

## **ECONOMICS**

### **Malthus Was Right: Explaining a Millennium of Stagnation**

by

**Jakob B. Madsen**

Business School, The University of Western Australia

**Peter E. Robertson**

Business School, The University of Western Australia

and

**Longfeng Ye**

RMIT University

**DISCUSSION PAPER 19.16**

# Malthus Was Right: Explaining a Millennium of Stagnation

Jakob B. Madsen\*

Peter E. Robertson

University of Western Australia

University of Western Australia

Longfeng Ye  
RMIT University

May 16, 2019

## Abstract

**Abstract:** We develop a simple Malthusian growth model with continuous productivity growth and derive the stationary steady-state equilibrium. We show that linearization around the steady-state gives an empirically tractable model of Malthusian wage and population behaviour using the familiar concept of  $\beta$  – convergence. Our empirical strategy addresses the concern in the literature over model identification and inconsistent parameter estimates. Based on newly constructed population data, we estimate wage and population growth models using panel data for up to 17 countries from 900CE to 1870CE. Our results provide the first time-series evidence of a strong Malthusian trap that was pervasive across countries and time before the 19<sup>th</sup> century industrial revolution.

**Keywords:** Economic Growth, Economic Development, Convergence, Technological Progress, Malthusian Trap.

**JEL:** O4, O47, O1, N00

---

\*Corresponding author: Jakob Madsen, Department of Economics, University of Western Australia, WA 6009, jakob.madsen@uwa.edu.au. Helpful comments and suggestions from: anonymous referees, seminar participants at the University of Western Australia; Elena Capatina and participants at the 5th Workshop of the Australasian Macroeconomic Society, Queenstown, Nov. 2018, are gratefully acknowledged. Financial support from the Australian Research Council is gratefully acknowledged (grants DP150100061 and DP170100339).

# 1 Introduction

From antiquity to the industrial revolution, the life of the broad population was characterized by a near subsistence existence. The conditions responsible for this economic stagnation, and what happened subsequently to permit over a century of growth, is a quintessential question in the study of Economics.

Since Malthus's (1798) "Essay on the Principle of Population", economists have realised that wage stagnation may be a consequence of the "Malthusian trap", whereby improvements in land or technology generated increases in population and so suppressed wage and income growth. The Malthusian trap remains a central tenant of unified growth theory and the related contemporary literature on long run growth (Kremer, 1993; Galor and Weil, 2000; Hansen and Prescott, 2002; Clark, 2007; Galor, 2011; Strulik and Weisdorf, 2008, 2014). Moreover, the ongoing failure of many countries to attain sustained increases in per capita incomes means that Malthusian population dynamics remain an important part of the literature on poverty and economic development.

Despite its widespread acceptance, however, the Malthusian trap hypothesis does not enjoy strong empirical support. While many studies of pre-industrial economies have found evidence of statistically significant population responses to changes in incomes, the size of the response is typically found to be very small. Consequently, there is relatively little evidence to show that Malthusian demographic responses were large enough to suppress wage growth (Persson, 2008; Allen, 2008; Guinnane, 2011). For example, the implied half-lives of adjustment from these small coefficients typically exceed several centuries. This means that a positive technological improvement would increase wages sharply in the short run, and this higher wage could persist for centuries before reverting to the steady-state. In this sense the "iron-law" of wages would be extremely weak (Lee, 1980; Crafts and Mills, 2009; see also discussion below).

In addition to finding small coefficients, the existing evidence also often finds that Malthusian mechanisms had completely broken down by the 18<sup>th</sup> century or even earlier, with estimates of demographic responses to wage incomes becoming insignificant in latter sample periods. This is at odds with the fact that there was an acceleration in population growth in the 18<sup>th</sup> and 19<sup>th</sup> centuries, and so casts further doubt on the relevance of the Malthusian model for understanding the changes in population growth.

There are also competing theories of pre-industrial stagnation. Specifically, a lack of technological progress will also generate a stationary wage rate, irrespective of whether there is a Malthusian population trap. A slower rate of technological progress may have

resulted from a variety of causes, including, an absence of market and legal institutions, a lack of population and scale effects or, a lack of human capital (Rosenberg and Birdzell, 1987; Acemoglu et al., 2005; Greif, 2006; Tabellini, 2010; Mokyr, 2011).

Thus the existing literature provides little evidence on whether the Malthusian trap was an important cause of economic stagnation. In this paper we therefore investigate the Malthusian stagnation hypothesis focusing on the key questions of whether there was a Malthusian trap that was sufficiently strong and widespread to be an important cause of the economic stagnation that permeated life before the industrial revolution. We also investigate how widespread the Malthusian model was across countries, and whether it existed up until the 19<sup>th</sup> century, or whether it broke down earlier?

We make the following contributions to the literature. The first innovation is our empirical strategy. We derive a model of household fertility choice and derive the steady-state income and wage levels. We show that linearizing around the steady-state yields a parsimonious empirical characterization of Malthusian population and wage dynamics. This single equation approach allows us to estimate the key underlying structural parameters using a standard error correction model, and thereby obtain direct estimates of the speed of convergence, as well as address simultaneity problems that have been a concern in the existing literature.

Second, to estimate the model, we construct a new per capita income and population data set for 17 European countries over the period 900CE-1870CE. As discussed below, population data have become available at 1, 5 and 10-year frequencies starting around 1600 or earlier for most countries in our sample. These provide a much more consistent time series than the standard sources of McEvedy and Jones (1978) and Maddison (2003), noting that Maddison (2003) draws heavily on McEvedy and Jones (1978) in his earlier estimates. Specifically, the data of McEvedy and Jones (1978) are only available on centennial or semi-centennial frequencies, thus making them less useful for econometric analyses of the Malthusian model.

Our larger data set offers several important advantages over the existing time series literature. First the large sample size increases the efficiency of the estimates. Second, it allows us to test the generality of the Malthusian mechanism by including up to 17 countries, whereas the existing literature relies almost entirely on single country data. This is important since it is often stressed that the age at marriage, which is regarded as one of the main methods of fertility regulation, differed significantly across European countries. (Hajnal, 1965). Third, our panel data also allow us to control for common time-varying unobserved effects. Fourth, the balanced nature of our panel enables us to allow for

contemporary cross-country correlation between the error terms based on the seemingly unrelated regression (SUR) method and, at the same time, allow for heteroscedasticity and serial correlation. This method not only provides substantially more efficient estimates than the OLS estimator, it also caters for unobserved endogenous effects (Egger, 2001).

We find strong evidence of a Malthusian trap across 17 countries for the thousand year period before the 19<sup>th</sup> century - with very rapid convergence of population wages and per capita incomes across our full sample. Thus, our empirical model and panel data yield the first time-series econometric evidence of widespread Malthusian forces in Europe, and other countries, that are sufficiently strong to explain the observed economic stagnation from the middle ages to the industrial revolution. Moreover, contrary to much of the existing literature, we find that there was no weakening of the Malthusian trap in the centuries leading up to the industrial revolution. This is consistent with unified growth models that suggest that late 19th century Europe was characterized by a post-Malthusian era - where wages began to rise due to technological progress, but income still had strong positive impacts on fertility (Galor, 2011).

The paper is organized as follows. Section 2 gives a brief overview of the relevant literature on estimating the convergence rate in Malthusian models. Section 3 then introduces a simple model of household fertility choice and derives the steady state equilibrium for wages and population. Section 4 discusses the transitional dynamics of the model and relates the rate of convergence to parameters of the utility and production functions. Section 5 describes the empirical model and gives a detailed description of the data. