Empirical Testing of the Malthusian Population Model
Evidence of the Positive and Preventive Checks in Sweden from 1751 to 1870

Burkett Evans

June 2017
Acknowledgements

This thesis has been a ongoing project dating back to a class I took in the winter of 2016, so I have quite a few people to thank for their help and support.

First, to Professor Joel Mokyr, my thesis advisor. This paper began when we covered Malthusian models in week 3 of his course on European Economic History and I had to present on two papers by Hans-Joachim Voth and Nico Voigtländer in front of a room of graduate students. I ended up writing my paper for his course on Malthusian models, and that paper developed into this thesis. Professor Mokyr provided valuable assistance and feedback throughout the process.

Second, to Professor Robert J Gordon. Prior to my work on this thesis, I spent two years working as a research assistant for Professor Gordon. This was my first exposure to rigorous economic research, with a focus on economic history and the causes of long-run economic growth. This was an incredibly beneficial experience for me in developing as a researcher and in piquing my particular research interests, and I’m very thankful to have had the opportunity to work for Professor Gordon.

Third, to Professor Mark Witte. Professor Witte has been one of my primary mentors at Northwestern, and provided helpful feedback on an earlier version of this paper. I’m inspired by his devotion to his students, his deep passion for economics, and his unceasing Game of Thrones references.

Finally, to my friends and family, without whose help I surely would have had multiple mental breakdowns throughout this process. Thank you for putting up with my endless complaints about my thesis. A special shout-out goes to Gwen for proofreading this thesis (sister, may I have some oats?) and Maria for her continuing emotional support.
1 Introduction

The sources of long-term economic growth have been an important subject in economic literature for centuries. In recent years, a number of theoretical models have been introduced, beginning with Galor and Weil (2000) among others, to explain the transition of societies from the stagnant growth and low standard of living experienced for most of human history to the incredible growth and high standard of living enjoyed in modern industrial nations over the last century. These models commonly use a Malthusian framework to explain stagnant growth in the past. Thus, the usefulness of these models are subject to the existence of Malthusian checks in pre-industrial societies.

Much empirical research has been done to test the existence (or lack thereof) of Malthusian checks in the past. Due to a lack of accurate historical demographic and wage data, this research is typically confined to European countries in the early modern era with reliable birth and death records. The results of this research are conflicted; some studies like Lee (1981) and Eckstein et al. (1984) find support for the existence of Malthusian checks, whereas others such as Lee and Anderson (2002), Nicolini (2007), and Crafts and Mills (2009) find that such checks are weak or non-existent. One major difficulty in modeling Malthusian checks is the inherent endogeneity in all the variables under analysis. VAR models, which are designed to capture the interactions of endogenous time series variables in a highly flexible framework, are a useful way of approaching this problem with aggregate data. Another approach is to use survival analysis methods on individual level data, where available. This paper will attempt to do both using aggregate birth and death rates for all of Sweden and individual level data for five parishes in Southern Sweden, all between 1750 and 1870.

This paper proceeds as follows: Section 2 examines the Malthusian hypothesis in detail. Section 3 gives an extensive review of the existing literature with a focus on those that apply VAR or survival analysis models to historical demographic time series data. Section 4 provides an overview of VAR models and their applicability to testing the Malthusian hypothesis. Section 5 presents a VAR analysis of Malthusian dynamics in pre-industrial Sweden. Section 6 provides an overview of proportional hazard models applied to grouped data. Section 7 describes the data used for our survival analysis, and section 8 presents that analysis. Finally, Section 9 presents conclusions of this analysis.

2 Malthusian Hypothesis

Overview

Thomas Malthus’ *An Essay on the Principle of Population*, published in 1798, was a groundbreaking work on population dynamics. In his essay, Malthus postulated that population growth diminishes the standard of living and that an increasing population level would prevent sustained economic growth. According to Malthus, when wages fall,

A foresight of the difficulties attending the rearing of a family acts as a preventive check, and the actual distresses of some of the lower classes, by which they are
disabled from giving the proper food and attention to their children, act as a positive check to the natural increase of population. (Malthus, 1798)

Thus, birth rates are an increasing function of real wages in a Malthusian model, as families either choose to delay having children due to diminished wages or are encouraged to have children as a result of heightened wages. Likewise, death rates are a decreasing function of real wages, as real wages are a measure of living standards which are a major determinant of life expectancy. These two processes are known as the "preventive check" and the "positive check" respectively.

The final component of a simple Malthusian model is real wages as a decreasing function of population. In most models, this is justified by treating arable land as fixed and setting wages equal to the marginal product of labor. This results in diminishing returns to labor under commonly used production functions and thus diminishing wages as a function of population.

**Model**

Let \( n_t, b_t, d_t, \) and \( w_t \) represent the logarithm of population, birth rates, death rates, and real wages respectively at time \( t \). A simple Malthusian model similar to that of Lee and Anderson (2002) can be defined by the following set of equations:

\[
\begin{align*}
    n_t &= n_{t-1} + b_t - d_t + e_t \\
    b_t &= k_b + \sum_{i=0}^{l} \beta_i w_{t-i} + u_{b,t} \\
    d_t &= k_d + \sum_{i=0}^{m} \delta_i w_{t-i} + u_{d,t} \\
    w_t &= k_w + \omega n_t + u_{w,t}
\end{align*}
\]

In the first equation, \( e_t \) captures primarily net immigration, as well as absorbing measurement error. In the other equations, \( k_b, k_d, \) and \( k_w \) are constants, \( l \) and \( m \) give the number of lags of real wages that impact birth and death rates respectively, \( \beta_i \) and \( \delta_i \) give the effect of real wages in time \( t-i \) on birth and death rates respectively, \( \omega \) gives the effect on population on wages, and \( u_{i,t} \) are random shocks corresponding to birth rates, death rates, and wages.

If the Malthusian hypothesis is true, we would expect \( \sum_{i=0}^{l} \beta_i > 0 \), which would imply that birth rates increase in the short term due to a shock in wages. We would also expect \( \sum_{i=0}^{m} \delta_i < 0 \), which would imply that death rates decrease in the short term due to shock in wages. Finally, we would expect \( \omega < 0 \), which would imply that wages are a decreasing function of population.

**Application**

Malthusian models provide a potential explanation for why the standard of living in Northern and Western Europe was stagnant until recent years, despite evidence of technological progress. The remarkable growth in the standard of living over the last hundred and fifty years in Europe is attributed in large part to the incredible scientific and technological developments underlying
the Industrial Revolution, which boosted labor productivity. However, technological development did occur prior to this period, albeit at a much slower rate, and the standard of living remained stagnant.

Under a Malthusian framework, any shocks which boosts labor productivity would lead to a short term increase in wages. This increase in wages would lead to elevated birth rates and diminished death rates, which in turn would boost population levels. This increase in population feeds back into the wage equation and lowers wages back to their initial levels. Thus, the long run impact of technological development in a Malthusian economy is an increase in population, but no change in the standard of living.

In modern developed European nations, birth and death rates are not determined primarily by wages. Thus, increases in labor productivity typically imply long term increases in the standard of living and not necessarily any effect on population levels. It’s clear that Malthus’ theory no longer describes population dynamics in developed nations. For the purposes of this paper, we’re interested primarily in three questions:

- Does the Malthusian hypothesis adequately explain population dynamics in pre-industrial European nations?
- How are these Malthusian regimes characterized (ie does adjustment occur primarily through the preventive check or the positive check)?
- When did these nations transition out of a Malthusian regime?

3 Literature Review

Introduction

In this section, we review existing research that examines the Malthusian hypothesis, both empirical results for and against the hypothesis and theoretical developments of the base model. Among the most important empirical developments are Eckstein et al. (1984) which was the first to apply a VAR methodology to the Malthusian hypothesis, Nicolini (2007) which applied a recursive VAR to English demographic data, and Crafts and Mills (2009) which applied generalized impulse responses to a similar data set. We will also review other developments that use VAR methods to estimate Malthusians models.

Important Research in the Malthusian Literature

Two major books, Galor (2011) and Clark (2008), provide an excellent overview of the modern uses of Malthusian models. Both use Malthusian models to explore the major questions in economic history: why economic growth was non-existent at most times and places in human history, and why growth began and continued when and where it did. The so-called Unified Growth Theory in Galor (2011) is a endogenous model which explains the transition from low-growth, pre-industrial economies to industrialization with both significant population and
income growth, to the demographic transition with a stable population and high income growth. In Galor's model, the "rate of technological progress depends on the size of the population" (Dinopoulos, 2012, 215), which causes the multiple steady states. Clark (2008) seeks to explain why England in the early nineteenth century was the first site of sustained increases in the standard of living through industrialization. His answer of "downward mobility" has the cultural values of Britain's middle-class disperse over hundreds of years into the rest of the population. Given Malthusian conditions in pre-industrial Britain, Clark argues that the wealthy and literate were likely to have more children than the poor. Due to limited economic opportunities, these surviving children would often be poorer than their parents. Through many generations of this mechanism, Clark argues, Britain became a "more patient, less violent, hardworking, literate, and more thoughtful society" (Clark, 2008, 183), with these characteristics driving Britain to be the first nation to industrialize. In either case, Malthusian models are used to characterize pre-industrial society, with varying methods for breaking out of the Malthusian trap.

Newer cutting-edge research expands on the base Malthusian model in various ways. Particularly noteworthy are a pair of papers published recently, Voigtländer and Voth (2012) and Voigtländer and Voth (2013). Voigtländer and Voth (2012) presents a model where, for a certain range of incomes, death rates respond positively to an increase in income. The authors' justification for this is war and urbanization. In their model, manufactured goods are luxury items produced only in cities. European cities in this time had incredibly high mortality rates, so as a country became richer, demand for luxury goods would increase. This in turn would lead to migration from agrarian jobs in the country to manufacturing jobs in the city, which would lead to an increase in death rates. Additionally, richer countries were more likely to engage in wars with their neighbors, causing both direct increases in mortality through combat and indirect increases in mortality through disease, which armies carried and spread in abundance. The result is a "ratchet effect", where there are two steady states, one at a lower level of income and another at a higher level of income. A sufficiently large positive shock to income would begin a process of convergence from the lower income steady state to the higher income steady state, which would manifest itself as a substantial long-term growth in income per capita despite the country remaining in a Malthusian trap. The authors claim that the Black Death was exactly such a shock to income, and that the "Horsemen Effects" of increased mortality were uniquely present in Europe. Thus, their model provides an Malthusian explanation for why Western European incomes on the eve of industrialization were substantially higher than in Asia and elsewhere.

