this would not be the case if small changes in parameters imply radical changes in predictions. We observe some changes in the predicted moments; however, these changes do not alter the main predictions of our model.



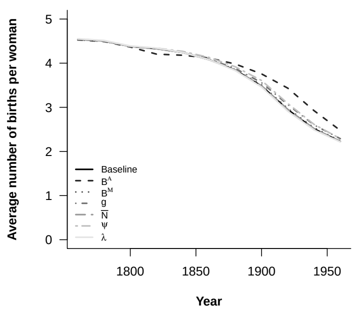
(a) Technology parameters



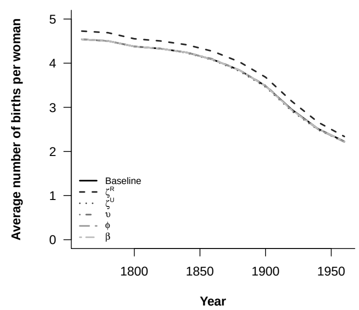
(b) Fertility parameters



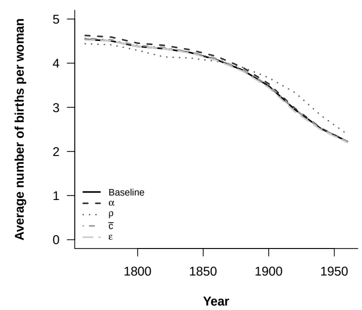
(c) Preference parameters



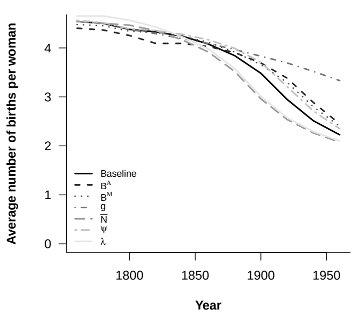
(d) Technology parameters



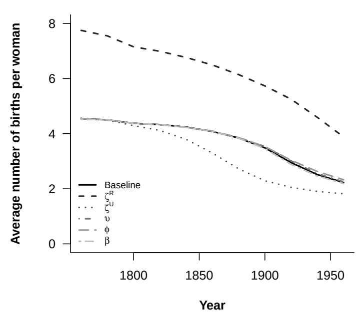
(e) Fertility parameters



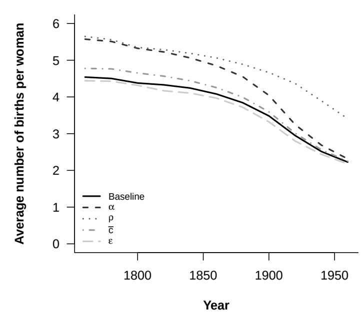
(f) Preference parameters



(g) Technology parameters

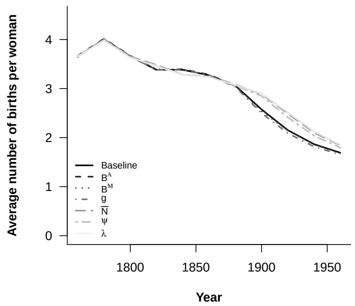


(h) Fertility parameters

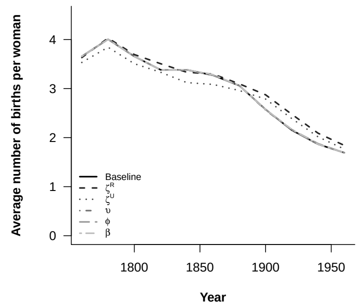


(i) Preference parameters

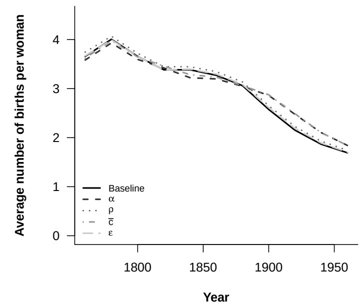
Figure 27: Sensitivity analysis and rural fertility.
The values of the parameters are increased by 5% in the first row, decreased by 5% in the second row, and divided by two in the bottom row.



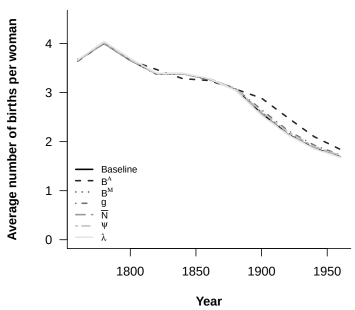
(a) Technology parameters



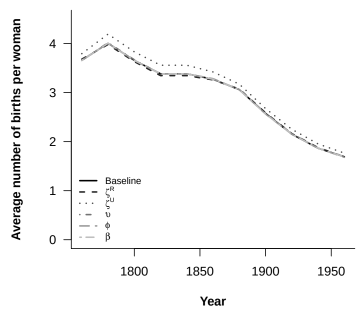
(b) Fertility parameters



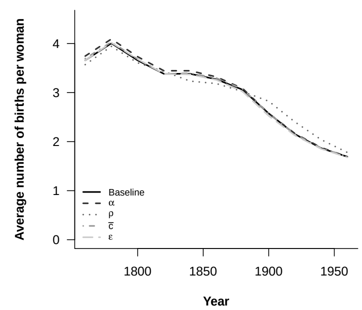
(c) Preference parameters



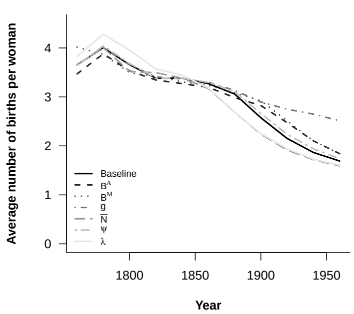
(d) Technology parameters



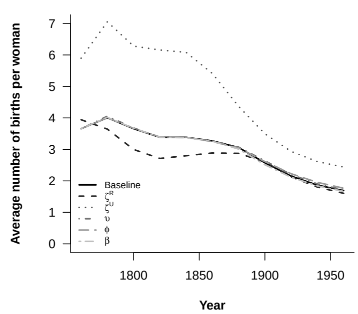
(e) Fertility parameters



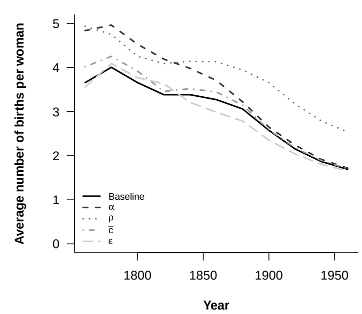
(f) Preference parameters



(g) Technology parameters

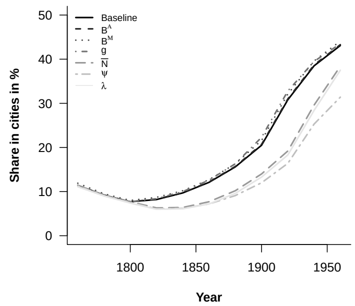


(h) Fertility parameters

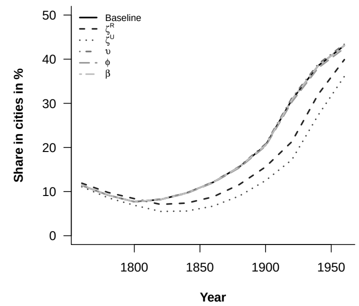


(i) Preference parameters

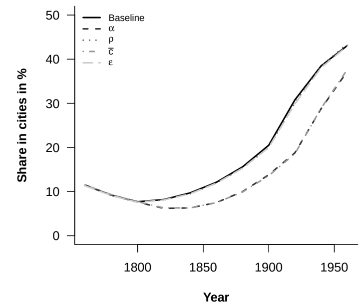
Figure 28: SENSITIVITY ANALYSIS AND URBAN FERTILITY.
The values of the parameters are increased by 5% in the first row, decreased by 5% in the second row, and divided by two in the bottom row.



