

Rathke, Alexander; Sarferaz, Samad

Working Paper

Malthus and the industrial revolution: Evidence from a time-varying VAR

KOF Working Papers, No. 351

Provided in Cooperation with:

KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Rathke, Alexander; Sarferaz, Samad (2014) : Malthus and the industrial revolution: Evidence from a time-varying VAR, KOF Working Papers, No. 351, ETH Zurich, KOF Swiss Economic Institute, Zurich,
<https://doi.org/10.3929/ethz-a-010073763>

This Version is available at:

<https://hdl.handle.net/10419/102969>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

KOF Working Papers

Malthus and the Industrial Revolution:
Evidence from a Time-Varying VAR

Alexander Rathke and Samad Sarferaz



KOF

ETH Zurich
KOF Swiss Economic Institute
WEH D 4
Weinbergstrasse 35
8092 Zurich
Switzerland

Phone +41 44 632 42 39
Fax +41 44 632 12 18
www.kof.ethz.ch
kof@kof.ethz.ch

Malthus and the Industrial Revolution: Evidence from a Time-Varying VAR*

Alexander Rathke[†]

ETH Zurich

Samad Sarferaz[‡]

ETH Zurich

Abstract

In the process of economic development economies grow through various regimes, each characterized by different demographic-economic interactions. The changes in these interactions are key elements in different explanations of the escape from Malthusian stagnation. We employ time-varying vector autoregressions, an approach that allows tracking this transition for England in the period between 1541 and 1870. The empirical findings suggest that the link between real wages and population growth was at work until the 19th century. Furthermore, we document changes in the propagation mechanism from real wages on population growth over time that feature prominently in Unified Growth Theory. Most remarkably, in contrast to earlier empirical literature we find strong effects of income on mortality after the 1750s.

JEL: C32, J13, N13, O11

Keywords: Industrial Revolution, Malthusian Trap, Time-Varying Vector Autoregression, Unified Growth Theory

*We are grateful to Julia Casutt-Schneeberger, Martin Dribe, Oded Galor, Mathias Hoffmann, Jan-Aaron Klaassen, John Komlos, Ronald Lee, Jim Malley, Heiner Mikosch, Esteban Nicolini, Ulrich Pfister, Wolfgang Reichmuth, Albrecht Ritschl, Peter Rosenkranz, Paul Sharp, Tobias Straumann, Hans-Joachim Voth, Jacob Weisdorf, Ulrich Woitek, and participants in seminars at the Humboldt University, University of Münster, Bodensee Symposium, the 2011 EHES congress, and the 2012 CESifo Area Conference on Macro, Money and Finance for very helpful comments. A previous version of this paper has circulated under the title “Malthus Was Right: New Evidence from a Time-Varying VAR”.

[†]KOF Swiss Economic Institute, ETH Zurich, Weinbergstrasse 35, CH - 8092 Zurich; email: rathke@kof.ethz.ch.

[‡]KOF Swiss Economic Institute, ETH Zurich, Weinbergstrasse 35, CH - 8092 Zurich; email: sarferaz@kof.ethz.ch.

1 Introduction

Economic growth as observed in modern industrialized societies is a relatively new phenomenon. For most of history, economic development was stagnant, with barely growing population figures and relatively stable real wages. Not until the sectoral shift from agriculture to industry did this period of stasis come to an end, marking the beginning of one of the most important events in world economic history, the Industrial Revolution.

Thomas Robert Malthus provided an economic-demographic framework for the period of stagnation that laid the foundation for the so-called “Malthusian” theory (Malthus, 1798). The Malthusian framework can be understood as a self-equilibrating system of population and economic activity in which the population level is stable at a subsistence wage level. Three assumptions are crucial to perpetuating the equilibrium. First, real wages are a decreasing function of population due to the existence of a fixed factor land. In other words, the production function exhibits decreasing returns to labor. Second, fertility positively depends on real wages (preventive check); for instance, increasing wages allow for a higher number of marriages. Third, mortality negatively depends on real wages (positive check). A decrease in wages causes malnutrition and deteriorating health conditions, leading to higher mortality rates.

Devising a framework that includes the transition from a stagnant Malthusian economy to a prosperous economy is the task of unified growth theory (UGT).¹ Many different aspects have been emphasized in the theory thus far, including human capital, demography, technological progress, and the influence of institutions. The prevalent explanations for the transition from stagnation to growth stress the role of human capital accumulation, technol-

¹An extensive overview can be found in Galor (2011). Recent contributions include Jones (2001); Galor and Moav (2002); Hansen and Prescott (2002); Lucas (2002); Lagerlöf (2003); Doepke (2004); Lagerlöf (2006); Doepke and Zilibotti (2008); Galor (2009). Other related contributions that model only particular aspects of the transition include Tamura (2002); Cervellati and Sunde (2005, 2006); Artzrouni and Komlos (1985); Komlos and Artzrouni (1990) and Kremer (1993).

ogy and population dynamics. In particular, according to Galor and Weil (2000) population growth accelerates the rate of technological change, making human capital more valuable. A virtuous cycle develops involving technological progress and human capital formation, as parents start to substitute "quantity for quality", reducing the number of children while increasing their education. First, increasing income allows households to have more children; later, the substitution effect dominates and fertility declines. Other studies suggest several complementary mechanisms that additionally spur the process of human capital formation through an increase in demand for human capital. For instance, higher life expectancy (lower mortality) through medical technology increases the individuals' ability to benefit from human capital accumulation.²

If we take these theoretical considerations seriously, they have implications for empirical analysis. The relationship between income and fertility - and hence, the parameters involved - changes from positive to negative as the economy evolves away from the Malthusian, through the post-Malthusian regime to modern growth.³ Presumably, other linkages are also time-varying; changes in the occupational structure, health environment, legal and social environment affect the relationship between the demographic and economic variables as transition takes place.⁴

In this paper, we add to the empirical literature by focusing on changes in the interaction of demography and income during the initial phase of industrialization.⁵ A vector autoregression (VAR) with time-varying coefficients and stochastic volatilities, as described in Cogley and Sargent (2005) and

²See Galor (2011) for an overview. A discussion of unified growth theories and a historical appraisal can be found in Mokyr and Voth (2007) and Broadberry (2007).

³See Galor (2005) for empirical evidence; see also Galor and Weil (1999); Galor and Moav (2002) and Lagerlöf (2003).

⁴See also Kalemli-Ozcan (2002); Lagerlöf (2003); Voigtländer and Voth (2013) and Voigtländer and Voth (2012).

⁵Note that in the literature on Malthus our real wage-fertility link is referred to as "preventive check" and our real wage-mortality link as "positive check".

Primiceri (2005) is applied.⁶ Time-varying parameters reflect long-run developments like changes in the “quantity-quality” trade-off, whereas changing variances account for variations in the size and nature of shocks (e.g. weather or epidemics).⁷

The model is applied to English data on birth and death rates and on real wages in the period from 1541 to 1870. This observation period captures the transition from the Malthusian economy to an industrialized one. The time-varying VAR analysis reveals that consistent with UGT the impact of real wages on population growth is found to be positive over the whole sample. Interestingly, the real wage-fertility link was mainly driving population growth until the Industrial Revolution, whereas the real wage-mortality link was more important in the period after 1750. Even though, the former results are roughly consistent with earlier findings, the latter results are in sharp contrast to the recent literature. As for the feedback effect of mortality on real wages, we find an overall downward trend for the entire sample. Moreover, the apparent decline in the volatility of the birth and death rate series is confirmed by our stochastic volatility results. The evidence that mortality reacted strongly to real wages during the initial phase of industrialization is our most striking result. We propose an explanation that emphasizes the increasing importance of wage dependence in conjunction with a change in the epidemiological environment.⁸

Related Literature and Our Empirical Approach

In a cross-country study for the time period 1-1500 CE, Ashraf and Galor (2008) find that technological improvements led to a higher population den-

⁶VAR models allow to solve the endogeneity problem by treating all variables as endogenous. It also allows to identify shocks and hence to calculate dynamic causal effects, e.g. Sims (1980).

⁷Lagerlöf (2003) assumes that human capital mitigates epidemic shocks, lowering the variance of mortality as the level falls.

⁸Our results imply that the classical Malthusian diagram (Lee, 1973, p. 593), is appropriately depicted by a rather horizontal wage-mortality schedule and an increasing wage-fertility schedule.

sity. They also find no effects of technological improvements on the standard of living which leads them to conclude that the Malthusian mechanisms were at work during this earlier period.⁹ By using time series data for England, several complementary studies test the Malthusian framework for the time period after 1540. Although the time series studies lack the cross-country dimension, they allow to model the dynamic relations between the variables more thoroughly. Lee and Anderson (2002) estimate a structural Malthusian model that allows for some time-variation in the coefficients to analyze data for the same period (1541 to 1870). Their results highlight a stronger real wage-fertility than a real wage-mortality link. The negative feedback effect of population on real wages has also been found to be weaker than suggested in earlier studies. Nicolini (2007) uses VAR techniques focusing on the short-run relationship between death and birth rates and real wages. For the period from 1541 to 1840, he finds evidence that the real wage-mortality link appeared only up until the 17th century, whereas the real wage-fertility link disappeared before 1740. Crafts and Mills (2009) repeat the analysis conducted by Lee and Anderson (2002) and Nicolini (2007), using a new real wage series compiled by Clark (2005). They also come to the conclusion that the Malthusian equilibrium did not exist in late pre-industrial England.¹⁰

The above mentioned studies clearly indicate changes in parameters and variances as implied by UGT. Nicolini (2007) uses sub-samples to capture the time variation in the parameters of his model. Lee and Anderson (2002) recognize the importance of time variation. However, their approach accounts only for long-run exogenous influences on mortality and fertility that are independent of the wage channel.¹¹ As discussed above, other developments that affect for instance the “quantity-quality” trade-off are also likely. Hence, our approach treats all parameters as time dependent while retain-

⁹Kelly and Ó Gráda (2010) present evidence that in the 13th century the effect of bad harvests on mortality was strong for all strata of society (unfree tenants and nobility).

¹⁰Nicolini even asserts that “perhaps the world before Malthus was not so Malthusian after all” (Nicolini, 2007, p. 116).

¹¹They model the intercept terms of the equations that describe the relationship between fertility, mortality and real wages as random walks in their state-space model. However, the crucial parameters connecting birth rates and death rates to wages are constant.

ing the non-restrictive VAR structure of Nicolini.¹² The transition from the Malthusian to an industrialized economy is modeled exploiting the information included in the whole sample. Hence, a to some degree arbitrary sample split is not necessary as the time-varying coefficients capture continuously the changes in the propagation mechanism. The time-varying variances capture the change in the impact and nature of the shocks, allowing us to model the apparent decline in volatility (Figure 1).

The rest of the paper is structured as follows. The next section introduces the time-varying VAR model and describes the estimation approach. Section 3 discusses the data used. Our prior assumptions are revealed in Section 4. The results are presented in Section 5 followed by a discussion of the model and the prior assumptions at play. Finally, Section 6 offers some concluding remarks.

2 Time-Varying Vector Autoregression

The following section describes the model proposed by Primiceri (2005), which is closely related to that of Cogley and Sargent (2005). The model allows for time-varying coefficients and volatilities, thereby providing extreme flexibility in tandem with a parsimonious specification.

2.1 Model Setup

Consider the following VAR model with time-varying coefficients and stochastic volatilities

$$\mathbf{y}_t = \mathbf{C}_t \mathbf{x}_t + \sum_{j=1}^p \mathbf{A}_{j,t} \mathbf{y}_{t-j} + \mathbf{u}_t, \quad (1)$$

¹²The model of Lee and Anderson (2002) is restrictive in the sense that, for example the structural model lacks a direct link between mortality and fertility as emphasized by Kalemli-Ozcan (2002). There is only an indirect channel through population.

where \mathbf{y}_t is a $n \times 1$ vector of endogenous time series, \mathbf{C}_t is a time-varying $n \times z$ matrix of parameters including a constant, \mathbf{x}_t is a $z \times 1$ vector of exogenous variables and $\mathbf{A}_{j,t}$ are p time-varying $n \times n$ parameter matrices.¹³ Rewriting the model in matrix form yields

$$\mathbf{y}_t = \begin{pmatrix} \mathbf{C}_t & \mathbf{A}_{1,t} & \dots & \mathbf{A}_{p,t} \end{pmatrix} \begin{pmatrix} \mathbf{x}_t \\ \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p} \end{pmatrix} + \mathbf{u}_t, \quad (2)$$

$$\mathbf{y}_t = \mathbf{A}_t \mathbf{Z}_{t-1} + \mathbf{u}_t.$$

Vectorizing both sides gives

$$\mathbf{y}_t = (\mathbf{Z}'_{t-1} \otimes \mathbf{I}_n) \mathbf{a}_t + \mathbf{u}_t, \quad (3)$$

where $\mathbf{a}_t = \text{vec}[\mathbf{C}_t \ \mathbf{A}_{1,t} \ \dots \ \mathbf{A}_{p,t}]$.¹⁴ The errors \mathbf{u}_t are assumed to be normally distributed with zero mean and time-varying covariance matrix $\mathbf{\Omega}_t$. The error term is structured as follows

$$\mathbf{u}_t = \mathbf{B}_t^{-1} \mathbf{\Sigma}_t \boldsymbol{\epsilon}_t,$$

where \mathbf{B}_t is a lower triangular matrix with ones on the main diagonal, $\mathbf{\Sigma}_t$ is a diagonal matrix, and $\boldsymbol{\epsilon}_t$ is assumed to be distributed as standard normal

$$\mathbf{B}_t = \begin{pmatrix} 1 & 0 & \dots & 0 \\ b_{21,t} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ b_{n1,t} & \dots & b_{n(n-1),t} & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{\Sigma}_t = \begin{pmatrix} \sigma_{1,t} & 0 & \dots & 0 \\ 0 & \sigma_{2,t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_{n,t} \end{pmatrix}.$$

¹³The following notation is used throughout: subscripts denote period t values of a variable (or vector of variables), and superscripts denote the history of a variable (or vector of variables) up to time t , e.g. $\mathbf{d}^t = [\mathbf{d}_1 \ \mathbf{d}_2 \ \dots \ \mathbf{d}_t]$.