Voigtländer and Voth (2013) tackles the same question of the divergence in income between Western Europe and the rest of the world prior to industrialization using a Malthusian framework; however, this paper explores the European Marriage Pattern, the tendency for a large proportion of women in Western Europe to marry later in life, and a significant minority to not marry at all. Under a Malthusian model, this fertility restriction is another explanation for why European incomes were substantially higher than elsewhere, where women generally married and had children near the biological age of reproduction. Voigtländer and Voth (2013) provides an economic argument for why delaying marriage could be an optimal economic decision for women in Western Europe. Similarly to Voigtländer and Voth (2012), there is a two sector economy; however, in this model, the luxury goods are animal products such as wool, meat, and milk. Women in this economy have a comparative advantage in land-intensive animal husbandry, as
it requires less physical strength. A negative shock to population, like the Black Death, would increase wages and land per worker. The increased wages would increase demand for the luxury animal products, and the increased land per worker favors the land-intensive animal husbandry industry. Since women have a comparative advantage in this industry, there is more incentive for women to earn money working and postpone marriage. The result is a lower fertility rate, which in the Malthusian model leads to a new steady state at a higher level of income. In both Voigtländer and Voth (2012) and Voigtländer and Voth (2013), the high incomes enjoyed by Western Europeans in the eighteenth century is postulated to be a contributing factor to the Great Divergence between these nations and the rest of the world.

Two other recently released papers find further evidence for both the positive and preventive checks in pre-Industrial England. Kelly and Gráda (2014) examines the positive check in England over some five hundred years using property transfer records to construct a mortality series for England for the century prior to the Black Death. The authors find that over this period, both peasants and nobles exhibited increased mortality in response to poor harvests, resulting in a positive check that spanned all levels of society. As with many other studies on England, they find little evidence of the positive checks from the seventeenth century onward in rural England. However, the authors do find evidence of the positive check for another century in London. As the disappearance of the positive check in both regions corresponds with the introduction of poor relief, the authors suggest that the disappearance of the positive check may be attributed to public charity. Kelly and Gráda (2012) examines the preventive check, in particular using merchets, fees paid by peasants for their daughters to marry. The size of the merchet corresponds to the wealth of the paying peasant, allowing for the authors to examine how variation in household wealth impacted the response of marriage (and by extension, fertility) to harvest shocks. The authors find evidence for the preventive check in both medieval and pre-industrial England as merchet issuance decreased in responses to poor harvests. However, this effect was limited to the poorer peasant, as wealthier peasants actually exhibited an increase in merchets following poor harvests. The authors postulate that for wealthier peasants with significant land holdings, the high grain prices following a poor harvest could be beneficial, as these peasant could sell excess grain on the market. Thus, they may have been more likely to marry off their daughters due to their increased income following a poor harvest. In both of these cases, the authors find tentative evidence of Malthusian pressure acting through both fertility and mortality dating back to medieval times in England.

VAR Testing of the Malthusian Model

Eckstein et al. (1984) was the first study to adopt a VAR framework to historic demographic data. This seminal paper uses demographic data from Sweden's Tabellverket, a national census which has been conducted annually since 1749, to analyze birth and death rate fluctuations in Sweden over the period 1749-1869. This paper is also unique in that it incorporates measures of weather (mean winter, spring, summer, and autumn temperature and precipitation) alongside real wages and crop yields in the VAR analysis. The authors also use data from the Tabellverket to separate the death rate into an infant and non-infant death rate.

In this paper, the authors estimate a reduced form VAR with exogenous variables, namely
the five different weather variables mentioned above. They begin by assuming that shocks to the different endogenous variables are contemporaneously uncorrelated. After settling on a four lag specification, the authors reexamine this assumption. Following estimation, they find large correlations between shocks to their endogenous variables, in contrast to their working assumption. The authors attempted different decompositions in a previous working paper, Eckstein et al. (1982), which found that although the results were robust to different specifications with the exogenous variables, real wages, and crop yields, the results were sensitive to orderings in the birth and death rate. Although the authors "maintain the assumption of zero contemporaneous correlation among the variables rather than impose a temporal ordering on the endogenous variables" (Eckstein et al., 1984, 305), they admit that they "have less confidence in the robustness of our results with respect to alternate specifications and interpretations of the contemporaneous relationships between the birth rate and the death rates" (Eckstein et al., 1982, 43).

After estimating the reduced form VAR, they compute the variance decomposition for each variable. They find in general that the endogenous variables are more sensitive to shocks in the exogenous variables than they are to shocks in other endogenous variables. They go on to generate impulse response figures and cumulative impulse responses of endogenous variables to shocks in exogenous and endogenous variables. Most shocks stabilize quickly (within five to ten years) which leads to asymptotes in the cumulative impulse responses.

Among the major results of these impulse responses are

- Positive shocks to the standard of living (either directly through real wages or indirectly through crops yields or favorable weather) led to increased birth rates and decreased death rates in the short term.

- Shocks to infant mortality led to an increase in the birth rate over the following two years but a negligible cumulative impulse response, suggesting that families performed short-term fertility planning.

- Shocks to non-infant mortality also led to an increase in the birth rate, but in this case the response was significantly persistent.

Eckstein et al. (1984) was a seminal paper in the adoption of VAR models to historical demographics and its use of exogenous weather data is particularly interesting. However, the model’s lack of identification structure and strong assumption of contemporaneously uncorrelated shocks leaves room for further developments.

Nicolini (2007) applies a VAR framework to the reconstructed demographic data from Wrigley and Schofield (1981) and Wrigley et al. (1997) and wage data from Allen (2001). The author examines birth rates, death rates, and real wages over the period 1541-1840, which he divides into three centuries to examine differences in estimated coefficients over time. To confirm the applicability of a standard VAR model, he applies two different unit root tests to confirm that the data are stationary.
The significant development in Nicolini (2007) is a recursive VAR ordering for testing the Malthusian hypothesis. The author proposes the following ordering:

\[
X_t = \begin{bmatrix} CBR_t \\ CDR_t \\ W_t \end{bmatrix}
\]

(5)

where \( CBR \) is the crude birth rate, \( CDR \) is the crude death rate, and \( W \) is the log of real wage. Since pregnancy lasts around nine months, any factors which affect a family's decision to have additional children should operate on around a year long lag, when the time to achieve conception is included. Thus, it is argued that fertility rates in period \( t \) should be exogenous to shocks in other variables at time \( t \). The author assumes that crude death rates are contemporaneously exogenous to wages without any particular justification. Since the price data used for wages is generally a yearly average, it is quite possible that a shock to grain prices, particularly earlier in the year, could affect mortality in the same year.

The author presents impulse responses, cumulative impulse responses, and variance decompositions from the recursively estimated VAR. He finds evidence that although the positive check was significant prior to 1640, it was virtually irrelevant afterwards. While the preventive check was persistent through 1740, after 1740 birth rates responded negatively to shocks in real wages. Thus, this paper would suggest that population dynamics in England in the late seventeenth century were quite different than Malthus thought.

Nicolini (2007) provides interesting evidence against the existence of a Malthusian regime in England in the centuries prior to industrialization by estimating a recursive VAR of demographic data. However, there are some puzzling results that emerge from the author's analysis. In particular, real wages respond positively to a shock in birth rates, and death rates respond negatively to a shock in birth rates, in various sub-samples of the data. He suggests that this may be a failure of his identification strategy, particularly the assumption that birth rates are contemporaneously exogenous.

Crafts and Mills (2009) applies a variety of time series methods to evaluate the presence of a Malthusian regime in early modern Europe. The demographic data are obtained from Wrigley and Schofield (1981), as with Nicolini (2007). However, Crafts and Mills (2009) examine two different real wage series for their VAR modelling: \( WS \) from Wrigley and Schofield (1981) and \( Clark \) from Clark (2005), with \( Clark \) being the authors' preferred series. Although the authors also explore wage trending and structural modeling in this paper, we will only consider their use of VAR modelling.

The authors begin by confirming stationarity of their birth and death rate data and their measures of real wages with ADF test statistics. Their model is a VAR of birth rates, death rates, and real wages, with one version using \( Clark \) and the other using \( WS \). They determine a lag length of 5 using \( Clark \) for real wages and 4 using \( WS \). The authors choose to use generalized impulses responses as described in Pesaran and Shin (1998) instead of applying identifying restrictions.
The authors present accumulated 25 year impulse response functions over rolling sub-intervals of their data, examining the evolution of the positive check (response of death rates to an impulse in wages) and preventive check (response of birth rates to an impulse in wages) over 1541-1799. They find no statistically significant cumulative impact of this impulse in real wages on death rates over any subinterval from 1541 onward, which they present as evidence that there was no positive check functioning in England in the early modern era. Even more condemning for Malthus, they find that the sign on this accumulated impulse response is positive for many subintervals, when the Malthusian hypothesis would imply a negative sign. There is limited support from these accumulated impulse responses for the existence of a preventive check over certain periods of time, but the evidence presented indicates that by the 18th century, the preventive check was no longer functioning in England. Thus, Crafts and Mills (2009) affirms Nicolini (2007) in its findings that England at the time of Malthus had no Malthusian controls on the population.

One potential point of criticism in Crafts and Mills (2009) is the choice of Pesaran and Shin (1998) generalized impulse response functions instead of identifying restrictions. Nicolini (2007) also computed generalized impulse response functions and found results very similar to those of his identifying restrictions, including "the puzzling pattern of the interaction between fertility and mortality" which is potential evidence for mis-identification (115). Unfortunately, Crafts and Mills (2009) only present accumulated impulse response functions for impulse of wages on birth and death rates, so we cannot tell whether the author’s choices of generalized impulse response function present unusual results in the other variables. However, generalized impulse responses are sub-optimal when reasonable identification restrictions exist; although generalized impulse response functions are invariant to variable ordering, they will differ from recursively identified impulse response functions in general, and may do so significantly. Assuming that there are some true structural relationships between the variables, potentially erroneous results would result from not imposing the corresponding restrictions.

Bengtsson and Broström (1997) was another of the earlier works to apply VAR modeling to historical demographic data. In general, the paper discusses techniques for analyzing demographic time series data. As a case study in their paper, the authors apply a distributed lag model, a VAR model, and a VARMA model to the interactions of wages and death rates in Sweden from 1750 to 1850. Though intended only as a limited case study, their VAR model provides preliminary evidence for a positive check in Sweden.

There are two interesting papers that have come out in recent years which use alternative variable orderings to estimate a recursive VAR. Fertig and Pfister (2012) examine the Malthusian hypothesis in Germany using a VAR and a distributed lag model. In their VAR model, the authors make a strong case for reevaluating the Nicolini (2007) variable ordering. They point out two particular flaws in the Nicolini (2007) recursive ordering:

- Lee (1981) found a strong negative correlation between wheat prices in England and birth rates in the same year. As mentioned earlier, the effect of birth rates on wheat prices (which can be treated as a proxy for real wages in the short term) ought to work on a substantial lag, as an increase in birth rates will only increase the working age population after a lag of
over a decade, which will in turn lead to increased competition driving down real wages. We find the same significant correlation between real wages and crude birth rates in our data set in the following section. Lee (1981) interpreted this correlation as working through fetal mortality, as a decrease in income will lead to an increase in miscarriages due to stress or malnutrition. Although this may be reflected as a component of the preventative check in some models, it is not truly preventive in the Malthusian family planning sense and is rather the positive check applied to the unborn.

- Nominal wages were very sticky in the pre-industrial era. Thus, in the short term, real wages are primarily determined by prices (particularly grain prices), which are in turn determined primary by weather and environmental conditions, political developments (such as wars), and other factors which are exogenous to contemporary birth and death rates.

Furthermore, when attempting to use the Nicolini (2007) variable ordering, the authors found additional evidence of mis-identification, as wages responded positively to birth rate shocks and negatively to death rate shocks over various subintervals of German data.

Thus, Fertig and Pfister (2012) suggest the following variable ordering:

$$X_t = \begin{bmatrix} W_t \\ CDR_t \\ CBR_t \end{bmatrix}$$

(6)

For the reasons mentioned above, it follows that wages are contemporaneously exogenous to death rates and birth rates. To justify the ordering of death rates and birth rates, the authors use non-infant death rates instead of overall death rates. Since the main effect of birth rates on death rates is the increased death rates that result from infant mortality following a shock in birth rates, it follows that using non-infant death rates will cause $CDR$ to be contemporaneously exogenous to $CBR$.

By estimating the recursive VAR, the authors find evidence of a Malthusian regime dominated by a strong positive check prior to the nineteenth century. During the nineteenth century, the positive check was no longer statistically significant, although the preventive check persisted later. The significant effect of non-infant death rates on birth rates suggests that what is commonly interpreted as deliberate family planning in response to lowered income may in fact be what the authors describe as an "unmeasured health effect on fertility" (21).

Fertig and Pfister (2012) makes a strong contribution to solving the identification problem in the Malthusian VAR literature. One point of criticism in their analysis is their use of non-infant death rates. Due to a lack of data, the authors estimate their non-infant death rate series as a linear function of death rates in the current period and birth rates in the current and previous periods, using somewhat arbitrary coefficients taken from earlier studies. It would clearly be preferable to use actual non-infant death rates, as the authors’ method is essentially the same as imposing an additional restriction on the coefficients estimates in their VAR.
The other paper is Edvinsson (2015), which uses a distributed lag model, an ARMAX model (a generalization of the autoregressive-moving average model to incorporate exogenous variables), and a VAR model to examine Malthusian dynamics in Sweden. Using previous research, the author estimates these models using birth rates, death rates, marriage rates, and harvest data under the hypothesis that per capita agricultural production is a better measure for standard of living than real wages. There is certainly merit to this argument; (Söderberg, 2010, 463) notes that "declining real wage rates of male labourers may to some extent have been compensated by increased labour market participation by women and children" in eighteenth century Sweden, making real wages an inaccurate measure of the real income available to families. However, it's debatable how well the author's harvest index captures yearly variation in the standard of living.

The author identifies his VAR with a similar recursive structure to Fertig and Pfister (2012), where per capita harvest output takes the place of real wages and marriage rates are between crude death rates and crude birth rates. Unfortunately, the author doesn't elaborate on his choice of identification structure; as he is using total death rates instead of non-infant death rates, we would expect contemporary death rates to be endogenous to birth rates, whereas he assumes no contemporaneous effect of birth rates on death rates. Thus, his identification structure seems implausible.

4 VAR Model

Introduction

The vector autoregression (VAR) model was introduced in Sims (1980) as a flexible framework for capturing the interactions between multiple time series. Sims argued that it was more effective to model multivariate time series data as jointly endogenous rather than imposing unrealistic exogeneity assumptions or restrictions with limited justification. VAR and related models have found widespread adaption in macroeconomics and other fields, as they provide a consistent approach for analyzing a set of endogenous variables over time.

Framework

This section draws on Stock and Watson (2001) and Kilian (2013). Let $X_t$ be a vector of $n$ endogenous variables at time $t = 1, ..., T$ where

$$X_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{n,t} \end{bmatrix} \quad (7)$$

Assume that each $X_t = [x_{i,1}, ..., x_{i,T}]$ is standardized to mean zero and standard deviation of one. A VAR of order $p$, denoted VAR($p$), is assumed to have the structural form

$$A_0 X_t = A_1 X_{t-1} + ... + A_p X_{t-p} + U_t \quad (8)$$

11
where \( A_i \) is a \( n \times n \) matrix for each \( i = 0, 1, ..., p \) and

\[
U_t = \begin{bmatrix}
u_{1,t} \\
u_{2,t} \\
\vdots \\
u_{n,t}
\end{bmatrix}
\]  \((9)\)

We assume additionally that \( E[U_t] = 0 \) and \( E[U_t U_t'] = I_n \), where \( U_t' \) denotes the transpose of \( U_t \) and \( I_n \) is the \( n \times n \) identity matrix. Thus, each of the "structural shocks" are mean zero with a standard deviation of one and are mutually uncorrelated. We define \( E[U_t U_t'] = \Sigma_U \).

We can multiply both sides of our equation by \( A_0^{-1} \) to get

\[
A_0^{-1} A_0 X_t = A_0^{-1} A_1 X_{t-1} + ... + A_0^{-1} A_p X_{t-p} + A_0^{-1} U_t
\]  \((10)\)

By defining \( A_0^{-1} A_i \equiv B_i \) for \( i = 1, ..., p \) and \( A_0^{-1} U_t \equiv E_t \), we get

\[
X_t = B_1 X_{t-1} + ... + B_p X_{t-p} + E_t
\]  \((11)\)

which is known as the reduced form VAR. This equation can be consistently estimated using ordinary least squares (Stock and Watson, 2001, 102). We define \( E[E_t E_t'] = \Sigma_E \).

**Identification**

If \( A_0^{-1} \) is not a diagonal matrix, which is generally the case, then our shocks in the reduced form equation will be correlated. Thus, we want to find a way to recover \( A_0^{-1} \) from our estimates of \( B_1, ..., B_p \).

One common approach, which we will focus on, is recursive identification. We begin by noting that since \( E_t = A_0^{-1} U_T \), we know that

\[
\Sigma_E = A_0^{-1} \Sigma_U A_0^{-1}' = A_0^{-1} A_0^{-1}'
\]  \((12)\)

By performing a Cholesky decomposition, we know that there exists a \( n \times n \) lower triangular matrix \( C \) such that \( CC' = \Sigma_E \). It follows that setting \( C = A_0^{-1} \) allows us to recover the structural shocks, as \( U_t = C^{-1} E_t \). This process is known as "orthogonalizing the shocks".

Though this approach is frequently used in VAR modeling, it’s important to note that there are strong assumptions underlying this process. In particular, the Cholesky decomposition requires imposing a specific causal interpretation on contemporary variables. To see why, note that by (Lütkepohl, 2006, 58-59), the decomposition is equivalent to estimating the following equation:

\[
X_t = D_0 X_t + D_1 X_{t-1} + ... + D_p X_{t-p} + U_t
\]  \((13)\)

where

\[
D_0 = \begin{bmatrix}0 & 0 & \cdots & 0 \\
d_{2,1} & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
d_{n,1} & \cdots & d_{n,n-1} & 0
\end{bmatrix}
\]  \((14)\)
Thus, the results of recursive identification are highly dependent on the ordering of our variables. By imposing this particular form, we assume:

- $x_2, ..., x_n$ have no contemporary effect on $x_1$.
- $x_3, ..., x_n$ have no contemporary effect on $x_2$, but $x_1$ may have a contemporary effect on $x_2$.
- $x_4, ..., x_n$ have no contemporary effect on $x_3$, but $x_1$ and $x_2$ may have a contemporary effect on $x_3$.
- Analogous properties apply for $x_4, ..., x_n$.

By imposing these restrictions, we can gain causal interpretations from our VAR results. However, these interpretations are only as strong as the theory or assumptions underlying the restrictions. If a credible case cannot be made for a particular ordering of variables, the results lose significance.

### Applicability to the Malthusian hypothesis

Since VAR models are designed for studying the interactions between endogenous variables over time, they have the potential to be very useful in testing the Malthusian hypothesis. As discussed in the previous section, VAR models testing the Malthusian hypothesis typically include a range of demographic variables (birth rates, death rates, population, marriage rates), indicators of standard of living (real wages, harvest output, grain prices), and sometimes exogenous variables (precipitation, average temperature).

The most basic setup, similar to the one used in our analysis in section 5, uses real wages, death rates, and birth rates. The Malthusian hypothesis predicts that birth rates respond positively and death rates respond negatively to a shock in real wages (the preventive and positive check respectively) and that wages respond negatively to population over time (diminishing returns to labor). Since this basic model does not include the population level as a variable, this third prediction is not directly tested in such a model. Furthermore, we expect that changes in the population affect wages on a longer lag than can be adequately evaluated with a VAR model; in particular, an increase in the birth rate will only manifest an effect on wages years later as the larger cohort ages, enters the workforce, and drives down wages. Thus, such a VAR model is primarily effective at evaluating positive and preventive checks.

### 5 VAR Analysis

#### Introduction

In this section, we estimate a recursive VAR using Swedish demographic and wage data from 1751 to 1870. Results are presented in impulse response functions, cumulative impulse responses, and variance decompositions. We also perform rolling regressions to test the stability of these estimates over time. We find evidence of a high pressure Malthusian regime (a Malthusian equilibrium driven primarily by the positive check) in Sweden until the early to mid eighteenth century.
Data

All of the demographic data used in the analysis come from Statistics Sweden (Statistiska centralbyrån, abbreviated SCB). As with Eckstein et al. (1984), this data was tabulated from the Tabellverket, Sweden’s national census. This time series, dating back to 1749, is among the most reliable sources of pre-industrial demographic data available on the national level. The starting date for our analysis is constrained by the data; the SCB infant death rate series begins in 1751, which is our starting date as well. While estimates of birth rates and death rates exist for Sweden prior to 1749, they are certainly less accurate than those in the Tabellverket, and estimating the infant death rates in a manner similar to Fertig and Pfister (2012) using these birth and death rate estimates would presumably be very inaccurate year to year. The ending date for our analysis was chosen to be 1870. This is consistently identified in existing literature as a cutoff year for the onset of industrialization and sustained economic growth in Sweden (Sandberg and Steckel, 1997, 128). There are two primary reasons for this specific date. The first is that after 1870, there was sustained growth in GDP per capita at a significantly higher rate than prior to that date. Sandberg and Steckel (1997) suggests that after 1870, GDP growth per capita exceeded 2% per year, whereas it had been 1% or less in the decades prior. The second is that in 1868, Sweden experienced a significant crop failure which led to elevated death rates. As the last significant famine in Swedish history, this event lends further support to a cutoff date of 1870 for industrialization in Sweden.

We calculate the crude birth rate (CBR) by multiplying the births in each year by 1,000 and dividing by the population in that same year. The crude death rate (CDR) is calculated in an analogous way for deaths in each year. To calculate the crude non-infant death rate (CNDR), we first calculate the non-infant deaths in each year by subtracting infant deaths (defined as deaths of children under one year of age) from the total deaths in each year. We then calculate the crude rate in the same manner as CBR and CDR.

Figure 1 presents CBR, CDR, and CNDR from 1751 to 1870. An ADF test gives test statistics of -5.331, -5.596, and -6.119 respectively, which is sufficient to reject the null hypothesis of a unit root at a 1% significance level for all three series. Thus, we can treat these series as stationary for the following analysis.