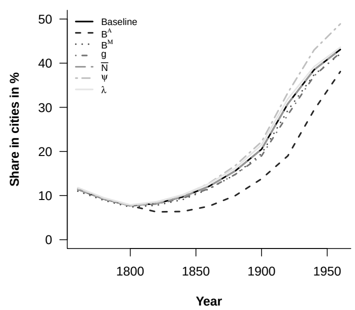
(a) Technology parameters



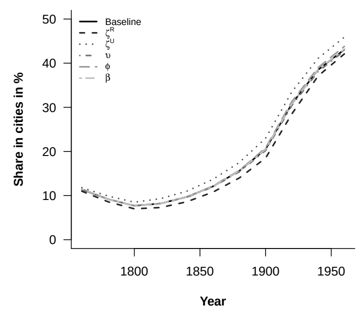
(b) Fertility parameters



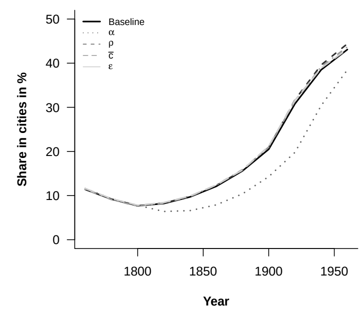
(c) Preference parameters



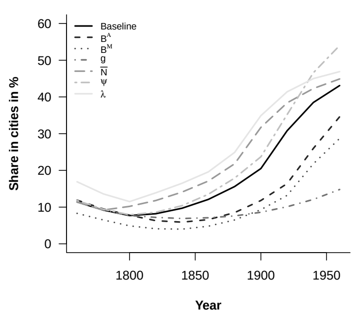
(d) Technology parameters



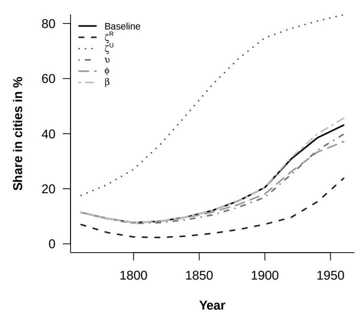
(e) Fertility parameters



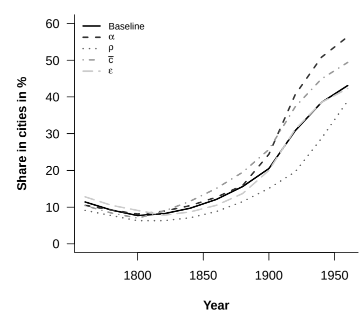
(f) Preference parameters



(g) Technology parameters

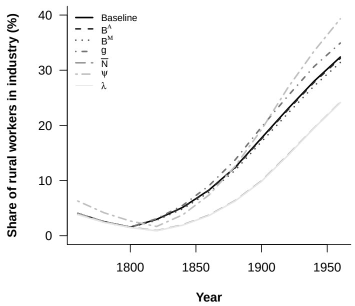


(h) Fertility parameters

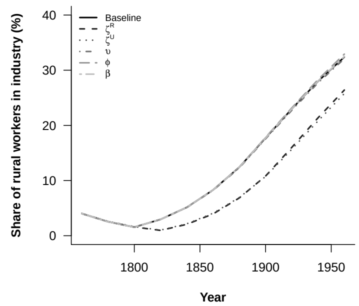


(i) Preference parameters

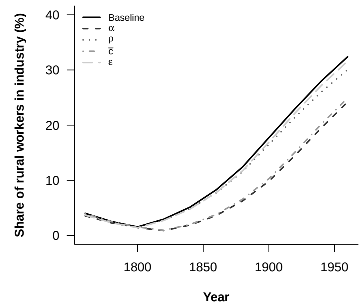
Figure 29: Sensitivity analysis and urbanization.
The values of the parameters are increased by 5% in the first row, decreased by 5% in the second row, and divided by two in the bottom row.



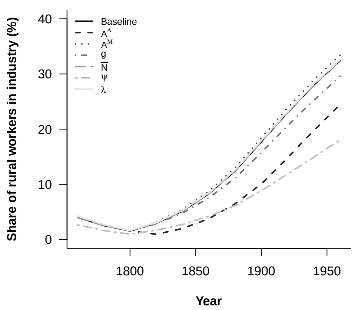
(a) Technology parameters



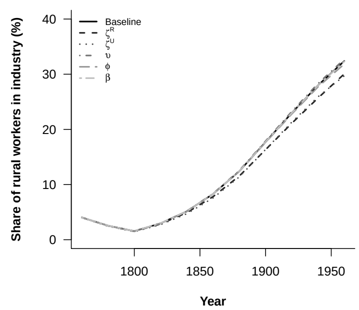
(b) Fertility parameters



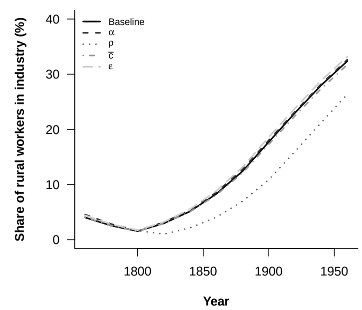
(c) Preference parameters



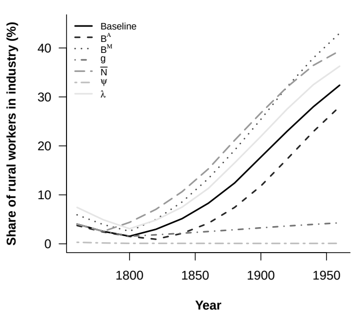
(d) Technology parameters



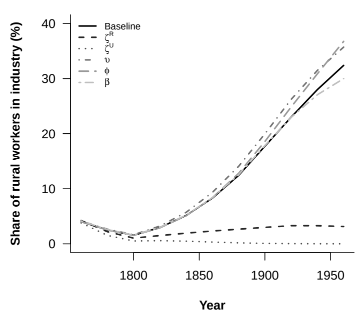
(e) Fertility parameters



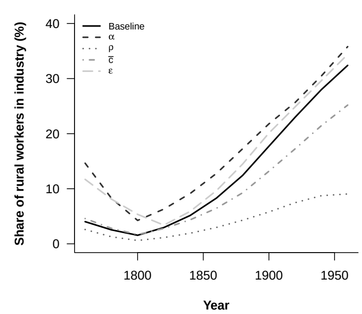
(f) Preference parameters



(g) Technology parameters



(h) Fertility parameters



(i) Preference parameters

Figure 30: Sensitivity analysis and rural industrial labor.
The values of the parameters are increased by 5% in the first row, decreased by 5% in the second row, and divided by two in the bottom row.

C.4 Internal migration

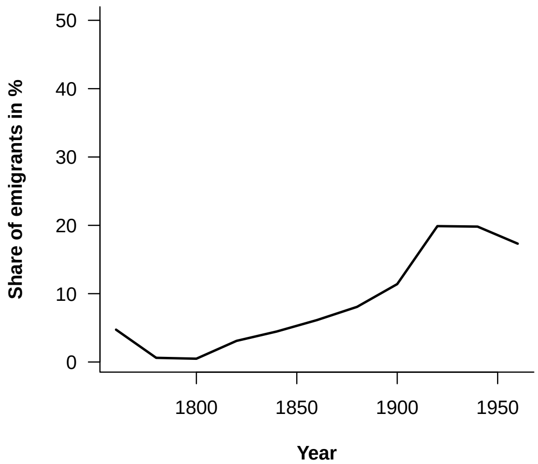
This appendix illustrates the evolution of the intensity of the rural exodus in our benchmark scenario and in the historical experiments in Section 5. Figures 31 and 32 display the emigration (δ_t) and immigration (Δ_t) rates, respectively. They are computed as follows:

$$\delta_t \equiv \frac{\text{Number of adults born in R moving to U}}{\text{Number of adults born in R moving to U + adults born in R staying in R}}$$

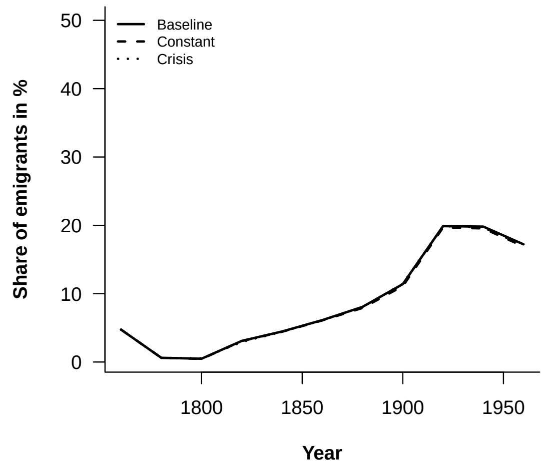
$$\Delta_t \equiv \frac{\text{Number of adults born in R moving to U}}{\text{Number of adults born in U staying in U}}$$

Figures 31a and 32a highlight our estimation of the intensity of the rural exodus during the Swedish take-off. Except for the initial reallocation, the exodus is negligible until the beginning of the nineteenth 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). Until 1840, the exodus mainly prevented deurbanization. Thereafter, it provided the modern industrial sector with the labor force it needed to enable the acceleration of growth and the Industrial Revolution. By 1920, approximately 40% of all young adults living in cities were born in the countryside, a fraction that declined to approximately one-quarter by 1960.

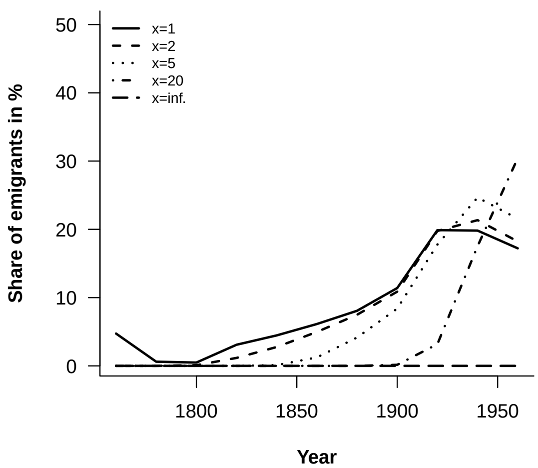
In particular, changing the migration costs and initial productivity levels in rural industry substantially alters internal migration. Reducing migration costs increases the exodus at the beginning, while higher costs lead to a postponement of internal migration. As soon as income increases sufficiently to pay κ', we observe a catch-up effect, and the exodus exceeds its initial level; see Figure 32c. The findings are similar for a higher level of initial productivity in rural industry shown in Figure 31e and 32e. Higher productivity levels increase earning capacity in the countryside and reduce the size of the exodus. However, the impact of lower productivity levels is more limited for emigration (31f) and immigration (32f). The effects of infant mortality crises are limited overall.



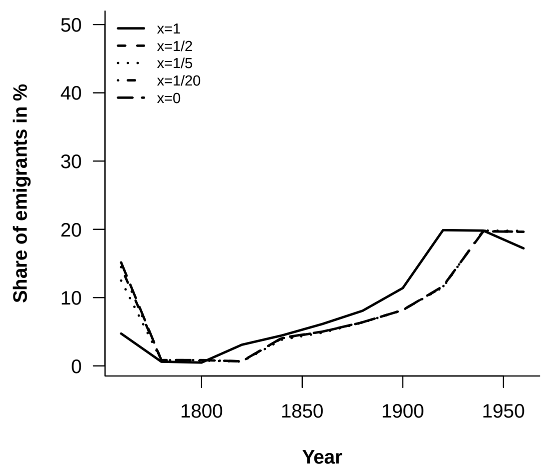
(a) Baseline



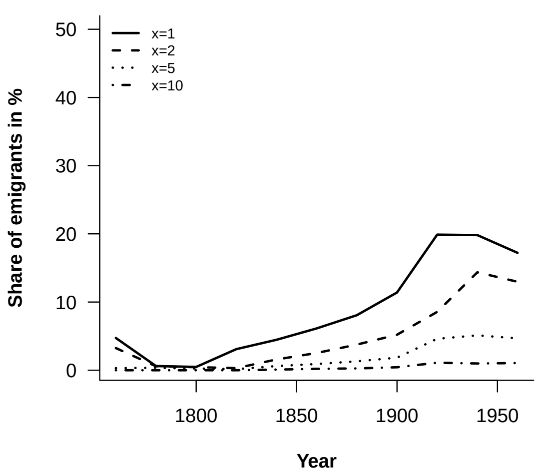
(b) Mortality shocks



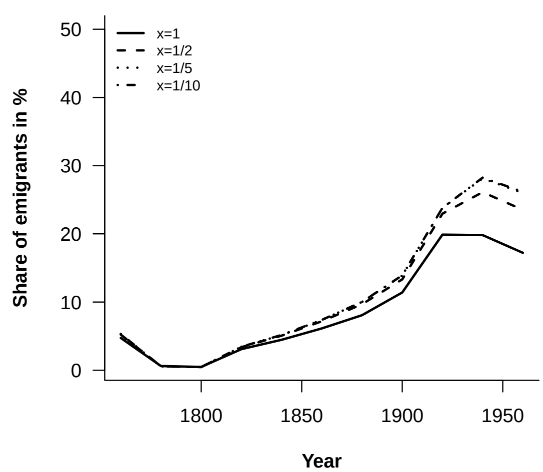
(c) Higher migration costs



(d) Lower migration costs



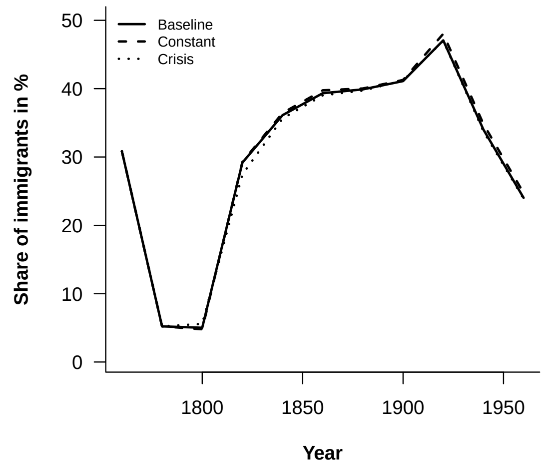
(e) Higher TFP in rural industry



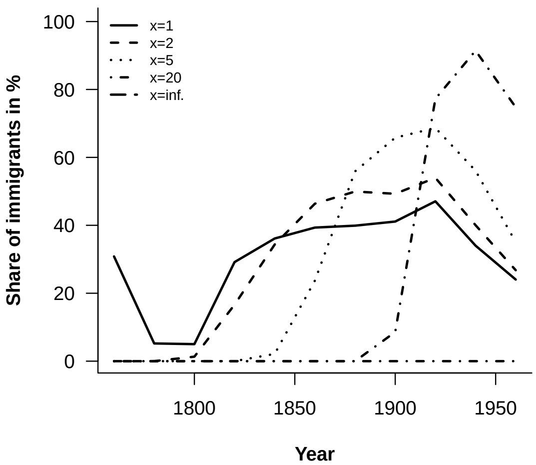
(f) Lower TFP in rural industry

Figure 31: Magnitude of emigration across different scenarios

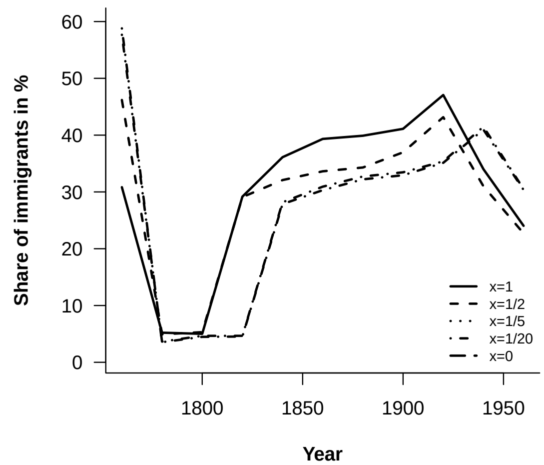
(a) Baseline



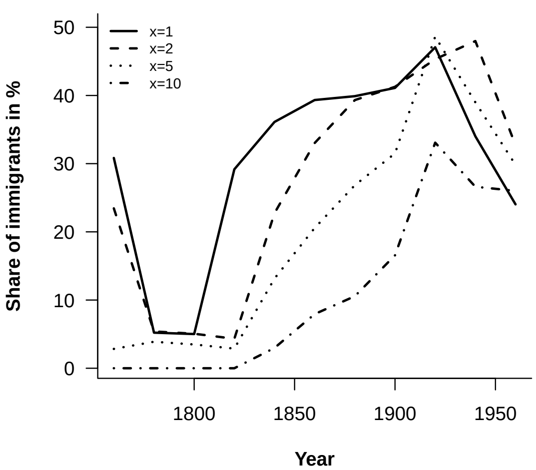
(b) Mortality shocks



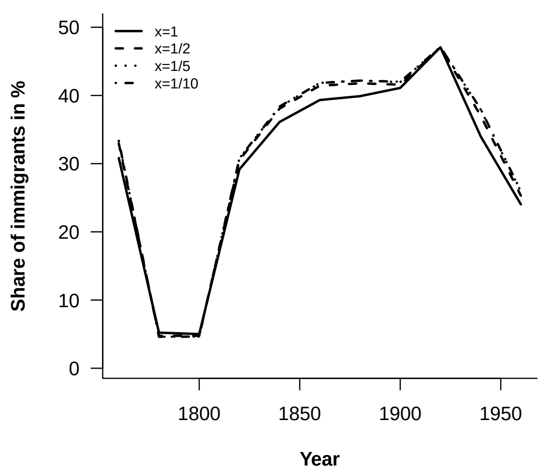
(c) Higher migration costs



(d) Lower migration costs



(e) Higher TFP in rural industry

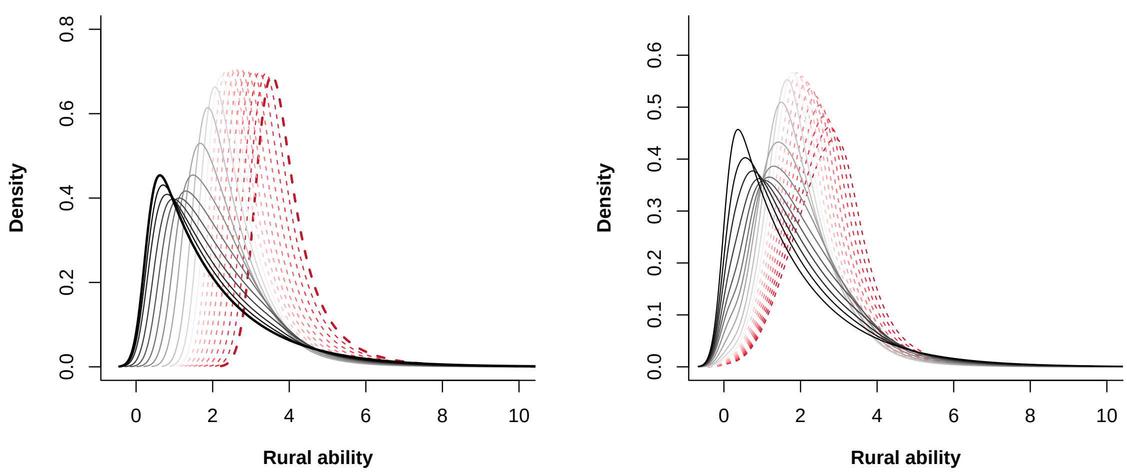


(f) Lower TFP in rural industry

Figure 32: Magnitude of immigration across different scenarios

C.5 Evolution of rural abilities and human capital

In this section, we document how the distributions of both ability to work in one of the rural sectors and human capital evolve. Figure 33 illustrates the distribution of rural abilities in the countryside and the overall economy.[61] As the dynasty-specific rural ability a^i is constant, the evolution is determined by the population dynamics only. As long as dynasties live in a Malthusian regime, higher rural abilities are linked to more offspring in the rural area what makes the distribution in Figure 33a shifting to the right. Switching to the interior regime due to gains in income, more and more dynasties in the rural area substitute fertility by education. Dynasties with a higher a^i have less children, which shifts the distribution to the left. Still, this effect is potentially counterbalanced by migration: those dynasties with lowest rural abilities leave the countryside and move to the cities, what makes the density shift to the right. As can be seen in Figure 33a, the right shift of the density dominates even after 1960 (dashed years). On top of this, the critical level of rural abilities required to live in rural places increase all along our dynamics as can be seen on Figure 35a.



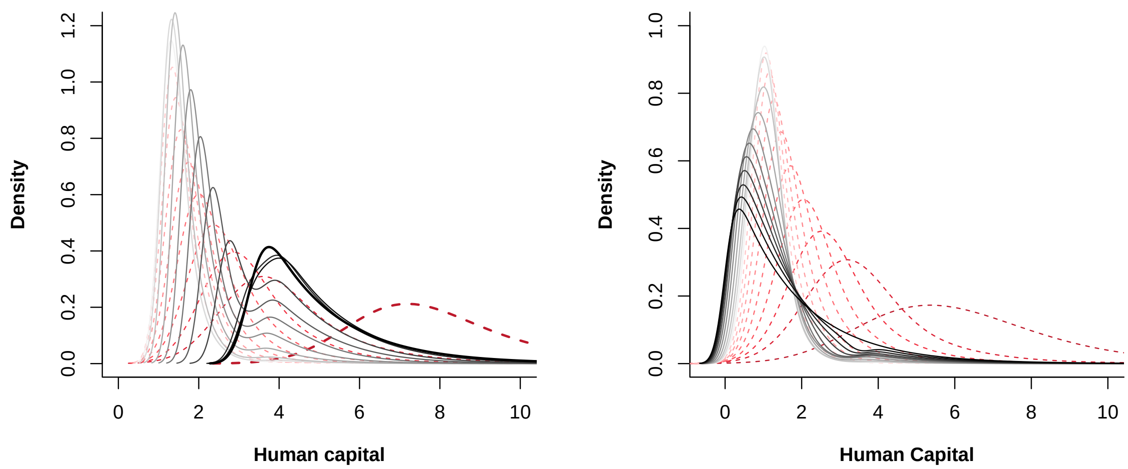
(a) Countryside

(b) Overall population

Figure 33: EVOLUTION OF THE DENSITY OF AGRICULTURAL ABILITIES FROM 1760 (SOLID BLACK) TO 2140 (DASHED RED).

[61]Figures show the kernel density estimation of the 5,000 dynasties weighted by their sizes at each point in time. From black to light gray, lines cover the period 1760-1960, every 20 years: the darker the lines the sooner the periods. Red dashed lines cover the period 1980-2140, every 20 years. The lighter the color the sooner the period.

Turning to the whole country, we also observe the shift of rural abilities to the right. This movement is driven by two main mechanisms: the right shifting to the right happening in the countryside (see previous paragraphs) and selection into migration. Rural migrants arrive in cities with relatively low levels of human capital. Once these dynasties live in interior regimes, they tend to have more children than more ancient urban dynasties what gives them an evolutionary advantage. Their weight in the urban population increases with time while they have higher rural skills.

