A Model of Historical Evolution of Output and Population*

Michael Bar[†] Oksana M Leukhina[‡]

July 28, 2005

Abstract

All industrialized countries have experienced a transition from high birth rates and stagnant standards of living to low birth rates and sustained growth in per capita income. What factors contributed to this transition and to what extent? Were output and population dynamics driven by common or separate forces? We develop a general equilibrium model with endogenous fertility in order to quantitatively assess the impact of changes in young-age mortality and technological progress on the demographic transition and industrialization in England. We find that the decline in young-age mortality accounts for 60% of the fall in the General Fertility Rate that occurred in England between 1700 and 1950. Over the same period, changes in productivity account for 76% of the increase in GDP per capita and nearly all of the decline of land share in total income. Furthermore, we find that changes in productivity are quantitatively insignificant in accounting for the observed patterns in fertility behavior, while mortality changes are quantitatively relevant only to population dynamics, not to the other quantities predicted by the model.

JEL classification: J10, O1, O4, E0

1

*Both authors are grateful to Larry Jones, Michele Boldrin, and V.V. Chari for their excellent advice. We also greatly benefited from suggestions by Matthias Doepke.

[†]Department of Economics, University of Minnesota, Minneapolis, MN 55455; and Research Department, Federal Reserve Bank of Minneapolis, Minneapolis, MN 55480; E-mail: mbar@econ.umn.edu

[‡]Department of Economics, University of Minnesota, Minneapolis, MN 55455; and Research Department, Federal Reserve Bank of Minneapolis, Minneapolis, MN 55480; E-mail: oksana@econ.umn.edu

Introduction

The ideas of Thomas Malthus published in 1798 [26] appear to be consistent with most of human history. Malthus expressed a rather dim outlook for the evolution of population and output. In the absence of sustained technological progress, Malthus claimed that standards of living would always remain constant. According to his theory, any one-time technological improvement would translate into temporarily higher population growth until the standards of living returned to their original level. Cross-country differences in the state of technology would simply translate into variation in population size, not into a disparity in per capita incomes. The main driving forces of Malthusian theory are land-labor technology with diminishing returns to labor due to a fixed supply of land and the assumption that population growth increases with per capita consumption.

Fortunately, all industrialized countries experienced a transition from stagnant standards of living to sustained growth in per capita income, thus escaping the Malthusian trap. This transition coincided with the demographic transition from high birth and mortality rates to low birth and mortality rates. Notably, in most countries, there was a lag between the drop in death rates and the drop in birth rates, which resulted in a hump in the population growth rate. Furthermore, resources reallocated from rural production to non-rural production, and the importance of land's income share in total production significantly declined over the same period of time. These key observations motivated this paper.

Why did these events take place? What are the main forces that drove this transition? Is there a common explanation for economic and demographic changes, or were output and population driven by separate forces? These questions are of pressing importance, especially in the view of current economic conditions in many sub-Saharan African countries that have not yet undergone the demographic transition. These countries' staggering poverty necessitates effective policy recommendations.

In order to answer the questions posed above, we develop a general equilibrium model with endogenous fertility capable of generating the transition from Malthusian stagnation to modern growth. Within the framework of our model, which is calibrated to match some key moments at the beginning of 17th century England, we quantitatively assess the importance of two factors in shaping the demographic transition and industrialization in England: changes in young-age mortality and technological progress. More precisely, we examine the model dynamics that result when changes in young-age mortality and total factor productivity (TFP) in the rural and non-rural sectors vary over time in accordance with historical data.

This paper contributes to the recent trend in growth literature that attempts to explain economic development over long time scales. In a recent work, Lucas [26] emphasizes the importance of this line of research: "... I think it is accurate to say that we have not one but two theories of production: one consistent with the main features of the world economy prior to the industrial revolution [Malthusian theory] and another roughly consistent with the behavior of the advanced economies today [Solow growth theory]. What we need is an understanding of the transition."

The mechanism that we use is a dynamic general equilibrium model with endogenous fertility. It has two important components. First, production is modeled as in Hansen and Prescott [22]. The final good can be produced using two different technologies, the Malthusian, which uses capital, labor, and land as inputs, and, the Solow, which employs capital and labor only. Since land is a fixed factor, it essentially introduces decreasing returns to scale to capital and labor in the Malthusian sector. We associate the Malthusian technology with rural production that took place on small individual farms. In contrast, the Solow technology is associated with urban production. This choice of modeling production allows us to investigate the implications of changes in young-age mortality and TFP for resource allocation between the two technologies. In this paper we refer to the fraction of non-rural output in total output as the level of industrialization and the fraction of labor employed by the non-rural sector in total labor as the level of urbanization.

The second important part of our mechanism is endogenous fertility. As in Barro and Becker [3], we assume that parents place value on both the number of surviving children and their children's well-being. Thus, there is a quantity-quality trade-off explicit in our model. Parents face a trade-off between having many children with small inheritance in the form of capital and land for each child and having a few children but endowing each with a larger piece of land and more capital.

How do changes in young-age mortality and TFP propagate in our model? We would like to highlight a few effects here, in particular, the effect of these changes on birth rates and the level of industrialization. There are two channels through which changes in young-age mortality affect the fertility choices made by households. On one hand, with a higher number of children surviving to adulthood, fewer births are needed to achieve the desired number of surviving children. On the other hand, as the probability of survival increases, the cost of raising a surviving child declines, and hence, induces higher birth numbers. This method of modeling the time cost of raising children follows Boldrin and Jones [6], as well as Doepke [11]. In short, an increase in the probability of survival always leads to a reduction in birth rates and a temporary increase in the number of surviving children. Also, depending on parameter specification, an increase in the probability of survival may or may not lead to a permanent increase in the number of surviving children. We find that changes in young-age mortality represent a very important force behind the demographic change in England. In particular, they can account for as much as 60% of the fall in birth rates between 1700 and 1950.

Similarly, a switch to a faster growing TFP affects the return on children. On one hand, children are normal goods, and hence, higher income growth induces higher fertility. On the other hand, with faster growing TFP, the opportunity cost of raising children measured in terms of foregone parents' wages also grows faster. This has a dampening effect on fertility. This latter channel has been considered vitally important for the decline in birth rates by many economists, demographers, and historians. Among those who argue that the technological progress was a dominant factor behind the demographic change are Galor and Weil [14], Greenwood and Seshadri [17], Hansen and Prescott [22]. In fact, our findings qualitatively agree with this effect. Changing the growth rate of TFP in the two sectors according to our estimates, does lead to a decline in both, birth rates and population growth rate. However, this effect turns out to be quantitatively unimportant.

The take-off in the non-rural TFP growth as well as the decline in young-age mortality can each generate full resource reallocation towards the non-rural sector. As the Solow TFP begins to grow faster than the Malthusian TFP, the Solow sector attracts a higher proportion of resources every period. The effect of young-age mortality on the level of industrialization, however, is less clear. As the probability of survival increases, the time cost of raising a surviving child declines. As a result, the time available for work increases and leads to the relative expansion of output in the Solow sector, which uses labor intensively. This train of thought follows the argument of the famous Rybczynski Theorem. So once again, as in the discussion about birth rates, it is unclear apriori which of the two exogenous changes is quantitatively more important for the process of urbanization and industrialization. We find that only changes in the TFP growth rates are quantitatively relevant for the time period studied here.

To summarize the main results, we find that the decline in young-age mortality accounts for 60% of the fall in the General Fertility Rate[1] that occurred in England between 1700 and 1950. Over the same period, changes in productivity account for 76% of the increase in GDP per capita and for nearly all of the decline of land share in total income. Interestingly, both experiments generate a transition from Malthus to Solow. However, changes in TFP do so in a manner consistent with empirical observations, driving the share of the Malthusian technology to nearly zero in the period from 1600 to 2000. Changes in the probability of survival lead to a much slower transition, predicting that even in 2400, the output produced by the Malthusian technology would comprise 10% of total output. We also find that changes in total factor productivity alone can account for long term trends in the observed patterns of factor income shares. This is due to resource reallocation between sectors with different but constant factor intensities.

One important question raised was whether the forces driving the economic and demographic changes could be separated out. We find that this is indeed the case; the explanations for changes in output and population need not be entirely common. In fact, we find that changes in productivity are quantitatively insignificant in accounting for the observed patterns in fertility behavior, while mortality changes are quantitatively relevant only to population dynamics, and not to the other quantities predicted by the model. Our finding does not rule out the possibility that there are important interactions between the two changes that we treat as exogenously given here, or that there exists some third force that is responsible for changes in technological progress and young-age mortality. This may very well be true. What our findings suggest is that the quantitatively relevant channels through which this third force acts to generate the demographic and economic change are different. It generates the demographic change through its effect on young-age mortality and it generates the economic change through its effect on technological progress.

The most important contribution of our work is the quantitative analysis of young-age mortality and changes in TFP within a general equilibrium framework with two sectors of production and explicit reproduction choice. We carry out careful analysis of historical data for England and Wales. We work with mortality and fertility data provided by Wrigley, Davies, Oeppen, and Schofield [36], Mitchell [31], and the Human Mortality Database [23]. The survival probabilities used in the model represent their actual historical estimates. In fact, these probabilities do not change monotonically, as the reader may conjecture. Similarly, we estimate TFPs in the rural and urban sectors using the dual-approach. This approach requires time series data on wages in the two sectors, land and capital rental rates, and the GDP deflator. These time series were either taken directly or inferred from three of Gregory Clark's papers [10], [8], [9].

Another important contribution is our analysis of transitional dynamics from one type of balanced growth path towards another, triggered by the observed changes in mortality rates and/or relative TFP growth rates. In many of the earlier works, the prevalent analysis of the exogenous changes was performed by comparing steady states. This was done because of the difficulties as-

[1] General Fertility Rate is the number of live births per 1,000 women ages 15-44 or 15-49 years in a given year.

sociated with solving for equilibrium paths in this type of non-stationary environment. Although we fully appreciate the importance of comparative statics analysis, we also find that a great deal of insight can be lost by leaving the transition path out of consideration. Indeed, we find that convergence in the benchmark model is quite slow.

The rest of this work is organized as follows. Chapter 2 reviews the English case data and elucidates the up to date accomplishments in related literature. In Chapter 3 we set up the model, discuss some equilibrium properties and report the main results. Chapter 4 contains all of the technical details.

2 Data and Literature Review

2.1 Some Facts about England and Wales

We choose to focus on the case of England and Wales due to data limitations for other countries. Next, we examine some fundamentally important facts that motivated this work. We will return to the data discussion in section 4. Extensive reviews of the English case are given in Boldrin and Jones [6] and Fernandez-Villaverde [20]. Galor and Weil [14] and Hansen and Prescott [22] provide detailed accounts of the regime switches experienced by developed countries. Galor [13] also reviews historical data on output and demographic changes and makes a strong case in favor of developing unified growth theory that can account for the process of development over long time scales.

Figure **??** depicts the evolution of the index of the real GDP per capita for England and Wales. The data sources are Clark [9] for the period from 1560 to 1860 and Maddison [28] from 1820 to 1992. Observe that per capita real GDP is roughly stagnant for centuries, until it takes off in the beginning of the 19th century.

The demographic transition in England and Wales is depicted in Figure **??**. This figure plots Crude Birth Rates[2] and Crude Death Rates[3] provided by Wrigley, Davies, Oeppen, and Schofield [36] for the time period up to 1871 and continued using Mitchell's [31] data. Before 1750, England experienced high fertility and high mortality rates. Average population growth in the first half of the 18th century was around 0.4% per year. In the second half of the 18th century, mortality began to decline, and this was accompanied by rising (or at least persisting high) fertility rates. In the second half of the 19th century, birth rates began to fall, while mortality continued to decline. Eventually, both stabilized at a new low level in the first half of the 20th century.

Note that the lag between the drop in death rates and the drop in birth rates implies a hump in the population growth rate. A number of studies attempt to generate the drop in population growth rates. It is, however, useful to notice that the drop in the population growth is just the later part of the hump, and the rise in population growth rate deserves just as much attention. Also, observe that the drop in fertility rates was so rapid that it indicates the importance of economic forces in governing the demographic transition, in contrast to usually slowly evolving cultural changes. It is also interesting that fertility remained high for about 80 years after the beginning of sharp growth in the real per capita GDP.

[2] Crude Birth Rate is the number of births in a given year per 1000 people.

[3] Crude Death Rate is the number of deaths in a given year per 1000 people.

Figure **??** depicts Crude Birth Rates together with the probability of survival to the age of twenty five. The latter time series we calculate based on age-specific mortality rates available from Wrigley, Davies, Oeppen, and Schofield [36] and the Human Mortality Database [23]. Notice that the timing of a sharp increase in the probability of survival coincides with the drop in birth rates. Since this observation is not unique to the case of England, changes in young-age mortality are often cited as the primary driving force behind the demographic transition.

It is also significant that this period is associated with the trend of people moving out of the rural sector and into the industrial capital-intensive sector. As depicted in Figures **??** and **??**, the share of the urban GDP in the total GDP rose from around 30% in the 1550s to roughly 98% in the 1990s. Similarly, the share of employment in non-rural production dramatically increased from around 40% to 98% over the same period. The data on the level of industrialization and urbanization up to 1860 are taken from Clark's papers [9] and [8]; the time series are continued using Maddison's [28] data.

2.2 Related Literature

At this point, it is instructive to review related literature and to elucidate the accomplishments made to this time with respect to the objectives stated above.

Most of the literature in this field is theoretical. Examples include the pioneering works of the dynamic formulation of the dynastic model of fertility choice carried out by Ben-Zion and Razin [32] and Barro and Becker [3], [4]. In both models, parents are altruistic toward their children and decide on the number of children as well as the amount of bequests. In Ben-Zion and Razin model, an increase in the productivity of capital tends to cause a decrease in the population growth. Barro and Becker use a model similar to that of Ben-Zion and Razin, but it also assumes exogenous labor augmenting technological progress. Their model is more standard and is more easily comparable with the National Income and Product Accounting.

There are two main experiments performed in quantitative studies employing the Barro-Becker model in an attempt to match the observed data on the demographic transition. The first is based on the belief that the increase in income that accompanied the industrial revolution started the demographic transition. Fernandez-Villaverde [20] finds that increasing productivity in a version of the Barro-Becker model leads to a rise in both fertility and the net reproduction rate. The second experiment reported in the literature is based on the observation that infant mortality rates fell during the demographic transition. For example, Doepke [11] finds that in several versions of the Barro-Becker model, as infant mortality declines, the total fertility rate falls, but the number of surviving children increases. He concludes that "factors other than declining infant and child mortality are responsible for the large decline in net reproduction rate observed in industrialized countries over the last century".

Another example of a purely theoretical model aimed at investigating the historical evolution of output and population are Galor and Weil [14]. Galor and Weil argue that technological progress is skill-biased. Hence, parents respond to technological progress by having fewer, higher quality children. The growing stock of human capital feeds back into higher technological progress, thus reenforcing this mechanism.

Becker, Murphy, Tamura [5] and Hansen and Prescott [22] represent valuable steps in the direction of understanding the transition. Becker, Murphy, and Tamura [5] emphasize the quantity-quality trade-off. The main driving force in their model is the assumption that the return on human capital is increasing in the stock of human capital. Their model has two possible stable equilibria, one characterized by high fertility, low income, and low stock of human capital, and one described by low fertility, high income, and high stock of human capital. The authors do not attempt to model the transition between these equilibria. Instead, they conclude that "luck" must play an important role in triggering the shift between them.

Hansen and Prescott [22] suggest a different mechanism that might have triggered the demographic transition. They propose a model with two technologies identical to those we choose, one Malthusian technology, which requires land, labor, and capital as inputs, and one Solow technology, which requires labor and capital as inputs. The transition from Malthus to Solow is an equilibrium property of their model brought about by technological progress in the Solow technology. As in the Malthusian model, population growth is postulated to be a function of per capita consumption. Hansen and Prescott calibrate the parameters of this function in order to match the demographic transition in Europe. They obtain a non-monotonic function: for low levels of per capita income, income and population growth are positively correlated and for income levels above a certain threshold, this correlation becomes negative. Hansen and Prescott's work suggests that explanations based on technological progress are promising. However, it provides no economic insight for why the relationship between population growth and the level of per capita consumption is non-monotonic. Our work is most closely related to Hansen and Prescott [22], as we use the same technological assumptions. However, unlike Hansen and Prescott, we explicitly model fertility choice.

Greenwood and Seshadri [17] use a two sector model with endogenous fertility to study the U.S. demographic transition. They find that changes in TFP alone can account for both the decline in fertility rates and the increase in GDP per capita that occurred in the U.S. Their results stand in contrast to ours. We find that the effect of changes in TFP are insignificant in accounting for fertility patterns in England.

Alternatives to quantity-quality trade-off stories have also been developed in the literature. One alternative is based on the idea that children provide old-age security for parents. This theory reverses the direction of altruism, as children are now the ones who care about their parents. In these models, parents have children because they expect to be cared for when they become old. Excellent references are Boldrin and Jones [6] and Ehrlich-Lui [19].

Boldrin and Jones perform an experiment to elucidate the implication of the decline in infant and child mortality and conclude, similarly to Doepke [11], that it leads to a decline in birth rates but not in the number of surviving births. Population growth can potentially be reduced in an extended version of the Boldrin-Jones model that would study equilibria in which children do not cooperate. Boldrin, DeNardi, and Jones [7] find that the increase in the size of the Social Security system leads to changes in fertility behavior that are consistent with empirical evidence whenever the Boldrin-Jones framework is used. By contrast, when they use the Barro-Becker dynastic framework, they find the effect of changes in the Social Security system to be quantitatively unimportant.

Two more excellent quantitative investigations are Doepke [12] and Fernandez-Villaverde [20]. Utilizing data on the timing and duration of demographic changes that took place in Brazil and Korea, Doepke concludes that government policies that impact the opportunity cost of education, such as education subsidies and child-labor laws, have a direct effect on the speed of the demographic transition. His mechanism is also based on a quantity-quality trade-off.

Fernandez-Villaverde explores the fall of relative capital prices and finds it to be quantitatively important in accounting for the observed patterns of fertility and per capita income.

3 Model and Simulation Results

This is a one sector overlapping generations model with two technologies, exogenous technological progress, and endogenous fertility.