¹⁴Note that, in contrast to Primiceri (2005), the matrix is vectorized along the column dimension.

The vectors $\mathbf{b}_t = [b_{21,t}, (b_{31,t} \ b_{32,t}), \dots, (b_{n1,t} \ \dots \ b_{n(n-1),t})]'$ are the equation-wise stacked free parameters of \mathbf{B}_t , and $\mathbf{h}_t = \log(\text{diag}(\boldsymbol{\Sigma}_t))$. The laws of motion for the parameters are assumed to be independent random walks,

$$\begin{aligned}\mathbf{a}_t &= \mathbf{a}_{t-1} + \boldsymbol{\nu}_t, \\ \mathbf{b}_t &= \mathbf{b}_{t-1} + \boldsymbol{\zeta}_t, \\ \mathbf{h}_t &= \mathbf{h}_{t-1} + \boldsymbol{\eta}_t.\end{aligned}\tag{4}$$

Because we are focusing on permanent changes to the parameters, mean reverting processes like stable autoregressive processes are not well suited to our research. Clearly, other even more flexible processes with a unit root could be applied that imply the same permanency of parameter changes.¹⁵ However, the conventional random walk assumption proved to be sufficiently flexible; we exploit the prior to control the smoothness of the stochastic processes by influencing their variances (Section 5.7.3).

The variance-covariance matrix of the innovations is block diagonal

$$\begin{pmatrix} \boldsymbol{\epsilon}_t \\ \boldsymbol{\nu}_t \\ \boldsymbol{\zeta}_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim N(\mathbf{0}, \mathbf{V}), \mathbf{V} = \begin{pmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W} \end{pmatrix}.\tag{5}$$

For simplicity's sake, it is assumed that matrix \mathbf{S} is also block-diagonal with respect to the parameter blocks for each equation.¹⁶

2.2 Estimation

Estimating a time-varying VAR model is challenging due to its high complexity and nonlinearities, which can potentially cause multiple peaks in low

¹⁵An example is an AR(2) process with a unit root that would involve the estimation of one additional parameter. The I(1) and I(2) processes are special cases of this more general process.

¹⁶See Primiceri (2005) for a discussion about relaxing this assumption.

probability regions of the likelihood.¹⁷ Classical estimation procedures have difficulties to cope with these problems. They lack the possibilities to rule out certain unreasonable regions of the parameter space sometimes suggested by the statistical model. In addition the multi-dimensional numerical maximization of the likelihood function can be problematic.

The Bayesian framework, however, provides the necessary tools to tackle these issues efficiently; for example, inadequate regions of the parameter space (e.g., regions implying highly explosive behavior) can be down-weighted using prior distributions. Following Cogley and Sargent (2005, 2001), Primiceri (2005) and Del Negro and Primiceri (2013), we adopt a Bayesian approach and employ Markov Chain Monte Carlo (MCMC) methods to estimate the time-varying VAR model. Applying the Gibbs sampling algorithm, a special MCMC method, allows one to split the estimation procedure for the high-dimensional model into smaller pieces that are easier to handle.¹⁸ Appendix A.1 explains the general concept and provides details of the Gibbs sampling algorithm used.

3 Data

For the empirical analysis, we use well-known annual data (1541–1870) on crude birth and death rates¹⁹ and a real wage index. The demographic variables were compiled by Wrigley and Schofield (1981) and the laborer wage series is taken from Allen (2001).²⁰ The raw data series are displayed in Figure 1.

The structural VAR approach usually has the disadvantage that the empirical results may depend on the ordering of the variables in the system due to the

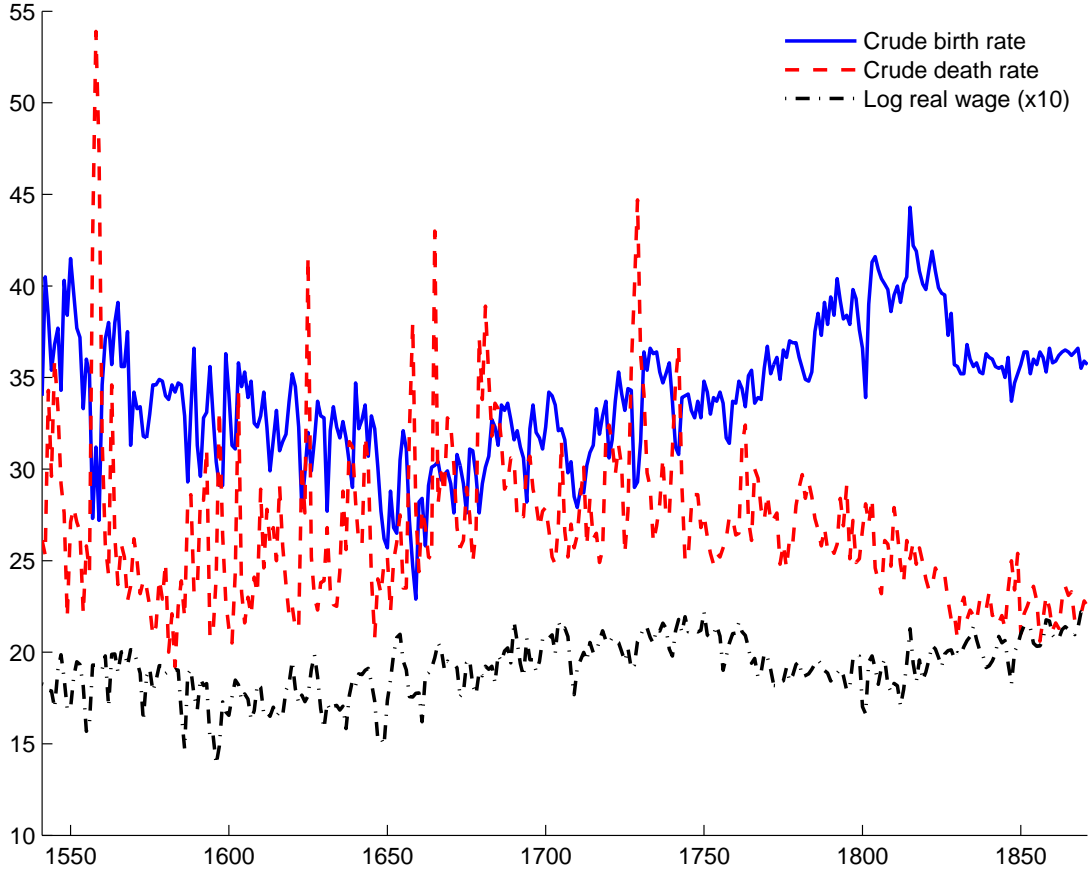
¹⁷See Primiceri (2005) for a more thorough discussion of these problems.

¹⁸The programs used to estimate the model were written by Ulrich Woitek and the authors in Matlab.

¹⁹The conventional unit is the number of births (deaths) per one thousand people.

²⁰We also used an alternative real wage series, see Section 5.7.4.

Figure 1: Demographic and Economic Data Series: 1541 to 1870



recursive identification scheme. Fortunately, this same identification strategy fits very well in our setting because the ordering of the demographic variables is predetermined by biology (Nicolini, 2007). The time interval between the decision to have a child and conception is on average over seven months (Bongaarts, 1978). Adding this interval to the nine months of pregnancy, we can see that the endogenous reaction of the birth rate is delayed by over a year. Assuming that wages affect mortality rather slowly through nutritional status, it is also more likely that real wages can react faster to population changes, than mortality reacts to wage changes. Death shocks affect the labor supply per definition instantaneously when at least some members of the working population are affected.²¹ As a result, it is safe for us to order

²¹See Lee and Anderson (2002, p. 210) for evidence on a significant positive instanta-

the crude birth rate first, the crude death rate second, and the real wage last.²² This also allows us to compare our results to those of Nicolini (2007) and Reichmuth (2008), who use the same ordering. Thus, the following data vector is used for the VAR

$$\mathbf{y}_t = [CBR_t, CDR_t, w_t]',$$

where CBR denotes the crude birth rate, CDR the crude death rate, and w the log of the real wage.

4 Priors and Model Specification

To find some of the prior parameters we use a sample from the beginning of the observation period (1541–1600) as applied by Cogley and Sargent (2005) and Primiceri (2005). We estimate a time-invariant VAR model using ordinary least squares (OLS) and use the point estimates to calibrate some of the prior distributions. Because the crude death series contains some spikes caused primarily by epidemics and wars in the initial sample, we used dummies to account for these. Given the high dimensionality of the model the priors are by construction informative. However, the priors are chosen in such a way that they are as diffuse as possible.

neous effect of population changes on wages and an insignificant contemporaneous impact of wages on mortality.

²²Reordering crude death rates and real wages does not affect our principal findings.

The prior choices can be summarized as follows.

$$\begin{aligned}
\mathbf{a}_0 &\sim N(\hat{\mathbf{a}}^{OLS}, 4 \text{Var}(\hat{\mathbf{a}}^{OLS})) \\
\mathbf{b}_0 &\sim N(\hat{\mathbf{b}}^{OLS}, k_b^2 \mathbf{I}_3) \\
\log \boldsymbol{\sigma}_0 &\sim N(\log \boldsymbol{\sigma}^{OLS}, k_h^2 \mathbf{I}_n) \\
\mathbf{Q} &\sim IW(k_Q^2 \text{Var}(\hat{\mathbf{a}}^{OLS}), T_Q) \\
\mathbf{W} &\sim IW(k_W^2 \mathbf{I}_n, T_W) \\
\mathbf{S}_1 &\sim IW(k_S^2 \mathbf{I}_1, T_{S_1}) \\
\mathbf{S}_2 &\sim IW(k_S^2 \mathbf{I}_2, T_{S_2})
\end{aligned} \tag{6}$$

where k_i for $i = b, h, Q, W, S$ are the usual scaling factors that are set to 0.01. The shape parameters T_j are set to $\dim(j) + 2$ for $j = Q, W, S_1, S_2$.

Bayesian model selection based on posterior odds ratios strongly suggests a time-varying model with 4 lags. This lag length is also often used for time invariant models. We will conduct a detailed discussion of prior and model choice and a sensitivity analysis of the results in Section 5.7.

5 Empirical Results

The Gibbs sampling algorithm described in Primiceri (2005) is used to generate 100,000 draws from the posterior, with the first 80,000 discarded as burn-in. To ensure convergence we conduct several tests, some visual and others numerical. All convergence diagnostics were satisfactory.²³

In the following, we present the empirical results. The time variation in the variances is discussed first. Then, the three implications of the Malthusian

²³We repeatedly started from different overdispersed starting values and experimented with different numbers of draws. We also looked at traceplots, running mean plots, and autocorrelations at various lags and calculated Geweke's χ^2 test (e.g. Geweke, 1999). The traceplots, running mean plots, and the detailed numerical convergence diagnostics are available on request. The statistical tests were calculated using Matlab code provided by James P. LeSage.

model are explored using impulse response analysis. We track how the real wage-fertility/mortality link and the reaction of wages to population changes evolved in the period from 1601 to 1870.²⁴ Unlike earlier researchers, we do not have to rely on sub-samples because our approach enables the computation of impulse responses for each year.

5.1 Declining Volatilities

We begin our analysis by investigating the estimated volatilities. The posterior means for the standard deviations of the structural shocks in the period from 1601 to 1870 are displayed in Figure 2.²⁵ A clear decrease in all variances is visible over the centuries but is most pronounced in the mortality series and least pronounced in the real wage series. Moreover, the estimated volatilities clearly capture changes in the nature and severity of shocks.

The results are similar to those observed by Wrigley and Schofield (1981). According to these authors, the amplitude of decadal fluctuations around a 25-year moving average in the demographic series declined over time, and variations in the mortality series were always more pronounced than those in the fertility series (Wrigley and Schofield, 1981, pp. 313-320).²⁶ Moreover, it is well known that major surges in mortality due to epidemic diseases or starvation ceased to occur.

Presumably, the quality of the historical series deteriorates as it goes back in time.²⁷ One more advantage of our econometric framework is that changes in

²⁴Note that due to the beginning of civil registration the demographic data is differently constructed after 1837 (see, Wrigley and Schofield, 1981, p. 3-4, pp. 130-135). We have estimated the model with the shorter sample 1541-1837. The main findings are robust to this truncation.

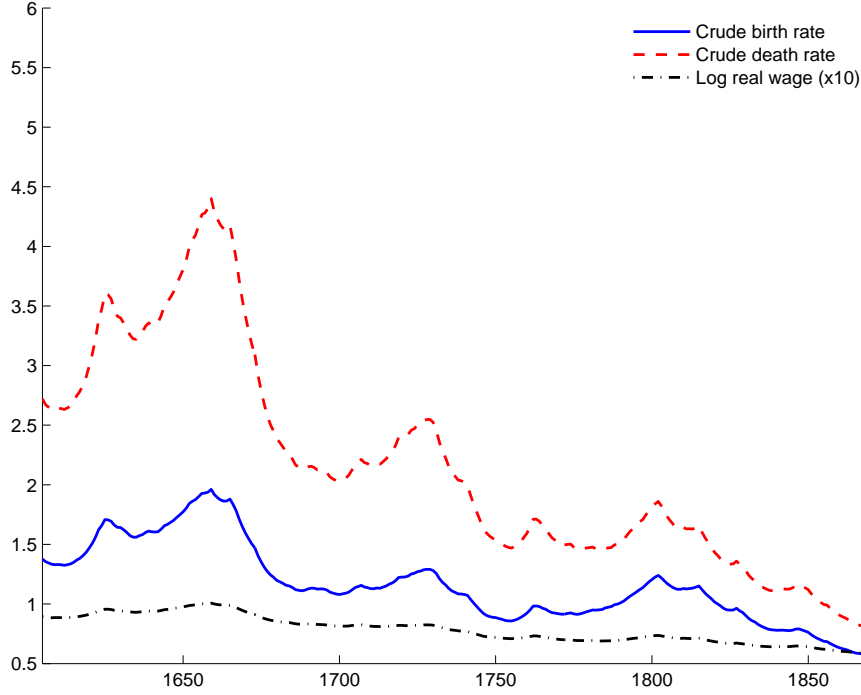
²⁵A comparison of posterior and prior means is depicted in Figure 10 in Appendix B. It can be seen that the prior is updated by information in the observed data.

