Bringing Unified Growth Theory to the Data

James Foreman-Peck[*] and Peng Zhou[†]

Abstract

A unified growth model is estimated from English economy data for up to six hundred years. At the core of the overlapping generations, rational expectations model is household choice about target number and quality of children, as well as female age at marriage. The moments of births, deaths, population and the real wage, are closely matched by the estimated model. The association of the first Industrial Revolutions and the late female age at first marriage in North-West Europe is ultimately explained by the ending of the High Mortality Regime of the 14th and 15th centuries and the distinctive contribution of late marriage to human capital accumulation broadly interpreted. Without the Western European Marriage Pattern in England there would have been no sustainable growth in real wages over the period traditionally assigned to the Industrial Revolution.

Key Words: Economic Development, Demography, Unified Growth, Overlapping Generations, English Economy

JEL Classification: O11, J11, N13

[*] Cardiff Business School, Cardiff University, CF5 3EU. foreman-peckj@cardiff.ac.uk.
[†] Cardiff Business School, Cardiff University, CF5 3EU. zhoup1@cardiff.ac.uk.

Bringing Unified Growth Theory to the Data

Unified growth theory offers an explanation for the transition from Malthusian stagnation to sustained economic growth (Galor and Weil, 2000; Galor and Moav, 2002; Galor, 2010, 2011). It also claims to explain the demographic transition from high to low fertility. In total it provides an analytical framework for the divergence in national incomes per capita in the last two centuries. At the centre are two equations governing human capital formation and technological progress, where technology depends on population size (Galor, 2011).

Until now unified growth models typically bear only a descriptive relation to historical experience; tests of specific components have not been undertaken. The closest links of demographic-economic growth models to the available historical data have been made with calibration. Some models emphasise human capital accumulation driven by mortality changes (Boucekkine et al, 2003; Lagerlof, 2003; Cervellati and Sunde, 2005), others structural transformation (Desmet and Parente, 2012), physical capital deepening (Voigtlander and Voth, 2006), culture (Chen, 2012) or female empowerment (Diebolt and Perrin, 2013a, 2013b). Some papers consider several historical economies (Lagerlof, 2003; Boucekkine et al, 2003; Cervellati and Sunde, 2015). Others restrict themselves to one (Cervellati and Sunde, 2005; Voigtlander and Voth, 2006; Mourmouras and Rangazas, 2009; Desmet and Parente, 2012). More detailed statistical analysis tends to be substantially separated from the model (Voigtlander and Voth, 2013; Diebolt and Perrin, 2013a, 2013b).

The appropriate choice of particular models for specific places and times is therefore unclear. The present paper is the first to address this problem by estimating formally a unified growth model, utilising the longest series of available data and matching mortality differences between generations in the greatest detail. It therefore provides more convincing links between the framework of unified growth theory and history, in particular the process of human capital accumulation over almost six centuries in one economy that culminates in the first Industrial Revolution. The simulated method of moments allows for more rigorous model testing than does deterministic calibration, by maximising the model's ability to match the moments (variances as well as means) of the observable series. Also it permits distinguishing the consequences of different shocks to elicit more detailed outcomes of the model[1].

[1] Unlike Lagerlof (2003) we model multiple mortality rates and shocks.

In fitting the data moments the estimated model can explain the endogenous transition from apparent Malthusian stagnation to a sustained rise in real wages for England. In addition, it can match fragmentary historical evidence such as female age at first marriage, celibacy rates, mortality regime shifts and human capital (patent data). As a continent-specific extension to Galor and Weil (2000), agents in each phase or generation make decisions on the marriage age, as well as on the number of children, the quality of children and consumption. The engine of growth, human capital, is endogenously determined and results from the individual decision over the quantity and quality of children (along with marriage age) in the previous phase. Ours is a unified growth model in the sense that it can explain different growth regimes without a structural break in the deep parameters of technology and preferences. The only break occurs to match history with a mortality regime shift early in the Malthusian period.

Western Europe showed a uniquely high female age at first marriage from the 15th century and also experienced the earliest modern economic growth (Hajnal, 1965; de Moor and van Zanden, 2010; Foreman-Peck, 2011). The connection in our model of the English economy is that a higher female marriage age contributes to lower costs of child quality and thus to greater future human capital. It is the key mechanism of growth that we test in this paper.

The overlapping generations model treats human capital accumulation as a matter of direct parental preference rather than a matter of efficient investment. Human capital may be maintained and increased even when it is unprofitable. After the High Mortality Regime interest rates and skill premia did not return to their previous levels despite population growth and increasing land scarcity. The explanation is that the emergence of the European marriage pattern induced changes in desired child quality and greater savings (Van Zanden, 2009 p162).

Unified Growth (UG) modelling is here enriched by the exogenous role of mortality; the lower mortality regime triggers a Malthusian preventative check, the rise in the age at first marriage of females in Western Europe, reducing the target number of births. This is a steady state effect in the model. There is also a stimulus to greater human capital accumulation that ultimately more than offsets diminishing returns to population growth, thereby raising wage growth. In the shorter term target child numbers, child quality, and consumption are all increased. A fall in mortality does not guarantee a high age at marriage, because target numbers of children will increase due to the income effect (surviving children are less costly so demand for them expands). But in the mortality regime shift we simulate the first effect of mortality dominates to achieve the higher marriage age.

The intensity and frequency of mortality crises (shocks) diminish with the success of Western European quarantine regulations from the early 18th century (Chesnais, 1992 p141). Demographic consequences, more surviving children, are realised more rapidly than the effects operating through human capital accumulation. This is because greater child quality must take longer to be translated into higher wages than the stronger demand for children's numbers takes to trigger an increase in population.

Our link between demography and economy allows the model to show that in the long term increasing productivity from human capital accumulation raises the demand for children, boosting population. Eventually this technical progress associates rising population and real wage growth. But, contrary to traditional UG theory, the greater population is not necessary for the sustained wage growth that we identify as the key consequence of the Industrial Revolution. Although our data stop in 1870 we simulate the model to show that the rise in real wages continues after this date.

Counterfactual simulations demonstrate how our model might explain some of the varieties of long term demographic-economic growth experience. We show how the absence of the European marriage pattern would have precluded the rise in English wages in the Industrial Revolution. We gain further insights into Asia-Europe divergences by postulating stronger preferences for numbers of children and a continuation of the High Mortality Regime (Jones 1981). Under these conditions we demonstrate that there is no rise in the marriage age and no increase in real wages.

The Dutch economy of the first half of the 17th century was probably the most productive in the world. But subsequently it did not experience a similar timed upsurge to England's because of strong negative shocks so that real wages actually fell (Van Zanden and van Leuween, 2012). We simulate the English model with large negative productivity shocks from the 1650s to the 1820s and show how this could have caused real wages to decline if England had comparable experiences.

We demonstrate the dependence of English economic growth on other features of the economy and why perhaps other economies with the European marriage pattern did not experience the sustained rise in real wages of England at the same time by postulating a 10 percent lower human capital coefficient in the production function. If the guilds had been more able to restrict activity in England as has been alleged for Germany (Denison and Ogilvie, 2014) this type of reparameterisation would have been appropriate. The lower human capital productivity is sufficient to defer the sustained rise in real wages in England until after 1870.

Section 1 sets out the model, section 2 discusses the data and their patterns and section 3 presents the results. Section 4 simulates the model with and without the marriage age/human capital accumulation process to test the contribution of the 'Western European Marriage Pattern' in the English context. Other section 4 simulations show the outcomes of changing selected shocks and model parameters. The robustness of the results is evaluated in section 5, in particular with a different price deflator for money wages (Allen, 2007) and by introducing patents as a measure of human capital (Madsden et al, 2010).

1 The Model

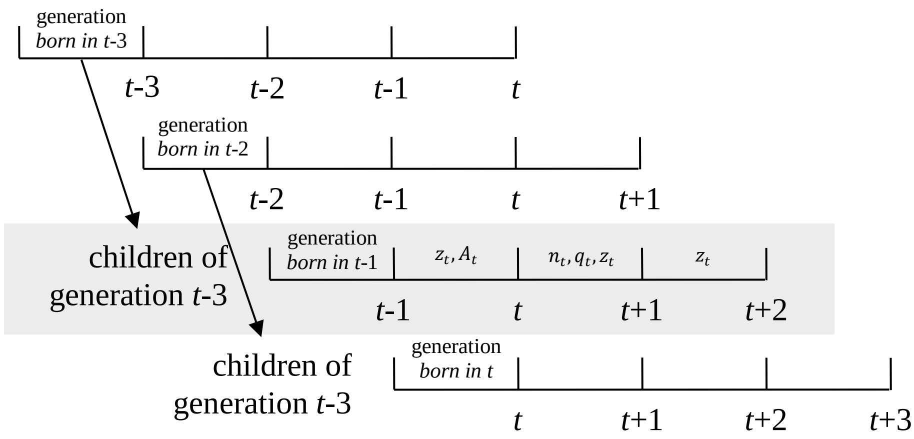
A theoretically meaningful and empirically measurable model of the interaction between population and the economy must allow for fertility choice and differential mortality chances of life stages. The traditional two period life cycle[1] would imply at least a 30-year 'generation' duration, which would require transforming the annual data to 30 year averages, resulting in a considerable loss of information. On the other hand, a more refined generation structure such as a period length of 5 years or even one year would result in colossal computation burden. The four-period life cycle is therefore the most appropriate compromise[2].

The present model phases or 'generations' are: '**childhood**' (0-15 years) when the costs of maintenance and training are borne by the parents, '**young adulthood**' (16-30 years) when the individual joins the workforce and decides on marriage and children, '**parenthood**' (31-45 years) when the majority of the costs of bringing up the children are paid by the working parents, and finally '**seniority**' (46-60) when the children have left home and begun working. Mortality in the first phase is higher than in the next two. In the last phase the assumed mortality rate is 100%.

Taking the generation born in period $t-1$ as an example, all the decisions are made in their young adulthood, including the female age at marriage (A_t), the target number of children (n_t), the target quality of children relative to the parents (q_t), and the lifetime consumption flows (z_t). The time subscript denotes when the variable is determined, so the variables with subscript t may be realised in different periods. For example, A_t occurs in period t; n_t and q_t take effect in period $t+1$; while z_t is relevant throughout periods t to $t+3$. The overlapping structure is illustrated in Figure 1.

[1] Following Galor and Weil (2000).
[2] A five-generation model was tried, but it does not significantly differ from the four-generation model.

Figure 1 Generation Structure of the Model



1.1 Household

The representative agent of the generation born in period $t-1$ maximises her CES utility in period t (during her young adulthood):

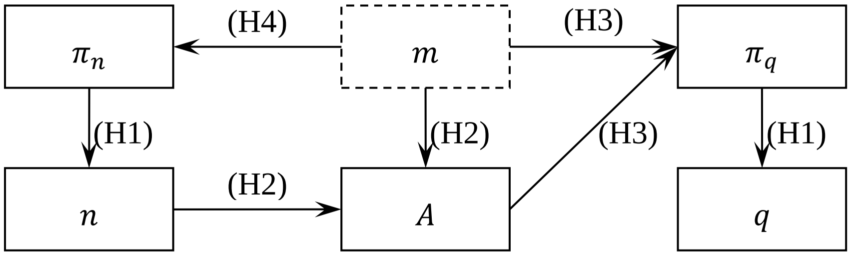
$$U(n_t, q_t, z_t) \equiv \left[\alpha \cdot \left(\frac{n_t}{2} \right)^{\frac{s-1}{s}} + \beta \cdot (q_t)^{\frac{s-1}{s}} + \gamma \cdot \left(\frac{z_t}{z_{t-1}} \right)^{\frac{s-1}{s}} \right]^{\frac{s}{s-1}}$$

The specification has the realistic merit of not imposing a unit elasticity of substitution between child quality and quantity; Clark and Cummins (2016) find a minimal trade-off between them in England 1770-1880. The utility function $U(\cdot)$ has three features in a dynamic setting: (i) The target number of children n is divided by 2, because the utility of n children is shared by the two parents. (ii) The target quality of children q is defined as the ratio of children's to parents' human capital levels. (iii) Consumption flow z_t enters the utility as a relative ratio rather than an absolute level. This 'habit persistence' in material consumption has a justification from empirical psychology (Scitovsky, 1992); changes in consumption, not the level, affect utility[1]. There are four constraints[2] of this maximisation problem:

(H1) Budget Constraint: $\sum_{i=0}^{2} w_{t+i} = \frac{1}{2}\pi_{n,t+1} n_t + \frac{1}{2}\pi_{q,t+1} q_t n_t + \sum_{i=0}^{2} z_{t+i}$;
(H2) Marriage Age Equation: $A_t = A(n_t, m1_{t+1}, m2_{t+2}; a_0, a_1)$;
(H3) Price of Children Quality Equation: $\pi_{q,t} = \pi_q\left(w_t, A_{t-1}, m1_t, m2_{t+1}; b_0, b_1, \eta\right)$;
(H4) Price of Children Quantity Equation: $\pi_{n,t} = \pi_n\left(w_t, m1_t, m2_{t+1}; c_0\right)$.

[1] Because the other two utility inputs (n and q) are stationary, the third utility input must be also. Becker, Murphy and Tamura (1990) instead invoke parental altruism.
[2] All these four equations are linear in parameters.

The following chart summarises the topology of the household's constraints.



Constraint (H1) specifies that lifetime earnings are equal to the lifetime expenditures (by 'earnings' we also mean earnings in kind from domestic production consumed in the household). Expenditure on children is shared by both husband and wife, so it is halved for the representative agent[1].

Constraint (H2) imposes a restriction on female's marriage age A_t, which negatively depends on the total planned births, taking premature deaths in childhood and young adulthood ($m1$, $m2$) into consideration[2]. a_0 is the age women usually stopped planning for, or expecting, new children (assumed around 35, Flinn, 1981 p33), and a_1 is the average gap between births (assumed between 2 to 3 years).

Constraint (H3) states that the *relative* price of educating each surviving child depends positively on mortality rates, and negatively on mother's marriage age (which is an indicator of the level of the mother's human capital). The parameter b_0 can be interpreted as the average proportion of income spent on q, while b_1 is the lower bound of marriage age. $\eta \equiv \left|\frac{d \ln \pi_q}{d \ln A}\right|$ measures the elasticity of human capital accumulation with respect to marriage age. This type of transmission has been widely discussed (DeTray, 1973; Marshall, 1961 p469). Historical evidence of the process is found in South Carolina before the US civil war (Murray, 2004) and in Victorian Britain (Mitch, 1992 ch. 4). More contemporary period research shows maternal education and maternal age have positive impacts both on cognitive skills and behavioural problems (Carneiro et al, 2007; Sutcliffe et al, 2012). Mothers' age at marriage and education affect child quality in India and offspring adult health in the US (Gaiha and Kulkarni, 2005; Myrskylä and Fenelon, 2011). Earlier in England, late age at marriage and the associated pattern of service was likely to be a widespread and critical learning experience (Kussmaul 1981). There is present day evidence for the relationship between female

---

[1] Agents make their choices in the expectation (or hope) of living for the full four periods of the model, even though they may be wrong. Remarriage was a common means of maintaining the household in the event of the premature death of a spouse.

[2] Family targeting plans are assumed to be based on survival of children to (approximately) age 30, perhaps to support parents in old age.

later marriage age and greater education (Georgiadis and Manning, 2011; Field and Ambus, 2008).

Constraint (H4) indicates that the *relative* price of feeding and clothing each surviving child increases with mortality rates, where c_0 is the average proportion of income spent on n. We explicitly include the monetary costs of childcare in our model, while the other studies either only include time costs of childcare (Galor and Weil, 2000; Lagerlof, 2003; Cervellati and Sunde, 2015) or ignore the costs as a whole (Voigtlander and Voth, 2006; Strulik and Weisdorf, 2008).

1.2 Demography

Based on the quantities and prices involved in the individual maximisation problem, the aggregate demographic variables can also be defined in terms of existing variables. The law of motion for population (P_t: population stock at time t) is:

(D1) $P_t = P_{t-1} - D_t + B_t$.