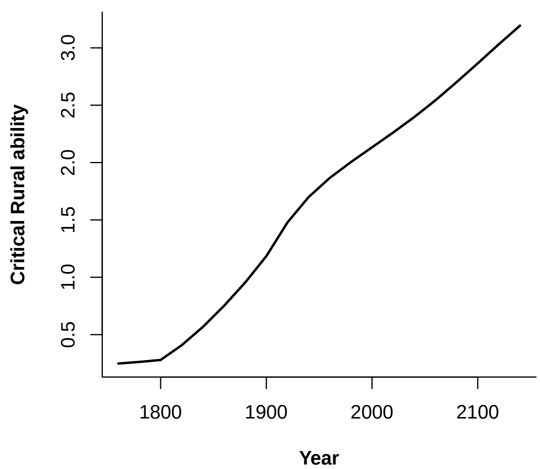


(a) Cities

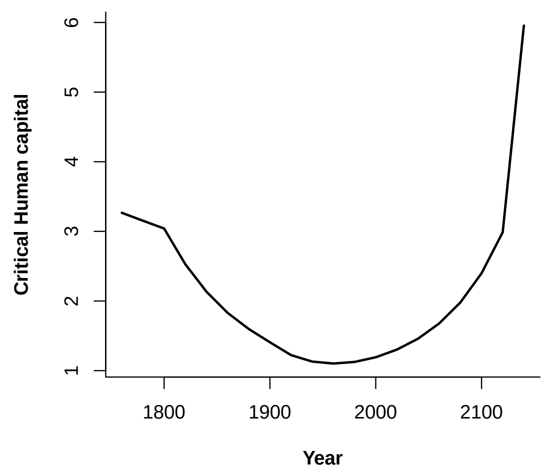
(b) Overall population

Figure 34: Evolution of the density of human capital from 1760 (solid black) to 2140 (dashed red).

Selection into migration plays a key role for the evolution of the distribution of human capital h_t^i both in cities and at the country level (Figure 34). Initially, high levels of human capital are concentrated in the hands of a small urban elite and slightly shifts to the right in 1780 and 1800. However, with the expansion of the rural exodus afterwards, the selection into migration mechanism starts to operate: human capital of immigrants reduces the average level of human capital in cities what pushes the distribution to the left in Figure 34a. This de-skilling of urban areas slows down in the twentieth century when the human capital of migrants stepwise increases. Figure 35b documents the U-shape pattern of the critical human capital with a minimum in the twentieth century. Since the middle of the twentieth century, the distribution of human capital shifts rightwards and becomes wider. This movement corresponds to the convergence toward the BGP where the accumulation of human capital is the unique growth engine while urbanization rates stabilize, what gives less importance to the mechanism of selection into migration.



(a) Threshold on rural abilities to live in rural areas



(b) Threshold on human capital to live in urban areas

Figure 35: Evolution of the critical value of abilities.

C.6 Mortality and net fertility

Coale's index relates the number of births in 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 approximate the number of surviving children (TFR_{net}) from Coale's index according to:

$$\hat{n}_t^{\text{net}j} = 12.44 * q_t^j \frac{\sum_{i \in S_t^j} n_t^i \ell_t^i}{N_t^j \frac{5}{2} \sum_o n_o^H}$$

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

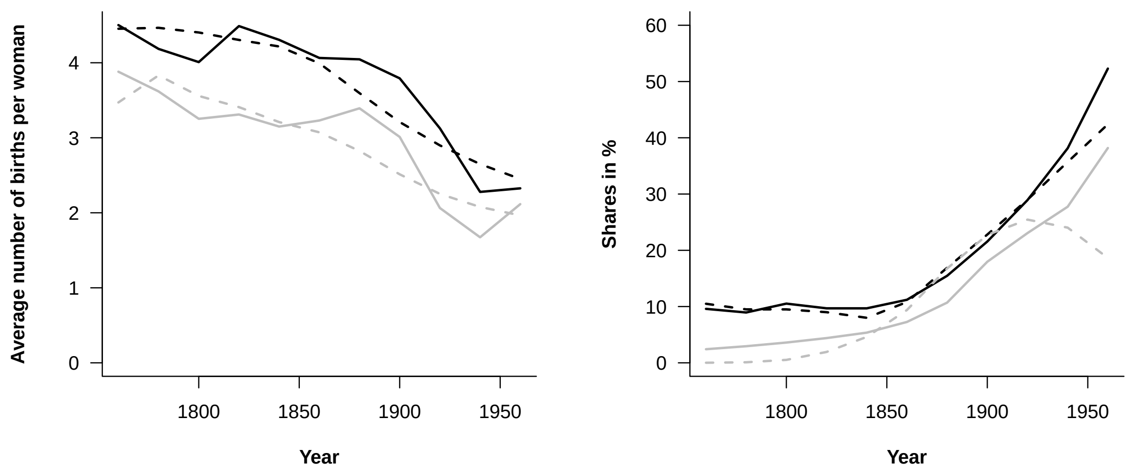
Table 9: Gross and net fertility in the benchmark, no improvement in the IMR and a mortality crisis scenario.

	Benchmark estimation				No improvements in IMRs				Crisis in IMRs in 1800			
	Rural		Urban		Rural		Urban		Rural		Urban	
	TFR_{gr}	TFR_{net}	TFR_{gr}	TFR_{net}	TFR_{gr}	TFR_{net}	TFR_{gr}	TFR_{net}	TFR_{gr}	TFR_{net}	TFR_{gr}	TFR_{net}
1760	4.541	3.617	3.657	2.501	4.541	3.617	3.657	2.501	4.541	3.617	3.657	2.501
1780	4.503	3.575	4.006	2.768	4.491	3.577	4.043	2.764	4.503	3.575	4.006	2.768
1800	4.379	3.535	3.657	2.729	4.429	3.527	3.956	2.705	4.379	3.535	3.657	2.729
1820	4.329	3.584	3.384	2.590	4.478	3.567	3.732	2.552	4.329	3.584	6.792	2.264
1840	4.242	3.595	3.384	2.623	4.466	3.557	3.769	2.577	4.242	3.595	3.321	2.575
1860	4.080	3.536	3.272	2.587	4.391	3.498	3.720	2.543	4.093	3.546	3.284	2.597
1880	3.844	3.417	3.060	2.539	4.230	3.369	3.620	2.475	3.856	3.428	3.073	2.549
1900	3.483	3.181	2.575	2.280	3.919	3.121	3.210	2.195	3.496	3.192	2.575	2.280
1920	2.948	2.761	2.152	2.008	3.384	2.695	2.824	1.931	2.948	2.761	2.165	2.020
1940	2.513	2.415	1.866	1.805	2.948	2.348	2.538	1.735	2.513	2.415	1.866	1.805
1960	2.227	2.198	1.692	1.671	2.675	2.130	2.326	1.591	2.227	2.198	1.693	1.671

D Alternative Frameworks

D.1 No variation in rural abilities

In our main framework, the initial distribution of rural skills is a permutation of the distribution of initial human capital h_0. This appendix drops this assumption and instead, assumes that all dynasties have the same rural abilities $a_0^i = \bar{a}$. To avoid an extension of the parameter space, we set the rural abilities to the predicted average human capital in 1760 $\bar{a} = \lambda$. The remaining theoretical framework and the calibration strategy are exactly the same as in the main paper.