Technology, firms

Firms are endowed with one of two possible technologies that can produce the same good as in Hansen and Prescott [22]. We subscript the Malthusian technology that requires capital, labor, and land as inputs by 1 and associate it with production taking place in the rural sector. The Solow technology that employs capital and labor as inputs is subscripted by 2 and associated with production taking place in the cities. Both technologies exhibit constant returns to scale, which allows us to assume that there are two aggregate competitive firms, one using the Malthusian technology, and another using the Solow technology. The outputs of these two firms are given by

$$\begin{aligned} Y_{1t} &= A_{1t} K_{1t}^{\phi} L_{1t}^{\mu} \Lambda_t^{1-\phi-\mu}, \\ Y_{2t} &= A_{2t} K_{2t}^{\theta} L_{2t}^{1-\theta}, \end{aligned}$$

where K_j, L_j denote capital and labor employed by technology $j \in \{1, 2\}$ and Λ_t denotes land employed by the Malthusian technology.

We assume exogenous technological progress in both technologies, that is,

$$A_{1t} = A_{10} \prod_{\tau=0}^{t} \gamma_{1\tau} \text{ and } A_{2t} = A_{20} \prod_{\tau=0}^{t} \gamma_{2\tau},$$

where $\gamma_{i\tau}$ represents the time τ exogenous growth rate of technology i TFP.

Formally, the profit maximization problems of Firms 1 and 2 are given by

$$\begin{aligned} &\max_{K_{1t}, L_{1t}, \Lambda_t} A_{1t} K_{1t}^{\phi} L_{1t}^{\mu} \Lambda_t^{1-\phi-\mu} - w_t L_{1t} - r_t K_{1t} - \rho_t \Lambda_t, \\ &\max_{K_{2t}, L_{2t}} A_{2t} K_{2t}^{\theta} L_{2t}^{1-\theta} - w_t L_{2t} - r_t K_{2t}, \end{aligned}$$

where w_t, r_t, and ρ_t denote time t wage, capital rental rate, and land rental rate respectively.

Preferences, households, dynasties

There is measure 1 of identical dynasties. Two people belong to the same dynasty if they have a common predecessor. We denote the number of households belonging to dynasty j and alive at time t by $N_t(j)$. Then the total number of households in the economy alive at time t is given by $N_t = \int_{[0,1]} N_t(j)\, dj$, where j is a uniform measure on $[0, 1]$ indexing the dynasties.

Households live for two periods, childhood and adulthood. An adult household belonging to dynasty j and alive at time t derives utility from its own consumption $c_t(j)$, the number of its surviving children $n_t(j)$ (young households), and its children's average utility. The general form of households' preferences is given by

$$U_t(j) = u(c_t(j), n_t(j)) + \beta U_{t+1}(j).$$

It should be noted that for this utility choice, altruism per child is implicitly set to $\frac{1}{n_t}$, which is the Barro and Becker [4] altruism with $\varepsilon = 1$[4].

A fraction π_t of children born, f_t, survives to adulthood. We denote the number of surviving descendants by n_t,

$$n_t(j) = \pi_t f_t(j). \tag{1}$$

There is a time cost associated with raising children. A household spends fraction a of its time per each born child and an additional fraction b of its time per each child who survives to adulthood. We assume that for each newborn child, households pay the expected cost of raising him with certainty. Thus, the total time cost of raising $f_t(j)$ newborn children is given by

$$[a + \pi_t b] f_t(j) = \left(\frac{a}{\pi_t} + b\right) n_t(j),$$

where we used (1) to factor out the number of surviving children. Hence, $\frac{a}{\pi_t} + b$ represents the time cost of raising a surviving child. Denote this cost by q_t, that is,

$$q_t = \frac{a}{\pi_t} + b. \tag{2}$$

Observe that the time cost of raising surviving children is a decreasing function of the survival probability. Intuitively, as more newborn children survive to adulthood it becomes cheaper to raise a surviving child [5].

An adult household belonging to dynasty j rents its land holdings $\lambda_t(j)$ and inelastically devotes the time not spent raising children to work. It chooses its own consumption $c_t(j)$, the number of his surviving children $n_t(j)$, and the amount of bequests $k_{t+1}(j)$ to be passed on to each surviving child in the form of capital, and divides its land holdings equally among its descendants

Formally, a household belonging to dynasty j solves

$$\begin{aligned}
\max_{c_t(j), n_t(j), \lambda_{t+1}(j), k_{t+1}(j) \geq 0} U(k_t(j), \lambda_t(j)) &= u(c_t(j), n_t(j)) + \beta U_{t+1}(k_{t+1}(j), \lambda_{t+1}(j)) \\
& \text{s.t.} \\
c_t(j) + k_{t+1}(j) n_t(j) &= (1 - q_t n_t(j)) w_t + (r_t + 1 - \delta) k_t(j) + \rho_t \lambda_t(j), \\
\lambda_{t+1}(j) &= \frac{\lambda_t(j)}{n_t(j)}.
\end{aligned}$$

Each household takes sequences of wages, capital rental rates, and land rental rates as given and takes into account the effect that his choices today have on the average utility of his descendants. Parents face the quantity-quality trade-off between having many children with small inheritance in the form of capital and land for each child and having a few children but endowing each with a larger piece of land and more capital.

[4] Barro and Becker (1989) assume $U_t = u(c_t) + \beta n_t^{1-\varepsilon} U_{t+1}$ to ensure that for given utility per child U_{t+1} parental utility is increasing and concave in the number of children.

[5] If we modeled the cost of raising children to be paid in terms of the final good, the results would not change. In that case, for the existence of a balanced growth path along which per capita variables grow at a positive rate, we would need to assume that the goods cost grows in proportion to income.

Population dynamics

The number of adult households of dynasty j evolves according to

$$N_{t+1}(j) = \int_{[0,N_t(j)]} n_t(i,j)\,di, \tag{3}$$

where i is a uniform measure on $[0, N_t(j)]$ indexing households of dynasty j. Hence, the total number of adult households alive evolves according to

$$N_{t+1} = \int_{[0,1]} N_{t+1}(j)\,dj = \int_{[0,1]} \left(\int_{[0,N_t(j)]} n_t(i,j)\,di \right) dj. \tag{4}$$

We assume that all dynasties have identical initial conditions, that is, the same initial size $N_0(j) = N_0\ \forall j$, the same endowment of capital per household $k_0(j) = k_0 = K_0/N_0\ \forall j$, and the same endowment of land $\lambda_0(j) = \lambda_0 = \Lambda/N_0\ \forall j$. We only consider symmetric equilibria, that is, equilibria with the property that all households in the model economy behave identically. In particular, $n_t(i,j) = n_t\ \forall i,j$. Under this assumption, (3) becomes

$$N_{t+1}(j) = n_t N_t(j),$$

and since all dynasties are identical initially, it follows that $N_t(j) = N_t\ \forall j$ and the total number of households in the economy is given by

$$\int_{[0,1]} N_t dj = N_t.$$

The total number of children per household in the economy is

$$\int_{[0,1]} n_t dj = n_t.$$

Then (4) becomes

$$N_{t+1} = n_t N_t.$$

Market Clearing

In the equilibria considered in this paper, all decisions are symmetric across households, i.e., $c_t(j) = c_t$, $k_t(j) = k_t\ \forall j$, and hence, the aggregates for the economy are given by

$$\begin{aligned}
C_t &= \int_{[0,1]} c_t N_t dj = c_t N_t = c_t N_t, \\
K_t &= \int_{[0,1]} k_t N_t dj = k_t N_t = k_t N_t, \\
\Lambda &= \int_{[0,1]} \int_{[0,N_t(j)]} \lambda_t(i,j)\,di dj = \lambda_t N_t.
\end{aligned}$$

The feasibility constraint and market clearing conditions in the capital, labor, and land markets are as follows:

$$\begin{aligned}
C_t + K_{t+1} &= A_{1t} K_{1t}^{\phi} L_{1t}^{\mu} \Lambda_t^{1-\phi-\mu} + A_{2t} K_{2t}^{\theta} L_{2t}^{1-\theta} + (1-\delta) K_t, \\
K_{1t} + K_{2t} &= K_t, \\
L_{1t} + L_{2t} &= (1 - q_t n_t) N_t, \\
\Lambda_t &= \Lambda.
\end{aligned}$$

3.1 Equilibrium

Definition 1 *A symmetric competitive equilibrium, for given parameter values and* (k_0, N_0)*, consists of allocations* $\{c_t, n_t, \lambda_t, k_{t+1}, k_{1t}, k_{2t}, l_{1t}, l_{2t}, N_{t+1}\}_{t=0}^{\infty}$ *and prices* $\{w_t, r_t, \rho_t\}_{t=0}^{\infty}$ *such that households and firms solve their maximization problems, and all markets clear.*

Consider the following Social Planning problem (SP) :

$$\max_{\{C_t, N_{t+1}, K_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t u\left(\frac{C_t}{N_t}, \frac{N_{t+1}}{N_t}\right) \tag{SP}$$

$$\begin{aligned}
&\text{s.t.} \\
C_t + K_{t+1} &= F(K_t, L_t; t) + (1-\delta)K_t, \\
L_t &= N_t - q_t N_{t+1}, \\
C_t, K_{t+1}, N_{t+1} &> 0,\ K_0, N_0 \text{ given, where} \\
F(K_t, L_t; t) &= \max_{K_{1t}, L_{1t}} \left[A_{1t} K_{1t}^{\phi} L_{1t}^{\mu} \Lambda^{1-\phi-\mu} + A_{2t}(K_t - K_{1t})^{\theta}(L_t - L_{1t})^{1-\theta}\right] \qquad (5) \\
\text{s.t. } 0 &\leq K_{1t} \leq K_t,\ 0 \leq L_{1t} \leq L_t
\end{aligned}$$

There are difficulties associated with defining efficiency in models with endogenous fertility. The Social Planning problem defined above corresponds to the $\mathcal{A}$-efficiency concept as defined by Golosov, Jones, Tertilt [16]. According to this concept, when comparisons are made across allocations, the positive weight is put only on those households that are alive in all possible allocations. Analyzing concepts of efficiency in models of endogenous fertility, however, is beyond the scope of this paper. We define this Social Planning problem because it compactly states the optimization problem at hand, and it illustrates the sense in which the competitive equilibrium allocation is efficient. For convenience, we also use the definition of output as a maximum value function (5) to numerically determine the optimal allocation of resources when solving for equilibrium time paths.

Proposition 2 *The competitive equilibrium in the decentralized economy corresponds to the solution of the Social Planning problem.*

Proof. See the Technical Appendix[6]. ■

We follow Lucas [26] and choose

$$u(c_t, n_t) = \alpha \log c_t + (1-\alpha) \log n_t,$$

so the Barro and Becker assumption that for a given utility per child U_{t+1} parental utility, $U_t = \alpha \log c_t + (1-\alpha)\log n_t + \beta U_{t+1}$, is increasing and concave in the number of children is maintained. In the Appendix we show that this utility represents the same preferences as represented by the Barro and Becker utility $U_t = c_t^{\sigma} + \beta n_t^{1-\varepsilon} U_{t+1}$ with $\sigma \to 0$ and $\frac{1-\sigma-\varepsilon}{\sigma} = \frac{1-\alpha-\beta}{\alpha\beta}$.

Proposition 3 *Under the assumption that* $u(c_t, n_t) = \alpha \log c_t + (1-\alpha)\log n_t$*, the objective function in (SP) can be replaced by* $\sum_{t=0}^{\infty} \beta^t (\alpha \log C_t + (1-\alpha-\beta)\log N_{t+1})$.

[6] Available through the authors' websites or upon request.

Proof. See the Appendix. ■

Notice that continuity of the objective function together with compactness and non-emptiness of the constraint set guarantees existence of a solution. We assume that $1 - \alpha - \beta > 0$ to guarantee that the objective function is strictly concave. Since the constraint set is convex, the solution is unique. In the Appendix we show that this solution can be characterized using first order, feasibility, transversality, and optimal resource allocation conditions. These conditions are derived in the Appendix.

Proposition 4 *If sequences $\{C_t, N_t, K_t, L_{1t}, Y_t\}_{t=0}^{\infty}$ satisfy the first order, feasibility, and optimal resource allocation conditions for (SP), C_t/Y_t is bounded away from 0 and Y_t exhibits growth as $t \to \infty$, then the transversality conditions are also satisfied.*

Proof. See the Appendix. ■

The above proposition allows us to replace the transversality condition by a simple check of the limiting behavior of C_t/Y_t and Y_t. Notice that any balanced growth behavior of the equilibrium time paths such that C_t/Y_t is constant guarantees that transversality conditions hold.

From the Social Planner's perspective, both capital and children are investment goods. By choosing more children today, N_{t+1}, production can be increased tomorrow although at the expense of decreasing production today due to the time cost of raising children. Another interesting trade-off clear from the set-up of the Social Planning problem is the trade-off between consumption and children today. Indeed, both C_t and N_{t+1} enter the objective function in the Social Planning problem. Hence, children are both consumption and investment goods.

It is instructive to review the intuition that can be obtained from the first order conditions derived for the Social Planning problem:

$$\frac{C_{t+1}}{C_t} = \beta \left(r_{t+1} + 1 - \delta\right), \tag{6}$$

$$\frac{(1-\alpha-\beta)\, C_t}{\alpha N_{t+1}} = q_t w_t - \frac{w_{t+1}}{r_{t+1} + 1 - \delta}, \tag{7}$$

$$C_t + K_{t+1} = F\left(K_t, N_t - q_t N_{t+1}; t\right) + (1-\delta) K_t, \tag{8}$$

where w_t denotes the marginal product of labor, i.e., $w_t = F_2\left(K_t, N_t - q_t N_{t+1}; t\right)$. The first equation, (6), is a standard Euler equation that describes the intertemporal trade-off in aggregate consumption. The second, (7), is the intratemporal trade-off between consumption and children since N_{t+1} denotes the number of adults in $t+1$, or equivalently, the number of children today. It represents the condition that the marginal rate of substitution between children and consumption is given by their relative price. The price of a child in terms of the final good is measured by the forgone output resulting from having to spend time to raise this child less the present value of the child's earnings in $t+1$. The last equation, (8), is the feasibility condition.

Proposition 5 *The Malthusian technology operates for all t as long as $K_t, L_t > 0$.*

Proof. Suppose on the contrary that there is time t such that $Y_{1t} = 0$. Since resources are allocated efficiently, this means that $K_{1t} = L_{1t} = 0$ and

$$\max_{K_t, L_{1t}} \left[A_{1t} K_{1t}^{\phi} L_{1t}^{\mu} \Lambda^{1-\phi-\mu} + A_{2t} \left(K_t - K_{1t}\right)^{\theta} \left(L_t - L_{1t}\right)^{1-\theta} \right] = A_{2t} K_t^{\theta} L_t^{1-\theta}. \tag{9}$$

Consider reallocating $(\varepsilon K_t, \varepsilon L_t)$ to the Malthusian technology, where $\varepsilon \in (0,1)$. Next we show that for ε small enough,

$$A_{1t}(\varepsilon K_t)^{\phi}(\varepsilon L_t)^{\mu}\Lambda^{1-\phi-\mu} + A_{2t}((1-\varepsilon)K_t)^{\theta}((1-\varepsilon)L_t)^{1-\theta} > A_{2t}K_t^{\theta}L_t^{1-\theta}. \tag{10}$$

Simplifying this inequality gives $\frac{1}{\varepsilon^{1-\phi-\mu}} > \frac{A_{2t}K_t^{\theta}L_t^{1-\theta}}{A_{1t}K_t^{\phi}L_t^{\mu}\Lambda^{1-\phi-\mu}}$. Since $\lim_{\varepsilon\to 0}\frac{1}{\varepsilon^{1-\phi-\mu}} = \infty$ and the right hand side is a finite number, $\exists\varepsilon > 0$ that ensures (10) is satisfied. Hence, we arrive at contradiction with (9). ■

Less formally, due to decreasing returns to scale in capital and labor, the marginal products of inputs in the Malthusian technology become very large when its capital and labor inputs converge to zero as long as all land is employed. This guarantees that the Malthusian technology is always used in production.

It is possible that the Solow technology does not operate in equilibrium. It is also possible that both technologies always operate side by side. Finally, it is possible that the Solow technology does not operate only for a number of time periods.

Limiting Behavior of Equilibrium Time Paths

The behavior of the solution to the model depends on the choice of the parameters and the initial conditions. We can identify three possible types of limiting behavior of the equilibrium time paths: (1) The solution exhibits the property that the level of output in both sectors converges to some constant positive fraction of total output, (2) The solution exhibits the property that the level of output in the Solow sector converges to 0, (3) The solution exhibits the property that the level of output in the Malthusian sector relative to the total output converges to 0. We refer to these types of limiting behavior of equilibrium time paths as convergence to Malthus-Solow Balanced Growth Path (BGP), Malthus BGP and Solow BGP respectively.

We do not include the discussion of how the choice of parameters and initial conditions affects the limiting behavior of equilibrium time paths in this paper, but a detailed discussion formulated in terms of propositions and proofs is available in the technical appendix through the authors' websites. Here we do, however, need to point out that on a Malthus-Solow BGP, both, population growth and per capita output growth, are determined by the TFP growth rates in the two sectors[7]:

$$\gamma = \gamma_2^{\frac{1}{1-\theta}},\ n = \left(\gamma_1\gamma_2^{-\frac{1-\phi}{1-\theta}}\right)^{\frac{1}{1-\phi-\mu}}. \tag{11}$$

The growth rate of per capita output increases in the Solow TFP and is independent of the Malthusian TFP. Population growth increases in the Malthusian TFP growth rate and decreases in the Solow TFP growth rate. Interestingly, the time cost of raising children does not enter these two equations. This means that increasing the probability of survival while keeping all other parameters fixed would directly result in the proportional reduction of fertility $(n = \pi f)$. For this class of simulations, we found that during the transition from the original to the new balanced growth path, population growth exhibits a hump, and that this transition is lengthy. Therefore, it is misleading to conclude from these comparative statics exercises that mortality changes do not affect population growth. It is important to notice that this analysis is only valid as long as the new value of π does not alter the type of limiting behavior of equilibrium paths,

[7] This result comes from the constancy of the interest rate on any balanced growth path and equality of the marginal products of capital in the two sectors. Hence, it is robust to the choice of the objective function.

i.e., as long as it does not preclude convergence to a new Malthus-Solow BGP. In fact, in the simulation results of the benchmark economy that are presented below, both of the exogenous changes (one is changes in γ_1 and γ_2 and one is changes in π) that are fed into the model imply that the economy converges to a Solow BGP.