Total deaths (D_t: death flow in period t) are the sum of premature and natural deaths:

(D2) $D_t = m1_t \cdot B_t + m2_t \cdot G1_{t-1} + m3_t \cdot G2_{t-1} + 1 \cdot G3_{t-1}$.

Here, Gi_{t-1} denotes the population of the generation born in period $t-i$ still living at time $t-1$ (at the beginning of period t):

(D3) $G1_t = (1 - m1_t) B_t$;

(D4) $G2_t = (1 - m1_{t-1})(1 - m2_t) B_{t-1}$;

(D5) $G3_t = (1 - m1_{t-2})(1 - m2_{t-1})(1 - m3_t) B_{t-2}$.

Total births in period t (B_t: birth flow in period t) can be defined as:

(D6) $B_t = \dfrac{(1-\mu_t) G2_{t-1}}{2} \dfrac{n_{t-1}}{(1 - m1_t)(1 - m2_{t+1})}$.

The variable μ_t captures the proportion of females without any children, or the celibacy rate. It includes both never-married females and married females with no offspring[1]. The unmatched proportion is a function of search and matching costs (Keeley, 1977; Choo and Siow, 2006) which are positively correlated with marriage age ($\tau_1 > 0$) and

[1] The extramarital births are such a small proportion of the total that they can be ignored for present purpose.

negatively with wage ($\tau_2 < 0$). The term $\bar{\mu}$ is the average celibacy rate, while A_t^* and w_t^* are the expected level of marriage age and wage in absence of all shocks.

$$\text{(D7)} \ \mu_t = \bar{\mu} + \tau_1 \frac{A_t - A_t^*}{A_t^*} + \tau_2 \frac{w_t - w_t^*}{w_t^*}.$$

Based on B_t and D_t, we can define birth rate (b_t) and death rate (d_t) to link with the observable historical data. In particular, the birth rate is the birth flow during period t divided by the population stock at the beginning of period t (or at time $t-1$). The death rate is the death flow during period t divided by the sum of the population at $t-1$ and the birth flow during period t, because the death rate takes into account the premature deaths of those in their childhood period.

$$\text{(D8)} \ b_t \equiv \frac{B_t}{P_{t-1}};$$

$$\text{(D9)} \ d_t \equiv \frac{D_t}{P_{t-1} + B_t}.$$

1.3 Production

Where Y is output, H is human capital per capita, $\bar{F}$ is a fixed production factor, especially land, and x_t a random productivity shock, the representative production unit's (farm's) problem is:

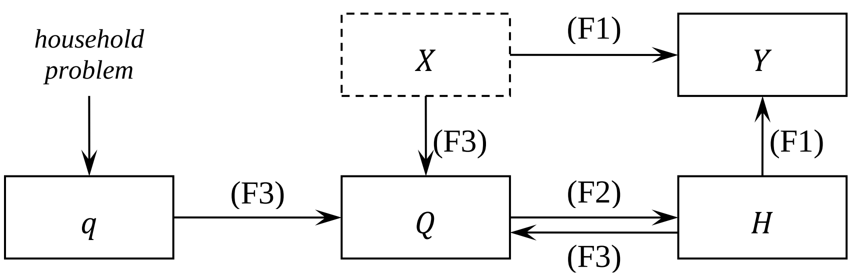
$$\max \Pi_t = Y_t - w_t P_{t-1} - \text{fixed costs}, \text{ subject to:}$$

(F1) Production Function: $Y_t = e^{x_t} P_{t-1}^{\theta_1} H_{t-1}^{\theta_2} \bar{F}^{1-\theta_1-\theta_2}$;

(F2) Aggregate Human Capital: $H_t = \frac{G1_t}{P_t} Q_t + \frac{G2_t}{P_t} Q_{t-1} + \frac{G3_t}{P_t} Q_{t-2}$;

(F3) Generational Human Capital: $Q_t = e^{x_t} H_{t-1}^{\varepsilon} \left(Q_{t-2} q_{t-1}\right)^{1-\varepsilon}$.

Note that the labour force of period t is the population stock (P_{t-1}) excluding the generations in their childhood during period t. The following chart summarises the topology of the farm's constraints.



Constraint (F1) describes the production technology with three separate inputs. H_{t-1} is the human capital per capita influencing period t, by which the 'raw' labour force (P_{t-1}) is augmented. The two variables are lagged because the contemporaneous values would include children. Human capital, the accumulation of skills and of work discipline, augments the productivity of labour, but so too can the building of trust, of reciprocal obligations and social networks, sometimes dubbed 'social capital' (Putnam, 1993, 1995). Social capital augments the productivity of the workforce or population regardless of their level of skills. The relative magnitudes of the coefficients on labour and human capital therefore show the balance of impact of these two types of capital. A higher θ_1 indicates that the social capital is derived more from the quantity of labour, while a higher θ_2 indicates a more important role for the quality of labour. $\bar{F}$ represents the factors that cannot be reproduced such as land. For simplicity, it can be set equal to 1 without loss of generality and it is equivalent to normalising output by land.

Constraint (F2) defines the average human capital (H_t), which is a weighted average of the generational human capital in the labour force (the three working 'generations').

Constraint (F3) describes how each generation's human capital level (Q_t) is formed. Apart from the parents' influence ($Q_{t-2}q_{t-1}$: the target quality of children formed by 'family education'), the generation born in period t is also affected by the average human capital of existing generations H_{t-1} (the parameter ε captures the contribution share of 'nonfamily education' including formal schooling and apprentice training)[1]. The productivity shock x_t also affects the formation of Q_t, on the grounds that a society with higher productivity tends to be more effective in the transmission of knowledge.

Finally, to complete the system of equilibrium conditions, a competitive labour market is postulated[2], so that the marginal product of labour equals the marginal cost:

$$\text{(F4)} \quad MPL_t \equiv \theta_1 e^{x_t} P_{t-1}^{\theta_1 - 1} H_{t-1}^{\theta_2} \bar{F}^{1-\theta_1-\theta_2} = w_t \equiv MC_t \, .$$

1.4 Shock Structure

Lee (1993) maintains that exogenous shocks were principally responsible for the approximately 250-year European demographic cycle. Mortality shocks drive Lagerlof's (2003) Industrial Revolution, while weather-induced shocks to agricultural productivity cause changes in prices and quantities, and affect wages in Voigtlander and Voth's (2006) model. We include both types of shock, distinguishing one type of productivity

---

[1] The effect of schooling/training depends crucially on the overall human capital of the time (embodied in the teacher).

[2] Clark (2007) provides evidence that the wage divided by product prices does indicate the marginal product of labour even in 1300.

shock and three mortality shocks. Runs of poor harvests (such as the Great European Famine of 1315-17) and livestock disease constitute a negative productivity shock. Epidemic diseases such as bubonic plague, typhus and smallpox were mortality shocks. We assume the exogenous productivity process is AR(1), where ex_t is white noise with standard deviation equal to σ_x.

(S1) $x_t = \rho_x \cdot x_{t-1} + ex_t$.

Similarly, there are three exogenous conditional premature mortality rates specific to each of the three generations:

(S2) $m1_t = \overline{m1} \cdot \exp(mm1_t)$, where $mm1_t = \rho_1 \cdot mm1_{t-1} + emm1_t$;

(S3) $m2_t = \overline{m2} \cdot \exp(mm2_t)$, where $mm2_t = \rho_2 \cdot mm2_{t-1} + emm2_t$;

(S4) $m3_t = \overline{m3} \cdot \exp(mm3_t)$, where $mm3_t = \rho_3 \cdot mm3_{t-1} + emm3_t$.

Here, $\overline{mi}$'s ($i = 1,2,3$) are the steady state mortality rates and $emmi_t$'s are Gaussian white noises with standard deviations equal to σ_i. We set the steady states of mortality rates to match the historical life expectancy pattern over the six centuries as shown in Table 1[1]. A life expectancy at birth of 23 (as Hatcher 1986 calculates for the Canterbury monastery) is assumed for the High Mortality Regime (level 3, Wrigley and Schofield, 1989 p714), and a life expectancy of 38 (level 9) for the Low Mortality Regime.

Table 1 Calibration of Steady States of Mortality Rates

	High Mortality Regime	**Low Mortality Regime**
$\overline{m1}$	50.00%	30.00%
$\overline{m2}$	30.27%	12.08%
$\overline{m3}$	30.27%	12.40%
$\overline{m4}$	100%	100%
Implied Life Expectancy	**23.54**	**38.23**

1.5 Stationarisation and Steady States

The above system is non-stationary because of growth in human capital and population. But standard numerical methods for solving this dynamic equation system require that we stationarise many of the original conditions. $U_t, n_t, q_t, A_t, \mu_t$ are stationary by definition, so for them no change is necessary. The non-stationary endogenous variables

[1] The calibrated rates should not be interpreted as a strict mapping from the life tables in Wrigley and Schofield (1989, p714) or any other sources. Since the model only has four generations with a maximum life of 60 years, the mortality rates in each cohort are inevitably different from the life table which has a maximum life of 90 years. The details of how we calibrate the mortality rates are described in Appendix I.

can be categorised into four groups in terms of their balanced growth path rates, or of their deflators. Where a hat ‘^’ indicates a stationarised variable:

Deflated by P_{t-1}: $\hat{B}_t \equiv \frac{B_t}{P_{t-1}} \equiv b_t$, $\hat{D}_t \equiv \frac{D_t}{P_{t-1}}$, $\hat{Gi}_{t-1} \equiv \frac{Gi_{t-1}}{P_{t-1}}$, $\hat{P}_t \equiv \frac{P_t}{P_{t-1}} = 1 + g_{Pt}$;

Deflated by H_{t-1}: $\hat{H}_t \equiv \frac{H_t}{H_{t-1}} = 1 + g_{Ht}$, $\hat{Q}_t \equiv \frac{Q_t}{H_{t-1}}$;

Deflated by $P_{t-1}^{\theta_1} H_{t-1}^{\theta_2}$: $\hat{Y}_t \equiv \frac{Y_t}{P_{t-1}^{\theta_1} H_{t-1}^{\theta_2}}$;

Deflated by $P_{t-1}^{\theta_1 - 1} H_{t-1}^{\theta_2}$: $\hat{w}_t \equiv \frac{w_t}{P_{t-1}^{\theta_1 - 1} H_{t-1}^{\theta_2}}$, $\hat{\pi}_{n,t}$, $\hat{\pi}_{q,t}$, $\hat{z}_t$.

The model is solved by perturbation method around a steady state (as described in Appendix III). It involves log-linearisation of the original nonlinear equations around the steady state. The existence of the steady state solution of the model is guaranteed because there is a solution—if there is no solution under a certain combination of parameter values, the estimation procedure automatically filters out that combination.

The uniqueness of the steady state is less of a problem for the perturbation method, because we only focus on the steady state in the neighbourhood of the actual history. If a global solution method (e.g. a projection method) is used, then there might be multiple steady states far away from the actual history, resulting in possibly implausible equilibria. This also marks a difference between our model and the Galor and Weil (2000) model. The latter has two equilibria (two solutions) from a single parameterisation, with one being a Malthusian regime and the other a modern growth regime. In contrast, our model has two sets of shock structure parameters for the two mortality regimes, resulting in two steady states. But each parameterised regime only has one unique steady state in the neighbourhood of the data.

2 Data

A central characteristic of the English economy in long term perspective is the population collapse in the 14th century. We adopt the Broadberry et al (2015) estimates of population. These show a fall from a peak of 4.81 million to a trough of 1.9 million in 1450. Our other source of demographic data is Wrigley and Schofield’s (1989) mortality and fertility crude rates derived from 404 Anglican parish registers of baptism, marriage and burial. English parish registration generally began in 1541. These data are also utilised for Broadberry et al’s later population series.

15th century high mortality rates are apparent from longitudinal studies of late medieval monasteries (Hatcher, 1986; Bailey, 1996; Hatcher et al 2009). We distinguish a High Mortality Regime from mortality shocks by the regime's sustained decline in population over around a century. The Great European famine of 1315-17 and the animal epidemics indicate a beginning for the High Mortality Regime. Wage data that peaks around the middle of the 15th century implies that population decline in the face of the high mortality must have ceased a little earlier; in the absence of major changes in birth rates, the High Mortality Regime then ended around 1420. Cummins' (2014) finding of a structural break in European noble life spans around 1400, when longevity increased from about 50 to 55, is broadly consistent with this characterisation[1].

We assume that real wages are the ultimate performance measure for an economic system; if we can explain the onset of the sustained increase in England then we have captured a critical stage of economic development. Clark's (2013) series begins in 1200. By 1450 real wages were higher than in 1800, an outcome conventionally attributed to labour scarcity. The series is constructed from separate indices of male farm day wages, male coal mining day wages, and male building wages (to represent both the secondary and tertiary sectors). Then these are aggregated into a national male wage using estimates of the share of males employed in each sector. Women's earnings are assumed to be 25 percent of men's (as they were estimated to be in 1866). As well as adopting Clark's series we also use alternative real wage data from Allen (2007) for the Industrial Revolution period to provide a robustness check. A detailed description of all the data used in this paper can be found in Appendix I.

3 Results

The model is solved, simulated and estimated using the methods described in Appendix III. One practical issue is that the unit period of the model is 15 years while the sample period is annual. If we transform the annual data to 15-year frequency, we lose information and there is no clear reason as to which year should begin the first period. To solve both problems, we note that there are only 15 distinctive ways of defining the unit period, i.e. the first period can start from 1301, 1302, 1303, etc. until 1315. We 'slice' the sample in these 15 ways and transform each slice into a 15-year frequency. Then, we obtain the moments based on each 'slice', and the averages of the moments can be used in the objective functions. In fact, it turns out that the moments are quite robust to how the sample is sliced.

[1] De la Croix and Licandro's (2012) 'famous people' data are insufficiently accurate over this period to use as mortality change evidence.

3.1 Parameter Estimation

Most parameters in Table 2 are estimated to minimise the squared distance between the simulated and the observed mean values and standard deviations of the data series (population, births, deaths and wages)[1] and fragmentary data (marriage age and celibacy rates) in both mortality regimes.

Table 2 Calibrated and Estimated Parameters of the Model

	Meaning	Estimates		C/E
a_0	Average stopping birth age	35		C
a_1	Gaps between births	2.5		C
b_0	average % of income spent on q	0.139		E
b_1	Lower bound of marriage age	16		C
c_0	average % of income spent on n	0.138		E
α	Utility weight for children quantity	0.253		E
β	Utility weight for children quality	0.393		E
s	Elasticity of substitution	0.154		E
$\bar{\mu}$	Average celibacy rate	0.155	0.138	E
τ_1	Elasticity of A on μ_t	0.177		E
τ_2	Elasticity of w on μ_t	-0.833		E
θ_1	Income share of raw labour	0.335		E
θ_2	Income share of human capital	0.425		E
ε	Contribution (share) of 'non-family' education	0.411		E
η	– Elasticity of A effect on π_q	1.185		E
ρ_1	AR coefficient of shock mm1	0.539	0.006	E
ρ_2	AR coefficient of shock mm2	-0.494	0.190	E
ρ_3	AR coefficient of shock mm3	0.869	0.074	E
ρ_x	AR coefficient of shock x	0.793	0.206	E
σ_1	Standard deviation of shock emm1	0.404	0.277	E
σ_2	Standard deviation of shock emm2	0.074	0.412	E
σ_3	Standard deviation of shock emm3	0.030	0.629	E
σ_x	Standard deviation of shock ex	0.019	0.050	E

Three parameters are calibrated according to such historical information as is available, while others are estimated within the theoretical upper and lower bounds ('C' for Calibrated and 'E' for Estimated) using the Simulated Method of Moments described in the Appendix III.