²⁶The declining volatility was also noted by Bailey and Chambers (1993), Nicolini (2007), and Reichmuth (2008).

²⁷For instance, due to the back projection method used by Wrigley and Schofield (1981). See, for instance Lindert (1983), Razzell (1994), and Woods (2000) for a more detailed discussion of the Wrigley/Schofield data series.

the quality of the data can be captured by the time variation in the variances. Hence, part of the observed decrease in variance could also be attributed to a decrease in measurement error.

Figure 2: Estimated Volatilities

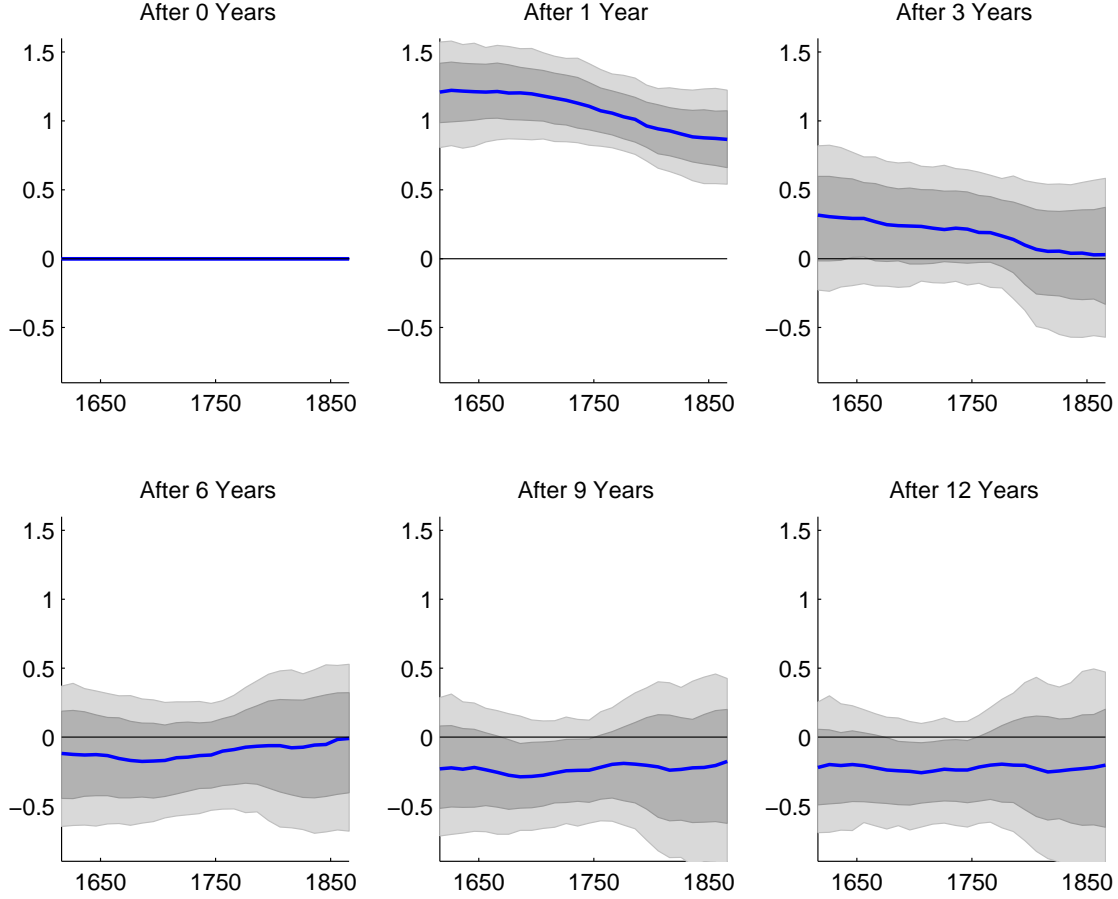


Notes: The graph shows posterior means of the estimated standard deviation of the structural shocks. For clearer presentation the standard deviation of the real wage shock was rescaled by a factor of 10.

5.2 The Real Wage-Fertility Link

First, we turn to the reaction of fertility to exogenous changes in the wage rate. Figure 3 shows the posterior means of the impulse responses of the crude birth rate to real wage shocks over time. Each graph displays the impulse response to a wage shock after a specific number of years, calculated for each decade from 1601 to 1870. To be able to see only the changes in the propagation mechanism, we calculate the impulse responses in all time periods to the same shock set to one standard deviation of the underlying

Figure 3: Response of CBR to a Wage Shock, 1601-1870.

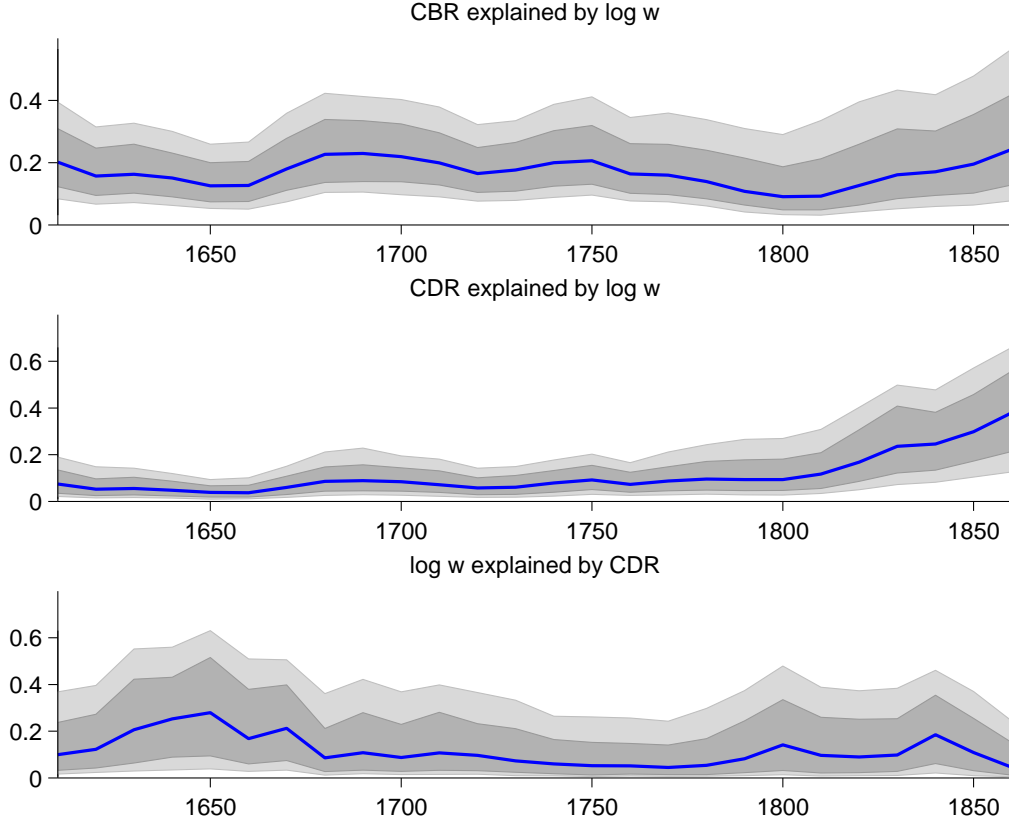


Notes: Posterior means of the impulse response to standardized real wage shock after a specified number of years for each decade from 1601 to 1870. The gray (dark gray) area indicates the 5th (16th) and 95th (84th) percentiles of the impulse responses.

data series. The upper-left graph depicts the instantaneous responses of the birth rate to a wage shock. The impulse responses at impact are equal to zero for all time periods, which follows directly from our identification strategy. However, this changes in the first year after the shock. The graph in the upper middle panel clearly reveals a positive effect of a wage shock on fertility for all time periods. The impulse responses are hump-shaped in each decade and come back to zero after about three years.²⁸

²⁸See Figure 9 in the appendix for the more conventional depiction of impulse responses over different horizons for some specific years.

Figure 4: Forecast Error Variance Decompositions, 1601-1870.



Notes: Means of forecast error variance decompositions for each decade from 1601 to 1870. The gray (dark gray) area indicates the 5th (16th) and 95th (84th) percentiles of the variance decompositions.

Helping us to assess the importance of real wage shocks on fertility, the uppermost panel of Figure 4 shows the contribution of real wage shocks to the forecast error variance of birth rates for each decade from 1601 to 1870. Note that forecast error variance decompositions can be interpreted as partial R^2 s for each forecast horizon.²⁹ The figure reveals that the wage shock explains a roughly constant fraction of forecast error variance of fertility throughout the sample. The posterior mean share of explained variance slightly oscil-

²⁹Note that to evaluate the contributions of shocks to the unconditional variance the forecast horizon is set to infinity.

lates between 10 to 20 percent reaching its low point around 1800. Nicolini (2007) finds an increasing contribution of real wage shocks over time in his subsamples.

Like Wrigley and Schofield (1981), Lee (1981), Galloway (1988), and Bailey and Chambers (1993) we find a positive relation between real wages and fertility. However, while the forecast error variance decompositions are almost time-invariant, the impulse responses of fertility to a real wage innovations change over time. This is similar to the findings of Nicolini (2007), and Crafts and Mills (2009), where the real wage-fertility link is also found to be unstable over time.³⁰

Apart from affecting fertility through nutrition, frequency of intercourse, and proportion of anovulatory menstrual cycles, the positive effect of real wages on fertility mainly acted via an increase in marriages (Wrigley and Schofield, 1981, Chapter 8).³¹ Interestingly, the strongest change in the dynamics of the income-fertility relationship starts with the breakdown of the European Marriage Pattern (EMP) at the beginning of the 18th century, which combined late marriage or celibacy with unrestricted fertility within marriage Hajnal (1965).³² As a result, fertility increased due to the occurrence of earlier marriage in the 18th century (Wrigley et al., 1997). In this particular period, fertility was starting to become detached from variations in wages as suggested by UGT.³³

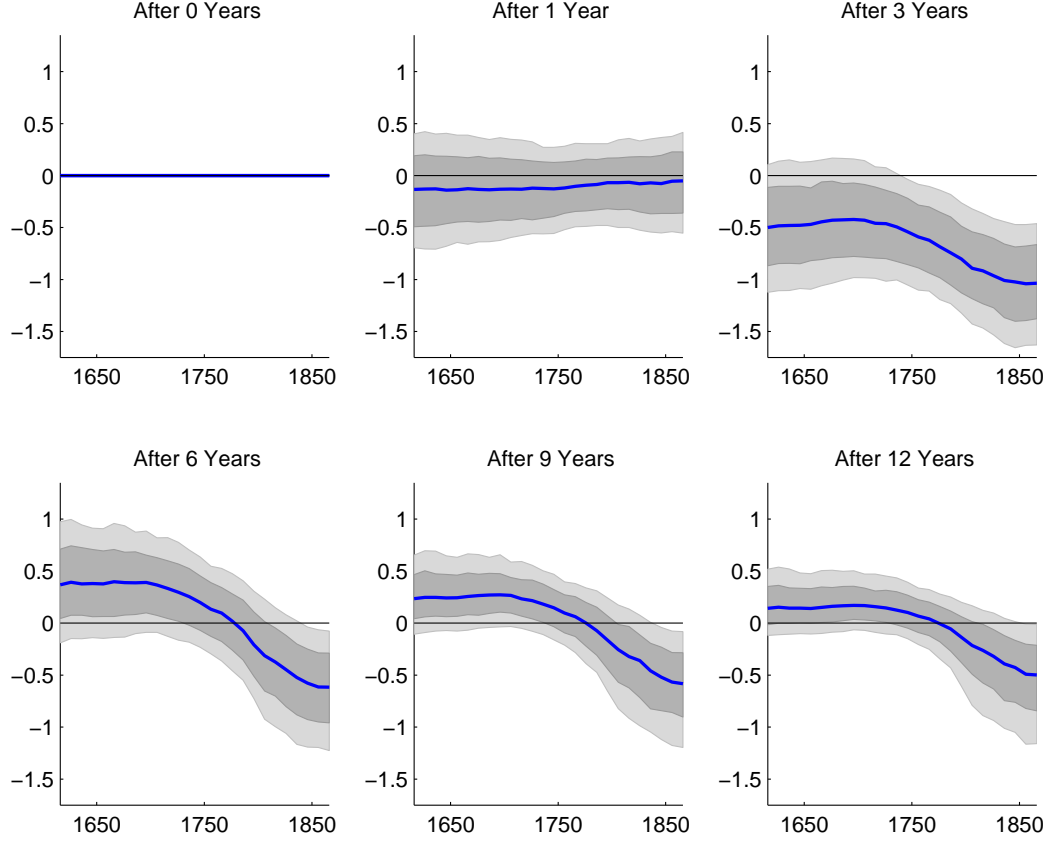
³⁰When using disaggregated data from the 404 parishes of Wrigley and Schofield (1981), Kelly and Ó Gráda (2011) find that a correlation between real wages and fertility up until the early 18th century that diminished thereafter.

³¹See also (Clark, 2007a, Chapter 4). However, Lee (1975) cannot find supporting evidence.

³²Postponing or eschewing marriage could limit fertility to less than half the biological maximum Clark (2007a). Voigtländer and Voth (2013, 2009) explain the formation of the EMP as an endogenous response to a large income shock caused by the Black Death, which caused a major change in the structure of agricultural production from “corn to horn” resulting in more employment opportunities for women as servants in animal husbandry.

