

Rural exodus and fertility at the time of industrialization*

Thomas Baudin[†] and Robert Stelter[‡]

September 21, 2016

Abstract

We propose a unified model of economic growth where people decide not only about fertility and education but also whether or not to migrate to the city or alternatively to the countryside. Using the simulated method of moments and an original set of Danish data, we calibrate our model to evaluate the relative contributions of the rural exodus and mortality reductions to economic growth and the fertility transition. We find, in line with Galor (2005), that the reduction of infant mortality has not been a major driver of the Danish economic take-off while the rural exodus, complementing technological progress, has been a workhorse of the economic and demographic revolution of Denmark.

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

JEL Classification Numbers: J11, J13, O41.

*We would like to thank Annika Westberg for helping us with SHIPs data and Hans Chr. Johansen for sending us data on Denmark. We are deeply grateful for the computational infrastructure provided by Max Planck Computing and Data Facility and the assistance offered by Markus Rampp. The paper greatly benefited from discussions with Jim Oeppen, David de la Croix, Nico Voigtlaender, Carl Mosk, Philippe Bocquier and Fabian Kindermann. We also thank participants to seminars and conferences in Louvain la Neuve, Paris and Hambourg for their helpful comments. The traditional disclaimer applies.

[†]Lille Economie Management, UMR 9221, Université de Lille and IRES, Université catholique de Louvain, E-Mail: thomas.baudin@uclouvain.be.

[‡]Max Planck Institute for Demographic Research, IRES, Université catholique de Louvain and University of Rostock, E-Mail: stelter@demogr.mpg.de.

1 Introduction

In line with the tradition instigated by Galor and Weil (2000), we propose a unified theory capturing the interactions between urbanization, the demographic transition, and the take-off to modern economic growth. We then use our theory to evaluate, in a quantitative way, the respective role of mortality decline and rural exodus for the Danish economic and demographic revolutions.

The Unified Growth Theory (UGT) provides a consistent framework explaining how economies of Western Europe switched from Malthusian stagnation to modern growth and how this take-off relates to the demographic transition. Starting with the seminal contribution of Galor and Weil (2000), the UGT posits, generally speaking, that technological progress has to take place to incite parents to invest in the human capital of their children and to reduce the number of children they have. Thus, the economic take-off is linked to massification of education and the transition from high to low fertility rates. A recent literature has emerged to show that some necessary conditions should be met to allow the general mechanism of the UGT to take place. For instance, Squicciarini and Voigtländer (2014) show that a rising total factor productivity is not necessarily sufficient to induce human capital accumulation and the economic take-off. They highlight the role of upper tail human capital and scientific knowledge.¹ We complement this literature by emphasizing the rural exodus as a pre-requirement for industrialization.

Historical demographers like Dyson (2011) and economic historians like Bairoch (1997) show some regularities of urbanization, industrialization and the demographic transition. First, before and during the demographic transition, cities “killed” people in the sense that mortality was significantly higher than in the countryside.² Second, the countryside was more fertile than cities as the number of births per women in rural areas exceeded those in urban areas.³ Third, despite cities “killed” more people and “produced” less newborns than the countryside, urbanization took place. Since neither international immigration nor emigration were able to sustain urbanization in

¹The Unified Growth Theory has produced a remarkable number of other refinements. Galor (2011) provides an enlightening review of this literature. Franck *et al.* (2015) also document, using French data, that industrialization has been a workhorse of the demographic transition confirming one key assumption of the UGT.

²The differential mortality at the beginning of the industrial revolution is widely documented, see e.g. Woods *et al.* (1988), Woods (2003) or Knodel (1974).

³Among others, Bocquier and Costa (2015) and Mosk (1980) discuss a differential fertility in cities and countryside.

European countries at that time, urbanization was only possible because of a massive and sustained rural exodus.⁴ Fourth, as cities were the natural ecosystem of industries,⁵ increasing shares of workers living in cities may have been a pre-requirement for industrialization.⁶

To assess if the rural exodus has really been crucial for the process of industrialization, the economic take-off, and if it has been as important as mortality reductions, we have first collected data for Denmark between 1840 and 1940. We evaluate fertility and mortality differentials between cities and countryside as well as the area specific dynamics of income per capita. As explained in Section 2, because the age structures of cities and the countryside were quite different during the 19th and 20th centuries, we compare fertility and mortality between cities and countryside using Coale's index of fertility and infant mortality rates (IMR). We evidence that the demographic transition was initiated in the Danish countryside where IMR decreased as soon as 1840 while fertility was rather constant. Fertility started to decrease with a lag, around 1880. In cities, reductions of IMR appeared only 50 years later, around 1890 at the same time as reductions in fertility (see Figure 1).

Let's turn to urbanization that we define as the increase of the proportion of people living in cities.⁷ For Denmark until 1850, the share of individuals in cities was quite stable and no urbanization is observable. Around 20% of all individuals lived in cities in 1850, twice the number from Sweden. The proportion increased to 54.5% in 1940. Both agricultural and industrial output experienced rather constant growth rates since 1818. As growth became higher in industries, it dominated agricultural output as soon as 1930. Despite some temporary fluctuations, the price of agricultural relative to industrial goods increased during the take-off to modern growth until the first world war. Afterwards, the relative price mainly stagnated (see Figure 3 of Section 2). Finally, the data we have gathered for education as well as stylized facts from papers like Meyer *et al.* (1992), de la Croix *et al.* (2008) and Benavot and Riddle (1988), indicate that enrollment rates to primary education were rather low at the beginning of the 19th

⁴We do not argue here that international immigration did not exist at that time but they were far too low to sustain the demographic and urban growth. In fact, we present some evidence for the significant role of the rural exodus in Denmark in Appendix A.2.2. For a more general discussion, see, for instance, Chesnais *et al.* (1992) or Ravenstein (1885; 1889) on this issue.

⁵See, for instance, Bertinelli and Black (2004), Bairoch (1997) and Lesthaeghe (1977).

⁶This is neither a necessary nor a sufficient condition for industrialization. Indeed, technological progress could be strong and asymmetric enough (in favor of industries) to ensure industrialization of the production without any form of exodus.

⁷This definition implies that experiencing a rural exodus is neither a necessary nor a sufficient condition to experience urbanization as natural population growth of cities and countryside differ.

century. This changed radically after the reform of 1814.⁸ We evidence a s-shaped pattern for enrollment rates in primary school.

Linking differential fertility and mortality to sector specific productivity and freedom of movements, our theory gives a salient role to the rural exodus in the interactions between the industrial revolution, urbanization and the demographic transition. We construct an overlapping-generation model that distinguishes explicitly between cities and the countryside. Modern industries settle in cities and agrarian production in the countryside.⁹ Additionally, mortality is area specific. Adults decide either to stay in their area of birth or to move to the alternative one. While this implies the opportunity of a rural exodus as well as an exodus from the cities, individuals can only choose their place of residence once at the beginning of adulthood. Return migration is excluded. Once they have settled down, they decide on their number of births, how much they educate their surviving children and how much to consume from the two alternative goods: the agricultural and the industrial good. Like in Kögel and Prskawetz (2001), individuals need to consume a subsistence level of the agricultural good before they have the choice to enjoy children and consumption of the manufactured goods.¹⁰ As parents value children's human capital, they educate them as soon as they are sufficiently rich, even if only urban industries require human capital.

In early stages of our economy, both the share of individuals in cities and the contribution of industrial production to GDP are low. Most parents are too poor to invest in human capital. As a result, the vast majority of parents do not face any trade-off between the quality (education) and the quantity of children such that average fertility

⁸The reform that has been implemented in 1814 may be considered as endogenous. It could be the case that a significant part of the population wanted education to become cheaper and more efficient such that the Danish state decided to make it public and compulsory. For sake of simplicity, we decided not to tackle this issue. Our theory nevertheless predicts that industrialization reduced the cost of education.

⁹Obviously, assuming that industries settled exclusively in cities is a rather strong assumption. Indeed, Bairoch (1997) as well as Ogilvie and Cerman (1996) and Mokyr *et al.* (2012) document a significant proto-industrialization in Western Europe's countryside during and even before the industrial revolution. This is why we consider industrialization as the development of modern industries needing mass production, efficient work organization, large labor force and human capital. It may then be true that in some cases, urbanization has been a kind of consequence of industrialization: because a big industry settled in countryside, a city developed around. Nevertheless, this last point does not contradict the fact that modern industries required densely populated areas. Lampe and Sharp (2015) evidence the role of the milk industry in the Danish economic modernization. Nevertheless, they make no statement about the localization of dairies providing no evidence in favor of a significant industrialization of the Danish countryside.

¹⁰While agricultural consumption does not generate utility in Kögel and Prskawetz (2001) and its demand is fixed to the subsistence level if income exceeds this latter, our individuals value the agricultural good and its consumption always increases with income.

increases with income. The economy is trapped in a Malthusian regime. Partly mimicking Galor and Weil (2000), our economy starts to experience an exogenous technological progress once the overall workforce reaches a threshold. The technological level starts to increase in cities and countryside and more and more parents are able to educate. Nevertheless, it is not sufficient that investing in human capital becomes economically desirable to enter in the process of industrialization; living in cities and so leaving the countryside has to be desirable, too. What can prevent this? The mortality penalty that existed in cities: urban mortality crises that push people out of urban areas or refrain people from the countryside to move. Mortality peaks in cities could significantly lower urbanization and industrialization and require a more intense rural exodus while this latter is less desirable.¹¹ The economic cost of the exodus to cities may also be crucial.¹² When this cost is high, the prevailing technological progress may not allow development, simply because it is not high enough to attract peasants to urban areas. This is the main mechanism we quantify in this paper.

To assess the quantitative accuracy of our theory, we calibrate our model to match the observed urban and rural trends in Denmark between 1840 and 1940. Even if not perfectly, our theory reproduces the main characteristics of the time series we try to match. Our model replicates the dynamics of output in the agricultural and industrial sector, the dynamics of fertility in urban and rural areas, dynamics of education as well as the share of persons in each area. As we do not try to replicate the evolution of relative prices, we use their dynamics as an overidentification check of our theory. Moreover, we propose two historical experiments: a first counterfactual experiment evaluates the role of the rural exodus by alternatively making the exodus impossible or increasing its costs substantially. Second, we simulate the dynamics of the model in the case where reductions of IMRs did not occur anywhere and in the case where it occurred only in the countryside.

Not allowing any form of rural exodus would have changed Denmark’s economic and demographic development completely. While the fertility transition in cities would have

¹¹The literature on urbanization is a very rich one. In line with most models on urban-rural migration, income differentials are a driving force of urbanization in our model, see for instance Todaro (1969) and Harris and Todaro (1970). Recent contributions either posit urbanization as a pre-requirement for industrialization (see Bertinelli and Black (2004)) or highlight cases where urbanization does not require an industrialization, due to the emergence of “consumption cities” in resource-exporting countries (see Gollin *et al.* (2016)). As we focus on early Western Europe, urbanization appears clearly as a potential necessary condition for industrialization.

¹²Different production technologies combined with costs to move link our theory to the “new economic geography” which studies agglomeration exclusively taking into account factor mobility and/or transportation costs between areas. Ottaviano and Thisse (2004) offer a review on this literature.

accelerated dramatically, the countryside would have been trapped in the Malthusian regime until 1940, a regime where fertility would have increased. Total population size would have grown much faster but cities would have vanished over time. Simultaneously, total GDP would have accelerated in the agricultural sector but shrunk in the industrial sector. Nevertheless, at the individual level, due to population dynamics, welfare in cities would have increased tremendously while populations in the countryside would have suffered increasing poverty. The transition to mass education would not have happened. Only the very low share of individuals in cities would have educated more and more. The rural exodus then appears as a mechanism that may prevent inequality as it allows populations from the countryside to benefit from industrial opportunities. Increasing migration costs goes along with similar but less extreme consequences. As soon as the migration costs are tripled, the educational, demographic and economic revolutions of Denmark would have vanished. It would have decreased GDP per capita in rural areas by almost 40% in 1940. Such violent adjustments do not occur for smaller increases of migration costs. For instance, we estimate that doubling the costs of migration would have delayed the massification of education by around 20 years, left the income per capita in cities almost unchanged but would have reduced income per capita in the countryside by 21.5% in 1940.

Compared to limitations in the exodus, changing the dynamics in the survival probability has quite limited effects. Fertility would have been almost unchanged and thus the higher IMRs would have reduced population growth. Additionally, the lower number of surviving children would have decreased costs to educate children and mass education would have inserted earlier. GDP per worker would have slightly increased and thus compensated individuals for the higher mortality. By contrast, if only cities would have suffered from a stagnation in IMRs, GDP per worker would have increased in cities (by 5% in 1940) but shrunk in the countryside (by only 2%). Still, even if the higher income differential would have compensated for the lower share of surviving children and intensified the rural exodus, urbanization would have diminished. This rather weak effect of mortality reduction is in line with the results of Galor (2005) and Doepke (2004).¹³

¹³Other papers treat the impact of mortality on the take-off in a radically different way leading to different results. Cervellati and Sunde (2011), Strulik and Werner (2012) or Strulik and Werner (2013) highlight a positive link between increasing life expectancy and human capital accumulation. Boucekkine *et al.* (2002) illustrate that a rising longevity might explain the take-off to sustained growth as well as the brake down of the positive relation between income and fertility in a model with vintage human capital. Strulik and Weisdorf (2014) emphasize the interaction of child mortality with fertility and economic development. Voigtländer and Voth (2013b) figure out, that mortality crises, like the black death, may explain divergent developments between Europe with “killer cities” and China. We

Except for Greenwood and Seshadri (2002) and Adams (2015) we are among the few economists who investigate the role of urban-rural differential demography in the process of economic development. Using a framework close to ours, Greenwood and Seshadri (2002) focus on the rural-urban fertility differential in the US. In contrast to our theory, they do not consider mortality and their economy has only one area which produces two goods with different technologies. While they show that technological progress is key for the reduction of fertility in rural and urban sectors, they do not study the exact role of the exodus conditions and mortality, neither theoretically nor quantitatively. The closest paper to ours in the literature is the one of Adams (2015) as we share the same initial objective that consists in estimating the deep parameters of a growth model where rural-urban differences are key.^{14,15} While Adams focuses on the qualitative reproduction of aggregated rural-urban differentials like urban wage premiums and differential fertility in England, we reproduce the dynamics of GDP and fertility in both sectors of the Danish economy. Adams (2015) tries to validate the main predictions of his theory thanks to cross-country estimations¹⁶, while we use our estimations to evaluate the relative importance of the rural exodus and the reduction of child mortality for the economic take-off. Finally, and maybe most importantly, compared to Adams (2015), we reach opposite conclusions about the role of mortality for the transition to modern economic and demographic conditions. The impact of improvements in infant mortality rates was quite limited – in particular in relation to the rural exodus.

The rest of the paper is organized as follows. Section 2 presents the Danish data we have collected. Section 3 describes our theoretical model. Afterwards, a calibration exercise in Section 4 enables to evaluate the fit of our theory before Section 5 offers counterfactual experiments. Section 6 concludes.

complement their approach by showing that simple adjustments of infant mortality is not sufficient to provoke the take-off.

¹⁴We also share some theoretical assumptions like the existence of fixed land supply, the use of human capital only in industries that exclusively settle in cities and the existence of a trade-off between quality and quantity of children. Nevertheless, we do not impose any kind of rule limiting a priori migration flows what is not the case in Adams (2015).

¹⁵If we are the first to estimate the deep parameters of a unified growth model, it is not the first one using calibration technics. Lagerlöf (2006) presents a numerical exercise for the model of Galor and Weil (2000) by choosing parameters either from the literature or currently observed moments. Cervellati and Sunde (2005) as well as Strulik and Weisdorf (2014) present illustrative numerical simulations to highlight the role of longevity and child survival, respectively. Cervellati *et al.* (2013) use panel data for 59 countries since 1880 for a structural estimation in a unified growth model.

¹⁶His three main predictions are that a country will grow more rapidly if it has “a lower initial urban share, a higher initial income growth rate, or a higher initial population growth rate”.

2 Danish historical data

For Western Europe, demographers and economists collected demographic and economic data at national as well as local levels. Such data have been widely used by the Unified Growth Theory to reconstruct a global demo-economic history of Western Europe. We complement this approach by providing a dataset for Denmark¹⁷ that makes a clear distinction between rural and urban areas. It provides information that go beyond simple demographic objects such as crude births and deaths rates. If these concepts are well suited for analyzing the global demographic dynamics of a country, they suffer strong weaknesses at a finer geographical scale. In complement to demographic data, we also build data on the evolution of GDP both in current and constant prices for both areas. From these latter, we infer the evolution of the relative price of agricultural goods compared to industrial goods.¹⁸

Demographic side

Coale (1969) proposed an index that normalizes the number of births by the number of births that would have occurred if the population under study would experience the age-specific fertility of Hutterites reputed to be the closest to biological limits. This index then should belong to $(0, 1)$ and it is positively correlated to fertility of the population under study. This is an efficient way to normalize the number of births by the age structure of the population as the age structure of cities and countryside in Denmark were quite different.¹⁹ We denote $I_t^{\mathcal{A}}$ as the value of Coale's index for area \mathcal{A} at time t . It is calculated as follows:

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

¹⁷The same dataset is also offered for Sweden in Appendix A.3. It allows us to show that the data we have collected for Denmark share common characteristics with other European countries. Indeed, even if Danish and Swedish data differ in many aspects, they offer a quite consistent history for industrialization, demographic transition and urbanization.

¹⁸Notice that a general issue arises with data comparability. Each data source that we use here may use alternative definitions for what is urban and what is rural; furthermore, even though we use one unique data source like statistical yearbooks, the definition may change. We discuss the issues raised by the successive changes in the definition of what is a city and what is a countryside in Appendix A.2.1.

¹⁹We briefly discuss alternative fertility indicators and the role of the Danish age structures in Appendix A.2.3.

where $B_t^{\mathcal{A}}$ denotes the number of births in area \mathcal{A} at time t while $\varepsilon_{o,t}^{\mathcal{A}}$ is the number of persons in area \mathcal{A} who belongs to age group o at time t . n_o^H finally denotes the Hutterite age-specific fertility.

The Princeton Project offers values of the Coale index in cities, in the countryside and for overall Denmark from 1852 to 1960 with ten years intervals (Treadway, 1980; Matthiessen, 1984). We report the values of the Coale index on the left panel of Figure 1.²⁰ Fertility started to decrease sharply around 1880 in cities while it remained relatively constant in the countryside until 1901. Surprisingly, the fertility differential between cities and countryside is much smaller before the industrial revolution than at the end of our period of observation.

