Online Appendix

to

Understanding per-capita income growth
in preindustrial Europe

by

Nils-Petter Lagerlöf

Abstract: This Online Appendix provides some additional material related to the manuscript "Understanding per-capita income growth in preindustrial Europe" (Lagerlöf 2018), henceforth referred to as "the paper." It is organized into several subsections. Section 1 proposes some micro foundations of the model that was simulated in the paper. Section 2 walks through the calibration in more detail. Section 3 extends the time horizon of the simulations to 2010, and compare the results to modern data. Section 4 compares the birth and death rates from our benchmark simulation to data from England and Sweden. Section 5 incorporates marriage into the model, and illustrates how the results change in a couple of simulations. Section 6 walks through some of the robustness checks discussed in the paper in more detail. Section 7 explores the stability properties of a linearized dynamical system. Section 8 derives analytical expressions for total output and output per capita in steady state when fertility has a wage elasticity of one. Finally, Section 9 compares simulated population trends to data.

1 Micro foundations for the model

This section proposes a model, similar to the simplified one in Section 4.1 in the paper, that allows us to derive the fertility behavior in (9) in the paper, as well as an expression describing consumption of all agents in the economy.

The model builds on a few, more or less plausible, assumptions. First, agents care about births, rather than the stock of born (and living) children. Second, there are no means of saving. These first two assumptions allow us to abstract from intertemporal allocation decisions over long time horizons.

Third, children carry only goods costs, and no time costs. This is probably not important, but simplifies the analysis, since working agents can then be assumed to supply labor inelastically.

Fourth, agents derive utility from transfers to dependent agents in the economy, rather than, e.g., transfers to their own children, or their own children's consumption. This allows

orphans to consume; recall that agents face mortality risk through their whole lives, so some dependent children will have no living parents.

Finally, some remaining assumptions, e.g. log utility, are quite standard.

One can discuss how plausible each of these assumptions is, but together they do generate the fertility behavior described by (9) in the paper, and may at any rate serve as a good starting point.

Consider thus the following set-up. In each model period t, agents in life periods $j \in \{\underline{B}, ..., \overline{B}\}$ allocate the wage $w_{j,t}$ between own consumption in the same period, $c_{j,t}$, total transfers to non-active agents, $\tau_{j,t}$, and spending on conceiving $n_{j,t}$ children, at total cost $q_j n_{j,t}^{1/\delta}$. As in Section 4.1 in the paper, δ is the inverse of the elasticity of the child-cost function, and q_j is a parameter that shifts the cost function and is here age dependent. This gives the budget constraint

$$c_{j,t} = w_{j,t} - \tau_{j,t} - q_j n_{j,t}^{\frac{1}{\delta}}. \tag{A.1}$$

For simplicity, utility is defined over the economically and reproductively active phase of life only; all other agents consume their incomes or transfers received, thus making no choices. In particular, agents who work but have no children ($j \in \{\overline{B}+1, ..., R\}$) consume all their income and do not make transfers. This can easily be relaxed, by defining their utility in terms of transfers and consumption only, but not fertility.

An agent who becomes active in period t has utility

$$U_t = \sum_{j=\underline{B}}^{\overline{B}} \beta^{j-\underline{B}} u_{j,t+(j-\underline{B})}, \tag{A.2}$$

where $u_{j,t}$ is direct utility (j being the period of life, t the model period), and $\beta \in (0,1)$ is a discount factor. The agents care about their own consumption, transfers, and the number of children conceived, with weights $\eta > 0$ and $\widetilde{\gamma} > 0$ on the latter two, where $\eta + \widetilde{\gamma} < 1$. The period-utility can thus be written

$$u_{j,t} = (1 - \eta - \widetilde{\gamma}) \ln(c_{j,t}) + \eta \ln(\tau_{j,t}) + \widetilde{\gamma} \ln(n_{j,t}). \tag{A.3}$$

Because agents have no means of allocating resources intertemporally, they simply maximize the direct utility in every period, as given by (A.3), subject to the budget constraint in (A.1). The first-order conditions with respect to $n_{j,t}$ and $\tau_{j,t}$ state that

$$(1 - \eta - \widetilde{\gamma}) \left[w_{j,t} - \tau_{j,t} - q_j n_{j,t}^{\frac{1}{\delta}} \right]^{-1} \left(\frac{1}{\delta} \right) q_j n_{j,t}^{\frac{1}{\delta}-1} = \widetilde{\gamma} \left[n_{j,t} \right]^{-1}, \tag{A.4}$$

and

$$(1 - \eta - \widetilde{\gamma}) \left[w_{j,t} - \tau_{j,t} - q_j n_{j,t}^{\frac{1}{\delta}} \right]^{-1} = \eta \left[\tau_{j,t} \right]^{-1}. \tag{A.5}$$

These can be solved to give consumption as a constant fraction of income:

$$c_{j,t} = \left[\frac{1-\eta-\widetilde{\gamma}}{1-\widetilde{\gamma}(1-\delta)}\right] w_{j,t}. \tag{A.6}$$

Using (A.1), (A.4), and (A.6), spending on children also becomes a constant fraction of income, $q_j n_{j,t}^{1/\delta} = \{\delta\widetilde{\gamma}/\left[1-\widetilde{\gamma}(1-\delta)\right]\} w_{j,t}$, so optimal fertility can be written as in (9), where

$$\gamma_j = \left(\frac{\delta\widetilde{\gamma}}{q_j\left[1-\widetilde{\gamma}(1-\delta)\right]}\right)^{\delta}. \tag{A.7}$$

for $j \in \{\underline{B}, ..., \overline{B}\}$ and otherwise $\gamma_j = 0$.

Moreover, using (A.4) and (A.5), transfers become a constant fraction of income, $\tau_{j,t} = \{\eta/\left[1-\widetilde{\gamma}(1-\delta)\right]\} w_{j,t}$, so that the factor $\eta/\left[1-\widetilde{\gamma}(1-\delta)\right]$ effectively functions as a voluntary tax rate.

For internal consistency of the model, we may also specify consumption of economically non-active agents. Dependent children simply rely on transfers from the active generations. Because the recipients of these transfers do not reproduce, the transfers play no role for the results. To model consumption of retired agents we can make several assumptions. For example, they could be partial recipients of the same transfers that children receive. Here, we let retired agents be owners of land from which they earn rental income. This interpretation means that agents receive an amount of land when they stop working as inheritance from landowners who die. (This includes all agents in period T of life, but also some younger landowning agents, since agents die throughout the life cycle.) Like dependent children, old agents are also reproductively non-active so their incomes and consumption play no role for the results.

More compactly, this can be written as:

$$c_{j,t} = \begin{cases} c_{j,t}^{\text{dependent}} & \text{if } j \in \{1, ..., \underline{B}-1\}, \\ c_{j,t}^{\text{reproductive}} & \text{if } j \in \{\underline{B}, ..., \overline{B}\}, \\ w_{j,t} & \text{if } j \in \{\overline{B}+1, ..., R\}, \\ c_{j,t}^{\text{retired}} & \text{if } j \in \{R+1, ..., T\}, \end{cases} \tag{A.8}$$

where the notation is self-explanatory:

$$c_{j,t}^{\text{dependent}} = \frac{\sum_{j=\underline{B}}^{\overline{B}} \tau_{j,t} P_{j,t}}{\sum_{j=1}^{\underline{B}-1} P_{j,t}} = \left[\frac{\eta}{1-\widetilde{\gamma}(1-\delta)}\right] \frac{\sum_{j=\underline{B}}^{\overline{B}} w_{j,t} P_{j,t}}{\sum_{j=1}^{\underline{B}-1} P_{j,t}}, \tag{A.9}$$

$$c_{j,t}^{\text{reproductive}} = \left[\frac{1-\eta-\widetilde{\gamma}}{1-\widetilde{\gamma}(1-\delta)}\right] w_{j,t}, \tag{A.10}$$

$$c_{j,t}^{\text{retired}} = \frac{Y_t - \sum_{j=\underline{B}}^{R} w_{j,t} P_{j,t}}{\sum_{j=R+1}^{T} P_{j,t}} = \frac{\alpha Y_t}{\sum_{j=R+1}^{T} P_{j,t}}. \tag{A.11}$$

The last of these uses (13) and (14) in the paper, to note that total payments to workers can be written $\sum_{j=\underline{B}}^{R} w_{j,t} P_{j,t} = (1-\alpha) Y_t$.

2 Data and calibration

2.1 Data on wages, fertility, and age-specific mortality

The age-specific fertility data in Figure 2 in the paper are from Statistics Sweden (1969, Table 34). They refer to children born per woman in seven different five-year age brackets, and are reported by decade, the first decade being 1751-1760. We use averages from 1751 to 1800. To translate the per-woman birth rates in the data to births per agent in our one-sex model we divide by two.

The youngest age bracket refers to 15-19-year-olds, and the oldest to 45-49-year-olds. We need to be careful to distinguish both an agent's age from his/her period of life, and conception from birth. The data refer to the mother's age when giving birth, while the period-of-life subindex (j in $n_{j,t}$) refers to the period of life in which the child is conceived. Because a child is born in the period after it is conceived, the model and the data coincide numerically: for example, children conceived in period of life $j = 15$ are born when the mother is of age 15 (period of life 16).

The age-specific wages in Figure 2 in the paper are from Statistics Sweden (1942, Table 11). These are annual wages in 1940 for male rural workers (*lantarbetare*), who were paid some of their wages in kind (*arbetare i kost*), and the wage includes the in-kind payment. 1940 is the earliest year available in these data, and coincides with the Second World War, a period when Sweden was largely cut off from trade, which may partly help mimic preindustrial conditions.

The wage rates are defined for slightly different age brackets than for fertility. The youngest and oldest brackets are open ended, labelled 18 and below, and 65 and above, respectively. We assume that the youngest and oldest workers are of ages 14 and 69, respectively, corresponding to 15 and 70 in terms of periods of life; recall that we set $\underline{B} = 15$ and $R = 70$. The wage of the youngest working age group (in life periods 15-19) is normalized to one.

The age-specific survival rates in Table 2 in the paper are from Statistics Sweden (1969, Table 40) and refer to the Swedish total population 1751-1800. The rates are defined for different age brackets, the youngest of age 0 (the first year of life), the second youngest of ages 1-2 (second and third years of life), and so on. The last age bracket is 80 and above, which is here translated to ages 80-89, meaning that the 90th period of life is the last; recall that we set $T = 90$.

2.2 Setting γ_j

Let w_i^{data} be the wage in the data referring to an agent in the ith period of life (of age $i-1$), and let n_i^{data} be the corresponding number from the fertility data (the number of children conceived per woman in the ith period of life, divided by two). The age-dependent fertility parameters are first set such that, for $j \in \{15, ..., 19\}$, γ_j is given by

$$\gamma_j = \frac{\sum_{i=15}^{19} n_i^{\text{data}}}{\sum_{i=15}^{19} \left(w_i^{\text{data}}\right)^{\delta}}, \tag{A.12}$$

and analogously for the other age groups. This means that if the model can match the data on wages it will also match the data on fertility.

As can be understood intuitively from the simplified model in Section 4.1 in the paper, the γ_j's are inversely related to steady-state wages. Fertility rates in the model adjust to whatever makes population growth proportional to productivity growth, which in turn determines the wage rates. Therefore, the γ_j's can be scaled uniformly to shift the age-wage profile. In the benchmark specification no such scaling is needed to match the simulated wages to data.

Values of γ_j for select age groups are shown in Table 3 in the paper.

2.3 Setting β_j and ρ

We set $\rho = .9$, close to perfect substitutability.

We let β_j increase by 1% per year of life from $\underline{B} = 15$ until period 32 (age 31), after which β_j decreases by 1% per year of life and drops to zero after period $R = 70$, all subject to $\sum_{j=\underline{B}}^{R} \beta_j = 1$. Values for select age groups are shown in Table 3 in the paper.

The period of life when β_j peaks is chosen to roughly coincide with the peak in the age-wage profile in the data. Note from Figure 2 in the paper that wages decline more slowly with age than they rise. This reflects that older cohorts are smaller, due to mortality throughout the life cycle, making older agents scarcer than young, and thus their marginal products higher; this is where imperfect substitutability ($\rho < 1$) plays a role.

2.4 Elasticities

Setting $\delta = .15$ translates directly to an elasticity of *conceived* children with respect to same-year wages of .15; see (9) in the paper. To compare this to data, we compute the elasticity of the Crude Birth Rate in a given year—the number of births over population—with respect to wages in the previous year. The CBR in year t equals the number newborns (agents who are in the first period of life), over the total population in the same period, excluding the

newborns themselves:

$$\text{CBR in year } t = \frac{P_{1,t}}{\sum_{i=2}^{T} P_{i,t}}. \tag{A.13}$$

The Crude Death Rate is computed as the total number of agents who die in a given year, divided by total population:

$$\text{CDR in year } t = \frac{\sum_{i=1}^{T}(1 - s_{i.t})P_{i,t}}{\sum_{i=1}^{T} P_{i,t}}. \tag{A.14}$$

Wages in year t are computed as the (unweighted) average across all age groups of working age:

$$\text{Wage in year } t = \frac{\sum_{i=\underline{B}}^{R} w_{i,t}}{R + 1 - \underline{B}}. \tag{A.15}$$

We then simulate the model, under the benchmark parametrization, for 501 years (corresponding to 1300-1800), and 100 times. The result is a panel dataset with variation in wages, CBR, and CDR across different (artificial) economies and over time. We then run a number of fixed-effects regressions with the logarithms of CBR and CDR as dependent variables, and log wages at different lags as independent variables.

The results are shown in Table A.1 for log wages at different lags. The reported coefficient estimates represent elasticities since all variables are logged. Regardless of specification, the estimates fall around .13 and $-.2$ for CBR and CDR, respectively. As discussed is Section 5.1 in the paper, this is similar to the empirical literature (e.g., Lagerlöf 2015, Klemp and Møller 2016).

Note that we have not imposed anything on the model to produce the measured effects from variables lagged two periods back. Rather, these effects are driven by the serially correlated productivity shocks and the way the shocks are propagated through the model.

The CDR elasticity is guided by κ, which was set to .01, and measures the elasticity of the survival rate from starvation, s_t^y, with respect downward changes in per-capita incomes; see (20) in the paper. The estimated CDR elasticity is much larger in absolute terms than the survival elasticity, because the survival rates are close to one: small percentage changes in survival rates correspond to large percentage changes in mortality rates.