³³Voigtländer and Voth (2013) also argue that the vanishing of late marriage might be explained by the decline in employment opportunities for women in husbandry.

Figure 5: Response of CDR to a Wage Shock, 1601-1870.



Notes: Posterior means of the impulse response to a standardized real wage shock after a specified number of years for each decade from 1601 to 1870. The gray (dark gray) area indicates the 5th (16th) and 95th (84th) percentiles of the impulse responses.

5.3 The Real Wage-Mortality Link

Figure 5 shows the impact of real wage shocks on mortality (See also 9 in the appendix). We find that mortality reacts negatively to a real wage shock throughout the entire sample. However, we observe an interesting change in the dynamics over the centuries. At higher forecast horizons we can see also some probability for the existence of a rebound effect. A period of unusually high mortality is followed by a period of lower mortality if the weakest individuals have perished in the initial wave (Lee and Anderson, 2002, p. 212). However, this rebound effect vanishes over time and the negative effect of

real wages on mortality becomes more pronounced and persistent over time. Thus, a clear Malthusian positive check enters the stage after 1750. While some probability exists that the impact of real wages on mortality was of modest size, if at all present, in the beginning of the sample, it becomes stronger in the mid-18th century, reaching a trough in 1840. The effects decay slowly and are still active after twelve years (see the second row of Figure 5). This implies a relatively persistent effects of a real wage shock. The second row of Figure 4 displays the corresponding contribution to the forecast error variance of mortality. As we can see, the fraction of the variance of mortality explained by innovations in real wages is small until the 18th century. After 1800, the whole probability distribution shifts massively upwards, pushing the mean explained variance to around above 40 percent by the mid-19th century.

The findings up until 1750 are mostly in agreement with the empirical Malthus literature. The evidence predominantly does not support a link between real wages and mortality during that time.³⁴ However, the results for the period after 1750 differ from the rest of the recent literature: neither Nicolini (2007) and Crafts and Mills (2009) nor Lee and Anderson (2002) identify a negative link between real wages and mortality for this period. The same is true for Wrigley and Schofield (1981) and Lee (1981) who find that the a negative relation between real wages and crude death rates disappeared sometime in the early 17th century. We find strong effects after 1750 that even increase in importance afterwards. Astonishingly, we find the strongest relationship between real wages and mortality for the period from 1750 to 1870, when one is commonly supposed not to have been present.

The emergence of a strong real wage-mortality link during the initial phase of the Industrial Revolution can be explained by the rapid structural changes

³⁴Nicolini (2007) and Crafts and Mills (2009) find supportive evidence up until 1640. They also do not necessarily contradict the results of Lee and Anderson (2002) and Møller and Sharp (2008), who find only minor or insignificant results for the positive check. Note that Møller and Sharp (2008) focus on the period from 1560 to 1760, whereas Nicolini (2007) uses data covering the period from 1541 to 1840.

in the economy that occurred during that period. One result of the shift from the primary to the secondary sector was that the share of wage earners in the labor force increased. The agricultural share of labor decreased from 35 percent in 1801 to 22 percent in 1851, whereas the industrial labor share increased from 29 percent to 42 percent.³⁵ Hence, people could no longer rely on subsistence farming during times of hardship.³⁶ Not only did a shift from agriculture to manufacturing occur but also the structure of agriculture itself changed. The major change in agriculture was that large scale capitalist farms that used wage-dependent labor replaced peasant cultivators (Allen, 2004, 2009). Moreover, families started to depend more and more on the earnings of more erratically employed fathers as female earning opportunities declined as industrialization spread (Allen, 2004). The process of rapid growth of industrial areas was also accompanied by urbanization and the corresponding side effects, such as overcrowding, an increase in crime, inadequate access to sanitation and other infrastructure, inadequate access to safe water and poor-quality working-class housing.³⁷ At the same time, diseases where the outcome depends on nutritional status (e.g. measles, typhoid fever)

³⁵See Deane and Cole (1969, p. 143). The share of the agricultural sector's contribution to the total output was almost halved from 1790 to 1850, decreasing from around 40 percent to 21 percent, whereas the share of the industrial sector's contribution almost doubled, from around 21 percent in 1790 to 35 percent in 1850 (Mitchell, 2003, p. 934).

³⁶In contrast, Galloway (1988) finds in a cross-country study of prices and mortality a positive correlation between the share of population employed in agriculture and the mortality-prices link.

³⁷Note that the Wrigley and Schofield parish data do not include London as the authors supposed that the high turnover in its population precluded reconstitution (Wrigley and Schofield, 1981, p. 6). Therefore, we cannot simply assume that the aggregate data reflects the data for the urban center of England. Moreover, Kelly and Ó Gráda (2010) find no impact of wages on mortality for a subsample from 1750–1800 using a distributed lag model for only London. They explain the disappearance of the Malthusian positive check by the systematic introduction of poor relief during this period. However, when inflating the data to national frequency the data was adjusted to account for London (Wrigley and Schofield, 1981, pp. 77–88). We also used the bills of mortality data (taken from Hull, ed (1986), Wrigley and Schofield, p. 79 and Marshall (1832), p. 67) and looked for correlations between the aggregate data and the London (corrected and uncorrected) data. We found significant correlations ranging from 0.8 to 0.9 for Burials/Deaths and 0.5 to 0.71 for Baptisms/Births. Overall, the claim by Wrigley-Schofield that “In general, years of heavy mortality outside London found some reflection within the capital” cannot be rejected (Wrigley and Schofield, 1981, p. 83). Still, every interpretation concerning the effects of urbanization has to be taken with the necessary precaution.

became relatively more important as highly virulent nutrition independent epidemics like plague and smallpox became less frequent (Antonovsky, 1967; Bengtsson and van Poppel, 2011).

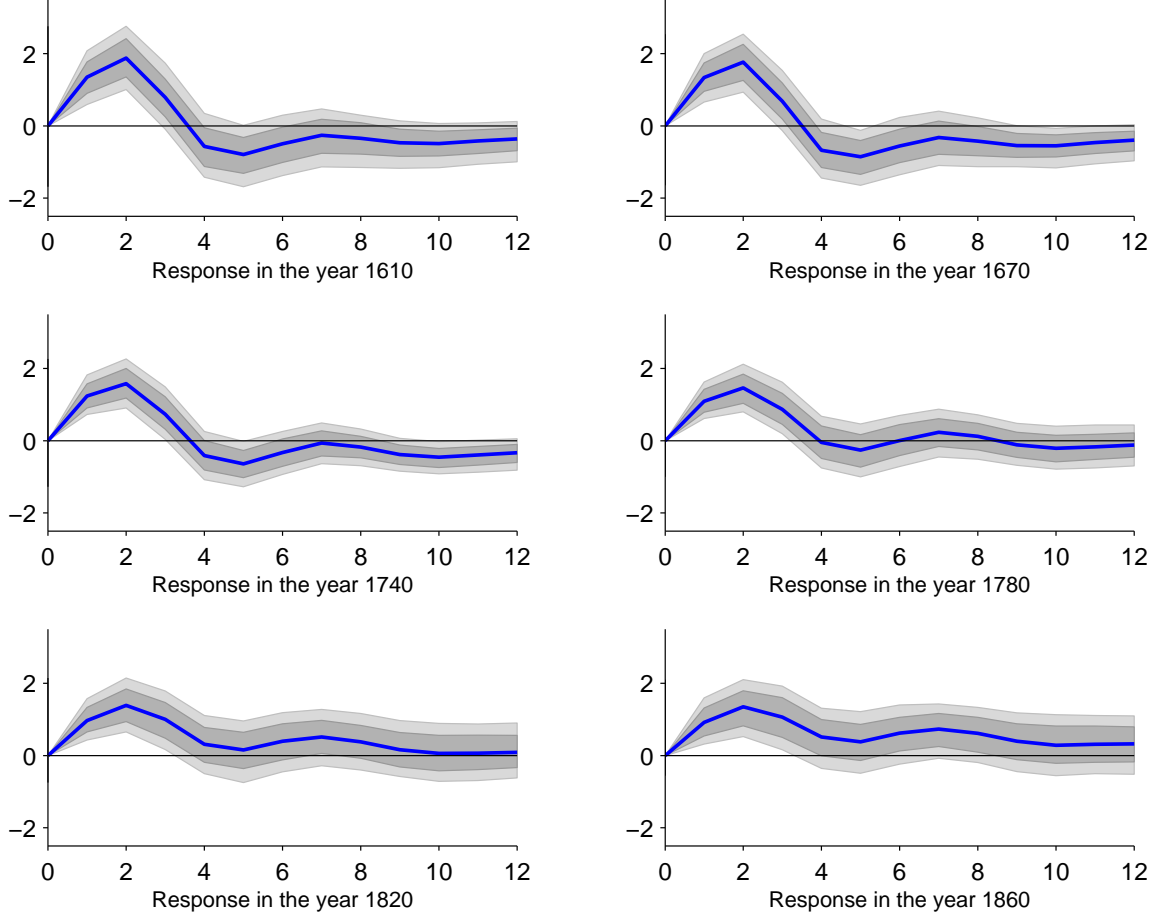
The increasing importance of wage-related income as a result of the sectoral shift and structural changes in the economy in combination with changes in the epidemiological environment provide a possible explanation for the intensification of the relationship between real wages and mortality during the period from 1750 to 1850. When examining disaggregated data from the 404 parishes of Wrigley and Schofield, Kelly and Ó Gráda (2010) find a negative correlation between real wages and mortality in parishes which had the highest share of manufacturing during the period from 1750 to 1800. Interestingly, the relative importance of the manufacturing sector made no major difference during the earlier period from 1700 to 1750. In this subsample, the strength of the correlation between real wages and mortality was found to be strong in most parishes.

How the standard of living and inequality evolved during the industrial revolution is a very contentious issue because different indicators give different answers depending on the specific environment and time under consideration; see, e.g. Hobsbawm (1962, Chapter 11), Lindert (1994, Chapter 14), and Voth (2004, Chapter 10). However, our estimates are in agreement with the evidence found in the biological standard-of-living literature, in which human stature is used to measure economic status. This proxy measure declines in the early stages of the industrial revolution.³⁸ Komlos (1998) refers to this phenomenon as the “early industrial growth puzzle”.³⁹

³⁸See, e.g. Komlos (1993, 1998), Nicholas and Steckel (1991), Johnson and Nicholas (1995), and Floud et al. (1990). A critical discussion can be found in Bodenhorn et al. (2013).

³⁹He concludes that “The human organism did not always thrive as well in its newly created socioeconomic environment as one might be led to believe on the basis of purchasing power at the aggregate level” (Komlos, 1998, p. 793).

Figure 6: Real Wage-Population Growth Link, selected years.



Notes: Posterior means of the impulse response to a standardized real wage shock. The gray area (dark gray) area indicates the 5th (16th) and 95th (84th) percentiles of the impulse responses.

5.4 The Real Wage-Population Growth Link

One major stylized fact modeled in Unified Growth Theories is that population growth increases in income up until the Post-Malthusian regime (Galor, 2011, Chapter 2). To gain insights into the real wage-population growth link, we combine the effects of real wages on mortality and fertility (see Figure 6). Before 1750 the impact of wages on mortality is rather unclear and population growth was rather driven by fertility. However, during the early phase

of industrialization the effect of wages on mortality starts to become more important. Thus, while the propagation mechanism of real wage shocks on population growth varied over time, the real wage-population growth link was always positive. When parents utility is defined over the number of surviving children population growth can be driven by either a rise in fertility or a decline in mortality.⁴⁰

5.5 Long-Run Values

At first, our findings may seem surprising. In a period in which mortality was decreasing and living standards were supposedly rising, we find that real wage shocks impacted stronger then ever before on mortality. To interpret these results in the right way, we have to consider that more pronounced impulse responses of mortality to a real wage innovations do not imply a high level of mortality or a low level of real wages per se. Instead, it indicates a strong reaction of mortality to exogenous real wage changes. The implied long-term values can be inferred from the infinite moving-average representation of the model for each point in time:⁴¹

$$\mathbf{y}_t = \boldsymbol{\mu}_t + \sum_{i=0}^{\infty} \mathbf{C}_{t,i} \boldsymbol{\epsilon}_{t-i},$$

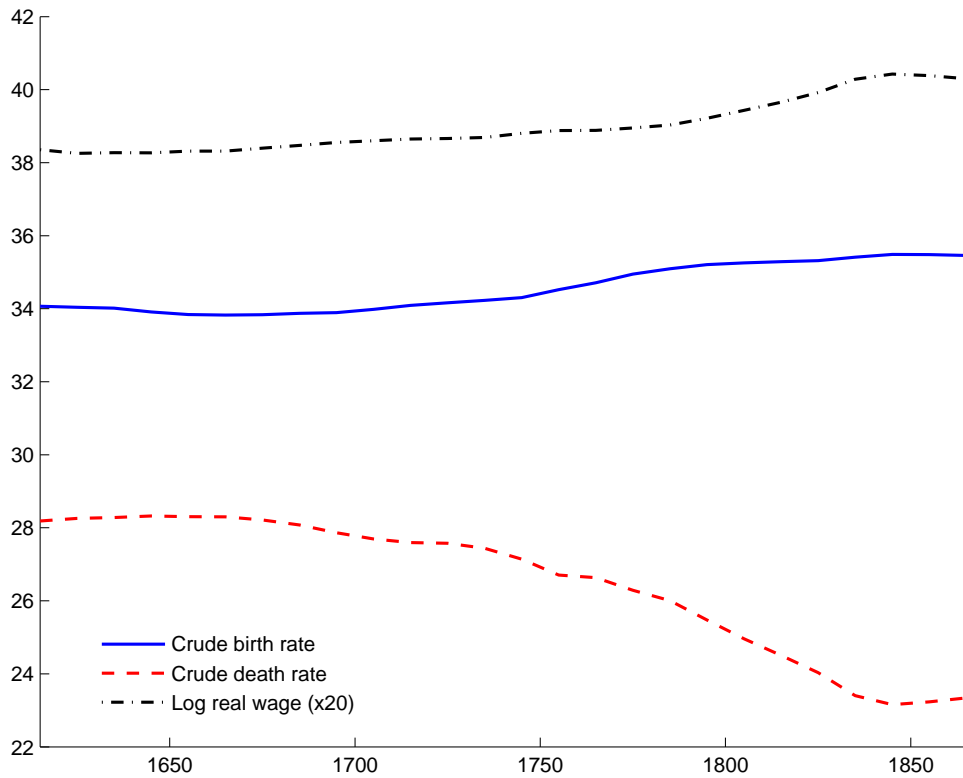
where $\mathbf{C}_{i,t}$ represents the impulse response matrices. Any changes in the level of the variables will be captured by changes in the constant $\boldsymbol{\mu}_t$, which reflects the long-run equilibrium levels of the variables. Figure 7 shows the median of these levels. As can be seen, the inferred long-run values do indeed imply a decline in mortality and an increase in real wages after 1800. Thus, a strong real wage-mortality link does not necessarily imply that the level of mortality is very high, but that some change in the structure of

⁴⁰We thank an associate editor for pointing this out.