A structural break is assumed only for the shock structure around 1420 dividing the whole sample period into the High Mortality Regime (1300-1420, 8 periods) and Low Mortality Regime (1420-1870, 30 periods), but the deep parameters related to preferences and technologies are kept constant across regimes. The persistence (measured by

[1] Patent data is also used as a measurement of human capital in the robustness test.

the magnitude of autoregressive coefficients) of the mortality shocks in the Low Mortality Regime are all higher than those of High Mortality Regime, and the opposite signs imply very different dynamics. The volatility (measured by the standard deviations) of mortality shock affecting the childhood phase in the High Mortality Regime is much higher than that in the Low Mortality Regime.

3.2 Matching the Moments

Using the estimated/calibrated values of the parameters above, the model can simulate the first moments (sample means or theoretical steady states) and the second moments (standard deviations) of the endogenous variables[1]. The simulated moments are contrasted with the observed sample moments in Table 3.

Table 3 Simulated Moments versus Data Moments

		Simulated Steady State	**Data Moments**	**Simulated SD**	**Data Moments**
High Mortality Regime	A	**18.02**	**18~20**	2.68	
	g_P	0.000	**-0.074**	0.133	**0.135**
	b	1.258		0.664	
	d	0.557		0.128	
	g_W	0.075	**0.095**	0.106	**0.105**
	n	2.368		0.274	
	q	1.376		0.139	
	$\hat{z}$	0.236		0.022	
	π_n	0.132		0.064	
	π_q	0.116		0.068	
Low Mortality Regime	A	**24.14**	**24~26**	0.24	
	g_P	0.073	**0.097**	0.076	**0.081**
	b	0.617	**0.627**	0.104	**0.104**
	d	0.336	**0.313**	0.039	**0.039**
	g_W	0.007	**0.007**	0.082	**0.083**
	n	2.673		0.058	
	q	1.518		0.031	
	$\hat{z}$	0.272		0.007	
	π_n	0.075		0.010	
	π_q	0.046		0.006	
	μ	0.138	**0.129**	0.043	**0.044**

[1] The simulated periods (600 years, 40 periods) are longer than the sample periods of birth/death data (330 years, 22 periods).

The model can successfully match the second moments of observed data, but there are slight discrepancies between the simulated steady states and sample means. This is because the simulated steady states are not equivalent to sample means in the first place. The sample means measure the average levels of the variables in real history, which is never in a steady state due to continual shocks. In contrast, the steady states are defined as the long-run levels when the effects of shocks die away.

Comparison of steady state regimes (Table 3) shows that two types of effect of the drop in mortality rates can be distinguished. The first is the 'wealth effect'; greater human capital accumulation more than offsets diminishing returns, raises wages growth (g_w), target child number (n), child quality (q), and consumption (z), and lowers marriage age (A). There is also a direct steady-state effect which requires fewer births (b) to achieve a given target family size and therefore raises the marriage age. This second effect strongly dominates the wealth effect in marriage age determination. It is the 'preventative check' to population (higher marriage age, lower births) of Malthusian theory.

Both target family size (n) and child quality (q) are permanently higher after the end of the High Mortality Regime. Meanwhile, the price of child quality (π_q) falls further than does the price of child quantity (π_n). One reason is the rise in the age of female first from 18 to 24; it is this that lowers the price of child quality because of what extra older mothers bring to child rearing.

The steady state growth rate of real wages (g_w) apparently drops when mortality rates fall but this is a consequence of an assumption necessary to estimate a steady state. The High Mortality Regime can only reach a steady state when population growth is set to zero, whereas historically it was negative; unsustainable in a steady state. The growth in wages of the Low Mortality Regime is accompanied by a sustainable population growth, which is not possible in the Higher Mortality Regime.

3.3 Impulse Response Functions

This stochastic model can show the responses of endogenous variables, such as growth rates of population, human capital and wages, to one standard deviation of the mortality shocks (Figure 2). The impulse response functions (IRFs) exhibit firstly, an *oscillating* feature (negative eigenvalues) which may rise from the heterogeneous generation structure of the model. Oscillation is also apparent in the data (Appendix I, Figure 14). The model successfully matches this data feature. Secondly, the effects of mortality shocks are highly *persistent* with a half-life of 3 periods (45 years) for population, because one high mortality shock is likely to be followed by another.

If we regard the High Mortality Regime as a big mortality shock at the beginning of the 14th century, then the IRF in Figure 2 can be interpreted as changes of an overall trend[1]. According to the condition that wage is equal to the marginal product of labour (F4), the dynamics of wages over the six century in Figure 2 are a result of the two opposite effects between population (negative pull) and human capital (positive push). The relative strengths of the two forces divide the six hundred years into three stages:

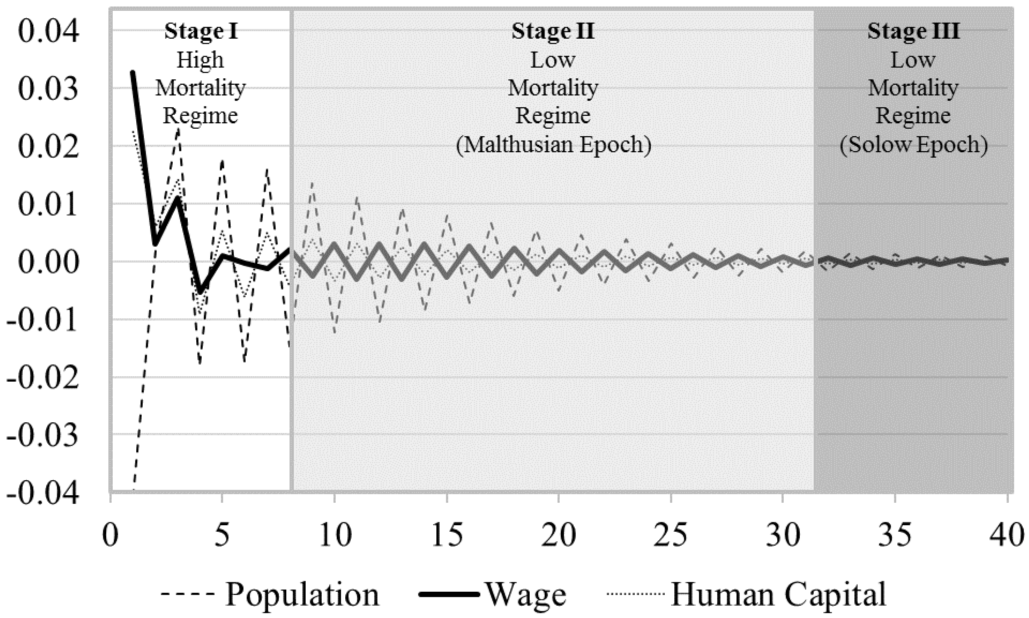
- **Stage I** (High Mortality Regime). During the High Mortality Regime population drops and wage rises sharply due to scarcity of labour. Subsequently, human capital and population change in the same direction; they are consequences of the two complementary goods q and n (with a low elasticity of substitution) formed in the previous period. The fluctuations of wages are dominated by human capital rather than population in the High Mortality Regime because the contribution of population has fallen so much.
- **Stage II**. After about 8 periods (120 years), the economy enters a Malthusian epoch or an early Low Mortality Regime, where population and wage change in opposite direction because the population pull effect dominates—population takes off earlier than wage because human capital needs longer to accumulate before it is fully effective.
- **Stage III**. In the long run, after the effects of the big shock finally vanish, both wage and population start growing at a sustainable equilibrium rate as in the Industrial Revolution (the late Low Mortality Regime).

The IRF of wage fits the timing of the Industrial Revolution—the 32nd period is the first period when deviations from the steady state are less than 1%. That is to say, the effect of a shock at the beginning of the High Mortality Regime will vanish around 1780 ($= 1300 + 32 \times 15$), the traditional dating of the onset of the first Industrial Revolution.

The transitional dynamics illustrated in the IRFs (Figure 2) show that the signs of the effects can alter in the short run, medium run and long run. A fall in mortality does not guarantee a high age at marriage for always, because target numbers of children will increase (surviving children are less costly so demand for them expands, similar to a substitution effect). The temporary balance between these forces determines the transitional ages at marriage before it converges to the steady state level.

[1] Because the other shocks are sufficiently small to ignore compared to this big shock.

Figure 2 Impulse Responses of Mortality Shocks[1]



4 Simulation

The previous section summarises the theoretical characteristics of the parameterised model. To show how well the model 'retrodicts' we simulate the six centuries' demographic and economic history based on the final form (or 'reduced form') of the model. We go on to demonstrate the importance of marriage age for human capital accumulation by a counterfactual simulation without this effect. Finally, we offer some simulations to suggest why other countries' experiences was different from England's.

4.1 Simulated History

In Figure 3-Figure 6, the grey broken lines are the simulated series (the conditional expected values based on the current information set, i.e. $\mathbf{E}_t[\cdot]$) and the black solid lines are the observed actual series[2]. Most of the fluctuations in the data can be well matched. One novelty of the simulation method we adopt here (described in Appendix III) is that we can make full use of the unbalanced data—the real wage series (Clarks, 2013) has more observations than the birth and death rate series (Wrigley and Schofield, 1981). Moreover, the simulated growth rates of population and wage (stationary) are translated into levels (non-stationary) for a more intuitive data comparison.

---

[1] The impulse responses in Figure 2 are the deviations from the steady states, so they return to zero in the long run.

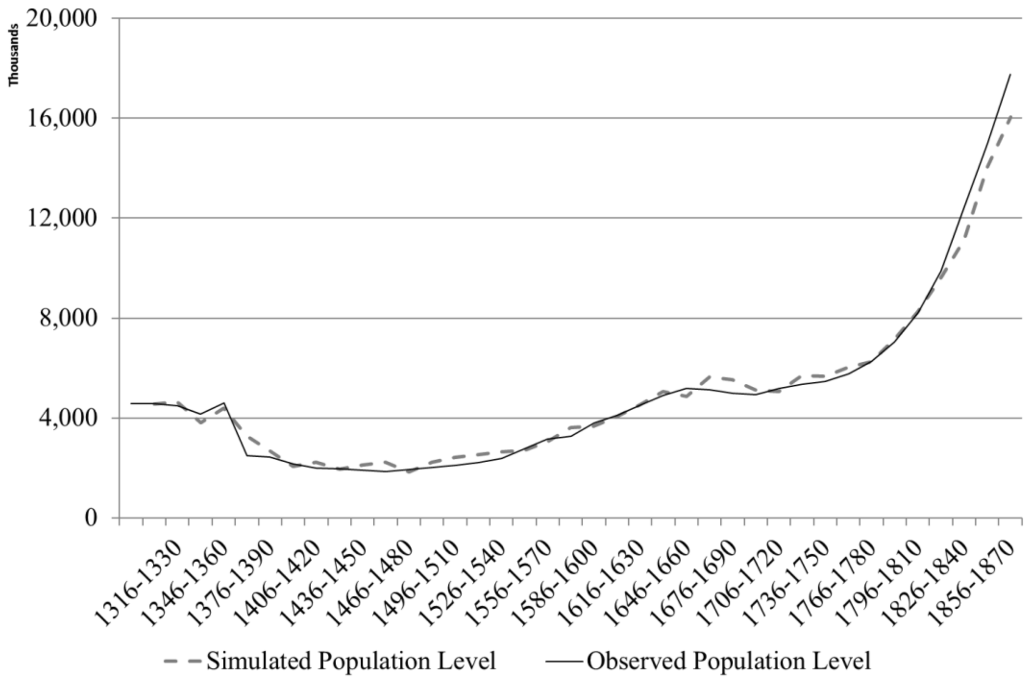
[2] This conditional simulation is consistent with our rational expectations model. Moreover, applying unconditional simulation to non-stationary data such as ours is likely to be highly erratic.

Both observed and simulated population growth turn upwards after the first quarter of the 18th century (Figure 3), whereas real wage growth begins to rise permanently much later (Figure 4). Rising population growth was facilitated by the success of Western European quarantine regulations from the early 18th century reducing the number and severity of mortality crises (Chesnais, 1992 p141) (death rate, Figure 6). Population increases earlier than the effects operating through human capital accumulation. Higher child quality takes longer to raise wages than the stronger demand for children's numbers takes to trigger an increase in population (birth rate, Figure 5).

Female marriage age jumps from around 18 to about 25 in the mid-15th century (Figure 7), by which time the High Mortality Regime is assumed over and the accumulation of human capital starts to speed up.

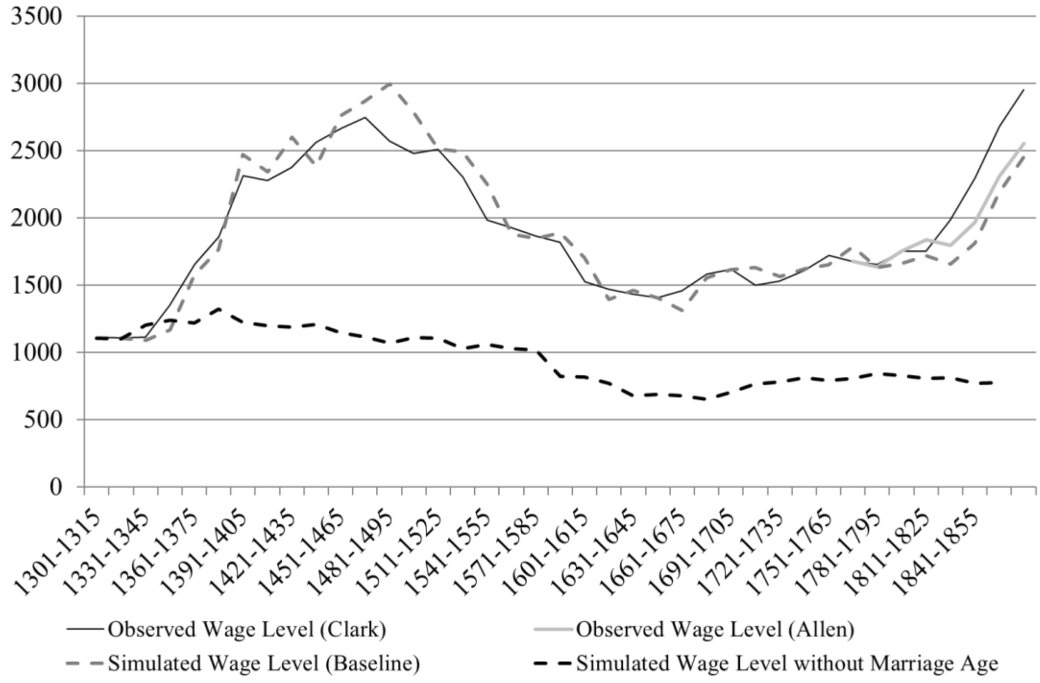
Figure 4 indicates that our model favours Allen's interpretation of the timing of the Industrial Revolution wage 'take-off' over Clark's. The simulated death rate of Figure 6 is higher than the observed for two reasons. On the one hand, the death rate is likely to be underestimated in the data due to unreported infant deaths and miscarriages. On the other, for the sake of modelling parsimony, we assume everyone has to die by 60 years old, which upwardly biases the death rate.