3 Extended time horizon

The analysis in the paper was based on GDP per capita data from 1300 to 1800. Figure A.1 shows the time paths of GDP per capita from 1300 all the way to 2010, for the same five European countries considered in the paper. Before 1800, the source is the same as in the paper (Fouquet and Broadberry 2015). For 1801-2010, we rely on Bolt and van

Zanden (2014), who update estimates initially compiled by Angus Maddison (2003).[11] These two sources report per-capita GDP data in the same units (1990 Geary-Khamis dollars). However, for some of the countries there are discrepancies for overlapping years; they are particularly large for Sweden, where Fouquet and Broadberry (2015) report a 33% higher number for 1800 than Bolt and van Zanden (2014). For some years and countries, data are also missing in the post-1800 period, in particular for Portugal 1801-1864. For these reasons, we may want to take the drop in levels right after 1800 with a grain of salt, but we can arguably put more trust in the numbers reported for more recent years.

Figure A.1 shows that a simultaneous take-off for all five countries some time around or after 1800. This take-off is larger in magnitude than the expansions and contractions which preceded it in each of the individual countries. In other words, the period up until 1800 looks stagnant in comparison to what follows.

We can also assess how well, or poorly, our Malthusian model performs when compared to data over the whole period 1300-2010. To do this, we first assume that μ_t (the expected growth rate in land productivity, A_t) stays constant at 1.25% per year from 1800 and onwards. All other parameter values are kept as in the benchmark setting. We then simulate the model for 1000 economies and 710 years (210 more years than the benchmark simulation). For each year, we calculate the median and the 5th and 95th percentiles of GDP per capita across the 1000 simulated economies, which can be compared to the country data in Figure A.1. We normalize both the empirical and simulated log per-capita GDP series to equal zero when averaging over time, and across either the five countries, or the 1000 simulated economies, respectively. Note that this shifts the pre-1800 paths down compared to those in Figure 5 in the paper.

The results are shown in Figure A.2, which corresponds to Figure 5 in the paper, but here considers the full period 1300-2010. As seen, toward the end of the sample period, the paths of GDP per capita for all five countries overshoot the 95th percentile from the simulation. They also dip below the 5th percentile shortly after 1800, which may be partly due to the discrepancies in the data sources discussed above.

The simulated per-capita GDP levels do trend upwards several centuries after 1800, but not enough to explain the data. If we were to run the model a few more centuries, the simulated GDP per capita levels would stabilize at a non-growing plateau, associated with the productivity growth rate of 1.25% per year. This level would be higher than in 1300, when productivity growth was zero. Of course, the associated simulated population levels would also continue to grow at a rate of 1.25% per year on average, which would not match the data either.

Table A.2 shows the fraction country-years falling outside the 5th and 95th percentiles

[11] See www.ggdc.net/maddison/maddison-project/home.htm

of the simulated data. For the entire period 1300-2010, about 17% of the country-years fell below the 5th percentile and almost 10% above the 95th percentile. These numbers are not too extreme in and of themselves. What stands out is the time profile. Almost all of the country-years above the 95th percentile happened after 1800. For 1801-2010, 28% of the country-years fell above the 95th percentile, compared to 0.5% in the pre-1800 era.

Among the five countries, Italy has the smallest fraction years above the 95th percentile after 1800, namely 22%. The corresponding numbers for the other four countries range from 24% (the Netherlands) to 34% (England). Only about 6.6% of the 1000 simulated economies fell above the 95th percentile in 22% of the years, or more, after 1800. That is, under the null hypothesis that our model captures the true data generating process, the probability of drawing one economy as extreme as Italy after 1800 is about 6.6%. This may not seem too rare. However, the probability of drawing *five* countries at least as extreme as Italy by this measure is $0.066^5 \approx 1.25 \times 10^{-6}$, or about 1.25 in a million.

The fraction country-years below the 5th percentile is high both before and after 1800, although the post-1800 numbers might be partly due to the artificial drop in 1801. Recall that, for years where they overlap, our post-1800 source (Bolt and van Zanden 2014) tends to report lower GDP per-capita numbers than our pre-1800 source (Fouquet and Broadberry 2015).

4 Data on Crude Birth and Death Rates

This section compares the simulation results to data on Crude Birth and Death Rates (CBR and CDR) from England and Sweden. The data sources are largely the same as those used in the empirical literature discussed in Section 2.4 above. The English data are from Wrigley et al. (1997, Table A9.1, p. 614), and are available on five-year intervals starting in 1541. This gives us 52 years (from 1541 to 1796) that overlap with the simulation period 1300-1800. The Swedish data can be found in Statistics Sweden (1969, Table 28, pp. 91-93), and are available annually from 1749, thus resulting in equally many years (52) that overlap with our simulations, from 1749 to 1800.[12]

As is standard, the sources define CBR and CDR as total birth and deaths, respectively, per 1000 population. As in Section 2.4, we here normalize them to per-capita terms. For example, a CBR of 9.4 per 1000 is here reported as 0.0094.

Figure A.3 compares the time paths of log CBR in the English and Swedish data to the

[12] The English data are updates of those originally published by Wrigley and Schoefield (1981, Table A3.1, pp. 528-529). Like the Swedish data, these can be accessed through the online publication Our World in Data sponsored by Oxford University (Roser and Ortiz-Ospina 2017); see OurWorldInData.org, Section II.2 under the tab World Population Growth.

mean, and 5th and 95th percentiles, of log CBR in our simulations. We show the period from 1541, since we do not have any earlier demographic data. The main insight is that the Swedish and English data are similar both to each other, and to the mean across the 1000 simulated economies. The mean also shows a mild upward trend, reflecting that income levels are rising due to accelerating growth rates in land productivity; cf. Figure 5 in the paper.

We also note that log CBR levels fluctuate a great deal, and often fall outside the 5th and 95th percentiles from the simulations. For England, 17% (9/52) of the years exceed the 95th percentile, and equally many fall below the 5th percentile; the corresponding numbers for Sweden are approximately 17% and just below 10%. In other words, fertility seems to fluctuate less in the model than in the data. But these differences are by no means extreme. Among the 1000 simulated economies, 8.4% had CBR levels above the 95th percentile as often as England did (i.e., in 17% of the years, or more), and 10.4% had CBR levels below the 5th percentile as often England (also in 17% of the years, or more).

Moreover, if we consider possible measurement errors the deviations between model and data may not seem too large, especially since we did not try to directly match these particular data.

Figure A.4 presents the corresponding comparison of CDR data to the mean, and 5th and 95th percentiles, of the simulated equivalent. Here the pattern is the reverse compared to the CBR data: the fluctuations are smaller in the data than in the simulations. Log CDR levels would be expected to fall outside the 5th and 95th percentiles in about 10% of the years, and the data points for Sweden and England never do. But again, this result is not really extreme either: as many as 7.8% of the simulated economies never fall outside these boundaries. Similar to the CBR data, the model does really well in terms of mean levels.

There are some simple mechanical adjustments that one can make to the model to allow it to better match these patterns. For example, one can add a random factor into the equation determining fertility [see (9) in the paper], thus allowing it to vary independently of wage fluctuations. One can also change the parameters δ and κ, which determine how sensitive the fertility and death rates are to shocks to wages, but at the cost of making the fit of the per-capita GDP fluctuations slightly worse. At any rate, the match is not perfect, but not completely off either.

In Section 5 below we consider another extension that can feed more fertility fluctuations into the model, namely letting marriage rates depend on wages.

5 Marriage

This section extends the benchmark model to allow for marriage. Since we are working with a one-sex model, this requires some imaginative interpretations. When an agent in our model "marries" we simply mean that (s)he enters a state in which the rate of fertility is higher, holding constant age and income. The Malthusian mechanism we seek to capture here is that the probability of a transition into marriage increases with income.

Recall from (16) in the paper that the population dynamics for all ages $j > 1$ can be written

$$P_{j+1,t+1} = s_{j,t} P_{j,t}. \tag{A.16}$$

Let $M_{j,t}$ be the total number of agents in period j of life who are married in period t, and let $r_{j,t}$ be the fraction of non-married agents in period j of life who become married if they survive to the next period, i.e., the marriage rate. Abstracting from divorce, this implies

$$M_{j+1,t+1} = s_{j,t} M_{j,t} + r_{j,t} s_{j,t} (P_{j,t} - M_{j,t}). \tag{A.17}$$

Dividing (A.17) by (A.16), and letting $m_{j,t} = M_{j,t}/P_{j,t}$ be the fraction agents in period j of life who are married in period t, gives

$$m_{j+1,t+1} = m_{j,t} + r_{j,t}(1 - m_{j,t}). \tag{A.18}$$

To be able to simulate the model, we need to impose assumptions about how the marriage rate is determined. To that end, we let

$$r_{j,t} = r_j^{\text{age}} \widetilde{r}_{j,t}, \tag{A.19}$$

where r_j^{age} is an age-dependent factor (varying only across periods of life, j), and where $\widetilde{r}_{j,t}$ is a factor that depends on stochastic shocks and wages. Next let

$$r^* = \max_{j \in \{1,\ldots,T\}} r_j^{\text{age}}, \tag{A.20}$$

and

$$\widetilde{r}_{j,t} = \min \left\{ \frac{1}{r^*}, \eta_t w_{j,t}^{\xi} \right\}, \tag{A.21}$$

where η_t is drawn from a uniform distribution on $[0, 2]$ in each period t, and thus has mean one, and where (recall) $w_{j,t}$ is the (age and time dependent) wage rate, and $\xi > 0$ an elasticity parameter. It is easy to see from (A.19) to (A.21) that the marriage rate falls between zero and one, i.e., $r_{j,t} = r_j^{\text{age}} \widetilde{r}_{j,t} \in [0, 1]$.

The positive relationship between the marriage rate ($r_{j,t}$) and the wage ($w_{j,t}$) embeds into the model the Malthusian feature that a positive shock to land productivity and thus wages induce more marriage, and thus higher fertility.

Finally, we let fertility among married and non-married agents be $n_{j,t}^{\text{marr}}$ and $n_{j,t}^{\text{unmarr}}$, respectively. Analogously to (9) in the paper, we then let

$$n_{j,t}^{x} = \upsilon_j^x w_{j,t}^{\delta}, \tag{A.22}$$

where υ_j^x is a fertility parameter that varies with period of life, j, and marriage state, $x \in \{\text{marr,unmarr}\}$. While the levels of fertility at any given wage do vary with marriage status and age through υ_j^x, for all agents we assume the same elasticity of fertility with respect to wages as in the benchmark model, i.e., δ. The overall fertility rate is thus given by

$$n_{j,t} = \left[m_{j,t}\upsilon_j^{\text{marr}} + (1 - m_{j,t})\upsilon_j^{\text{unmarr}}\right] w_{j,t}^{\delta}. \tag{A.23}$$

Note that setting $m_{j,t} = 1$ and $\upsilon_j^{\text{marr}} = \gamma_j$ makes fertility in (A.23) identical to that in the setting without marriage; see (9) in the paper.

In principle, it is straightforward to simulate the model by simply using (A.18) to (A.23) to update the fraction agents married in each period. The challenge is to find parameter values that allow the model to match the data, as explained below.

5.1 Quantitative examination of model with marriage

5.1.1 Parameter values

The ambition here is merely to get a rough idea about how the predictions of this extended model differ from our benchmark setting without marriage. To that end, we keep all applicable parameters (and functional forms) as in our benchmark simulation, and then set the new parameters as follows.

First, we note that the mean of the stochastic part of the marriage rate ($\widetilde{r}_{j,t}$) is approximately one. This follows because η_t has mean one, and the wage rate ($w_{j,t}$) by normalization fluctuates around approximately one for the working (and marrying) ages; for ξ not too large, it then holds that $w_{j,t}^{\xi}$ is approximately one as well.

This means that we can set the age-dependent factor, r_j^{age}, to match the age-specific rates that we observe in the data. The earliest period for which we have age-specific marriage rates is around 1861-1870 (Statistics Sweden 1969, Table 31). We want to translate these numbers to the period 1751-1800 (which is the period that our age-specific fertility rates refer to in the benchmark model). To that end, we use the gap in the crude marriage rates for Sweden between the periods in 1861-1870 and 1751-1800, averaged across sexes, to scale the age-specific marriage rates. (The numbers are found in Statistics Sweden 1969, Table 30.) This implicitly assumes that the decline in marriage rates observed between the mid 1700's and the 1860's was the same across age groups, which is probably at best a decent approximation.

Second, we set the parameter that determines the elasticity of marriage rates with respect to wages, ξ, to the same as the corresponding parameter for fertility. That is, $\xi = \delta = .15$.

Finally, consider the parameters υ_j^x. These are set roughly analogously to how we set the values for γ_j in the setting without marriage; see (A.12). More precisely, we first set υ_j^x to match the age-specific fertility rates for married and unmarried women, respectively, as observed in Swedish data for 1751-1800 (Statistics Sweden 1969, Table 35). We then find that overall age profile of fertility from the simulations, as implied by these marriage rates and marriage-specific fertility rates, is higher than that found in the data for 1751-1800, and used in our benchmark calibration (Statistics Sweden 1969, Table 34). One reason could be the imprecise way in which we have imputed the marriage rates above; another could be that not all marriages in the data are first marriages. To better match the data, we therefore scale the values for υ_j^x downwards by a factor of .85. As shown below, this allows the model to match the age profile of fertility in Sweden 1751-1800, thus also bringing it closer to the benchmark calibration.

5.1.2 Simulation results

Age profiles The resulting age profiles in the simulations and the data are shown in Figure A.5. Similar to Figure 2 in the paper, the model generated profiles are averages over 1000 runs for the last 50 years of the simulation, corresponding to 1751-1800 in the data. The bottom two panels in Figure A.5 refer to the marriage rate ($r_{j,t}$) and the fraction agents being married ($m_{j,t}$). The marriage rate profile in the data is simply the variable r_j^{age}, imputed as described above; the fraction being married is then generated from (A.18), setting $r_{j,t} = r_j^{\text{age}}$. Not surprisingly, the model simulations can match these (imputed) data well, reflecting that the stochastic factor affecting the marriage rate ($\widetilde{r}_{j,t}$) averages roughly one in the simulations.

The somewhat worse fit of the fertility (and wage) profiles in the top panels can partly be understood from how we imputed marriage rates in Sweden 1751-1800 from marriage rates in the 1860's. In reality, agents probably married younger in the 1700's that the 1800's. As discussed, even with more detailed data, it would be difficult to calibrate the model, since agents have only one sex, and since there is no dissolution of marriages (e.g., through death of one spouse). That said, the match between model and data are arguably close enough to make at least some comparisons meaningful.