An examination of Figure 1 will reveal that CDR and CNDR are highly correlated. A roughly constant, high proportion of infants died in the first year of life, which causes very similar movement in the two series. The death rate series is quite volatile and has many noticeable spikes - in particular, 1758, 1763, 1773, 1789, 1800, 1809, 1829, and 1857, among others, stand out as years with unusually high mortality. Many of these mortality shocks overlap famines in Swedish history; according to Lancaster (1990), 1756-1757, 1761-1762, 1771-1772, and 1798-1800 were all years with significant crop failures. He notes that "famine conditions ... contribute largely to mortality, even though the effects might appear indirectly" though heightened rates of disease and infant mortality (Lancaster, 1990, 405). For example, one of the highest rates of typhus in Swedish history was in 1773, corresponding to the famine and contributing to the elevated mortality in that year. However, death rates could spike for reasons not directly connected to the food supply; for example, 1809 had elevated death rates due to poor sanitation among Swedish
soldiers involved in the Napoleonic Wars; these soldiers spread typhus and dysentery at high rates and brought them home after demobilization, leading to widespread outbreaks of disease (Lancaster, 1990, 405).

The final data series used in our analysis is a measure of real wages over this time period. We obtained nominal wages for Swedish agricultural workers over the period 1751 to 1870 from Jörberg (1972). We then converted these wages to a real wage index using the Swedish Consumer Price Index from Edvinsson and Söderberg (2010), setting real wages in 1870 equal to 100. Although Söderberg (2010) contains a real wage index as well, this index is based on urban workers in Stockholm. Pre-industrial Sweden was highly rural; as late as 1850, over 93% of the population lived in towns with a population of fewer that 5,000 people (Bairoch and Goertz, 1986, 288). Thus, the wages of agricultural workers are more reflective of the standard of living in Sweden in a given year. Furthermore, urban wages tended to be more volatile than agricultural wages, and the real wage index from Söderberg (2010) does fluctuate much more than our real wage index over 1751-1870.

When examining the effect of population increases on wages, it’s important to know at what age people in Sweden in this period entered the work force, as this determines the lag with which shocks to the birth rate will affect wages. Although children in rural Sweden would begin performing household chores for their parents at a very young age, they would typically not enter the workforce until thirteen or fourteen; guild regulations in eighteenth century Sweden would allow boys to begin an apprenticeship at the age of fourteen (Olsson, 2009, 593). Children from poorer families often started working at a younger age, and children as young as six would be hired by factories in eighteenth century Sweden. However, when we examine agricultural wages, we would expect a shock in the birth rate to have its primary depressing effect on wages occur roughly thirteen to fifteen years later. This is consistent with the empirical results from Nicolini

---

**Figure 1: Crude Birth Rate, Crude Death Rate, and Crude Non-Infant Death Rate, Sweden, 1751-1870**

![Graph showing birth and death rates over time](image)

---
(2007), which found that wages dropped thirteen years after a shock to birth rates.

Figure 2 presents the real wage index $RW I$ and the log real wage index $L_{RW I}$ from 1751 to 1870. An ADF test gives tests statistics of -3.715 and -3.641 respectively, which is sufficient to reject the null hypothesis of a unit root at a 1% significance level for both series.

Table 1 presents the correlations, means, and standard deviations of $L_{RW I}$, $CNDR$, and $CBR$. Non-infant death rates are negatively correlated with birth rates and wages, while birth rates and wages are positively correlated. While non-infant death rates have a significantly lower mean than birth rates, they are significantly more volatile than birth rates.

![Figure 2: Real Wages (Left Scale, 1870=100) and Log of Real Wages (Right Scale), Sweden, 1751-1870](image)

**Table 1: Correlations, Means, and Standard Deviations**

<table>
<thead>
<tr>
<th>Variables:</th>
<th>$L_{RW I}$</th>
<th>$CNDR$</th>
<th>$CBR$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L_{RW I}$</td>
<td>-</td>
<td>-0.453</td>
<td>0.333</td>
</tr>
<tr>
<td>$CNDR$</td>
<td>-</td>
<td>-</td>
<td>-0.330</td>
</tr>
<tr>
<td>Mean</td>
<td>4.537</td>
<td>18.973</td>
<td>32.694</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.113</td>
<td>4.157</td>
<td>2.260</td>
</tr>
</tbody>
</table>

**Estimating the VAR**

We estimate a VAR with the recursive ordering

$$X_t = \begin{bmatrix} L_{RW I_t} \\ CNDR_t \\ CBR_t \end{bmatrix}$$  \hspace{1cm} (15)
We justify this ordering along similar lines to Fertig and Pfister (2012). Firstly, we assume that real wages are contemporaneously exogenous to birth and non-infant death rates. Nominal wages, especially in the eighteenth century, were very sticky in the short term; the median annual growth in the nominal wage series we use was zero, which occurred in 37% of years. Thus, real wages in the short run are determined primarily by prices, especially food prices which make up the bulk of any pre-industrial CPI. The Malthusian hypothesis claims that birth rates and death rates do affect real wages in the long run through their effect on the population of workers, but this is a long-run effect especially in the case of birth rates, working over many years and not contemporaneously.

Secondly, we assume that crude non-infant death rates are contemporaneously exogenous to birth rates. We know that death rates are contemporaneous endogenous to birth rates through the elevated infant mortality following an increase in birth rates. By eliminating infant deaths from our crude death rates, crude non-infant death rates become contemporaneously exogenous to birth rates by construction. We get similar results using either real wages or log real wages; however, using log real wages allows us to interpret wage variation in our model in terms of percentages instead of level. Since our wage data in this model is an index, it makes more sense to think about percentage variation in our index rather than level. This is consistent with Nicolini (2007) among others.

The next step is to determine the lag length. Over the period 1751-1870, the likelihood ratio test, final prediction error, and Akaike's information criterion all yield a recommended lag length of 3, whereas Hannan-Quinn and Schwarz's information criteria yield a recommended lag length of 2 and 1 respectively. In absence of a compelling reason to pick a particular lag length, we've chosen a lag length of 3.

Results

The most common way to display VAR results is through impulse response functions. An impulse response function graphs the path of a given variable to a one unit standard deviation shock in the error term of another variable. Figures 3a-f plot the impulse response functions of different combinations of variables. In all cases, the impulse response function is presented for ten years. Additionally, 95% confidence bands are graphed in each figure. These figures are in the appendix at the end of this document.

Figure 3a presents the response of \(C_{NDR}\) to a shock in \(L_{RWI}\), and Figure 3b presents the response of \(CBR\) to a shock in \(L_{RWI}\). Figure 3a shows that death rates are highly responsive to shocks in wages. After a shock in wages, death rates are depressed for two years at -1.5 following the shock, before returning to zero by year six and remaining slightly below zero for the remaining years. Figure 3b shows that birth rates are likewise responsive to real wages. However, the magnitude of the response is much smaller; they rise to only 0.7-0.8 for the two years following the shock. Furthermore, the response of birth rates is negative from year 4 onward. This suggests that shocks to real wages may affect the timing of births without affecting the total number of births in the years following the shock. We will address this point further when looking at cumulative impulse response functions.
Figure 3c presents the response of $CBR$ to a shock in $CNDR$. Consistent with Fertig and Pfister (2012), this figure captures an unobserved health effect on fertility, namely how a shock to mortality affects birth rates in the current year (presumably through heightened miscarriage rates) and in the following years. The contemporary effect is significantly negative; increased non-infant mortality has the immediate effect of reducing the birth rate. However, by year two the birth rate is significantly positive, before returning to zero. This figure both shows that there were real and substantial unobserved health effects on birth rates during this period and provides further support to our interpretation of Figure 3b, that there was intentional family planning in response to shocks to the birth rate.

Figures 3d and 3e present the response of $L_{RWI}$ to a shock in $CNDR$ and $CBR$ respectively. As discussed earlier, the VAR model used in this analysis is ineffectively at capturing the effect of birth rates and death rates on real wages due to its lag structure. Thus, it’s not unexpected that the response of $L_{RWI}$ quickly becomes insignificant to a shock in $CNDR$ and $CBR$. However, the response is significant for the first year after the shock and with the correct sign in both cases (real wages respond positively to a shock in death rates and negatively to a shock in birth rates).

Finally, Figure 3f presents the response of $CNDR$ to a shock in $CBR$. The sign is positive as expected; although rates of infant mortality were much higher than the rest of the population, mortality rates for children aged 1-5 were also very high, as young children are especially susceptible to disease. Thus, we would expect non-infant death rates to increase after a shock to birth rates as the population of young children increased.

Another common way of presenting VAR results is by cumulative impulse responses functions. These are the sum of the responses of a given variable to a one unit standard deviation shock in the error term of another variable over a given period of time. Table 2 presents the cumulative impulse response of each variable to its own shock and the shock of other variables. Asterisks denote significance at a 95% level.

<table>
<thead>
<tr>
<th>Table 2: Cumulative Impulse Responses after Ten Years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impulse Variables:</strong></td>
</tr>
<tr>
<td><strong>Response of:</strong></td>
</tr>
<tr>
<td>$L_{RWI}$</td>
</tr>
<tr>
<td>$CNDR$</td>
</tr>
<tr>
<td>$CBR$</td>
</tr>
</tbody>
</table>

These cumulative impulse responses support our general findings from the impulse response figures. In particular, we see that crude death rates have a significantly negative cumulative response to a shock in real wages. We can interpret this as strong evidence for a positive check; a standard deviation negative shock in real wages led to an approximately 0.5% decline in population over the following decade as a result of elevated death rates. The response of birth rates to a shock in real wages is positive but not statistically significant, which may indicate a weak or even non-existent preventive check.
Finally, another useful way to present VAR results is a variance decomposition. The variance decomposition shows the percentage of variance of a given variable due to shocks in the other variables over a given time period. Thus, variance decomposition is another way of examining the interactions occurring between different variables.

Table 3 presents the variance decomposition of each variable due to its own shocks and the shocks of other variables after ten years. Evidence from Table 3 supports our conclusions from the impulse responses and cumulative impulse responses. In particular, shocks to $L_{RWI}$ explain roughly the same proportion of variance in $CNDR$ and $CBR$, although $CBR$ doesn’t have a significant cumulative impulse response to $L_{RWI}$ over the same period. This suggests that although shocks to real wages do impact birth rates, this effect is only one of timing; the total number of births in the decade following a shock to real wages is unchanged, even though the timing of the births may be affected by the shock.

<table>
<thead>
<tr>
<th>Impulse Variables:</th>
<th>L_RWI</th>
<th>CNDR</th>
<th>CBR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance of:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L_{RWI}$</td>
<td>90.8%</td>
<td>5.3%</td>
<td>3.8%</td>
</tr>
<tr>
<td>$CNDR$</td>
<td>31.1%</td>
<td>59.8%</td>
<td>9.1%</td>
</tr>
<tr>
<td>$CBR$</td>
<td>30.5%</td>
<td>11.7%</td>
<td>57.8%</td>
</tr>
</tbody>
</table>

We are also interested in seeing how the positive and preventive checks changed over time. In particular, we want to know whether the cumulative response of $CNR$ and $CNDR$ to a shock in $L_{RWI}$ remained the same or evolved over time. To accomplish this, we performed estimations of our VAR over rolling 50 year sub-samples of our whole period. Because the sample size was severely reduced in these rolling estimations, we reduced the number of lags to 2 for these estimations. Furthermore, we reduced our confidence bands to 90% in the following figures.