Figure 3 Simulated and Observed Population (15-Year)



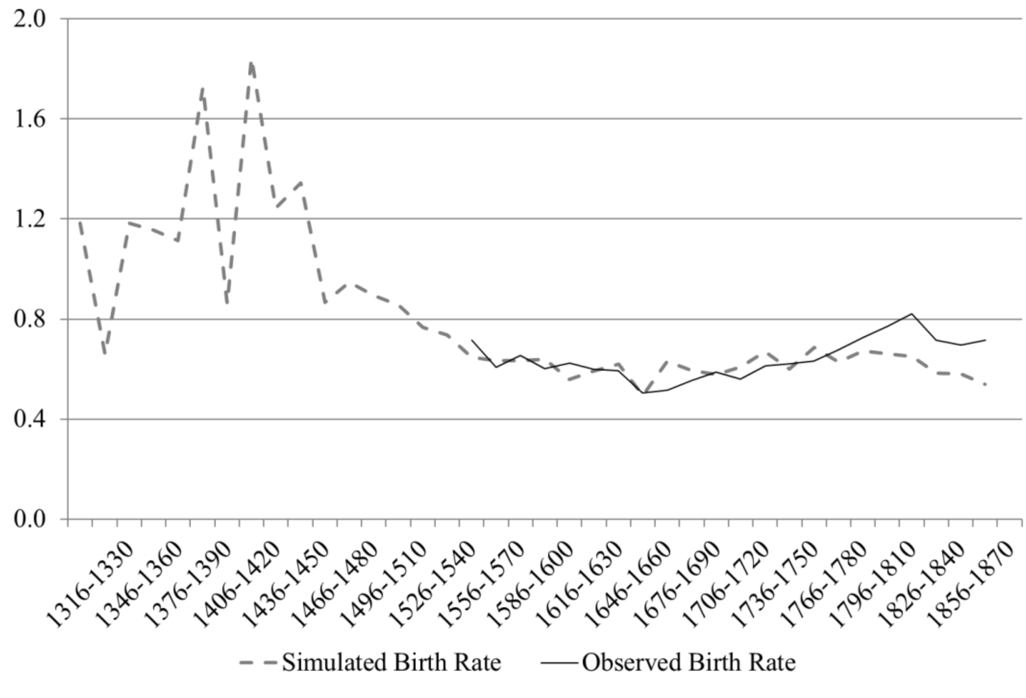
Notes: The population data combines the 1300-1540 series in Broadberry et al (2015) and the 1541-1870 series in Wrigley and Schofield (1981).

Figure 4 Simulated and Observed Real Wage (15-Year)



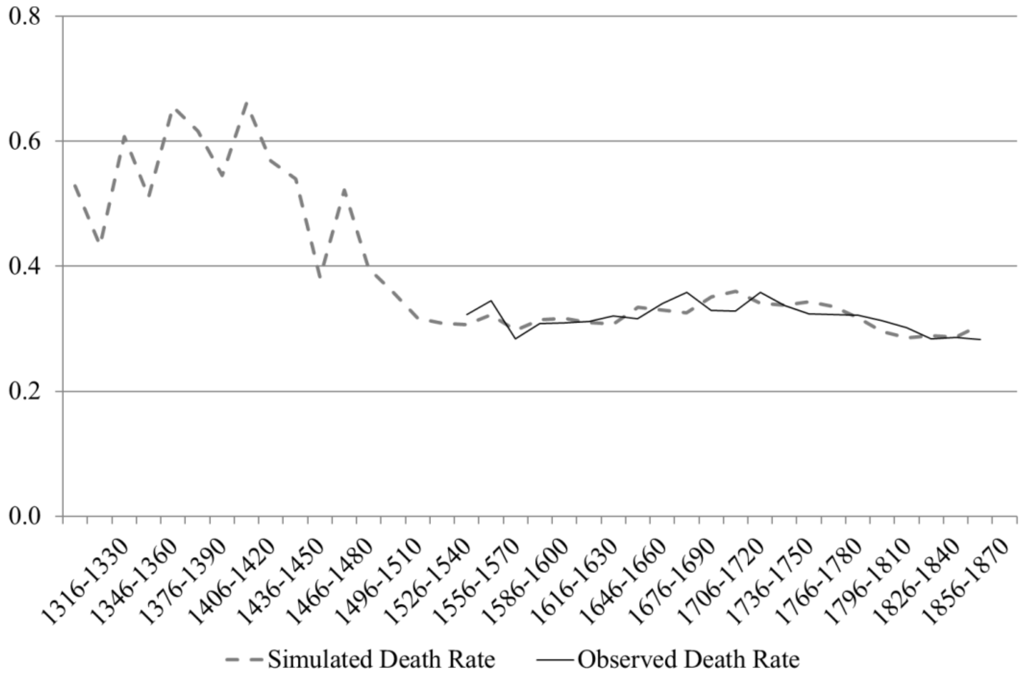
Notes: The wage data is based on Clarks (2013) for the baseline and Allen (2007) for robustness.

Figure 5 Simulated and Observed Birth Rate (15-Year)



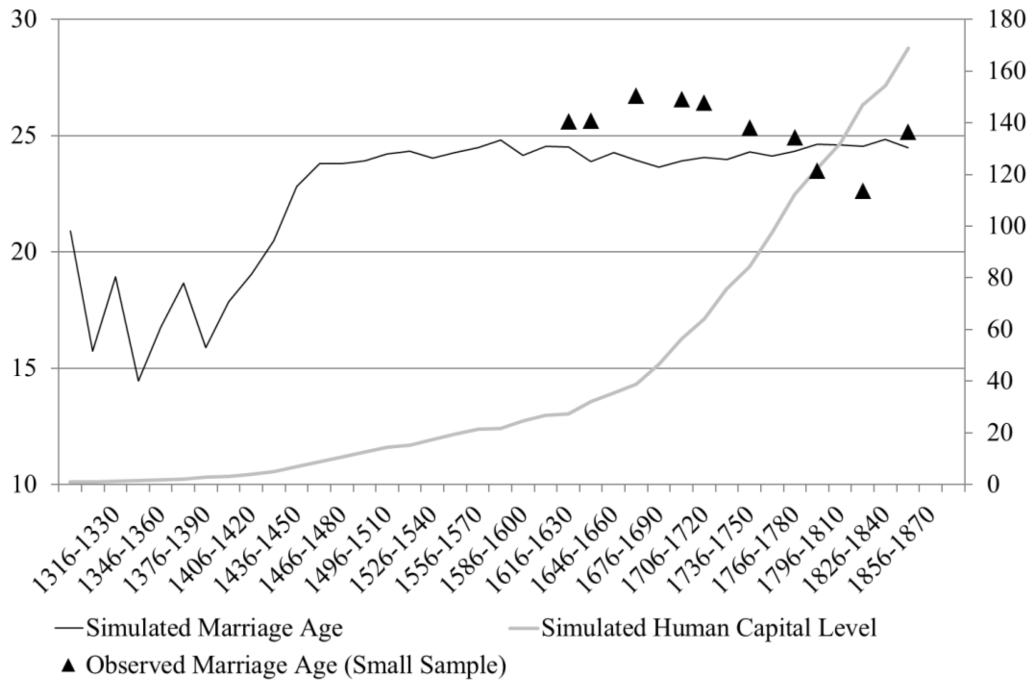
Notes: The birth data is based on Wrigley and Schofield (1981).

Figure 6 Simulated and Observed Death Rate (15-Year)



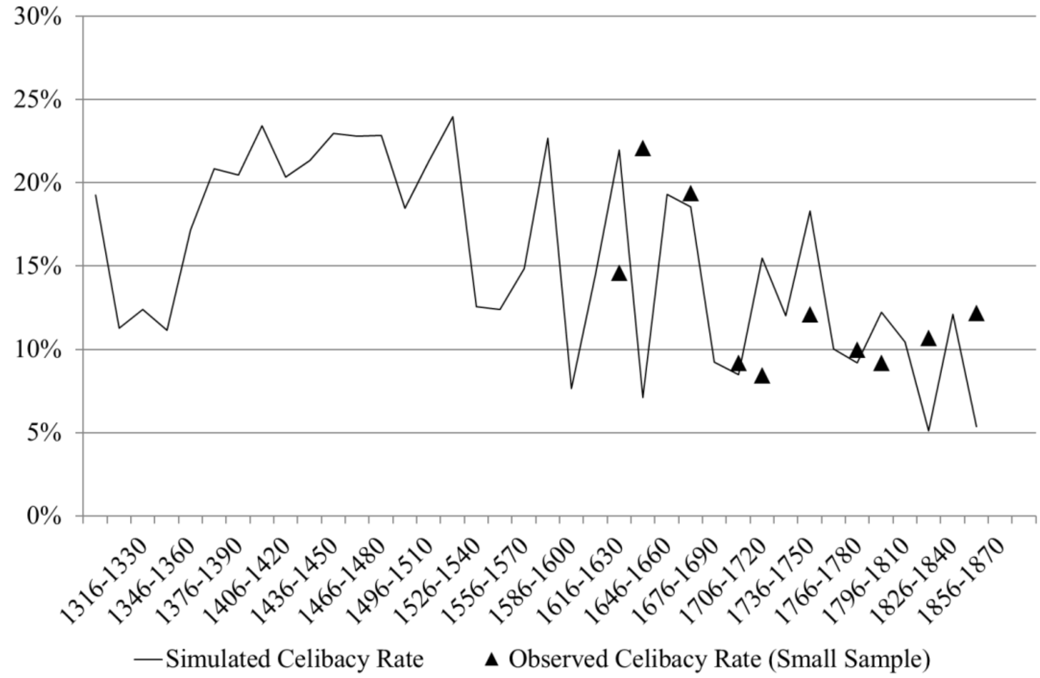
Notes: The death data is based on Wrigley and Schofield (1981).

Figure 7 Simulated Female Age at Marriage and Human Capital



Notes: The marriage age data is based on Wrigley et al (1997), whose data are collected from only 26 parishes that may not be representative of the entire country. In the robustness section, we also use patent data based on Madsen et al (2010) and Mitchell (1962).

Figure 8 Simulated Proportion of Celibacy Rate



Notes: The celibacy data is based on Schofield (1985).

Assuming constant marital fertility, historical demographers have examined the relationship between the gross reproduction rate of a small number of parishes on the one hand, and celibacy and female age at first marriage on the other (Weir, 1984; Schofield, 1985). They concluded that celibacy movements dominated in explaining gross reproduction rate before 1700 or 1741, and afterwards the age at marriage mattered more. In the late 17th century unregistered clandestine marriage adds complexity to the fragmentary statistics.

Our model-consistent series does not indicate a fall in the age at marriage in the late 18th century (Figure 7) as implied by the small sample parish analysis. Dennison and Ogilvie's (2014) finding of a rising marriage age in their broad European sample from the sixteenth to the nineteenth centuries might cast doubt on the representativeness of the English sample.

Although we assume that the proportion unmarried is stationary over six centuries it is possible to see shorter term trends in the fluctuations (Figure 8). The proportions fluctuate between 13 and 15 percent, while the marriage age is stable at between 24 and 25 after the late 15th century, having begun at 20 in 1300 and dropped during the high mortality years. This last is consistent with the little that is known (for example Hallam 1985). The timing of the human capital build-up (Figure 7) is consistent with numeracy and literacy statistics (Baten and van Zanden, 2008; A'Hearn et al, 2009; Clark, 2005). These show an accumulation process long before the Industrial Revolution.

An eyeball test suggests that all simulated series fit the data quite well[1]. To assess quantitatively how well the model fits the data we employ both the Theil (1966) measure and a tool similar to R-squared: if y_i is the actual level, $\hat{y}_i$ is the simulated level, then the goodness of fit for each observation is: $1 - |\hat{y}_i - y_i|/y_i$. The mean goodness of fit is the average of this measure across all observations. Therefore, there are four goodness-of-fit values corresponding with the four observables. This measure of goodness of fit is more flexible than R-squared in the sense that it can capture the performance of the model in matching different variables of an unbalanced dataset. As shown in Table 4, the birth rate prediction performs the least well but the simulated values based on the model (the conditional expectations) can still account for around 90% of the observed birth rates.

Table 4 Goodness of Fit of the Model

Observable Variables	**Theil (1966) Measure** $V = 1 - U$	**Alternative Measure** W
Population Level	92.42%	92.96%
Birth Rate	87.56%	90.49%
Death Rate	94.71%	95.56%
Wage Level	89.96%	93.38%

Notes: The Theil (1966) measure (V) is based on the inaccuracy rate $U = \sqrt{\sum_i (\hat{y}_i - y_i)^2} / \sqrt{\sum_i y_i^2}$, where $\hat{y}_i$ is the fitted value of the observed value y_i. The alternative measure (W) is defined as $W = 1 - \sum_i (|\hat{y}_i - y_i|/y_i)$, which is based on the absolute errors rather than the mean squared errors.

4.2 Marriage Age and Human Capital Accumulation

The importance of the marriage age for human capital in this model can be demonstrated by a simulation that excludes the effect of marriage age on the price of child quality. This is achieved by setting $\eta = 0$ in (H3) (where there is no inter-generational transmission of human capital through the family). The simulation reported below in Table 5 compares the steady states under the scenario with the baseline scenario.

If we assume the mechanism is missing during the Low Mortality Regime, then the growth rates of wages would have been negative. Population and output would also grow more slowly. The last column of Table 5 shows how much lower the four variables would have been than the actual levels if the marriage age mechanism was missing.

[1] But the discrepancies between simulated and observed levels of population are actually greater than it looks in Figure 3, because the level of population is some millions.

In particular, the wage would have been a quarter lower and GDP would have been halved, due to the reduction in both labour/population and human capital.

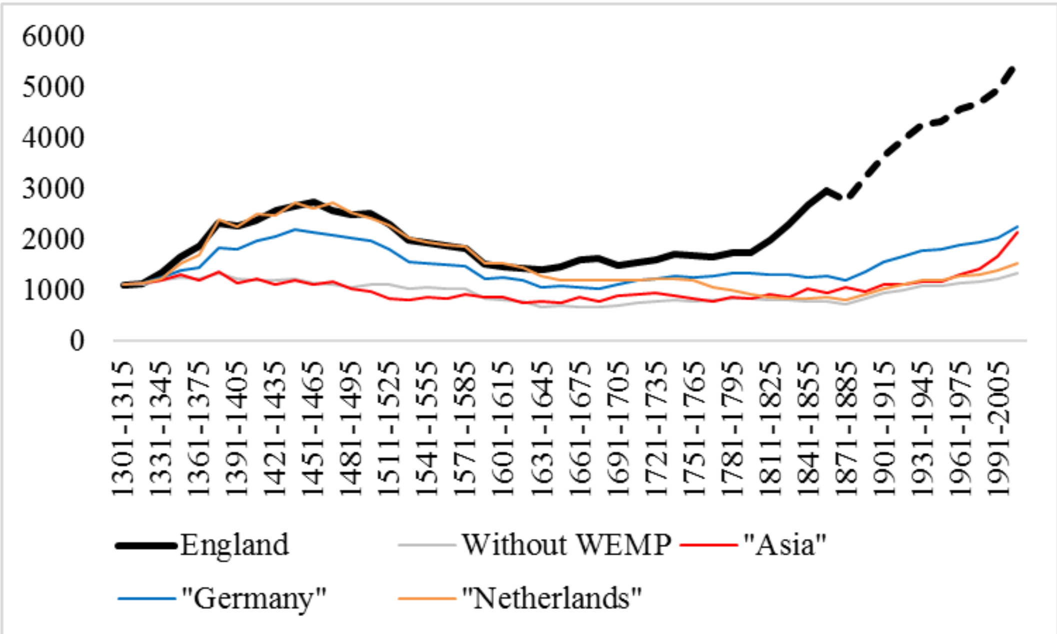
Table 5 Key Variables with and without Marriage Age Mechanism

Growth Rate of	Baseline	No Marriage Age Effect	How much lower in 1870
Population	0.073	0.067	15.78%
Wage	0.007	-0.0004	18.80%
Output	0.081	0.067	31.61%
Human Capital	0.136	0.106	53.15%

4.3 Other 'Countries'

To assess the capability of the model to explain the growth experiences in other economies, we simulate the model by re-calibrating the parameters and changing the path of shocks to match hypothesized key features of other economies than England. A caveat is due here—the simulations are NOT intended to match exactly other countries growth of real wages. Rather, the purpose is to show how the English economy might have deviated from the actual historical path if some parameters or the shock path were different. That is why the 'countries' are in quotation marks; only single possible characteristics of these countries are reflected in the simulations. To allow for delayed take-off timing in these simulations, we extend the simulation horizon another 10 periods (150 years).