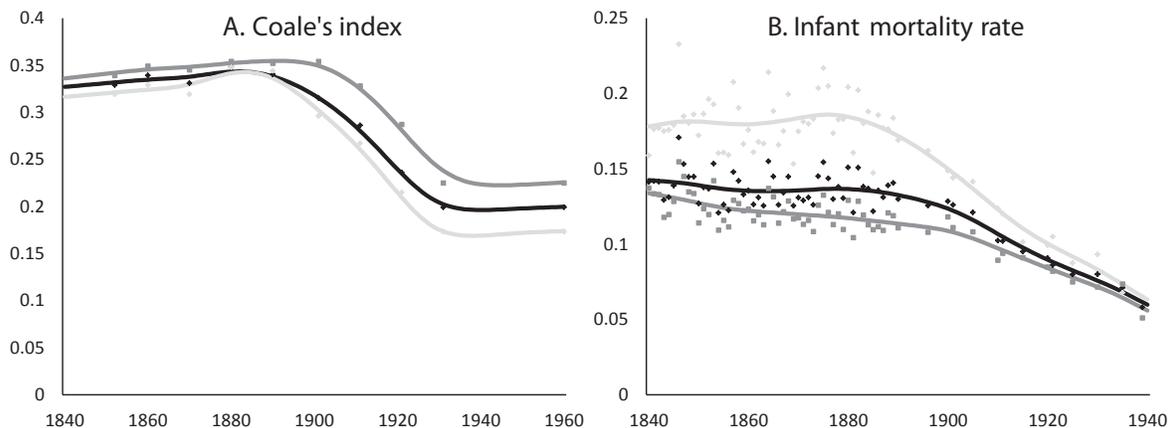


Figure 1: COALE'S INDEX AND INFANT MORTALITY RATES IN DENMARK

Figure A: Coale's index in rural areas (dark gray), cities (light gray) and total Denmark (black). Figure B: Infant mortality rates in rural areas (dark gray), cities (light gray) and total Denmark (black). For these graphs and all subsequent ones, solid lines represent HP filtered trends. Sources: Treadway (1980), Matthiessen (1984) and Johansen (2002); own calculations.

Demographers like Woods (2003) have emphasized a mortality penalty in cities compared to countryside at the time of the industrial revolution. In a perfect framework,

²⁰Two potential problems arise with these data from the Princeton Project. First, for the year 1890, Coale's index for overall Denmark is smaller than the indices in cities and the countryside, which might be due to different age-specific fertility patterns in the two areas. Second, an alternative source of data comes from Johansen (2002). When we use his data to compute general fertility rates, we find a strong fertility differential between cities and the countryside prior to the industrial revolution what is not the case for the Princeton Project. This last point could drastically change our vision of the demographic dynamics of Denmark before and during industrialization. We discuss this point in Appendix A.2.3.

one would like to reconstruct the dynamics of mortality at all ages to get a complete picture of this phenomenon. Nevertheless, restricted data availability and quality do not allow to build such a time series for a country like Denmark in the 19th century. We then focus on the local and global dynamics of infant mortality rates (IMR), the right panel of Figure 1 exposes infant mortality rates. These data come from Johansen (2002) for the years 1840 to 1890 on a yearly basis and from 1896 to 1939 every five years. Additionally, we add observations on infant mortality provided by the Princeton Project to this time series. Focusing on infant mortality is important because infants and young children are particularly sensitive to environmental conditions and also because it was sufficiently large to affect national levels and trends (see e.g. Woods *et al.* (1988)). Furthermore, it is an important determinant of the number of births within a family as well as of the number of surviving children. In our structural approach of the long-run dynamics of Denmark, we only focus on this type of mortality. We discuss consistency of our dataset in Appendix A.

Data on infant mortality are quite remarkable: in addition to be much higher in cities than in countryside, infant mortality in cities remained rather constant up to 1880 while it decreased almost monotonously in the countryside from 1840 to 1940. Between 1880 and 1920, mortality in cities has been divided by almost two making mortality differential between cities and countryside almost vanishing. In other words, no lag can be observed between fertility and mortality reduction in cities while a large one is detectable in the countryside.

Turning to urbanization, Johansen (2002) provides data for the year 1660, relying on calculations of Lassen (1965).²¹ He also provides historical evidence indicating that during the 17th and the beginning of the 18th centuries, the share of persons living in cities has been rather stable in Denmark. The rest of the data comes from Statistiske Undersøgelser, vol. 19. (Departement, 1966). In fact, no significant urbanization can be observed in Figure 2A before 1850 where a break occurred.²² Between 1850 and 1940, the urbanization rate has been multiplied by more than 2.5 increasing from around 20% to 54.5% what corresponds to an explosion of the population in cities as in the meantime, the size of the total Danish population has more than doubled.

²¹This data point is estimated mainly by two sources: the existing portion of the Peasants Survey in 1660 mainly covers rural areas. Population figures on towns have been taken from the 1672 census. The first modern census was taken 15.08.1769. Unfortunately, most of this census has been lost. The next census was done 1.7.1787. Still, the first reliable census was done 1.2.1801 (Lassen, 1966).

²²At the end of our period of observation, a reduction of urbanization rates can be observed. As our calibration-simulation exercise stops in 1940, this phenomenon that we suspect to originate from changes in the definition of cities and countryside is not too important.

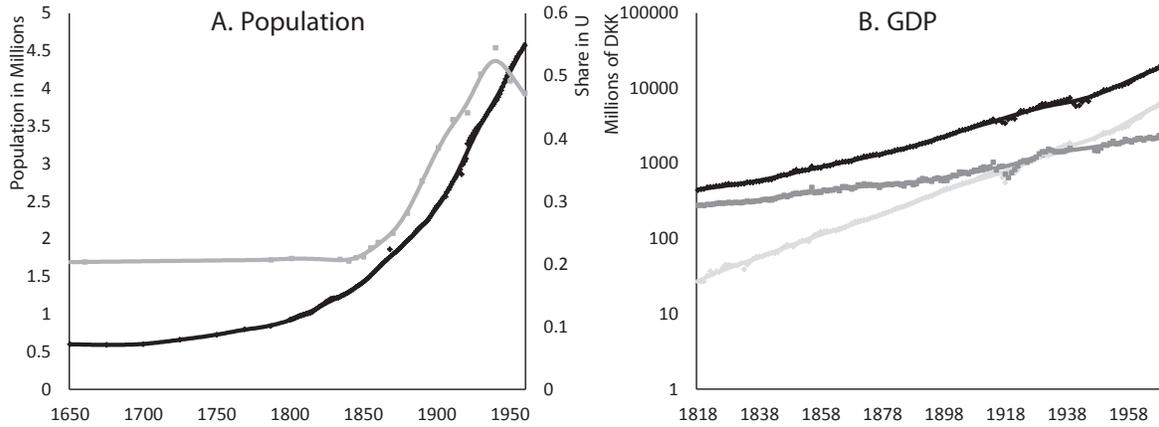


Figure 2: POPULATION DYNAMICS AND GDP IN DENMARK

Figure A: Share of persons living in urban areas (U) in gray on the right y-axis and total population of Denmark (black) on the left y-axis. Figure B: Denmark’s agricultural (dark gray), industrial (light gray) and total GDP (black). Sources: Departement (1966), Johansen (1985; 2002); own calculations.

Economic side

Data on production and relative prices are obtained from Johansen (1985). To approximate the agricultural production, we use the sector “Agriculture, Forestry, Fishing” of Johansen while we approximate industrial production thanks to the production in the sector called “Trade, industry and public utilities”. We do not include the production of sectors like Commerce, Transport and Communication as well as Public services and Building and Construction in any of the two sectors. In other words, our definition of industrial production is quite conservative. Industrial production (Figure 2B) started to grow well before urbanization took place and became more important than agricultural production between the two World Wars.²³ As Johansen (1985) provides GDP in constant and current prices, we can easily compute prices in each sector and deduce the price of industrial goods relative to agricultural goods, plotted in Figure 3A. Not surprisingly, the relative price of industrial goods decreased in the long-run despite a stagnation can be detected after 1918.

²³Based on Westergaard (1922) and Kindleberger (1951), Lampe and Sharp (2015) evoke a rapid catch up of the Danish economy to world’s leading economies after around 1864. This catch up may have been driven by the development of milk and butter industries. In our data, we are neither able to identify a break in growth of total GDP nor in growth rates of the industrial or agricultural sector since 1818.

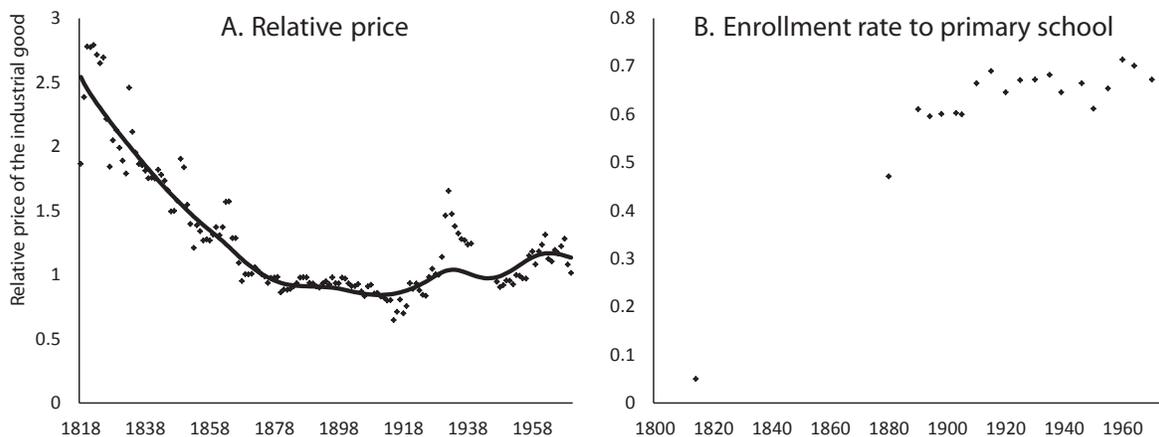


Figure 3: RELATIVE PRICE AND ENROLLMENT RATES TO PRIMARY SCHOOL IN DENMARK

Figure A: Dynamics in the relative prices of industrial compared to agricultural goods. Figure B: Primary education enrollment rates in Denmark. Sources: Johansen (1985), Flora (1983), Benavot and Riddle (1988) and de la Croix *et al.* (2008); own calculations.

In Denmark like everywhere else in Western Europe, economic modernization has been accompanied by a massification of education, this is highlighted for instance by Galor (2005). Even if not perfect, we approximate the dynamics of education thanks to the enrollment rates to primary school, defined as % of pupils in age group 5–14. Since no entire time series capturing 1840–1940 exists, we had to gather many sources.

While compulsory primary education was introduced in 1814 in Denmark,²⁴ enrollment rates remained first pretty low. We use data points offered by Flora (1983) starting in 1894 and add the years 1880 and 1890 from Benavot and Riddle (1988). No reliable information is available before 1880. As enrollment rates were generally quite low before the introduction of compulsory primary education and their development afterwards were quite similar across Scandinavian countries²⁵ we apply the Swedish information gathered from de la Croix *et al.* (2008) before 1880. This leads to right panel of Figure 3 where we only display data points.²⁶

Summarizing our stylized facts, the urbanization of Denmark started around 1850 while its demographic transition impulsed from the countryside where IMRs started to de-

²⁴Children were given the right of 7 years of education (Flora, 1983).

²⁵See Soysal and Strang (1989) page 281 for a discussion about the similarity that exist between the Swedish and the Danish histories of education massification.

²⁶Notice that Morrisson and Murin (2009) use an alternative dataset from Mitchell delivering a slightly different view of the education democratization process in Denmark.

crease in 1840. Fertility declined in the countryside only around 1880. The demographic transition of cities started later, around 1880, but without time lag between the reduction of mortality and fertility. Industrial production started to dominate agricultural production around 1930 while urbanization begun around 1850.

3 The model

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

3.1 Production sectors

Concerning production technologies, we assume that the only activity existing in rural areas is agriculture while only manufactured production settles in cities.²⁷ In the countryside, firms produce the agricultural output Y_t^R with unskilled labor and land following a technology with constant returns to scale:

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

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

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

²⁷We discuss this assumption and the related critics of Clark *et al.* (2012) in Appendix A.3.3.

The manufactured good is produced in the urban areas using efficient labor H_t :

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

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

We assume the existence of a general knowledge. It is used and a source of technological progress in both sectors. We denote the stock of this knowledge as A_t , and it starts to increase exogenously when the size of the population becomes large enough:

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

This assumption is close to the one of Galor and Weil (1999) and formalized in Galor and Weil (2000).²⁸ Furthermore, we assume that the respective sector specific technological situation is as follows:

$$A_{R,t} = B_R A_t \quad (4)$$

$$A_{U,t} = B_U A_t H_t^{1-\gamma}. \quad (5)$$

For all \mathbb{A} , $B_{\mathbb{A}} \in \mathbb{R}^+$ is a scale parameter allowing for initial technological differences between industries and agriculture. Hand in hand with Equation (2), Equations (4)–(5) are at the core of our theory: benefiting from an (ex-post) *AK* technology, human capital is valued only in the cities meaning that urbanization may be a pre-requirement to industrialization.

²⁸An alternative would consist in assuming that exogenous technological progress appears when the size of urban populations reaches a threshold. This would have made population growth in cities a pre-requirement for the take-off what is not the case in our model.

3.2 Individuals

We assume the existence of a subsistence level for income, denoted $\bar{c} > 0$, below which people cannot do anything else than consuming the agricultural good c_t . The latter is chosen as numeraire.²⁹ This is a state of extreme poverty. When their income increases sufficiently, people can potentially consume industrial goods d_t , give birth to n_t children, educate them e_t and also migrate. Migration entails fix costs which are denoted $\bar{\kappa} > 0$ and arise at the beginning of adult life. Children are born after potential migration³⁰ and costly in terms of time: giving birth to a child requires a part ξ of the parental time unit. Furthermore, with exogenous probability $q_t^{\mathbb{A}}$, the child survives to age one. Then, it additionally needs area specific time from its parents denoted $\zeta^{\mathbb{A}}$, with $(\xi, \zeta^{\mathbb{A}}) \in [0, 1]^2$, such that the labor force participation of a parent living in area \mathbb{A} and having n_t children is $I_t = 1 - (\xi + \zeta^{\mathbb{A}} q_t^{\mathbb{A}}) n_t$. Finally, education entails industrial good cost such that each unit of education costs βp_t to the parents and their budget constraint writes:

$$c_t + p_t d_t + \beta q_t^{\mathbb{A}} p_t n_t e_t = \left(1 - [\xi + \zeta^{\mathbb{A}} q_t^{\mathbb{A}}] n_t\right) \omega_t^{\mathbb{A}} - \kappa. \quad (6)$$

If the person moves from one area to another, κ equals $\bar{\kappa}$ and is zero otherwise. Because raising children takes time, the number of surviving children, a mother can have, is limited to $n_M \equiv \frac{1}{\xi + \zeta^{\mathbb{A}} q_t^{\mathbb{A}}}$. $\omega_t^{\mathbb{A}}$ is the potential labor income of a parent living in area \mathbb{A} at date t . Potential income is only determined by labor income such that:

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

with $w^{\mathbb{A}}$ denoting the wage per efficient unit of labor in sector \mathbb{A} . Similar to de la Croix and Doepke (2003), human capital accumulates according to:

$$h_{t+1} = (v + e_t)^\phi h_t^{1-\phi}, \quad (7)$$

where $v > 0$ ensures that in absence of educational investment, human capital does not converge to zero. Each dynasty is endowed with a specific initial level of human capital

²⁹For reasons of clarity, we neglect the superscript for the individuals i hereinafter.

³⁰We make this assumption for sake of simplicity. Allowing people to have kids before migrating would require to use a 3-period OLG-model what would complexify our analysis without changing its main results.

$h_0 > 0$. This initial value differs from one dynasty to another and is drawn from an exponential distribution with average λ . There is an intrafamily transmission of human capital, whereat $1 - \phi$ denotes the elasticity by which parents' human capital enters human capital accumulation.

One should notice that parents face the same budget constraint as well as human capital accumulation technology in cities and in the countryside. The same holds for preferences. Utility of an agent living in area \mathbb{A} is denoted $W_t^{\mathbb{A}}$.³¹

$$W_t^{\mathbb{A}} = \alpha \ln(c_t - \bar{c}) + (1 - \alpha) \ln(d_t + \varepsilon) + \rho [\ln(n_t q_t^{\mathbb{A}}) + \ln h_{t+1}]. \quad (8)$$

Parents are characterized by “warm glow preferences”, hence, they do not value the future well-being of their children, but they care about future human capital h_{t+1} .

Parental behavior

We assume that young adults have to decide between migrating or staying where they were born before having children, consuming, working, etc. Once they know where they live, they have to determine how many children they have, if any, and how they allocate their income between consumption of agricultural goods, industrial goods and children's education. We solve this problem backwards. Future parents calculate their optimal behavior in each location given the prevailing economic conditions. Then, they decide where to live by comparing their expected indirect utilities.

In Table 1, we list the six potential regimes adults can live in. In a subsequent proposition, we describe the conditions under which these regimes are a best response to prevailing conditions in cities and countryside. Regimes are obtained by maximizing the parental utility function (Eq. (8)) with respect to budget constraint (Eq. (6)), maximal fertility and minimal consumption constraints.

³¹The fact that people value industrial goods in their preferences is not a sufficient condition for urbanization to arise. We show this result in Subsection 5.

Table 1: REGIMES IN WHICH ADULTS CAN LIVE

Regime	c_t	n_t	d_t	e_t	Exodus
R1	$\omega_t^{\mathbb{A}}$	0	0	0	No
R2	$\frac{\alpha\omega_t^{\mathbb{A}}}{\alpha+\rho} + \frac{\rho}{\alpha+\rho}\bar{c}$	$\frac{\rho(\omega_t^{\mathbb{A}}-\bar{c})}{(\alpha+\rho)(\xi+\zeta^{\mathbb{A}}q_t^{\mathbb{A}})\omega_t^{\mathbb{A}}}$	0	0	No
R3	$\frac{\alpha(\omega_t^{\mathbb{A}}-\kappa)}{\alpha+\rho} + \frac{\rho}{\alpha+\rho}\bar{c}$	$\frac{\rho(\omega_t^{\mathbb{A}}-\kappa-\bar{c})}{(\alpha+\rho)(\xi+\zeta^{\mathbb{A}}q_t^{\mathbb{A}})\omega_t^{\mathbb{A}}}$	0	0	Yes
R4	$\frac{\alpha\tilde{\omega}_t^{\mathbb{A}}}{1+\rho} + \tilde{c}$	$\frac{\rho(\tilde{\omega}_t^{\mathbb{A}}-\bar{c})}{(1+\rho)(\xi+\zeta^{\mathbb{A}}q_t^{\mathbb{A}})\omega_t^{\mathbb{A}}}$	$\frac{1-\alpha}{1+\rho} \frac{(\tilde{\omega}_t^{\mathbb{A}}-\bar{c})}{p_t} - \varepsilon$	0	Yes
R5	$\frac{\alpha(\omega_t^{\mathbb{A}}-\kappa)}{\alpha+\rho} + \frac{\rho}{\alpha+\rho}\bar{c}$	$\frac{\rho(1-\phi)(\omega_t^{\mathbb{A}}-\kappa-\bar{c})}{(\alpha+\rho)\pi_t}$	0	$\frac{\pi_t}{(1-\phi)\beta q_t^{\mathbb{A}}}$	Yes
R6	$\frac{\alpha\tilde{\omega}_t^{\mathbb{A}}}{1+\rho} + \tilde{c}$	$\frac{\rho(1-\phi)(\tilde{\omega}_t^{\mathbb{A}}-\bar{c})}{(1+\rho)\pi_t}$	$\frac{1-\alpha}{1+\rho} \frac{(\tilde{\omega}_t^{\mathbb{A}}-\bar{c})}{p_t} - \varepsilon$	$\frac{\pi_t}{(1-\phi)\beta p_t q_t^{\mathbb{A}}}$	Yes

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

The last column Exodus in Table 1 indicates if migration is affordable.³² Poverty in regime R1 imposes childlessness and very low consumption levels.³³ In regime R2, having children is possible and optimal but their number is limited by \bar{c} and it is impossible to move from one area to another, to consume industrial goods and to educate children. Moving from one area to another becomes additionally possible in regime R3. In regime R5, the price of the industrial good is relatively high. Parents can educate children but they cannot consume industrial goods. By contrast, if the price of the industrial good is relatively low, parents can consume the industrial good but not educate children. This is regime R4. Finally, regime R6 is a state where, except for budget, no constraint binds.