Figure 4a presents the ten year cumulative impulse response of $CNDR$ to a shock in $L_{RWI}$ over rolling 50 year sub-samples. The year on the x-axis represents the starting year of the sample; thus, the -2.29 plotted as the first value on the graph represents the ten year cumulative impulse response of $CNDR$ to a shock in $L_{RWI}$ over the period 1751-180. Figure 4a shows a significantly negative response of $CNDR$ to a shock in $L_{RWI}$ in all but the first few sub-periods. More interestingly, the responses becomes more and more negative, falling from a high of -0.69 over 1758-1807 to a low of -7.68 over 1800-1849. The response of $CNDR$ rapidly increases in the following sub-periods, but the responses remains significantly significant at a 10% level. This suggests that sometime in the early to mid nineteenth century, Sweden transition from having a strong positive check with death rates very responses to real wages to having a weak positive check.

Figure 4b presents the ten year cumulative impulse response of $CBR$ to a shock in $L_{RWI}$ over rolling 50 year sub-samples in the same manner as Figure 4a. Firstly, the magnitude of the the cumulative responses is significantly lower. The cumulative response of $CBR$ is never greater.
than 3, whereas the cumulative response of CNDR is less than -3 all but a few sub-periods. Furthermore, the cumulative responses is insignificant for many of the sub-samples. There's some limited evidence for a statistically significantly response roughly over 1760-1810, but at best we can conclude that there may have been a weak preventive check over this time period. There does not appear to be a significant break at any point in time as there was with CNDR in the early to mid eighteenth century.

Finally, one unexplained component of these figures is the statistical insignificance of the cumulative responses of both variables in the early sub-samples. It seems reasonable that if a positive or preventive check were functioning over any sub-sample, it would be more likely to function over earlier ones primarily in the eighteenth century rather than later ones entirely in the nineteenth century. One possible reason for this surprising result is measurement error. The older data, especially close to 1750, tend to be less precise and accurate than the nineteenth century data, which could cause inaccurate estimates.

6 Grouped Proportional Hazard Model

Introduction

The preceding sections have provided evidence for a strong positive check in Sweden. However, the aggregate death rate statistics provide very little explanation for how exactly the positive check functioned with regards to age and socioeconomic status. To explore how people of different age and class in pre-industrial Sweden responded to shocks in real wages, we will use survival analysis techniques on individual-level data.
Model Description

This section draws from Allison (1982) and Jenkins (2005). Let \( t, T \in \{0, 1, 2, \ldots \} \) be discrete time periods and \( i \) be an individual in the population. \( t \) measures the age of individual \( i \) and \( T \) indicates the period of death. Let \( d(t) = P(T = t) \), the probability that death occurs in period \( t \). Let

\[
F(t) = P(T < t) = \sum_{j=0}^{t-1} d(j) \tag{16}
\]

be the probability that \( T < t \) (i.e. the probability that death occurs prior to time \( t \)). We can then define a survival functions \( S(t) = P(T \geq t) = 1 - F(t) \), the probability that death occurs on or after period \( t \). Using these two functions, we can define the hazard function

\[
h(t) = P(T = t | T \geq t) = \frac{d(t)}{S(t)} \tag{17}
\]

The hazard function represent the conditional probability of death at period \( t \) given survival prior to period \( t \), and will be the subject of our estimation.

Let \( \mathbf{x}_{i,t} \) be a \( k \times 1 \) vector of possibly time-varying explanatory variables. Then our hazard function for individual \( i \) at time \( t \) is \( h(i, t) = P(T_i = t | T_i \geq t, \mathbf{x}_{i,t}) \). The next step is to choose a function relating the hazard rate to our explanatory variables. There are a number of functions which can be used for this purpose. In this paper, we will be using the following function:

\[
h(i, t) = 1 - \exp(- \exp(\alpha_t + \beta \mathbf{x}_{i,t})) \tag{18}
\]

where \( \beta \) is a \( 1 \times k \) vector of coefficients and \( \alpha_t \) is a series of constant terms dependent on \( t \). We can rearrange this equation as follows:

\[
1 - h(i, t) = \exp(- \exp(\alpha_t + \beta \mathbf{x}_{i,t})) \tag{19}
\]
\[
\log(1 - h(i, t)) = -\exp(\alpha_t + \beta x_{i,t}) \quad (20)
\]
\[
\log(-\log(1 - h(i, t))) = \alpha_t + \beta x_{i,t} \quad (21)
\]

There are a number of advantages to using this model, usually referred to as complementary log-log, over more typical logit models. One practical reason is that, since the complementary log-log function is asymmetric, it is better for fitting low-probability events such as the probability of death in a given year. Another reason is, in using discrete survival models, it is often assumed that there is some actual continuous underlying process from which we obtain discrete observations. For such models, we would have a probability density functions \( g(t) \) and a cumulative density function \( G(t) \) for time of death \( T \). The continuous hazard rate in that case would be \( \lambda(t) = \frac{g(t)}{1 - G(t)} \).

Given a vector of possibly time-varying explanatory vectors \( x(t) \), the standard approach in estimating this continuous hazard rate, from Cox (1972), is the proportional hazard model

\[
\log(\lambda(t, x)) = \lambda_0(t) + \beta x(t) \quad (22)
\]

where \( \lambda_0(t) \) is an arbitrary function. The partial likelihood method introduced by Cox (1972) allows for estimations of \( \beta \) without making assumptions about the functional form of \( \lambda_0(t) \). Per Prentice and Gloeckler (1978), the estimates of \( \beta \) in the discrete model are equivalent to \( \beta \) in the continuous model in the case that the continuous data is grouped into discrete data. Hence, the discrete model is often referred to as a grouped proportional hazard model.

The Cox proportional-hazard model is semi-parametric, in that the functional form of \( \lambda_0(t) \) is not specified. For the complementary log-log model to maintain analogous properties, the \( \alpha_t \) term should simply be a series of dummy variables corresponding to each discrete age \( t \). However, per Allison (1982), there may be compelling reasons to impose a specific functional form on \( \alpha_t \) as a function of \( t \).

To estimate this model, Jenkins (2005) and others recommend a technique known as event splitting. To accomplish this, assume that the original data structure has, for each person \( i \), a length of time under observation \( T_i \in \{1, 2, \ldots\} \), a variable \( c_i \) indicating right censoring (\( c_i = 0 \) if the individual is right censored), \( u_i \in \{0, 1, 2, \ldots\} \) as the period of entry for left truncated data (\( u_i = 0 \) if the data is not left truncated), and whatever time and/or individual dependent explanatory variables are used in the analysis.

The first step is to reorganize the data into person-period format. To do that, for each person \( i \), we create \( T_i \) observations. Each observation corresponds to one time period 1, 2, \ldots, \( T_0 \) for which the individual is alive, with a new variable \( t_{i} \) designating that person’s age in that time period. Next, we drop all observations for which \( u_i \geq t_i \), thus omitting person-periods for which the person in question has not yet entered observation. Next, we generate whatever time and/or individual dependent explanatory variables we wish to use in this new period format. Among these should be variable(s) corresponding to \( \alpha_t \), either dummies for each age \( t \) or some function of \( t \). We then will create a death indicator variable \( y_{i,t} \), which is equal to 1 only when both \( t_i = T_i \) and \( c_i = 1 \), and 0 otherwise. Finally, we can perform a standard complementary log-log regression with \( y_{i,t} \) as the binomial dependent variable on the reorganized data.
7 Data for Survival Analysis

Overview

The data used in this study comes from Bengtsson et al. (2014), hereafter referred to as SEDD, the Scanian Economic Demographic Database. The SEDD covers five rural parishes (Halmstad, Hög, Kågeröd, Kävlinge and Sireköpinge) in Scania, a region in Southern Sweden. Data in the SEDD comes from a combination of birth registers, death registers, parish registers, and census registers, and individuals are grouped into families when identifiable. Altogether, the SEDD spans several centuries of Swedish history and combines demographic and economic variables over time, which makes it an excellent source of evidence for testing the existence of a positive check in pre-industrial Sweden.

The historic region of Scania covers the southern tip of Sweden, which is separated from Denmark by the narrow Öresund Strait. Scania was originally part of Denmark, which gained control of the strait and by extension much of the Baltic Sea trade. As a result, Scania became a focal point for conflict between Sweden and Denmark. The Treaty of Roskilde in 1658 following the Second Northern War passed control of Scania from Denmark-Norway to Sweden. By the signing of the Treaty of Frederiksborg in 1720, which concluded the Great Northern War between Sweden and Denmark-Norway, Scania was firmly established as Swedish territory.

Births, Deaths, and Migrations

For our analysis, we used four main variables to identify the status of the individuals under observation. The variable BirthDate identifies the birth date and DeathDate identifies the death date. For each individual under observation, we identified their birth and death years and created an observation for each year that the individual was alive, with an age variable which counts the number of years from birth.

A cursory observation of the SEDD will reveal that the number of individuals with an identified birth date far exceeds the number of individuals with an identified death date. The reason for this is migration. One variable InmigSCBcode_Parish records the parish from which an individual immigrated into observation, whereas OutmigSCBcode_Parish records the parish to which an individual emigrated out of observation. For the vast majority of individuals, we have a listed birth date, as individuals born outside of observation who immigrated into observation will have their birth date recorded, and we can thus calculate their age for the period in which they were under observation.

Thus, to incorporate migration into our event splitting, we used roughly the following process: for each individual in each year, we identified their most recent status variable and their next upcoming status variable, if either existed. We would then flag individuals as under observation if their previous status was birth or immigration and next status was death or emigration. We had a similar process in place for more complicated cases, where there were multiple status change variables in the same year.
One of the primary reasons we believe a discrete model to fit this data better is the issue of migration. While the SEDD records precise dates for births and deaths, migration variables are often fixed at July 1 of the given year. Thus, we assume that individuals are under observation beginning in the year when immigration occurs, and are no longer under observation in the year emigration occurs.

**Socioeconomic Variables**

There are a number of socioeconomic variables that can be used for additional analysis. The analysis by Bengtsson (2004) primarily utilized LandTypeID and SizeFraction. The SizeFraction variable records the mantal of the property of the individual under observation. A mantal is a Swedish tax unit which measure productive capacity of a given farm; per Bengtsson (2004), a farm with a mantal of \( \frac{1}{16} \) or greater in the nineteenth century was considered sufficient to support a family. Thus, Bengtsson (2004) considers the lowest, fourth social tier to be those without any land and the third social tier to be the semi-landed.