There is no analytical solution, similar to (11), for the growth rate of population and per capita output on Malthus BGP and Solow BGP. The comparative statics results show that for both, Malthus BGP and Solow BGP, increases in the TFP growth rate lead to a decline in the population growth rate and an increase in per capita output growth rate. For Malthus BGPs, increases in probability of survival lead to exactly the opposite effect. In contrast, for Solow BGPs, increases in survival probabilities lead to increases in population growth but do not affect the growth rate of per capita output, $\gamma = \gamma_2^{\frac{1}{1-\theta}}$.

This discussion contrasts the result obtained by Hansen and Prescott [22]. In Hansen and Prescott, as long as the growth rate of the Solow TFP is positive, all equilibria exhibit convergence to a Solow BGP. In our model, however, the limiting behavior of equilibrium time paths is determined by the particular parameterization and, possibly, initial conditions.

4 Calibration

The data for England and Wales was already briefly discussed in the introduction. The objective is to calibrate the parameters of the model to match some key data moments at the beginning of 17th century England. One important assumption that we make in order to map the data moments to the model is that in the beginning of the 17th century, the economy is on a Malthus-Solow BGP.

The data on population growth and mortality rates are available in Wrigley, Davies, Oeppen, and Schofield [36], Mitchell [31], and Human Mortality Database [23]. Most other data moments come from Clark's work [8] and [9]. We also need to estimate the time series of TFP in the rural and non-rural sectors. Unfortunately, we do not have the data on time series of inputs and outputs of the two sectors necessary for standard growth accounting. To get around this problem, we implement the dual-approach of TFP estimation, which uses the assumption of profit-maximization. This approach requires time series data on wages in the two sectors, land and capital rental rates, as well as the GDP deflator. These time series we either take directly or infer from three of Gregory Clark's papers [10], [8], [9].

We choose 25 years to represent the length of each time period. The parameters that we have to calibrate are the Malthusian parameters $A_{10}, \gamma_1, \phi, \mu$, the Solow parameters A_{20}, γ_2, θ, preference parameters α, β, cost of children parameters a, b, π, and the remaining parameters Λ and δ.

Land in the model is a fixed factor whose value we normalize to one ($\Lambda = 1$). Since A_{10} and Λ only enter the model as a product, $A_{10}\Lambda^{1-\phi-\mu}$, we are allowed the second degree of normalization, so we set $A_{10} = 100$. We also set $A_{20} = 100$ as there is no better way to infer it, and sensitivity analysis shows that there is a wide range for A_{20} that will not have any quantitative bearing on the results. It only has the impact on whether the Solow technology is being used in production of output. We have 11 parameters left to calibrate. In order to pin them down we use 11 pieces of information presented in Table 3.1 below[8].

The general idea is to rewrite the balanced growth path equations in terms of moments and parameters only, then solve for the model parameters using the information about the corresponding moments in the data. For the description of calibration as a solution to a system of linear equations see the Appendix.

Table 1: England Around 1600: Data Moments Used for Calibration

Moment	Value	Description
δ	0.723 (0.05)	Depreciation
π	0.67	Probability of survival to 25
$\frac{l_1}{l}$	0.6	Fraction of rural labor in total labor
$\frac{y_1}{y}$	0.67	Fraction of rural output in total output
$\frac{rk}{y}$	0.16	Capital share in total income
$\frac{wl}{y}$	0.6	Labor share in total income
$r+1-\delta$	2.666 (1.04)	Interest rate
qn	0.42	Fraction of time spent with children (or not working)
$\frac{a+b}{a}$	4	Average time cost of surviving children relative to that of non-surviving children
$\gamma_{1,1600}$	1.0402 (1.0016)	Growth of rural TFP around 1600
$\gamma_{2,1600}$	1.0088 (1.00035)	Growth of non-rural TFP around 1600

Notice that we do not aim to match per capita output growth and population growth in our model because, although stationary, these moments are quite volatile around the beginning of the 17th century. These moments, however, will be compared with their counterparts predicted by the calibrated model. Depreciation and nominal interest rate are given as 25 year rates, the corresponding annual rates are indicated in parenthesis. Historical estimates of annual depreciation rates range from 2.5% (Clark [8]) to over 15% (Allen [1]). We set $\delta = 0.723$ to match 5% annual depreciation. The probability of surviving to age of 25 around 1600 was roughly constant at the level of 67%. This number comes from Wrigley, Davies, Oeppen, and Schofield [36]. Hence, π is also pinned down directly by the data.

Clark [9] provides labor and capital shares in total output produced in England as well as relative labor and relative levels of output in the two sectors. The nominal interest rate also comes from one of Clark's papers [10]. The fraction of time spent raising children, qn, is set to 0.42 by us and will be discussed below in this section. Recall that a is the fraction of time spent on each newborn child while b represents the additional time cost incurred when a child lives to become an adult. We set $\frac{a+b}{a}$ to 4 using an assumption of a linear declining functional form for the instantaneous cost function of raising children in conjunction with the data on young-age mortality rates. See the Appendix for a more detailed explanation of how we arrive at this quantity. The discussion of how $\gamma_{1,1600}$ and $\gamma_{2,1600}$ are obtained follows below.

Calibrating ϕ, μ, θ

[8]Numbers in parenthesis indicate annual rates

We can determine the labor share μ of the Malthusian technology using $\frac{y_1}{y}$, $\frac{l_1}{l}$, $\frac{wl}{y}$ and the equilibrium property that wages equal the marginal product of labor in the Malthusian sector, $w\frac{l}{y} = \left(\frac{\mu y_1}{l_1}\right)\frac{l}{y}$. This implies $\mu = 0.537$.

Now that we know μ, we can pin down the capital share θ of the Solow technology by using $\frac{y_1}{y}$, $\frac{wl}{y}$, and the equilibrium identity that the total labor income is given by the sum of the income paid to the labor employed by the Malthusian technology and labor employed by the Solow technology, $\mu\frac{y_1}{y} + (1-\theta)\frac{y_2}{y} = \frac{wl}{y}$. This determines $\theta = 0.273$.

Similarly, we obtain the capital share ϕ of the Malthusian technology by using $\frac{y_1}{y}$, $\frac{rk}{y}$, and the equilibrium property that the total income paid to capital is the sum of rental income paid to the capital employed in the Malthusian sector and capital employed in the Solow sector, $\phi\frac{y_1}{y} + \theta\frac{y_1}{y} = \frac{rk}{y}$. This gives $\phi = 0.104$.

Calibrating γ_1, γ_2 and estimating TFP time series

Now we are ready to explain how we obtained $\gamma_{1,1600}$ and $\gamma_{2,1600}$ that appear in Table 3.1. We first estimated TFP time series for each sector for the time period of 1585-1915. Then for each sector we fit a trend consisting of two parts each characterized by a constant growth rate. The growth rates characterizing the first part of the TFP trends in the two sectors are denoted by $\gamma_{1,1600}$ and $\gamma_{2,1600}$.

In order to estimate the TFP time series we need to know the factor income shares in the two sectors, ϕ, μ, θ. This fact is responsible for this estimation taking place in the middle of calibration. Next we explain more precisely how the time series for the Malthusian and Solow TFP were constructed and how the trends were obtained.

Recall that we associate the output produced by Malthusian technology with output produced in the rural sector and the Solow technology with output produced in the non-rural sector.

From profit maximization of the firms, using the dual-approach of estimating TFP, we derive

$$A_{1t} = \left(\frac{r_t}{\phi}\right)^{\phi}\left(\frac{w_{1t}}{\mu}\right)^{\mu}\left(\frac{\rho_t}{1-\phi-\mu}\right)^{1-\phi-\mu}, \quad (12)$$

$$A_{2t} = \left(\frac{r_t}{\theta}\right)^{\theta}\left(\frac{w_{2t}}{1-\theta}\right)^{1-\theta}, \quad (13)$$

where r_t (%) is the rental rate on capital, w_t is the real wage measured in units of the final good per unit of labor, and ρ_t is the land rental price measured in units of the final good per acre. Since the data available from Clark is the time series of r_t (%), nominal wages ω_{1t} and ω_{2t} (£)[9], $\tilde{\rho}_t$ (% return on land rents), $P_{\Lambda t}$ (price of land in £/acre), and the GDP deflator P_t, we infer the

[9]To be more precise, we infer w_2 using Clark's time series for the total wage bill in the economy $w_1L_1 + w_2L_2$, the bill in the rural sector w_1L_1, fraction of rural labor in total labor $\frac{L_1}{L}$, and the following identity:

$$\frac{w_1L_1 + w_2L_2}{w_2L_2} = \frac{w_1L_1}{w_2L_2} + 1,$$

$$w_2 = \frac{w_1}{\frac{w_1L_1+w_2L_2}{w_2L_2} - 1}\frac{1}{\frac{1}{\frac{L_1}{L}} - 1}.$$

real wages w_{it} and the real land rental price ρ_t by using

$$w_{it} = \frac{\omega_{it}}{P_t} \text{ and } \rho_t = \frac{\tilde{\rho}_t P_{\Lambda t}}{P_t}.$$

Substituting these into (12) and (13), we obtain the equations that allow us to estimate the Malthusian and Solow TFP time series using the available data:

$$\begin{aligned} A_{1t} &= \left(\frac{r_t}{\phi}\right)^{\phi} \left(\frac{\omega_{1t}}{\mu}\right)^{\mu} \left(\frac{\tilde{\rho}_t P_{\Lambda t}}{1-\phi-\mu}\right)^{1-\phi-\mu} P_t^{\phi-1}, \\ A_{2t} &= \left(\frac{r_t}{\theta}\right)^{\theta} \left(\frac{\omega_{2t}}{1-\theta}\right)^{1-\theta} P_t^{\theta-1}. \end{aligned}$$

Figure ?? below is a plot of these time series together with the fitted trends.

Both, the rural and non-rural TFP time series exhibit a regime switch. Next we explain how we find the two trends.

Let x_t represent the data and y_t its trend, which we restrict to be of the following form:

$$y_t = \begin{cases} y_0 g_1^t & 0 \le t \le \tau \\ y_0 g_1^{\tau} g_2^{t-\tau} & \tau \le t \le T \end{cases},$$

where g_1 is the growth rate in the first regime and g_2 is the growth rate in the second regime. To find the trend we solve

$$\min_{y_0, g_1, g_2, \tau} \sum_{t=0}^{T} (y_t - x_t)^2 .$$

Notice that this procedure determines the two growth rates as well as the timing of the regime switch. Applying this methodology to both of the TFP time series we obtain the trends. That is, we obtain the TFP growth rates characterizing the first part of the trends, $\gamma_{1,1600}$ and $\gamma_{2,1600}$, as well as the endpoint growth rates that we denote by $\gamma_{1,1900}$ and $\gamma_{2,1900}$. We thus pin down two more of the parameters, $\gamma_{1,1600} = 1.0402$ and $\gamma_{2,1600} = 1.0088$. These growth rates are given in 25 year terms. Note the level of TFP in the estimated time series contains no information, it is just an index since GDP deflator used is an index. Hence, the only relevant information that we obtain from this estimation procedure is the growth rates. These are the growth rates that we use to perform one of the experiments. In fact, for the purpose of performing this experiment, we take an extra step and smooth out the transition from $\gamma_{i,1600}$ to $\gamma_{i,1900}$ by fitting the logistic function to the endpoint growth rates.

Interestingly, γ_1 and γ_2 give prediction to the growth rate of population and per capita output. Recall that the balanced growth path values for n and γ are determined by γ_1 and γ_2. Hence, the obtained values for the growth rates of the Malthusian and Solow TFP imply that $n = \left(\gamma_1 \gamma_2^{-\frac{1-\phi}{1-\theta}}\right)^{\frac{1}{1-\phi-\mu}} = 1.083$ (or 0.32% in annual terms)and $\gamma = \gamma_2^{\frac{1}{1-\theta}} = 1.0121$ (or 0.048% in annual terms). These predictions are roughly consistent with the data. Indeed, the population in the beginning of the 17th century England grew at the annual rate of 0.4%, while output per capita remained roughly stagnant.

Calibrating the remaining parameters

The preference parameter β is given by the Euler equation $\gamma = \frac{\beta}{n}[r + 1 - \delta]$ after we substitute for γ, n, and the gross interest rate. This yields $\beta = 0.411$.

We set the total fraction of time spent raising children qn at 0.42. There is no obvious way to infer qn from the data, but a simple example may be illustrative. Say a person has 100 hours of productive time endowment per week. He works 40 hours, rests 30 hours and spends 30 hours with all of his children. Since there is no leisure in our model, this pattern of time allocation would imply $qn = \frac{30}{30+40} \cong .429$. The sensitivity analysis shows that the results are robust to changes in qn.

We also set $\frac{a+b}{a} = 4$. Recall that a is the fraction of time spent raising each newborn, and b is the additional cost incurred on children that survive to adulthood. We pin down fraction $\frac{a+b}{a}$ by assuming the instantaneous cost function of raising a child to be linear and declining with the child's age. We then use data on age-specific mortality rates around 1600 to infer the relative size of b to a. We also perform sensitivity analysis for this fraction and find that the results are very robust to changes in $\frac{a+b}{a}$. Hence, $qn = 0.42$ and $\frac{a+b}{a} = 4$ determine $a = 0.086$ and $b = 0.259$.

The balanced growth path feasibility equation gives prediction for $\frac{c}{k} = r\frac{y}{rk} + 1 - \delta - \gamma n$. Using $\frac{c}{k}, n, \gamma, qn, \frac{l_1}{l}$ along with the data moments, $r, \frac{rk}{y}, \frac{y_1}{y}$, in the remaining balanced growth path equation, $\frac{(1-\alpha-\beta)(1-qn)}{\alpha\mu}\frac{y}{y_1}\frac{1}{r}\frac{rk}{y}\frac{l_1}{l}\rho = qn - \frac{\gamma n}{(r+1-\delta)}$, allows us to calibrate α to 0.583.

The calibrated parameters are summarized in Table 3.2.

Table 2: Summary of Calibrated Parameters

	Value	Description
Malthusian Technology Parameters		
A_{10}	100	Initial level of TFP
γ_1	1.04	TFP growth rate
ϕ	0.104	Capital share
μ	0.537	Labor share
Solow Technology Parameters		
A_{20}	100	Initial level of TFP
γ_2	1.0088	TFP growth rate
θ	0.273	Capital share
Preference Parameters		
α	0.583	Weight on consumption
β	0.411	Discount rate
Cost of Children		
a	0.086	Fraction of time spent on each life birth
b	0.259	Additional time spent on each surviving child
Other parameters		
δ	0.723	Depreciation
Λ	1	Land

5 Solution Method

The equilibrium time paths are sequences of allocations and prices that satisfy

$$\frac{C_{t+1}}{C_t} = \beta\left(r_{t+1}+1-\delta\right), \tag{14}$$

$$\frac{(1-\alpha-\beta)\,C_{t+1}}{N_{t+1}} = \alpha\beta\left[\left(r_{t+1}+1-\delta\right)q_t w_t - w_{t+1}\right], \tag{15}$$

$$C_t + K_{t+1} = F\left(K_t, L_t; t\right) + (1-\delta)K_t, \tag{16}$$

$$L_t = N_t - q_t N_{t+1}, \tag{17}$$

where

$$F\left(K_t, L_t; t\right) \equiv \max_{K_{1t},L_{1t}} \left[A_{1t}K_{1t}^{\phi}L_{1t}^{\mu}\Lambda^{1-\phi-\mu} + A_{2t}\left(K_t-K_{1t}\right)^{\theta}\left(L_t-L_{1t}\right)^{1-\theta}\right] \tag{18}$$

$$\text{s.t. } 0 \leq K_{1t} \leq K_t,\ 0 \leq L_{1t} \leq L_t$$

and

$$r_t = F_1\left(K_t, L_t; t\right),\ w_t = F_2\left(K_t, L_t; t\right). \tag{19}$$

Notice that the time cost of raising a surviving child q_t as well as $A_{1t} = A_{10}\prod_{\tau=0}^{t}\gamma_{1\tau}$ and $A_{2t} = A_{20}\prod_{\tau=0}^{t}\gamma_{2\tau}$ are indexed by t. The experiments that we perform involve changing $\{q_t, \gamma_{1t}, \gamma_{2t}\}_{t=1600}^{t=2000}$ in accordance with historical data.

Conditions (14) – (19) rewritten in per household terms become

$$\frac{c_{t+1}}{c_t} = \beta\frac{N_t}{N_{t+1}}\left(r_{t+1}+1-\delta\right), \tag{20}$$

$$(1-\alpha-\beta)\,c_{t+1} = \alpha\beta\left[\left(r_{t+1}+1-\delta\right)q_t w_t - w_{t+1}\right], \tag{21}$$

$$c_t + k_{t+1}\frac{N_{t+1}}{N_t} = f\left(k_t, l_t\right) + (1-\delta)k_t, \tag{22}$$

$$l_t = 1 - q n_t, \tag{23}$$

where

$$f\left(k_t, l_t\right) = \max_{k_{1t},l_{1t}} \left\{A_{1t}k_{1t}^{\phi}l_{1t}^{\mu}\left(\frac{\Lambda}{N_t}\right)^{1-\phi-\mu} + A_{2t}\left(k_t-k_{1t}\right)^{\theta}\left(l_t-l_{1t}\right)^{1-\theta}\right\} \tag{24}$$

$$\text{s.t. } 0 \leq k_{1t} \leq k_t,\ 0 \leq l_{1t} \leq l_t,$$

and

$$r_t = f_1(k_t, l_t),\ w_t = f_2(k_t, l_t). \tag{25}$$

Since the equilibrium time paths exhibit exponential growth, it is difficult to directly search for the numerical solution that satisfies the above conditions. As is commonly done in practice, we work with efficiency, or detrended, variables defined as follows:

$$c_t^* = \frac{c_t}{\prod_{\tau=0}^{t-1} g_\tau},\ k_t^* = \frac{k_t}{\prod_{\tau=0}^{t-1} g_\tau},\ k_{1t}^* = \frac{k_{1t}}{\prod_{\tau=0}^{t-1} g_\tau},\ N_t^* = \frac{N_t}{\prod_{\tau=0}^{t-1} g_{n\tau}}, \tag{26}$$

$$l_t^* = l_t,\ l_{1t}^* = l_{1t},\ w_t^* = \frac{w_t}{\prod_{\tau=0}^{t-1} g_\tau},\ r_t = r_t^*,$$

where we assume that $\prod_{\tau=0}^{-1} g_\tau = \prod_{\tau=0}^{-1} g_{n\tau} = 1$ and g_t and g_{nt} represent the balanced growth rates of y_t and N_t respectively that correspond to the parameters at time t. For the discussion of determining the balanced growth path growth rates for a given parameter choice see the previous section on balanced growth. The reason why we use products of growth rates to detrend the original variables instead of powers of the original growth rate is again the fact that changing parameters might (and actually does) lead to a change of the limiting growth rates. Hence, detrending the original variables by powers of the growth rates along the initial balanced growth path will not be sufficient to eliminate exponential growth of the unknown time paths.