Figure 9 Simulated Wage of Other 'Countries'



In this way we can show that real wages in England are predicted by the model to go on rising after 1870 (when the data used ends). The black solid line is the observed (Clark's) wage series, and the dotted line is the extended forecast based on bootstrapped shocks: (1) the mortality shocks are bootstrapped from the actual shocks during the entire Low Mortality Regime; (2) the productivity shocks are bootstrapped from the actual shocks during the last 6 periods (1780-1870). The latter is to reflect that there might be a structural break in productivity shock structure ('Modern Growth Regime').

To capture a non-European environment with a low marriage age (Hajnal, 1965, especially Table 4 and discussion) and low real wages that we term 'Asia' we broadly follow Jones (1981) in the model adjustments. We keep the simulation in the High Mortality Regime throughout and increase the utility weight of the number of children by 40% to capture a possible cultural difference[1]. Figure 9 shows that under these conditions the simulated model of the English economy yields a real wage that is very low and barely rises at all until the last decade of the twentieth century. Also female marriage age remains below 20 and population oscillates wildly.

In the first half of the 17th century the Netherlands was probably the most productive economy in the world. Then the country experienced a series of strong negative productivity shocks from 1650 to around 1820. Cromwell's Navigation Acts of the 1650s and Colbert's protectionist policies of the 1660s and 1670s undermined Dutch supremacy on the high seas and international markets with very adverse consequences for Dutch trade and shipping and for Dutch real wages (van Zanden and van Leuween, 2012). We bootstrap are model with the negative shocks replacing the actual shocks during that period. The simulation shows that, had England been subject to a sequence of strong negative shocks like the Netherlands, real wages would have fallen as they did in the Netherlands (Figure 9).

Germany's economy may have been more constrained by guild power than England's (Denison and Ogilvie, 2015); the production function then was less efficient. This could explain the lag behind England in the sustained rise of real wages, despite the European Marriage Pattern. In this simulation, θ_2, the output elasticity of human capital, is adjusted down by 10% to mimic a "German" type of stronger guild institutions (on the supply side). The simulation (Figure 9) shows no increase in real wages until after 1870.

[1] The shocks are therefore bootstrapped from the actual shocks during the High Mortality Regime, but the productivity shocks after 1945 are bootstrapped from the period 1780-1870.

5 Robustness

The robustness of the conclusions for England is checked against three uncertainties, model specification, data sources and estimation method. Regarding the model specification, we check different variations of the model, such as extending the demand side to five 'generations' and restricting the supply side without 'social capital' (setting $\theta_1 = \theta_2$). For robustness to the data sources, we use Allen (2007) real wage data as an alternative to Clark (2013) and also add patents data as an extra observable for human capital. To check robustness to estimation method, we re-estimate the model by maximum likelihood with a Kalman filter, in addition to simulated method of moments.

We focus on the robustness check for the data source, because it turns out that all model specification variants and estimation methods give very similar conclusions[1]. If we replace the Clark's wage data for the period 1770-1870 by Allen's wage data, then obviously the results for the High Mortality Regime are not affected. The estimated parameters for the model in the Low Mortality Regime are contrasted in Table 6. The differences are not substantial for most parameters, but a lower wage and a later 'Solow epoch' in Allen's series imply a different set of shocks for the Low Mortality Regime.

Table 6 Comparison of Estimated Parameters under Different Data Sources

	Clark's Data (as in Table 2)	**Allen's Data**	**Patents Data**
ρ_1	0.006	0.168	-0.025
ρ_2	0.190	-0.153	0.292
ρ_3	0.074	0.108	-0.024
ρ_x	0.206	0.425	0.201
σ_1	0.277	0.277	0.280
σ_2	0.412	0.416	0.412
σ_3	0.629	0.631	0.595
σ_x	0.050	0.050	0.050

Since all other conclusions, such as matching the moments, variance decomposition and impulse responses, are derived from the parameter values, these implications are fairly robust. A key reason is that our estimation method is to maximise the model's ability to match the data moments, rather than to match the data itself. A different data source may have different realisations of a variable, but the moments (mean and standard deviation) are very stable. This is another advantage over maximum likelihood (a more detailed discussion can be found in Appendix III).

[1] Results are available from the corresponding author on request.

A possible source of human capital data is patents sealed in England since 1617 (Madsen et al, 2010; Mitchell, 1962 p268). If this additional observable is used, we need to introduce a fifth structural shock (for simplicity, we use a measurement error here) to avoid 'stochastic singularity'. In the state space form of the model, one more measurement equation is added to describe the relationship between the growth rates of accumulative patents and human capital:

(S5) $g_{patents,t} = \kappa \cdot g_{H,t} + measurement\_error_t$.

The structural parameters are again not significantly different from the previous two choices of data, with κ equal to 0.322 and the standard deviation of the measurement error equal to 0.062. The estimation results of shock persistence and volatilities are shown in the last column of Table 6.

However, the use of patents to measure human capital has drawbacks. First, there was apparently no patent protection before 1450, but human capital must have existed then. Moreover, patent protection in the early years was often a grant of monopoly powers over existing assets rather than a protection for those newly created. Therefore, the link between patents and human capital is difficult to capture in a measurement equation. Even when the regular series begins there are some years with no patents sealed at all (1643-1659 for example due to civil war and the Protectorate) but human capital is unlikely to have been unchanged over this period. To ensure the relationship in the measurement equation is as stable as possible, we discard most of the early years of patents data but this also weakens the value of the additional observable. Hence, we only use patents data as a robustness check.

6 Conclusion

The structure of our unified growth model for England differs from that of Galor and Weil (2000) and Galor (2005, 2011) in its greater historical specificity. A distinctive response to catastrophic mortality sets off the process that eventually gives rise to the break out from the Malthusian epoch but there was no necessity for the particular response. The model shows how the Western European Marriage Pattern emerged as a result of the high mortality of the 14th and 15th centuries (we conclude that Chaucer's 14th century fictional 'Wife of Bath' was married unusually young at 12, even for the time).

Whereas Malthus saw the consequence of the late marriage 'preventative check' as being simply population control, the model proposed here represents the change as more fundamental. It can demonstrate how the high female age at first marriage ultimately contributed to the shift away from the Malthusian equilibrium—through increasing ‘child quality’ and family-based human capital accumulation. However, there was little substitution between child quality and child quantity, judging by the low value of the elasticity of substitution (0.154). Apparently child quality and quantity were complementary, consistent with Clark and Cummins (2016).

Mortality crises and high mortality levels eventually diminished so that a new stage of development began. This period initially exhibits Malthusian inverse fluctuations between wage and population growth. Around 1780 the model shows the economy entering a third stage in which first population and then real wages grow secularly, the Industrial Revolution. Economic growth lags behind demographic growth. As the intensity and frequency of mortality crises diminish, more children survive and are planned immediately. The consequence of their higher quality, greater human capital, takes longer to work through the economy.

The cyclical behaviour of the data series is matched by a similar pattern in the impulse response functions. This is due to the multiple overlapping-generation structure of the model. Each generation has its own smoothed impulse responses (a feature of rational behaviour), but they act in their own interest with different timing. The aggregation of the four generations in each period creates the final shape of the overall impulse responses. Given the regular and persistent pattern, this behaviour must be rooted in the structure of the model.

The model implications for positive and preventative checks in pre-industrial England conform generally with those of other recent researchers (e.g. Crafts and Mills, 2009). However, the present approach draws attention to the difference between transition dynamics of these checks, and the steady state relations between the key variables. The steady state predicts a negative relationship between mortality rate and marriage age but the transition can give rise to a positive relation, depending on whether wealth effects dominate in the transition.

The explanation of how human capital accumulation triggered the English Industrial Revolution is necessarily pitched at a highly abstract and general level. More detailed historical accounts, such as Allen (2009), can be entirely consistent with the model estimated here. Human capital in the form of greater skill and discipline warranted higher wages. These high wages incentivised the search for innovations that would reduce labour costs. Now more widespread skills triggered important innovations that made industrial breakthroughs (Khan, 2015).

The model presented here is less ambitious than theoretical unified growth models in that we have not extended the estimation to the fertility decline period. In principle rising incomes and the time opportunity cost of children embedded in the model could bring about such a transition. However, the structural change in the English economy by 1870 seems highly likely to have changed key parameters, necessitating abandoning the simplicity of our one sector model that captures data characteristics back to 1300.

References

A'Hearn, B., J. Baten and D. Crayen (2009). 'Quantifying Quantitative Literacy: Age Heaping and the History of Human Capital', *Journal of Economic History*, vol. **69**(**3**), pp. 783-808.

Allen, R. C. (2007). 'Pessimism Preserved: Real Wages in the British Industrial Revolution', *Oxford University Department of Economics Working Paper*.

Allen, R. C. (2009). *The British Industrial Revolution in Global Perspective*, Cambridge: Cambridge University Press.

Ashraf, Q. and O. Galor (2011). 'Dynamics and Stagnation in the Malthusian Epoch', *American Economic Review*, vol. **101**(**5**), pp. 2003-2041.

Bailey, M. (1996). 'Demographic Decline in Late Medieval England: Some Thoughts on Recent Research', *Economic History Review*, vol. **49**(**1**), pp. 1-19.

Bailey, M. (2014). *The Decline of Serfdom in Late Medieval England: From Bondage to Freedom*, Woodbridge: Boydell Press.

Baten, J. and J. L. Van Zanden (2008). 'Book Production and the Onset of Modern Economic Growth', *Journal of Economic Growth*, vol. **13**(**3**), pp. 217-235.

Becker, G. S., K. M. Murphy and R. Tamura (1990). 'Human Capital, Fertility, and Economic Growth', *Journal of Political Economy*, vol. **98**(**5**), pp. S12-S37.

Behrman, J. R., A. D. Foster, M. R. Rosenzweig and P. Vashishtha (1999). 'Women's Schooling, Home Teaching and Economic Growth', *Journal of Political Economy*, vol. **107**, pp. 682-714.

Blanchard, O. and C. Kahn (1980). 'The Solution of Linear Difference Models under Rational Expectations', *Econometrica*, vol. **48**(**5**), pp. 1305-1312.

Boucekkine, R., D. De La Croix and O. Licandro (2003). 'Early Mortality Declines at the Dawn of Modern Growth', *Scandinavian Journal of Economics*, vol. **105**(**3**), pp. 401-418.

Broadberry, S. N. Bruce M. S. Campbell, Alexander Klein, Mark Overton, and Bas van Leeuwen (2015) *British Economic Growth, 1270-1870*, Cambridge University Press.

Carneiro, P., C. Meghir and M. Parey (2007). 'Maternal Education, Home Environments and the Development of Children and Adolescents', *Institute for Fiscal Studies Working Papers*.

Cervellati M. and U. Sunde (2015) The Economic and Demographic Transition, Mortality, and Comparative Development *American Economic Journal: Macroeconomics* 7(3): 189–225

Cervellati, M. and U. Sunde (2005). 'Human Capital Formation, Life Expectancy, and the Process of Development', *American Economic Review*, vol. **95**(**5**), pp. 1653-1672.

Chen, K (2012) Analysis of the Great Divergence under a Unified Endogenous Growth Model,*Annals of Economics and Finance*,. 13,. 2, pp. 317-53
Chesnais, J. C. (1992). *The Demographic Transition: Stages, Patterns and Economic Implications; A Longitu-dinal Study of Sixty-Seven Countries Covering the Period 1720-1984*, Oxford.
Choo, E. and Siow, A. (2006), 'Who Marries Whom and Why', *Journal of Political Economy*, **114** (**1**), pp. 175-201.
Christiano, L., M. Eichenbaum and C. Evans (2005). 'Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy', *Journal of Political Economy*, vol. **113**(**1**), pp. 1-45.
Clark, G. (1987). 'Why Isn't the Whole World Developed? Lessons from the Cotton Mills', *Journal of Economic History*, vol. **47**(**1**), pp. 141-173.
Clark, G. (2005). 'The Condition of the Working Class in England, 1209 to 2004', *Journal of Political Economy*, vol. **113**, pp. 1307-1340.
Clark, G. (2007). 'The Long March of History: Farm Wages, Population, and Economic Growth, England 1209-1869', *Economic History Review*, vol. **60**(**1**), pp. 97-135.
Clark, G. (2013). 'What were the British Earnings and Prices Then?' http://www.measuringworth.com/ukearncpi/.
Clark, G. and Cummins, N., (2016). "The Child Quality-Quantity Tradeoff, England, 1780-1880: A Fundamental Component of the Economic Theory of Growth is Missing," CEPR Discussion Papers 11232,
Crafts, N. and T. C. Mills (2009). 'From Malthus to Solow: How did the Malthusian economy really evolve?'*Journal of Macroeconomics*, vol. **31**(**1**), pp. 68-93.

Cummins, N. (2014) 'Longevity and the Rise of the West: Lifespans of the European Elite, 800-1800', *EHES Working Papers in Economic History*, 64.

de la Croix, D and Licandro, D O (2012) ‘The Longevity of Famous People from Hammurabi to Einstein’, *Barcelona GSE Working Paper*666.
De Moor, T. and J. L. Van Zanden (2010). 'Girl Power: The European Marriage Pattern and Labor Markets in the North Sea Region in the Late Medieval and Early Modern Period', *Economic History Review*, vol. **63**(**1**), pp. 1-33.
Dedola, L. and L. Giovanni (2012). 'Financial Frictions, Financial Integration and the International Propagation of Shocks', *Economic Policy*, vol. **27**, pp. 319-359.
Dennison T and Ogilvie S (2014) Does the European Marriage Pattern Explain Growth? *Journal of Economic History* 74: 651-692.
Desmet, K and Parente, S L. (2012) The Evolution of Markets and the Revolution of Industry: A Unified Theory of Growth, *Journal of Economic Growth*, September 17, 3, pp. 205-3.
DeTray, D. N. (1973). 'Child Quality and the Demand for Children', *Journal of Political Economy*, vol. **81**(**2**), pp. S70-S95.
Diebolt, C. and F. Perrin (2013a). 'From Stagnation to Sustained Growth: The Role of Female Empowerment', *AFC Working Paper*.
Diebolt, C. and F. Perrin (2013b). 'From Stagnation to Sustained Growth: The Role of Female Empowerment', *American Economic Review Papers and Proceedings*, vol. **103**(**3**), pp. 545-549.
Field, E. and A. Ambrus (2008). 'Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh', *Journal of Political Economy*, vol. **116**(**5**), pp. 881-930.
Flinn, M. W. (1981). *The European Demographic System 1500-1820*, Brighton: Harvester.

Foreman-Peck, J. (2011). 'The Western European marriage pattern and economic development', *Explorations in Economic History*, vol. **48**(**2**), pp. 292-309.

Gaiha, R. and V. Kulkarni (2005). 'Anthropometric Failure and Persistence of Poverty in Rural India', *International Review of Applied Economics*, vol. **19**(**2**), pp. 179-197.

Galor, O. (2011) *Unified Growth Theory*, Princeton University Press.

Galor, O. (2005), "From Stagnation to Growth: Unified Growth Theory," in Philippe Aghion and Steven N. Durlauf (eds.)*Handbook of Economic Growth*, North-Holland, vol**1A**, pp. 171-293.

Galor, O. (2010) The 2008 Lawrence R. Klein Lecture--Comparative Economic Development: Insights from Unified Growth Theory, *International Economic Review*, 51, 1, pp. 1-44.

Galor, O. and D. N. Weil (2000). 'Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and beyond', *The American Economic Review*, vol. **90**(**4**), pp. 806-828.

Galor, O. and O. Moav (2002). 'Natural Selection and the Origin of Economic Growth', *Quarterly Journal of Economics*, vol. **117**(**4**), pp. 1133-1191.

Georgiadis, A. and A. Manning (2011). 'Change and Continuity among Minority Communities in Britain', *Journal of Population Economics*, vol. **24**, pp. 541-568.

Grantham, G. (1999). 'Contra Ricardo: On the Macroeconomics of Pre-Industrial Economies', *European Review of Economic History*, vol. **3**(**2**), pp. 199-232.

Hajnal, J. (1965). 'European Marriage Patterns in Perspective', in (Glass, D. V. and D. E. C. Eversley Eds.), *Population in History: Essays in Historical Demography*: Edward Arnold.