Regression results Table A.3 reports the results when regressing crude birth, death, and marriage rates—CBR, CDR, and CMR, all in logs and defined as in Section 2.4 (or equivalently for marriages)—on log wages of various time lags, using simulated data. These correspond to those in Table A.1, but here reporting results also for marriage rates. As seen, the estimated coefficients in the birth rate regressions are now larger. A shock to wages

has a direct positive effect on fertility, as in the benchmarks setting, but now also induces more marriage, which in turn raises fertility, since married agents have higher fertility than unmarried. Note that we keep the parameters guiding fertility and mortality (δ and κ, respectively) as in the benchmark setting, rather than recalibrating them.

A related observation in column (4) is that birth rates now show positive correlation with wages at a two-year lag, when controlling for current wages and wages at a one-year lag. This contrasts with the result in the benchmark setting without marriage. Here positive wage shocks have positive effects on fertility further into the future by raising current marriage rates.

Comparisons to data from England and Sweden Figures A.6-A.8 compare the mean and 5th and 95th percentiles of the simulated crude marriage, birth, and death rates to data from England and Sweden, using the same sources discussed in Section 4 above (Wrigley et al. 1997, Statistics Sweden 1969). Figures A.6 and A.7 correspond to Figures A.3 and A.4 in the benchmark setting. As seen in Figure A.6, this extension captures more of the variation in log CBR in both the Swedish and English data than the benchmark setting. This is intuitive, since we now have another source of volatility in the birth rate.

As seen in Figure A.8, the model does a worse job matching the marriage rate, log CMR. However, this seems to reflect the inherent problems discussed already associated with matching this model to data. First of all, to make the crude marriage rate in the data comparable to the simulations we have multiplied it by two, since each marriage implies that two agents enter into the state of being married. (We divide per-woman fertility rates in the data by two for similar reasons; see Section 2.1 above.) At the same time, different from the model, not all marriages in the data are first marriages, which tends to make the marriage rate as measured in the data mechanically greater than that in the model.

6 Robustness checks

This section provides more details about the robustness checks in Section 5.4 of the paper.

6.1 Only temporary shocks

The first exercise was to close down all permanent shocks by setting $\sigma_A = 0$, and then adjust the standard deviation in the temporary shocks by setting $\sigma_X = .35$. This serves to position the distribution of simulated standard deviations somewhere in the middle of the same measures for the five countries; see the top-right panel of Figure A.9. That is, we try to make the model-data match in standard deviations as good as possible, given the constraint that all shocks are temporary.

As seen in Figure A.9, there is now too little serial correlation in the model-generated distributions to match data, and the mean per-capita outcomes have far less dispersion than observed between the five countries. Permanent shocks to productivity are thus needed for the model to match the data. This is quite intuitive, since any standard Malthusian model would predict that changes to *levels* in productivity have no effect on steady-state living standards, but changes in productivity *growth rates* can, as we learned in Section 4.1 of the paper.

This insight in a sense resembles that from simulating real-business cycle models, where the model's internal propagation mechanisms are often so weak that all persistence comes from the shocks fed into the model (Gogley and Nason 1995). However, Figure A.9 also shows that there is some propagation force at play that generates persistence in our model, since even without any serially correlated shocks, the simulated distributions of measured serial correlation have most of the mass above zero. In other words, the Malthusian model's internal mechanisms do matter for the results.

6.2 Using 1560 as start year

Data coverage starts in different years for each country, the latest (Sweden) lacking data before 1560. As a result, the moments are calculated over different periods for the five different countries. The easiest way to correct for this is to consider only the period 1560-1800, for which all five countries have data. Figure A.10 illustrates the outcomes when doing this for the data and the model simulations; the associated 5th and 95th percentiles from the simulations are shown in Table 5 in the paper.

The means in the simulated series now show much larger dispersion, easily overlapping with the means observed in data. This is because they are calculated on smaller samples (over shorter time periods). Compared to Figure 6 in the paper, the moments for four of the countries shift slightly, since they too are calculated over a different time period. The model-data fit for standard deviations is similar to that in the benchmark case, with Italy as the marginal case. Also similar to the benchmark case, the model cannot quite account for some countries' low levels of serial correlation; that of the Netherlands is lowest of them all over this period.

However, the gaps between model and data are not huge. Note also that we here keep σ_X and σ_A at their benchmark values. The fit in Figure A.10 would be better if we were to recalibrate these when using 1560 as start year.

6.3 Higher elasticities

Two crucial parameters are those which guide the elasticities of fertility (δ) and mortality (κ). Table 5 in the paper shows how the 5th and 95th percentiles of the simulated moments change when doubling each of these over their benchmark values: δ from .15 to .3. and κ from .01 to .02. Either of these changes shrinks the dispersion in means of per-capita incomes across simulated economies, and lowers the levels overall of measured standard deviations. Intuitively, the larger is either elasticity, the smaller are the effects of (both permanent and temporary) productivity shocks, since population levels adjust faster when wages increase, making the effects on per-capita incomes less prolonged. There is little effect on the percentiles of the distributions of the serial correlation coefficients in Table 5 in the paper.

6.4 Higher disease mortality

The parameter ϕ guides the mortality component that we associate with the disease environment, s_t^d. A higher ϕ implies lower expected survival rates from disease, as well as higher variance in the survival rate. Table 5 in the paper shows what happens when we double ϕ from .01 to .02. Mean outcomes become more dispersed, and the distribution of standard deviations shifts up, and so do (at least marginally) the distributions of the serial correlation coefficients.

This is quite intuitive; feeding larger shocks into the model makes outcomes vary more. The increase in serial correlation can be understood from the way in which a standard Malthusian model reacts to a disease-induced drop in population levels. After an immediate increase in per-capita incomes, the economy converges gradually back to its Malthusian steady state, showing up as persistence time series. A higher ϕ implies larger such shocks, and thus more serial correlation.

6.5 Lower land share in production

The last row of Table 5 in the paper considers a reduction in the land share of output, α, from .4 to .2. This shrinks the variation in the shocks to land productivity, as can be seen from (10) in the paper. It also makes wages less sensitive to population changes (i.e., a given rise in the labor force reduces wages less), which in turn implies more persistent effects of productivity shocks. These effects tend to pull in opposite directions, in the sense that productivity shocks have smaller direct effects, but those effects last longer. The net effect here is that mean outcomes become less dispersed and standard deviation levels decline, but there is also a small increase in serial correlation.

The reason the increase in persistence does not show up much quantitatively seems to be that changes in α have small impact when δ is small. This can be understood from the simplified two-period Malthusian model in Section 4.1 of the paper, were we saw that the exponent on lagged wages in the dynamic wage equation equalled $1-\delta\alpha$; see (7).

7 The linearized dynamical system

The dynamics can be described in terms of the log population size of each of the T age groups, $\ln(P_{j,t})$, where $j \in \{1,...,T\}$. The joint evolution of these log population levels is guided by a T-dimensional dynamical system, which we can linearize around a non-growing steady state equilibrium (see, e.g., Azariadis 1993, Ch. 6).

This section derives the Jacobian matrix associated with that linearized dynamical system. To get there we need to impose some parametric assumptions. First we close down all shocks to productivity and mortality ($\sigma_A = \sigma_X = \phi = \kappa = 0$). Then we close down the deterministic growth in productivity ($\mu_t = 0$). This means productivity and mortality are constant over time, $A_t = X_t = 1$, and $s_{j,t} = s_j$, for all t.

Let Q_t be a $T \times 1$ column vector, each element containing the deviation of $\ln(P_{j,t})$ from its steady state level, $\ln\left(\overline{P}_j\right)$, where we use bars to denote steady state levels of all time dependent variables. That is,

$$Q_t = \begin{bmatrix} \ln(P_{1,t}) - \ln\left(\overline{P}_1\right) \\ \ln(P_{2,t}) - \ln\left(\overline{P}_2\right) \\ \vdots \\ \ln(P_{T-1,t}) - \ln\left(\overline{P}_{T-1}\right) \\ \ln(P_{T,t}) - \ln\left(\overline{P}_T\right) \end{bmatrix}. \tag{A.24}$$

Next, let Π be the $T \times T$ Jacobian matrix, whose element for each row j and each column k equals $\partial \ln(P_{j,t+1}) / \partial \ln(P_{k,t})$ evaluated in steady state, i.e.,

$$\Pi = \begin{bmatrix} \left.\frac{\partial \ln(P_{1,t+1})}{\partial \ln(P_{1,t})}\right|_{P_{j,t}=\overline{P}_j \forall j} & \cdots & \left.\frac{\partial \ln(P_{1,t+1})}{\partial \ln\left(P_{T,t}\right)}\right|_{P_{j,t}=\overline{P}_j \forall j} \\ \vdots & \vdots & \vdots \\ \left.\frac{\partial \ln\left(P_{T,t+1}\right)}{\partial \ln(P_{1,t})}\right|_{P_{j,t}=\overline{P}_j \forall j} & \cdots & \left.\frac{\partial \ln\left(P_{T,t+1}\right)}{\partial \ln\left(P_{T,t}\right)}\right|_{P_{j,t}=\overline{P}_j \forall j} \end{bmatrix} \tag{A.25}$$

Now the linearized dynamical system can be written

$$Q_{t+1} = \Pi Q_t. \tag{A.26}$$

The task is to find analytical expressions for the elements of Π , and then study the magnitude of its eigenvalues.

7.1 Finding the elements of the Jacobian matrix

7.1.1 Second and higher rows

For all rows $j > 1$, it holds that

$$P_{j+1,t+1} = s_j P_{j,t}, \tag{A.27}$$

where we use the population dynamics in (16) in the paper, setting $s_{j,t} = s_j$ for all t. Thus, for all j and k such that $j > 1$ and $k \neq j-1$, it holds that $\partial \ln(P_{j,t+1}) / \partial \ln(P_{k,t}) = 0$; and for all $j > 1$ and $k = j-1$, it holds that $\partial \ln(P_{j,t+1}) / \partial \ln(P_{k,t}) = 1$. We can thus write Π as

$$\Pi = \begin{bmatrix} \frac{\partial \ln(P_{1,t+1})}{\partial \ln(P_{1,t})}\Big|_{P_{j,t}=\overline{P}_j \forall j} & \cdots & \frac{\partial \ln(P_{1,t+1})}{\partial \ln\left(P_{T,t}\right)}\Big|_{P_{j,t}=\overline{P}_j \forall j} \\ 1 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}. \tag{A.28}$$

For example, the second row of Π contains 1 in the first column and zeros in all others. This follows since $\partial \ln(P_{2,t+1}) / \partial \ln(P_{1,t}) = 1$ and $\partial \ln(P_{2,t+1}) / \partial \ln(P_{k,t}) = 0$ for all $k > 1$. The bottom row contains zeros in all columns, except column $T-1$, which takes the value one.

Next we examine the first row of Π.

7.1.2 The first row

The difference equation in (A.27) can be solved to give

$$P_{j,t} = S_j P_{1,t-j+1}, \tag{A.29}$$

where

$$S_j = \prod_{i=1}^{j-1} s_i \tag{A.30}$$

is the (time invariant) probability that an agent survives to period j in life, and where $S_1 = 1$ and $P_{1,0}$ are given. Imposing steady state on (A.29) gives

$$\overline{P}_j = S_j \overline{P}_1, \tag{A.31}$$

where (recall) steady state levels are denoted by bars. Next we use the expression for the effective workforce, L_t, given by (13) in the paper. Evaluated in steady state, and using (A.31), we see that

$$\overline{L} = \overline{P}_1 \left(\sum_{j=1}^{T} \beta_j S_j^{\rho} \right)^{\frac{1}{\rho}}. \tag{A.32}$$

Note that we here sum from 1 to T, rather than from $\underline{B}$ to R, which does not change anything, since $\beta_j = 0$ for all $j \notin \{\underline{B}, ..., R\}$.

From (A.31) and (A.32) we also note for later use that

$$\beta_j \left(\frac{\overline{P}_j}{\overline{L}}\right)^{\rho} = \beta_j S_j^{\rho} \left(\frac{\overline{P}_1}{\overline{L}}\right)^{\rho} = \frac{\beta_j S_j^{\rho}}{\sum_{j=1}^{T} \beta_j S_j^{\rho}}. \tag{A.33}$$

The wage is given by (14) in the paper, here restated as

$$w_{j,t} = \frac{\partial Y_t}{\partial L_t}\frac{\partial L_t}{\partial P_{j,t}} = (1-\alpha)\frac{Y_t}{L_t}\beta_j \left(\frac{L_t}{P_{j,t}}\right)^{1-\rho}. \tag{A.34}$$

Recall that Y_t is total output, which is here given by

$$Y_t = (A_t X_t)^{\alpha} L_t^{1-\alpha} = L_t^{1-\alpha}, \tag{A.35}$$

since A_t and X_t are constant and equal to one. Fertility is given by (9) in the paper, i.e.,

$$n_{j,t} = \gamma_j w_{j,t}^{\delta}. \tag{A.36}$$

Using (A.34) to (A.36), we can write the number of agents conceived in period t by agents living in period j of life as follows:

$$\begin{aligned} n_{j,t}P_{j,t} &= \gamma_j (1-\alpha)^{\delta} \left(\tfrac{Y_t}{L_t}\right)^{\delta} \beta_j^{\delta} \left(\tfrac{L_t}{P_{j,t}}\right)^{\delta(1-\rho)} P_{j,t} \\ &= \gamma_j (1-\alpha)^{\delta} L_t^{\delta(1-\alpha-\rho)} \beta_j^{\delta} P_{j,t}^{1-\delta(1-\rho)}. \end{aligned} \tag{A.37}$$

Recalling (15) in the paper, this gives the total newborn population in period $t+1$:

$$\begin{aligned} P_{1,t+1} &= \textstyle\sum_{j=1}^{T} n_{j,t}P_{j,t} \\ &= \textstyle\sum_{j=1}^{T} \gamma_j (1-\alpha)^{\delta} L_t^{\delta(1-\alpha-\rho)} \beta_j^{\delta} P_{j,t}^{1-\delta(1-\rho)} \\ &= (1-\alpha)^{\delta} L_t^{\delta(1-\alpha-\rho)} \textstyle\sum_{j=1}^{T} \gamma_j \beta_j^{\delta} P_{j,t}^{1-\delta(1-\rho)}. \end{aligned} \tag{A.38}$$

where we recall that $\gamma_j = 0$ for all $j \notin \left\{\underline{B}, ..., \overline{B}\right\}$, allowing us to write the sum from 1 to T. Taking the natural logarithm of (A.38) gives

$$\ln(P_{1,t+1}) = \delta \ln(1-\alpha) + \delta(1-\alpha-\rho)\ln(L_t) + \ln\left[\sum_{j=1}^{T} \gamma_j \beta_j^{\delta} \overbrace{\exp\{[1-\delta(1-\rho)]\ln(P_{j,t})\}}^{=P_{j,t}^{1-\delta(1-\rho)}}\right]. \tag{A.39}$$