We rewrite conditions (20) – (25) in terms of efficiency variables to obtain

$$\frac{c_{t+1}^* g_t g_{nt}}{c_t^*} = \beta \frac{N_t^*}{N_{t+1}^*} \left(r_{t+1}^* + 1 - \delta \right), \tag{27}$$

$$(1-\alpha-\beta)\, c_{t+1}^* g_t = \alpha\beta \left[\left(r_{t+1}^* + 1 - \delta \right) q_t w_t^* - w_{t+1}^* g_t \right], \tag{28}$$

$$c_t^* + k_{t+1}^* \frac{N_{t+1}^*}{N_t^*} g_t g_{nt} = f(k_t^*, l_t^*) + (1-\delta) k_t^*, \tag{29}$$

$$1 - q_t \frac{N_{t+1}^* g_{nt}}{N_t^*}, \tag{30}$$

where

$$\begin{aligned} f(k_t^*, l_t^*) &= \max_{k_{1t}^*, l_{1t}^*} \left\{ \tilde{A}_{1t} k_{1t}^{*\phi} l_{1t}^{*\mu} \tilde{\Lambda}_t^{1-\phi-\mu} + \tilde{A}_{2t} \left(k_t^* - k_{1t}^* \right)^\theta \left(l_t^* - l_{1t}^* \right)^{1-\theta} \right\} \quad (31) \\ \text{s.t. } 0 &\le k_{1t}^* \le k_t^*, \ 0 \le l_{1t}^* \le l_t^*, \text{ where} \\ \tilde{\Lambda}_t &= \frac{\Lambda}{N_t^*}, \\ \tilde{A}_{1t} &= A_{10} \left(\prod_{\tau=0}^{t-1} \gamma_{1\tau} \right) \left(\prod_{\tau=0}^{t-1} g_\tau \right)^{\phi-1} \left(\prod_{\tau=0}^{t-1} g_{n\tau} \right)^{\mu+\phi-1}, \\ \tilde{A}_{2t} &= A_{20} \left(\prod_{\tau=0}^{t-1} \gamma_{2\tau} \right) \left(\prod_{\tau=0}^{t-1} g_\tau \right)^{\theta-1} \quad (32) \end{aligned}$$

and

$$r_t^* = f_1(k_t^*, l_t^*), \ w_t^* = f_2(k_t^*, l_t^*). \tag{33}$$

Hence, we search for equilibrium time paths of efficiency variables that satisfy conditions (27) – (33) using the original steady state efficiency variables as the initial guess. Once the equilibrium efficiency variables are obtained, we use (26) to back out the equilibrium time paths of the original variables.

6 Simulation Results

The model is thus calibrated to match some key moments in the beginning of the 17th century. Next we perform the two experiments. The first experiment is changing the total factor productivity of the two technologies according to our estimates as discussed above. The second experiment is changing the probability of surviving to adulthood according to its historical estimates. The data for the second experiment is obtained from Wrigley, Davies, Oeppen, and Schofield [36] up to 1841 and Human Mortality Database [23] for later years. Each period in the model corresponds to a specific year. We feed in the exogenous changes in accordance with historical data and solve for the model dynamics under the assumption of perfect foresight.

As noted in the discussion above, different types of limiting behavior of equilibrium time paths are possible in our model. Although we start the economy off on a Malthus-Solow BGP, both of the experiments generate convergence to a Solow BGP, characterized by the Malthusian share of output converging to zero. In terms of the earlier discussion about limiting behavior of the equilibrium time paths and the parameter space segmentation depicted in Figure **??**, as parameters $\hat{\theta}$ change during the experiments, a switch occurs from S_1^* to S_3.

6.1 Changes in the Growth Rates of TFP

The first experiment performed is changing the growth rates of the total factor productivity in the Malthusian and Solow sectors in accordance to our estimation described in the section on calibration. Until the second half of the 18th century, the rural technology enjoyed a somewhat higher TFP growth relative to that of the non-rural technology. Around 1750, the growth rate of the Solow TFP overtook the Malthusian TFP growth. According to our estimates, the agricultural revolution took place after the industrial revolution and on a smaller scale. In annual terms, the growth rate of the Malthusian TFP switches from 0.158% to 0.52% while the growth rate of the Solow TFP undergoes a slightly more significant change, from 0.035% to 0.64%. The TFP time series that are fed into the model are depicted in Figure **??**.

Figure **??** illustrates that this experiment generates labor reallocation from Malthus to Solow in a manner consistent with the data. As the Solow sector becomes continuously more productive relative to the Malthusian sector it employes a higher fraction of the available resources. The equilibrium path converges to the asymptotic balanced growth path on which the fraction of the Malthusian output relative to total output converges to zero. Observe that in our experiment changes in TFP growth rates first take place in 1750, hence, this experiment fails to generate any resource reallocation towards the Solow sector prior to 1750.

The results for the industrialization as depicted in Figure **??** are similar. As TFP in the Solow technology becomes sufficiently large, resources reallocate towards the Solow technology and the fraction of Solow output in total output converges to 1. It is important to notice that urbanization and industrialization are imperfect data counterparts of l_2/l and y_2/y in our model. The main reason is that we associate the Malthusian sector with rural production and Solow sector with non-rural production. However, in the data rural output is not a perfect substitute of the non-rural output while in the model the Malthusian good is a perfect substitute to the Solow good. It is nonetheless instructive to make these comparisons.

It is clear from Figure **??** that changes in the TFP growth rates generate a transition from the Malthusian stagnation to modern growth. Around 1600, the growth rate of per capita GDP is near zero, or more precisely, GDP per capita grows at the annual rate of 0.048%. It then takes off around 1800 and exhibits a sustained growth of nearly 1% per year. In the time period from 1700 to 1950, this experiment accounts for roughly 76% of the increase in per capita GDP in the data. The reason why we are not able to generate a higher sustained growth is possibly the fact that we stop the TFP estimation in 1915 and work under the assumption that the growth rate of TFP remained constant since 1915. If, however, the TFP growth increased after 1915, this experiment would generate higher limiting growth in GDP per capita, which would be more consistent with the data.

It should be clear that since resources are reallocated towards the Solow sector, the land share in total income declines. This happens simply because the Solow sector's land share is 0. The observed changes in capital and labor share are also well captured by this experiment.

See Figures **??** and **??** for land and labor shares. Notice that factor shares in the two technologies are fixed at the calibrated levels. We conclude that changes in TFP alone can account for long term trends in the observed factor income shares. This occurs as a result of resource reallocation between sectors with different factor intensities.

Notice from Figure **??** that changes in the TFP growth rates have a very small quantitative impact on fertility rates. Interestingly, this experiment generates first a rise and then a fall in fertility rates. Recall that productivity increases affect birth rates through two different channels in our model. On one hand, rising productivity translates into higher income. Since surviving offsprings are normal goods, the income effect induces higher birth rate. On the other hand, since the time cost is measured in terms of wages, the opportunity cost of raising children increases. Hence, the substitution effect that puts downward pressure on birth rates is also present here. Clearly, the income effect dominates before the second half of the eighteenth century, and for later years the substitution effect becomes stronger. But in any case, the quantitative effect of changes in the TFP growth rates on fertility is very small, which is a very interesting finding.

Comparison of the population growth rate in the data to the one in the model is similar. As depicted in Figure **??**, starting at the calibrated level of 0.32% annual rate, population growth increases first, but then decreases converging to 0.3% annual rate in the limit. This experiment does generate a small hump in the population growth rate, but the timing of this hump is premature compared to that in the data, and it is quantitatively insignificant.

This experiment leads us to conclude that changes in the productivity in the two sectors represent an important force behind the observed patterns in per capita income, the level of industrialization and urbanization, as well as patterns of labor, capital, and land shares in total income. By contrast, changes in productivity are quantitatively unimportant in driving fertility behavior.

The limiting behavior of equilibrium time paths can be summarized by $y_{t+1}/y_t \to 1.0088$, $\frac{N_{t+1}}{N_t} \to 1.003$, $r_t \to 1.04476$, and $c_t/k_t \to 0.396$, which are given in annualized rates. Only one of the eigen values (0.290631074) for this dynamical system is less than 1. Notice that if the Malthusian technology does not operate, N_t is no longer a state variable for the rest of the system, which means that the only state variable is k_t and exactly one eigen value needs to be less that 1 for local stability. Hence, the Solow balanced growth path to which the equilibrium time paths converge as a result of changes in TFP growth rates is locally stable.

6.2 Changes in Young-Age Mortality

In the second experiment, changes occur in the probability of surviving to the age of 25. The data on these survival probabilities are taken from Wrigley, Davies, Oepenn, and Schofield [36] up to 1841 and from the Human Mortality Database [23] for later years. The series is then smoothed by using a 3 period moving average. The time series of the survival probability used for this experiment is depicted in Figure **??**. Over this time period, survival probability changes from 67% in 1550 to 98% in 2000. Notice that the probability of survival declined until about 1700 when it started to rapidly increase. The most rapid rise began in the second half of the 19th century, most likely related to discoveries of pasteurization in 1864.

Figures **??** and **??** illustrate that changes in young-age mortality have large quantitative effects on fertility behavior. When the probability of survival increases, it becomes less costly to produce a surviving child (q declines). As a result, the number of surviving children always goes up in our model, at least temporarily. On the other hand, fertility always drops since fewer births are needed to achieve the desired number of surviving children. We find that in the period from 1700 to 1950, changes in the probability of survival roughly account for 60% of the drop in the Crude Birth Rate as well as the General Fertility Rate that occurred in England.

Figures **??** and **??** present the time series of the level of urbanization in the model and in the data data but using different time scales. As the probability of survival declines, the time spent on raising surviving children declines, which frees up time available for work. Then the intuition is similar to the statement of the Rybczynski Theorem from trade, which states that as the endowment of one factor increases, the relative output in the sector that uses that factor intensively also rises. Similarly, as more time becomes available for work, the output in the sector that uses labor more intensively (the Solow sector for our calibration) increases relative to the Malthusian output. In the long run, resources reallocate towards the Solow technology, although it takes a very long time. Even in 2400, this experiment predicts that 10% of total labor is still employed by the Malthusian technology. The model versus data patterns of the level of industrialization look very similar to those of urbanization and therefore are not included.

Figure **??** illustrates the result that changes in the probability of survival are quantitatively insignificant in accounting for patterns in GDP per capita. Instead, they account for around 60% of the drop in the Crude Birth Rate and the General Fertility Rate.

The population growth rate does go up from the annual rate of 0.32% to the annual rate of 0.75%, as observed in Figure **??**, but the increase is small quantitatively. What is more bothersome is that neither of the two experiments were able to generate a quantitatively significant hump in the population growth rate despite the fact that it was theoretically possible in our model. We conclude that factors other than young-age mortality and changes in growth rates of total factor productivity are responsible for generating the hump in the population growth rate.

The limiting behavior of equilibrium time paths can be summarized by $y_{t+1}/y_t \to 1.0005$, $\frac{N_{t+1}}{N_t} \to 1.007$, $r_t \to 1.0404$, and $c_t/k_t \to 0.3567$, which are given in annualized rates. Only one of the eigen values (0.29256638) for this dynamical system is less than 1. Hence, the Solow balanced growth path to which the equilibrium time paths converge as a result of changes in young-age mortality is locally stable.

It turns out that there is no special interaction between the two exogenous changes studied in this section in the sense that when a joint experiment is performed with both, TFP growth rates and young-age mortality, changing according to their historical estimates, the effects presented here for the two separate experiments essentially add up. Hence, we do not present the results of the joint experiment.

7 Sensitivity Analysis

Barro and Becker Preferences

As proved in the Appendix, the parental utility that we assumed, $U_t(c_t, n_t, U_{t+1}) = \alpha \log c_t + (1-\alpha)\log n_t + \beta U_{t+1}$, is a special case of the Barro and Becker parental utility, $U_t(c_t, n_t, U_{t+1}) = c_t^{\sigma} + \beta n_t^{1-\varepsilon} U_{t+1}$, if $\sigma \to 0$ and $\frac{1-\varepsilon-\sigma}{\sigma} = \frac{1-\alpha-\beta}{\alpha\beta}$. Notice that this implies that $\varepsilon \to 1$. Hence, a natural question is whether our main results remain if we use the Barro and Becker parental utility form with $\sigma > 0$ and $\varepsilon < 1$.

We follow the same steps to recalibrate the model under the assumption of the Barro and Becker utility, except that this calibration procedure does not pin down both ε and σ. Instead, it pins down $\frac{1-\varepsilon-\sigma}{\sigma} = 0.0233$, thus allowing us the choice of ε. We performed the experiments with different values of ε in the admissible range of $(0,1)$. When $\varepsilon = .9$, which implies that $\sigma = 0.0977$, the results are very close to the results reported for the Lucas utility, that is, they are not very sensitive to the choice of ε. However, here we report the results for a more extreme case, with $\varepsilon = .6$ and the implied $\sigma = 0.39089$.

In short, the main results, that changes in young-age mortality drive mainly the demographic changes while changes in the growth rates of TFP drive all the other quantities, still remain. The only difference is that with Lucas utility (see Figure **??**) changes in the TFP growth rates generate a very slight drop in birth rates, while with Barro and Becker utility and $\varepsilon = .6$, the same experiment generates a slight increase in birth rates for the relevant time period as depicted in Figure **??**.

Figure **??** presents the effect of changes in young-age mortality on CBR. We observe that once again the fall in birth rates is driven by changes in mortality rates although not quite to the same extent as in the case with Lucas utility (see Figure **??**).

Sensitivity to δ, $(\alpha + b)/a$, *and* qn.

All the quantitative results are extremely robust to changes in δ. Since the estimates of δ vary from 2.5% to 15% in the literature as mentioned above, we tried δ in this range. The model v. data plots are practically indistinguishable from the plots in the Simulation Results Section.

Recall that $(\alpha + b)/a$ is an estimate of the average time cost of surviving children relative to that of non-surviving children. In the Appendix we show in detail how we arrive at the estimate of $(\alpha + b)/a = 4$. It is still fair to say that it is unclear what this fraction should be and hence sensitivity analysis is required here. Notice that this fraction only affects the values of a and b, it does not affect the time cost of raising surviving children q. In particular, a decreases and b increases in $(\alpha + b)/a$. The results are again very robust to the assumptions on $(\alpha + b)/a$. Higher $(\alpha + b)/a$ slightly increases the importance of π in driving the birth rates in England while lower $(\alpha + b)/a$ slightly decreases the importance of π. Overall, we tried values for $(\alpha + b)/a$ ranging from 1 to 7 and the results were not affected significantly.

Finally, recall that we set the fraction of time spent raising children, qn, to equal .42. Unfortunately, for $qn \leq .411$, we have $1-\alpha-\beta < 0$, or equivalently $1-\varepsilon-\sigma < 0$ for the Barro-Becker preferences, which implies that the Planner's utility decreases in the size of population. Although this does not mean that the Planner will set the population size to 0 as households would still be valued as a factor of production, we would not be able to guarantee strict concavity of the objective function. Thus, we do not perform any experiments with a value of qn lower than .42. Raising qn to a higher value does not change the results much at all. We analyzed the results when qn is as high as .7.

8 Conclusions

To summarize, we developed a unified framework capable of generating a transition from stagnation to growth. Two exogenous changes appear to have played an important role in this transition. These are a decline in mortality rates and changes in the growth rates of total factor productivity in the rural and non-rural sectors. During the transition we are able to reproduce the stylized facts of the demographic transition: the transition from high levels to low levels of both, mortality and fertility rates. The transition is accompanied by resource reallocation towards the Solow sector which is consistent with the data.

We find that the decline in young-age mortality accounted for 60% of the fall in General Fertility Rate that occurred in England between 1700 and 1950. Over the same time period, changes in productivity accounted for 76% of the increase in GDP per capita and for nearly all of the decline of land share in total output. Interestingly, both experiments generate a transition from Malthus to Solow. However, changes in TFP do so in a manner consistent with empirical observations, driving the share of the Malthusian technology to nearly zero in the period from 1600 to 2000. Changes in the probability of survival generate a much slower transition, predicting that even in 2400 the output produced in the rural sector would still comprise 10% of total output. We also find that changes in TFP alone can account for long term trends in the observed factor income shares. This occurs as a result of resource reallocation between sectors with different factor intensities that remain constant over time.

One of the questions we raised was whether some common forces induced both changes in output and population. In other words, was there a deeper link between the demographic and economic change, or was their joint timing a mere coincidence? Our quantitative results suggest that the explanation for changes in output and population need not be entirely common. In fact, we find that changes in productivity are quantitatively insignificant in accounting for the observed patterns in fertility behavior, while mortality changes are quantitatively relevant only to population dynamics, and not to the other quantities predicted by the model.

An important contribution of this work is the quantitative analysis of the equilibrium time paths within the framework that is capable of generating a transition from Malthusian stagnation to modern growth. We feed the exogenous changes into our model according to historical data. Every period in our model corresponds to a particular date in the data. We perform careful data analysis of the historical time series for England and Wales, working with mortality and fertility data provided by Wrigley, Davies, Oepenn, and Schofield [36], Mitchell [31], and Human Mortality Database [23]. We also estimate total factor productivity in the rural and urban sectors using the dual-approach. This approach requires time series data on wages in the two sectors, land and capital rental rates, as well as the GDP deflator. These time series we either take directly or infer from three of Gregory Clark's papers [10], [8], [9].

We find that solving for the entire transition can lead to results that are quite different from those obtained through comparative statics analysis.