In order to determine the optimal behavior of potential parents facing the system of prices $\{p_t, w_t^U, w_t^R\}$ and mortality rates $\{q_t^U, q_t^R\}$, we introduce the following critical values on wages and the relative price: $\bar{\omega}_t$ is the potential income above which a parent living in area \mathbb{A} and who had paid κ , decides to educate his or her children:

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

³²As investing in human capital costs goods rather than time in our framework, rich people invest in education before poor people. As a result, rural-urban migrants are better educated than those who stay in their area of origin. This is in line with the findings of Funkhouser (1998).

³³Notice that this poverty driven childlessness is the only form of childlessness our model is able to account for. As explained in Baudin *et al.* (2015), alternative types of childlessness exist and can have a crucial impact on the global demographic dynamics. We, nevertheless, disregard these alternative forms for simplicity.

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

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

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

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

Proposition 1

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

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

such that:

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

when $p_t \leq p^*(\kappa)$, optimal decisions of people are:

Not moving person (AA):

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

Moving person (AB):

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

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

Not moving person (AA):

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

Moving person (AB):

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

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

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

Not moving person (AA):

$$\begin{aligned} \text{if } \omega_t^{\mathbb{A}} \leq \bar{c} \quad \text{then} \quad & R1 \\ \text{if } \omega_t^{\mathbb{A}} \in]\bar{c}, \hat{\omega}_t] \quad \text{then} \quad & R5 \\ \text{if } \omega_t^{\mathbb{A}} \geq \hat{\omega}_t \quad \text{then} \quad & R6 \end{aligned}$$

Moving person (AB):

$$\begin{aligned} \text{if } \omega_t^{\mathbb{A}} \in [\bar{c} + \bar{\kappa}, \hat{\omega}_t] \quad \text{then} \quad & R5 \\ \text{if } \omega_t^{\mathbb{A}} > \hat{\omega}_t \quad \text{then} \quad & R6 \end{aligned}$$

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

Not moving person (AA):

$$\begin{aligned} \text{if } \omega_t^{\mathbb{A}} \leq \bar{c} \quad \text{then} \quad & R1 \\ \text{if } \omega_t^{\mathbb{A}} \in]\bar{c}, \bar{\omega}_t] \quad \text{then} \quad & R2 \\ \text{if } \omega_t^{\mathbb{A}} \in]\bar{\omega}_t, \hat{\omega}_t] \quad \text{then} \quad & R5 \\ \text{if } \omega_t^{\mathbb{A}} \geq \hat{\omega}_t \quad \text{then} \quad & R6 \end{aligned}$$

Moving person (AB):

$$\begin{aligned} \text{if } \omega_t^{\mathbb{A}} \in [\bar{c} + \bar{\kappa}, \bar{\omega}_t] \quad \text{then} \quad & R3 \\ \text{if } \omega_t^{\mathbb{A}} \in [\bar{\omega}_t, \hat{\omega}_t] \quad \text{then} \quad & R5 \\ \text{if } \omega_t^{\mathbb{A}} > \hat{\omega}_t \quad \text{then} \quad & R6 \end{aligned}$$

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

Not moving person (AA):

$$\begin{aligned} \text{if } \omega_t^{\mathbb{A}} \leq \bar{c} \quad \text{then} \quad & R1 \\ \text{if } \omega_t^{\mathbb{A}} \in]\bar{c}, \hat{\omega}_t] \quad \text{then} \quad & R2 \\ \text{if } \omega_t^{\mathbb{A}} \in]\hat{\omega}_t, \bar{\omega}_t] \quad \text{then} \quad & R4 \\ \text{if } \omega_t^{\mathbb{A}} > \bar{\omega}_t \quad \text{then} \quad & R6 \end{aligned}$$

Moving person (AB):

$$\begin{aligned} \text{if } \omega_t^{\mathbb{A}} \in [\bar{c} + \bar{\kappa}, \hat{\omega}_t] \quad \text{then} \quad & R3 \\ \text{if } \omega_t^{\mathbb{A}} \in [\hat{\omega}_t, \bar{\omega}_t] \quad \text{then} \quad & R4 \\ \text{if } \omega_t^{\mathbb{A}} > \bar{\omega}_t \quad \text{then} \quad & R6 \end{aligned}$$

Proof. See Appendix B ■

Impact of wages on individual fertility

In this subsection, we highlight the non-monotonicity that characterizes the relationship between fertility and income. For an illustration purpose, we consider the situation of a non-moving person in the case where $v > \bar{v}_t$ and $p_t \in]p^*(\kappa), \tilde{p}(\kappa)]$. In this situation, we get that:

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

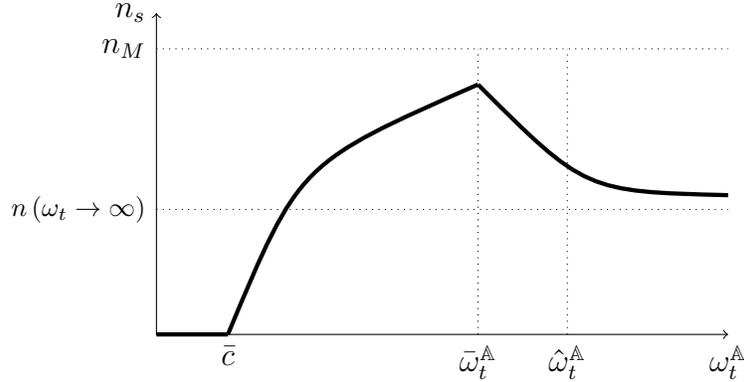


Figure 4: RELATION BETWEEN FERTILITY AND INCOME WHEN $v > \bar{v}_t$ AND $p_t \in]p^*(\kappa), \tilde{p}(\kappa)]$

The relationship between optimal fertility and wages is represented in Figure 4. For our role model who decided not to migrate, having an income lower than \bar{c} makes her unable to give birth. For higher levels of wage, she can have children in a Malthusian context where fertility is limited by earning capacities. Any increase in $\omega_t^{\mathbb{A}}$ will allow this person having more children. When income exceeds $\bar{\omega}_t^{\mathbb{A}}$, educating children becomes affordable and a trade-off between quality and quantity of children takes place. The higher $\omega_t^{\mathbb{A}}$ the higher the cost of quantity relative to the cost of quality. Increases in wage rates incite parents to have fewer children better educated. Once $\omega_t^{\mathbb{A}}$ becomes higher than $\hat{\omega}_t^{\mathbb{A}}$, investing in the education of children can be done without renouncing to the consumption of industrial goods.

Should I stay or should I go?

To decide where she wants to live, a person born in area \mathbb{A} has to compare the value to live in a city to the value to live in the countryside. She will decide to live in a city if:³⁴

$$V [U|\mathbb{A}, h_t, \omega_t^U, q_t^U] \geq V [R|\mathbb{A}, h_t, \omega_t^R, q_t^R]. \quad (9)$$

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

We can also show that there exist 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, $\omega_t^R = w_t^R$ does not depend on h_t . Hence, an increase in h_t will change the way the person will allocate his or her total investment in children between quality and quantity but it will not change the potential income of this person. Said differently, a higher h_t does not change the set of choice faced by an individual living in the countryside. By contrast, the set of choices of a person living in a city enlarges when h_t increases, it implies that the value to live in a city monotonously increases with h_t while it is not the case for the value of living in the countryside. The threshold \check{h}_t then follows.

Importantly, the value of living in one place decreases with the child mortality rate that prevails in this place. For this reason, a higher child mortality in cities will refrain rural exodus. The returns to human capital will have to be greater to attract new workers when child mortality is higher. It could also be the case that mortality crisis in cities provoke an urban exodus – pushing persons toward the countryside. In this case, the persons with the lowest levels of human capital in cities will be the first to leave to the countryside.

³⁴By convention, we assume that if a person is indifferent between living in a city or in the countryside, she decide to live in a city.

3.3 Market equilibria

As now individual decisions have been determined, we can turn to the description of market equilibria. In this economy, we have four markets: the labor market in cities, the labor market in the countryside, the industrial good market and the agricultural good market. Thanks to the Walras' law, we know that markets are clearing if the three following conditions are met:

$$w_t^U = A_t^U p_t \quad (10)$$

$$\sum_{\mathcal{I}_R} I_t^i = X \left(\frac{A_t^R}{w_t^R} \right)^{\frac{1}{\theta}} \quad (11)$$

$$\sum_{\mathcal{I}_U} d_t^i + \sum_{\mathcal{I}_R} d_t^i = A_t^U \sum_{\mathcal{I}_U} h_t^i I_t^i \quad (12)$$

Eq. (10) ensures the equilibrium on the urban labor market. Because of the AK production technology, wages do not depend on labor supply. Eq. (11) ensures the equilibrium on the rural labor market. Demand is a decreasing function of wages, which are equal to the average product since land is a public good. Labor supply depends on the available time of persons living in the countryside. The more children they have, the less they participate to the labor market. Finally, condition (Eq. (12)) describes the equilibrium on the industrial good market. The demand comes from both the cities and the countryside while the supply is ensured only by the cities. As for the agricultural production in the countryside, the higher the fertility of workers in cities, the lower the total production of industrial goods.

We can finally notice that the resource constraint of the economy writes as follows:

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

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

4 Identification

After having described the long-run economic and demographic dynamics of Denmark and our theory, we combine them to see how well our theory fits the data. Therefore, some of our parameters will be set a priori using the literature, while others will be fixed such as minimizing the distance between our empirical and theoretical moments. Before discussing the way we proceed for each type of parameter, we introduce firstly the way we compute Coale’s index and GDP from the model and secondly how human capital is initially distributed among the 1000 dynasties that will generate our simulated moments and the dynasties on cities and the countryside.

Coale’s index

As our model does not predict Coale’s index directly, we approximate its value following the main idea of Coale (1969). We relate fertility in the model to the maximum fertility of the Hutterite:

$$\hat{I}_t^{\mathbb{A}} = \frac{\sum_{\mathcal{I}^{\mathbb{A}}} n_{i,t}^{\mathbb{A}}}{\frac{N_t^{\mathbb{A}}}{2} 5 \sum_o n_o^H},$$

where:

Table 2: AGE-SPECIFIC FERTILITY RATES OF HUTTERITES

o	15–19	20–24	25–29	30–34	35–39	40–44	45–49
n_o^H	0.3	0.55	0.502	0.447	0.406	0.222	0.061

Source: Coale (1969).

We divide the total population in area \mathbb{A} by two as we consider a single sex population. Furthermore, we multiply the age-specific fertility rates of Hutterites by five to account for the 5-year age groups. Within the model, we define the enrollment rate to primary education as the share of educated children related to the surviving number of children. Additionally, we normalize observed and predicted GDP by dividing their values by the GDP in the agricultural sector in the year 1840.

Initial distribution of human capital and dynasties

For each of our 1000 dynasties, we draw an initial human capital h_0^i from an exponential distribution with λ as average. To ensure that the distribution is ex-post fulfilled, we determine the initial value of each household by the inverse of the distribution:

$$h_0^i = \frac{\ln\left(1 - \frac{i}{N_0+1}\right)}{\lambda}. \quad (13)$$

We assume that the adults with the highest initial human capital locate in cities. The initial share of adults in cities is determined by $\delta_0 = 20.7\%$; the observation in 1840.

4.1 A priori fixed parameters

We calculate trends on infant mortality and pick values every 20 years presented in Table 3³⁵ and apply the following formula to obtain survival probabilities: $q_t^A = 1 - \text{IMR}_t^A$.

Table 3: INFANT MORTALITY RATES

Year	1840	1860	1880	1900	1920	1940
Rural	0.134	0.122	0.117	0.109	0.084	0.056
Urban	0.178	0.180	0.184	0.150	0.100	0.063

Three additional parameters are fixed a priori: θ , ϕ and ξ . $1 - \theta$ denotes the labor income share in the agricultural sector, or said differently, the partial output elasticity of the Cobb-Douglas function with respect to labor. In the literature values of $1 - \theta$ vary between 0.6 in Adamopoulos (2008) and 0.71 in Desmet and Parente (2012).³⁶ We fix the elasticity of labor ($1 - \theta$) to 0.65 and hence $\theta = 0.35$. As our human capital accumulation process is similar to de la Croix and Doepke (2003), we use their elasticity between education and human capital of parents $\phi = 0.635$. Additionally, we assume that a share $\xi = 0.04$ of each period is required to give birth. This assumption corresponds to a little less than 10 months for each pregnancy, including the time to recover from giving birth. Implicitly, it fixes the maximum number of pregnancies to 25. Finally, we normalize the size of land X and A_0 to 1.

³⁵We describe the general procedure in more detail in Appendix A.1.

³⁶Desmet and Parente (2012) cite Clark (2002) and Hayami *et al.* (1971) as sources for their parameter value.

4.2 Moments to be matched and minimum distance

A set of 14 parameters remains in the model, listed in Table 5. We choose this set of parameters to minimize the distance between the 36 observed and predicted moments, according to:

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

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

Table 4 presents the 36 moments to be matched. As explained in Section 2, we have been able to collect data on varying time periods for our variables of interest. While GDP is available on a yearly basis since 1818, the number of observations for Coale's index or the enrollment rates to primary school are quite limited. Except for education, we have first interpolated and smoothed our data using the Hodrick-Prescott filter (see Hodrick and Prescott (1997)). Then we have selected observations of the smoothed time series each 20 years from 1840 to 1940. For education, in order to interpolate the s-shaped evolution found in our data, we use a logistic function:

$$e_t^{\text{pr}} = \frac{e_0^{\text{pr}} e_{\text{MAX}}^{\text{pr}}}{e_0^{\text{pr}} + (e_{\text{MAX}}^{\text{pr}} - e_0^{\text{pr}}) \exp(-k e_{\text{MAX}}^{\text{pr}} t)},$$

with e_t^{pr} as enrollment rate to primary education at date t and $e_{\text{MAX}}^{\text{pr}}$ as the maximum value. Since 5–7 years old children are in pre-school age and the highest value observable within our observation is 72.5% we set $e_{\text{MAX}}^{\text{pr}} = 0.75$. To approximate the enrollment rate when primary education became compulsory, we apply the Swedish value in 1814, $e_0^{\text{pr}} = 0.05$. Given our assumptions, we apply the Generalized Method of Moments to determine the k that minimizes the distance between observations and estimated values.

On the demographic side, taking infant mortality rates as given, we will try to match the values of Coale's index in cities and countryside as well as the distribution of the population between both types of areas. On the economic side, we calibrate the model to match total GDP in both areas as well as enrollment rates to primary education. All other data exposed in the previous part is used as overidentification checks.

Table 4: MOMENTS TO BE MATCHED BY OUR MODEL

Year	Coale's index		Share in cities	Total GDP*		Enrol. rates
	Rural	Urban		Rural	Urban	
1840	0.336	0.316	0.207	275.9	27.1	0.135
1860	0.345	0.324	0.231	326.5	56.8	0.256
1880	0.352	0.341	0.285	433.9	113.5	0.413
1900	0.351	0.306	0.379	516.3	212.6	0.558
1920	0.286	0.215	0.454	647.8	435.4	0.655
1940	0.223	0.169	0.524	916.7	785.7	0.707

* GDP in Millions of DKKs

Then we follow the strategy of Baudin *et al.* (2015) to minimize the distance between empirical and predicted moments. We first apply PIKAIA (Charbonneau, 2002) to search the region in the parameter space where the global maximum is located. Given this information as initial values, we identify the precise parameters by UOBYQA (Powell, 2002). However, compared to Baudin *et al.* (2015) one major difference exists: in order to get the fit of our model in each parametric situation, we need to compute the general equilibrium dynamics of our economy. To do that, we use the `fminsearch` algorithm which is also based on Powell (2002) and proposed by Fehr and Kindermann (2015). For each parametric condition explored by PIKAIA and UOBYQA, this algorithm minimizes the distance between supplies and demands on markets in order to find equilibrium prices and individual behaviors at the equilibrium. As this numerical exercise is extensively demanding in terms of computational power, a numerical algorithm is used inside a numerical algorithm, we apply a hybrid MPI-openMP programming strategy within Fortran 90.³⁷

³⁷Our MPI-openMP programming strategy bases upon the following idea: Starting from the master computer, PIKAIA allocates its parameter sets to the slaves or nodes by MPI (distributed memory system). Each node of the cluster then distributes the number of dynasties on its available cores in a shared memory model by openMP.

4.3 Results

Following the calibration strategy described above, the model reproduces the dynamics of the most important moments of Denmark. Table 5 first presents the optimal parameter set. Afterwards, we discuss the fit of the model in details.

Table 5: PARAMETERS IN THE DANISH CALIBRATION EXERCISE

Parameter	Sym.	Value
Initial share of adults in cities	δ_0	0.207
Elasticity in HK accumulation	ϕ	0.635
Part. output elasticity of land	θ	0.35
Time to give birth	ξ	0.04
Elasticity of the agricultural good	α	0.3804
Price per unit of education in terms of d	β	0.4201
Stone-Geary element for d	ε	1.7594
Time to rear a surviving child in R	ζ_R	0.0698
Time to rear a surviving child in U	ζ_U	0.0612
Costs of moving	$\bar{\kappa}$	1.5137
Average human capital in 1840	λ	1.5163
Basic human capital	ν	1.1093
Preference for educated children	ρ	0.1840
Initial TFP in R	B_0^R	1280.5
Initial TFP in U	B_0^U	6.6830
Minimum consumption of food	\bar{c}	0.0023
Exogen. growth rate of A	g	0.1011
Threshold on population size	\bar{N}	2898.7

If individuals want to leave the area of origin, in general they move from the countryside into cities, they have to pay 1.51 units of the agricultural good what represents 8.2% of the simulated income per capita in the countryside for the year 1840. As outlined above, it might be optimal to pay this cost, as both areas vary in production technology and infant mortality. In the period of stagnation the total factor productivity in the agricultural sector is 1,280.5 and hence much higher than the total factor productivity in industries (6.68) as land is normalized to 1. Once the adult population becomes large enough and exceeds the critical population size of $\bar{N} = 2,898.7$ exogenous growth starts, then, at each period, B_t^A increases by 10.1%.³⁸

³⁸2,898.7 has to be compared to the initial population size which is assumed to be 1000.

Individual preferences for consumption are characterized by three parameters: a minimum consumption in terms of the agricultural good of $\bar{c} = 0.002$, a Stone-Geary element of $\varepsilon = 1.76$ for the industrial good and an elasticity of 0.38 of the agricultural good. Hence, the subsistence level is low compared to initial incomes in both cities and countryside. The weight of agricultural goods within the utility function α is equal to 0.38 what is almost twice the weight of educated children as $\rho = 0.18$. Giving birth requires 4.0% of the available time. If the child survives, 7.0% of the time are additionally required to bring it up in the countryside and 6.1% in cities. In the absence of infant mortality, this would lead to a maximal fertility of 10 children in cities and 9 in the countryside. Initially, the human capital is distributed across dynasties with an average of 1.8254. If parents are sufficiently rich they like to educate children as they value children's human capital. For each unit of education, they have to pay 0.42 times the price of the industrial good. In absence of investment in education, human capital converges to the basic level of skills ($v = 1.11$) in the long-run.

The parameter set enables our model to reproduce the characteristics of the Danish dynamics between the years 1840 and 1940. Figure 5 shows calibration results in more detail. Solid lines represent the moments to be matched and dashed lines mark dynamics of the simulated moments.