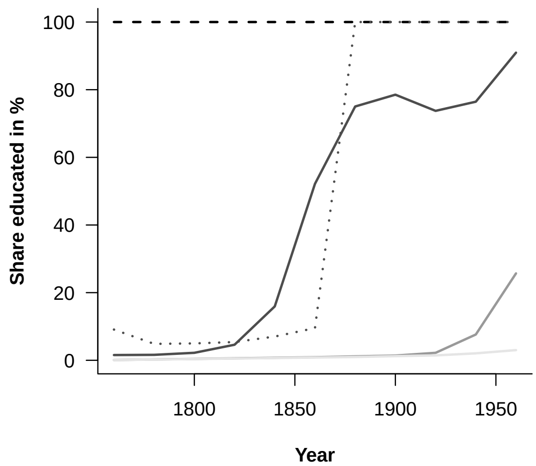


(a) Urban (gray) and rural (black) fertility

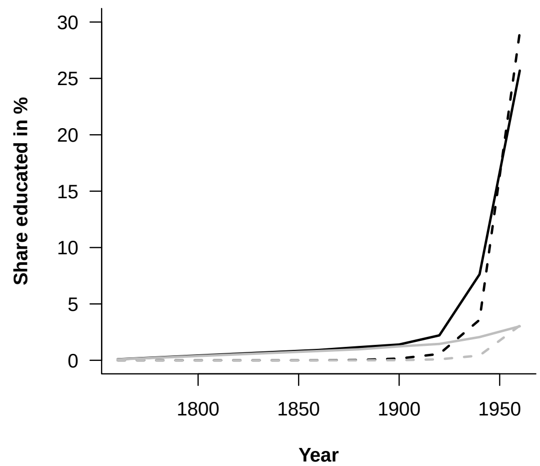
(b) Share of rural workers in industry (gray) and share of population in cities (black)

Figure 36: Observed (solid) and simulated (dashed) demographic dynamics and allocation of labor in Sweden without heterogeneity in rural abilities, 1760–1960.

Fig. 36a depicts predicted and estimated fertility trends. Our model predicts the fertility differential and the fertility transition in a convincing way – even without any kind of heterogeneity in rural abilities. Fig. 36b documents the rural industrialization and urbanization that we also find in this alternative framework. However, without the heterogeneity in rural abilities, the share of rural workers linked to industries already starts to decline in the middle of the twentieth century which reduces the overall fit in quantitative terms.



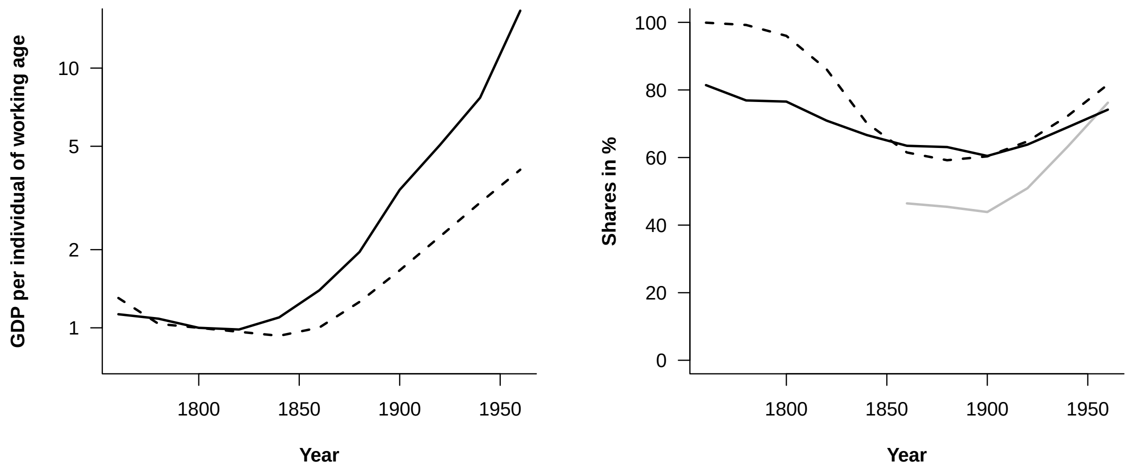
(a) Observed (solid) and predicted shares without (dashed) and with educational thresholds $\bar{e}^P$ (dotted), enrolled into primary (black), integrated secondary (medium gray) and tertiary (light gray) education.



(b) Integrated secondary (black) and tertiary (gray) education with educational thresholds $\bar{e}^S$ and $\bar{e}^T$, respectively.

Figure 37: Observed enrollment rates into education (solid) and simulated shares of educated agents (dashed).

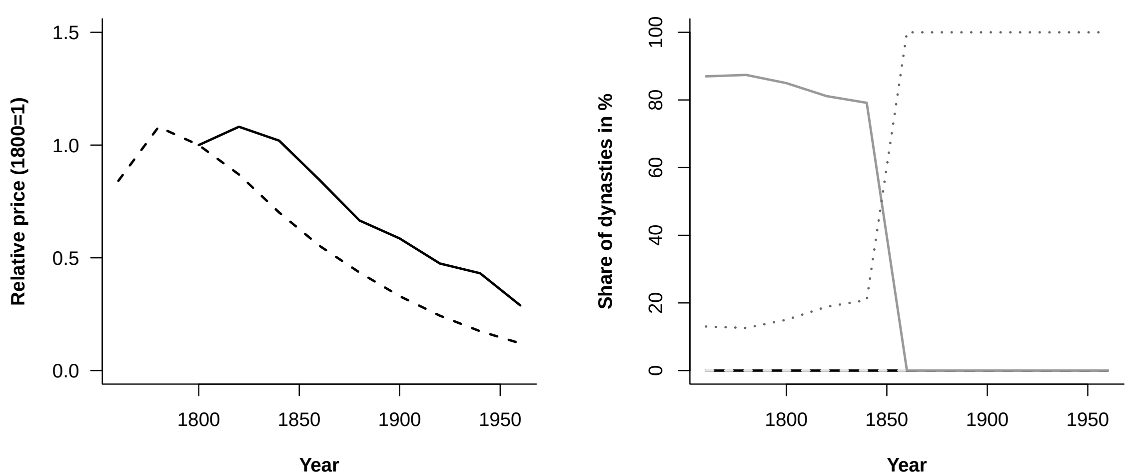
The predicted and observed shares in education without heterogeneity in rural abilities, displayed in Fig. 37, elicit the reason for the decline in the share of rural workers and a major difference to our baseline framework: the absence of heterogeneity in rural productivities prevents any economic and demographic heterogeneity in the countryside; said differently, all rural households adopt the same behavior once they decide to stay in the countryside. It implies that, in the present specification, as soon as from the first period, all dynasties educate their offspring. All dynasties in the countryside are in Regime 5, while dynasties settling in cities already enjoy the interior regime. In order to improve the situation, we define a minimum educational investment ($e^P = 0.515$) as critical threshold for primary education; it enables to fit the timing of the transition to mass education much better. Still, the dynamics of education remains very "rugged" as all dynasties in the countryside would start to educate exactly at the same moment. The fact that dynasties in the countryside switch between the regimes together also explains why the share of rural labor devoted to industries already starts to decline in 1920 (Fig. 36b): because the entire population starts to accumulate human capital earlier, the relative productivity of rural industries compared to urban ones tend to decline sooner. We illustrate the dynamics of regime changes in Fig. 39b.



(a) GDP per individual of working age

(b) Share of industrial workers in cities

Figure 38: Simulated (dashed) and observed (solid) total GDP per adult of working age and share of industrial workers located in urban areas: raw data (gray), adjusted data (solid black), and model predictions (dashed black).



(a) Observed (solid) and simulated (dashed) evolution of the relative industrial price

(b) Share of dynasties in regime 1–3 (dashed black), regime 4 (dashed light gray), regime 5 (solid medium gray) and regime 6 (dotted dark gray)

Figure 39: Evolution of relative prices and the share of dynasties by Regimes of Tab. 1

Our additional over-identification checks emphasizes that the model with identical rural abilities also performs reasonable along the relevant dimensions we do not try to fit. GDP per individual of working age normalized to one in 1800 initially stagnates before the take-off in the middle of the nineteenth century. However, as the predicted GDP per worker stagnates for an additional period, we under-predict the growth rate in GDP per worker, see Fig. 38a. Fig. 38b shows that the model predicts the expected U-shaped pattern in the share of industrial workers in cities even if the predicted magnitude of U-shape pattern exceeds the observed one.