⁴¹Note that $\mathbf{C}_{t,i} = \mathbf{J} \mathbf{A}_{t,0}^i \mathbf{J}' \mathbf{B}^{-1} \boldsymbol{\Sigma}$ and $\boldsymbol{\mu}_t = \mathbf{J}(\mathbf{I} - \mathbf{A}_{t,0})^{-1} \begin{pmatrix} \mathbf{c}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}$, where $\mathbf{A}_{t,0}$ is the companion matrix and $\mathbf{J} = [\mathbf{I}_n \ \mathbf{0}, \dots, \mathbf{0}]$.

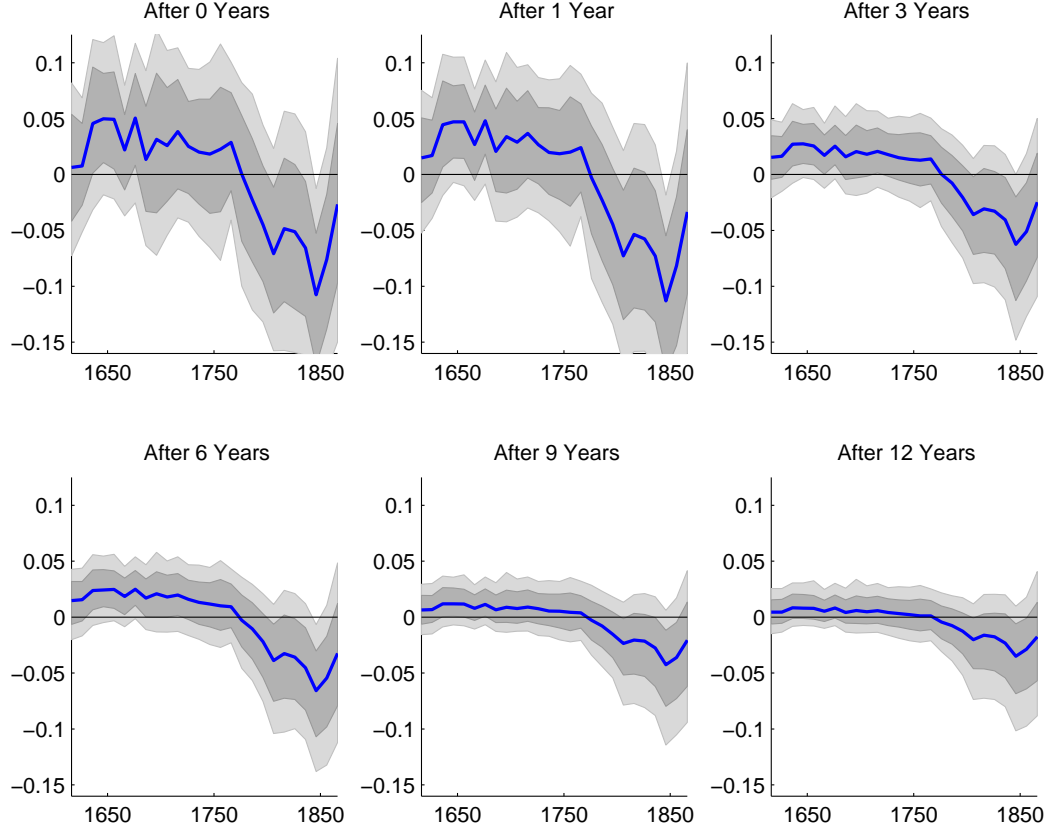
the economy (the transmission mechanism) made the population on average more susceptible to wage changes. The figure shows nicely how England was trapped in the Malthusian straitjacket up until the middle of the 18th century. Then the transition began; population growth accelerated, and the decrease in mortality was more pronounced than the increase in fertility. Equilibrium real wages also started to rise.

Figure 7: Implied Long-Run Values for Vital Rates and Real Wages, 1601–1870.



Notes: Median of the posterior distribution for $(\mathbf{I} - \mathbf{A}_{t,0})^{-1}\mathbf{c}_t$.

Figure 8: Response of real wage to a *CDR* shock, 1601-1870.



Notes: Posterior means of the impulse response of the real wage to a standardized shock to the death rate after a specified number of years for each decade from 1601 to 1870. The gray (dark gray) area indicate the 5th (16th) and 95th (84th) percentiles of the impulse responses.

5.6 The Feedback Effect

We next consider the feedback effect of population on wages. Malthusian theory predicts that real wages should decline with population increases.⁴² Figure 8 depicts the effect death rates on real wages. Based on the impact of a mortality shock at impact (after 0 years), an increase in real wages can be observed in the beginning of the 17th century. Around the time of the

⁴²Due to the higher dimensionality of the resulting system and identification problems we could not include population in our framework. See Lee (1987, 1993), and Lee and Anderson (2002) for an analysis of the direct link between population and wages.

English Civil War (1642-1651), the feedback effect is most pronounced, Afterwards, while there is some probability for the existence of the feedback effect up to 1770, after which the dynamics change fundamentally and no evidence for the existence can be found any more. A similar pattern is evident in the lower panel of Figure 4, which depicts the forecast error variance decompositions. While the explanatory power of the mortality shocks can be rather high (up to 60 percent) in the 17th century, it is smaller afterwards. After 1750 when there is clearly no evidence for a Malthusian feedback effect any more the explained variance increases again slightly.

Thus, the results provide some evidence that there was a positive effect of a smaller labor force on the real wage rate (e.g. Lee, 1973; Lee and Anderson, 2002) at least during the time up until mid eighteenth century. This early breakdown is also consistent with the findings of Nicolini (2007). Nevertheless, even in the beginning of the sample the zero line is often included.

Some caveats should be noted in conjunction with the interpretation of these results. Although an increase in the death rate reduces the population, it does not necessarily imply that the workforce is also decreasing. Perhaps only infant mortality is increasing, in which case there will be a very delayed impact on the labor supply. Furthermore, due to shifts in women's participation rates or migration, the relevant labor supply may stay constant or even increase causing a change in the sign of the impulse response from the late 18th century onward.

5.7 Discussion of Prior and Model Assumptions

In the following, we discuss our prior choice and lag specification.

5.7.1 Model Selection

To choose the appropriate lag specification for the time varying model, we use Bayesian model selection procedures. The time-varying VAR was estimated

Table 1: Log posterior odds ratios for different lag order

	p_1	p_2	p_3	p_4
p_1	0	837	-893	-2049
p_2	-837	0	-1730	-2886
p_3	893	1730	0	-1156
p_4	2049	2886	1156	0

The element in row i and column j represents the log posterior odds ratio of model p_i over model p_j .

for the lag lengths $p = 1$ up to $p = 4$. More specifically, we have computed the posterior odds ratio between two competing models.⁴³ From the comparison of the four competing model specifications the model with a lag length of 4 performed best in terms of its out of sample fit. Hence, the posterior odds strongly suggests to use $p = 4$ model.

5.7.2 Initial Sample

Because the crude death series contains some spikes, we made sure that our results were not influenced by potential outliers in the initial sample. We estimated various time invariant VAR models to calibrate the prior distributions. To cope with potential outliers caused primarily by epidemics and wars we used two alternative strategies. First, a dummy variable that is based on mortality crises as defined in Wrigley and Schofield (1981, p. 334) was used.⁴⁴ We also discarded the period with extreme observations and used the period from 1562–1600 instead. Using different priors elicited from the various specifications of the time invariant model did not change the impulse responses and forecast error variance decompositions of the time varying model sub-

⁴³The posterior odds ratios were calculated using the approach described in Chib (1995). Note that the Schwarz criterion gives a rough approximation of the posterior odds ratio with ex-ante equally likely models.

⁴⁴The dummy is active either as 3 star crisis dummy: 1557, 1558, 1559, 1625, 1638, 1639, 1658, 1665, 1681, 1727, 1728, 1729, 1730, 1742; or as 3 and 2 star crisis dummy: 1545, 1546, 1557, 1558, 1559, 1588, 1593, 1597, 1603, 1624, 1625, 1638, 1639, 1643, 1657, 1658, 1665, 1681, 1727, 1728, 1729, 1730, 1742, 1762, 1763.

stantially. All estimated specifications allow the same conclusions.⁴⁵

Increasing the variance of \mathbf{a}_0 by a factor of one hundred and the variance of \mathbf{b}_0 by factor of one thousand hence giving them less weight in the posterior did also not change the results.

5.7.3 The Choice of k_i , for $i = Q, W, S$

As has already been noted by Primiceri (2005), the choice of the values for k_i for $i = Q, W, S$ can be crucial to the results. These values reflect prior beliefs about the amount of time variation in the parameters.⁴⁶ For instance, even a small value of 0.001 for k_Q implies an average cumulative change of 341 percent for the elements of \mathbf{a}_t over the sample period. Hence our choice is a prior that does not particularly punish time variation of the parameters as it implies 100 times more time variation. For prior parameters favoring too much time variation, the model produces impulse responses that are explosive and have counterintuitive signs (Primiceri, 2005, pp. 841-843). We generated 1000 artificial data sets for $k_Q \in \{0.001, 0.01, 0.1\}$, and we compared the implied number of explosives draws. When increasing k_Q from 0.01 to 0.1 the share of explosive draws increased from 4 percent to 98 percent making meaningful estimation impossible.

The k parameters were hence calibrated to values that avoid explosive behavior for the demographic variables, while still allowing for as much time variation as possible. The main conclusion remains unchanged when k_S and k_W were set to 0.1.

A comparison of the prior with the posterior distributions of the traces of **Q**, **W**, and **S** can be found in Figure 11 in Appendix B.

⁴⁵The results for the time-invariant models and corresponding time-varying models are available upon request.

⁴⁶However, the prior choice for the covariance matrices is rather uninformative. Note that the degrees of freedom parameters determine the weight in the posterior. Loosely speaking we can think of an additional number of observations equal to the degrees of freedom.

5.7.4 Alternative Real Wage Series

We also used an alternative real wage series that was provided in Clark (2007b) which is based on English farm workers. The time series constructed in Clark (2007b) and the one constructed in Allen (2001) mainly differ in the period after 1800. While the Clark series implies a more optimistic view of the period of the Industrial Revolution, the Allen series is rather pessimistic. However, our main conclusions are similar when using the Clark real wage series.

When using the real wage series constructed by Clark (2007b) we find similar patterns. In the benchmark specification, the real wage-fertility link is, according to the forecast error variance decompositions, present throughout the whole sample (see Figure 13 in Appendix D). Similar to the results obtained with the Allen series, the impulse responses change over time (see Figure 12 in Appendix D). However, the changes we observe using the Clark series are much more gradual.

Regarding the real wage-mortality link, the impulse responses show that the real wage-mortality link was already present during the 17th century. However, the forecast error variance decompositions imply again that the explanatory power of real wage shocks is negligible up until mid 18th century. Similar to the results obtained using the Allen series the link between mortality and real wages strengthened during the 18th and 19th century (see Figure 13 in Appendix D). With the exception that the transformation with the Clark series reaches its peak already earlier.

6 Conclusion

“Here,, one of the most important advantages of the availability of long runs of data lay in the opportunity they afforded to go beyond the discovery of general relationships between two or more variables to a study of how the relationship changed over time.” (Wrigley and Schofield, 1981, p. 9)

This paper has documented the changes in the interaction between vital rates and real wages during the period from 1541 to 1870 in England. The time-varying VAR approach takes into account the endogeneity of the variables and the drifts and volatilities contained in the data. Hence, we are able to show how secular changes affect the interaction between mortality/fertility and real wages over this time period.

The results underline the importance of time variation. We confirm the apparent changes in the volatility of the birth and death rate series by calculating the stochastic volatilities. We find considerable changes in the real wage-fertility and real wage-mortality linkages over time.

How do the empirical results relate to Unified Growth Theory? One of the main assumptions of UGT is that a rise in income translates into higher population growth before the demographic transition. This is clearly supported by our findings. When decomposing the effect of real wages on population growth into effects on mortality and fertility we find that before 1750 the impact of wages on mortality is negligible and population growth is rather driven by fertility. However, this changes in the initial phase of the industrialization when the effect of wages on mortality becomes more important. For the Malthusian model also the negative feedback loop of population on income is important. We also provide some evidence that the ‘iron law’ started to crumble after 1750.

Although our findings are mostly consistent with the findings present in the previous empirical literature, there is one major exception. The evidence

that mortality reacted most strongly to real wages during the initial phase of industrialization. This result indicates a substantially different relationship between real wages and mortality during the period of industrialization than has been indicated by previous econometric studies. However, this finding is consistent with historical evidence and with the findings of the biological standard-of-living literature.

That a substantial real wage-mortality link emerged during the time of transition when living standards were rising and mortality was decreasing is a fascinating finding. One interesting question is whether this is only a reflection of the sectoral change taking place during that time or if the strengthening of the real wage-mortality link could also reflect an underlying cause for the transition. This question is left open for future research.