To distinguish the two higher social tiers, Bengtsson (2004) considers the property rights of those rural Swedes with a mantal > \( \frac{1}{16} \) farm. Per Myrdal and Morell (2011), Swedish farm land could be grouped into three different legal statuses: freehold, crown, and exempt. Freehold land was land owned by Swedish farmers who paid taxes directly to the crown. Crown land was owned by the crown and farmed by peasants, who had increasingly strong property rights as tenants over this period. Exempt land was land owned by nobility who rented it to tenants; as these tenants generally had fewer right than crown tenants, Bengtsson (2004) placed the noble tenants in the second social tier and grouped freeholders and crown tenants as the first social tier. The LandTypeID variable records the legal status of the property.

In our analysis, we will consider another measure of social status, occupation of the head of household as recorded by FamHeadOccEvents. The occupations can be linked to HISCLASS codes from HISCO, the Historical International Standard Classification of Occupations. HISCLASS sorts all occupations into 12 categories based on their manual/non-manual status, skill level, and supervisory responsibilities. For the purposes of our study, we group all non-manual and supervisory roles (HISCLASS 1-6) into one category. The remaining 6 categories correspond to medium-skilled, lower-skilled, and unskilled workers in agricultural or non-agricultural work.

**Real Wages**

For this paper, we constructed three different measures of real wage based on three staple crops consumed by poorer Swedes in this period. First, we obtained day laborer wages from Jörberg (1972) for region 11, Malmöhus County. Although Scania technically includes this region and region 10, Kristianstad County, the parishes included in the SEDD all fall within Malmöhus County. The day laborer wages are not available until 1781; thus, we use 1781 as a start date for our analysis involving real wages.
To deflate this nominal wage series, we also obtained grain prices in region 11 for rye, barley, and oats from Jörberg (1972). These three crops were commonly grown in Scania and would all be used in various dishes consumed by poorer Swedes. Bengtsson (2004) only uses rye prices to deflate wages. Rye and barley were in many ways complementary grains; they were both commonly grown in Sweden as they were more cold-resistant than wheat, they were similarly priced, and could both be used to make Swedish breads. Oats, on the other hand, were substantially cheaper, were often used as fodder, and, especially in the mid to late 18th century, were largely exported. Thus, we would expect similar results from rye wages and barley wages, while oat wages may behave differently. For the period prior to 1803, grain prices are listed per tunna, a Swedish unit of volume equivalent to 165 liters. Although Sweden had three different currencies during this time period, no currency conversion were necessary as we divided the day laborer’s wage by the price of each grain; the only conversion needed was for volume.

Figure 5 depicts the three real wage series from 1781 to 1870. Rye and barley are similarly priced and highly correlated (0.873), with rye being slightly more expensive per liter. Barley is quite a bit more expensive than oats, however they too are highly correlated at 0.853. Rye and oats are less correlated at 0.67.

8 Survival Analysis

Kaplan–Meier

One common way of analyzing survival data is by the Kaplan–Meier estimator. If \( n_i \) is the total population of individuals at age \( i \) and \( d_i \) is the total number of deaths at age \( i \), the estimator for
the survival function is

$$\hat{S}(t) = \prod_{i < t} \left(1 - \frac{d_i}{n_i}\right)$$

(23)

where \(t\) represents age. The result is a decreasing step function.

Figure 6a depicts the Kaplan-Meier estimator for the whole period 1750-1870. As this figure clearly shows, infant mortality in Scania was extremely high, with one in five infants dying before their first birthday. Mortality rates decreased significantly after the first year of life, however. Conditional on surviving their first year of life, Swedes at this time had a better than 50% chance of surviving to their 55th birthday, which was around the age most people stopped working. However, due to the high infant and childhood mortality rates, life expectancy from birth over the whole time period was approximately 38 years.

Figure 6b depicts the same Kaplan-Meier estimator for three sub-periods: 1750-1800, 1800-1835, and 1835-1870. As these estimates show, the major change over this period was the decline in infant mortality. While three in ten infants died before the first birthday before 1800, this value decreased to just over one in ten for the last sub-period. Driven largely by this decline in infant and childhood mortality, life expectancy climbed from 30 years in the first period to 37 in the second and 44 in the third. However, mortality rates remained relatively unchanged for adults. In the first sub-period, a 15 year old had a 63.4% chance of surviving to his or her 55th birthday. In the second sub-period, this same probability actually decreased to 63.4%, and increased only a minor amount to 68.9% in the third sub-period.
Effect of Real Wage Shocks by Age

The next part of our investigation was to examine the effect of real wages shocks on mortality. We estimated the complementary log-log model described above with the binary outcome of death, a baseline hazard rate $a_t$, and vector of explanatory variables $x_{i,t}$. For all six of the following tables, we looked at two different forms of $a_t$. For the first variation, we had a dummy variable for each age up to 85, and a dummy variables for ages greater than or equal to 86. The second variation involved grouping age groups by their Kaplan-Meier estimated survival rates. This produced eight different dummies. The first three dummies ageind_group0, ageind_group1, and ageind_group2 correspond to ages 0, 1, and 2 respectively. Over the whole time sample 1750-1870, these groups had estimated hazard rates of 19.4%, 5.6%, and 3.2%. These high yet distinct rates justified placing each of these ages into their own groups. ageind_group3 covers ages 3-6, where hazard rates vary from 1% to 2%. ageind_group4 captures the group of ages with the lowest hazard rate, 7-35, where hazard rates are below 1% in almost every year. Ages 36-45, 46-55, and 56+ are covered by ageind_group5, ageind_group6, and ageind_group7 respectively. Ages 36-45 had hazard rates between 1% and 1.5%, ages 46-55 were between 1.5% and 2.5%, and hazard rates rapidly increase after the age of 55.

For brevity, the following tables all depict the second form of $a_t$. The age mortality patterns captured by our Kaplan-Meier estimator are reflected in our regressions whether we use actual ages or grouped ages as described above. Additionally, our estimates and p-scores for $\beta$ are consistent between the two forms of $a_t$.

All of the following regressions cover the period 1781, when real wage data becomes available, to 1870. Unfortunately, many of the variables available for analysis in the SEDD do not began recording until the early 19th century - thus, we can only control for age and gender in our
analysis. The wage data used in all of the following regressions is lagged a year, as the crop price data was usually recorded during the fall harvest.

Table 4a: Effect of Rye Deflated Wages

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>female_ind</td>
<td>-0.0474</td>
<td>0.0234</td>
<td>4.0965</td>
<td>0.0430</td>
</tr>
<tr>
<td>ln_RW</td>
<td>-0.3374</td>
<td>0.0502</td>
<td>45.1419</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group0</td>
<td>-0.9448</td>
<td>0.1078</td>
<td>76.7529</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group1</td>
<td>-2.2245</td>
<td>0.1150</td>
<td>374.0837</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group2</td>
<td>-2.7859</td>
<td>0.1227</td>
<td>515.7626</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group3</td>
<td>-3.5377</td>
<td>0.1158</td>
<td>932.5542</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group4</td>
<td>-4.3747</td>
<td>0.1103</td>
<td>1571.8992</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group5</td>
<td>-3.8145</td>
<td>0.1163</td>
<td>1075.5866</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group6</td>
<td>-3.2337</td>
<td>0.1133</td>
<td>813.9154</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group7</td>
<td>-1.9914</td>
<td>0.1075</td>
<td>343.4351</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Table 4a depicts the effects of variation in rye wages on mortality. In interpreting these results, the binary outcome is 1 if death occurs and 0 if not, so negative values represent a decrease in the probability of death. All variables presented here are significant at a 5% level, and all but the female indicator are significant at a 1% level. When interpreting these tables, note that there is no constant term. Rather, the age indicator dummy variables provide the base hazard rate for each individual given their age. This hazard rate is then increased or decreased based on the coefficient and value of the explanatory variables. Since the probability of any particular person dying in a given year is quite low (with the exception of infants and the elderly), the coefficients on our age indicator dummies will all be negative, with the age indicators corresponding to prime-age people being large in absolute value. If $\alpha_t + \beta x_{i,t} = 0$, our model would estimate a 63% probability of death in year $t$ which is exceptionally high. If $\alpha_t + \beta x_{i,t} = -3$, the probability of death in year $t$ drops to less than 5%. Thus, the base hazard rate $\alpha_t$ will generally be highly negative.

Recall that our estimated hazard rate from the complementary log-log function is $h(i, t) = 1 - \exp(-\exp(\alpha_t + \beta x_{i,t}))$. The average real rye wage over this period was 11.67 liters, corresponding to a log real wage of 2.46. Thus, the estimated hazard rate for an adult male in ageind_group4 when wages are at their historic average is $1 - \exp(-\exp(-4.3747 + 0.3374 \times 2.46)) = 0.55\%$. If wages were to drop by 20%, the new log real wage of 2.26 would yield an estimated hazard rate of $1 - \exp(-\exp(-4.3747 + 0.3374 \times 2.26)) = 0.59\%$, an approximately 7% increase in the hazard rate. A similar analysis for infants in ageind_group0 yields an initial hazard rate of 15.6% and a new hazard rate of 16.6%, for an increase of approximately 6%. Given the high volatility of rye wages over this period, this suggests that morality rates year to year were quite dependent on real wages, and is thus evidence of the positive check operating in Scania.

This second table repeats the analysis with barley wages. We find the same statistically significant results with the same sign on each coefficient; however, the magnitude of the response to real wages drops from -0.338 to -0.178. Thus, the positive check appears less strong when using
Table 4b: Effect of Barley Deflated Wages

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>female_ind</td>
<td>-0.0480</td>
<td>0.0234</td>
<td>4.1970</td>
<td>0.0405</td>
</tr>
<tr>
<td>ln_RW</td>
<td>-0.1768</td>
<td>0.0509</td>
<td>12.0586</td>
<td>0.0005</td>
</tr>
<tr>
<td>ageind_group0</td>
<td>-1.2207</td>
<td>0.1267</td>
<td>92.7716</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group1</td>
<td>-2.5007</td>
<td>0.1328</td>
<td>354.6691</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group2</td>
<td>-3.0624</td>
<td>0.1394</td>
<td>482.5291</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group3</td>
<td>-3.8146</td>
<td>0.1334</td>
<td>818.1962</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group4</td>
<td>-4.6501</td>
<td>0.1288</td>
<td>1302.5364</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group5</td>
<td>-5.1371</td>
<td>0.1422</td>
<td>1305.2238</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group6</td>
<td>-5.5581</td>
<td>0.1399</td>
<td>1061.6158</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group7</td>
<td>-3.3173</td>
<td>0.1353</td>
<td>601.1655</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

barley wages. The reason for this is likely that barley was a secondary crop to rye in Southern Sweden. Although the two were to some degree substitutable, an increase in rye prices was more likely to affect the standard of living than an increase in barley prices.