Relative prices normalized to 1 in 1800 in Fig. 39a are rather similar to our baseline predictions. The main difference occurs in the eighteenth century. Opposed to the declining relative price at the beginning in the baseline predictions, the relative price now initially increases.

D.2 Differential rural-urban preference on children's human capital

Only the manufacturing sector in cities uses human capital as input, implying that human capital does not matter for production in rural industries and agriculture. This motivates another alternative framework where we assume that individuals living in urban and rural areas value human capital of their offspring differently. We limit the presentation of the theoretical considerations to the individual perspective, as the remaining model is not altered in this alternative framework.

D.2.1 Individual's maximization problem

In symmetry to the baseline framework, utility of the representative agent of a given dynasty is still denoted W_t and holds almost unchanged except that we now introduce a parameter η^j, which denotes the weight attached to the human capital of children. Parents still have "warm glow preferences" as they care about the future human capital h_{t+1} of their children. We assume that people living in cities $(j = U)$ value the human capital of their children more than those living in the countryside $(j = R)$: $\eta^U > \eta^R = 1$. This is the main difference in this alternative framework.

Accordingly, we assume that when $\omega_t \leq \bar{c}$, $W_t \to -\infty$ by definition and $c_t = \omega_t$, $n_t =$

$e_t = d_t = 0$. In contrast, when $\omega_t > \bar{c}$, W_t, we adjust the utility in the following way:

$$W_t^j = \alpha \ln (c_t - \bar{c}) + (1-\alpha) \ln (d_t + \varepsilon) + \ln \left(n_t q_t^j \right) + \eta^j \ln h_{t+1}. \tag{34}$$

As the budget constraint (Equation 14) and human capital accumulation (Equation 17) remain unaffected, we directly turn our attention to the individual decision. Table 10 summarizes the six adjusted potential regimes adults can live in.

Table 10: Alternative regimes in which adults can live

Regime	c_t^j	n_t^j	d_t^j	e_t^j
R1	ω_t^j	0	0	0
R2	$\frac{\alpha \omega_t^j + \bar{c}}{1+\alpha}$	$\frac{\omega_t^j - \bar{c}}{(1+\alpha)\left(\xi + \zeta^j q_t^j\right)\omega_t^j}$	0	0
R3	$\frac{\alpha(\omega_t^j - \mathbb{1}_\kappa \kappa p_t) + \bar{c}}{1+\alpha}$	$\frac{\tilde{w}_t^j}{(1+\alpha)\left(\xi + \zeta^j q_t^j\right)\omega_t^j}$	0	0
R4	$\bar{c} + \frac{\alpha}{2}(\tilde{w}_t^j + \varepsilon p_t)$	$\frac{\tilde{w}_t^j + \varepsilon p_t}{2\left(\xi + \zeta^j q_t^j\right) w_t^j}$	$\frac{1-\alpha}{2} \frac{\tilde{w}_t^j}{p_t} - \frac{1+\alpha}{2}\varepsilon$	0
R5	$\bar{c} + \frac{\alpha \tilde{w}_t^j}{1+\alpha}$	$\frac{1-\eta^j \phi}{1+\alpha} \frac{\tilde{w}_t^j}{\pi_t}$	0	$-\upsilon + \frac{\phi \eta^j}{1-\phi \eta^j} \frac{\pi_t}{\beta q_t^j p_t}$
R6	$\bar{c} + \frac{\alpha}{2}\left(\tilde{w}_t^j + \varepsilon p_t\right)$	$\frac{1-\phi \eta^j}{2} \frac{\tilde{w}_t^j + \varepsilon p_t}{\pi_t}$	$\frac{1-\alpha}{2} \frac{\tilde{w}_t^j}{p_t} - \frac{1+\alpha}{2}\varepsilon$	$-\upsilon + \frac{\phi \eta^j}{1-\phi \eta^j} \frac{\pi_t}{\beta q_t^j p_t}$

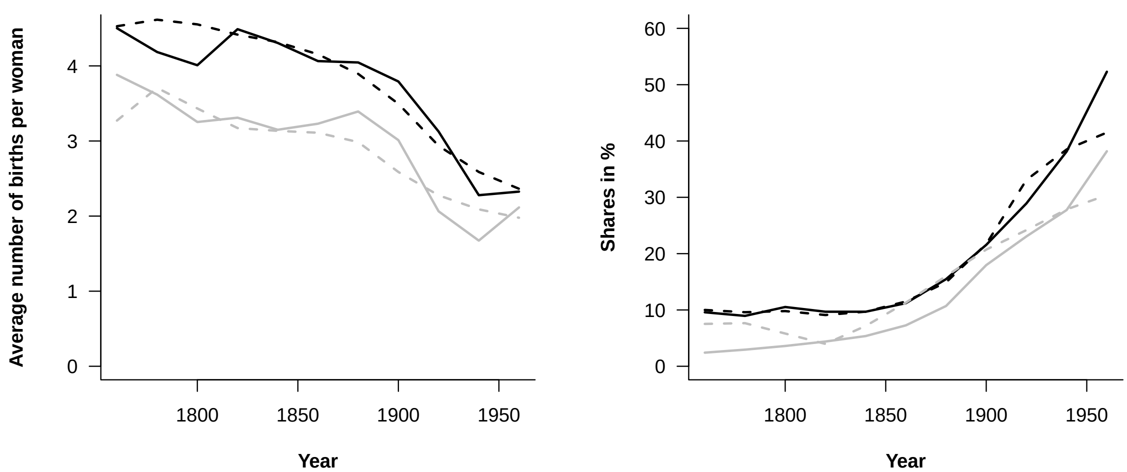
with $\tilde{\omega}_t^j = \omega_t^j - \mathbb{1}_\kappa \kappa p_t - \bar{c}$ and $\pi_t = \left(\xi + \zeta^j q_t^j\right)\omega_t^j - \beta p_t q_t^j \upsilon$

Provided that $\rho = 1$, regimes R1 to R4 are unaffected by the adjusted preferences. The higher valuation of children's human capital of parents living in urban areas magnifies their quantity-quality trade-off. A higher η^j reduces fertility in regime R5 and R6 and increases educational investments. Additionally, the introduction of η^j alters critical prices $p^*(\mathbb{1}_\kappa \kappa)$ and $\tilde{p}(\mathbb{1}_\kappa \kappa)$ as well as the thresholds $\bar{\upsilon}_t(\mathbb{1}_\kappa \kappa)$ and $\bar{\bar{\upsilon}}_t(\mathbb{1}_\kappa \kappa)$. Still, the characterization of individual behavior for any configuration of our set of parameters is symmetrical to Appendix B.1. Thus, we can determine the regime in which each person lives conditional on the area she decides to live in, given the system of relative price p_t, potential incomes $\{\omega_t^U, \omega_t^R\}$ and mortality rates $\{q_t^U, q_t^R\}$.

The description of the location choices and the general equilibrium perfectly corresponds to the baseline model. The general calibration strategy is also the same in this modified framework, with one exception: instead of estimating ρ, we estimate the era specific valuation of human capital η^j.

D.2.2 Calibration results

If we better predict the share of individuals living in cities (Fig. 40b), it is the reverse for both the share of rural workers in industries (Fig. 40b) and fertility (Fig. 40a). In detail, the average distance between predicted and observed moments declines from 12.7% to 8.6% for what concerns urbanization rates while it increases from 3.1% to 5.1% for the share of rural workers in industries, 10.5% to 11.4% for fertility in urban areas and 6.1% to 7.4% for fertility in rural areas.