Hallam, H. E. (1985). 'Age at First Marriage and Age at Death in the Lincolnshire Fenland 1252-1478', *Population Studies*, vol. **39**, pp. 55-69.

Hatcher, J. (1986). 'Mortality in the Fifteenth Century: Some New Evidence', *Economic History Review*, vol. **39**(**1**), pp. 19-38.

Hatcher, J. and M. Bailey (2001). *Modelling the Middle Ages*, Oxford.

Humphries, J. and J. Weisdorf (2015). 'The Wages of Women in England 1260-1850'. *Journal of Economic History*

Jones E L (1981) *The European Miracle:: Environments Economics and Geopolitics in the History of Europe and Asia*, Cambridge University Press

Keeley, M. C. (1977) ‘The Economics of Family Formation’ *Economic Inquiry*, 15(2) 238–250.

Khan, Z. (2015). 'Knowledge, Human Capital and Economic Development: Evidence from the British Industrial Revolution, 1750-1930', *NBER Working Paper Series*.

Kollmann, R. (2013). 'Global Banks, Financial Shocks, and International Business Cycles: Evidence from an Estimated Model', *Journal of Money, Credit and Banking*, vol. **45**, pp. 159-195.

Lagerlof, N.-P. (2003). 'Mortality and Early Growth in England, France and Sweden', *Scandinavian Journal of Economics*, vol. **105**(**3**), pp. 419-439.

Lee, R. D. (1985). 'Inverse Projection and Back Projection: Comparative Results and Sensitivity Tests for England', *Population Studies*, pp. 233-248.

Lee, R. D. (1993a). 'Accidental and Systematic Change in Population History: Homeostasis in a Stochastic Setting', *Explorations in Economic History*, vol. **30**, pp. 1-3.

Lee, R. D. (1993b). 'Inverse Projection and Demographic Fluctuations: A Critical Assessment of New Methods', in (Reher, D. and R. S. Schofield Eds.), *Old and New Methods in Historical Demography*, New York: Oxford University Press.

Lee, R. D. and M. Anderson (2002). 'Malthus in State Space: Macroeconomic-Demographic Relations in English History, 1540-1870', *Journal of Population Economics*, vol. **15**, pp. 195-220.

Maddison, A. (2014). 'Statistics on World Population, GDP and Per Capita GDP, 1-2008 AD', *http://www.ggdc.net/maddison/oriindex.htm*.

Madsen, J. B., J. B. Ang and R. Banerjee (2010). 'Four Centuries of British Economic Growth: The Roles of Technology and Population', *Journal of Economic Growth*, vol. **5**, pp. 263-290.

Marshall, A. (1961). *Principles of Economics*: MacMillan.

Mitch, D. F. (1992). *The Rise of Popular literacy in Victorian England: The Influence of Private Choice and Public Policy*, Philadelphia: University of Pennsylvania Press.

Mitchell, B. R. (1962). *Abstract of British Historical Statistics*, Cambridge.

Mourmouras, A and Rangazas, P. (2009) Reconciling Kuznets and Habbakuk in a Unified Growth Theory Journal of Economic Growth, 14, 2, pp. 149-81

Murray, J. E. (2004). 'Family Literacy and Skill Training in the Antebellum South: Historical-Longitudinal Evidence from Charleston', *Journal of Economic History*, vol. **63**(**3**), pp. 773-779.

Myrskylä, M. and A. Fenelon (2011). 'Maternal Age and Offspring Adult Health: Evidence from the Health and Retirement Study', *Max Planck Institute for Demography Working Paper*.

Nordblom, K. (2003). 'Is Increased Public Schooling Really a Policy for Equality? The Role of Within-the-Family Education', *Journal of Public Economics*, vol. **87**(**9-10**), pp. 1943-1965.

Parker, G. (2013) *Global Crisis: War, Climate Change and Catastrophe in the Seventeenth Century*, Yale University Press.

Peng, J.-Y. and J. Aston (2011). 'The State Space Models Toolbox for Matlab', *Journal of Statistical Software*, vol. **41**(**6**), pp. 1-26.

Phelps-Brown, E. H. and S. V. Hopkins (1956). 'Seven Centuries of the Prices of Consumables Compared with Builders' Wage Rates', *Economica*, vol. **23**, pp. 296-314.

Putnam, R. (1993). 'The Prosperous Community', *The American Prospect*, vol. **4**(**13**).

Putnam, R. (1995). 'Bowling Alone: America's Declining Social Capital', *Journal of Democracy*, vol. **6**(**1**), pp. 65-78.

Rogers, T. E. (1866). *A History of Agriculture and Prices in England.*

Ruge-Murcia, F. (2007). 'Methods to Estimate Dynamic Stochastic General Equilibrium Models', *Journal of Economic Dynamics and Control*, vol. **31**(**8**), pp. 2599-2636.

Ruggles, S. (1992). 'Migration, Marriage and Mortality: Correcting Sources of Bias in English Family Re-constitutions', *Population Studies*, vol. **46**, pp. 507-522.

Ruggles, S. (1999). 'The Limitations of English Family Reconstitution: English Population History from Family Reconstitution 1580-1837', *Continuity and Change*, vol. **14**(**1**), pp. 105-130.

Schmitt-Grohé, S. and M. Uribe (2004). 'Solving Dynamic General Equilibrium Models using a Second-Order Approximation to the Policy Function', *Journal of Economic Dynamics and Control*, vol. **28**(**4**), pp. 755-775.

Schofield, R. (1985) ‘English Marriage Patterns Revisited’ *Journal of Family History*Spring 10:1 2-20
Scitovsky, T. (1992). *The Joyless Economy: The Psychology of Human Satisfaction*, Oxford: Oxford University Press.
Smets, F. and R. Wouters (2003). 'An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area', *Journal of the European Economic Association*, vol. **1**(**5**), pp. 1123-1175.
Smets, F. and R. Wouters (2007). 'Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach', *The American Economic Review*, vol. **97**(**3**), pp. 586-606.
Stone, L. (1977). *Family, Sex and Marriage in England 1500-1800,* New York: Harper and Row.
Strulik, H. and Weisdorf, J (2008) Population, Food, and Knowledge: A Simple Unified Growth Theory *Journal of Economic Growth*, 13, 3, 195-216
Sutcliffe, A. G., J. Barnes and J. Belsky (2012). 'The Health and Development of Children Born to Older Mothers in the United Kingdom: Observational Study Using Longitudinal Cohort Data', *British Medical Journal*.
Uhlig, H. (1998). 'A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily', in (Marimon, R. and A. Scott Eds.), *Computational Methods for the Study of Dynamic Economies*, Oxford: Oxford University Press.
Van Zanden J L (2009) *The Long Road to the Industrial Revolution*, Brill
Van Zanden, J L and van Leeuwen, B (2012) ‘Persistent but not consistent: The growth of national income in Holland 1347–1807’, *Explorations in Economic History* 49 (2), 119-130.
Voigtländer, N. and H-J Voth (2006) Why England? Demographic factors, structural change and physical capital accumulation during the Industrial Revolution *Journal of Economic Growth* 11:319–361
Voigtländer, N. and H.-J. Voth (2013). 'How the West 'Invented' Fertility Restriction', *American Economic Review*, vol. **103**, pp. 2227-2264.
Weir, D. (1984) ‘Rather Never than Late: Celibacy and Age at Marriage in English Cohort fertility, 1541-1871’ *Journal of Family History* 9 341-355
Wrigley, E. A. and R. S. Schofield (1981). *The Population History of England, 1541-1871: A Reconstruction*, London: Arnold.
Wrigley, E. A., R. S. Davies, J. E. Oeppen and R. S. Schofield (1997). *English Population History from Family Reconstitution 1580-1837*: Cambridge University Press.

Appendix I: Graphs and Tables of Data

Demographic data are based on Wrigley and Schofield (1981) and Broadberry et al (2015), wage data are based on Clark (2013) and Allen (2007), and steady state mortality rates are calibrated using Wrigley and Schofield (1989, p714).

Figure 10 Birth Flows, Death Flows and Population Level (Annual)

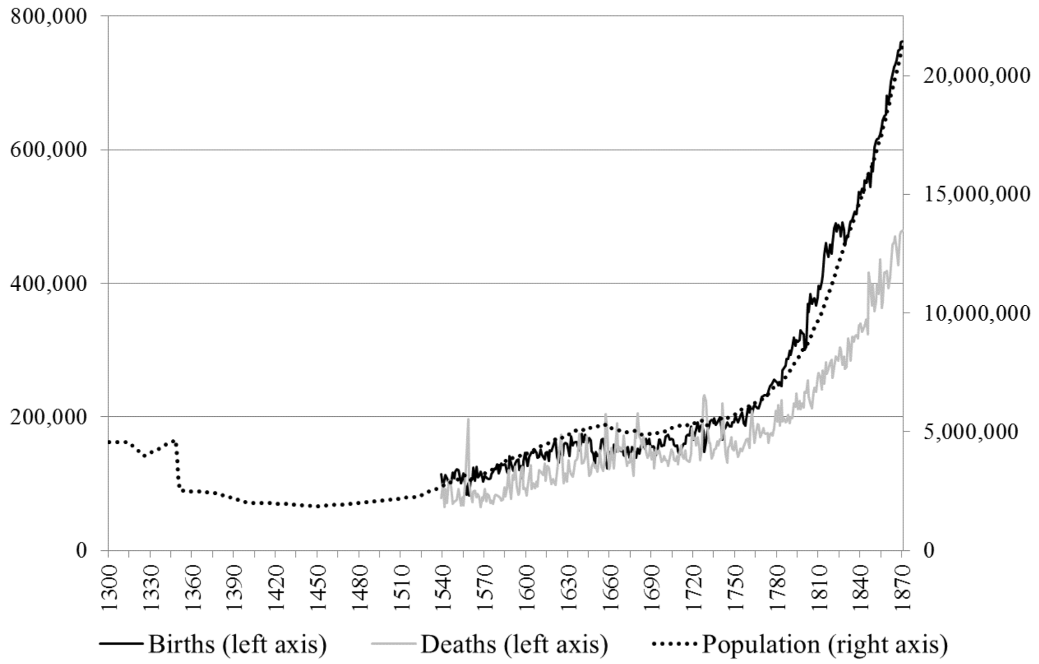


Figure 11 Birth Rate and Death Rate (Annual)

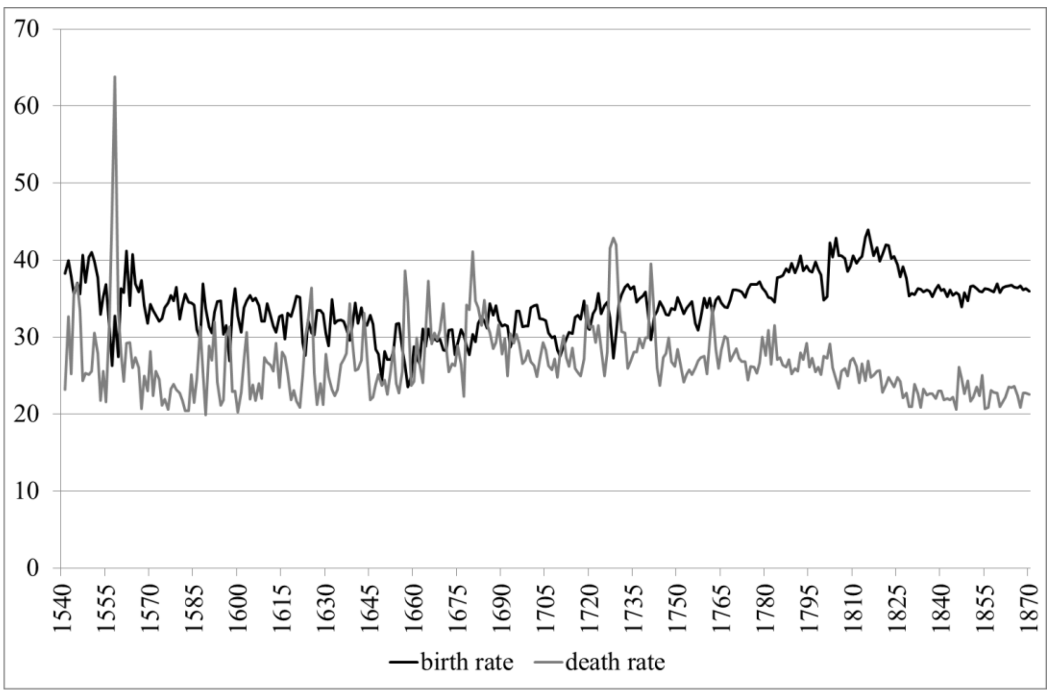


Figure 12 Clark Real Wage VS. Allen Real Wage (Annual)

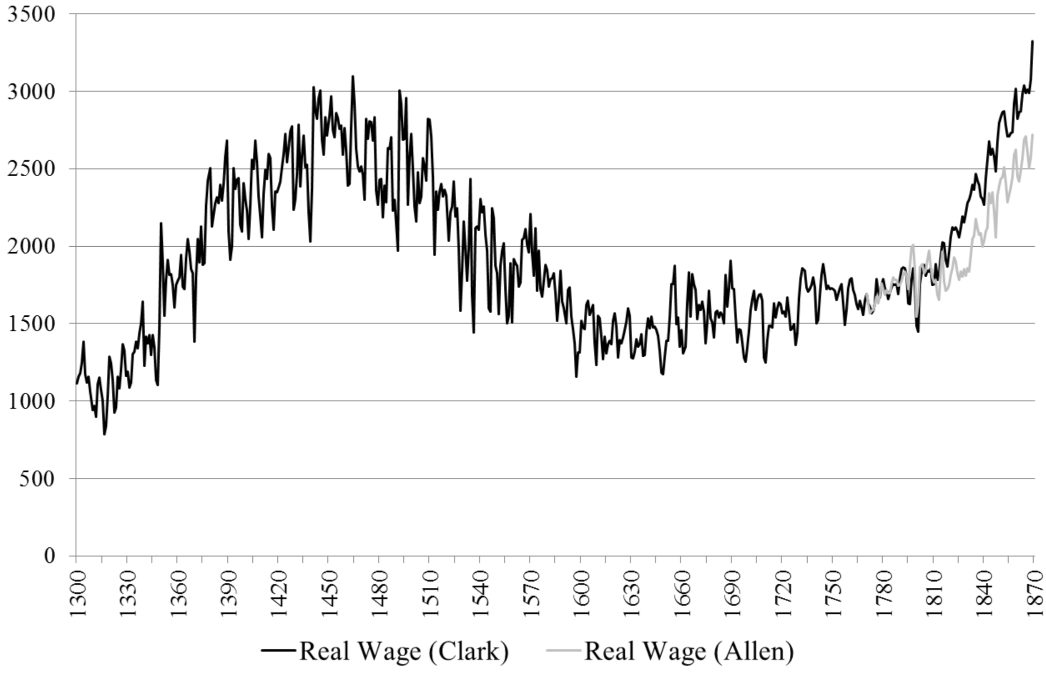


Figure 13 Population Growth Rate and Wage Growth Rate (Annual)