The model reproduces the fertility transition in both areas. Coale's index drops from 0.340 (resp. 0.258) in 1840 to 0.185 (resp. 0.159) in 1940 in the countryside (resp. cities), as illustrated in Figure 5A. The model reproduces the fertility differential even if the time to bring up a surviving child is lower in cities than in the countryside. Furthermore, it traces the general dynamic of fertility: around 1880, the number of births starts to decline significantly first in cities. One period later the sharp decline sets in in the countryside. However, while we slightly underestimate the number of births at the beginning in urban areas (by 18.4%), the simulated Coale index is slightly too low in the rural area at the end of the period under investigation (by 17%).

Figure 5D highlights urbanization. The simulated share of individuals increases over time, even if the slope is below the observed moments what leads us to overestimate the urbanization rate in 1840 by 31.4% and underestimate that of 1940 by 19.5%. Notice that, on the whole period of study (1840–1940), we predict an average urbanization rate of 34.75% what is close to the real value that is 34.67%.³⁹

³⁹Those values are simple non-weighted averages.

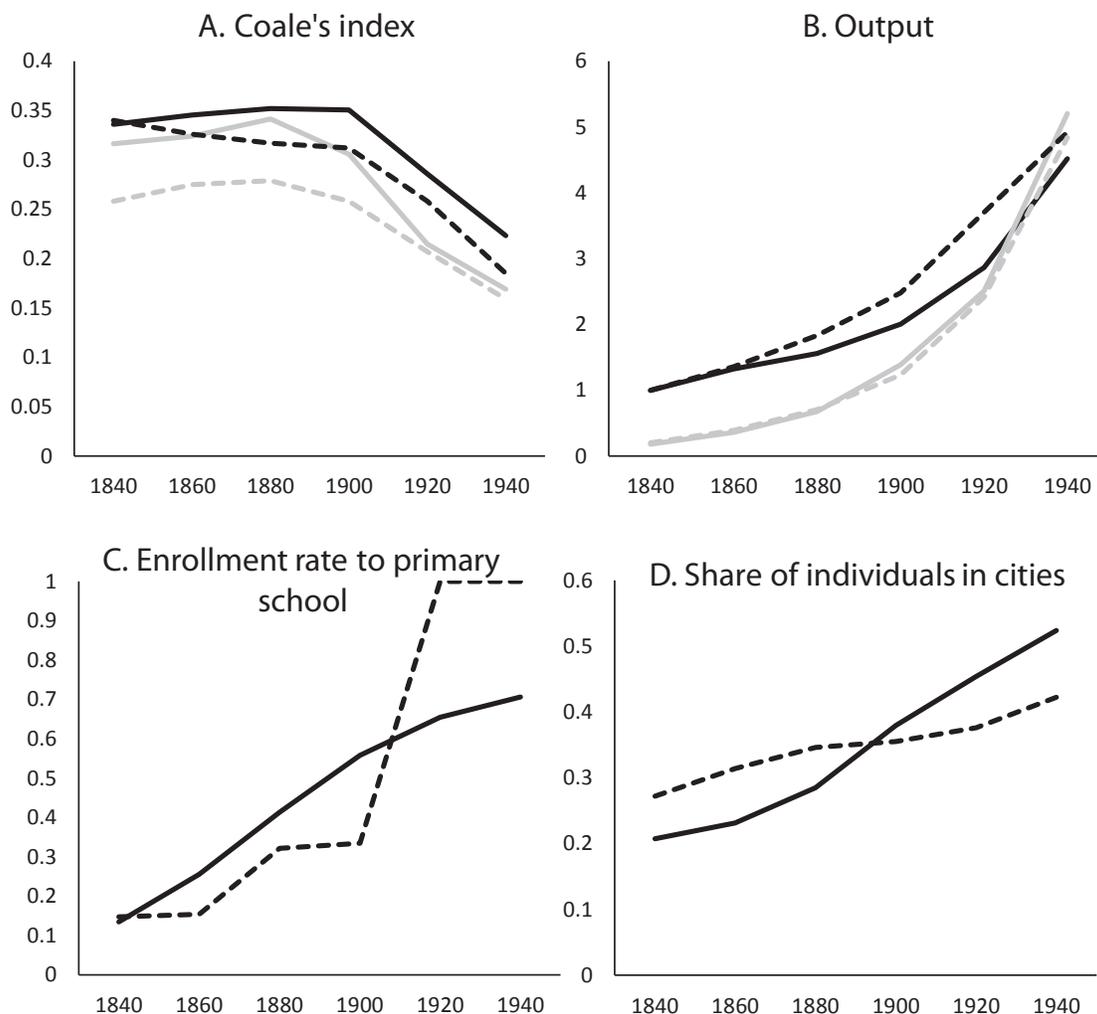


Figure 5: OBSERVED AND SIMULATED TRENDS IN DENMARK 1840–1940

Solid lines present observed trends in Denmark and dashed lines calibration results. Black lines mark rural and gray lines urban areas in Figures A and B.

Dynamics of simulated enrollment rates are even more extreme than in the smoothed estimations of observed enrollment rates to primary education, see Figure 5C. Both, estimated and simulated moments, start on a similar level of educated children in 1840 with 14.0% and 14.8%, respectively. While smoothed data follow an s-shape that reaches a value of 71% in 1940, dynamics in the model increases stepwise. Enrollment rates are almost constant in 1840–1860. Then they increase to around 32.2% in 1880. They almost remain on this level until 1900 (33.4%). Afterwards, mass education arises – enrollment rates jump dramatically and all children are educated since 1920.

Finally, we calibrate the model to fit dynamics of GDP and in particular the take-off in agricultural and industrial output. Figure 5B plots normalized output in both sectors. Initially, agricultural and industrial production fit almost perfectly. Over the periods, simulated agricultural production overestimate observed trends but with a maximum of 29% in 1920. The fit of industrial production is almost perfect as our highest gap with data is equal to 11.9% in 1840. Our model estimates that the date at which industries have produced more than the countryside is 1940 while it has been 1930 in reality, a pretty precise estimation.

4.4 Overidentification checks

To check the reliability of our calibration exercise, we compare the predicted and the observed moments of two additional time series. First of all, we have a look on the relative price presented in Section 2. To focus on the dynamics, we normalize all prices by the initial period. Figure 6 illustrates the comparison of both time series.

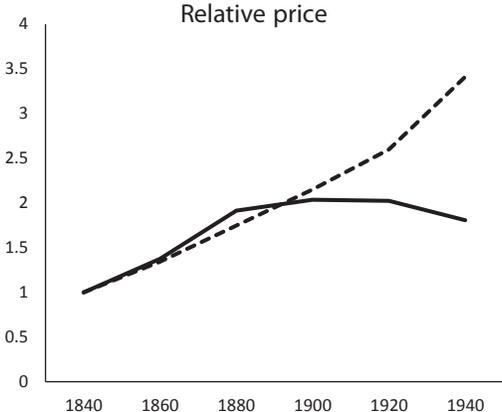


Figure 6: ESTIMATED AND SIMULATED DYNAMICS OF THE RELATIVE PRICE IN DENMARK 1840–1940

Relative price between agricultural and industrial goods. The solid line presents the normalized trend in Denmark and the dashed graph calibration results.

The relative price predicted by the model increases almost linearly. By contrast, the relative price calculated from the data has an inverse u-shaped pattern. Therefore, the predicted relative price has a good fit during the periods 1840 to 1900 but turns to be too high afterwards. The stagnation of relative prices following the industrial revolution may be explained by changes either in the varieties of goods (Grossman and Helpman,

1991) or in their quality (Aghion and Howitt, 1992), two dimensions not present in our model. A complementary explanation may be found in the history of Denmark’s international trade. At the time relative prices started to stagnate, dairy industries have shifted from trading with Britain via Hamburg to trading directly with Britain. It reduced costs of exports as middlemen were not included anymore and made Denmark both the leading country in international dairy industry (Lampe and Sharp, 2015) and more open to international trade.

Table 6: AVERAGE ANNUAL GROWTH RATES OF GDP PER CAPITA/WORKER IN %

	1840–59	1860–79	1880–99	1900–19	1920–39	1840–1939
1.	0.99	1.13	1.62	1.40	1.24	1.28
2.	0.70	0.39	1.07	1.01	1.91	1.02
3.	0.57	0.41	0.24	0.87	1.70	0.76

Rows present the following: (1.) Average annual growth rates of GDP per capita from the Maddison-project; (2.) Average annual growth rates of GDP per capita only including the agricultural and industrial sector and (3.) GDP per worker in prices of the industrial good predicted by the model. Source: The Maddison-Project (2013); own calculations.

Secondly, Table 6 presents average yearly growth rates of GDP per capita in percent from the The Maddison-Project (2013) (1. row) and compares its values to predicted growth rates of GDP per worker (3. row). Obviously, in Maddison’s database, all sectors of the economy are included what is not the case in our dataset where Transport and Communication or Commerce are for instance missing. For this reason, the second line of the table provides the growth rates of GDP per capita only including the industrial and agricultural sector as computed from our dataset. One can see that our model is characterized by good performances. Remaining differences are essentially induced by the higher population growth in the model.⁴⁰ The higher deviation between our data and those of Maddison is essentially due to our quite conservative definition of the industrial sector.

⁴⁰Remember that our calibration aims at reproducing the dynamics in fertility; only taking into account infant mortality. Excluding mortality in older ages and emigration, we overestimate population growth. On average, the simulated population increases by 1.97%, whereas the observed Danish population grew by 1.26% p.a. More precisely, the Danish population increased by around factor 3.5 between 1840 to 1940 while our population increases by factor 7.

5 Historical experiments

Up to now, we have assumed that, even if costly, movements between areas are not regulated. Freedom of settlement has not always been the rule. Institutions, like guilds, facilitated or hindered internal migration in Europe already before the industrialization (de la Croix *et al.* (2016), Ogilvie (2014)) and a country like China still implements restrictions on migrations from countryside to cities (Chan (1992), Chan and Zhang (1999)). In this section, we investigate how restrictions on internal migrations would have influenced the economic and demographic dynamics of Denmark. Answering this question in Section 5.1 will help us assessing the importance of the rural exodus in the process of economic and demographic modernization, its role for the transition from low to mass education and for the fertility transition. Then, in line with papers like Lagerlöf (2003), Weisdorf (2004) Strulik and Weisdorf (2014) or Bar and Leukhina (2010), we discuss the impact of IMRs on the demographic and economic dynamics of Denmark.

5.1 Restrictions on internal migration

To highlight the role of Denmark’s rural exodus we proceed in two steps. We first simulate the extreme case that does not allow any dynasty to move. All individuals have to stay at their place of birth. The second exercise is less extreme and implements a static comparative exercise on κ .

No-movement scenario

Excluding any kind of movement between urban and rural areas leads to dramatic changes in economic and population dynamics. It does not solely alter urbanization. In fact, the rural exodus appears as a key element for the fertility transition, the democratization of education as well as the increase and distribution of welfare. Using parameters estimated in Section 4, Figures 7–9 compare the benchmark case to the “no-movement” scenario.

The “no-movement” scenario strongly magnifies the urban and rural fertility differential as illustrated in Figure 7A. Coale’s index declines during the whole period in cities and hence starts to decrease much earlier than in the benchmark. By contrast, fertility in the countryside does not experience any kind of transition from high to low fertility. Indeed, Coale’s index increases over time and the number of births is higher in 1940 than

in 1840. While the number of births is 1.71 times higher in rural than in urban areas in 1840, the ratio increases up to 4.79 in 1940. Obviously, this changes the dynamics of urbanization and population size. The missing number of births in cities leads to a shrinking share of individuals in urban areas, what we call a counter-urbanization. Individuals in cities only count for 0.1% of the population in 1940. Because “killer cities” cannot absorb the rural population anymore and because this rural population does not refrain its fertility anymore, the total population growth rate exceeds that of the benchmark scenario. In 1940, the total population size is 9.83 times larger in the “no-movement” scenario than in the benchmark. Would it mean that, at the end, what we call the countryside becomes cities?

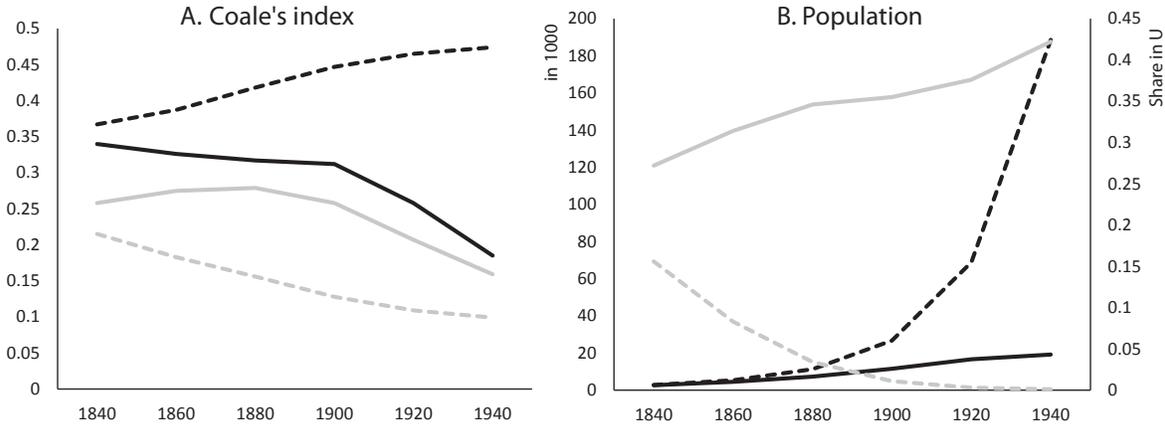


Figure 7: DANISH POPULATION DYNAMICS IN THE BENCHMARK AND “NO-MOVEMENT” SCENARIO 1840–1940

Solid lines present trends of the benchmark calibration and dashed lines the “no-movement” simulation. Black lines mark rural and gray lines urban areas in Figure 7A. The left y-axis in Figure 7B mark total populations (black) and the right y-axis the share in the urban area (gray).

The answer could be yes if the GDP per capita in the countryside would have increased over the period what is not true despite exogenous growth takes place earlier than in the benchmark scenario. Indeed, as shown in Figure 8, even if total agricultural GDP (in prices of the industrial good) accelerates due to the faster growth of rural population, the 1940 GDP per capita in the countryside under the “no-movement scenario” represents only 14.6% of its value under the benchmark scenario. Said differently, without migration, agricultural workers would have pauperized. Results are reversed for industrial production as GDP per capita would have been multiplied by 10.6 while total GDP would have represented only 17.6% of its benchmark value. Notice that, at the

aggregated level, the total GDP would have been much higher than in the benchmark scenario, a window-dressing as almost entire population would have pauperized.

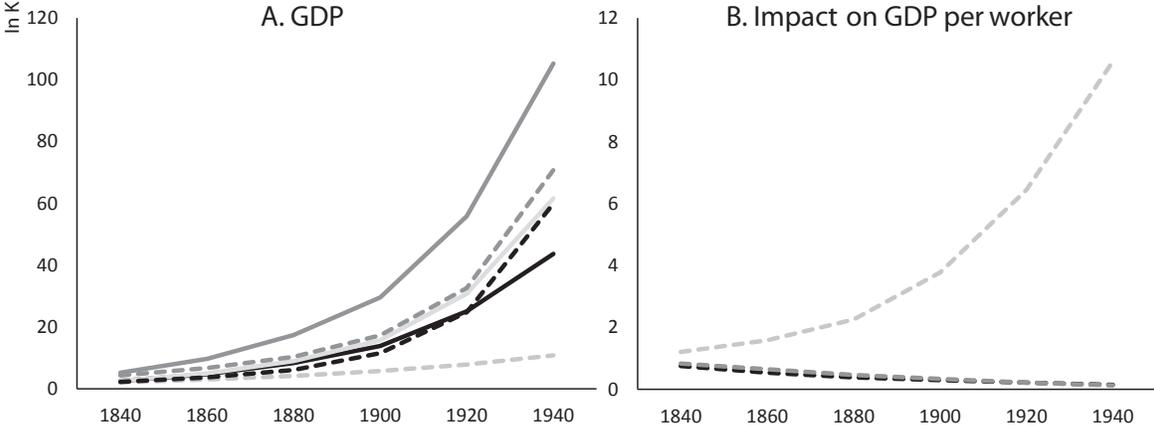


Figure 8: IMPACT OF A MISSING RURAL EXODUS ON GDP AND GDP PER WORKER IN DENMARK 1840–1940

Solid lines present trends of the benchmark calibration in Denmark and dashed lines the “no-movement” scenario. Black marks rural areas, gray urban ones and dark gray total Denmark.

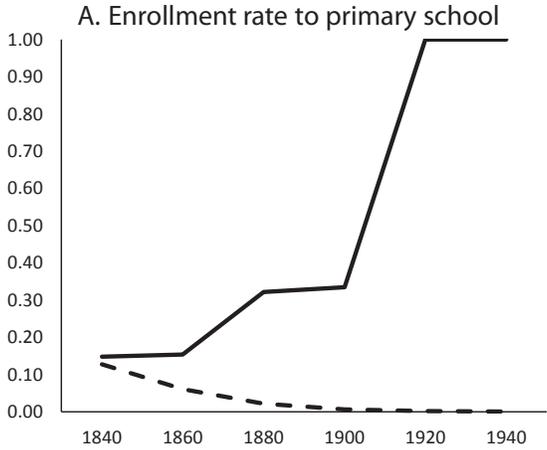


Figure 9: PRIMARY EDUCATION IN THE BENCHMARK AND “NO-MOVEMENT” SCENARIO IN DENMARK 1840–1940

Solid lines present trends of the benchmark calibration in Denmark and dashed lines the “no-movement” scenario.

As GDP per worker as well as real wages shrink,⁴¹ parents are not rich enough to educate their children and the take-off in primary education disappears (see Figure 9).

⁴¹Dynamics of real wages are presented in Appendix C.2.

Static comparative on κ

Here, we analyze the case where we gradually raise the costs of leaving the birthplace by factor 2 to 5. The increasing cost of moving to cities intensifies the demographic transition in cities step by step (Figure 10). In countryside, the higher cost of migrating may either delay or annihilate the fertility transition what can be explained by the dynamics of education. When κ is multiplied by 2, the massification of education still arises but later (by 1920 instead of 1900). Once κ is multiplied by 3, education remains confined to a minority of persons living in cities, no trade-off between quality and quantity of children takes place in the countryside which finally does not experience any form of fertility transition. At the aggregated level, education rates follow an inverted u-shaped pattern, see Figure 11. Obviously, all these effects strongly influence the dynamics of population size. The stall in the fertility transition of the countryside translates into a larger total population size. Furthermore, doubling κ annihilates the urbanization process while for larger increases, it even provokes a counter-urbanization. Interestingly enough, the absence of urbanization⁴² does not mean an absence of rural exodus, as shown in Figure 11. By contrast, one can see that the total number of movers from countryside to cities even increase when κ increases (size effect) but the weight of these persons within the total population is smaller. In other words, we highlight here that having a rural exodus is not a sufficient condition to get an urbanization of the society.

Turning to the analysis of GDP dynamics, our first finding is that total GDP growth slows down in the long-run when costs to move increase. In the most extreme case, where κ is multiplied by 5, the difference to the benchmark case is highest in 1880 when total GDP reaches only 85.3% of the benchmark GDP. In 1940, it is only 2.9% smaller than in the benchmark case what may be perceived as quite limited; nevertheless, this result hides dramatic changes. Still in the most extreme case where κ is multiplied by 5, the GDP per worker in cities in 1940 is increased by 30.1% while in the rural area, it shrinks to 55.7% of the benchmark case.⁴³ To evaluate the violence of these adjustments, one should remember that in the benchmark scenario, the 1940 GDP per worker in cities was only 1.78 times higher than in the countryside, the value more than

⁴²Urbanization defined as a higher proportion of the population living in cities.