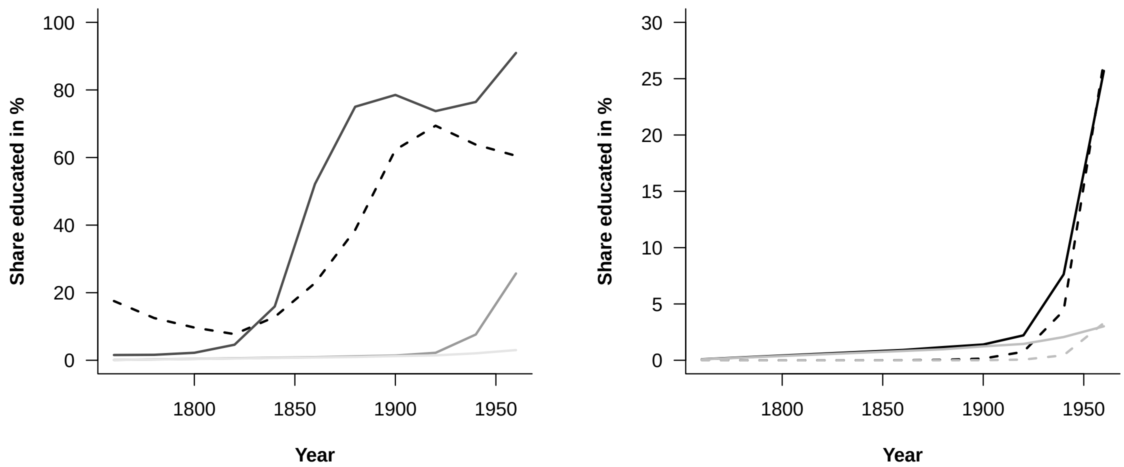


(a) Urban (gray) and rural (black) fertility

(b) Share of rural workers in industry (gray) and share of population in cities (black)

Figure 40: Observed (solid) and simulated (dashed) demographic dynamics and allocation of labor in Sweden with alternative preference on child quality, 1760–1960.

This being said, we capture better the timing of the transition to mass primary education, see Figure 41. Nevertheless, the share of educated children is higher before and lower after the transition compared to the data. Introducing critical levels of educational investments required to reach secondary ($e^S = 3.743$) and tertiary education ($e^T = 6.606$) leads to a similar fit as in the baseline model.

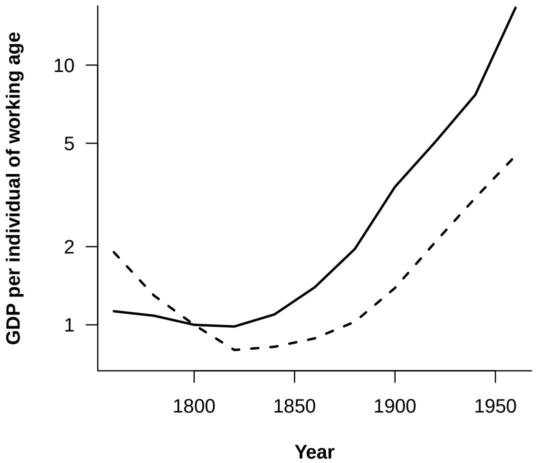


(a) Predicted share educated (black) and observed enrollment rates into primary (dark gray), integrated secondary (medium gray) and tertiary (light gray) education.

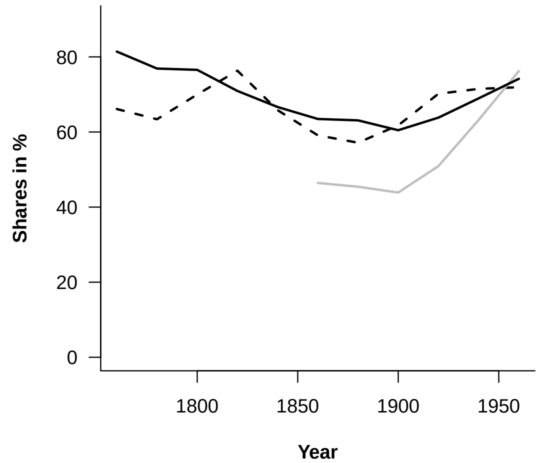
(b) Integrated secondary (black) and tertiary (gray) education with educational thresholds $\bar{e}^S$ and $\bar{e}^T$, respectively.

Figure 41: Observed enrollment rates into education (solid) and simulated shares of educated agents (dashed).

Other important differences show up in the overidentification checks. We document a relatively strong decline in the GDP per individual in working age at the beginning; reflecting the absence of TFP growth in combination with lower (non-normalized) relative prices of industrial goods, see 42a and 43a. Still, in more recent times, when TFP growth enters, growth rates of GDP per capita are predicted well. The U-shape in the share of industrial workers in cities is less clearly predicted in this alternative framework compared to the benchmark. As in the benchmark, we document a permanent decrease in (normalized) prices of the manufactured goods.

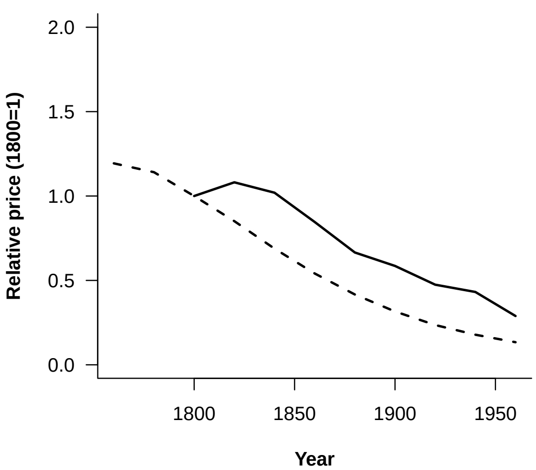


(a) GDP per individual of working age

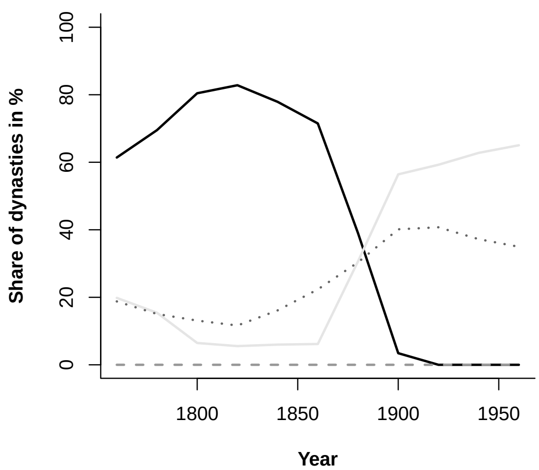


(b) Share of industrial workers in cities

Figure 42: Simulated (dashed) and observed (solid) total GDP per adult of working age and share of industrial workers located in urban areas: raw data (gray), adjusted data (solid black), and model predictions (dashed black).



(a) Observed (solid) and simulated (dashed) evolution of the relative industrial price



(b) Share of dynasties in regime 1–3 (solid black), regime 4 (solid light gray), regime 5 (dashed medium gray) and regime 6 (dotted dark gray)

Figure 43: Evolution of relative prices and the share of dynasties by Regimes of Tab. 1

Finally, Fig. 43b illustrates the dynamics in the allocation of dynasties across the regimes. At the beginning, the vast majority of dynasties, typically settling in the countryside, is in the Malthusian regimes 1–3. Remaining households allocate on the intermediate regime 4 (no education but consumption of manufactured goods) and the interior regime 6. Logically, the decline of the average GDP per capita until 1840 is accompanied by an enlargement in poverty in the sense that the share of dynasties living in Malthusian type regimes increases until that date. As soon as economic growth sets in around the second part of the nineteenth century, we observe a strong shift in the allocation of dynasties. The Malthusian regimes disappear and the dynasties are either in regime 4 or 6 by the end of the observation window.