Table 4c repeats the above analyses, but with oat wages. Here we find our first surprising results, that the coefficient on our real wage term is positive and statistically significant. Thus, an increase in oat wages would lead to higher mortality in this model. The most likely explanation for this is, as mentioned earlier, that oats were often used as fodder in addition to human consumption and were increasingly exported during the middle and late 19th century. Thus, there is a weaker link between standard of living and oat prices, as oat prices were likely to be affected by other factors. For example, it could be that when rye or barley prices were low, demand for meat or other animal products would increase, as more people could afford to include meat in their diets. As a result, the price for oats as fodder would increase to feed the livestock, leading to a decrease in oat wages. With regards to exports, oat wages were at their lowest levels toward the end of the time period likely due to increase foreign demand for Swedish oats and subsequently higher prices. This trend is not present in barley or oat wages, which
were more stable around their long-run average. As mortality rates in general tended to decline towards the end of the time period, our model’s positive relationship between oat wages and hazard rates may be picking up on the higher oat wages and lower mortality at the end of the period in question, rather than the effect of short-term wage variation on hazard rates.

### Table 4d: Effect of Rye Deflated Wages With Age Interactions

<table>
<thead>
<tr>
<th>Parameter:</th>
<th>Estimate</th>
<th>SE</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>female_ind</td>
<td>-0.0477</td>
<td>0.0234</td>
<td>4.1463</td>
<td>0.0417</td>
</tr>
<tr>
<td>ln_RW_group0</td>
<td>-0.2094</td>
<td>0.0943</td>
<td>4.9298</td>
<td>0.0264</td>
</tr>
<tr>
<td>ln_RW_group1</td>
<td>-0.5886</td>
<td>0.1994</td>
<td>8.7129</td>
<td>0.0032</td>
</tr>
<tr>
<td>ln_RW_group2</td>
<td>-0.5346</td>
<td>0.0974</td>
<td>30.1507</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ln_RW_group3</td>
<td>-0.5184</td>
<td>0.0978</td>
<td>28.1240</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ln_RW_group4</td>
<td>-0.3901</td>
<td>0.1822</td>
<td>4.5824</td>
<td>0.0323</td>
</tr>
<tr>
<td>ln_RW_group5</td>
<td>-0.2187</td>
<td>0.0908</td>
<td>5.7967</td>
<td>0.0161</td>
</tr>
<tr>
<td>ageind_group0</td>
<td>-1.2127</td>
<td>0.1995</td>
<td>36.9599</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group1</td>
<td>-1.7008</td>
<td>0.4169</td>
<td>16.6449</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group2</td>
<td>-2.3744</td>
<td>0.2123</td>
<td>125.0878</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group3</td>
<td>-3.1262</td>
<td>0.2085</td>
<td>224.9104</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group4</td>
<td>-3.9846</td>
<td>0.2033</td>
<td>384.2281</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group5</td>
<td>-3.4382</td>
<td>0.2090</td>
<td>270.7194</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group6</td>
<td>-3.1238</td>
<td>0.3817</td>
<td>66.9654</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group7</td>
<td>-2.2391</td>
<td>0.1915</td>
<td>136.6514</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

For Table 4d, we’ve expanded our analysis to include interaction effects. Thus, instead of having a single real wage term, we created six real wage variables, where \( \ln RW_{groupi} \) is equal to \( \ln RW \) if and only if the individual is in \( groupi \), and 0 otherwise. Thus, this regressions captures the effect of wage variation on different age groups. The grouping for this regressions varies somewhat from the \( a_t \) grouping. \( ln_RW_{group0} \) and \( ln_RW_{group1} \) cover ages 0 and 1 respectively, while \( ln_RW_{group2} \) covers ages 2 to 15. Prime age adults (16-45) are covered in \( ln_RW_{group3} \), while older working adults (46-55) are in \( ln_RW_{group4} \) and adults who generally no longer worked (56+) are in \( ln_RW_{group5} \). As with the first three tables, we had similar results using dummies for \( a_t \) as opposed to \( a_t \) grouping.

These results suggest that there was differences in the effect of rye wage variation on people of different ages. While the effect of variation in real wages on all age subgroups is significant at the 5% level, we find the largest magnitude in the middle aged groups, while both infants and the elderly have lower mortality responses to wage variation. This is consistent with previous findings in Bengtsson (2004) and others.

When running the same regression with barley wages (Table 4e), we obtain similar results. As before, we find an overall lower magnitude of response to variation in barley wages. Additionally, like with rye wages, we find a more significant interaction effect for wages with middle aged groups. In fact, the response of infant mortality to wage variation is actually positive, although statistically insignificant.
Table 4e: Effect of Barley Deflated Wages With Age Interactions

<table>
<thead>
<tr>
<th>Parameter:</th>
<th>Estimate</th>
<th>SE</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>female_ind</td>
<td>-0.0479</td>
<td>0.0234</td>
<td>4.1854</td>
<td>0.0408</td>
</tr>
<tr>
<td>ln_RW_group0</td>
<td>0.0556</td>
<td>0.0964</td>
<td>0.3327</td>
<td>0.5641</td>
</tr>
<tr>
<td>ln_RW_group1</td>
<td>-0.3749</td>
<td>0.2057</td>
<td>3.3215</td>
<td>0.0684</td>
</tr>
<tr>
<td>ln_RW_group2</td>
<td>-0.4599</td>
<td>0.0991</td>
<td>21.5224</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ln_RW_group3</td>
<td>-0.4450</td>
<td>0.0991</td>
<td>20.1821</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ln_RW_group4</td>
<td>-0.4450</td>
<td>0.0991</td>
<td>21.5224</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ln_RW_group5</td>
<td>-0.1195</td>
<td>0.0911</td>
<td>1.7194</td>
<td>0.1898</td>
</tr>
<tr>
<td>ageind_group0</td>
<td>-1.7897</td>
<td>0.2379</td>
<td>56.6093</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group1</td>
<td>-2.0184</td>
<td>0.5019</td>
<td>16.1717</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group2</td>
<td>-2.3740</td>
<td>0.2486</td>
<td>91.1860</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group3</td>
<td>-3.1267</td>
<td>0.2451</td>
<td>162.7078</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group4</td>
<td>-3.9841</td>
<td>0.2406</td>
<td>274.2138</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group5</td>
<td>-3.4371</td>
<td>0.2452</td>
<td>196.5250</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group6</td>
<td>-3.6683</td>
<td>0.4444</td>
<td>68.1528</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group7</td>
<td>-2.4054</td>
<td>0.2241</td>
<td>115.2372</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Finally, running the same regression with oat wages, as shown in Table 4f, we find as before contrary results to rye and barley wages. The large and statistically significant positive response of infant mortality to oat wages is potentially evidence for the export explanation, as infant mortality declined significantly from the beginning of the period to then end, when oat exports increased and oat prices rose substantially.

Effect of Real Wage Shocks by Class

For the final part of this analysis, we restrict our analysis to the years following 1813, when many of the demographic variables become available. There are two socio-economic variables in particular that we will use in our analysis. The first is HISCLASS, as mentioned earlier. HISCLASS 1-5 covers non-manual workers, and HISCLASS 6 covers foremen, so we restrict our analysis to HISCLASS 7-12.

HISCLASS 7 is medium skilled non-agricultural workers. In the SEDD subset under analysis, the most common occupations is HISCLASS 7 include *smed*, smiths, *mjölnare*, millers, and *skomakare*, shoemakers. HISCLASS 8 is medium skilled agricultural workers, usually referring to farmers. There are many different terms in Swedish for farmers which denote different distinctions in social class. The most common occupation is HISCLASS 8 is *åbo*, a term usually used for tenant farmers with inheritable leases, such as crown tenants. The next most common is *arrendator*, which is a general term for tenant farmer. HISCLASS 9 is lower skilled non-agricultural workers. The most common occupations in this group are soliders, including *rotesoldat* and *soldat*, both referring to soldiers, and *husar*, or hussars. HISCLASS 10 is lower skilled agricultural laborer. Common occupations in this group include *husman*, a somewhat generic term which
often referred to cottagers, farmers who owned a house but not land, and *torpare*, crofters with small plots of land not sufficient to raise a family on. HISCLASS 11 is unskilled non-agricultural laborers. By far the most commonly listed occupation in this group is *inhyses*, or lodger. This is a generic term for anyone living on a property which they do not own. While this could refer to actual lodgers, it often was a term used for parents who transferred ownership of a farm to their children while retaining the right to live on the farm, with their children often being obligated to provide them with food or other necessities. Finally, HISCLASS 12 covers unskilled non-agricultural laborers. The most common occupations in this group are variations on *dräng*, a generic term for farmhand or servant. For the number of occupations where a HISCLASS is unassigned, the most common "occupation" is *undantag*, which refers to free lodgers on land who are guaranteed certain necessities. Similar to *inhyses* but explicitly referring to those with free lodging, this was a common arrangement for the elderly farmer in return for transferring their farm to their children. Another unassigned HISCLASS is *änka*, or widow.

Another variable used in this analysis is *mantal*, a Swedish term for a taxable unit of land. Similarly to Bengtsson (2004), we create three dummies based on this variable: *mantal_ind_0* when *mantal* equals zero, indicating no land ownership, *mantal_ind_1* when *mantal* is non-zero but less than $\frac{1}{16}$, indicating land ownership but not sufficient to support a family, and *mantal_ind_2*, indicating land ownership of $\frac{1}{16}$, sufficient to support a family.

For the class analysis, we ran a number of regressions with different controls and interaction effects. Presented here is the most complete model using rye wages, with controls for gender, parish, HISCLASS, decade, and *mantal*, and wage interaction effects for HISCLASS and *mantal*. The results presented in this table are generally robust to the different regression specifications. For dummy variables, *parish_ind_1* is the parish base case (corresponding to Hög parish), HIS-