In the near future, we plan to extend our fertility model in two ways, both can be thought of as robustness analysis of our results. In the present model, the goods produced by the Malthusian and the Solow sectors are perfect substitutes. It would be instructive to relax this simplifying assumption and check whether the main results remain. Another way to modify our model would be to give the technological progress a better chance at accounting for fertility behavior. If technological progress was modeled to be capital-biased, changes in productivity would possibly have a stronger effect on fertility via quantity-quality trade-off. Alternatively, if technological progress was modeled to be skill-biased, and parents were allowed to invest into children's skills, we would also expect a stronger effect on fertility. Further sensitivity analysis is needed to test the robustness of our perfect foresight assumption and functional form choices.

9 Appendix

9.1 Data Sources

- Fraction of non-rural labor in total labor, L_2/L

[1565-1865] - Clark, [9]
[1820 - 1992] - Maddison [28], p. 253

- Index of Real GDP per capita, y

[1565-1865] - Clark [9].
[1820-1990] - Maddison [28], p. 194, rescaled to match Clark's index in 1850.

- Labor Share in Total Income, wL/Y

[1585 - 1865] - Clark [9]
[1924 - 1973] - Matthews, Feinstein, Odling-Smee [30], p. 164
Average for [1973 - 1982] - Maddison [27], p. 659
1992 - Gollin [18], p. 19, Table 2, adjustment 3 (average employee compensation used to impute compensation for the entire workforce)

- Land Share in Total Income, $\rho\Lambda/Y$

[1585 - 1865] - Clark [9]
[1873 - 1913] - Matthews, Feinstein, Odling-Smee [30], p.643
[1987 - 1998] - UK National Statistics [34]

- Capital Share in Total Income, $rK/Y = 1 - wL/Y - \rho\Lambda/Y$

- Fraction of non-rural output in total output, Y_2/Y

[1555-1865] - Clark [9]
[1788-1991] - Mitchell [31]

- Crude Birth and Crude Death Rates

[1541 - 1871] - Wrigley, Davies, Oepenn, and Schofield [36]
[1871 - 1986] - Mitchell [31]

- Population Growth Rate

[1541 - 1836] - Wrigley, Davies, Oepenn, and Schofield [36]
[1841 - 1999] - Human Mortality Database [23]

- Age-specific survival probabilities

[1580-1837] - Wrigley, Davies, Oepenn, and Schofield [36]
[1841 - 1999] - Human Mortality Database [23]

- Data used for TFP estimation

Clark [9]
Clark [8]

9.2 Proof of Proposition 3

With this functional form, the Social Planning objective function can be written as

$$u\left(\frac{C_t}{N_t}, \frac{N_{t+1}}{N_t}\right) = \alpha \log C_t + (1-\alpha)\log N_{t+1} - \log N_t$$

and the objective function becomes

$$\begin{aligned} U_0 &= \sum_{t=0}^{\infty} \beta^t \left[\alpha \log C_t + (1-\alpha)\log N_{t+1} - \log N_t\right] \\ &= (\alpha \log C_0 + (1-\alpha)\log N_1 - \log N_0) \\ &\quad +\beta(\alpha \log C_1 + (1-\alpha)\log N_2 - \log N_1) \\ &\quad +\beta^2(\alpha \log C_2 + (1-\alpha)\log N_3 - \log N_2) + ... \\ &= -\log N_0 + \sum_{t=0}^{\infty} \beta^t \alpha \log C_t + \sum_{t=0}^{\infty} \beta^t (1-\alpha-\beta)\log N_{t+1} \\ &= -\log N_0 + \sum_{t=0}^{\infty} \beta^t \left[\alpha \log C_t + (1-\alpha-\beta)\log N_{t+1}\right] \end{aligned}$$

Since N_0 is just a constant, the result holds.

9.3 Sufficiency of First Order, Feasibility, and Transversality Conditions

We first rewrite our Social Planner's problem in the form of a sequential problem in Stokey and Lucas (with Prescott) [33]. Let $x_t \equiv (K_t, N_t)$. Then our problem can be written as

$$\begin{aligned} &\max_{\{x_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t u(x_t, x_{t+1}) \\ &\text{s.t.} \\ x_{t+1} &\in \Gamma(x_t), \ t = 1, 2, ..., \\ x_0 &\geq 0 \text{ given,} \end{aligned}$$

where

$$\Gamma(K_t, N_t) =$$
$$= \left\{(K_{t+1}, N_{t+1}) \in \mathcal{R}_+^2 | 0 \leq K_{t+1} \leq F(K_t, N_t - q_t N_{t+1}) + (1-\delta)K_t, \ 0 \leq N_{t+1} \leq N_t/q_t\right\}$$

We can illustrate the set $\Gamma(K_t, N_t)$ graphically. See Figure **??**.

Notice that the frontier of the set is given by

$$K_{t+1} = F(K_t, N_t - q_t N_{t+1}) + (1-\delta)K_t$$

and we have

$$\begin{aligned} \frac{dK_{t+1}}{dN_{t+1}} &= -q_t F_2\left(K_t, N_t - q_t N_{t+1}\right) < 0 \\ \frac{d^2 K_{t+1}}{dN_{t+1}^2} &= q_t^2 F_{22}\left(K_t, N_t - q_t N_{t+1}\right) < 0 \end{aligned}$$

Thus, the frontier of this set is strictly decreasing and a strictly concave function, so the set $\Gamma\left(K_t, N_t\right)$ is convex. Under the assumptions that we made on u, Theorem 4.15 from Stokey and Lucas (with Prescott) applies and hence, first order and transversality conditions can be used to characterize the solution.

9.4 Proof of Proposition 4

With our choice of the objective function, the transversality conditions for (SP) are given by

$$\begin{aligned} \lim_{t\to\infty} \beta^t \frac{\alpha\left[F_1\left(K_t, N_t - qN_{t+1}\right) + (1-\delta)\right]}{F\left(K_t, N_t - qN_{t+1}\right) + (1-\delta)K_t - K_{t+1}} K_t &= 0, \\ \lim_{t\to\infty} \beta^t \frac{\alpha F_2\left(K_t, N_t - qN_{t+1}\right)}{F\left(K_t, N_t - qN_{t+1}\right) + (1-\delta)K_t - K_{t+1}} N_t &= 0. \end{aligned}$$

Consider the term inside the first transversality condition,

$$\frac{\left[F_1\left(K_t, N_t - qN_{t+1}\right) + (1-\delta)\right]}{F\left(K_t, N_t - qN_{t+1}\right) + (1-\delta)K_t - K_{t+1}} K_t = \frac{\phi Y_{1t} + \theta Y_{2t} + (1-\delta) K_t}{Y_t + (1-\delta)K_t - K_{t+1}}$$

Dividing the numerator and the denominator by the level of aggregate output gives

$$\frac{\phi Y_{1t}/Y_t + \theta Y_{2t}/Y_t + (1-\delta) K_t/Y_t}{1 - \left(K_{t+1} - (1-\delta)K_t\right)/Y_t}$$

The term $K_{t+1} - (1-\delta)K_t$ represents aggregate investment and $Y_t - \left(K_{t+1} - (1-\delta)K_t\right)$ represents aggregate consumption. Notice that there are two ways to violate this T.V.C. Either the numerator goes to ∞ fast enough or the denominator goes to zero fast enough (or both). Clearly, the numerator cannot go to ∞ since $Y_{1t}/Y_t \in [0,1]$, $Y_{2t}/Y_t \in [0,1]$, and $(1-\delta) K_t/Y_t \to \infty$ is impossible (Indeed, suppose $(1-\delta) K_t/Y_t \to \infty$, then $\delta K_t/Y_t \to \infty$ and $\exists t^*$ such that $\forall t > t^*$, $K_{t+1} - K_t = X_t - \delta K_t < 0$. This in turn implies that Y_t is shrinking which violates our assumption). The only way the T.V.C. can be violated is when $1 - \left(K_{t+1} - (1-\delta)K_t\right)/Y_t \to 0$ fast enough, which means that $\frac{C_t}{Y_t}$ goes to zero and again violates our assumption. The argument for the second T.V.C. is similar.

9.5 Cost of Raising Children, Measuring $(a+b)/a$

In this Appendix we explain how we arrived at $(a+b)/a = 4$, the moment that is used in calibration. Recall that this moment represents the average time cost of surviving children relative to that of non-surviving children.

Denote the momentary cost of raising a child by $p(t)$ and the total cost of raising a child to the age of τ by $c(\tau)$, that is,

$$c(\tau) = \int_0^{\tau} p(t)\, dt.$$

Then the expected cost of raising any newborn child is given by

$$\begin{aligned} E[c(\tau)] &= (1-\pi)a + \pi(a+b) = \\ &= (1-\pi)E[c(\tau)|NS] + \pi c(25), \end{aligned}$$

where NS stands for "Not Surviving." Thus,

$$a = \int_0^{25} c(\tau) f(\tau|NS)\, d\tau,$$

where $f(\tau|NS)$ is the density of deaths conditional on not surviving to 25.

Suppose that the instantaneous cost is given by a linear declining function, $p(t) = \alpha - \frac{\alpha}{25}t$. Then $c(\tau) = \alpha\tau - \alpha\frac{\tau^2}{50}$ and $a + b = c(25) = 12.5\alpha$. To find a we use the data around 1600 on age-specific mortality rates.

In Figure **??**, the age groups are 0-1, 1-5, 5-10, 10-15, and 15-25. The first point, for example, illustrates that out of all children who do not survive to the age of 25, 45% die in the first year of their life. This pattern of age-specific mortality conditional on dying before the age of 25 is persistent throughout the years and looks similar even today. It is important to note this here in order to understand that updating a/b as a part of Experiment 2 in this paper (changing the probability of surviving to 25) would not change the results significantly.

If we assume that the costs are incurred in the middle of the above age groups, we get

$$\begin{aligned} a &= 0.45c(0.5) + 0.22c(3) + 0.12c(7.5) + 0.05c(12.5) + 0.16c(20) = 4\alpha \\ b &= 12.5\alpha - 4\alpha = 8.5\alpha \end{aligned}$$

With these assumptions,

$$\frac{b}{a} = 2.125 \text{ and } \frac{a+b}{a} = 3.125$$

If we assume that the costs are incurred in the beginning of the age groups then this ratio is higher (about 3), and hence we have $\frac{a+b}{a} = 4$. It seems more reasonable to assume the latter as once again the data suggests that conditional on dying between any two given ages most children die at younger ages.

9.6 Calibration as a Solution to a System of Linear Equations

We calibrate the model under the assumption that the English economy around 1600 is on a Malthus-Solow BGP, or using the terminology introduced in the main text, we look for parameters $\hat{\theta}$ and initial condition (k_0, N_0) in S_1^*.

The idea is to rewrite the system of equations that describes the properties of a Malthus-Solow BGP (**??**) – (**??**) in terms of the available data moments.

Rewriting the Malthus-Solow BGP equations (??) – (??) in terms of available moments

The objective is to rewrite (**??**) – (**??**) in terms of available moments $\left(\frac{l_1}{l}, \frac{y_1}{y}, \frac{wl}{y}, \frac{rk}{y}\right)$ and parameters only. First, labor share is available and hence it directly pins down μ: $\frac{wl}{y} = \frac{\mu y_1}{l_1}\frac{l}{y} \Rightarrow \mu$. Equation six is $w_t = \frac{\mu y_{1t}}{l_1} = \frac{(1-\theta)y_{2t}}{(1-l_1-qn)}$. Combine it with an algebraic identity to get

$$\begin{aligned}
1-qn &= l_1 + 1 - l_1 - qn \\
wl &= wl_1 + w\left(1-l_1-qn\right) \\
\frac{wl}{y} &= \frac{\mu y_1}{l_1}\frac{l_1}{y} + \frac{(1-\theta)\,y_2}{(1-l_1-qn)}\frac{(1-l_1-qn)}{y} \\
\frac{wl}{y} &= \frac{\mu y_1}{y} + \frac{(1-\theta)\,y_2}{y}
\end{aligned}$$

This pins down θ. Next, equation five is just $r \equiv \frac{\phi y_1}{\rho_k k} = \frac{\theta y_2}{(1-\rho_k)k}$. Combining it with an algebraic identity we get

$$\begin{aligned}
k &= \rho_k k + (1-\rho_k)\,k \\
rk &= r\rho_k k + r\,(1-\rho_k)\,k \\
rk &= \frac{\phi y_1}{\rho_k k}\rho_k k + \frac{\theta y_2}{(1-\rho_k)\,k}\,(1-\rho_k)\,k \\
rk &= \phi y_1 + \theta y_2 \\
\frac{rk}{y} &= \frac{\phi y_1}{y} + \frac{\theta y_2}{y}
\end{aligned}$$

This allows us to get γ_1 and γ_2 and give prediction to γ and n. Then $\gamma = \frac{\beta}{n}\left[r+1-\delta\right]$ can be used together with $r+1-\delta$ in the data to get β. We then use the moment qn to get $a/\pi + b$. Separately employing the assumption on a/b and π we calibrate a and b. Finally, we combine equations four and seven

$$\begin{aligned}
\frac{(1-\alpha-\beta)}{\alpha n}\frac{\rho\phi l_1}{\mu r\rho_k} &= q - \frac{\gamma}{(r+1-\delta)} \\
\rho + \gamma n &= \frac{y}{k} + (1-\delta)
\end{aligned}$$

to get

$$\begin{aligned}
\frac{(1-\alpha-\beta)}{\alpha\mu}\frac{\phi l_1}{r\rho_k}\left(r\frac{y}{rk} + (1-\delta) - \gamma n\right) &= qn - \frac{\gamma n}{(r+1-\delta)} \\
\frac{(1-\alpha-\beta)\,(1-qn)}{\alpha\mu}\frac{\phi}{r\rho_k}\frac{l_1}{(1-qn)}\left(r\frac{y}{rk} + (1-\delta) - \gamma n\right) &= qn - \frac{\gamma n}{(r+1-\delta)}
\end{aligned}$$

Since

$$\begin{aligned}
r &= \frac{\phi y_1}{k_1} \\
\frac{k}{y_1} &= \frac{\phi}{r\rho_k}
\end{aligned}$$

and hence, the above becomes

$$\frac{(1-\alpha-\beta)(1-qn)}{\alpha\mu}\frac{y}{y_1}\frac{1}{r}\frac{rk}{y}\frac{l_1}{(1-qn)}\left(r\frac{y}{rk}+(1-\delta)-\gamma n\right)=qn-\frac{\gamma n}{(r+1-\delta)}$$

which pins down α.

We can think of solving the last equation as solving 2 equations in two unknowns: ρ and α. Hence, we solve the following equations

$$\frac{wl}{y} = \frac{\mu y_1}{l_1}\frac{l}{y} \tag{34}$$

$$\frac{wl}{y} = \frac{\mu y_1}{y}+\frac{(1-\theta)\,y_2}{y} \tag{35}$$

$$\frac{rk}{y} = \frac{\phi y_1}{y}+\frac{\theta y_2}{y} \tag{36}$$

$$\gamma = \gamma_2^{\frac{1}{1-\theta}} \tag{37}$$

$$n = \left(\gamma_1\gamma_2^{-\frac{1-\phi}{1-\theta}}\right)^{\frac{1}{1-\phi-\mu}} \tag{38}$$

$$\gamma = \frac{\beta}{n}\left[r+1-\delta\right] \tag{39}$$

$$\frac{qn}{n} = \frac{b}{\pi}\frac{a}{b}+b \tag{40}$$

$$a = b\frac{a}{b} \tag{41}$$

$$\frac{(1-\alpha-\beta)(1-qn)}{\alpha\mu}\frac{y}{y_1}\frac{1}{r}\frac{rk}{y}\frac{l_1}{(1-qn)}\rho = qn-\frac{\gamma n}{(r+1-\delta)} \tag{42}$$

$$\rho+\gamma n = r\frac{y}{rk}+(1-\delta) \tag{43}$$

Hence calibration can be summarized as a solution to a system of linear equations. In the system of linear equations above, $\pi,\delta,\gamma_1,\gamma_2$ are directly pinned down in the data, although γ_1 and γ_2 are pinned down only ϕ,μ,θ are determined. The system of equations consists of 10 equations in terms of 10 unknowns, 7 of which are parameters, μ, ϕ, θ, β, a, b, α, and 3 of which are moments that we do not take from the data: $\frac{c}{k}$, γ, n. Moments used are $\frac{wl}{y}$, $\frac{rk}{y}$, $\frac{y_1}{y}$, $\frac{l_1}{l}$, qn, $\frac{a}{b}$, r.

9.7 Mapping of the Model to the Data

Population Growth: Beginning and End of Period Calculations

Beginning of period t : $2N_t$ adults and $2f_tN_t$ live births.

End of period t : $2N_t$ adults and $2\pi_tf_tN_t$ children.

The link between the two periods is $N_{t+1}=\pi_tf_tN_t$.

Population growth when counted from the beginning of one period to the beginning of the next period is

$$\frac{2N_{t+1}+2f_{t+1}N_{t+1}}{2N_t+2f_tN_t}=\frac{2\pi_tf_tN_t+2f_{t+1}\pi_tf_tN_t}{2N_t+2f_tN_t}=\frac{\pi_tf_t\,(1+f_{t+1})}{1+f_t}=\frac{n_t\,(1+f_{t+1})}{1+f_t}.$$

Population growth when counted from the end of one period to the end of the next period is

$$\frac{2N_{t+1}+2\pi_{t+1}f_{t+1}N_{t+1}}{2N_t+2\pi_t f_t N_t}=\frac{2\pi_t f_t N_t+2\pi_{t+1}f_{t+1}\pi_t f_t N_t}{2N_t+2\pi_t f_t N_t}=\frac{n_t\left(1+n_{t+1}\right)}{1+n_t}.$$

Population Growth: Average Over Time Period Calculation

We need to estimate the average level of population in period t.

The number of adults is unchanged over the duration of a period.

The number of children does change. We use age-specific mortality rates to determine the average level of population.