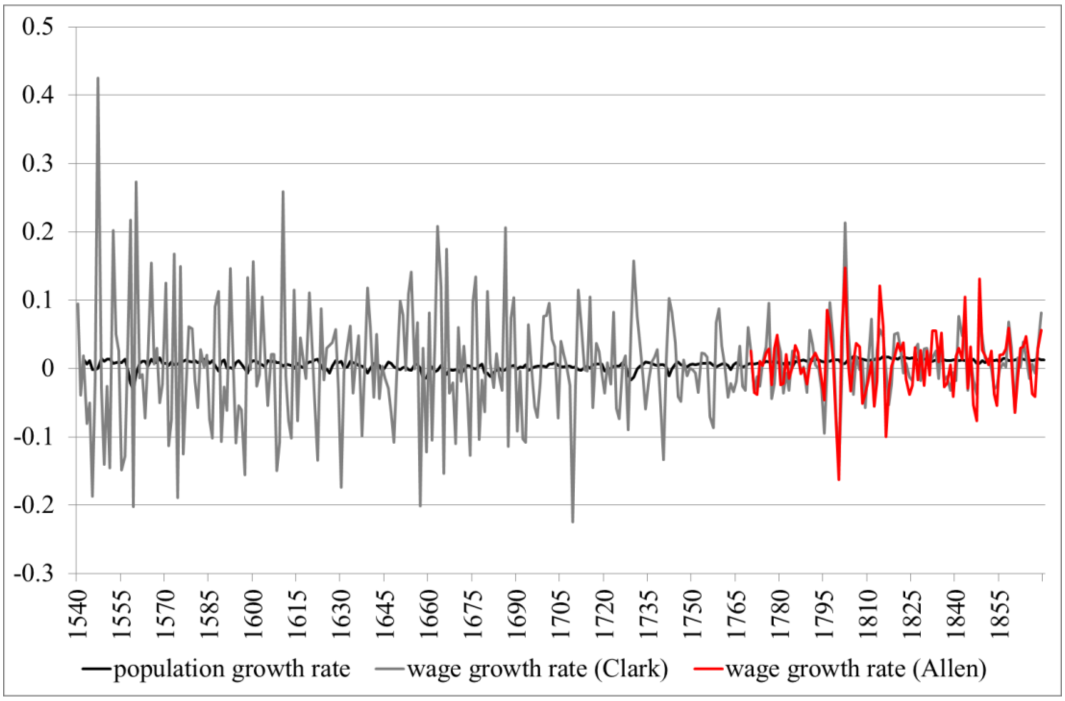


Figure 14 Historical Fluctuations of Key Variables (15-year)

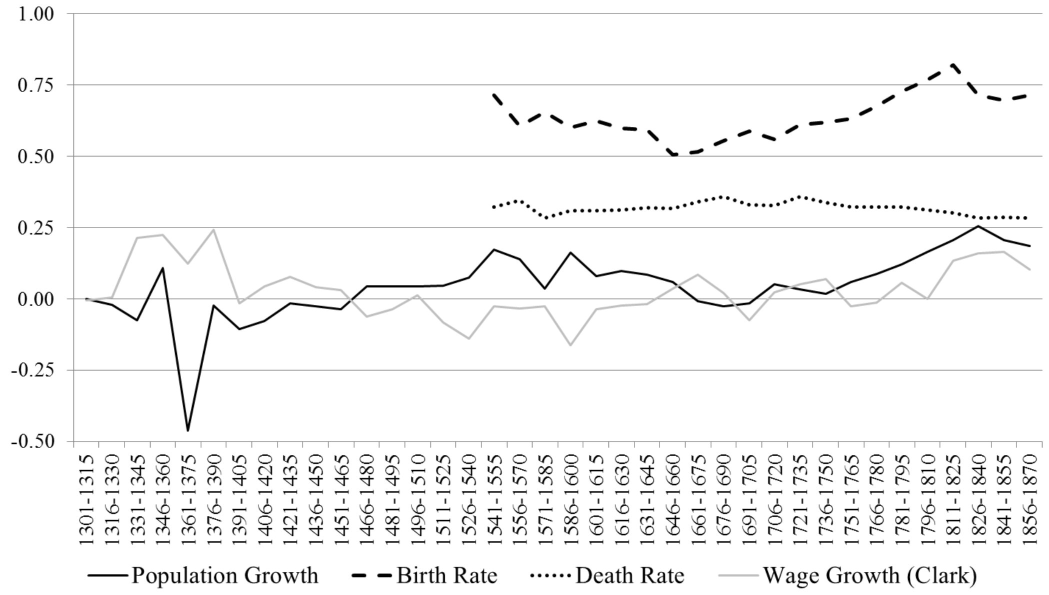


Table 7 Life Table and Life Expectancy

i	Age (A_i)	**W&S Life Table** d_i level 3	level 9	**Survival Rate** s_i level 3	level 9	**Unconditional** m_i level 3	level 9
1	0	279.2	164.1	72.08%	83.59%	27.92%	16.41%
2	1	265.7	148.6	52.93%	71.17%	19.15%	12.42%
3	5	91.9	51.7	48.06%	67.49%	4.86%	3.68%
4	10	49.3	28.2	45.69%	65.59%	2.37%	1.90%
5	15	58.3	34.8	43.03%	63.30%	2.66%	2.28%
6	20	77.6	47.2	39.69%	60.32%	3.34%	2.99%
7	25	86.3	52.2	36.27%	57.17%	3.43%	3.15%
8	30	95.4	57.4	32.81%	53.89%	3.46%	3.28%
9	35	106.5	63.8	29.31%	50.45%	3.49%	3.44%
10	40	118.4	72.1	25.84%	46.81%	3.47%	3.64%
11	45	135.0	83.4	22.35%	42.91%	3.49%	3.90%
12	50	158.2	101.6	18.82%	38.55%	3.54%	4.36%
13	55	205.8	132.6	14.94%	33.44%	3.87%	5.11%
14	60	275.2	182.6	10.83%	27.33%	4.11%	6.11%
15	65	377.8	258.1	6.74%	20.28%	4.09%	7.05%
16	70	527.2	370.7	3.19%	12.76%	3.55%	7.52%
17	75	682.4	503.7	1.01%	6.33%	2.17%	6.43%
18	80	790.1	638.3	0.21%	2.29%	0.80%	4.04%
19	85	888.4	783.2	0.02%	0.50%	0.19%	1.79%
20	90	944.8	886.3	0.00%	0.06%	0.02%	0.44%
	Life Expectancy	**23.54**	**38.23**				

The steady state mortality rates are calibrated based on the life table level 3 and level 9 (the two grey columns), which gives the conditional mortality rates per thousand (d_i)[1] at each specific age assuming survival from the previous age. To make use of this life table to calibrate the steady state mortality rates in our model, we first translate the conditional mortality rates into unconditional mortality rates (m_i). This is most easily done through survival rates (s_i):

$$s_i = \frac{1000 - d_i}{1000} \text{ for } i = 1;$$

$$s_i = \frac{1000 - d_i}{1000} \times s_{i-1} \text{ for } i = 2, \dots, 20;$$

which are linked with both conditional mortality rates and unconditional morality rates:

$$m_i = \frac{d_i}{1000} \text{ for } i = 1;$$

$$m_i = \frac{d_i}{1000} \times s_{i-1} \text{ for } i = 2, \dots, 20.$$

With these unconditional mortality rates (or distribution of age) at hand, we can then calculate the average age of those who die between 0 and 15 ($\overline{A1}$), between 16 and 30 ($\overline{A2}$), between 31 and 45 ($\overline{A3}$), and between 46 and 60 ($\overline{A4}$):

$$\overline{A1} = \sum_{i=0}^{5} \left(\frac{m_i}{\sum_{i=0}^{5} m_i} \times A_i \right); \ \overline{A2} = \sum_{i=6}^{8} \left(\frac{m_i}{\sum_{i=6}^{8} m_i} A_i \right); \ \overline{A3} = \sum_{i=9}^{11} \left(\frac{m_i}{\sum_{i=9}^{11} m_i} A_i \right); \ \overline{A4} = 60.$$

Note that $\overline{A4}$ is set to 60 because it would be above 60 if we use the formula:

$$\overline{A4} = \sum_{i=12}^{20} \left(\frac{m_i}{\sum_{i=12}^{20} m_i} A_i \right).$$

Lastly, we make use of these average ages to calculate the life expectancies, i.e. a weighted average with weights being the unconditional model mortality rates:

$$\bar{A} = w_1 \overline{A1} + w_2 \overline{A2} + w_3 \overline{A3} + w_4 \overline{A4},$$

where the weights are linked with the conditional model mortality rates in the model: $w_1 = \overline{m1}$, $w_2 = (1 - w_1)\overline{m2}$, $w_3 = (1 - w_1 - w_2)\overline{m3}$ and $w_4 = (1 - w_1 -$

[1] The subscript i is the row index, so the age of 0 corresponds to $i = 1$, the age of 1 corresponds to $i = 2$ and the age of 90 corresponds to the last row $i = 20$.

$w_2 - w_3)\overline{m4}$. We can calibrate these $\overline{mi}$'s to match the target life expectancies (23.54 for level 3 and 38.23 for level 9) using a numerical algorithm. The results are shown in the table below and the calibrated conditional mortality rates ($\overline{m1}, \overline{m2}, \overline{m3}$) correspond to the values in Table 1.

	Level 3			**Level 9**		
Mortality Rate	$\overline{mi}$	w_i	$\overline{Ai}$	$\overline{mi}$	w_i	$\overline{Ai}$
$\overline{m1}$	**50.00%**	50.00%	1.880	**30.00%**	30.00%	2.291
$\overline{m2}$	**30.27%**	15.14%	25.059	**12.08%**	8.46%	25.156
$\overline{m3}$	**30.27%**	10.55%	39.998	**12.40%**	7.63%	40.212
$\overline{m4}$	**100%**	24.31%	60	**100%**	53. 91%	60
Life Expectancy	**23.54**			**38.23**		

Appendix II: Linearised Equilibrium Conditions

$$U_t = \left[\alpha \cdot \left(\frac{n_t}{2} \right)^{\frac{s-1}{s}} + \beta \cdot \left(q_t \right)^{\frac{s-1}{s}} + \gamma \cdot \left(\frac{\hat{z}_t \hat{P}_{t-1}^{\theta_1 - 1} \hat{H}_{t-1}^{\theta_2}}{\hat{z}_{t-1}} \right)^{\frac{s-1}{s}} \right]^{\frac{s}{s-1}}$$

$$\alpha \left(\frac{2U_t}{n_t} \right)^{\frac{1}{s}} = \frac{1}{2} \hat{\lambda}_t \hat{P}_t^{\theta_1 - 1} \hat{H}_t^{\theta_2} \left[\hat{\pi}_{n,t+1} + \hat{\pi}_{q,t+1} q_t + \frac{\eta_A \hat{\pi}_{q,t+1}}{A_t} \frac{a_1}{\left(1 - m1_{t+1}\right)\left(1 - m2_{t+2}\right)} q_t n_t \right]$$

$$\beta \left(\frac{U_t}{q_t} \right)^{\frac{1}{s}} = \frac{1}{2} \hat{\lambda}_t \hat{P}_t^{\theta_1 - 1} \hat{H}_t^{\theta_2} \hat{\pi}_{q,t+1} n_t$$

$$\gamma \left(\frac{\hat{z}_{t-1} U_t}{\hat{z}_t} \right)^{\frac{1}{s}} = 3 \hat{\lambda}_t \hat{z}_{t-1} \left(\hat{P}_{t-1}^{\theta_1 - 1} \hat{H}_{t-1}^{\theta_2} \right)^{\frac{1-s}{s}}$$

$$\frac{\hat{w}_t}{\hat{P}_t^{\theta_1 - 1} \hat{H}_t^{\theta_2}} + \hat{w}_{t+1} + \hat{w}_{t+2} \hat{P}_{t+1}^{\theta_1 - 1} \hat{H}_{t+1}^{\theta_2} = \frac{1}{2} \left(\hat{\pi}_{n,t+1} n_t + \hat{\pi}_{q,t+1} q_t n_t \right) + \frac{3\hat{z}_t}{\hat{P}_t^{\theta_1 - 1} \hat{H}_t^{\theta_2}}$$

$$A_t = a_0 - a_1 \frac{n_t}{\left(1 - m1_{t+1}\right)\left(1 - m2_{t+2}\right)}$$

$$\hat{\pi}_{q,t} = \frac{b_0}{\left(1 - m1_t\right)\left(1 - m2_{t+1}\right)} \left(\frac{b_1}{A_{t-1}} \right)^{\eta} \hat{w}_t$$

$$\hat{\pi}_{n,t} = \frac{c_0}{\left(1 - m1_t\right)\left(1 - m2_{t+1}\right)} \hat{w}_t$$

$$\hat{P}_t = 1 - \hat{D}_t + \hat{B}_t$$

$$\hat{D}_t = m1_t \hat{B}_t + m2_t \hat{G}1_{t-1} + m3_t \hat{G}2_{t-1} + \hat{G}3_{t-1}$$

$$\hat{G}1_t = \left(1 - m1_t\right) \frac{\hat{B}_t}{\hat{P}_t}$$

$$\hat{G}2_t = \left(1 - m1_{t-1}\right)\left(1 - m2_t\right) \frac{\hat{B}_{t-1}}{\hat{P}_t \hat{P}_{t-1}}$$

$$\hat{G}3_t = \left(1 - m1_{t-2}\right)\left(1 - m2_{t-1}\right)\left(1 - m3_t\right) \frac{\hat{B}_{t-2}}{\hat{P}_t \hat{P}_{t-1} \hat{P}_{t-2}}$$

$$b_t \equiv \hat{B}_t = \frac{\left(1 - \mu_t\right) \hat{G}2_{t-1}}{2} \frac{n_{t-1}}{\left(1 - m1_t\right)\left(1 - m2_{t+1}\right)}$$

$$\mu_t = \bar{\mu} + \tau_1 \frac{A_t - A_t^*}{A_t^*} + \tau_2 \frac{w_t - w_t^*}{w_t^*}$$

$$d_t = \frac{\hat{D}_t}{1 + \hat{B}_t}$$

$$\hat{Y}_t = e^{x_t}$$

$$\hat{H}_t = \hat{Q}_t \hat{G}1_t + \frac{\hat{Q}_{t-1}}{\hat{H}_{t-1}} \hat{G}2_t + \frac{\hat{Q}_{t-2}}{\hat{H}_{t-1}\hat{H}_{t-2}} \hat{G}3_t$$

$$\hat{Q}_t = e^{x_t} \left(\frac{\hat{Q}_{t-2}}{\hat{H}_{t-1}\hat{H}_{t-2}} q_{t-1} \right)^{1-\varepsilon}$$

$$\theta_1 \hat{Y}_t = \hat{w}_t$$

$$1 + g_{w_t} = \frac{\hat{w}_t}{\hat{w}_{t-1}} \hat{P}_{t-1}^{\theta_1 - 1} \hat{H}_{t-1}^{\theta_2}$$

$$1 + g_{Y_t} = \frac{\hat{Y}_t}{\hat{Y}_{t-1}} \hat{P}_{t-1}^{\theta_1} \hat{H}_{t-1}^{\theta_2}$$

$$1 + g_{H_t} = \hat{H}_t$$

$m1_t = \overline{m1} \cdot \exp(mm1_t)$, where $mm1_t = \rho_1 \cdot mm1_{t-1} + emm1_t$

$m2_t = \overline{m2} \cdot \exp(mm2_t)$, where $mm2_t = \rho_2 \cdot mm2_{t-1} + emm2_t$

$m3_t = \overline{m3} \cdot \exp(mm3_t)$, where $mm3_t = \rho_3 \cdot mm3_{t-1} + emm3_t$

$$x_t = \rho_x \cdot x_{t-1} + ex_t$$

NB: For simplicity, we omit the expectation operators $\mathbf{E}_t[\cdot]$, but any variable with a time subscript beyond period t should be treated as the expected value of that variable based on information set available in period t.