⁴³In the scenario with 5 times κ the total working population exceeds critical value one period earlier. Thus, exogenous technological process starts earlier and pushes economic growth. Simultaneously, the related GDP per worker increases and enables a higher fraction of individuals to move in 1900, as observable in Figure 11B.

doubles (4.17) in the scenario 5 times κ .⁴⁴

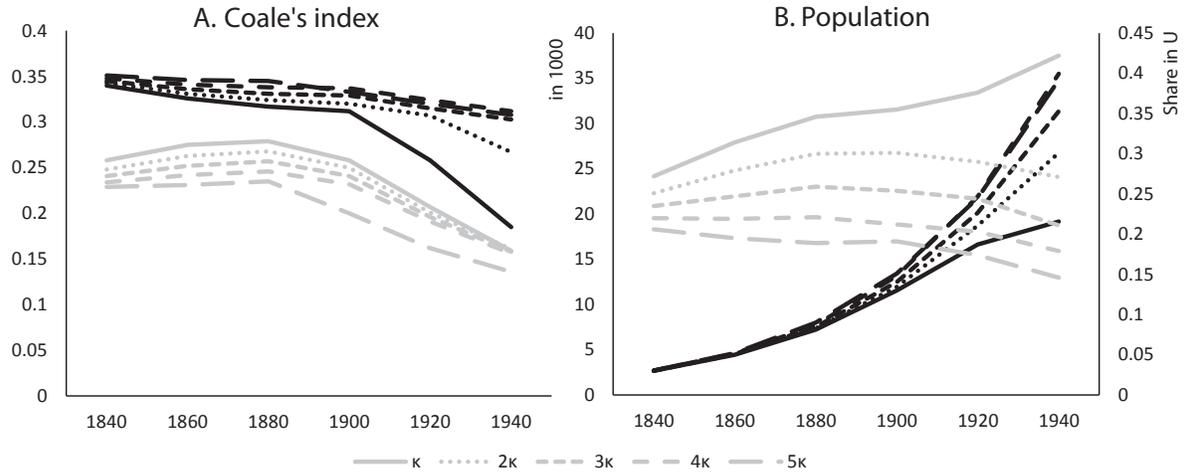


Figure 10: DANISH POPULATION DYNAMICS IN THE BENCHMARK CALIBRATION AND VARYING COSTS TO MOVE 1840–1940

Solid lines present trends of the benchmark calibration and the different types of dashed lines varying costs to move. Black lines mark rural and gray lines urban areas in Figure 7A. The left y-axis in Figure 7B mark total populations (black) and the right y-axis the share in the urban area (gray).

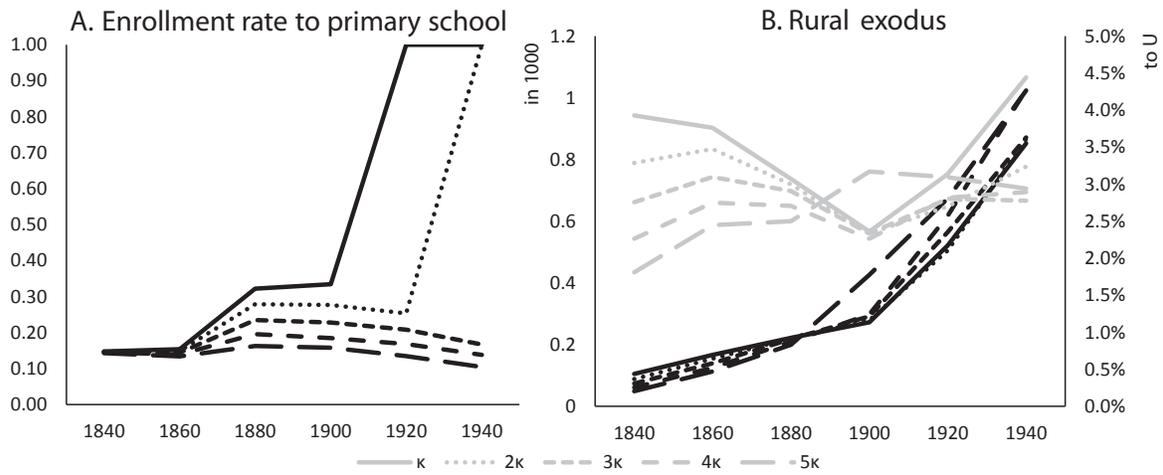


Figure 11: EDUCATION AND RURAL EXODUS IN THE BENCHMARK SCENARIO COMPARED TO HIGHER MOVING COSTS IN DENMARK 1840–1940

Solid lines present trends of the benchmark calibration and the different types of dashed lines the varying costs to move. In Figure 11B, black lines mark total number of movers from R to U plotted on the left y-axis and gray lines their share on the total population (right y-axis).

⁴⁴As in the “no-movement” scenario, total GDP in the agricultural sector is higher than in the benchmark because of the demographic expansion of rural areas.

On the whole, urban dynasties are the winner of increases of κ . The missing immigration from the countryside limits production of the industrial good and thus increases its relative price. Welfare growth in cities accelerates, driven by the quality-quantity trade-off. Fertility drops faster and parents invest more in education if the rural exodus is restricted. A development that clearly strengthens inequality between cities and the countryside.

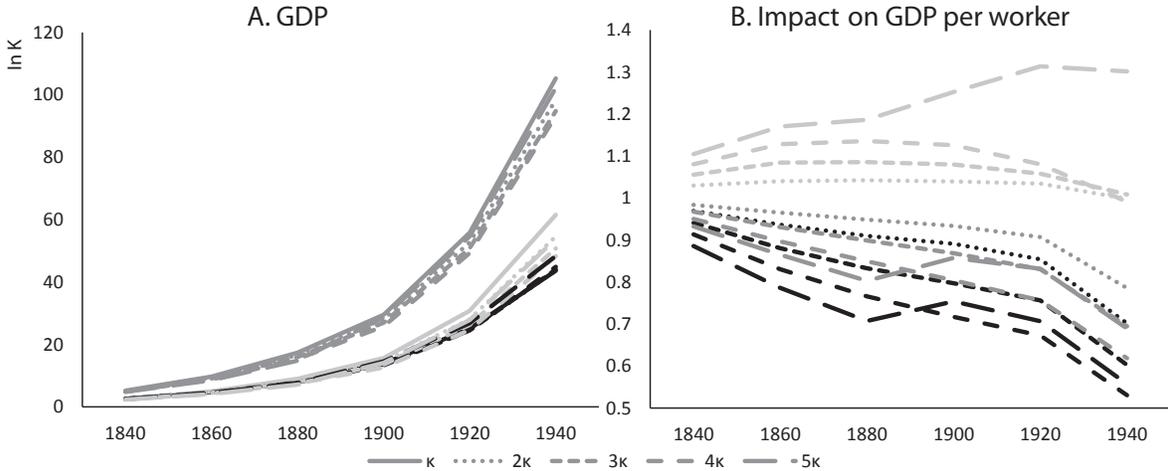


Figure 12: IMPACT OF HIGHER COST TO MOVE ON GDP AND LOSS OF GDP PER WORKER IN DENMARK 1840–1940

Solid lines present trends of the benchmark calibration and the different types of dashed lines simulations with varying costs to move. Black marks rural areas, gray urban ones and dark gray total Denmark.

Beyond doubt this numerical exercise underlines that the rural exodus does not only matter for urbanization. Indeed, limiting the rural exodus highlights that all dimensions of economic and demographic development are altered. A missing rural exodus expedites the demographic transition in urban areas. By contrast, this kind of transition may simply disappear in rural areas. The rural exodus supports the transition to mass education and a more egalitarian distribution of welfare. If the exodus is not allowed, we still have an economic take-off – but only for the elite. Thus, the rural exodus first of all acts as a mechanism to prevent an increasing inequality and secondly avoids situations, where elites become inefficiently small. It implies an important chance for “talented mass” from the countryside to move to cities and thus to employ resources efficiently.

5.2 Health traps

Young children suffered particularly from unhealthy and unhygienic situations in European “killer cities”. To evaluate the impact of infant mortality rates on the rural exodus and general economic and demographic dynamics, we freeze IMRs on their high level of 1840 and simulate model’s dynamics. In a second step, we allow an asymmetric dynamic in IMRs. The countryside experiences observed improvements in IMRs, while the survival probability stagnate on the initial level in cities.

Freezing IMRs to their initial values has almost no impact on fertility. Dashed and solid lines in Figure 13A coincidence. The effect of the infant mortality on the share in urban areas is likewise limited. However, since IMRs stagnates on the high level and the number of births is almost unchanged, population growth is obviously reduced. Total population is notably below the benchmark case as soon as 1920. However, the most remarkable change occurs in the transition to mass education, which takes place one period earlier (in 1900) as shown in Figure 14A. The lower survival probability reduces the critical wage ($\bar{\omega}_t$) that enables parents to educate. A smaller fraction of births survives and needs to be educated. The impact on the rural exodus is limited, too. The share of individuals moving to cities is below the benchmark calibration in 1900 and higher afterwards.

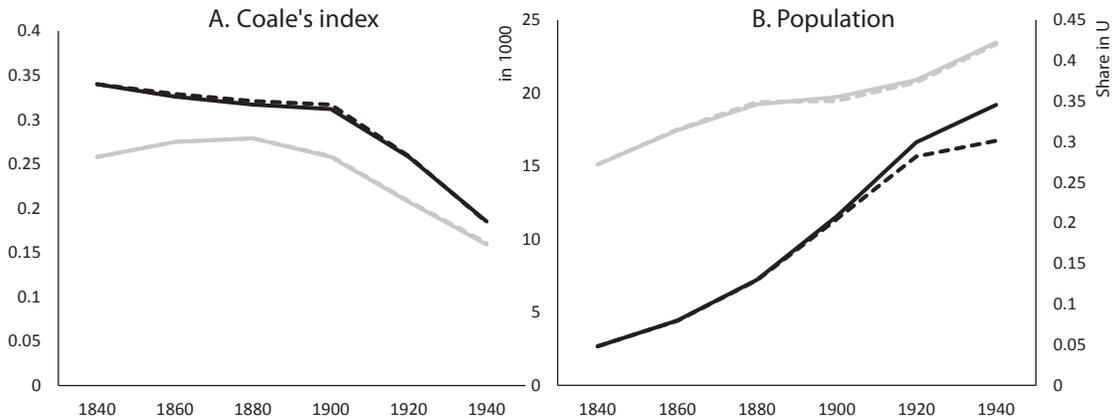


Figure 13: DANISH POPULATION DYNAMICS IN THE BENCHMARK AND CONSTANT IMRS SCENARIO 1840–1940

Solid lines present trends of the benchmark calibration and dashed lines the constant IMR simulation. Black lines mark rural and gray lines urban areas in Figure 13A. The left y-axis in Figure 13B shows total populations (black) and the right y-axis the share in the urban area (gray).

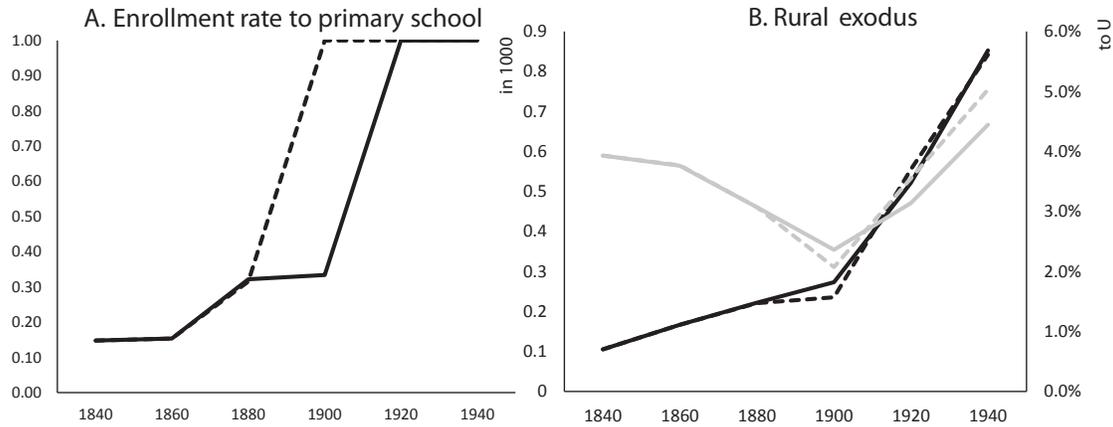


Figure 14: EDUCATION AND RURAL EXODUS IN THE BENCHMARK AND CONSTANT IMRS SCENARIO IN DENMARK 1840–1940

Solid lines present trends of the benchmark calibration in Denmark and dashed lines the constant IMR simulation. In Figure 14B, the left y-axis marks the total number of moving individuals (black) and the right y-axis their share on the total population (gray).

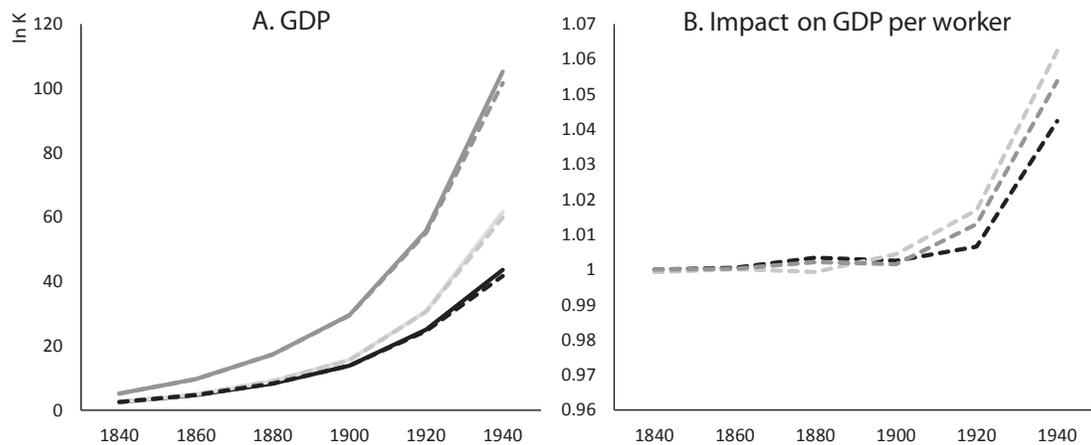


Figure 15: GDP AND GDP PER WORKER IN THE BENCHMARK AND CONSTANT IMRS SCENARIO IN DENMARK 1840–1940

Solid lines present trends of the benchmark calibration in Denmark and dashed lines the constant IMRs simulation. Black marks rural areas, gray urban ones and dark gray total Denmark.

Excluding improvements in survival probabilities has a very limited effect on the total GDP, as well as on the GDP in both sectors. Until 1900, the impact on the GDP per worker is also very limited. Afterwards, values are higher compared to the benchmark case (Figure 15). Gains of 4.2% in R and 6.2% in U compensate for the lower survival probability. This highlights the persistence of Malthusian mechanisms within this simulated economy.

If only the countryside experiences a reduction in IMRs, the most evident difference occurs in urbanization.⁴⁵ The share of individuals in cities shrinks such that a higher rural exodus would be required to increase urbanization. We can observe that the income differential between cities and countryside rises and partially compensates individuals in cities for the higher IMR what incites them to leave the countryside. Nevertheless, it is not sufficient to hold urbanization rates unchanged. The asymmetric improvements in IMRs reduce GDP per worker in rural areas after 1880. Simultaneously, GDP per worker in cities exceeds the benchmark case.

Overall, effects of the IMRs are quite limited compared to policies that reduce or eliminate the rural exodus. Anyway, the much stronger impact of the rural exodus might be related to simplifications in the model. There is no uncertainty in the model and our utility function is log-linear.⁴⁶ In particular an introduction of uncertainty related to child survival might alter findings in this section.

6 Conclusion

Grounding on the Unified Growth Theory, this paper is the first one evaluating the role that rural exodus played in the process of economic development in Western Europe. In the specific case of Denmark, we show that favorable conditions of exodus are a pre-requirement for both the fertility transition and the economic take-off, what is not the case of the health revolution making IMRs recede strongly. We estimate that doubling the cost of migration would not have been sufficient to prevent the take-off to modern growth but it would have delayed this latter and weakened its intensity. It would also have magnified the inequalities between cities and the countryside, inequalities initiated by the industrial revolution. While GDP per capita in cities would have remained unchanged, farmers would have severely impoverished. The pattern of the fertility transition would not have been altered in cities while it would have been weakened and delayed by 20 years in the countryside. We also show that, at some point, making the exodus even more costly would simply have put a stop to the Danish development, the countryside would have pauperized while a vanishing elite would have haunted dying cities.

⁴⁵Figures are presented in Appendix C.1.

⁴⁶Adams (2015) discusses the issue related to the choice of the utility function.

Importantly enough, this paper is among the few applying a full-fledged calibration-simulation strategy to a unified theory of growth and it is the first using the simulated method of moments. It offers new perspectives to evaluate the respective accuracy of refinements of the UGT, refinements that followed Galor and Weil (2000) seminal contribution.

References

- ADAMOPOULOS, T. (2008). Land inequality and the transition to modern growth. *Review of Economic Dynamics*, **11** (2), 257–282.
- ADAMS, J. J. (2015). Urbanization, long-run growth, and the demographic transition.
- AGHION, P. and HOWITT, P. (1992). A model of growth through creative destruction. *Econometrica*, **60** (2), 323–351.
- BAIROCH, P. (1997). *Victoires et déboires: histoire économique et sociale du monde du XVIIe siècle à nos jours*, vol. I, II and III. Gallimard.
- BAR, M. and LEUKHINA, O. (2010). Demographic transition and industrial revolution: A macroeconomic investigation. *Review of Economic Dynamics*, **13** (2), 424–451.
- BAUDIN, T., DE LA CROIX, D. and GOBBI, P. E. (2015). Fertility and childlessness in the united states. *The American Economic Review*, **105** (6), 1852–1882.
- BENAVIDES, A. and RIDDLE, P. (1988). The expansion of primary education, 1870-1940: Trends and issues. *Sociology of Education*, pp. 191–210.
- BERTINELLI, L. and BLACK, D. (2004). Urbanization and growth. *Journal of Urban Economics*, **56** (1), 80–96.
- BOCQUIER, P. 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., EDVINSSON, S. and ROGERS, J. (2002). Illegitimacy, infant feeding practices and infant survival in sweden 1750–1950: A regional analysis. *Hygiea Internationalis*, **3** (1), 13–52.
- CENTRALBYRÅN, K. S. (1914). Statistisk årsbok för sverige 1914. *Stockholm (and subsequent editions)*.