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>female_ind</td>
<td>-0.0492</td>
<td>0.0234</td>
<td>4.4091</td>
<td>0.0357</td>
</tr>
<tr>
<td>ln_RW_group0</td>
<td>0.5906</td>
<td>0.0848</td>
<td>48.5343</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ln_RW_group1</td>
<td>0.3359</td>
<td>0.1762</td>
<td>3.6354</td>
<td>0.0566</td>
</tr>
<tr>
<td>ln_RW_group2</td>
<td>-0.0046</td>
<td>0.0869</td>
<td>0.0028</td>
<td>0.9581</td>
</tr>
<tr>
<td>ln_RW_group3</td>
<td>0.0056</td>
<td>0.0865</td>
<td>0.0043</td>
<td>0.9477</td>
</tr>
<tr>
<td>ln_RW_group4</td>
<td>0.2682</td>
<td>0.1604</td>
<td>2.7967</td>
<td>0.0945</td>
</tr>
<tr>
<td>ln_RW_group5</td>
<td>0.0095</td>
<td>0.0812</td>
<td>0.0138</td>
<td>0.9064</td>
</tr>
<tr>
<td>ageind_group0</td>
<td>-3.3994</td>
<td>0.2536</td>
<td>179.7134</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group1</td>
<td>-3.9201</td>
<td>0.5222</td>
<td>56.3508</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group2</td>
<td>-3.4804</td>
<td>0.2622</td>
<td>176.1298</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group3</td>
<td>-4.2322</td>
<td>0.2588</td>
<td>267.4123</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group4</td>
<td>-5.0869</td>
<td>0.2552</td>
<td>397.2795</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group5</td>
<td>-4.5357</td>
<td>0.2598</td>
<td>304.8348</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group6</td>
<td>-4.7324</td>
<td>0.4779</td>
<td>98.0508</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group7</td>
<td>-2.7256</td>
<td>0.2410</td>
<td>127.9117</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Table 5: Effect of Rye Deflated Wages With Various Interactions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>female_ind</td>
<td>-0.0540</td>
<td>0.0326</td>
<td>2.7524</td>
<td>0.0971</td>
</tr>
<tr>
<td>Parish_ind_2</td>
<td>0.1836</td>
<td>0.0755</td>
<td>5.9202</td>
<td>0.0150</td>
</tr>
<tr>
<td>Parish_ind_3</td>
<td>0.2042</td>
<td>0.0711</td>
<td>8.2555</td>
<td>0.0041</td>
</tr>
<tr>
<td>Parish_ind_4</td>
<td>0.2343</td>
<td>0.0701</td>
<td>11.1655</td>
<td>0.0008</td>
</tr>
<tr>
<td>Parish_ind_5</td>
<td>0.2721</td>
<td>0.0645</td>
<td>17.8257</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>HisClass7</td>
<td>0.2420</td>
<td>0.7945</td>
<td>0.0928</td>
<td>0.7607</td>
</tr>
<tr>
<td>HisClass9</td>
<td>0.9593</td>
<td>0.9049</td>
<td>1.1239</td>
<td>0.2891</td>
</tr>
<tr>
<td>HisClass10</td>
<td>0.6364</td>
<td>0.5214</td>
<td>1.4898</td>
<td>0.2223</td>
</tr>
<tr>
<td>HisClass11</td>
<td>0.1070</td>
<td>0.5897</td>
<td>0.0329</td>
<td>0.8560</td>
</tr>
<tr>
<td>HisClass12</td>
<td>-0.3152</td>
<td>0.5510</td>
<td>0.3272</td>
<td>0.5673</td>
</tr>
<tr>
<td>HisClassMissing</td>
<td>0.3045</td>
<td>0.5795</td>
<td>0.2761</td>
<td>0.5993</td>
</tr>
<tr>
<td>decade_ind_1820</td>
<td>0.6017</td>
<td>0.0842</td>
<td>51.0593</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>decade_ind_1830</td>
<td>0.8464</td>
<td>0.0779</td>
<td>117.9678</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>decade_ind_1840</td>
<td>0.6520</td>
<td>0.0785</td>
<td>68.9692</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>decade_ind_1850</td>
<td>0.4944</td>
<td>0.0788</td>
<td>39.4040</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>decade_ind_1860</td>
<td>0.4253</td>
<td>0.0807</td>
<td>27.7483</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>mantal_ind_0</td>
<td>-0.1111</td>
<td>0.4634</td>
<td>0.0575</td>
<td>0.8105</td>
</tr>
<tr>
<td>mantal_ind_1</td>
<td>-0.3319</td>
<td>0.7996</td>
<td>0.1724</td>
<td>0.6780</td>
</tr>
<tr>
<td>ln_RW_HisClass7</td>
<td>-0.4189</td>
<td>0.3937</td>
<td>1.1322</td>
<td>0.2873</td>
</tr>
<tr>
<td>ln_RW_HisClass8</td>
<td>-0.2874</td>
<td>0.1909</td>
<td>2.2663</td>
<td>0.1322</td>
</tr>
<tr>
<td>ln_RW_HisClass9</td>
<td>-0.8869</td>
<td>0.4485</td>
<td>3.9104</td>
<td>0.0480</td>
</tr>
<tr>
<td>ln_RW_HisClass10</td>
<td>-0.5622</td>
<td>0.2773</td>
<td>4.1098</td>
<td>0.0426</td>
</tr>
<tr>
<td>ln_RW_HisClass11</td>
<td>-0.2222</td>
<td>0.3057</td>
<td>0.5283</td>
<td>0.4673</td>
</tr>
<tr>
<td>ln_RW_HisClass12</td>
<td>-0.1204</td>
<td>0.2822</td>
<td>0.1821</td>
<td>0.6696</td>
</tr>
<tr>
<td>ln_RW_HisClassMissin</td>
<td>-0.3469</td>
<td>0.3003</td>
<td>1.3345</td>
<td>0.2480</td>
</tr>
<tr>
<td>ln_RW_mantal_ind_0</td>
<td>0.0877</td>
<td>0.2490</td>
<td>0.1241</td>
<td>0.7246</td>
</tr>
<tr>
<td>ln_RW_mantal_ind_1</td>
<td>0.1810</td>
<td>0.3953</td>
<td>0.2096</td>
<td>0.6471</td>
</tr>
<tr>
<td>ln_RW_mantal_ind_2</td>
<td>0.0299</td>
<td>0.1234</td>
<td>0.0588</td>
<td>0.8084</td>
</tr>
<tr>
<td>ageind_group0</td>
<td>-2.2076</td>
<td>0.3199</td>
<td>47.6324</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group1</td>
<td>-3.4977</td>
<td>0.3254</td>
<td>115.5627</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group2</td>
<td>-3.9323</td>
<td>0.3293</td>
<td>142.5705</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group3</td>
<td>-4.7289</td>
<td>0.3250</td>
<td>211.7766</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group4</td>
<td>-5.4680</td>
<td>0.3207</td>
<td>290.6725</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group5</td>
<td>-4.8727</td>
<td>0.3248</td>
<td>225.0114</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group6</td>
<td>-4.3682</td>
<td>0.3237</td>
<td>182.1593</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>ageind_group7</td>
<td>-3.1435</td>
<td>0.3212</td>
<td>95.7653</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
CLASS8 is the HISCLASS base case, \textit{decade\_ind\_1810} is the decade base case (corresponding to 1813-1820), and \textit{mantal\_ind\_2} is the mantal base case.

For our various parish and decade dummies, we find significantly positive coefficients, indicating that our base parish and base decade are particularly low mortality. While our various HISCLASS and \textit{mantal} dummies have varying signs on their coefficients, none are statistically significant. When examining our wage interaction effects, the interactions with \textit{mantal} are slightly positive and not statistically significant. In general, there is a lack of statistical significance on many of these variables due to number of controls used in this regression.

The main wage interaction effects of interest are the interactions with HISCLASS. All of these effects are negative, but only HISCLASS 9 and HISCLASS 10 are statistically significant. Although they are not statistically significant, there is a larger response of HISCLASS 7 to variations in wages than HISCLASS 8. HISCLASS 9 and 10 both have larger and statistically significant negative responses. HISCLASS 11 and 12 both have smaller responses, as does the missing HISCLASS category. The small responses of HISCLASS 12 is confusing, as we would expect low skill agricultural laborers without land to be among the most responsive to real wage variations. However, the other values have a coherent interpretation. For the missing and HISCLASS 11, we know that individuals in this category are typically elderly, and as we found in the previous section, the elderly tend to have lower responses to wage variation. For HISCLASS 7, the response is somewhat larger that for HISCLASS 8; although both groups are on the higher end of the manual labor social hierarchy, those in HISCLASS 8 are farmers and are thus less dependent on the market price of grain, as they can grow their own. For HISCLASS 10, there is a larger and statistically significant response; however, since these individuals often had some limited farm land on which to grow their own crops, there may have been some insulation from real wage variation. HISCLASS 9 has neither the higher paying skills of HISCLASS 7 nor the land of HISCLASS 10, and they thus have the largest response to wage variation.

There are some possible explanations as to why HISCLASS 12 doesn't exhibit the level of response to real wages as would be expected. One reason is that agricultural laborers often received payment "in kind", i.e. in the form of food and accommodations. This may have insulated them from temporary shocks in real wages. Another reason may be that agricultural laborers were more likely than other groups to be migratory. Thus, it may be that local records were less likely to record agricultural laborers, particularly if they were on the road looking for work in response to spikes in the price of grain. Regardless, this lack of response in HISCLASS 12 deserves further research.

9 Conclusion

As stated earlier, we're interested in answering three questions about the Malthusian hypothesis - does it provide a compelling explanation for population dynamics in pre-industrial Sweden, did it function primarily through the positive or preventive check, and when did it end? Using a VAR model, we find that a Malthusian model does provide a good explanation for population dynamics in Sweden over the period 1751 to 1870. This Malthusian regime can be described as
high pressure, as most of the population adjustment comes from a large and statistically significant positive check, whereas the preventive check was smaller and not significant over many sub-periods. Finally, we provide evidence that Sweden began transitioning out of a Malthusian regime in the early to mid eighteenth century. By the end of our time period under consideration, the preventive check was statistically insignificant and the positive check was much smaller in magnitude and barely significant.

Although crude birth rates do not have a significant accumulated response to a shock in wages, the timing of birth rates are affected by these shocks; the population adjustment occurs primarily through the shocks’ effects on non-infant death rates. This finding is in accordance with Eckstein et al. (1984), which found that "wages affect predominantly the timing of births, whereas the persistent effect of wages on population growth arises from the reduction in mortality, and quantitatively the reduction in non-infant mortality" (313).

Finally, we provide compelling evidence that birth rates respond to shocks in wages at least in part through unobserved health effects, in accordance with Fertig and Pfister (2012). This would suggest that models which treat birth rates as contemporaneous exogenous to real wages may suffer from misidentification in interpreting the response of birth rates to shocks in wages as evidence of intentional family planning. Thus, such models may overestimate the existence of preventive checks in pre-industrial European nations.

We also used a grouped proportional hazard model to examine how the positive check functioned at an individual level in five Swedish parishes. We found that the positive check was present in this area over the same time span, and was much more significant for prime-age people than infants or the elderly. When examining the affect of shocks in real wages on people of different occupations, we saw mixed results; however, there was evidence that semi-landless Swedish farmers and Swedes engaging in non-agricultural manual labor were more susceptible to shocks in real wages.

There are two major conclusions we would draw from our results in this paper. First, this evidence presented here suggests that the Malthusian model is a very plausible candidate for describing pre-industrial societies. This in turns supports the work of Galor (2011) and Clark (2008) among others, whose work requires pre-industrial societies to be subject to Malthusian pressures. While some critics target the use of Malthusian models in this research agenda, we believe the evidence from Sweden suggests that this model have an important role to play in understanding long-run economic growth.

That being said, our second conclusion is that it is unreasonable to assume that population dynamics in all pre-industrial societies functioned in similar ways. In many ways, pre-industrial Sweden is an ideal candidate to find evidence for the Malthusian model. It was a poor, highly agrarian society without social structures designed to support the poor and vulnerable. If any nation with reliable records could present evidence of a positive or preventive check, it would be Sweden. Thus, we do not think our results here conflict with other researchers’ results that show England escaping the Malthusian trap by the seventeenth century. Pre-industrial Sweden and pre-industrial England were two very different countries, and it is unreasonable to expect the
same population models to describe them both. Thus, we would caution anyone from assuming that the population of any place at any given time, whether societies from hundreds of years ago or modern sub-Saharan countries, follows a Malthusian model without examining the evidence closely.

References


**Figure Appendix**