We thus find that

$$\frac{\partial \ln(P_{1,t+1})}{\partial \ln(P_{j,t})} = \delta(1-\alpha-\rho)\left[\frac{\partial \ln(L_t)}{\partial \ln(P_{j,t})}\right] + [1-\delta(1-\rho)]\left[\frac{\gamma_j \beta_j^{\delta} P_{j,t}^{1-\delta(1-\rho)}}{\sum_{j=1}^{T} \gamma_j \beta_j^{\delta} P_{j,t}^{1-\delta(1-\rho)}}\right]. \tag{A.40}$$

We first derive an expression for the first term on the right-hand side in (A.40). Using (13) in the paper again, we see that

$$\ln\left(L_t\right)=\left(\frac{1}{\rho}\right)\ln\left[\sum_{j=1}^{T}\beta_j P_{j,t}^{\rho}\right]=\left(\frac{1}{\rho}\right)\ln\left[\sum_{j=1}^{T}\beta_j\overbrace{\exp\{\rho\ln\left(P_{j,t}\right)\}}^{=P_{j,t}^{\rho}}\right], \tag{A.41}$$

and

$$\frac{\partial\ln\left(L_t\right)}{\partial\ln\left(P_{j,t}\right)}=\left(\frac{1}{\rho}\right)\underbrace{\left[\frac{1}{\sum_{j=1}^{T}\beta_j P_{j,t}^{\rho}}\right]}_{1/L_t^{\rho}}\rho\beta_j P_{j,t}^{\rho}=\beta_j\left(\frac{P_{j,t}^{\rho}}{L_t^{\rho}}\right)=\beta_j\left(\frac{P_{j,t}}{L_t}\right)^{\rho}. \tag{A.42}$$

Using (A.40) and (A.42) shows that

$$\frac{\partial\ln\left(P_{1,t+1}\right)}{\partial\ln\left(P_{j,t}\right)}=\delta(1-\alpha-\rho)\beta_j\left(\frac{P_{j,t}}{L_t}\right)^{\rho}+\left[1-\delta\left(1-\rho\right)\right]\left[\frac{\gamma_j\beta_j^{\delta}P_{j,t}^{1-\delta(1-\rho)}}{\sum_{j=1}^{T}\gamma_j\beta_j^{\delta}P_{j,t}^{1-\delta(1-\rho)}}\right]. \tag{A.43}$$

Now (A.43) and (A.33) give

$$\begin{aligned}\left.\frac{\partial\ln(P_{1,t+1})}{\partial\ln(P_{j,t})}\right|_{P_{j,t}=\overline{P}_j\forall j}&=\delta(1-\alpha-\rho)\beta_j\left(\frac{\overline{P}_j}{\overline{L}}\right)^{\rho}+\left[1-\delta\left(1-\rho\right)\right]\left[\frac{\gamma_j\beta_j^{\delta}\overline{P}_j^{1-\delta(1-\rho)}}{\sum_{j=1}^{T}\gamma_j\beta_j^{\delta}\overline{P}_j^{1-\delta(1-\rho)}}\right]\\&=\delta(1-\alpha-\rho)\left[\frac{\beta_j S_j^{\rho}}{\sum_{j=1}^{T}\beta_j S_j^{\rho}}\right]+\left[1-\delta\left(1-\rho\right)\right]\left[\frac{\gamma_j\beta_j^{\delta}S_j^{1-\delta(1-\rho)}}{\sum_{j=1}^{T}\gamma_j\beta_j^{\delta}S_j^{1-\delta(1-\rho)}}\right].\end{aligned} \tag{A.44}$$

Next define

$$\begin{aligned}X_j&=\frac{\beta_j S_j^{\rho}}{\sum_{j=1}^{T}\beta_j S_j^{\rho}}\in(0,1),\\Z_j&=\frac{\gamma_j\beta_j^{\delta}S_j^{1-\delta(1-\rho)}}{\sum_{j=1}^{T}\gamma_j\beta_j^{\delta}S_j^{1-\delta(1-\rho)}}\in(0,1).\end{aligned} \tag{A.45}$$

We can now write (A.44) as

$$\left.\frac{\partial\ln\left(P_{1,t+1}\right)}{\partial\ln\left(P_{j,t}\right)}\right|_{P_{j,t}=\overline{P}_j\forall j}=\delta(1-\alpha-\rho)X_j+\left[1-\delta\left(1-\rho\right)\right]Z_j. \tag{A.46}$$

Using (A.46) and (A.28), the matrix Π can thus be written

$$\Pi=\begin{bmatrix}\delta(1-\alpha-\rho)X_1 & \cdots & \delta(1-\alpha-\rho)X_T\\ +\left[1-\delta\left(1-\rho\right)\right]Z_1 & & +\left[1-\delta\left(1-\rho\right)\right]Z_T\\ 1 & \cdots & 0\\ \vdots & \vdots & \vdots\\ 0 & \cdots & 0\end{bmatrix}. \tag{A.47}$$

Using (A.45), computing values of the elements of the first row is straightforward, given values for δ, α, and ρ, as well as β_j, γ_j, and S_j for all $j\in\{1,...,T\}$. Once we have the

full T-dimensional matrix Π it is then straightforward to compute its T eigenvalues. Figure 8 in the paper plots these in the complex plane for the benchmark parameter values. All eigenvalues are located within the unit circle, implying that the linearized systems always converges to its steady state; see, e.g., Azariadis (1993, Theorem 6.2). That is, $P_{j,t}$ converges to $\overline{P}_j$ for each age group j.

At the same time, the imaginary parts are large, and most eigenvalues are relatively close to one in absolute terms, explaining why the dynamics also display a high degree of persistence.

The degree of persistence depends on δ, which governs the elasticity of fertility with respect to wages. Figure A.11 illustrates the time path of the log population gap of an economy, whose initial population is below its steady state level, for three different values of δ. The log gap is measured as $\ln(P_t/\overline{P})$, where $P_t = \sum_{j=1}^{T} P_{j,t}$ is total population in period t [see (17) in the paper], and $\overline{P} = \sum_{j=1}^{T} \overline{P}_j$ is the steady-state level of P_t. As seen, smaller δ implies slower transitions. Intuitively, when δ is small fertility is less sensitive to changes in wages, so it takes longer for the Malthusian forces to push the economy back to steady state. In the benchmark case, when $\delta = .15$, the gap closes by half after 356 years, compared 56 years when $\delta = 1$.

Figure A.12 makes a similar point, illustrating how the absolute value of the largest real eigenvalue of Π varies with δ over the interval $[0, 1]$. It is closer to one when δ is close zero, becoming equal to one when $\delta = 0$. The case when $\delta = 0$ implies that fertility is independent of wages, so that population levels never converge.

8 Analytical steady-state expressions with unit-elastic fertility

This section derives explicit analytical expressions for several variables in steady state, such as output per agent, which determines survival from starvation; recall (20) in the paper. First, we impose the same assumptions as in Section 7 above. That is, we close down all shocks to productivity and mortality, and deterministic growth in productivity ($\sigma_A = \sigma_X = \phi = \kappa = \mu_t = 0$), making productivity and mortality constant ($A_t = X_t = 1$, and $s_{j,t} = s_j$).

Here we also impose unit-elastic fertility, i.e., we let fertility be linear in wages ($\delta = 1$), which makes it easy to find nice closed-form solutions. With $\delta = 1$, fertility in (9) in the paper can be written

$$n_{j,t} = \gamma_j w_{j,t} = \gamma_j (1-\alpha) \frac{Y_t}{L_t} \beta_j \left(\frac{L_t}{P_{j,t}} \right)^{1-\rho}, \tag{A.48}$$

where the second equality uses the expression for wages in (14) in the paper.

Using (A.48), together with (13) and (15) in the paper, letting sums be written over age groups 1 to T (since $\beta_j = 0$ for all $j \notin \{\underline{B}, ..., R\}$ and $\gamma_j = n_{j,t} = 0$ for all $j \notin \{\underline{B}, ..., \overline{B}\}$), we can now write the new-born population in period $t+1$ as

$$\begin{aligned} P_{1,t+1} &= \sum_{j=1}^{T} \gamma_j (1-\alpha) \tfrac{Y_t}{L_t} \beta_j \left(\tfrac{L_t}{P_{j,t}}\right)^{1-\rho} P_{j,t} \\ &= (1-\alpha) Y_t \tfrac{\sum_{j=1}^{T} \gamma_j \beta_j P_{j,t}^{\rho}}{L_t^{\rho}} \\ &= (1-\alpha)\theta_t Y_t, \end{aligned} \tag{A.49}$$

where

$$\theta_t = \frac{\sum_{j=1}^{T} \gamma_j \beta_j P_{j,t}^{\rho}}{\sum_{j=1}^{T} \beta_j P_{j,t}^{\rho}}. \tag{A.50}$$

Next, using (A.50) and (A.29) gives

$$\theta_t = \frac{\sum_{j=1}^{T} \gamma_j \beta_j P_{j,t}^{\rho}}{\sum_{j=1}^{T} \beta_j P_{j,t}^{\rho}} = \frac{\sum_{j=1}^{T} \gamma_j \beta_j S_j^{\rho} P_{1,t-j+1}^{\rho}}{\sum_{j=1}^{T} \beta_j S_j^{\rho} P_{1,t-j+1}^{\rho}}, \tag{A.51}$$

where S_j is given by (A.30). In steady state, where $P_{1,t}$ is constant at $\overline{P}_1$, (A.51) can be written

$$\overline{\theta} = \frac{\overline{P}_1^{\rho} \sum_{j=1}^{T} \gamma_j \beta_j S_j^{\rho}}{\overline{P}_1^{\rho} \sum_{j=1}^{T} \beta_j S_j^{\rho}} = \frac{\sum_{j=1}^{T} \gamma_j \beta_j S_j^{\rho}}{\sum_{j=1}^{T} \beta_j S_j^{\rho}}. \tag{A.52}$$

Using (A.32), and imposing steady state on (A.35), it can be seen that

$$\overline{Y} = \overline{L}^{1-\alpha} = \overline{P}_1^{1-\alpha} \left[\sum_{j=1}^{T} \beta_j S_j^{\rho}\right]^{\frac{1-\alpha}{\rho}}. \tag{A.53}$$

Next, evaluating $P_{1,t+1} = (1-\alpha)\theta_t Y_t$ from (A.49) in steady state gives

$$\overline{P}_1 = (1-\alpha)\overline{\theta Y}. \tag{A.54}$$

Now, (A.53) and (A.54) give $\overline{Y} = [(1-\alpha)\overline{\theta Y}]^{1-\alpha} \left[\sum_{j=1}^{T} \beta_j S_j^{\rho}\right]^{(1-\alpha)/\rho}$, which can be solved for $\overline{Y}$ to give

$$\overline{Y} = \left[(1-\alpha)\overline{\theta} \left(\sum_{j=1}^{T} \beta_j S_j^{\rho}\right)^{\frac{1}{\rho}}\right]^{\frac{1-\alpha}{\alpha}}. \tag{A.55}$$

Now (A.55), together with (A.30) and (A.52), defines $\overline{Y}$ in terms of exogenous parameters, including the age-specific parameters γ_j, β_j, and s_j. Then the steady-state population of the new-born age group can be computed from (A.54), and those of the other age groups from (A.30) and (A.31).

Output per agent in any period is given by (17) and (18) in the paper, i.e., $y_t = Y_t/P_t = Y_t/\sum_{j=1}^{T} P_{j,t}$. Together with (A.54), and (A.55), steady-state output per agent can thus be written

$$\overline{y} = \frac{\overline{Y}}{\sum_{j=1}^{T} \overline{P}_j} = \frac{\overline{Y}}{\overline{P}_1 \sum_{j=1}^{T} S_j} = \frac{1}{(1-\alpha)\overline{\theta} \sum_{j=1}^{T} S_j}. \tag{A.56}$$

9 Population paths and artificial countries

The paper examined how well the model could match per-capita GDP trends 1300-1800 in five European countries. Here we look at population trends in the same five countries and over the same period. We rely on estimates by McEvedy and Jones (1978), which are available for all five countries over the whole period 1300-1800, and have often been used in earlier literature. These are little more than informed guesses, at least when going back as far 1300. Moreover, the countries for which population levels are reported are not precisely matched with per-capita GDP data. For example, GDP per capita data for Italy refer to its northern part, while the population data are for the whole of Italy.

With these caveats in mind, the exercise we undertake here is the following. For each country, we select the 25 simulated economies (out of 1000) whose time paths of log GDP per-capita most closely resemble those in the data from that country, as measured by the squared deviation summed across years. We then take the mean of log GDP per capita and log population among those 25 simulated economies, and refer to these paths as representing an "artificial" version of the country in question. This is done for each of the five countries. We then compare the artificial paths to those in the data. In Figure A.13, they are drawn as dashed red and solid black lines, respectively.

The population paths are the same as those in Figure 3 in the paper, normalized to equal one in 1300 (or zero in logs). As in Figure 5 in the paper, the GDP per capita data come from Fouquet and Broadberry (2015), and are normalized to equal zero when logged and then averaged over time and across countries.

As mentioned, both GDP per capita and population may be measured with error. One indication of this is that different sources sometimes report different GDP per capita numbers for the same year and country. Recall from Section 3 above that the numbers from Fouquet and Broadberry (2015) and Bolt and van Zanden (2014) differ by 33% for Sweden in 1800. Another example is the disagreements between Gregory Clark and Stephen Broadberry and coauthors about the reconstruction of the English GDP data; see Section 2 in the paper. If some of these measurement errors persist over time, it can matter a lot, because small differences in per-capita GDP levels in a Malthusian model tend to accumulate to large differences in population levels over long periods of time. To help address this, we also create artificial economies which mimic paths 20% above and below those in country data.

These are the blue dotted lines in Figure A.13. The number 20% is arbitrary but could give a sense of the magnitudes involved.

As seen from the panels in the top row of Figure A.13, the artificial (red dashed) paths for log GDP per capita closely follow the corresponding paths in the country data. This is not surprising, since they were constructed to do just that. Looking at the bottom row, we can see how the population levels of these artificial economies compare to those reported by McEvedy and Jones (1978). The results are somewhat mixed. The artificial paths mimic the data relatively well for the Netherlands and Sweden, overshoot the data for Italy, and undershoot it for England and Portugal.

Intuitively, if we take the simulated model and the per-capita GDP and population data seriously, then we should expect to see more variation in the population paths between countries than is reported by McEvedy and Jones (1978). Richer countries should see faster population growth than they do, and poorer countries should see slower growth. For example. Portugal should see a decline, which seems at odds with the data. However, the mismatch is not complete, in the sense that relatively small shifts in the GDP per capita paths (in the order of 20%) can account for much (if not all) of the variation in the population paths. If we take into account the possibility of measurement error in population levels as well, the mismatch is arguably not too bad, at least if we disregard Portugal.

A simple way to adjust the model to better match the population data is to assume that cross-country variation in living standards affect fertility and mortality differently in different countries. In particular, if the age-specific fertility parameter γ_j differs across countries, then they would have different per-capita income levels in steady state, even when otherwise identical. To compare such a model to data amounts to roughly the same thing as comparing the current model to demeaned country data, i.e., after normalizing the data so that all countries have the same mean when averaging across years.

Figure A.14 shows the results from an exercise identical to that in Figure A.13, but based on demeaned country data. That is, we first normalize log GDP per capita to average zero across all available years for each country. Then we create artificial economies, both of the demeaned country paths, and of the paths 20% above and below these. As seen in the bottom row of Figure A.14, the population paths of the artificial economies are much better aligned with the data.

Of course, a model where the countries have different steady state levels of GDP per capita would not really "explain" observed differences in mean per-capita incomes over time, as the benchmark setting could; cf. the top left histogram in Figure 6 in the paper. Those means are here zero for all countries by construction. However, the variance and serial correlation coefficients of the time series would be the same.

References

[1] Azariadis, C., 1993, Intertemporal Macroeconomics, Blackwell Publishers, Cambridge, Massachusetts.

[2] Bolt, J., and J.L. van Zanden, 2014, The Maddison Project: collaborative research on historical national accounts, Economic History Review 67, 627–651. Data available at: http://www.ggdc.net/maddison/maddison-project/home.htm

[3] Cogley, T., and J.M. Nason, 1995, Output dynamics in real-business-cycle models, American Economic Review 85, 492-511.

[4] Fouquet, R., and S. Broadberry, 2015, Seven centuries of European economic growth and decline, Journal of Economic Perspectives 29, 227-244.

[5] Klemp, M., and N.F. Møller, 2016, Post-Malthusian dynamics in pre-industrial Scandinavia, Scandinavian Journal of Economics 118, 841-867.

[6] Lagerlöf, N.P., 2015, Malthus in Sweden, Scandinavian Journal of Economics 117, 1091-1133.

[7] Lagerlöf, N.P., 2018, Understanding per-capita income growth in preindustrial Europe, mimeo, York University.

[8] Maddison, A., 2003, The World Economy: Historical Statistics, OECD, Paris.

[9] McEvedy. C., and R. Jones, 1978, Atlas of World Population History, Penguin UK.

[10] Statistics Sweden, 1969, Historical Statistics of Sweden, Part 1: Population 1720-1967, AB Allmänna Förlaget, Stockholm. Available from: www.scb.se/Grupp/Hitta_statistik/Historisk_statistik/_Dokument/Historisk-statistik-for-Sverige-Del-1.pdf

[11] Roser, M., and E. Ortiz-Ospina, 2017, World Population Growth, Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/world-population-growth

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

[13] 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, Cambridge, UK.

Online Appendix tables

Table A.1: Regressions based on simulated data: the benchmark model.

	Dependent variable:							
	Log birth rate year t				Log death rate year t			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log wage t	0.130^{***}			0.038^{***}	-0.192^{***}			-0.280^{***}
	(0.000)			(0.001)	(0.007)			(0.035)
Log wage $t-1$		0.132^{***}		0.113^{***}		-0.184^{***}		0.013
		(0.000)		(0.002)		(0.007)		(0.040)
Log wage $t-2$			0.129^{***}	-0.018^{***}			-0.180^{***}	0.078^{**}
			(0.000)	(0.001)			(0.007)	(0.035)
R^2	0.91	0.92	0.90	0.92	0.05	0.05	0.05	0.05
Number of obs.	50100	50000	49900	49900	50100	50000	49900	49900

Notes: Ordinary least squares regressions with standard errors in parentheses, based on 100 cases and 500 years of data from the benchmark simulation. All specifications include year and case fixed effects. * indicates $p<0.10$, ** $p<0.05$, and *** $p<0.01$.

Table A.2: Fraction of the years in which log GDP per capita fell outside the 5th and 95th percentiles across 1000 simulated economies when extending the model horizon to 2010. The results refer to all country-years over the periods 1300-1800, 1800-2010, and 1300-1800.

Period	**Fraction years below 5th percentile**	**Fraction years above 95th percentile**
1300-2010	.17	.096
1801-2010	.2	.28
1300-1800	.16	.0051

Table A.3: Regressions based on simulated data: the model extension with marriage.

	Dependent variable:											
	Log birth rate year t				Log death rate year t				Log marriage rate year t			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log wage t	0.176^{***}			0.036^{***}	-0.194^{***}			-0.281^{***}	0.061^{***}			0.235^{**}
	(0.002)			(0.009)	(0.008)			(0.035)	(0.021)			(0.097)
Log wage $t-1$		0.178^{***}		0.121^{***}		-0.186^{***}		0.013		0.053^{**}		-0.038
		(0.002)		(0.010)		(0.008)		(0.041)		(0.021)		(0.113)
Log wage $t-2$			0.175^{***}	0.023^{***}			-0.182^{***}	0.077^{**}			0.050^{**}	-0.140
			(0.002)	(0.009)			(0.008)	(0.035)			(0.021)	(0.097)
R^2	0.33	0.33	0.33	0.34	0.05	0.05	0.04	0.05	0.01	0.01	0.01	0.01
Number of obs.	50100	50000	49900	49900	50100	50000	49900	49900	50100	50000	49900	49900

Notes: Ordinary least squares regressions with standard errors in parentheses, based on 100 cases and 500 years of data from the simulation with marriage. All specifications include year and case fixed effects. * indicates $p<0.10$, ** $p<0.05$, and *** $p<0.01$.

Online Appendix figures

Figure A.1: Data over log GDP per capita 1300-2010, using data from both Fouquet and Broadberry (2015) and Bolt and van Zanden (2014). The series are normalized to zero when averaged over time and across countries.

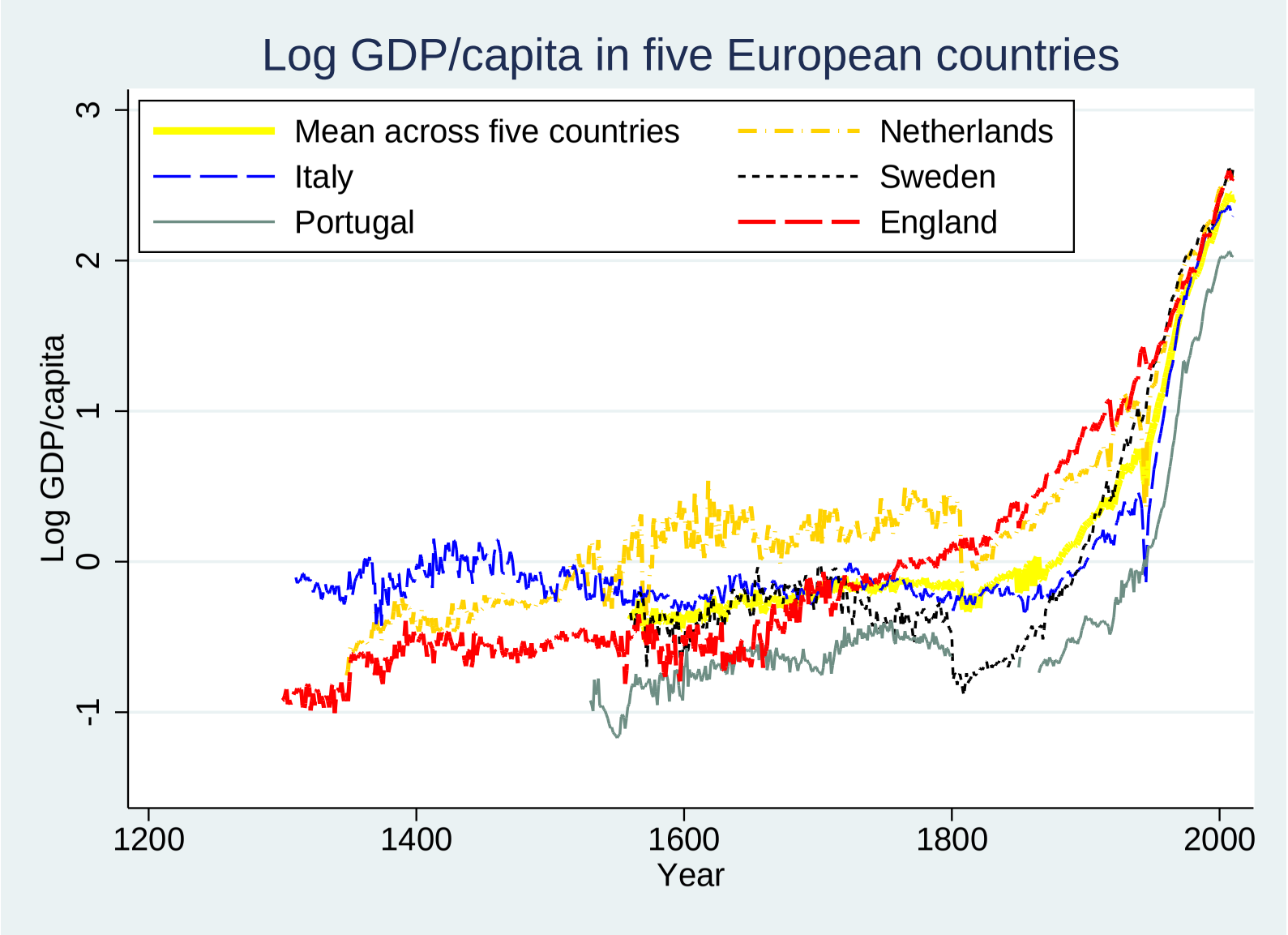

Figure A.2: This figure compares the GDP per capita data in Figure A.1 to the median and 5th and 95th percentiles for each year when extending the model horizon to 2010.

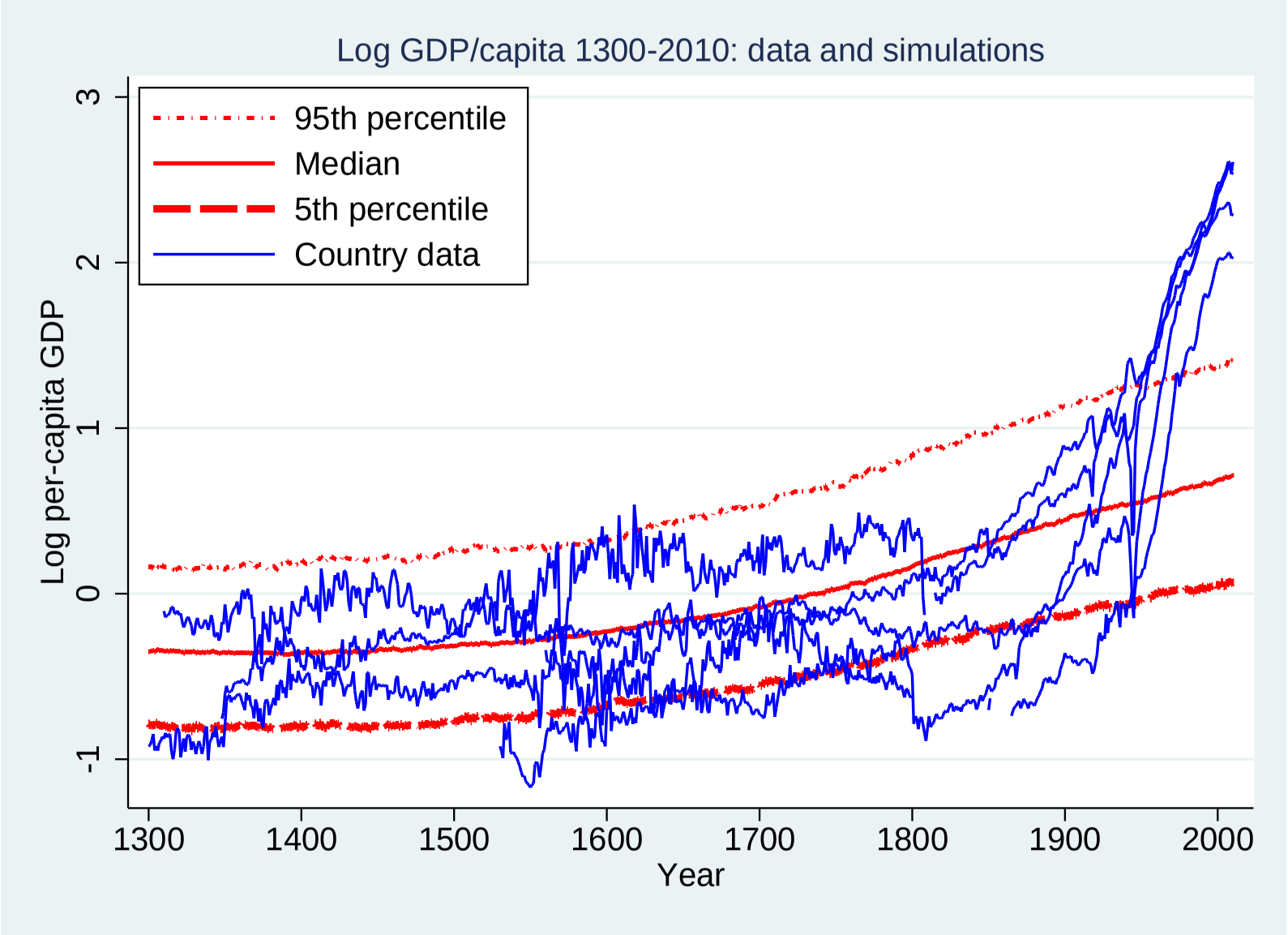

Figure A.3: Time paths for log of the Crude Birth Rate in data for England and Sweden, and the corresponding 5th and 95th percentiles from the benchmark simulation.