- CENTRALBYRAN, S. (1969). Historisk statistik for sverige. *Del I. Befolkning. Andra upplagan 1720–1967.*
- (1996). Tätorter 1995. *Statistiska meddelanden Be16SM9601.*
- 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. *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.
- CHAN, K. W. (1992). Economic growth strategy and urbanization policies in china, 1949–1982*. *International Journal of Urban and Regional Research*, **16** (2), 275–305.
- and ZHANG, L. (1999). The hukou system and rural-urban migration in china: Processes and changes. *The China Quarterly*, **160**, 818–855.
- CHARBONNEAU, P. (2002). An introduction to genetic algorithms for numerical optimization. *NCAR Tech. Note TN-450+ IA*, 74pp.
- CHESNAIS, J.-C. et al. (1992). *The demographic transition: Stages, patterns, and economic implications.* Oxford University Press.
- CLARK, G. (2002). Land rental values and the agrarian economy: England and wales, 1500–1914. *European Review of Economic History*, **6** (3), 281–308.
- , CUMMINS, J. and SMITH, B. (2012). Malthus, wages, and preindustrial growth. *The journal of economic history*, **72** (02), 364–392.
- COALE, A. (1969). The decline of fertility in europe from the french revolution to world war ii. In C. L. J. Behrman, S.J. and R. Freedman (eds.), *Fertility and Family Planning: A World View.*, University of Michigan Press, pp. 3–24.
- 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. (2016). Clans, guilds, and markets: Apprenticeship institutions and growth in the pre-industrial economy. *IREES Discussion Paper 2016-09.*

- , LINDH, T. and MALMBERG, B. (2008). Swedish economic growth and education since 1800. *Canadian Journal of Economics/Revue canadienne d'économique*, **41** (1), 166–185.
- DEMOGRAPHIC DATA BASE (). The demographic data base on Swedish historical population statistics. CEDAR, Umeå University.
- DEPARTEMENT, D. S. (1966). Befolkningsudvikling og sundhedsforhold 1901-60. *Statistiske undersøgelser*, **19**.
- 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.
- DYSON, T. (2011). The role of the demographic transition in the process of urbanization. *PoPulation and develoPment review*, **37** (s1), 34–54.
- FEHR, H. and KINDERMANN, F. (2015). *Introduction to Computational Economics using Fortran*. mimeo.
- 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.
- FRANCK, R., GALOR, O. *et al.* (2015). Industrialization and the fertility decline.
- FUNKHOUSER, E. (1998). Changes in the returns to education in costa rica. *Journal of Development Economics*, **57** (2), 289–317.
- GALOR, O. (2005). The demographic transition and the emergence of sustained economic growth. *Journal of the European Economic Association*, **3** (2-3), 494–504.
- (2011). *Unified growth theory*. Princeton University Press.
- and WEIL, D. N. (1999). From malthusian stagnation to modern growth. *The American Economic Review*, **89** (2), 150–154.
- and — (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American economic review*, pp. 806–828.

- 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.
- GROSSMAN, G. M. and HELPMAN, E. (1991). Quality ladders in the theory of growth. *The Review of Economic Studies*, **58** (1), 43–61.
- HARRIS, J. R. and TODARO, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review*, **60** (1), 126–142.
- HAYAMI, Y., RUTTAN, V. W. *et al.* (1971). *Agricultural development: an international perspective*. Baltimore, Md/London: The Johns Hopkins Press.
- 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.
- JOHANSEN, H. C. (1985). *Dansk historisk statistik, 1814-1980: Danish historical statistics, 1814-1980*. Gyldendal Copenhagen, Denmark.
- (2002). *Danish population history 1600-1939*. Hans Christian Andersen Center, University of Southern Denmark, University Press of Southern Denmark.
- KINDLEBERGER, C. P. (1951). Group behavior and international trade. *The Journal of Political Economy*, pp. 30–46.
- KNODEL, A. J. (1974). *The decline of fertility in Germany, 1871-1939, Office of Population Research and others*, vol. 2. Princeton University Press.
- KÖGEL, T. and PRSKAWETZ, A. (2001). Agricultural productivity growth and escape from the malthusian trap. *Journal of Economic Growth*, **6** (4), 337–357.
- 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.

- LASSEN, A. (1965). *Fald og fremgang: Traek af befolkningsudviklingen i Danmark 1645-1960*. Universitetsforlaget.
- (1966). The population of denmark, 1660–1960. *Scandinavian Economic History Review*, **14** (2), 134–157.
- LESTHAEGHE, R. J. (1977). *The decline of Belgian fertility, 1800-1970*. Office of Population Research and others, Princeton University Press.
- MATTHIESSEN, C. W. (1985). Danske byers folketal 1801-1981. *Statistiske Undersøgelser Nr. 42*.
- MATTHIESSEN, P. C. (1984). Some reflections on the historical and recent fertility decline in denmark. *Scandanavian Population Studies*, **6** (2).
- MEYER, J. W., RAMIREZ, F. O. and SOYSAL, Y. N. (1992). World expansion of mass education, 1870-1980. *Sociology of education*, **65** (2), 128–149.
- MOKYR, J., VOTH, H.-J. *et al.* (2012). Understanding growth in europe, 1700–1870: Theory and evidence. *Journal of Economic Sociology*, **13** (5), 57–102.
- MORRISSON, C. and MURTI, F. (2009). The century of education. *Journal of Human Capital*, **3** (1), 1–42.
- MOSK, C. (1980). Rural-urban fertility differences and the fertility transition. *Population studies*, **34** (1), 77–90.
- OGILVIE, S. (2014). The economics of guilds. *The Journal of Economic Perspectives*, **28** (4), 169–192.
- and CERMAN, M. (1996). *European proto-industrialization: an introductory handbook*. Cambridge University Press.
- OTTAVIANO, G. and THISSE, J.-F. (2004). Agglomeration and economic geography. *Handbook of regional and urban economics*, **4**, 2563–2608.
- POWELL, M. J. (2002). Uobyqa: unconstrained optimization by quadratic approximation. *Mathematical Programming*, **92** (3), 555–582.
- RAVENSTEIN, E. G. (1885). The laws of migration. *Journal of the Statistical Society of London*, **48** (2), 167–235.

- (1889). The laws of migration. *Journal of the royal statistical society*, **52** (2), 241–305.
- SCHÖN, L., KRANTZ, O. *et al.* (2012). Swedish historical national accounts 1560-2010. *Lund Papers in Economic History*, **123**.
- SOYSAL, Y. N. and STRANG, D. (1989). Construction of the first mass education systems in nineteenth-century europe. *Sociology of education*, **62** (4), 277–288.
- SQUICCIARINI, M. P. and VOIGTLÄNDER, N. (2014). Human capital and industrialization: Theory and evidence from 18th century encyclopédie subscriptions. *UCLA Anderson, manuscript*.
- STRULIK, H. and WEISDORF, J. (2014). How child costs and survival shaped the industrial revolution and the demographic transition. *Macroeconomic Dynamics*, **18** (01), 114–144.
- and WERNER, K. (2012). Life expectancy, labor supply, and long-run growth: Reconciling theory and evidence. *Center for European Governance and Economic Development Research Discussion Paper, no. 141*.
- and — (2013). 50 is the new 30-long-run trends of schooling and retirement explained by human aging. *Journal of Economic Growth*, pp. 1–23.
- THE MADDISON-PROJECT (2013). <http://www.ggdc.net/maddison/maddison-project/home.htm>.
- TODARO, M. P. (1969). A model of labor migration and urban unemployment in less developed countries. *The American economic review*, **59** (1), 138–148.
- TREADWAY, R. (1980). Princeton european fertility project - master data file.
- VOIGTLÄNDER, N. and VOTH, H.-J. (2013a). How the west “invented” fertility restriction. *American Economic Review*, **103** (6), 2227–64.
- and VOTH, H.-J. (2013b). The three horsemen of riches: Plague, war, and urbanization in early modern Europe. *The Review of Economic Studies*, **80**, 774–811.
- WEISDORF, J. L. (2004). From stagnation to growth: Revisiting three historical regimes. *Journal of Population Economics*, **17** (3), 455–472.

- WESTERGAARD, H. (1922). *Economic Development in Denmark before and during the World War*. Oxford: Clarendon.
- WOODS, R. (2003). Urban-rural mortality differentials: An unresolved debate. *Population and development review*, **29** (1), 29–46.
- WOODS, R. I., WATTERSON, P. A. and WOODWARD, J. H. (1988). The causes of rapid infant mortality decline in england and wales, 1861–1921 part i. *Population studies*, **42** (3), 343–366.

A Data

A.1 General procedure

To develop a dataset that enables us to check our theory, we gathered data from different sources, like original census data, books and available databases. Focusing on data over more than two centuries entails two main challenges: *(i)* definitions vary as they originate from different sources. Additionally, even those from official statistics changed over time. This issue is of particular interest with respect to the separation of urban and rural areas. We document and discuss applied definitions for Denmark in A.2.1 and Sweden in A.3.2. *(ii)* Countries' borders changed. In a next step, we combine the sources to construct time series, whereby we try to minimize biases due to definitions. Often different sources provide varying numbers in the same year. In this case, we choose those values that are either identified as the less biased in the literature, originate from official statistics or reduce biases in order to use the probably most reliable value.

Even if Danish data are rich, several observations are missing such that we generally have incomplete time series. In order to close gaps, we apply (linear-) interpolation. As we are not interested in short-term fluctuations, this approach is still defensible. We then smooth data. As time series include no missing values anymore, we apply the Hodrick-Prescott filter (Hodrick and Prescott, 1997). Finally, we pick values every 20 years and try to match these values.

A.2 Denmark

A.2.1 Reclassification of rural and urban areas

It is important to notice that official data on urban and rural populations in Denmark are subject to many changes in methods and definitions what limits the comparability over time. Beyond changes of borders, identification of urban and rural areas changed several times according to Matthiessen (1985). During the 19th century cities were classified following administrative criteria. Definitions differed across provinces as no standardized rules existed. In general, all municipalities with rights to trade and market places as well as Copenhagen were classified as urban areas. These cities enjoyed a certain autonomy as they were allowed to trade. Once a part of the parish was urban, the whole is treated as urban in the data.

Even after 1857 when trade was deregulated, the definition of cities associated with certain rights were kept until the 20th century. This leads to quite conservative urban population numbers. At the beginning of the 20th century the concept of cities changed and the number step by step increased. A minimum number of individuals was introduced that varied between 50 and 250 persons. Furthermore, introducing the concept of suburban areas in 1916, criteria became again less objective. No minimum population was required, only a continuous house construction. Hence, some suburban areas were quite similar to villages.

A.2.2 Distribution of the Danish population by place of birth and residence

Quality of Danish data is too poor to measure directly migration flows in the 19th and at the beginning of the 20th century. But we can use the distribution of the Danish population according to the place of birth and residence to provide some evidence on immigration and the role of the rural exodus (Johansen, 2002).

In Denmark, we observe the following area specific immigration pattern. A high share of individuals in the capital is either born in provincial cities or the countryside. Individuals in provincial towns are as well often born in rural districts. In several years, more than 40% of the population in provincial towns and up to one quarter in the capital were born in the rural areas. By contrast, only 6% to 8% of the population in the capital was born abroad. The countryside experienced almost no immigration; neither from abroad nor cities in Denmark. Table 7 documents the resident population in 1850, 1880, 1901, 1911, 1921, 1930 and 1940 by place of birth. Although we do not know Danish emigration, the distribution of the resident population by place of birth indicates the significant role of the rural exodus. Individuals moved from the villages in surrounding cities or the capital. Furthermore, those born in the provincial towns moved to the capital. Finally, depending on the economic conditions, Denmark lost some inhabitants abroad (Johansen, 2002).

Table 7: POPULATION BY PLACE OF BIRTH AND RESIDENCE IN %

	Place of birth					
	Copen- hagen	Town		Rural parish		Abroad
		Islands	Jutland	Islands	Jutland	
1850						
Copenhagen	63	8	4	15	3	7
Town (islands)	6	56	5	26	3	4
Town (Jutland)	3	2	61	1	28	5
Rural parish (islands)	2	2	1	93	1	1
Rural parish (Jutland)	0	0	2	1	95	2
1880						
Copenhagen	54	10	5	20	3	8
Town (islands)	5	53	2	35	1	4
Town (Jutland)	3	3	51	3	35	5
Rural parish (islands)	3	3	1	90	1	2
Rural parish (Jutland)	0	0	3	1	94	2
1901						
Copenhagen	53	15		24		8
Town	3	51		41		5
Rural parish	2	4		91		3
1911						
Copenhagen	55	15		22		8
Town	4	52		40		4
Rural parish	3	5		89		3
1921						
Copenhagen	55	16		22		7
Town	4	51		42		3
Rural parish	2	6		89		3
1930						
Copenhagen	56	15		23		6
Town	4	50		43		3
Rural parish	3	7		88		2
1940						
Copenhagen	55	16		23		6
Town	4	50		43		3
Rural parish	4	8		86		2

Source: Johansen (2002).

A.2.3 Age structures and differential fertility in Denmark

In general, crude birth rates (CBR) are used to study the demographic transition, see e.g. Dyson (2011) or Bocquier and Costa (2015). Figure 16A presents the trend of the Danish CBR between 1840–1940. It suggests that no fertility differential existed until the onset of the fertility transition around 1890. Afterwards, rural CBRs exceed urban ones. If we only consider women in fertile age, here age 20–50, the pattern significantly changes. Measuring fertility by the so called Generalized Fertility Rate (GFR) in Figure 16B shows a clear fertility differential. Fertility in the countryside was much higher than in the cities. The varying findings are induced by the differences in the age structures illustrate in Figure 17 using the example of Denmark 1901. The share of women in fertile age on the total population in the cities (23.5%) significantly exceeds the share in the countryside (18.0%). Focusing on the role of the rural exodus, this difference deserves our particular attention, as it is widely induced by the exodus. Figure 17 indicates a flow of women (at the beginning of or) in fertility age from the countryside into the cities.

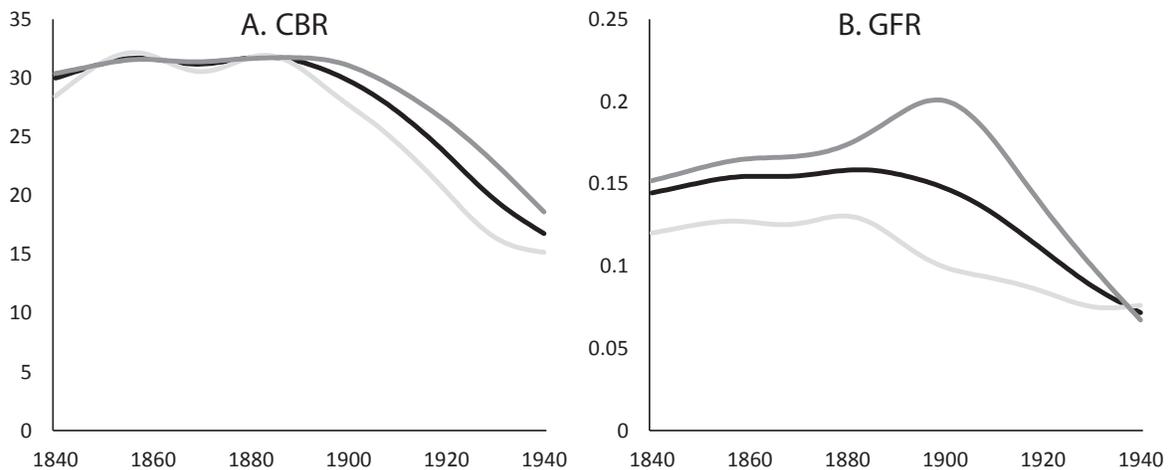


Figure 16: FERTILITY IN DENMARK 1840–1940 MEASURED BY CBR AND GFR

Black lines mark total Denmark, dark gray rural and light gray urban areas in Denmark. Figure 16A presents the number of births per 1000 persons (CBR) while 16B illustrates the number of births per women in age 20–50 (GFR). Sources: Johansen (2002) and Departement (1966); own calculations.

Since the exodus contributed to varying age structures, it is important to rely on a fertility indicator that controls for the age structure. While the GFR only excludes men and women not in fertile age, Coale’s index additionally offers a way to control

for differences within fertile age. Thus, we rely on Coale’s index which also indicates a significant fertility differential. Comparing both indicators, Coale’s index and the GFR, shows that age-specific fertility matters. The peak in GFRs around 1900 does not exist in Coale’s index.

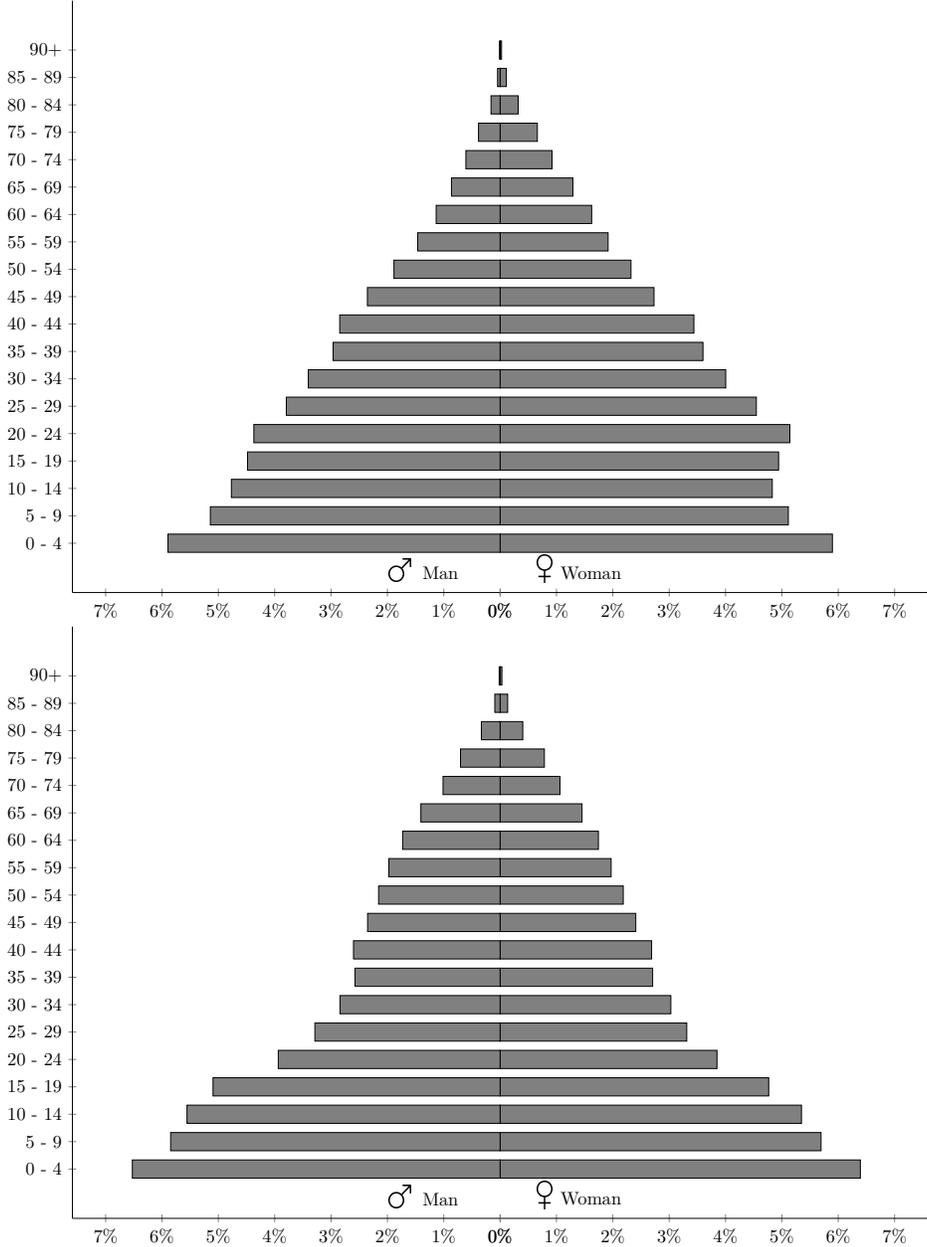


Figure 17: AGE STRUCTURES OF DENMARK’S URBAN AND RURAL POPULATION 1901
 Age structures of the urban (top) and rural (bottom) population by sex in Denmark 1901. Source: Departement (1966); own calculations.