References

- Allen, Robert C.**, “The Great Divergence in European Wages and Prices from the Middle Ages to the First World War,” *Explorations in Economic History*, October 2001, 38 (4), 411–447.
- , “Agriculture During the Industrial Revolution, 1700–1850,” in Roderick Floud and Paul Johnson, eds., *The Cambridge Economic History of Modern Britain: Industrialisation, 1700–1860*, Vol. 1, Cambridge: Cambridge University Press, 2004, pp. 268–294.
- , *The British Industrial Revolution in Global Perspective*, Cambridge: Cambridge University Press, 2009.
- Antonovsky, Aaron**, “Social Class, Life Expectancy and Overall Mortality,” *The Milbank Memorial Fund Quarterly*, 1967, 45 (2), pp. 31–73.
- Artzrouni, Marc and John Komlos**, “Population Growth Through History and the Escape from the Malthusian Trap: A Homeostatic Simulation Model,” *Genus*, 1985, 41, 21–39.
- Ashraf, Quamrul and Oded Galor**, “Dynamics and Stagnation in the Malthusian Epoch: Theory and Evidence,” Working Papers 2008-14, Brown University, Department of Economics 2008.
- Bailey, R. E. and M. J. Chambers**, “Long-Term Demographic Interactions in Precensus England,” *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 1993, 156 (3), pp. 339–362.
- Bengtsson, Tommy and Frans van Poppel**, “Socioeconomic Inequalities in Death From Past to Present: An Introduction,” *Explorations in Economic History*, 2011, 48 (3), 343 – 356.
- Bodenhorn, Howard, Timothy W. Guinnane, and Thomas A. Mroz**, “Problems of Sample-selection Bias in the Historical Heights Literature: A Theoretical and Econometric Analysis,” Working Papers 1023, Economic Growth Center, Yale University May 2013.

- Bongaarts, John**, “A Framework for Analyzing the Proximate Determinants of Fertility,” *Population and Development Review*, 1978, 4, 105–132.
- Broadberry, Stephen**, “Recent Developments in the Theory of Very Long Run Growth: A Historical Appraisal,” The Warwick Economics Research Paper Series (TWERPS) 818, University of Warwick, Department of Economics 2007.
- Carter, C. K. and R. Kohn**, “On Gibbs Sampling for State Space Models,” *Biometrika*, 1994, 81, 541–553.
- Cervellati, Matteo and Uwe Sunde**, “Human Capital Formation, Life Expectancy, and the Process of Development,” *American Economic Review*, December 2005, 95 (5), 1653–1672.
- and —, “Health, Development, and the Demographic Transition,” 2006 Meeting Papers 645, Society for Economic Dynamics December 2006.
- Chib, Siddhartha**, “Marginal Likelihood From the Gibbs Output,” *Journal of the American Statistical Association*, 1995, 90, 1313–1321.
- Clark, Gregory**, “The Condition of the Working Class in England, 1209–2004,” *Journal of Political Economy*, December 2005, 113 (6), 1307–1340.
- , *A Farewell to Alms: A Brief Economic History of the World*, Princeton, NJ: Princeton University Press, 2007.
- , “The Long March of History: Farm Wages, Population, and Economic Growth, England 1209–1869,” *Economic History Review*, 2007, 60, 97–135.
- Cogley, Timothy and Thomas J. Sargent**, “Evolving Post-World War II U.S. Inflation Dynamics,” in “NBER Macroeconomics Annual,” Cambridge, Massachusetts: MIT Press, 2001.
- and —, “Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S,” *Review of Economic Dynamics*, April 2005, 8 (2), 262–302.

- Crafts, Nicholas and Terence C. Mills**, “From Malthus to Solow: How Did the Malthusian Economy Really Evolve?,” *Journal of Macroeconomics*, March 2009, 31 (1), 68–93.
- Deane, Phyllis and W. A. Cole**, *British Economic Growth 1688–1959*, 2nd ed., Cambridge University Press, 1969.
- Del Negro, Marco and Giorgio Primiceri**, “Time-varying Structural Vector Autoregressions and Monetary Policy: a Corrigendum,” Staff Reports 619, Federal Reserve Bank of New York 2013.
- Doepke, Matthias**, “Accounting for Fertility Decline During the Transition to Growth,” *Journal of Economic Growth*, 09 2004, 9 (3), 347–383.
- **and Fabrizio Zilibotti**, “Occupational Choice and the Spirit of Capitalism,” *The Quarterly Journal of Economics*, 2008, 123, 747–793.
- Floud, Roderick, Kenneth Wachter, and Annabel Gregory**, *Height, Health and History: Nutritional Status in the United Kingdom, 1750–1980*, Cambridge: Cambridge University Press, 1990.
- Fuller, Wayne A.**, *Introduction to Statistical Time Series*, NJ: John Wiley & Sons, Inc., 1996.
- Galloway, Patrick R.**, “Basic Patterns in Annual Variations in Fertility, Nupiality, Mortality, and Prices in Pre-Industrial Europe,” *Population Studies*, 1988, 42, 275–303.
- Galor, Oded**, “From Stagnation to Growth: Unified Growth Theory,” in Aghion. Philippe and Steven N. Durlauf, eds., *Handbook of Economic Growth 1A*, Vol. 1A, Elsevier B. V., 2005, chapter 4, pp. 171–292.
- , “2008 Lawrence R. Klein Lecture – Comparative Economic Development: Insights from Unified Growth Theory,” Working Papers 2009-10, Brown University, Department of Economics 2009.
- , *Unified Growth Theory*, Princeton & Oxford: Princeton University Press, 2011.

- **and David N. Weil**, “From Malthusian Stagnation to Modern Growth,” *American Economic Review*, 1999, *89* (2), 150–154.
 - **and —** , “Population, Technology, and Growth: From Malthusian Stagnation to Demographic Transition and Beyond,” *American Economic Review*, 2000, *90*, 806–828.
 - **and Omer Moav**, “Natural Selection and the Origin of Economic Growth,” *The Quarterly Journal of Economics*, November 2002, *117* (4), 1133–1191.
- Gelfand, Alan E. and Adrian F. M. Smith**, “Sampling-Based Approaches to Calculating Marginal Densities,” *Journal of the American Statistical Association*, 1990, *85*, 398–409.
- Geman, Donald and Stuart Geman**, “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984, *PAMI-6*, 721–741.
- Geweke, John**, “Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication,” *Econometric Reviews*, 1999, *18* (1), 1–73.
- Hajnal, John**, “European Marriage Pattern in Historical Perspective,” in D V Glass and D E Eversley, eds., *Population in History: Essays in Historical Demography*, Chicago: Aldine Publishing Company, 1965, pp. 101–143.
- Hansen, Gary and Edward Prescott**, “Malthus to Solow,” *American Economic Review*, 2002, *92*, 1205–1217.
- Hobsbawm, Eric**, *The Age of Revolution: 1789–1848*, London: Weidenfeld and Nicolson, 1962.
- Hull, Charles Henry, ed.**, *The Economic Writings of Sir William Petty*, Fairfield NJ: Augustus M. Kelley, 1986.

- Johnson, Paul and Stephen Nicholas**, “Male and Female Living Standards in England and Wales, 1812-1857: Evidence From Criminal Height Records,” *The Economic History Review*, 1995, 48, 470–481.
- Jones, Charles**, “Was an Industrial Revolution Inevitable? Economic Growth Over the Very Long Run,” *Advances in Macroeconomics*, 2001, 1, 1028–1028.
- Kalemli-Ozcan, Sebnem**, “Does the Mortality Decline Promote Economic Growth?,” *Journal of Economic Growth*, December 2002, 7 (4), 411–39.
- Kelly, Morgan and Cormac Ó Gráda**, “Living Standards and Mortality since the Middle Ages,” Working Papers 201026, School of Economics, University College Dublin 2010.
- **and** —, “The Preventive Check in Medieval and Pre-industrial England,” Working Papers 201110, School Of Economics, University College Dublin May 2011.
- Kim, S., N. Shephard, and S. Chib**, “Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models,” *Review of Economic Studies*, 1998, 65, 361–393.
- Komlos, John**, “The Secular Trend in the Biological Standard of Living in the United Kingdom, 1730-1860,” *The Economic History Review*, 1993, 46, 115–144.
- , “Shrinking in a Growing Economy? The Mystery of Physical Stature during the Industrial Revolution,” *The Journal of Economic History*, September 1998, 58 (03), 779–802.
- **and Marc Artzrouni**, “Mathematical Investigations of the Escape from the Malthusian Trap,” *Mathematical Population Studies*, 1990, 2, 269–287.
- Kremer, Michael**, “Population Growth and Technological Change: One Million B.C. To 1990,” *Quarterly Journal of Economics*, 1993, 108, 681–716.

- Lagerlöf, Nils-Petter**, “From Malthus to Modern Growth: Can Epidemics Explain the Three Regimes?,” *International Economic Review*, 2003, 44, 755–777.
- , “The Galor-Weil Model Revisited: A Quantitative Exercise,” *Review of Economic Dynamics*, January 2006, 9 (1), 116–142.
- Lee, Ronald D.**, “Population in Preindustrial England: An Econometric Analysis,” *The Quarterly Journal of Economics*, 1973, 87 (4), 581–607.
- , “Natural Fertility, Population Cycles and the Spectral Analysis of Births and Marriages,” *Journal of the American Statistical Association*, 1975, 70 (350), 295–304.
- , “Short Term Variation: Vital Rates, Prices and Weather,” in Edward A. Wrigley and Roger S. Schofield, eds., *The Population History of England, 1541-1871: A Reconstruction*, London: Edward Arnold, 1981, pp. 285–355.
- , “Population Dynamics of Humans and Other Animals,” *Demography*, 1987, 24 (4), 443–465.
- , “Accidental and Systematic Change in Population History: Homeostasis in a Stochastic Setting,” *Explorations in Economic History*, 1993, 30 (1), 1–30.
- **and Michael Anderson**, “Malthus in State Space: Macro Economic-Demographic Relations in English History, 1540 to 1870,” *Journal of Population Economics*, 2002, 15 (2), 195–220.
- Lindert, Peter H.**, “English living standards, population growth, and Wrigley-Schofield,” *Explorations in Economic History*, April 1983, 20 (2), 131–155.
- , “Unequal Living Standards,” in Roderick Floud and Donald McCloskey, eds., *The Economic History of Britain Since 1700: 1700–1860*, Vol. 1, Cambridge: Cambridge University Press, 1994, pp. 357–386.

- Lucas, Robert E.**, “The Industrial Revolution: Past and Future,” in “Lectures on Economic Growth,” Cambridge, Massachusetts, and London, England: Harvard University Press, 2002, pp. 109–188.
- Malthus, Thomas Robert**, *An Essay on the Principle of Population*, St. Paul’s Church-Yard, London: J. Johnson, 1798.
- Marshall, John**, *Mortality of the Metropolis* 1832.
- Mitchell, Brian R.**, *International Historical Statistics: Europe 1750–1975*, NY: Palgrave Macmillan, 2003.
- Mokyr, Joel and Hans-Joachim Voth**, “Understanding Growth in Europe, 1700-1870: Theory and Evidence,” mimeo 2007.
- Nicholas, Stephen and Richard H. Steckel**, “Heights and Living Standards of English Workers During the Early Years of Industrialization, 1770–1815,” *Journal of Economic History*, 1991, 51, 937–957.
- Nicolini, Esteban A.**, “Was Malthus Right? A Var Analysis of Economic and Demographic Interactions in Pre-Industrial England,” *European Review of Economic History*, 2007, 11, 99–121.
- Primiceri, Giorgio E.**, “Time Varying Structural Vector Autoregressions and Monetary Policy,” *Review of Economic Studies*, 2005, 72, 821–852.
- Razzell, Peter**, *Essays in English Population History*, London: Caliban Books, 1994.
- Reichmuth, Wolfgang**, “Malthus in the Nordic Countries ? A Bayesian VAR Analysis of Economic-Demographic Interactions in the 18th and 19th Century,” 2008. Humbolt-University Berlin, mimeo.
- Sims, Christopher A.**, “Macroeconomics and Reality,” *Econometrica*, 1980, 48, 1–48.
- Tamura, Robert**, “Human Capital and the Switch from Agriculture to Industry,” *Journal of Economic Dynamics and Control*, 2002, 27, 207–242.

- Voigtländer, Nico and Hans-Joachim Voth**, “Malthusian Dynamics and the Rise of Europe: Make War, Not Love,” *American Economic Review: Papers & Proceedings*, 2009, 99, 248–254.
- **and** – , “The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe,” *The Review of Economic Studies*, 2012.
- **and** – , “How the West ”Invented” Fertility Restriction,” *American Economic Review*, 2013, 103 (6), 2227–64.
- Voth, Hans-Joachim**, “Living Standards and the Urban Environment,” in Roderick Floud and Paul Johnson, eds., *The Cambridge Economic History of Modern Britain: Industrialisation, 1700–1860*, Vol. 1, Cambridge: Cambridge University Press, 2004, pp. 268–294.
- Woods, Robert**, *The Demography of Victorian England and Wales*, Cambridge: Cambridge University Press, 2000.
- Wrigley, Edward A. and Roger S. Schofield**, *The Population History of England, 1541-1871: A Reconstruction*, London: Edward Arnold, 1981.
- , **R. S. Davis, James E. Oeppen, and Roger S. Schofield**, *English Population History from Family Reconstitution 1580-1837*, Cambridge: Cambridge University Press, 1997.

A Appendix

A.1 The Gibbs Sampling Procedure

The main objective of Bayesian econometrics is to characterize properties of posterior distributions. The Gibbs sampler (Geman and Geman, 1984; Gelfand and Smith, 1990) allows one to generate draws from posterior distributions, even when the joint posterior distribution is of a non-standard form. It subdivides the joint posterior distribution into a complete set of standard conditional distributions. By iteratively drawing from these conditional distributions, the sequence of draws converges to draws from the unknown true joint distribution. The posterior characteristics can then be summarized using graphical representations of the draws and estimating expectations of integrable functions by appropriate sample averages of these functions.