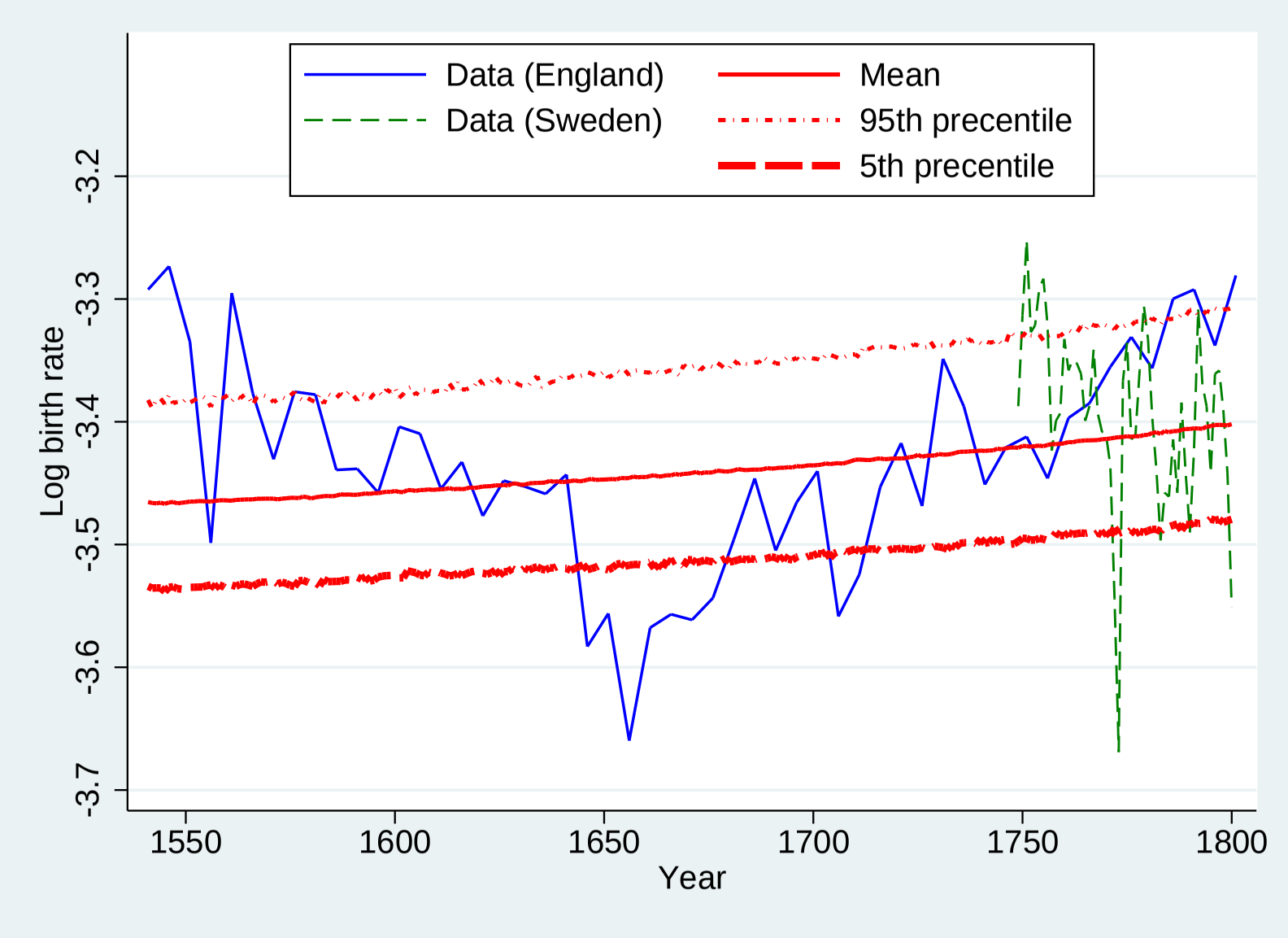

Figure A.4: Time paths for log of the Crude Death Rate in data for England and Sweden, and the corresponding 5th and 95th percentiles from the benchmark simulation.

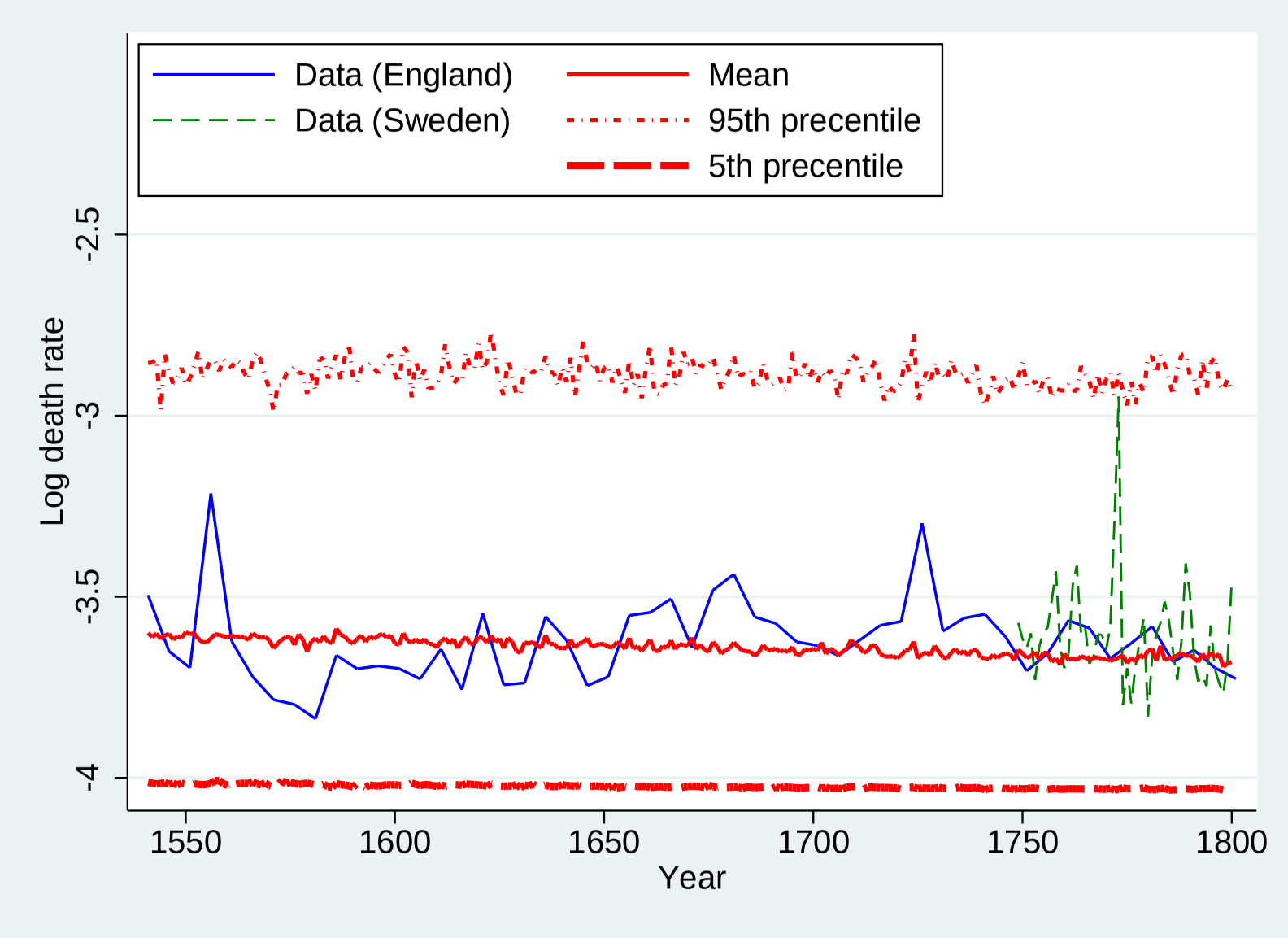

Figure A.5: Fertility, marriage, and wages by age group, when allowing for marriage. The top two figures correspond to those in Figure 2 in the paper. We have chosen v_j^{marr} and v_j^{unmarr} to match these profiles as closely as possible to Swedish data; see text. The model-generated profiles are averages over 1000 runs for the last 50 years of the simulation, corresponding to 1751-1800 in the data.

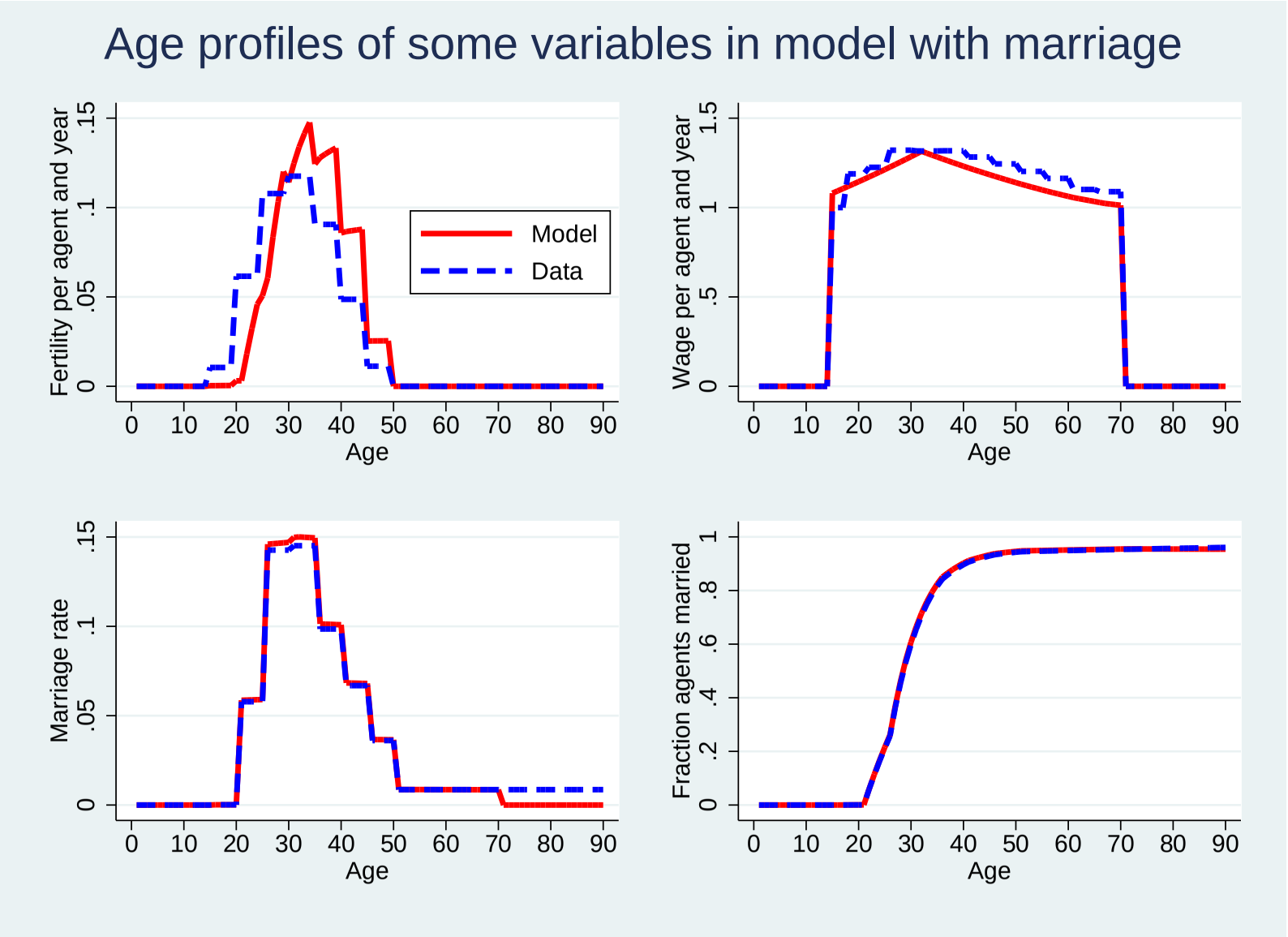

Figure A.6: Time paths for log of the Crude Birth Rate in data for England and Sweden, and the corresponding 5th and 95th percentiles from the simulation when allowing for marriage.

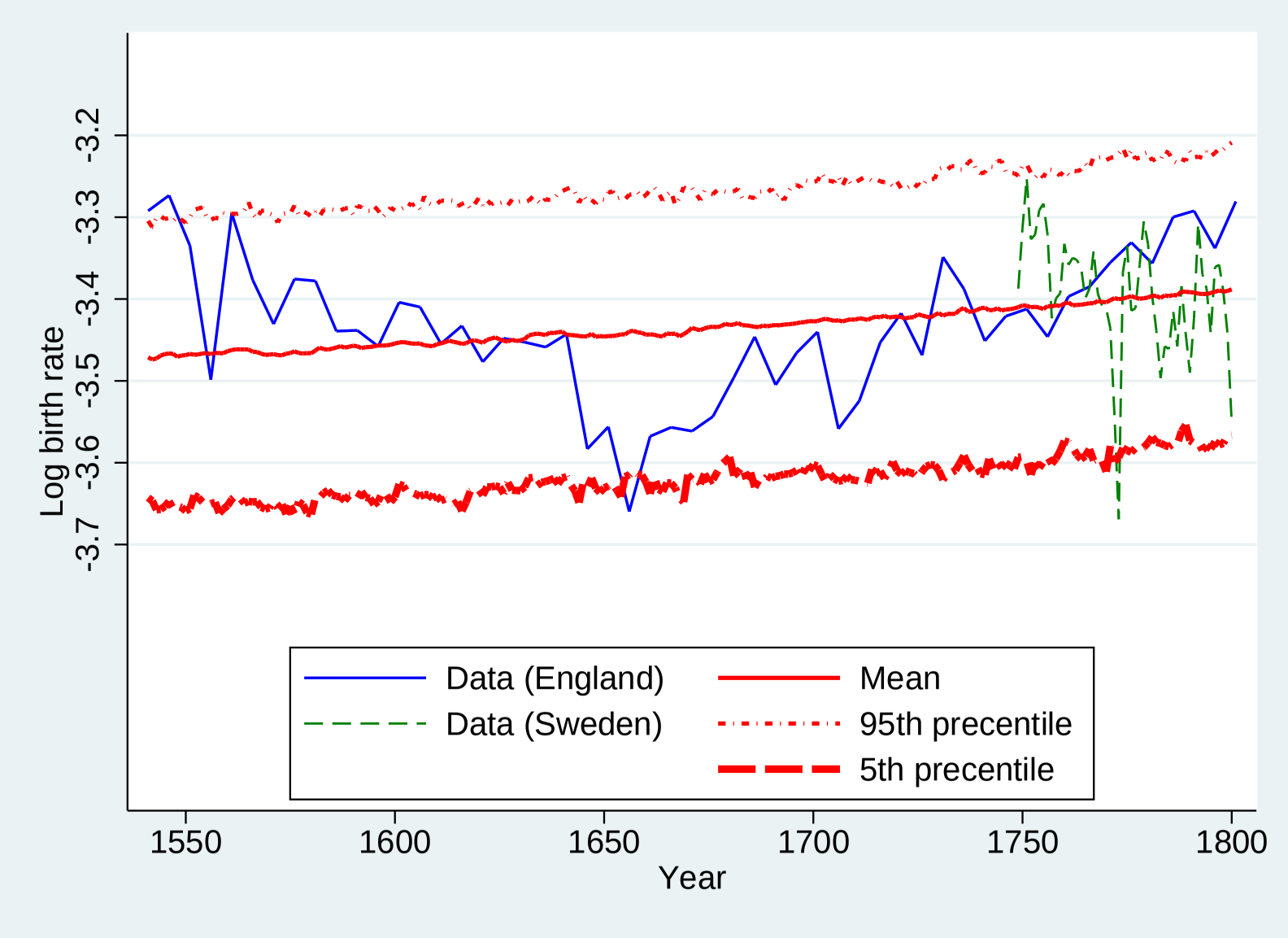

Figure A.7: Time paths for log of the Crude Death Rate in data for England and Sweden, and the corresponding 5th and 95th percentiles from the simulation when allowing for marriage.