Average number of children during period t is

$$\left(\frac{1}{25}+\frac{4}{25}\pi_0^1+\frac{5}{25}\pi_0^5+\frac{5}{25}\pi_0^{10}+\frac{10}{25}\pi_0^{15}\right)2f_tN_t.$$

Average population over period t is

$$POP_t=2N_t+\left(\frac{1}{25}+\frac{4}{25}\pi_0^1+\frac{5}{25}\pi_0^5+\frac{5}{25}\pi_0^{10}+\frac{10}{25}\pi_0^{15}\right)2f_tN_t.$$

Total Fertility Rate,General Fertility Rate, and Crude Birth Rate

Crude Birth Rate in the data is

$$CBR=1000\frac{\#\text{births}}{\text{Population}}.$$

Crude Birth Rate in the model is given by

$$CBR_t=1000\frac{2f_tN_t}{POP_t}=1000\frac{f_t}{1+\left(\frac{1}{25}+\frac{4}{25}\pi_0^1+\frac{5}{25}\pi_0^5+\frac{5}{25}\pi_0^{10}+\frac{10}{25}\pi_0^{15}\right)f_t}.$$

We want to determine the relationship between CBR over 25 years and CBR over 1 year.

$$\begin{aligned}CBR^1 &= 1000\frac{\#\text{births in 1 year}}{\text{Population}}\text{ and }CBR^{25}=1000\frac{\#\text{births in 25 year}}{\text{Population}},\\ CBR^{25} &= 1000\left(\frac{B_1+nB_1+...+n^{24}B_1}{\frac{1}{25}\left(N_1+nN_1+...+n^{24}N_1\right)}\right)=25\left(1000\frac{B_1}{N_1}\right).\end{aligned}$$

So, assuming population and number of newborns grow at some constant rate, we have $25\left(CBR^1\right)=CBR^{25}$.

TFR is defined as the number of births that a young woman would have during her fertile years if she followed age-specific fertility rates for that year. We do not have data for total fertility rate, instead, we compare the model to general fertility rate.

General fertility rate is

$$\begin{aligned}GFR &= 1000\frac{\#\text{births}}{\#\text{fertile women}}=1000\frac{\#\text{births/Population}}{\#\text{fertile women/Population}}=\\ &= \frac{\text{CBR}}{\text{fraction of fertile women in population}}.\end{aligned}$$

GFR in the model is

$$GFR_t = 1000\left(\frac{2f_tN_t}{POP_t}/\frac{N_t}{POP_t}\right) = 2000f_t.$$

Similarly to CBR, we have $25GFR^1 = GFR^{25}$.

Notice that GFR is different from TFR. They are only the same if age specific fertility rates are the same.

9.8 Barro and Becker v. Lucas Utility

Proposition 6 *Parental utility used by Lucas [26],*

$$U_t\left(c_t, n_t, U_{t+1}\right) = \alpha \log c_t + (1-\alpha)\log n_t + \beta U_{t+1},$$

represents the same preferences as represented by the Barro and Becker utility

$$U_t\left(c_t, n_t, U_{t+1}\right) = c_t^\sigma + \beta n_t^{1-\varepsilon}U_{t+1}$$

if $\sigma \to 0$ *and* $\frac{1-\varepsilon-\sigma}{\sigma} = \frac{1-\alpha-\beta}{\alpha\beta}$.

Proof. Let $\frac{1-\varepsilon-\sigma}{\sigma} = \frac{1-\alpha-\beta}{\alpha\beta}$. Consider the following transformation of the Barro and Becker utility, $W_t\left(c_t, n_t, U_{t+1}\right) = (1-\beta)\,U_t\left(c_t, n_t, U_{t+1}\right)$, i.e.,

$$W_t\left(c_t, n_t, W_{t+1}\right) = (1-\beta)\,c_t^\sigma + \beta n_t^{1-\varepsilon}W_{t+1}.$$

Next consider another transformation, $V_t\left(c_t, n_t, W_{t+1}\right) = W_t\left(c_t, n_t, W_{t+1}\right)^{\frac{\alpha}{(1-\beta)\sigma}}$, i.e.

$$\begin{aligned} V_t\left(c_t, n_t, V_{t+1}\right) &= \left[(1-\beta)\,c_t^\sigma + \beta n_t^{1-\varepsilon}V_{t+1}^{\frac{(1-\beta)\sigma}{\alpha}}\right]^{\frac{\alpha}{(1-\beta)\sigma}} = \\ &= \left[(1-\beta)\,c_t^\sigma + \beta\left(n_t^{\frac{1-\varepsilon}{\sigma}}V_{t+1}^{\frac{(1-\beta)}{\alpha}}\right)^\sigma\right]^{\frac{\alpha}{(1-\beta)\sigma}} = \\ &= \left(\left[(1-\beta)\,c_t^\sigma + \beta\left(n_t^{\frac{1-\varepsilon}{\sigma}}V_{t+1}^{\frac{(1-\beta)}{\alpha}}\right)^\sigma\right]^{\frac{1}{\sigma}}\right)^{\frac{\alpha}{(1-\beta)}}. \end{aligned}$$

Now taking limits as $\sigma \to 0$ while ε changes so that $\frac{1-\varepsilon-\sigma}{\sigma} = \frac{1-\alpha-\beta}{\alpha\beta}$ we have

$$\begin{aligned} \lim_{\sigma\to 0} V_t\left(c_t, n_t, V_{t+1}\right) &= \\ &= \left(\lim_{\sigma\to 0}\left[(1-\beta)\,c_t^\sigma + \beta\left(n_t^{\frac{1-\varepsilon}{\sigma}}V_{t+1}^{\frac{(1-\beta)}{\alpha}}\right)^\sigma\right]^{\frac{1}{\sigma}}\right)^{\frac{\alpha}{(1-\beta)}} = \\ &= \left(c_t^{1-\beta}\left(n_t^{\frac{1-\varepsilon}{\sigma}}V_{t+1}^{\frac{(1-\beta)}{\alpha}}\right)^\beta\right)^{\frac{\alpha}{(1-\beta)}}. \end{aligned}$$

Notice that $n_t^{\frac{1-\varepsilon}{\sigma}}$ and $V_{t+1}^{\frac{(1-\beta)}{\alpha}}$ remain fixed as $\sigma \to 0$. Finally, consider another transformation, $U_t(c_t, n_t, V_{t+1}) = \log V_t(c_t, n_t, V_{t+1})$,

$$U_t(c_t, n_t, U_{t+1}) = \frac{\alpha}{(1-\beta)}\left[(1-\beta)\log c_t + \frac{1-\varepsilon}{\sigma}\beta \log n_t + \frac{(1-\beta)}{\alpha}\beta U_{t+1}\right]$$

Simplifying and using our assumption that $\frac{1-\varepsilon-\sigma}{\sigma} = \frac{1-\alpha-\beta}{\alpha\beta}$, i.e., $\frac{1-\varepsilon}{\sigma} = \frac{(1-\alpha)(1-\beta)}{\alpha\beta}$, gives

$$\begin{aligned} U_t(c_t, n_t, U_{t+1}) &= \alpha \log c_t + \frac{\alpha}{(1-\beta)}\frac{(1-\alpha)(1-\beta)}{\alpha\beta}\beta \log n_t + \beta U_{t+1} \\ &= \alpha \log c_t + (1-\alpha)\log n_t + \beta U_{t+1} \end{aligned}$$

Therefore, the parental utility function assumed by Lucas and in this paper is a special case of the Barro-Becker utility. ■

References

[1] Allen, Robert C, "The Efficiency and Distributional Consequences of Eighteenth Century Enclosures," *The Economic Journal*, **92**, 1982, p. 937-953

[2] Allen, Robert C, *Enclosure and the Yeoman*, Clarendon Press, Oxford, 1992

[3] Barro, R. and G.S. Becker, "A Reformulation of the Economic Theory of Fertility," *Quarterly Journal of Economics,* **103**, 1988, p. 1-25

[4] Barro, R. and G.S. Becker, "Fertility Choice in a Model of Economic Growth," *Econometrica*, **57**, 1989, p. 481-501

[5] Becker, G., K.M. Murphy and R. Tamura, "Human Capital, Fertility, and Economic Fertility," *Journal of Political Economy*, **98**, 1990, p. 812-837

[6] Boldrin, M. and L., Jones, "Mortality, Fertility and Savings Decisions," *Review of Economic Dynamics,* **5**, 2002, 775-814

[7] Boldrin, M., DeNardi, M., Jones, L. "Fertility and Social Security," Working Paper, University of Minnesota, 2005

[8] Clark, Gregory, "The Agricultural Revolution and the Industrial Revolution: England, 1500-1912," Working paper, UC Davis, 2002

[9] Clark, Gregory, "The Secret History of the Industrial Revolution," Working paper, UC Davis, 2001

[10] Clark, Gregory, "Debt, Deficits, and Crowding out: England, 1727-1840," Working paper, UC Davis, 2002

[11] Doepke, M., "Child Mortality and Fertility Decline: Does the Barro-Becker Model fit the Facts?" Working Paper, UCLA, 2004

[12] Doepke, M., "Accounting for Fertility Decline during the Transition to Growth," *Journal of Economic Growth* **9(3)**, 2004, p. 347-383

[13] Galor, Oded "From Stagnation to Growth: Unified Growth Theory," Forthcoming in the Handbook of Economic Growth (P. Aghion and S. Durlauf eds.), North-Holland, 2005

[14] Galor, O. and D.N. Weil, "From the Malthusian Regime to Modern Growth," *American Economic Review*, **89**, 1999, p. 150-154

[15] Galor, O. and D.N. Weil, "Population, Technology, and Growth: From the Malthusian Regime to the Demographic Transition," *American Economic Review*, **110**, 2000, p. 806-828

[16] Golosov, M., Jones, L., and Tertilt, M "Efficiency with Endogenous Population Growth," NBER Working Paper 10231, 2005

[17] Greenwood, J., and A. Seshadri "The U.S. Demographic Transition," *American Economic Review (Papers and Proceedings)*, 2002, v.92, n.2, p.153-159

[18] Gollin, Douglus, "Getting Income Shares Right," mimeo, Williams College, 2001

[19] Ehrlich, I. and F. Lui, "Intergenerational Trade, Longevity, Intrafamily Transfers and Economic Growth," *Journal of Political Economy,* **99**, 1991, p. 1029-59

[20] Fernandez-Villaverde, J., "Was Malthus Right? Economic Growth and Population Dynamics," Working Paper, University of Pennsylvania, 2001

[21] Jones, Charles, "Was an Industrial Revolution Inevitable? Economic Growth over the Very Long Run," Working paper, UC Berkeley, 2001

[22] Hansen, G. and E. Prescott, "Malthus to Solow," *American Economic Review,* **93,** 2003, p. 1205-1217

[23] Human Mortality Database, http://www.mortality.org

[24] Hoffman, P.T., *Growth in a Traditional Society*, Princeton, Princeton University Press, 1996

[25] Livi-Bacci, Massimo, *Population and Nutrition: An Essay on European Demographic History,* Cambridge University Press, 1991

[26] Lucas, R., "The Industrial Revolution: Past and Future," in *Lectures on Economic Growth*, Harvard University Press, 2002

[27] Maddison, Angus, "Growth and Slowdown in Advanced Capitalist Economies: Techniques of Quantitative Assessment," *Journal of Economic Literature*, **XXV**, 1987, p. 649-698

[28] Maddison, Angus, *Monitoring the World Economy: 1820/1992*, OECD Development Center, 1995

[29] Malthus, T., *Population: The First Essay,* 1798, Reprinted by University of Michigan Press, 1959

[30] Matthews, R.C.O., Feinstein, C.H. and J.C. Odling-Smee, *British Economic Growth: 1856-1973,* Stanford University Press, Stanford, CA, 1982

[31] Mitchell, B.R., *European Historical Statistics 1750-1970*, Columbia University Press, Columbia, NY, 1978

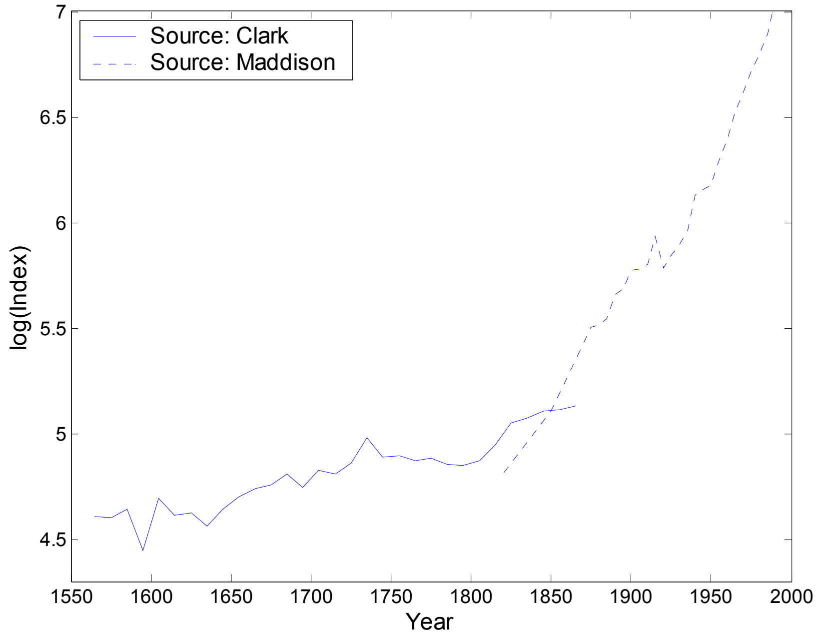
[32] Razin, A and U. Ben-Zion, "An Intergenerational Model of Population Growth," *American Economic Review,* **69**, p. 923-933

[33] Stokey, N. L. and R. E. Lucas (with Edward C. Prescott), *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, MA and London, England, 1989

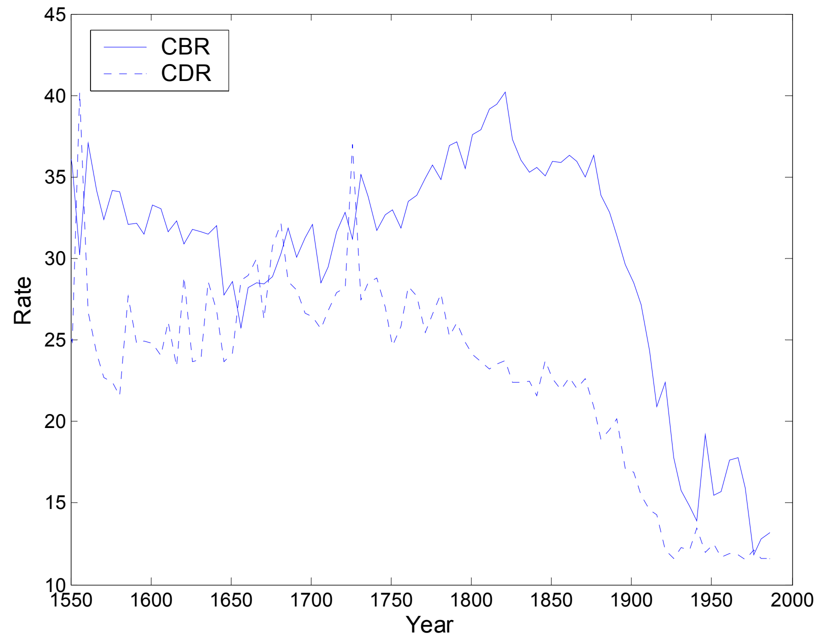
[34] UK National Statistics, http://www.statistics.gov.uk/

[35] Wrigley, E.A. and R.S. Schofield, *The Population History of England, 1541-1871: A Reconstruction*, Cambridge University Press, Cambridge, 1981

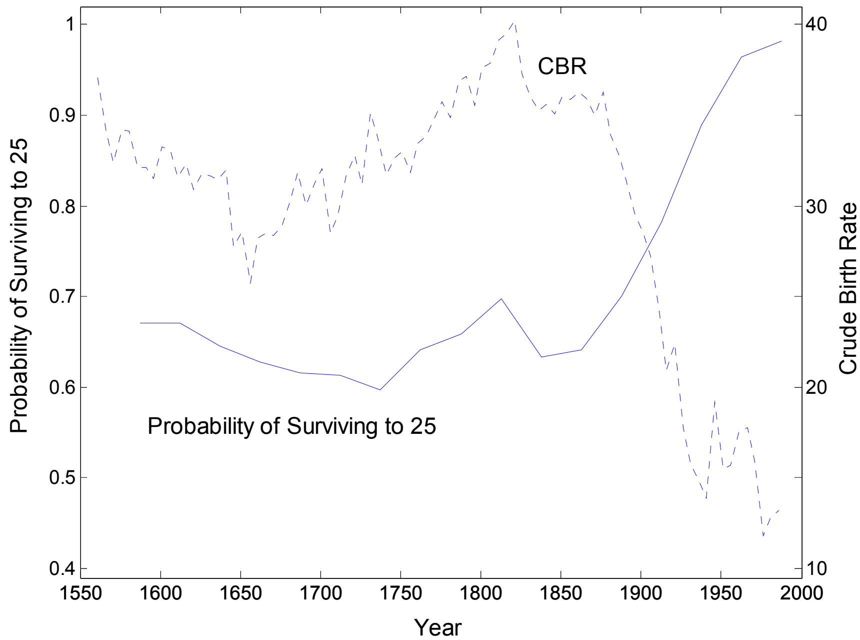
[36] Wrigley, E.A., Davies, R.S., Oeppen, J.E., and R.S. Schofield, *English Population History from Family Reconstitution 1580-1837*, Cambridge University Press, Cambridge, 1997



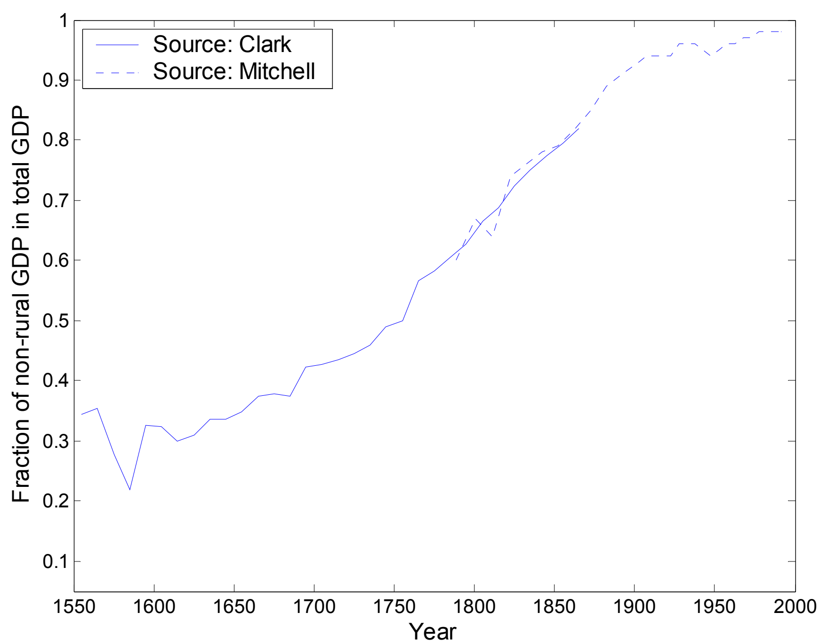
England and Wales: Log(Index of Real GDP per Capita)