In general, the Gibbs sampler proceeds as follows: assume that Ψ contains all parameters of the model and \mathbf{U} all latent variables.⁴⁷ Given some arbitrary chosen initial values $\bar{\mathbf{U}}$, the sampler can be started. The first draw for Ψ conditional on the starting value for \mathbf{U} is generated from the conditional density $p(\Psi|\bar{\mathbf{U}})$. The draw $\Psi^{(1)}$ can in turn be used to generate the first draw for \mathbf{U} from the conditional density $p(\mathbf{U}|\Psi^{(1)})$, where the superscripts in parenthesis denote draws from the conditional distributions. Repeating the steps indicated above using in each case the calculated draws from the previous cycle, $\Psi^{(w)}$ is drawn from $p(\Psi|\mathbf{U}^{(w-1)})$ and $\mathbf{U}^{(w)}$ from $p(\mathbf{U}|\Psi^{(w)})$. Thus, it is possible to generate the Gibbs sequence $\{\Psi^{(w)}, \mathbf{U}^{(w)}\}$. Under mild conditions, the Gibbs sequence converges (in distribution) to the true joint distribution at a geometric rate in w (Geman and Geman, 1984). Note that in general, $\Psi^{(w)}$ and $\mathbf{U}^{(w)}$ can themselves be Gibbs sequences, cycling over the parameter space and latent variables, respectively. In our case, Ψ consists of the covariance matrix \mathbf{V} and \mathbf{U} contains all time-varying parameters: \mathbf{a}^T , \mathbf{b}^T , and \mathbf{h}^T . The following gives an overview over the entire Gibbs sampling procedure suggested by Primiceri (2005) that was corrected in Del

⁴⁷In a simpler model without latent variables, one can think of Ψ and \mathbf{U} as representing the parameters split into two appropriate blocks.

Negro and Primiceri (2013) and then describes the individual steps in more detail.

A.1.1 Overview over the Gibbs Sampler

Gibbs sampling involves the following steps⁴⁸

- I Initialize $\mathbf{b}^T, \Sigma^T, \mathbf{s}^T, \mathbf{V}$.
- II Sample \mathbf{a}^T from $p(\mathbf{a}^T | \mathbf{y}^T, \mathbf{x}^T, \mathbf{b}^T, \Sigma^T, \mathbf{V})$.
- III Sample \mathbf{b}^T from $p(\mathbf{b}^T | \mathbf{y}^T, \mathbf{x}^T, \mathbf{a}^T, \Sigma^T, \mathbf{V})$.
- IV Sample \mathbf{s}^T from $p(\mathbf{s}^T | \mathbf{y}^T, \mathbf{x}^T, \mathbf{a}^T, \mathbf{b}^T, \mathbf{V})$.
- V Sample Σ^T from $p(\Sigma^T | \mathbf{y}^T, \mathbf{x}^T, \mathbf{a}^T, \mathbf{b}^T, \mathbf{s}^T, \mathbf{V})$.
- VI Sample \mathbf{V} , by sampling \mathbf{Q}, \mathbf{W} and \mathbf{S} from $p(\mathbf{Q}, \mathbf{W}, \mathbf{S} | \mathbf{y}^T, \mathbf{x}^T, \mathbf{A}^T, \mathbf{B}^T, \Sigma^T) = p(\mathbf{Q} | \mathbf{y}^T, \mathbf{x}^T, \mathbf{A}^T, \mathbf{B}^T, \Sigma^T) p(\mathbf{W} | \mathbf{y}^T, \mathbf{x}^T, \mathbf{A}^T, \mathbf{B}^T, \Sigma^T) \dots p(\mathbf{S}_1 | \mathbf{y}^T, \mathbf{x}^T, \mathbf{A}^T, \mathbf{B}^T, \Sigma^T) p(\mathbf{S}_{n-1} | \mathbf{y}^T, \mathbf{x}^T, \mathbf{A}^T, \mathbf{B}^T, \Sigma^T)$.
- VII Go back to II.

A.1.2 Step II: Drawing VAR Parameters \mathbf{a}^T

For given \mathbf{b}^T, Σ^T and \mathbf{V} , equations (3) together with (4) constitute the following linear Gaussian state-space system

$$\begin{aligned}\mathbf{y}_t &= (\mathbf{Z}'_{t-1} \otimes \mathbf{I}_n) \mathbf{a}_t + \mathbf{u}_t, \\ \mathbf{a}_t &= \mathbf{a}_{t-1} + \boldsymbol{\nu}_t.\end{aligned}$$

Therefore, the sampling procedure of Carter and Kohn (1994) can be applied to generate \mathbf{a}^T .

⁴⁸See Del Negro and Primiceri (2013)

A.1.3 Step III: Drawing Covariance States \mathbf{b}^T

For a given \mathbf{a}^T , it is possible to observe $\hat{\mathbf{y}}_t$, which is defined as

$$\hat{\mathbf{y}}_t = \mathbf{y}_t - (\mathbf{Z}'_{t-1} \otimes \mathbf{I}_n) \mathbf{a}_t.$$

Note that (3) can be expressed as

$$\mathbf{B}_t \hat{\mathbf{y}}_t = \Sigma_t \boldsymbol{\epsilon}_t. \quad (7)$$

Due to the diagonal structure of \mathbf{B}_t with ones on the main diagonal, the left hand-side of (7) can be written as $\hat{\mathbf{y}}_t + \tilde{\mathbf{L}}_t \mathbf{b}_t$, where $\tilde{\mathbf{L}}_t$ is given by

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 \\ \hat{y}_{1,t} & 0 & \cdots & \vdots \\ 0 & (\hat{y}_{1,t} \ \hat{y}_{2,t}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\hat{y}_{1,t} \ \hat{y}_{2,t} \ \cdots \ \hat{y}_{n,t}) \end{pmatrix}.$$

Therefore, the system of equations (7) can be rewritten as

$$\hat{\mathbf{y}}_t = \mathbf{L}_t \mathbf{b}_t + \Sigma_t \boldsymbol{\epsilon}_t,$$

where $\mathbf{L}_t = -\tilde{\mathbf{L}}_t$. Together with $\mathbf{b}_t = \mathbf{b}_{t-1} + \boldsymbol{\xi}_t$ this system has a state-space representation of this system that is Gaussian, but nonlinear since some of the dependent variables in the measurement equation also appear on the right-hand side. Because of the block diagonal structure of \mathbf{S} , this problem can easily be solved by applying the sampling procedure of Carter and Kohn (1994) equation by equation, starting with the second equation. Note that in that case all variables appearing on the right-hand side of the i th equation include only $\hat{y}_{k,t}$, $k < i$ which can be treated as predetermined due to the triangular structure.

A.1.4 Steps IV and V: Drawing Volatility States

Taking \mathbf{B}^T and \mathbf{a}^T as given, it is possible to observe \mathbf{y}_t^* , which is defined as

$$\mathbf{y}_t^* = \mathbf{B}_t(\mathbf{y}_t - (\mathbf{Z}'_{t-1} \otimes \mathbf{I}_n)\mathbf{a}_t) = \boldsymbol{\Sigma}_t \boldsymbol{\epsilon}_t. \quad (8)$$

This is a system of nonlinear measurement equations. It can be linearized by squaring and taking logarithms of each equation

$$\log (y_{i,t}^*)^2 = 2 \log \sigma_{i,t} + \log \epsilon_{i,t}^2, \quad i = 1, \dots, n.$$

As $(y_{i,t}^*)^2$ can be very small, an offset constant is introduced to make the estimation procedure more robust. This results in the following approximating linear state-space form

$$\begin{aligned} \mathbf{y}_t^{**} &= 2\mathbf{h}_t + \mathbf{e}_t, \\ \mathbf{h}_t &= \mathbf{h}_{t-1} + \boldsymbol{\eta}_t, \end{aligned} \quad (9)$$

where $y_{i,t}^{**} = \log[(y_{i,t}^*)^2 + \bar{c}]$, $e_{i,t} = \log \epsilon_{i,t}^2$, $h_{i,t} = \log \sigma_{i,t}$. The offset constant \bar{c} was introduced by (Fuller, 1996, pp. 494-7) and is set to 0.001. Although the representation is linear, it is not Gaussian, as the innovations in the measurement equation are distributed as $\log \chi(1)^2$. Since the covariance matrix of $\boldsymbol{\epsilon}_t$ is diagonal, the same is true for the covariance matrix of \mathbf{e}_t . This means that the Gaussian representation of the system in (9) can be found by approximating each element of \mathbf{e}_t by a mixture of normal densities as shown by Kim et al. (1998) for the univariate case. They match a number of moments of the $\log \chi(1)^2$ distribution using a mixture of seven normal densities with component probability q_j , and means m_j and variance v_j^2 , $j = 1, \dots, 7$, as tabulated in Table 2. Hence, each element of \mathbf{e}_t can be approximated as

$$f(e_{it}) \approx \sum_{j=1}^7 q_j f_N((e_{it}|m_j - 1.2704, v_j^2)).$$

Table 2: Selection of Mixing Distributions

ω	$q_j = Pr(\omega = j)$	m_j	v_j^2
1	0.00730	-10.12999	5.79596
2	0.10556	-3.97281	2.61369
3	0.00002	-8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	-1.08819	1.26261

Source: Kim et al. (1998).

An alternative way to express this is

$$\begin{aligned} e_{i,t}|s_{i,t} = j &\sim N(m_j - 1.2704, v_j^2), \\ Pr(s_{i,t} = j) &= q_j, \end{aligned} \tag{10}$$

where $\mathbf{s}^T = [\mathbf{s}_1, \dots, \mathbf{s}_T]$ is a matrix of unobserved indicator states $s_{i,t} \in 1, \dots, 7$, selecting at every period which member of the normal distribution mixture is used for the approximation of each element in \mathbf{e}_t .

Conditional on \mathbf{y}^{**T} and \mathbf{h}^T , it is possible to sample the indicator states \mathbf{s}^T . This is done by independently drawing each $s_{i,t}$ from the probability mass function defined by

$$Pr(s_{i,t} = j|y_{i,t}^{**}, h_{i,t}) \propto q_j f_N(y_{i,t}^{**}|2h_{i,t} + m_j - 1.2704, v_j^2),$$

with $j = 1, \dots, 7$, $i = 1, \dots, n$, and $t = 1, \dots, T$. Using the normal approximation to the $\log \chi(1)^2$ innovations transforms the system in (9) in a linear and Gaussian one, making the sampling algorithm of Carter and Kohn (1994) again applicable.

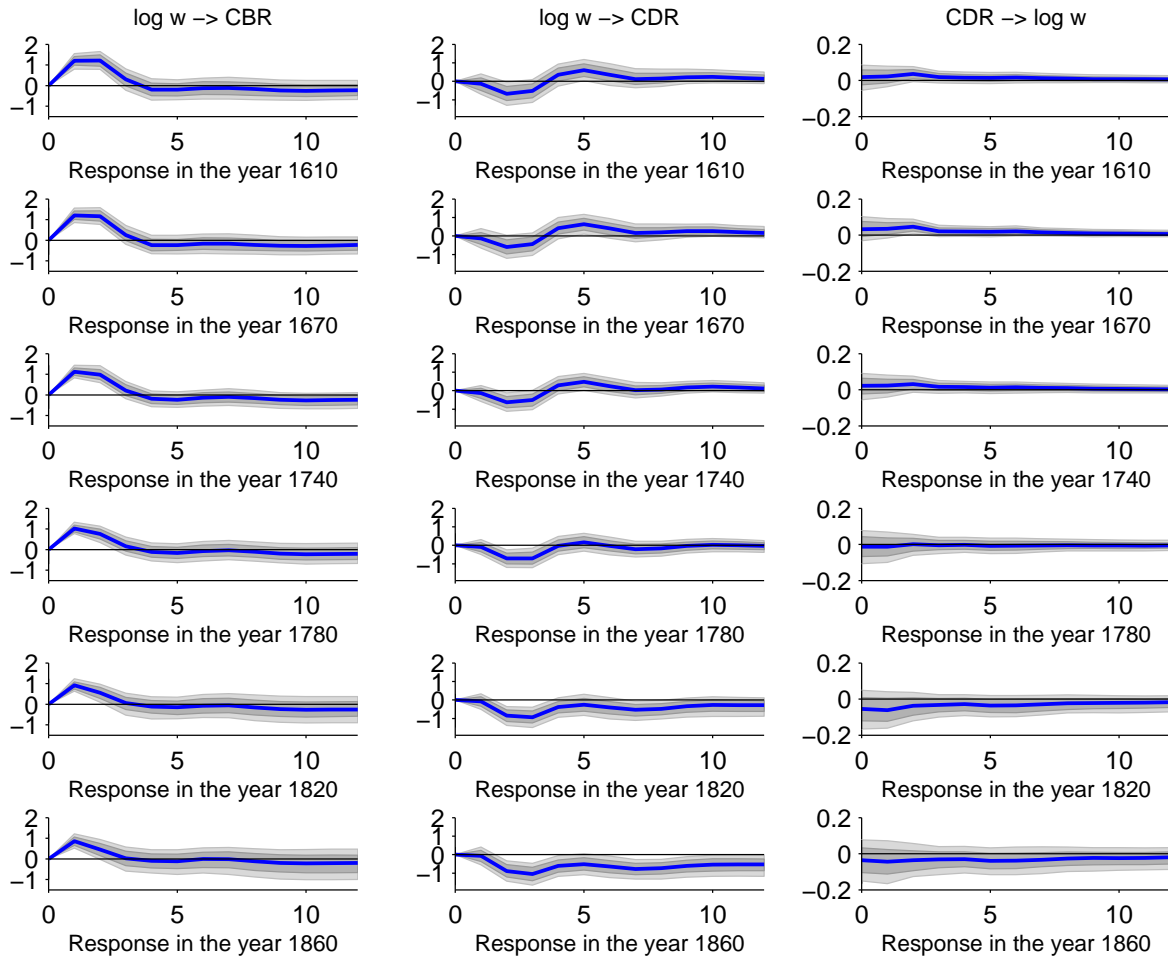
A.1.5 Step VI: Sampling Covariances \mathbf{V}

Conditional on \mathbf{a}^T , \mathbf{b}^T and \mathbf{h}^T all innovations in the transition equations $\boldsymbol{\nu}_t$, $\boldsymbol{\xi}_t$ and $\boldsymbol{\eta}_t$ are observable. With inverse Wishart priors, the conditional

posteriors of \mathbf{Q} , \mathbf{W} and the diagonal blocks of \mathbf{S} are also inverse Wishart.