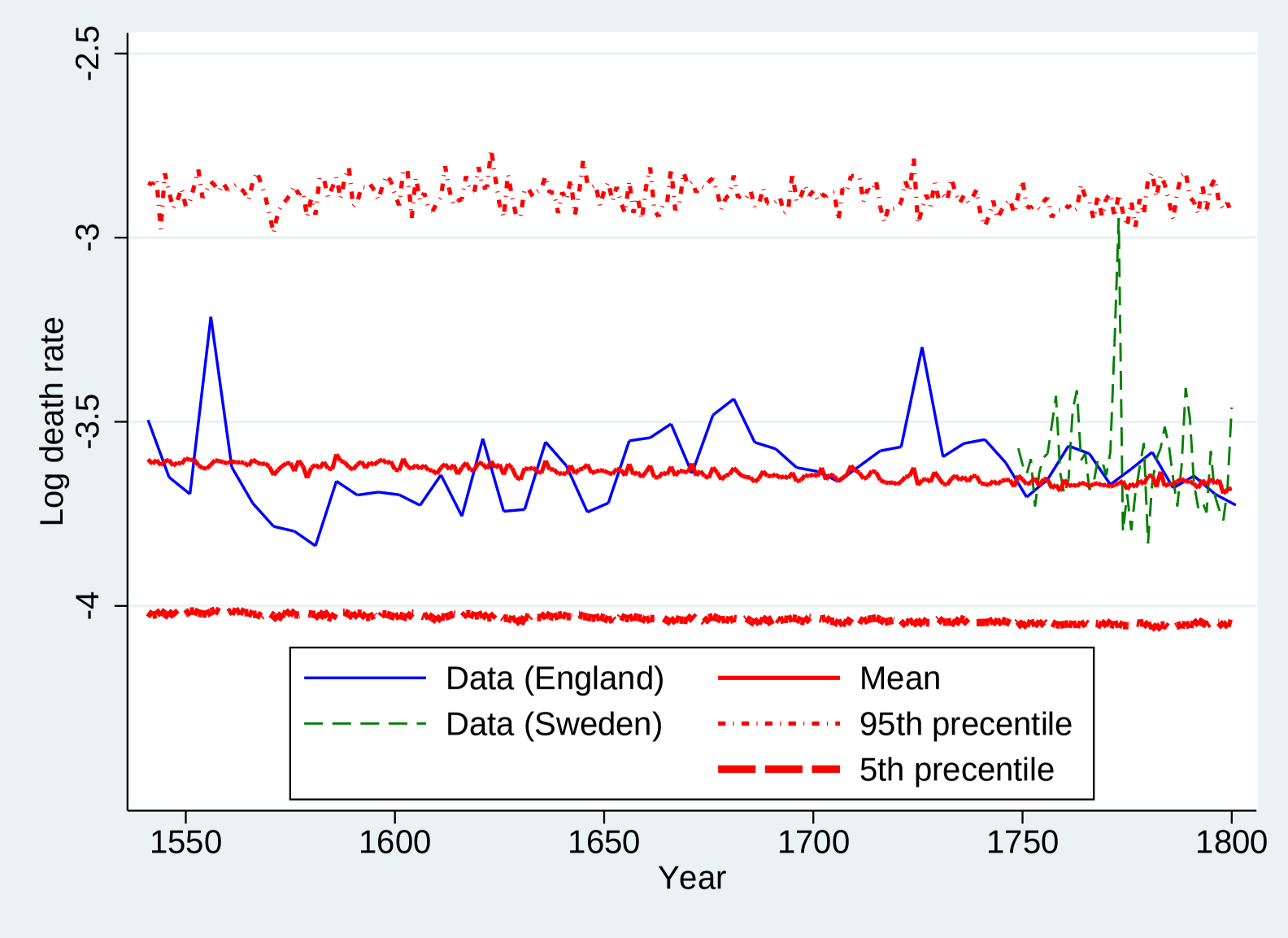

Figure A.8: Time paths for log of the Crude Marriage Rate in data for England and Sweden (adjusted as explained in the text), and the corresponding 5th and 95th percentiles from the simulation when allowing for marriage.

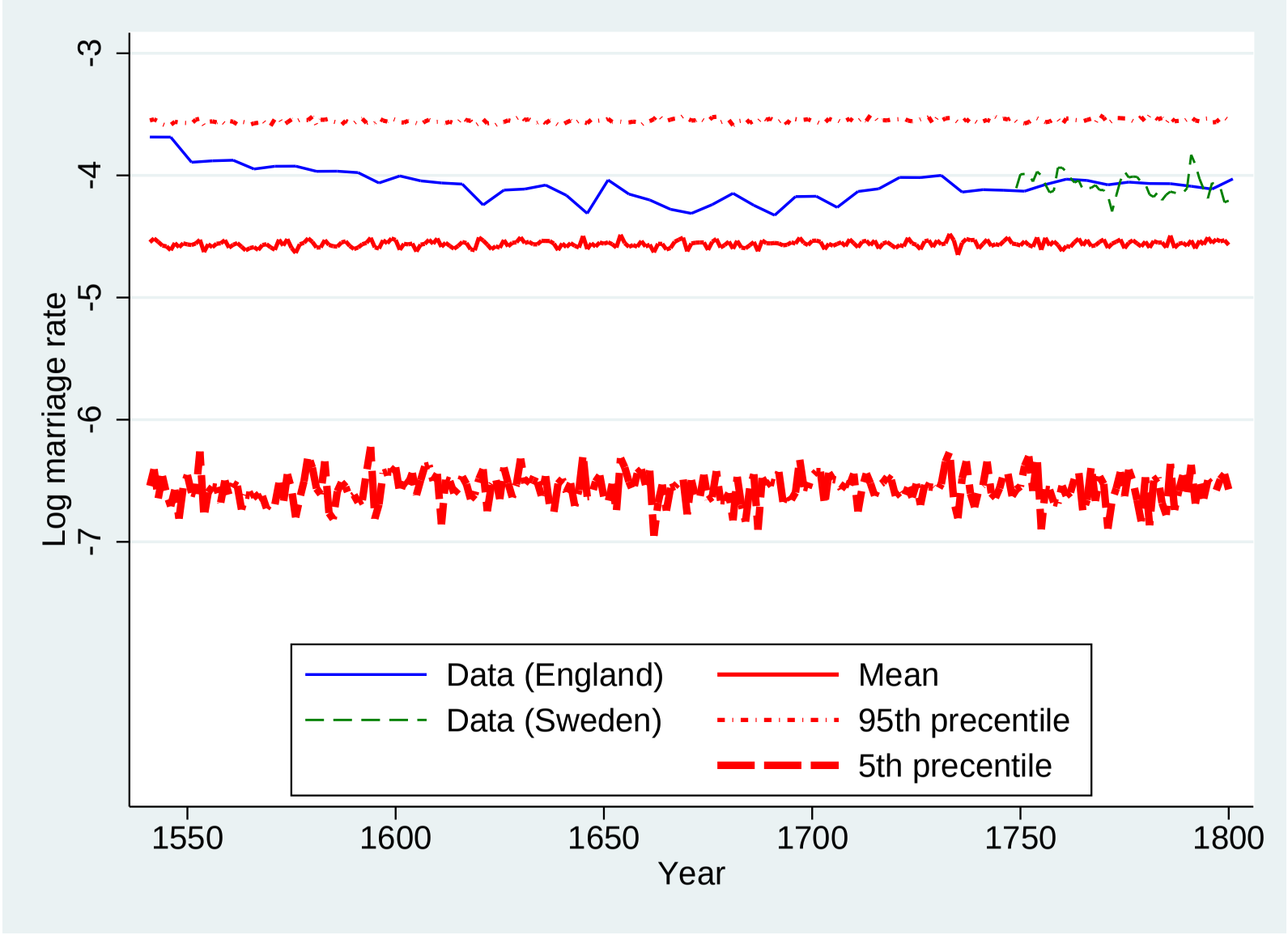

Figure A.9: The same histograms as in Figure 6 in the paper, when setting $\sigma_A = 0$ and $\sigma_X = .35$.

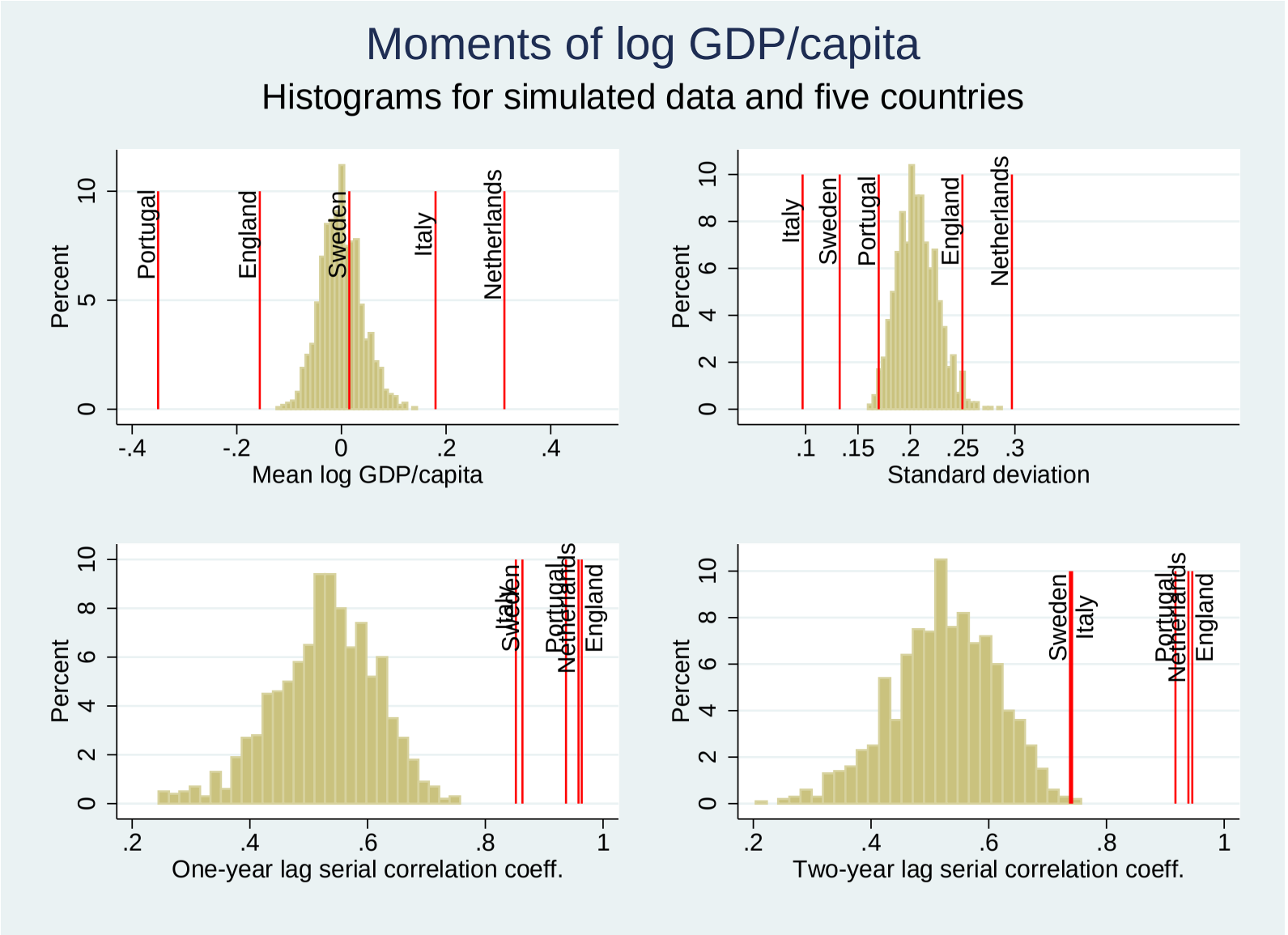

Figure A.10: The same histograms as in Figure 6 in the paper, considering the period 1560-1800, for which data are available for all five countries.