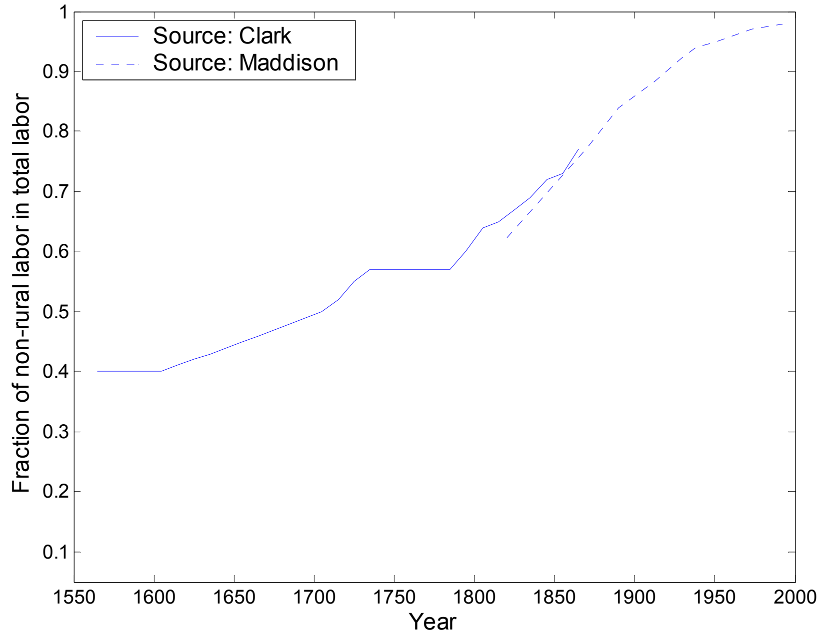
England and Wales: Demographic Transition



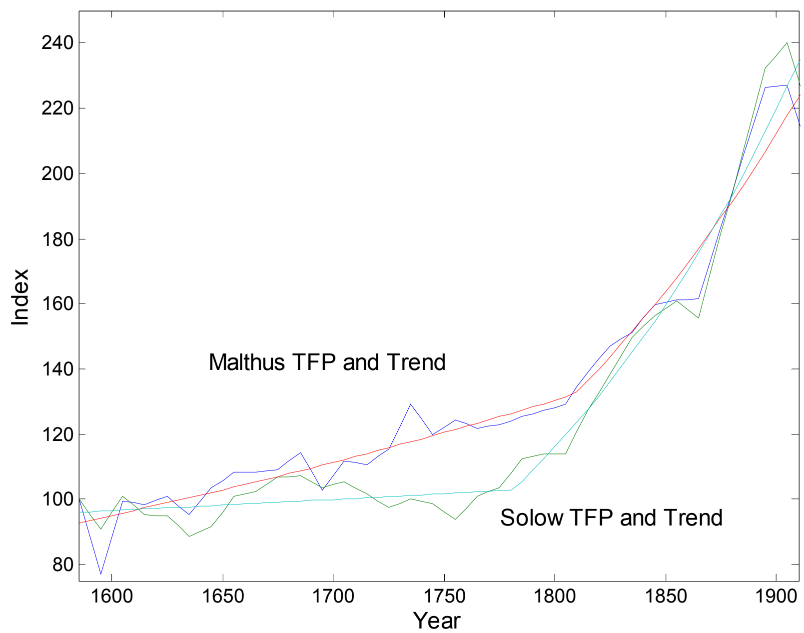
England and Wales: Birth Rates and Young-Age Mortality



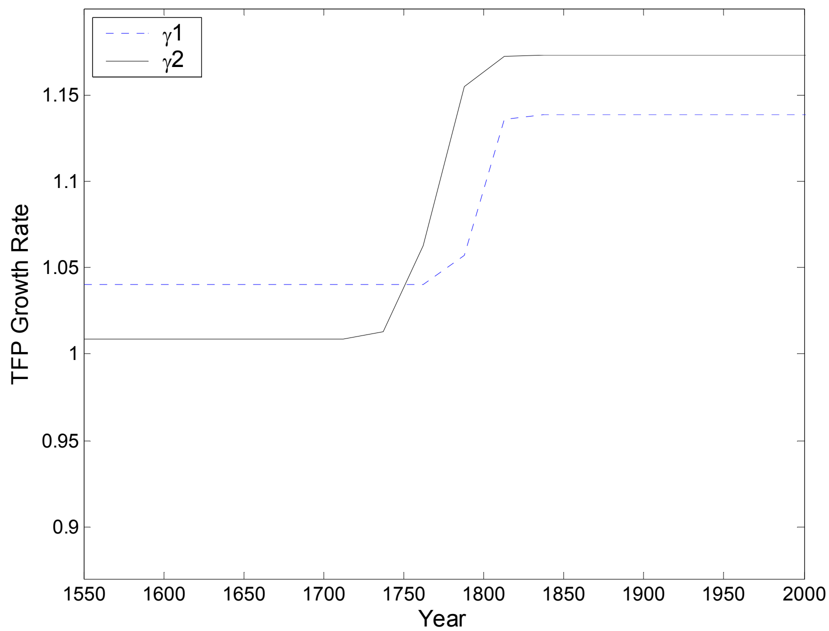
England and Wales: Industrialization



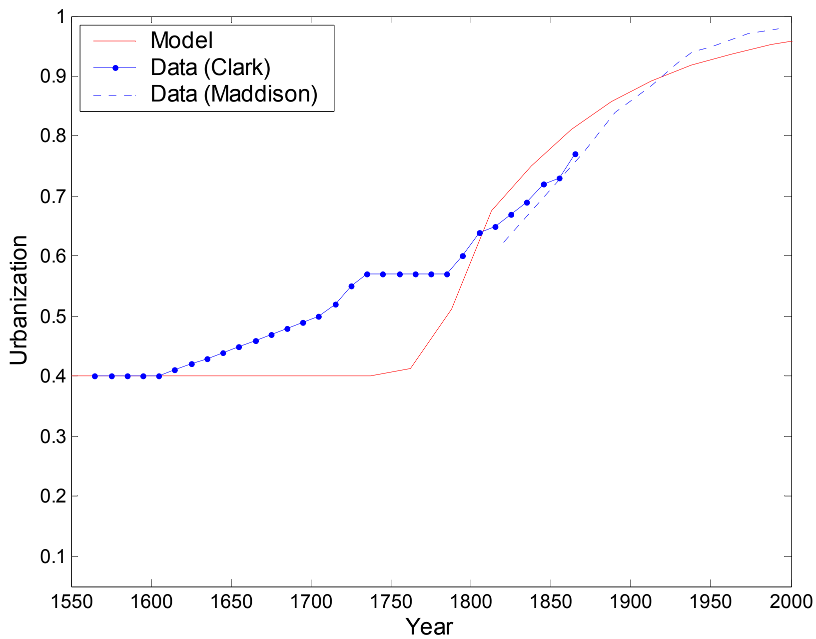
England and Wales: Urbanization



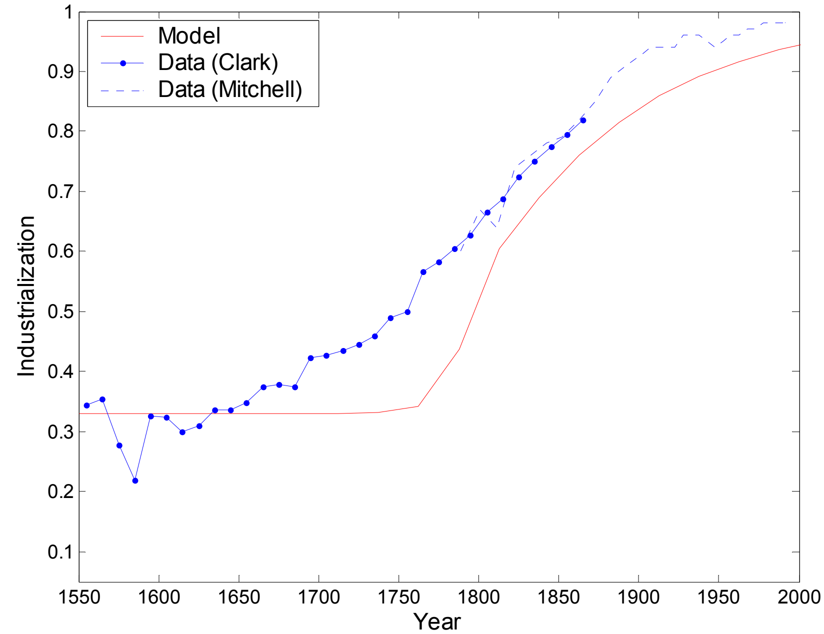
Estimated Malthus and Solow TFP



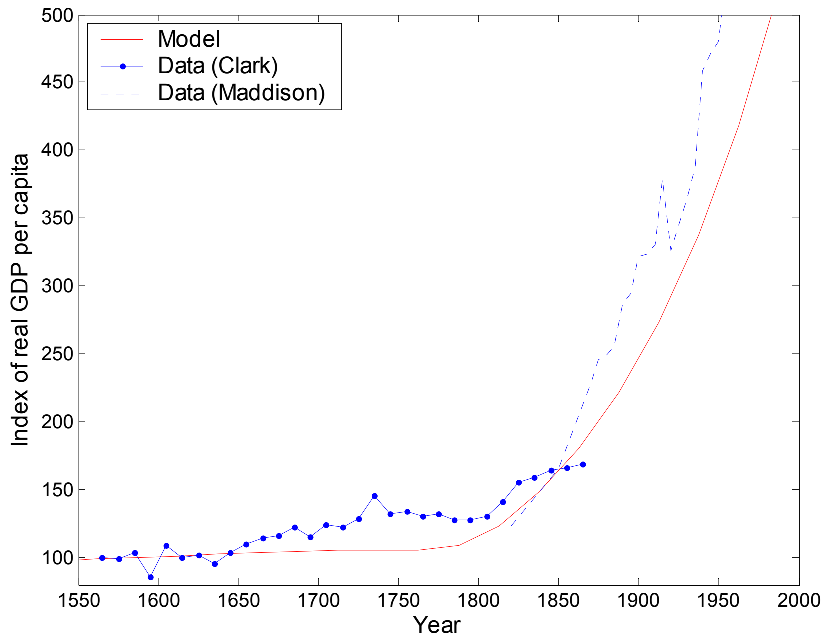
Experiment 1: Changing γ_1 and γ_2



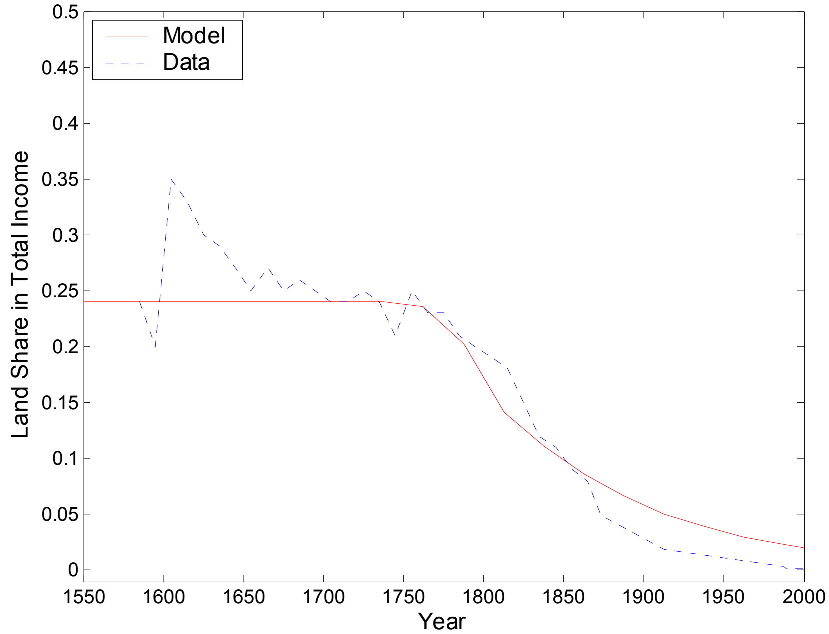
Experiment 1, Model v. Data: Urbanization



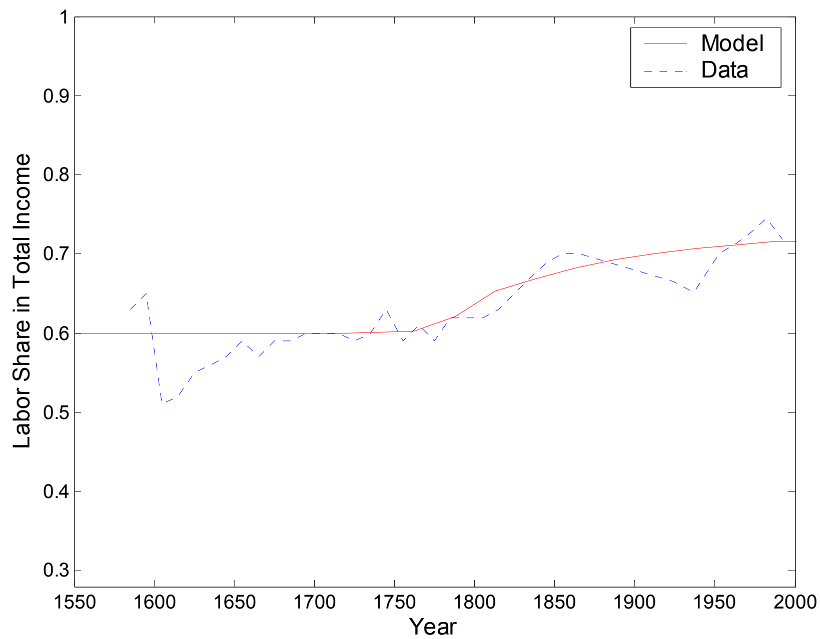
Experiment 1, Model v. Data: Industrialization



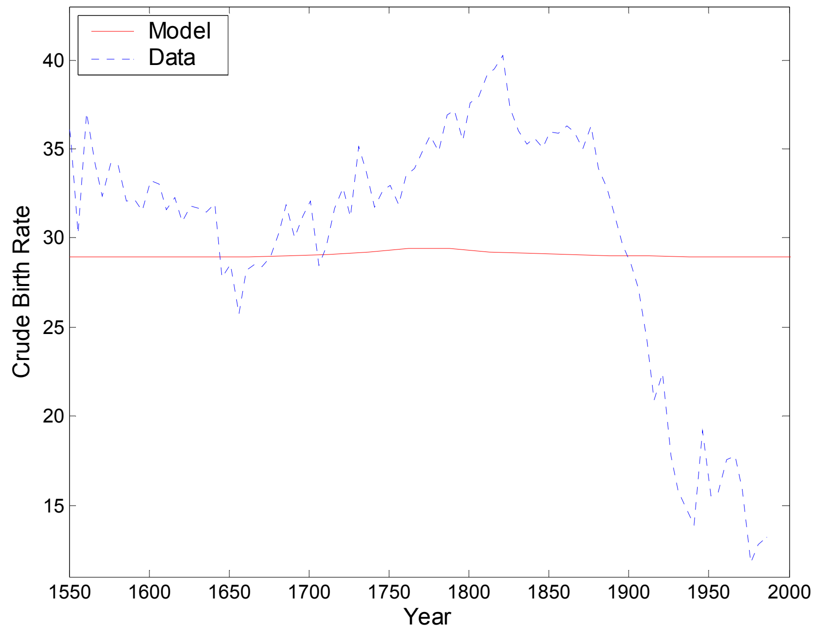
Experiment 1, Model v. Data: Index of Real GDP per Capita



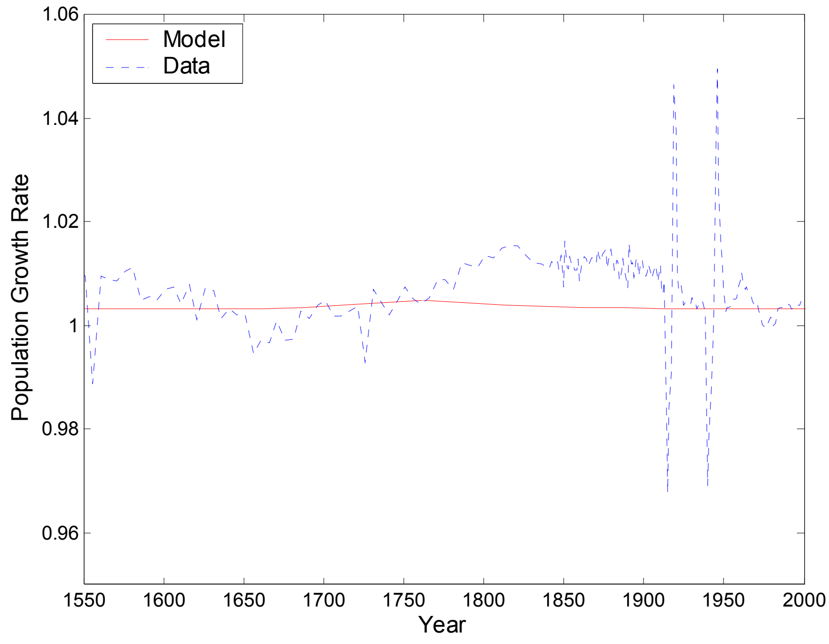
Experiment 1, Model v. Data: Land Share in Total Income