A.3 Sweden

A.3.1 Swedish historical data

To get a picture of demographic dynamics in urban and rural areas in a second country, we present the Swedish case from the mid-18th to the mid-20th century. We compute population size, Coale’s indices of fertility and infant mortality rates in both areas of Sweden. We combine information from *Historisk statistik för Sverige* (Centralbyran, 1969), the Demographic Data Base on Swedish Historical Population Statistics (SHiPs-Demographic Data Base) offered by the Umeå University, national statistic yearbooks and the Princeton European Fertility Project.

Demographic side

Historisk statistik för Sverige (Centralbyran, 1969) offers information on rural (“Landsbygd”) and urban (“Städer”) population sizes. From 1800 to 1855, information is available every 5 years while more recent population data (1856 to 1962) are given by year. To approximate earlier developments, we use SHiPs and compare the urban and rural populations. The SHiPs-Demographic Data Base contains a rich set of demographic variables for the period 1749–1859 for up to around 3000 Parishes and a classification of these Parishes by their urban vs rural status. This status is institutionally rather than demographically defined.⁴⁷ The proportion of the Swedish population that is covered by SHiPs data strongly varies from one year to another such that we only use data for the years covering at least 60% of the Swedish population.

We obtain Figure 18B in which we illustrate the dynamics of the Swedish population size as well as the evolution of the proportion of persons living in cities, our approximation of urbanization. Population size increased continuously during the whole period under the scope of our analysis, an acceleration occurred around 1815. The share of the population in cities remained constant over the first 100 years of observations. After 1850 urbanization takes off. From 1850 to 1960, the share of individuals living in urban areas increased up to 50%.

⁴⁷Above the obvious prevention of reclassification issues, see Bocquier and Costa (2015), we discuss the advantages of using this pure administrative distinction rather than the classical criterion of population size (e.g. 10,000 inhabitants in the definition of “Städer”, see e.g. Centralbyran (1996)) at the time of observations in Appendix A.3.2.

Like in the Danish case, we focus on the infant mortality rate in the urban and rural areas to describe dynamics of mortality.⁴⁸ Figure 18A illustrates the dynamics of infant mortality both at the global and local levels in Sweden from 1749 to 1960, whereat we apply the same sources as for population size, the SHiPs-Demographic Data Base and Historisk statistik för Sverige.

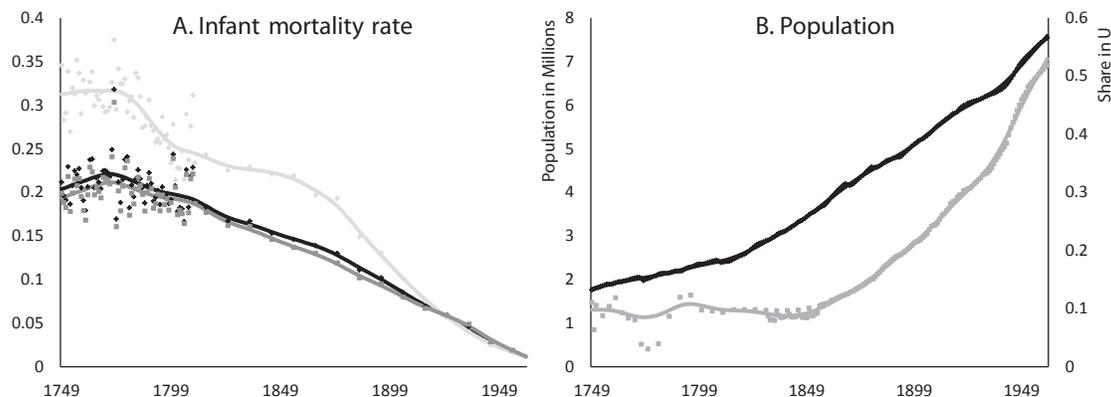


Figure 18: INFANT MORTALITY AND POPULATION DYNAMICS IN SWEDEN

Figure A: Infant mortality rates at the country level (black), in cities (light gray) and countryside (dark gray) by year. Figure B: Population size of Sweden (black) on the left y-axis and proportion of persons in cities (medium gray) on the right y-axis by year. Sources: Centralbyran (1969) and SHiPs-Demographic Data Base; own calculations.

In the middle of the 18th century infant mortality was quite high in Sweden. Driven by rural infant mortality, on average only four out of five children survived until age one. Situation was even worse in cities where the IMR was 10 percentage points higher. Around 1770, infant mortality starts to diminish. The trend of the IMR in rural Sweden is almost linear over almost the next 200 years and so does overall Sweden. The history of urban infant mortality is made of remarkable drops followed by stagnations. A first dramatic drop at the end of the 18th century shrank the rural-urban difference in mortality to 5%-points. However, afterwards gains in the survival probability in rural areas exceeded those in cities. In the middle of the 19th, the differential again reached 10%-points before a sustained convergence. In the early 20th century, the rural-urban-difference in IMR has vanished.⁴⁹ At the end of the period we consider, almost no difference is observable and IMRs are closed to zero.

⁴⁸Brändström (1993) presents Swedish infant mortality rates on a provincial level and figures out varying differences within provinces with similar urbanization rate. Brändström *et al.* (2002) investigate the role of illegitimacy and feeding practices on the parish level.

⁴⁹The probability to survive until age one is even higher in cities than in countryside just before the middle of the 20th century.

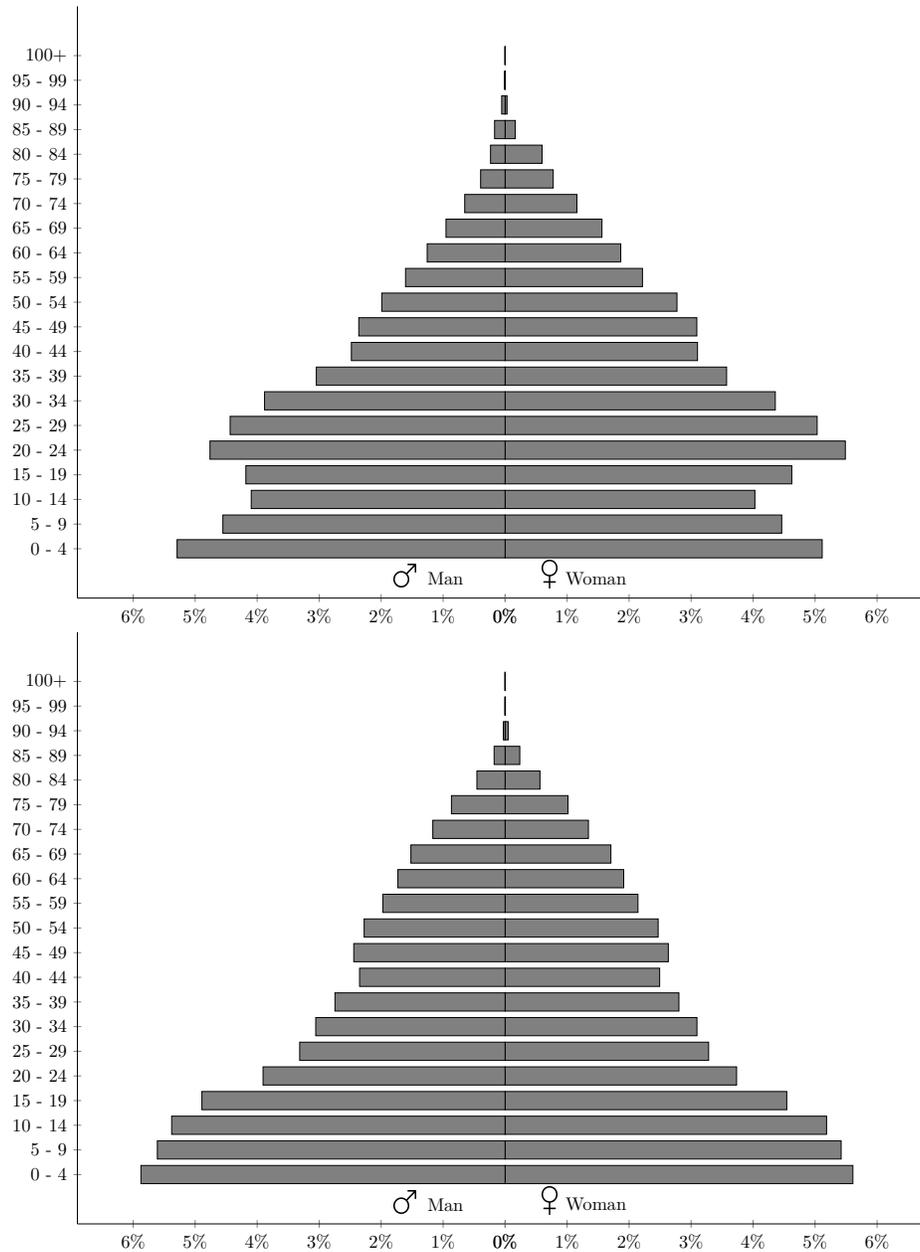


Figure 19: AGE STRUCTURES OF SWEDEN'S URBAN AND RURAL POPULATION IN 1910

Age structures of the Swedish population in urban (top) and rural (bottom) areas by sex. Source: statistical yearbook Sweden 1914; own calculations.

We now turn to the dynamics of fertility. As we focus on the distinction between cities and countryside, migrations between these areas may dramatically shape estimated local fertility rates. Indeed, migration influences the number of women in fertile age

in both areas and so the total number of births therein. As argued above, using crude births rate as an indicator of local fertility could lead to strongly biased estimations. Such an intuition is confirmed, for instance, by the inspection of age pyramids for the year 1910 (Figure 19) where the share of population in fertile age is much higher in cities than in countryside. This simple difference could lead to overestimating fertility in cities relative to countryside. As a result and like in the Danish case, we use Coale’s index of fertility instead of crude birth rate.

As for other variables, we merge data from SHiPs-Demographic Data Base in the period from 1749 to 1855 and several statistical yearbook from 1910 to 1970. Data about the age structure in specific areas are not available between 1855 and 1910, so we use the calculation of Mosk (1980) for the year 1880 to complement our data collection. As depicted in Figure 20, the Coale’s Index is equal to 0.372 in Sweden in 1749. Fertility is slightly higher in rural ($I^R = 0.376$) and lower in urban areas ($I^U = 0.326$). Thus rural fertility exceeds that of cities by around 15% at the beginning of our observations. Until 1790, fertility in both urban and rural areas shrinks and starts to increase in rural areas afterwards, while it stagnates in cities. From around 1880 to the end of our period of observation, fertility in both areas decrease despite they reach a maximal differential close to 60% around 1930. Then, both fertility rates increase and tend to converge to each other from 1930.

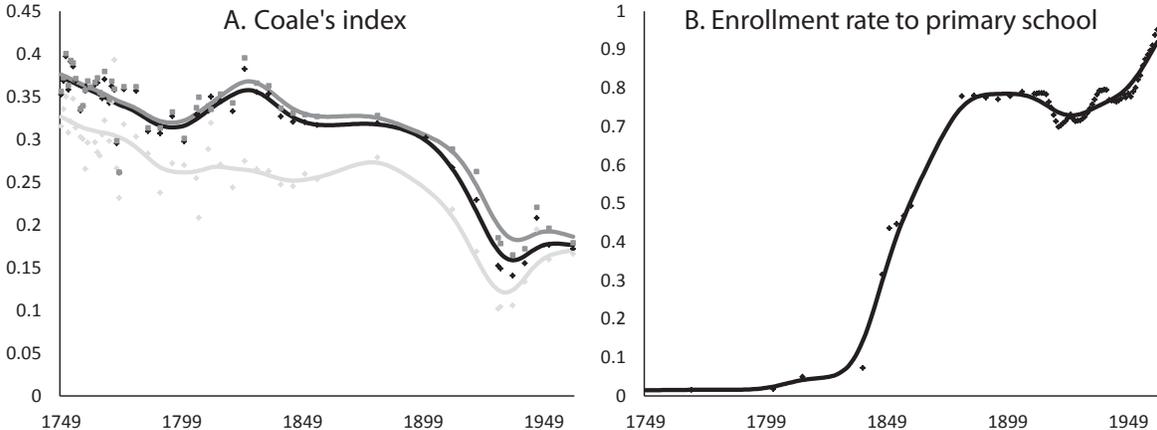


Figure 20: EVOLUTION OF FERTILITY AND EDUCATION IN SWEDEN

Figure A: Fertility approximated by Coale’s index for Sweden (black), countryside (dark gray) and cities (light gray) of Sweden. Figure B: Enrollment rates to primary education. Sources: SHiPs-Demographic Data Base, Mosk (1980), Swedish statistical yearbooks (several years) and de la Croix *et al.* (2008); own calculations.

Swedish dynamics in fertility and mortality are similar to the Danish patterns. Fertility was lower and mortality higher in cities than in countryside (Figure 20A and 18A). Then, without a significant rural exodus, the share of persons in the countryside should have increased what is apparently not the case looking at the right panel of Figure 18.

Enrollment rates to primary (and lower secondary) school have been extracted from de la Croix *et al.* (2008). Before the 19th century, primary education was negligible and remained low until 1842 when compulsory elementary education was introduced (Flora, 1983). Then, enrollment rates to primary education jumped to around 80% quickly and persisted on this level until the end of the second world war. Afterwards, a further increase is observable in Figure 20B such that almost all children are educated in 1960.

Economic side

The Swedish Historical National Accounts offer information on production per sector in Sweden since 1560. The economy is divided in seven main sectors: agriculture, manufacturing and industry, building and construction, transport and communication, private services, public services as well as services of dwellings, see Schön *et al.* (2012). We use the sector called “agriculture” to approximate output in rural areas, while we define the sector “manufacturing and industry” as the industrial production in cities. Time series are available in current prices as well as constant factor prices (Mill. SEK).⁵⁰ Figure 21 shows the evolution of the Swedish GDP in constant prices as well as the GDP in rural and urban areas.

As our data start very early, we are able to observe a period of stagnation even if as soon as 1610, mainly driven by industrial production, aggregate production started to grow slightly. At the end of the 16th century, agricultural production was around 14 times higher than the industrial production. This difference decreased to a factor of 8 until the beginning of the industrial revolution. Around 1820, a clear and impressive break occurred, production in the agricultural sector entered sustained growth until 1930 and then stagnated while the industrial production exploded and grew at the annual constant rate of 3.49% from 1820 to 2010.⁵¹

⁵⁰Until 1800 prices are in 1800 price level. Since 1800 the price level is 1910/1912.

⁵¹One would ideally like to know the evolution of GDP per capita in cities and countryside. Nevertheless, such data don't exist. The Swedish Historical National Accounts offer employment by sector what enable us to calculate the GDP per workforce, nevertheless, this information is available only from 1850. We report them in Appendix A.3.3 keeping in mind the critics of Clark *et al.* (2012) about

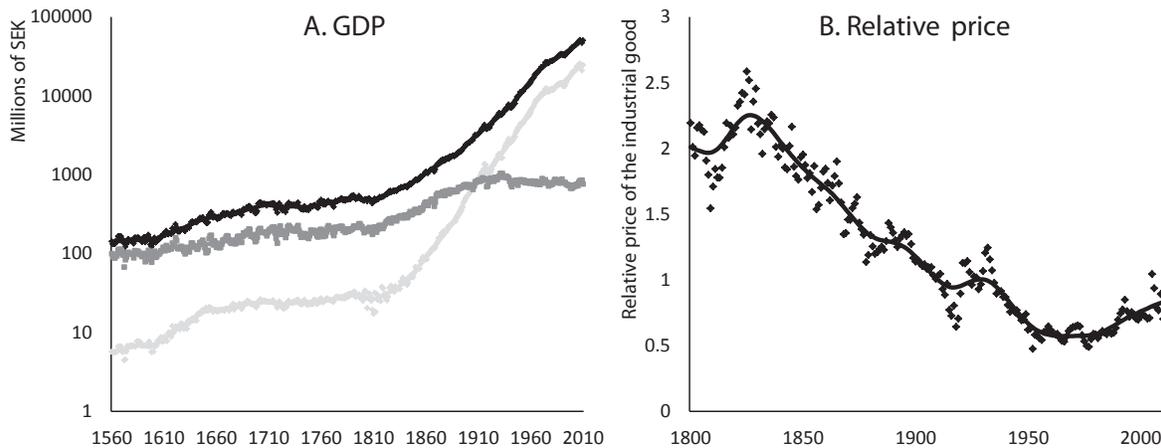


Figure 21: DYNAMICS OF SWEDISH GDP AND RELATIVE PRICE

Figure A: Total GDP of Sweden in black, countryside's total GDP in dark gray and cities' total GDP in light gray. Data is expressed in constant prices of 1910/12. Figure B: Relative price of the industrial goods. Source: Schön *et al.* (2012); own calculations.

In Sweden, industrialization came with a strong decrease of relative prices of manufactured goods compared to agricultural goods. Still using the Swedish National Accounts, we divide GDP in current prices by GDP in constant prices in each area what allows us to approximate the dynamics of prices in rural and industrial sectors since 1800. Like in Denmark, the relative price of industrial goods first slightly increases and shrinks after 1825.

A.3.2 Definition rural and urban areas

Information on urban and rural population originate from different sources. Very early data of our time series from 1749 to 1799 originates from SHiPs-Demographic Data Base and uses institutional criteria.⁵² More recent years since 1800 are added from official statistics and use demographic criteria.

this methodology.

⁵²We merge the data set from SHiPs with a Database Retrieval on urban-rural parishes both offered by the Umeå University.

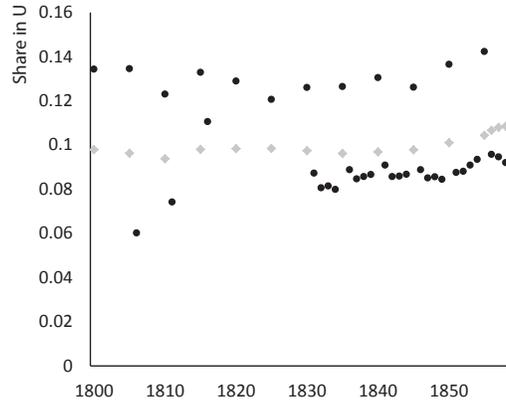


Figure 22: SWEDEN’S URBANIZATION RATES 1800–1859 FROM SHiPs AND OFFICIAL STATISTICS

Share of individuals in cities calculated from SHiPs (black) and official statistics (gray). Sources: SHiPs-Demographic Data Base and Centralbyran (1969); own calculations.

Based on municipal rights, the institutional criteria defines urban areas in early stages. They were intended to concentrate trade and artisanship in some locations and thus to support economic growth and facilitate control over the economy, an idea closed to the idea of a city in our model: places with modern technologies supporting economic growth. Additionally, the areas do not switch their class if population increases (or shrinks), which is in line with our model, too. By contrast, official statistics rely on demographic criteria. The classification of urban population, we apply in our data set, bases on “Landsbygd”(rural) and “Städer” (urban). Towns (“stad”) require more than 10,000 inhabitants. This definition fixes a kind of lower bound on the share of urban population, illustrated by Figure 22 and Table 8.

The overlapping period of observations, data from SHiPs is available until 1859, figures out that in the first half of the 19th century populations in urban areas are over-represented by SHiP’s data every five years, when official data is available.⁵³ In between the share of individuals implied by SHiPs is much lower and likely an underestimation.