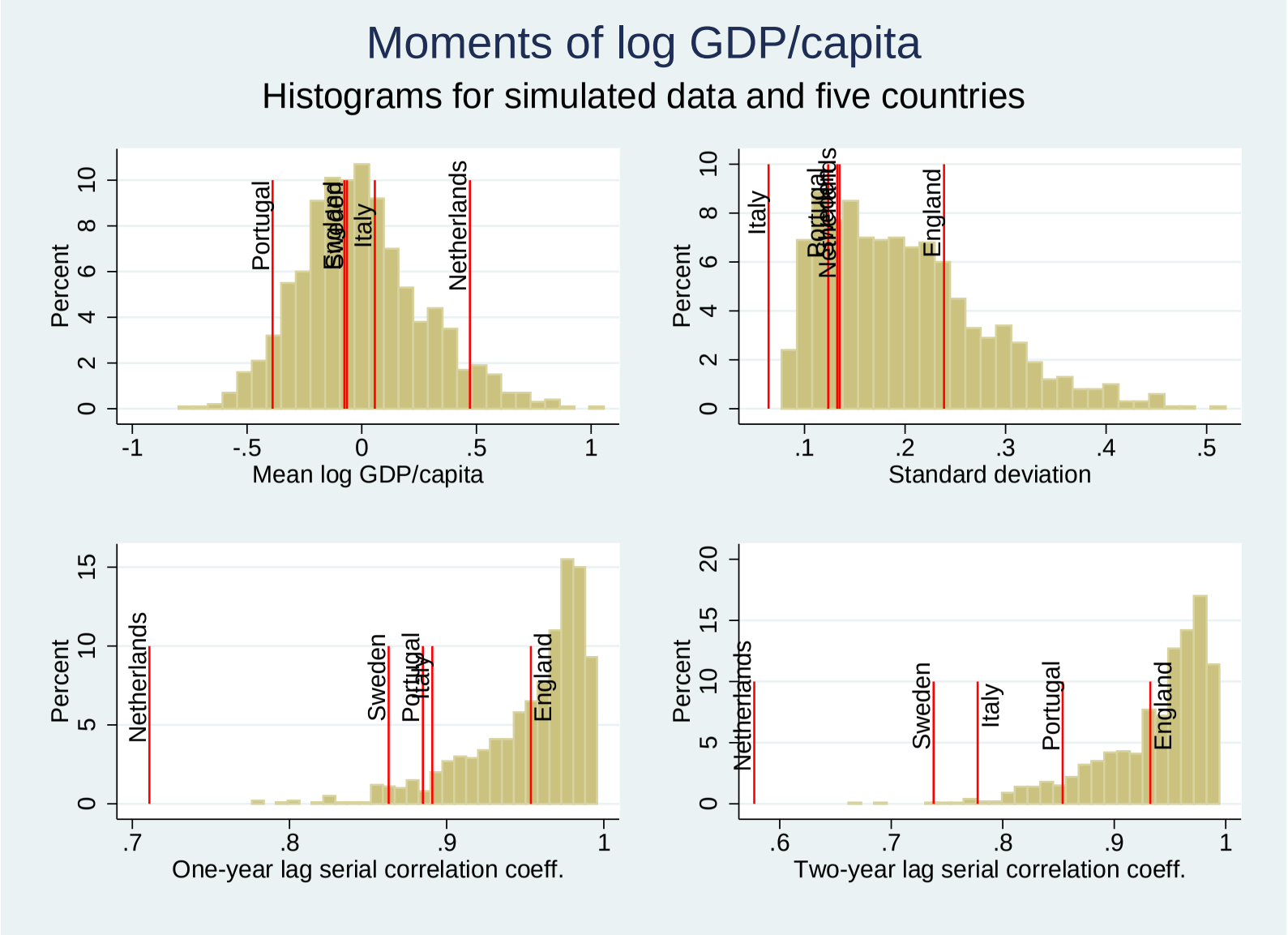

Figure A.11: Time path of the log population gap when starting off out of steady state, for different values of δ.

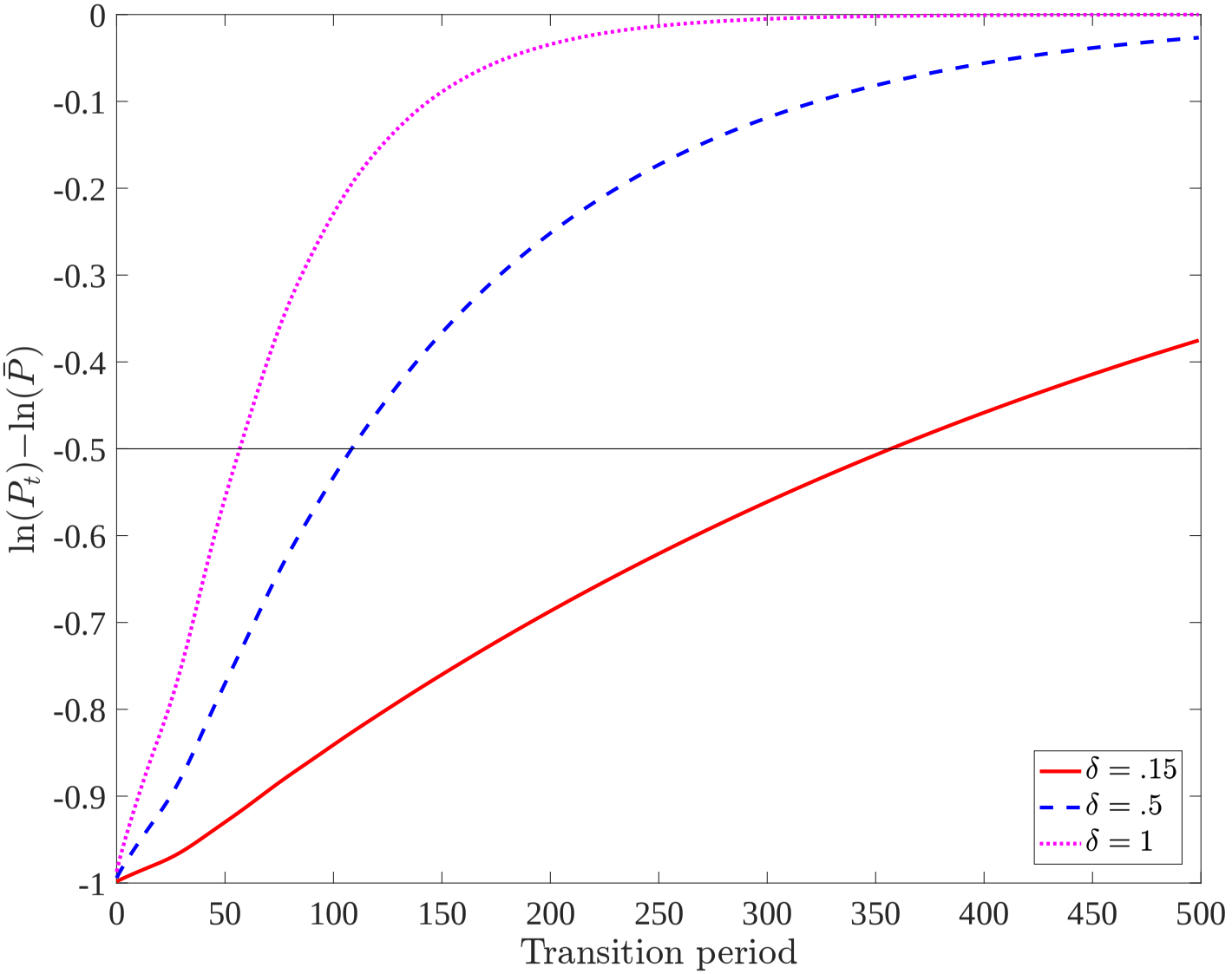

Figure A.12: How the largest eigenvalue of Π varies with δ.

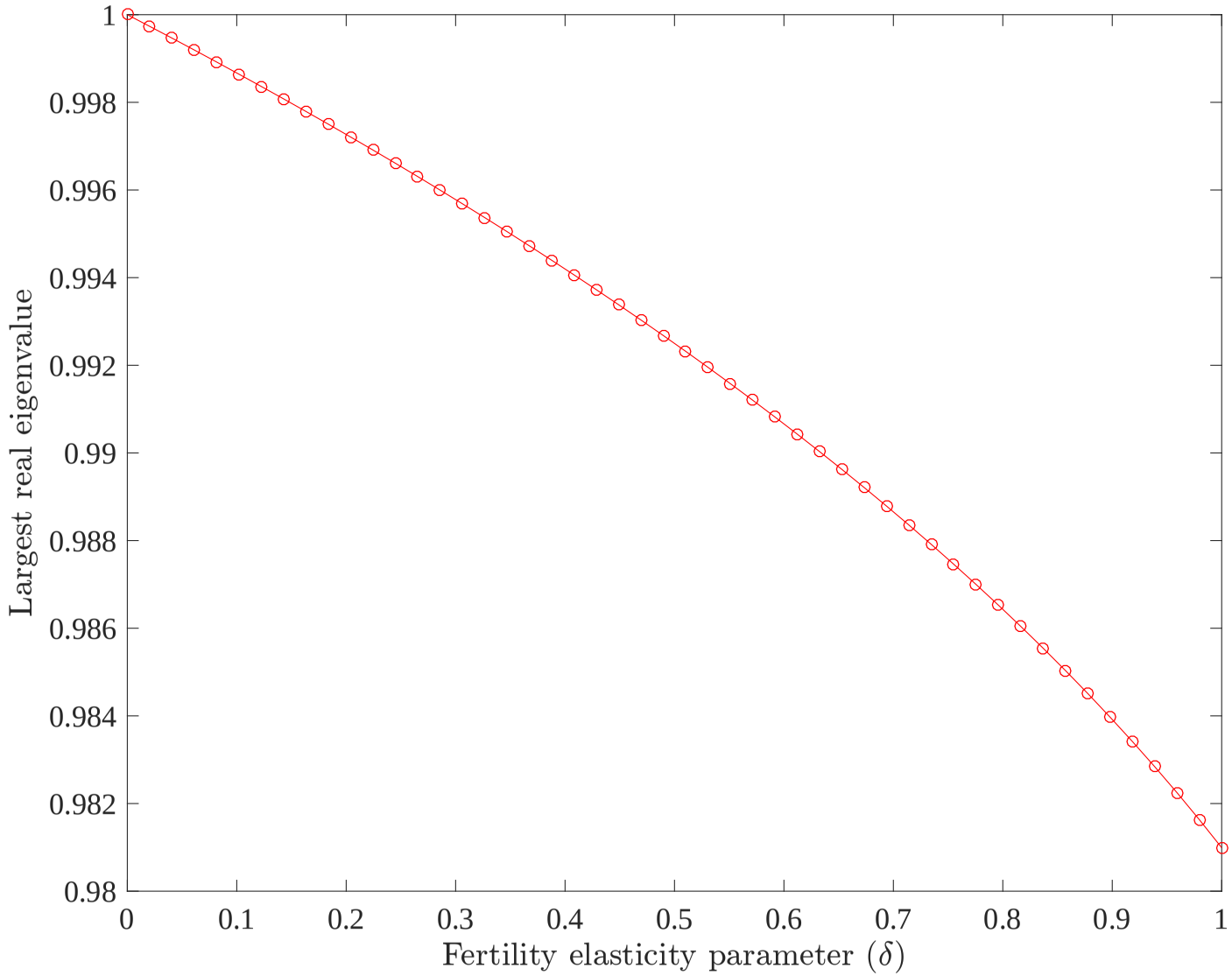

Figure A.13: Paths for log GDP/capita and log population for five actual countries, and five artificial ones. The artificial paths are constructed as the averages among the 25 simulated economies which are closest to the GDP/capita data for each country. The figure also shows paths artificially constructed to mimic levels 20% higher or lower than the GDP/capita levels observed in the data.

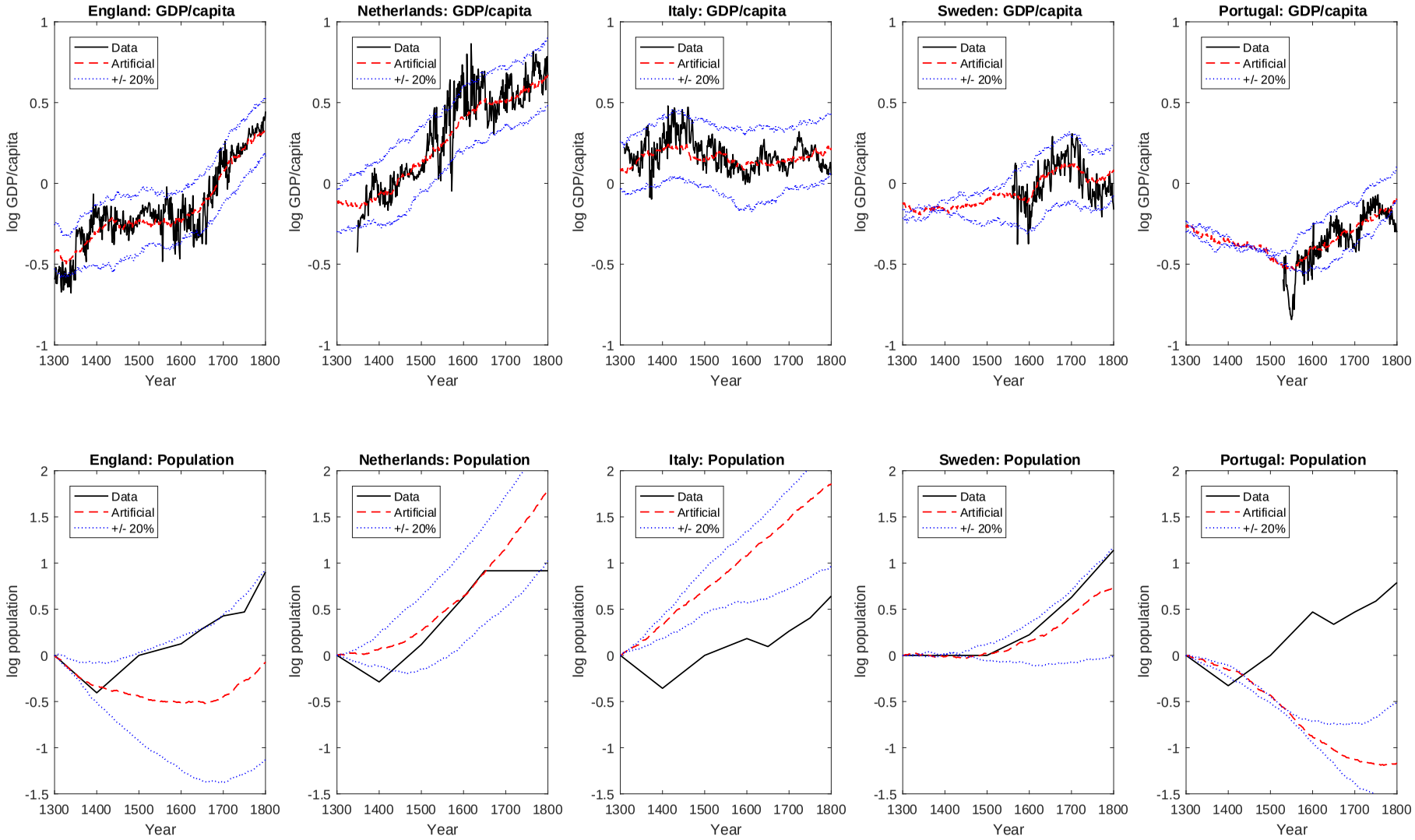

Figure A.14: The same paths as in Figure A.13, after demeaning the log GDP/capita data for each country to equal zero when averaging across years.