B Additional Figures

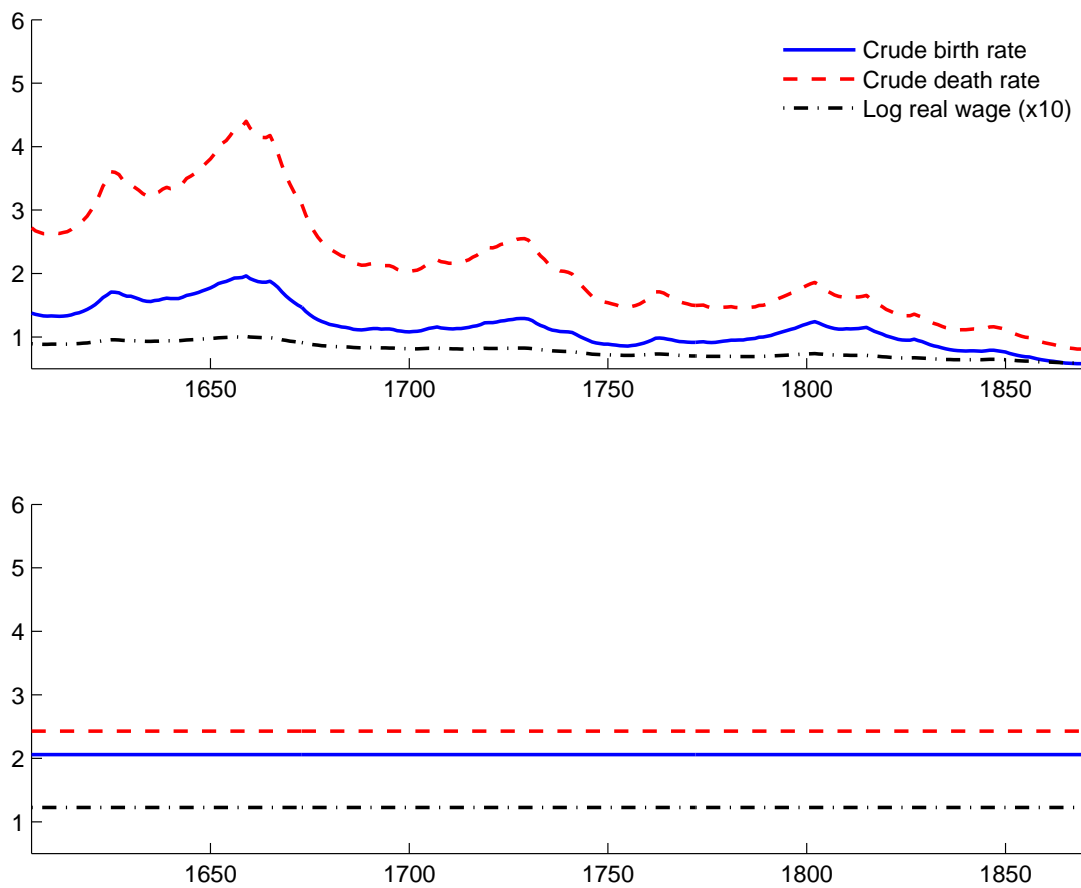
Figure 9: Impulse Responses for Various Years



Notes: Posterior means of the impulse response to standardized shocks. The gray (dark gray) area indicates the 5th (16th) and 95th (84th) percentiles of the impulse responses.

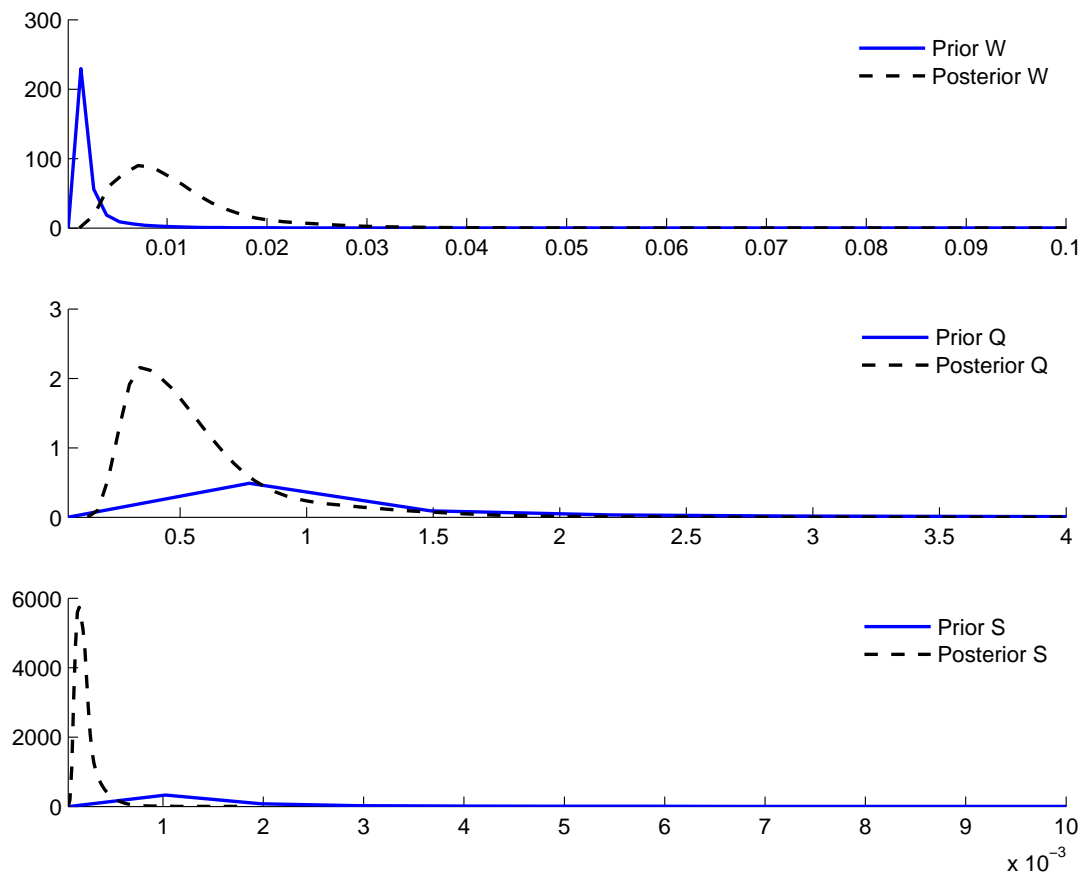
C Prior vs. Posterior

Figure 10: Estimated Volatilities: Prior vs. Posterior



Notes: The graph shows prior and posterior means of the estimated standard deviation of the structural shocks. For clearer presentation the standard deviation of the real wage shock was rescaled by a factor of 10.

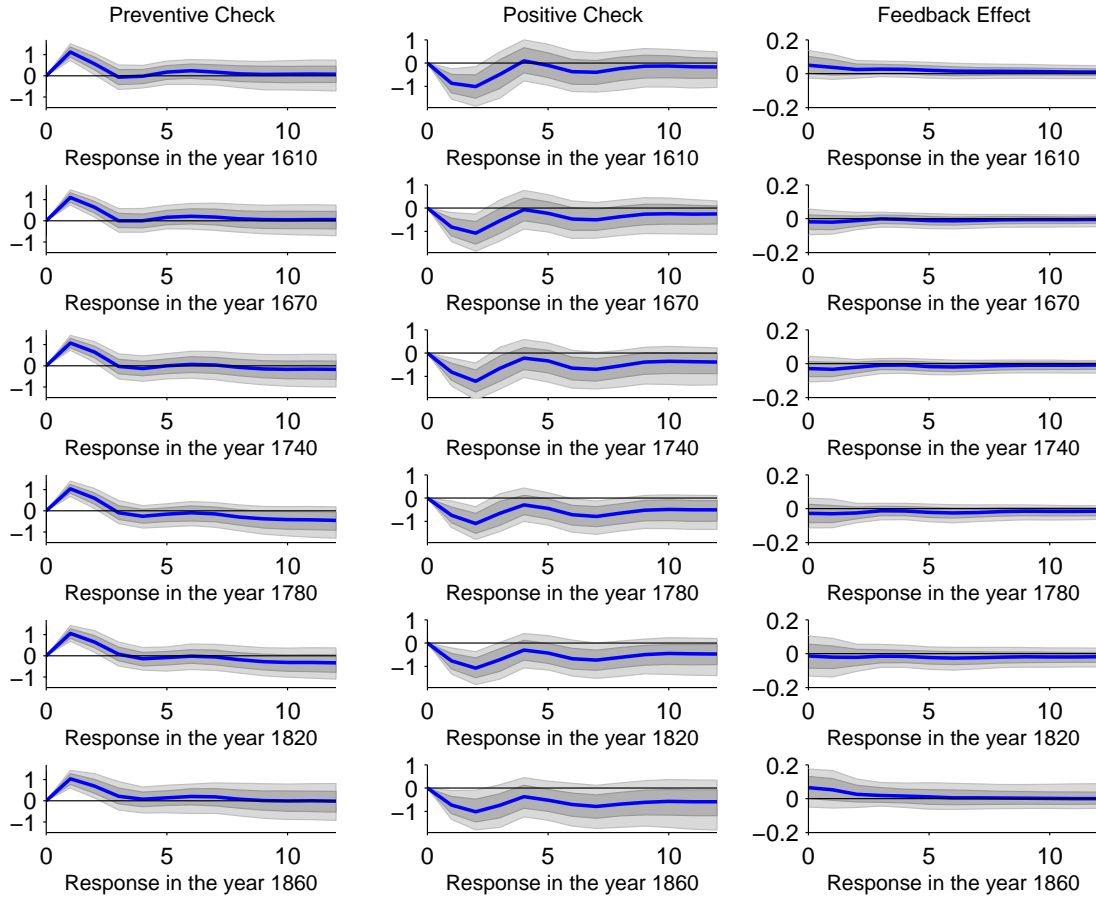
Figure 11: Trace of \mathbf{Q} , \mathbf{W} , and \mathbf{S} : Prior vs. Posterior



Notes: The graph shows prior and posterior density plots for the trace of \mathbf{Q} , \mathbf{W} , and \mathbf{S} .

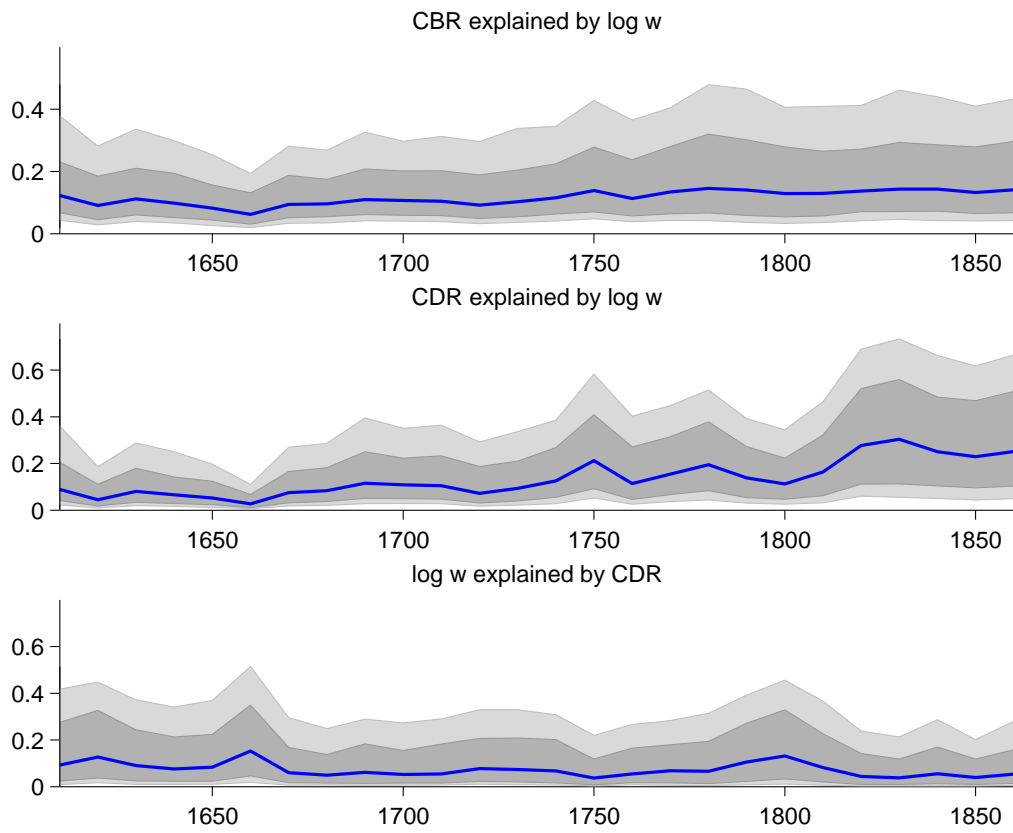
D Alternative Real Wage Series

Figure 12: Clark: Impulse Responses for Various Years



Notes: Posterior means of the impulse response to standardized shocks using Clark's real wage series. The gray (dark gray) area indicates the 5th (16th) and 95th (84th) percentiles of the impulse responses.

Figure 13: Forecast Error Variance Decompositions for Various Years



Notes: Posterior medians of the forecast error variance decompositions for the Clark real wage series. The gray (dark gray) area indicates the 5th (16th) and 95th (84th) percentiles of the forecast error variance decompositions.