⁵³This finding is inline with the intuition: Not all cities with municipal rights have more than 10,000 inhabitants. However, as the number of places with more than 10,000 individuals increased over time the difference should have vanished and even changed. Several places without municipal rights might exceed the threshold, such that the demographic criteria leads to higher values. Biases in the database, due to missing parishes are an alternative explanation for the lower official values. This argument is supported by the changing levels in Figure 22.

Table 8: SWEDEN'S URBAN AND RURAL POPULATION IN 1900 ACCORDING TO DIFFERENT DEFINITIONS

<i>“Landsbygd” versus “Städer”</i>				
	Towns	Market towns	LCR*	Rural areas
Inhabitants	1103951	25190	319706	3687594
Urban/Rural	1103951	4032490		
<i>Densely versus sparsely populated areas</i>				
	Densely populated areas		Sparsely populated areas	
	Towns	Market towns	Rural municipalities	
Inhabitants	1103951	25190	161078	3846222
Urban/Rural	1448847			3687594

*LCR marks areas with “Limited Communal Rights”. Source: Centralbyran (1996).

Table 8 confronts the applied definition to the more recent definition of localities. Whereas our definition, only taking into account towns with more than 10,000, leads to a value of 21.5% as lower bound, the share is much higher (28.2%) following the new definition. However, due to three reasons we apply the concept of “Landsbygd” and “Städer”: First, the current classification of densely and sparsely populated areas, by means of “tätorter” (localities) is only available since the beginning of the 20th century.⁵⁴ Secondly, most required demographic variables are distinguished by “Landsbygd” and “Städer”. Thirdly, the brake in time series on individuals in urban areas is sufficiently small, at the period our data sources change.

A.3.3 GDP per capita or GDP per employee?

One of the critics of Clark *et al.* (2012) is that “urbanization rates in preindustrial England provide a very poor guide to the share of the population employed in farming.” We apply Swedish data, to confront our assumption, that individuals in rural

⁵⁴“A locality consists of a group of buildings normally not more than 200 meters apart, and must fulfill a minimum criterion of having at least 200 inhabitants. In Sweden localities are defined as urban, and all areas outside the localities as non-urban.” (www.scb.se)

areas produce in the agricultural sector with Clark’s critic. Data on GDP and employee per sector starting in 1850 enable to compare time series of agricultural GDP per individual in countryside with that per employee in agricultural sector in Figure 23.

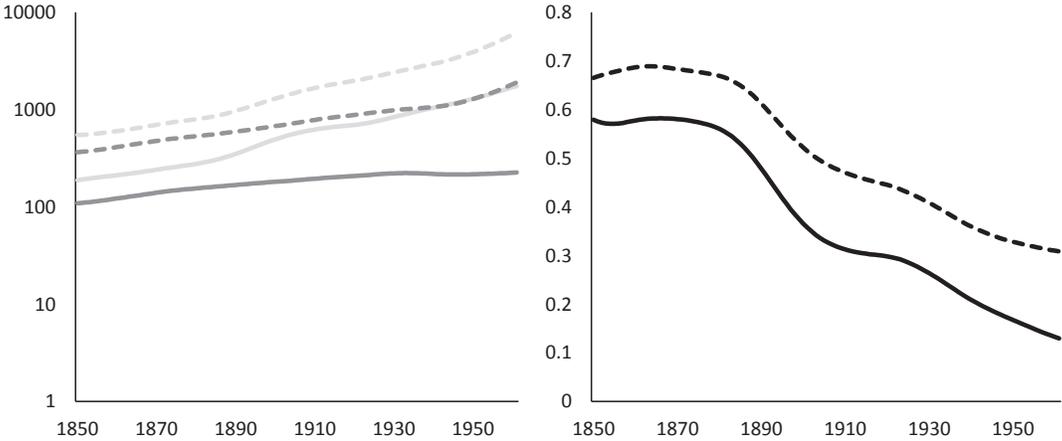


Figure 23: GDP PER PERSON IN URBAN AND RURAL AREAS VERSUS GDP PER EMPLOYEE

Left panel: Trend of GDP per person (solid line) and per employee (dashed line) in rural areas (dark gray) and urban areas (light gray) if the former produce the agricultural and the latter industrial output. Right panel: Dynamics of the relation between rural and urban output per person (solid line) or employee (dashed line). Source: Schön *et al.* (2012); own calculations.

The left panel of Figure 23 illustrates the dynamics of GDP per person supposing that exclusively agricultural goods are produced in rural and industrial goods in urban areas. Additionally, the output per employee in each sector is plotted. Development of both variables for U and R is rather similar over time, whereas a clear differential in the level exists. The right panel provides additional evidence that evolution of the ratios is not such a bad indicator: both the relation between agricultural and industrial output per employee develops similar to agricultural output per person in rural areas compared to industrial output per person in cities.

Table 9: SHARE OF INDIVIDUALS WORKING IN EACH SECTOR BY AREA

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

Source: Centralbyran (1969).

There are several explanations for the differences: First of all, not all individuals in R were employed in agricultural sector and vice versa not everybody in U worked in industries. An aspect inline with Clark’s critics. Table 9 illustrates the share of population related to agricultural or industrial sector. In particular in the rural area a kind of (proto)-industrialization is observable. Still, the share of individuals working in agricultural sector is very high. By contrast, in the urban area, agricultural employment is negligible during the whole period and the share of individuals in the industrial sector is almost constant. However, as retired persons are “enclosed to their ancient industry, domestic servants to the industry of their employer and dependents to the industry of their head of family” (Centralbyran, 1969), even this numbers are sensitive to socio-demographic characteristics. Our demographic investigation on Sweden illustrates differences in age structures and thus on the share of working individuals, see Figure 18.

Summarizing Figure 23 and Table 9, we conclude that first of all, urbanization rates are an acceptable indicator for the share of individuals working in the industrial sector in Sweden. Secondly, even if there was a slightly increasing share of employees in rural areas, urbanization was a driving force for the industrialization.

B Theory

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

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

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

- $n_1 = 0 < n_M$
- n_2 is continuous with respect to ω , $\forall \omega^{\mathbb{A}} \in \mathbb{R}^+$. Furthermore, $\frac{dn_2}{d\omega^{\mathbb{A}}} = \frac{c}{(\alpha+\rho)(\xi+\zeta q^{\mathbb{A}})\omega^{\mathbb{A}}} > 0$, $n_2(\bar{c}) = 0$ and $\lim_{\omega^{\mathbb{A}} \rightarrow +\infty} n_2(\omega^{\mathbb{A}}) = \frac{\rho}{\alpha+\rho} n_M < n_M$. We can then conclude that $n_2 < n_M$, $\forall \omega^{\mathbb{A}}$.
- n_3 is continuous with respect to ω , $\forall \omega^{\mathbb{A}} \in \mathbb{R}^+$. Furthermore, $\frac{dn_3}{d\omega^{\mathbb{A}}} = \frac{c+\kappa}{(\alpha+\rho)(\xi+\zeta q^{\mathbb{A}})\omega^{\mathbb{A}}} > 0$, $n_3(\bar{c} + \kappa) = 0$ and $\lim_{\omega^{\mathbb{A}} \rightarrow +\infty} n_3(\omega^{\mathbb{A}}) = \frac{\rho}{\alpha+\rho} n_M < n_M$. We can then conclude that $n_3 < n_M$, $\forall \omega^{\mathbb{A}}$.
- n_4 is continuous with respect to ω , $\forall \omega^{\mathbb{A}} \in \mathbb{R}^+$. Furthermore, it is a monotonous function of $\omega^{\mathbb{A}}$ as the sign of $\frac{dn_4}{d\omega^{\mathbb{A}}}$ does not depend on $\omega^{\mathbb{A}}$:

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

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

- Simple computations yield to the conclusion that:

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

■

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

- $\{c_1, n_1, d_1, e_1\}$ is, by assumption the solution of the individual maximization program $\forall \omega^{\mathbb{A}} \leq \bar{c}$.

- $\{c_2, n_2, d_2, e_2\}$ maximizes δ subject to \mathcal{B} , $n_i < n_M$, $c \geq \bar{c}$ and $n_i = d_i = e_i = 0$ when $\kappa = 0$.
- $\{c_3, n_3, d_3, e_3\}$ maximizes δ subject to \mathcal{B} , $n_i < n_M$, $c \geq \bar{c}$, $n_i > 0$ and $d_i = e_i = 0$ when $\kappa > 0$.
- $\{c_4, n_4, d_4, e_4\}$ maximizes δ subject to \mathcal{B} , $n_i < n_M$, $c \geq \bar{c}$, $n_i > 0$ and $d_i > 0$ and $e_i = 0$ $\forall \kappa \geq 0$.
- $\{c_5, n_5, d_5, e_5\}$ maximizes δ subject to \mathcal{B} , $n_i < n_M$, $c \geq \bar{c}$, $n_i > 0$ and $d_i = 0$ and $e_i > 0$ $\forall \kappa \geq 0$.
- $\{c_6, n_6, d_6, e_6\}$ maximizes δ subject to \mathcal{B} , $n_i < n_M$, $c \geq \bar{c}$, $n_i > 0$ and $d_i > 0$ and $e_i > 0$ $\forall \kappa \geq 0$.

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

■

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

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

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

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

$$\tilde{p}(\kappa) > p^*(\kappa) \Leftrightarrow \frac{\bar{c} + \kappa}{\frac{v\beta q^{\mathbb{A}}}{\phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}})} - \frac{\alpha + \rho}{1 - \alpha} \varepsilon} > \frac{(\bar{c} + \kappa) \phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}})}{v\beta q^{\mathbb{A}}} \quad (14)$$

If $v \leq \bar{v}$, the denominator of the left-hand side part of Inequality (14) is negative such that Inequality (14) is never satisfied. It means that $\tilde{p}(\kappa) \leq 0 < p^*(\kappa)$. By contrast, if $v > \bar{v}$, Inequality (14) summarizes to $\frac{\alpha + \rho}{1 - \alpha} \varepsilon \phi(\xi + \zeta^{\mathbb{A}} q^{\mathbb{A}}) > 0$ what is always satisfied meaning that if $v > \bar{v}$, $\tilde{p}(\kappa) > p^*(\kappa) > 0$.

■

Lemma 4

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

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

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

■

Inspecting Lemmas 1 to 4 yields to the conclusion that Proposition 1 describes well the solution to the maximization problem exposed in Subsection 3.2.

C Additional figures of historical experiments

C.1 Constant IMRs in cities

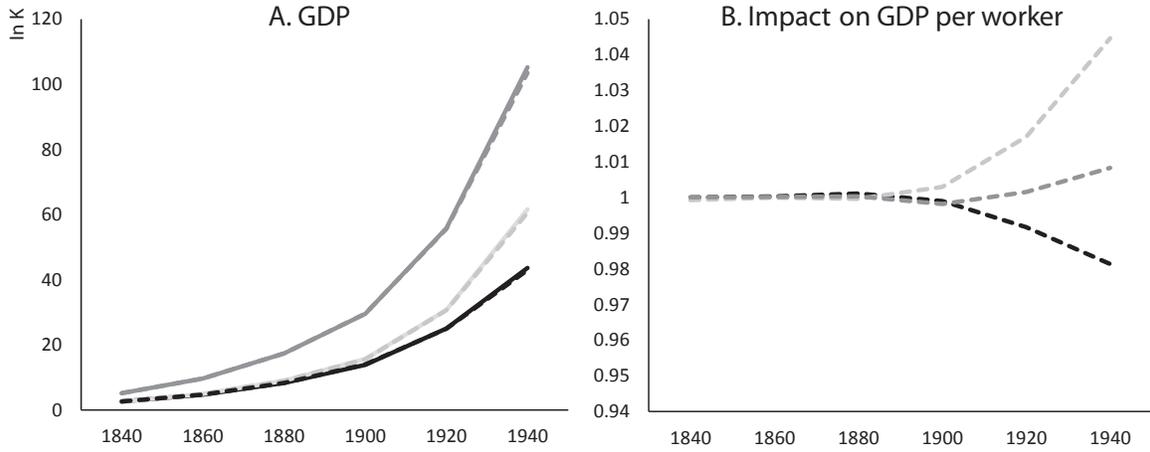


Figure 24: GDP AND THE IMPACT OF CONSTANT URBAN IMRS ON GDP PER WORKER

Solid lines present trends of the benchmark calibration in Denmark and dashed lines the simulation with constant IMRs in cities. Black marks rural areas, light gray urban ones and dark gray total Denmark.

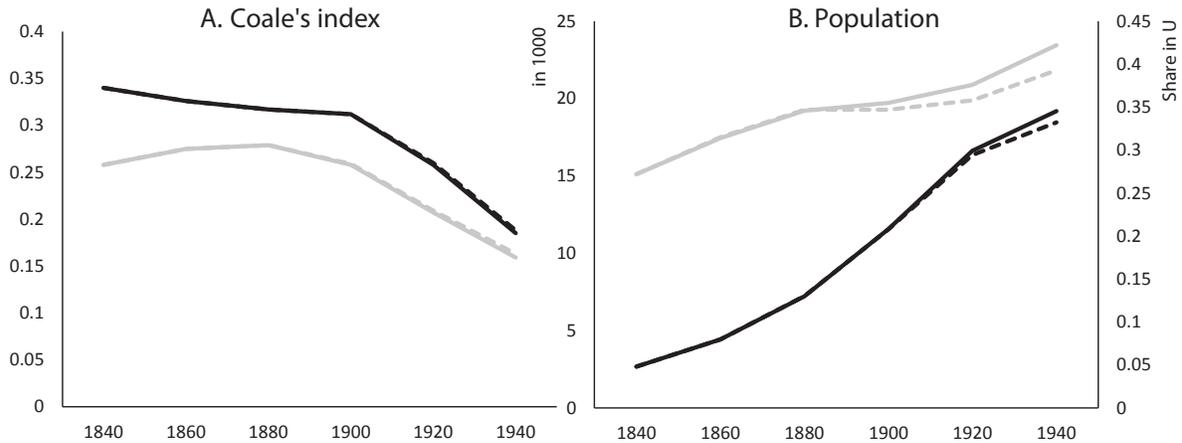


Figure 25: FERTILITY AND POPULATION DYNAMICS WITH CONSTANT IMRS IN CITIES

Solid lines present trends of the benchmark calibration and dashed lines the simulation with constant IMRs in cities. Black lines mark rural and gray lines urban areas in Figure 25. The left y-axis in Figure 25B shows total populations (black) and the right y-axis the share in the urban area (gray).

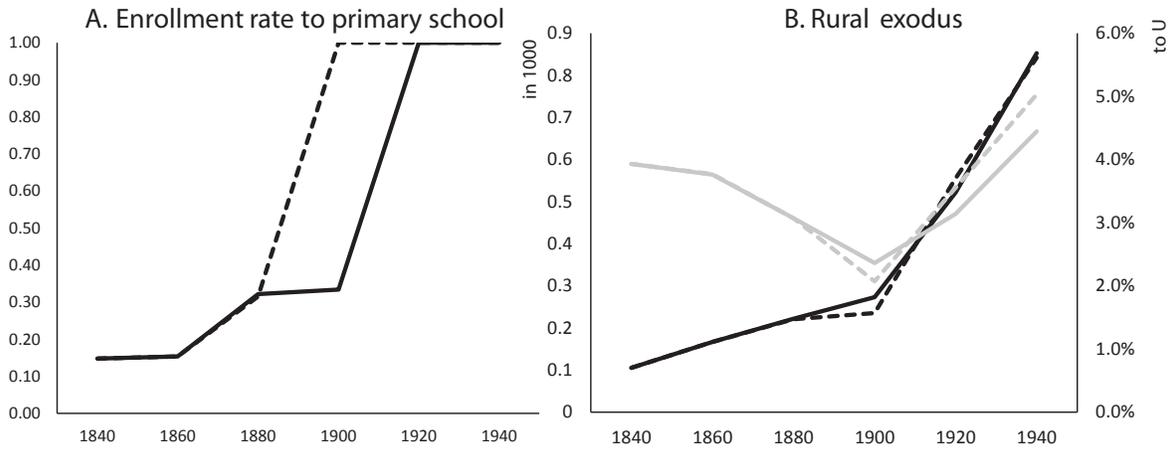


Figure 26: EDUCATION AND RURAL EXODUS WITH CONSTANT IMRS IN CITIES

Solid lines present trends of the benchmark calibration in Denmark and dashed lines the simulation with constant IMRs in cities. In Figure 26B, the left y-axis marks the total number of moving individuals (black) and the right y-axis their share on the total population (gray).

C.2 Wages

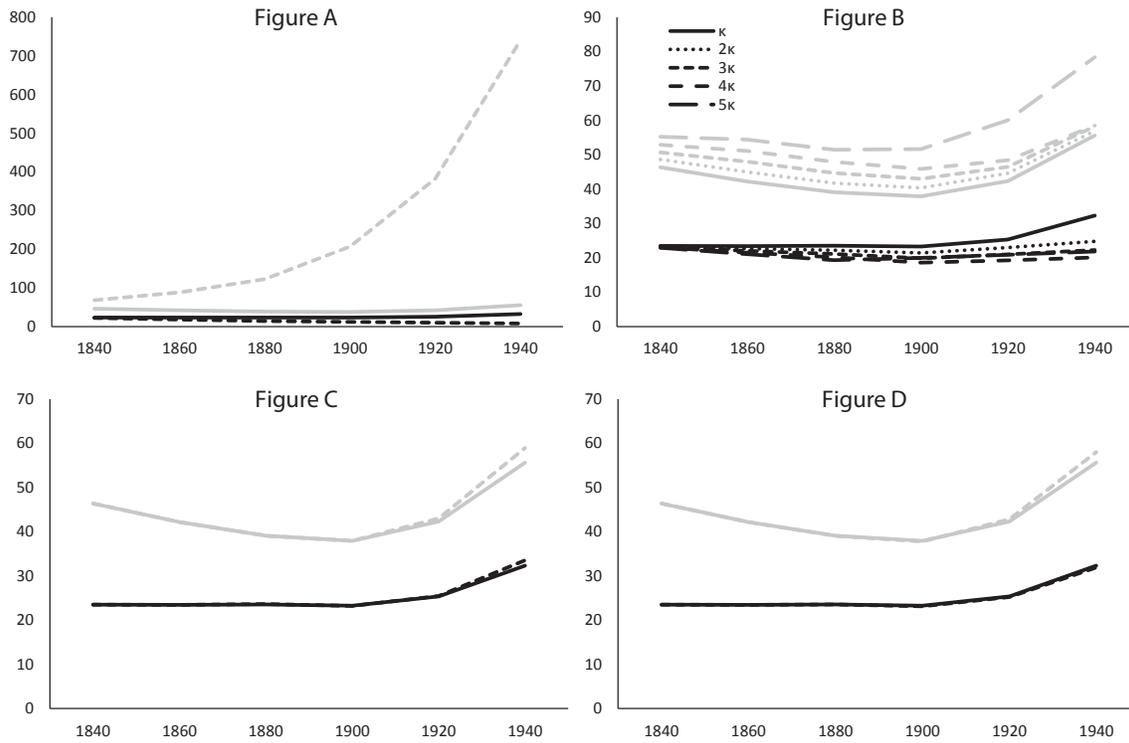


Figure 27: DYNAMICS OF REAL WAGES IN DIFFERENT HISTORICAL EXPERIMENTS

Solid lines present trends of the benchmark calibration and dashed lines the different simulations. Figure A presents the “no-movement” scenario and B the varying migration costs (legend see inside the figure). Lower figures illustrate the two scenarios with constant IMRs (Figure C) and constant IMRs in cities (Figure D). Black lines mark rural and gray lines urban areas.