To summarise, this nonlinear dynamic system has 27 endogenous variables, denoted as $\mathbf{y}_t$, and 4 exogenous shocks, denoted as $\mathbf{u}_t$, where:

$$\mathbf{y}_t \equiv \begin{bmatrix} U_t; n_t; q_t; \hat{z}_t; \hat{\pi}_{nt}; \hat{\pi}_{qt}; A_t; \hat{\lambda}_t; \mu_t; \\ \hat{P}_t; \hat{B}(b_t); \hat{D}_t; d_t; \hat{G}1_t; \hat{G}2_t; \hat{G}3_t; \\ \hat{Y}_t; \hat{Q}_t; \hat{H}_t; \hat{w}_t; g_{wt}; g_{Yt}; g_{H_t}; mm1_t; mm2_t; mm3_t; x_t \end{bmatrix}, \text{ and}$$

$$\mathbf{u}_t \equiv [emm1_t; emm2_t; emm3_t; ex_t].$$

Appendix III: Solution, Simulation and Estimation Methods

The DSGE model can be summarised by a nonlinear system of equilibrium conditions described in Appendix II. After transformation, all the endogenous variables are stationary so that steady states exist. Denote the vector of stationarised endogenous variables as $\mathbf{y}_t$, the vector of exogenous shocks as $\mathbf{u}_t$, and the vector of structural parameter as $\boldsymbol{\theta}$. Hence, the *nonlinear dynamic* system can be written as:

$$\mathbf{E}_t\left[f\left(\mathbf{y}_{t-1},\mathbf{y}_t,\mathbf{y}_{t+1},\mathbf{u}_t;\boldsymbol{\theta}\right)\right]=0.$$

AIII.1. Model Solution

In steady state, we assume the shocks are all equal to zero, and the endogenous variables are equal to their steady state levels $\bar{\mathbf{y}}$. The system in steady state is:

$$f\left(\bar{\mathbf{y}};\boldsymbol{\theta}\right)=0.$$

Since the number of equations are equal to the number of unknowns ($\bar{\mathbf{y}}$), the nonlinear equation system can be numerically solved using the Newton algorithm. Once the steady state levels ($\bar{\mathbf{y}}$) are solved (in terms of $\boldsymbol{\theta}$), we can then log-linearise the original system around the steady state. This leads to a *linear dynamic* system:

$$\mathbf{AE}_t\left[\mathbf{y}_{t+1}\right]+\mathbf{By}_t+\mathbf{Cy}_{t-1}+\mathbf{Du}_t=0, \text{ where } \mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D} \text{ are functions of} \boldsymbol{\theta}.$$

The system can be solved using perturbation method as described in Schmitt-Grohe and Uribe (2004). Note that to be solvable, the linearised system of difference equation needs to satisfy the Blanchard-Khan condition (1980), i.e. the number of forward-looking variables is equal to the number of unstable eigenvalues. The solution here is a first order approximation to the true solution to the original nonlinear system, which is actually a VAR process with restrictions:

$$\mathbf{y}_t=\bar{\mathbf{y}}+\mathbf{g}_{\mathbf{y}}\left(\mathbf{y}_{t-1}-\bar{\mathbf{y}}\right)+\mathbf{g}_{\mathbf{u}}\mathbf{u}_t, \text{ where } \mathbf{g}_{\mathbf{y}} \text{ and } \mathbf{g}_{\mathbf{u}} \text{ are functions of } \boldsymbol{\theta}.$$

AIII.2. Model Simulation

The solution can be used to simulate the moments of the endogenous variables and to track them, either by theoretical derivation based on the probability distributions of $\mathbf{u}_t$ or by Monte Carlo simulation. In this paper, we adopt the former procedure, so:

$$\mathbf{Var}\left[\mathbf{y}_t\right]=\mathbf{g}_{\mathbf{u}}\mathbf{Var}\left[\mathbf{u}_t\right]\mathbf{g}_{\mathbf{u}}'.$$

Based on this above equation, we can obtain the conditional variance-covariance matrix of the endogenous variables. We can also conduct a variance decomposition to analyse the contribution of each shock to the variance of each endogenous variable.

The impulse response functions can be derived based on the solution by rewriting the VAR into a VMA process:

$$\left(\mathbf{I}-\mathbf{g}_{\mathrm{y}}L\right)\mathbf{y}_t=\left(\mathbf{I}-\mathbf{g}_{\mathrm{y}}\right)\overline{\mathbf{y}}+\mathbf{g}_{\mathbf{u}}\mathbf{u}_t\text{, where } L \text{ is lag operator;}$$

$$\mathbf{y}_t=\overline{\mathbf{y}}+\left(\mathbf{I}-\mathbf{g}_{\mathrm{y}}L\right)^{-1}\mathbf{g}_{\mathbf{u}}\mathbf{u}_t\text{, if it is invertible;}$$

$$\mathbf{y}_t=\overline{\mathbf{y}}+\mathbf{\Phi}\left(L\right)\mathbf{u}_t\text{, where } \mathbf{\Phi}\left(L\right) \text{ is a polynomial of order } L;$$

$$\mathbf{y}_t=\overline{\mathbf{y}}+\sum_{i=0}^{\infty}\mathbf{\Phi}_i\mathbf{u}_{t-i}\text{, where } \mathbf{\Phi}_i \text{ is a function of } \boldsymbol{\theta} \text{ and } \overline{\mathbf{y}}.$$

The impulse response functions of a particular variable in $\mathbf{y}_t$ with respect to a particular shock in $\mathbf{u}_t$ can be extracted from the matrices $\mathbf{\Phi}_i$:

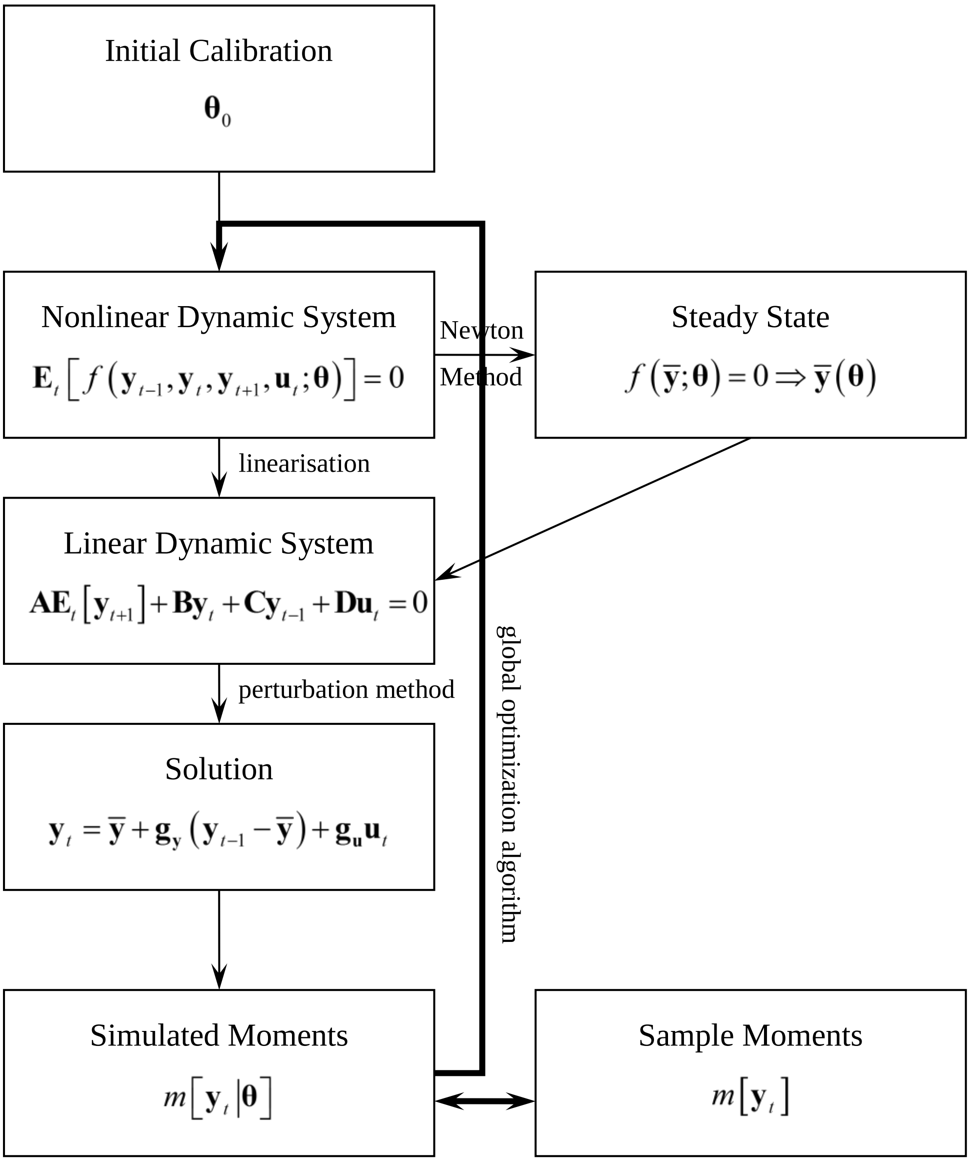
$$\mathbf{\Phi}_i=\frac{\partial\mathbf{y}_{t+i}}{\partial\mathbf{u}_t'}.$$

The solution and simulation procedures can be conveniently accomplished by the Dynare toolbox in Matlab or Octave environments.

AIII.3. Parameter Estimation

Assume the true value of the parameters ($\boldsymbol{\theta}$) is unknown, but we know an approximate value $\boldsymbol{\theta}_0$ as well as the upper bound $\overline{\boldsymbol{\theta}}$ and lower bound $\underline{\boldsymbol{\theta}}$. This initial value $\boldsymbol{\theta}_0$ is chosen to ensure the model has a valid start to converge. A popular way of judging the performance of a model is to check its ability to match the moments observed in the data. We have (stationarised) historical data of four series, i.e. population growth, birth rate, death rate and wage growth. It is easy to work out the sample mean and the sample variance of the four series. In addition, we also know that the marriage age is around 25 years during the sample period. To maximise the ability of the model to match these sample moments, we implement an optimization procedure based on a pattern search algorithm. The objective function is the gap—the sum of squared normalised differences—between model-simulated moments and the sample moments, while the choice variable is $\boldsymbol{\theta}$ and the constraints are the two bounds. The algorithm then searches the best $\widehat{\boldsymbol{\theta}}$ over the whole parameter space to minimise the objective function. For further references, see the Matlab manual on the global optimisation toolbox.

We use this global optimization algorithm, because the objective function cannot be differentiated—the simulated moments depend on the solution of the model which in turn depends on the *numerical* solution of steady state $\bar{\mathbf{y}}$. The following flow chart illustrates the links among solution, simulation and estimation procedures.



This estimation procedure is termed as optimal calibration, indirect estimation or simulated method of moments in DSGE literature.

AIII.4. Shock Estimation

In the subsection 4.1, we estimate the magnitude of mortality shocks and productivity shock to match the model-simulated series with the observed series. However, a key obstacle is that the real wage has longer observations than population, birth rate and death rate. Conventionally, one could sacrifice some observations for real wage so that all the variables have the common length of observations, so that we can apply maximum likelihood with Kalman filter dealing with the unobservable variables. However, there are two shortcomings of this approach. First, we will have to lose a lot of information, which is very important for the long history in which some variables are unobserved. Second, we will not be able to derive the endogenous variables or shocks occurring during that part of history. Third, maximum likelihood usually requires some subjective and restrictive assumption on the probability distribution function form, which might not always be standard.

Instead, we adopt a generalised moment-based approach, which addresses all the three shortcomings. Based on the estimated parameters ($\widehat{\boldsymbol{\theta}}$) as described in the previous subsection, we can express the final form of the model in numerical form, i.e. we know exactly what values $\bar{\mathbf{y}}, \mathbf{g}_{\mathbf{y}}, \mathbf{g}_{\mathbf{u}}$ take, because they depend on $\widehat{\boldsymbol{\theta}}$:

$$\mathbf{y}_t = \bar{\mathbf{y}} + \mathbf{g}_{\mathbf{y}} \left(\mathbf{y}_{t-1} - \bar{\mathbf{y}}\right) + \mathbf{g}_{\mathbf{u}} \mathbf{u}_t \, .$$

Some of the endogenous variables in $\mathbf{y}_t$ are linked with observable variables (population, birth rate, death rate and real wage), but the rest are not (e.g. human capital, marriage age, etc.). We assume the system is in steady state initially, i.e. we set $\mathbf{y}_0 = \bar{\mathbf{y}}$, then we will be able to express $\mathbf{y}_1$ according to the estimated final form:

$$\mathbf{y}_1^{(s)} = \bar{\mathbf{y}} + \mathbf{g}_{\mathbf{y}} \left(\mathbf{y}_0 - \bar{\mathbf{y}}\right) + \mathbf{g}_{\mathbf{u}} \mathbf{u}_1 = \bar{\mathbf{y}} + \mathbf{g}_{\mathbf{u}} \mathbf{u}_1 \, .$$

Note that we do not yet know the values of the shocks $\mathbf{u}_1$, so $\mathbf{y}_1^{(s)}$ is a function of $\mathbf{u}_1$. Some endogenous variables in $\mathbf{y}_1^{(s)}$ are observable, but some are not. We collect all the observable elements of $\mathbf{y}_1^{(s)}$ in a vector $\mathbf{z}_1^{(s)}$ (with population and wage transformed to levels), which is also a function of $\mathbf{u}_1$. We then carry on to period 2 to get $\mathbf{y}_2^{(s)}$and $\mathbf{z}_2^{(s)}$, and so on until the last observed period T. Again, all these $\mathbf{y}_t^{(s)}$ are functions of current and past shocks $\mathbf{u} \equiv \mathbf{u}_1, \dots, \mathbf{u}_T$. To summarise, we have collected T vectors, $\mathbf{z}_1^{(s)}, \mathbf{z}_2^{(s)}, \dots, \mathbf{z}_T^{(s)}$, in contrast to their observed counterpart in data, $\mathbf{z}_1^{(a)}, \mathbf{z}_2^{(a)}, \dots, \mathbf{z}_T^{(a)}$, which might have different lengths due to unbalanced data or missing observations.

Next, we stack the simulated data $\mathbf{z}^{(s)}(\mathbf{u}) \equiv \left[\mathbf{z}_1^{(s)}; \mathbf{z}_2^{(s)}; \dots; \mathbf{z}_T^{(s)}\right]$ and observed actual data $\mathbf{z}^{(a)} \equiv \left[\mathbf{z}_1^{(a)}; \mathbf{z}_2^{(a)}; \dots; \mathbf{z}_T^{(a)}\right]$. The objective function can be defined as a quadratic loss function (squared prediction errors) with a weighting matrix $\mathbf{W}$ to reflect the difference in magnitude, and the shocks are such that this objective function is minimised:

$$\mathbf{u} = \arg\min_{\mathbf{u}} L \equiv \left(\mathbf{z}^{(s)}\left(\mathbf{u}\right) - \mathbf{z}^{(a)}\right)' \mathbf{W} \left(\mathbf{z}^{(s)}\left(\mathbf{u}\right) - \mathbf{z}^{(a)}\right).$$

More generally, we can also put in some theoretical objective restrictions in the loss function L. For example, we theoretically assume the four shocks are uncorrelated. Another six elements (i.e. the six correlation coefficients) can be added in the objective function when we estimate the shocks.

AIII.5. An Alternative Method

One alternative approach to estimating the model and the shocks is maximum likelihood with Kalman filter, which can also handle unobservables and missing data (for example, see Peng and Aston, 2011). We therefore also used maximum likelihood to estimate the model and the shocks (results are available on request). The likelihood function is the sum of the 15 likelihood functions—because we have 15 distinct ways of slicing the annual data into 15-year frequency. This is similar to the way we make use of annual data in the simulated method of moments. The Kalman filter is used to estimate the unobservables and to construct the likelihood functions. The estimated parameters are chosen to maximise the composite likelihood function.

This method makes even more efficient use of the annual data because it maximises the probability of observing all the data points. However, there are two drawbacks of this distribution-based estimation method compared to the moments-based estimation method we finally adopted. First, maximum likelihood estimation requires more assumptions on distribution and is more sensitive to mis-specification, while the simulated method of moments is robust to these issues (see Ruge-Murcia 2007 for a complete discussion). Second, the Kalman filter which is part of the estimation procedure requires the model to be a Linear-Quadratic-Gaussian system, which accumulates the chances of errors. Moreover, the Kalman filter is designed for stationary systems, so the implied non-stationary series might be much further away from the actual series. In contrast, the method we adopted does not rely on any distributional assumption, and it performs better in matching the non-stationary series. Various robustness checks in section 5 also support the simulated method of moments.