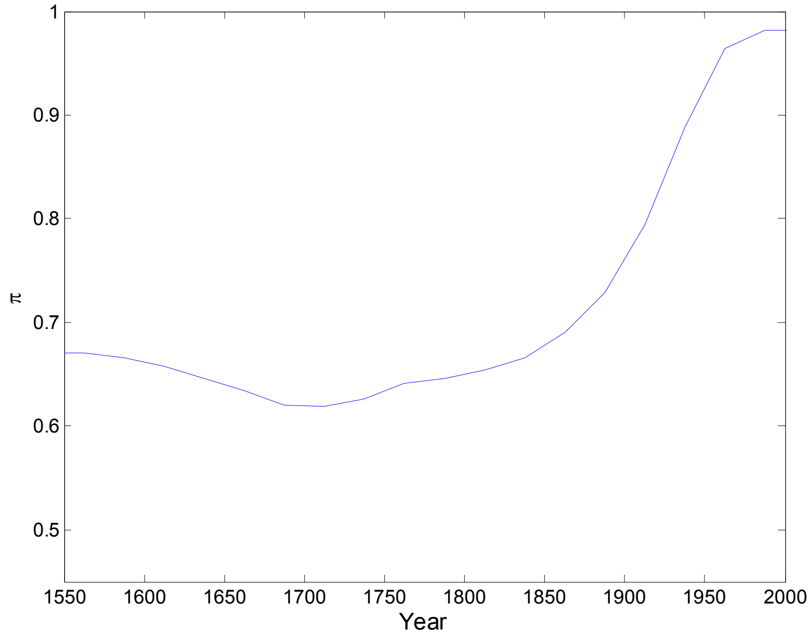
Experiment 1, Model v. Data: Labor Share in Total Income



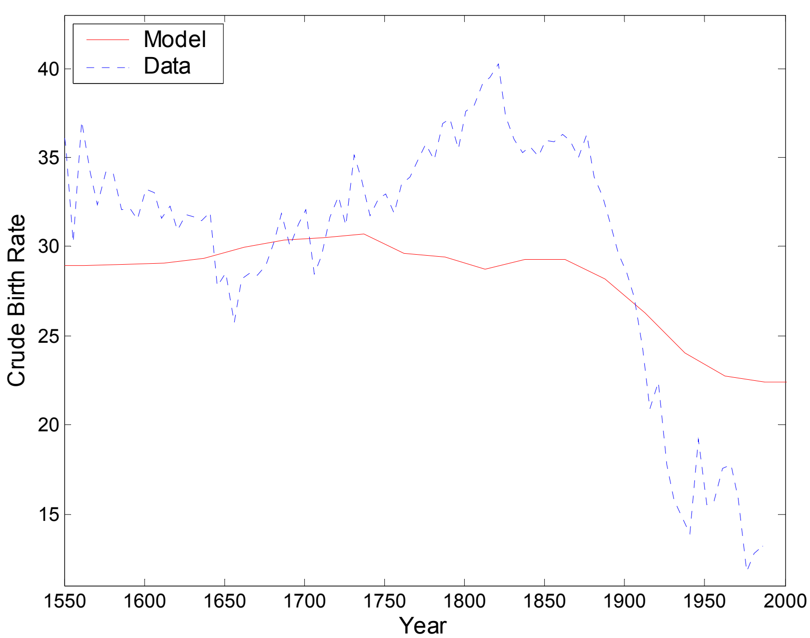
Experiment 1, Model v. Data: Crude Birth Rate



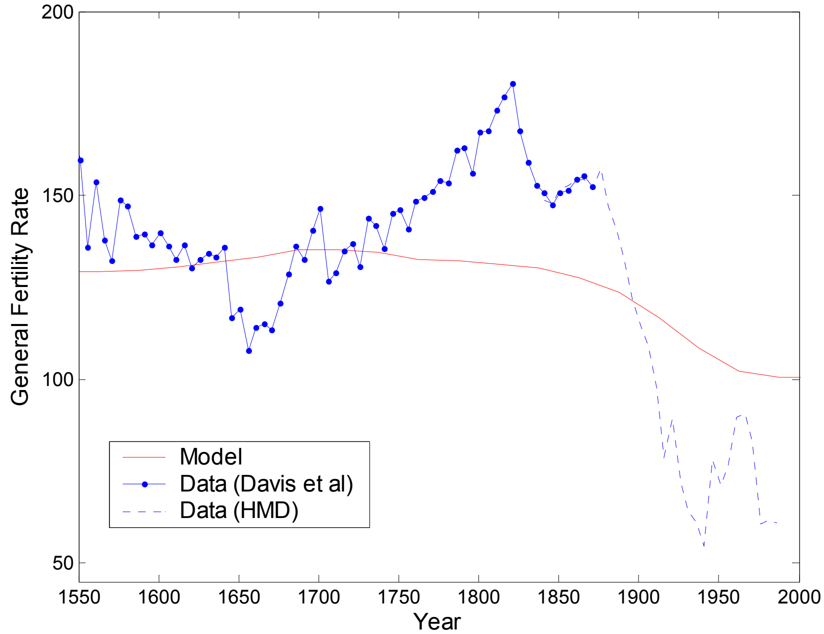
Experiment 1, Model v. Data: Population Growth Rate



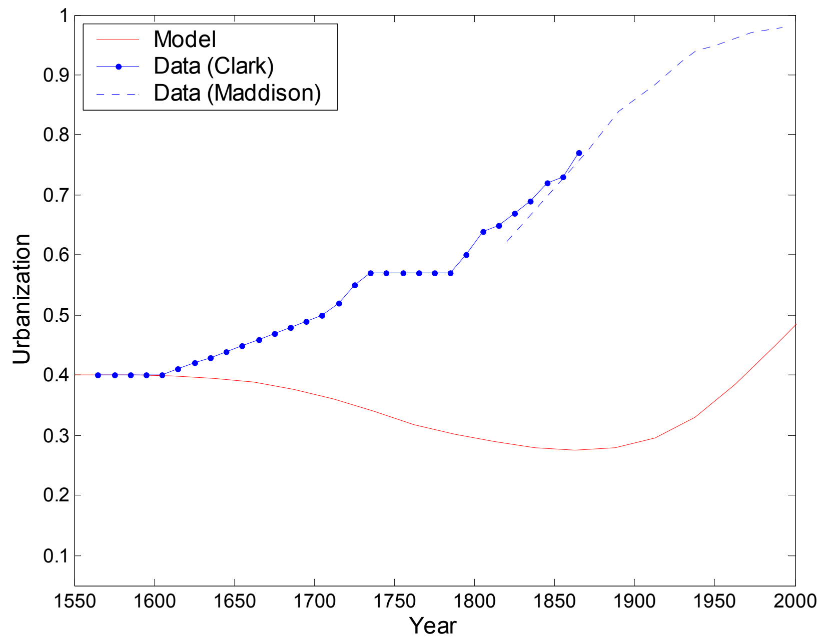
Experiment 2: Changing π



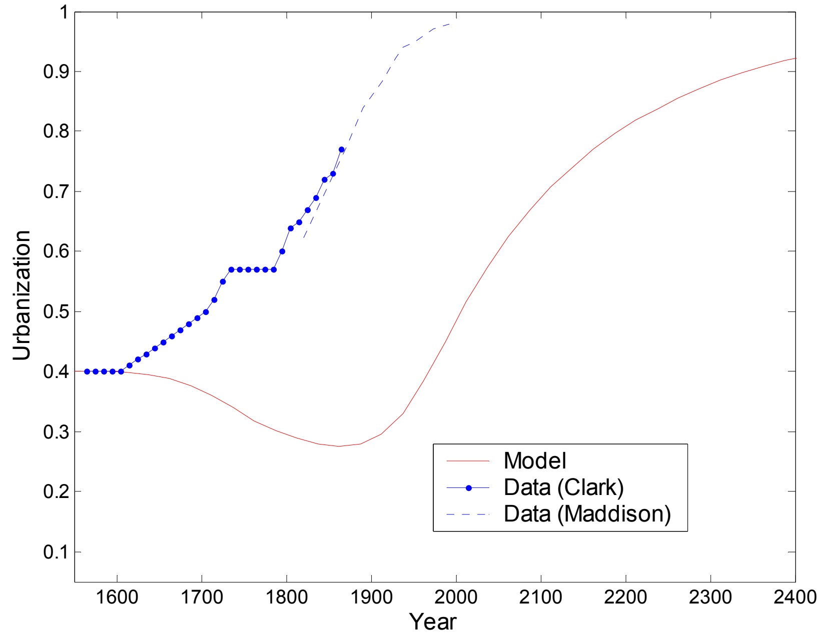
Experiment 2, Model v. Data: Crude Birth Rate



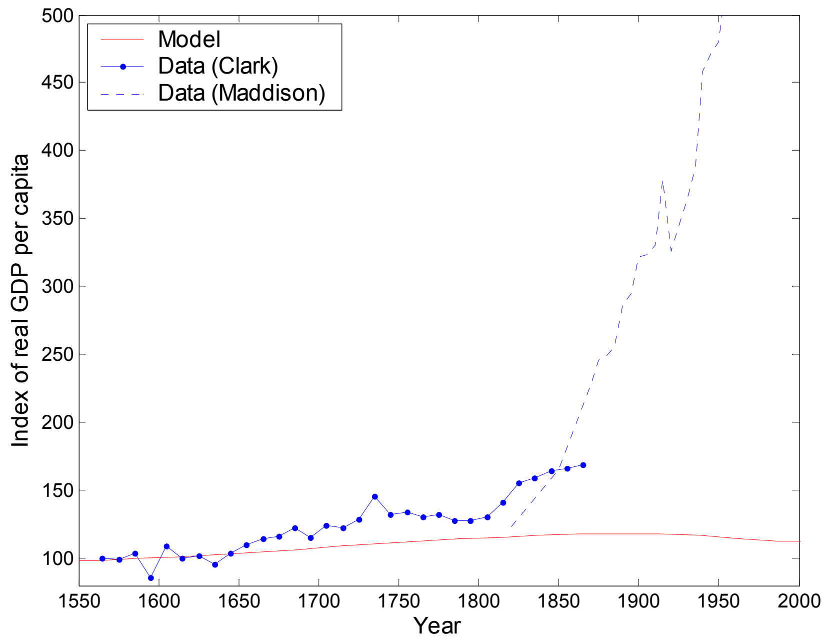
Experiment 2, Model v. Data: General Fertility Rate



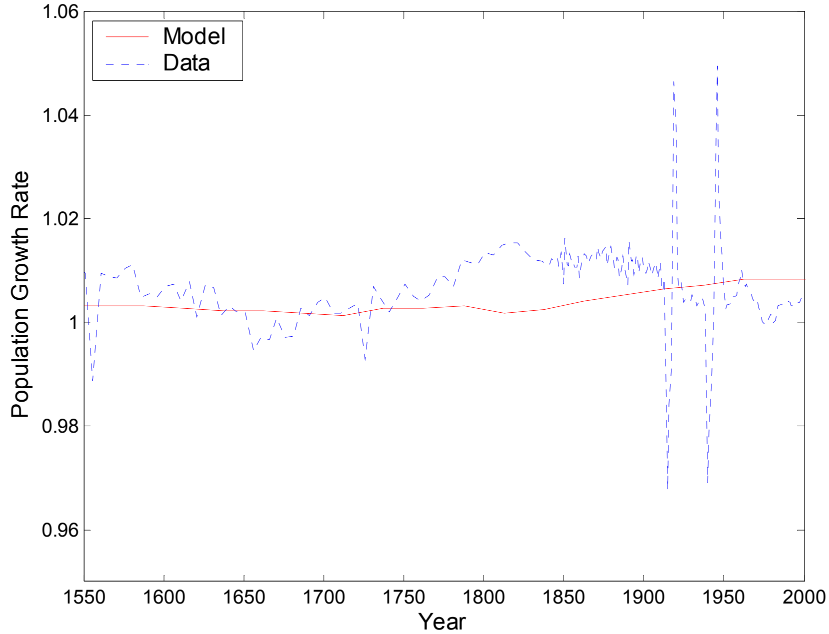
Experiment 2, Model v. Data: Urbanization



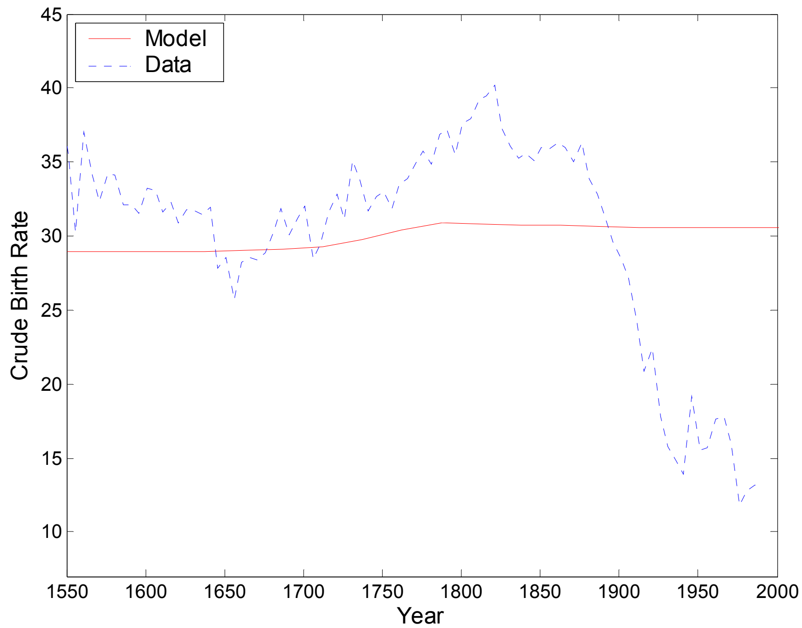
Experiment, Model v. Data: Urbanization, Longer Time Scale



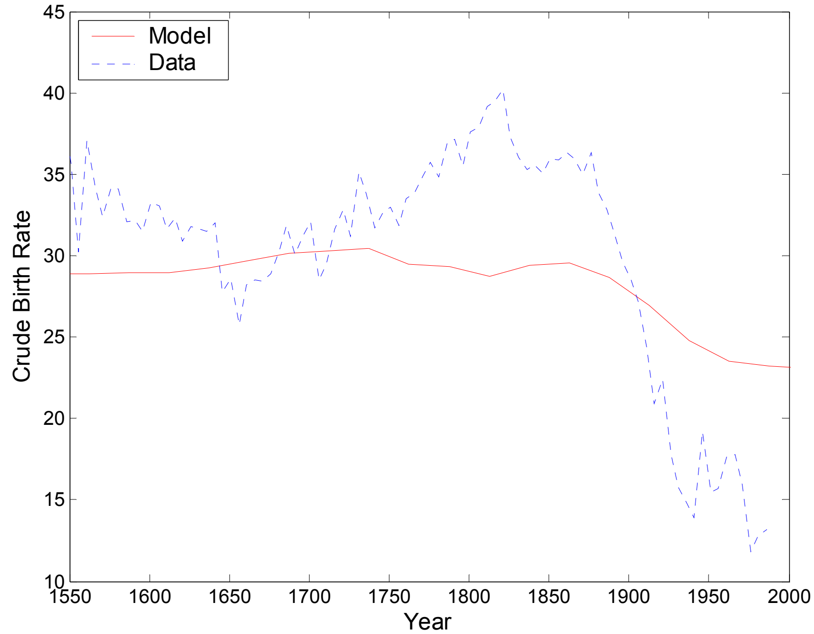
Experiment 2, Model v. Data: Index of Real GDP per Capita



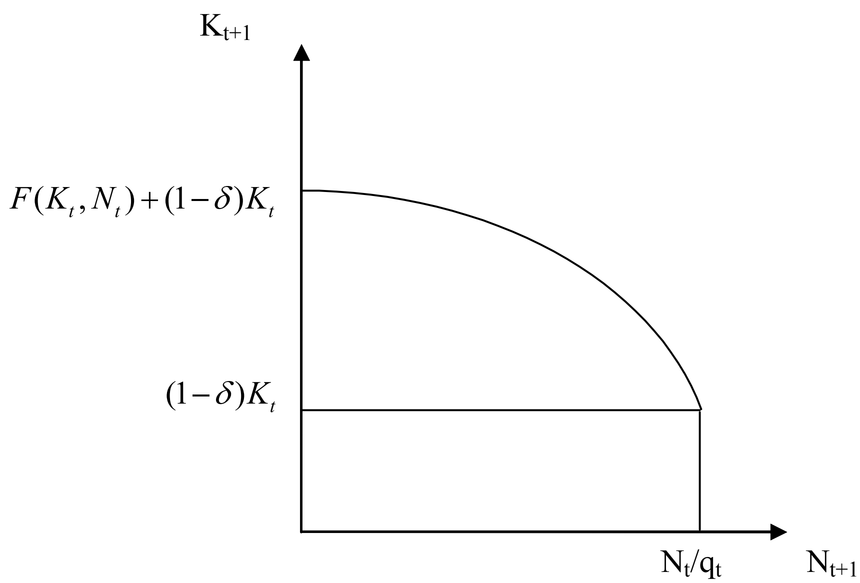
Experiment 2, Model v. Data: Population Growth Rate



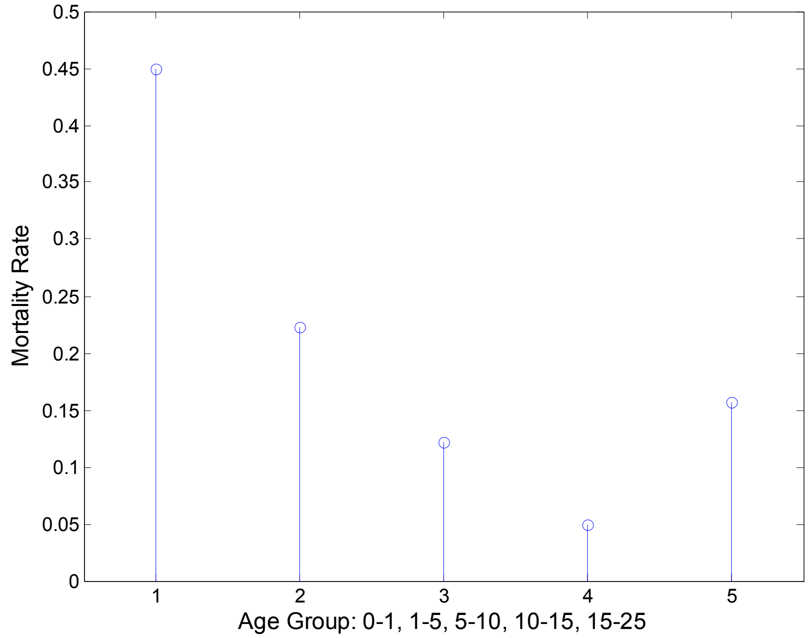
Barro and Becker Utility with $\varepsilon = .6$, Experiment 1: CBR



Barro and Becker Utility with $\varepsilon = .6$, Experiment 2: CBR



Constraint Set at time t



Age-Specific Mortality Conditional on Not Surviving to 25.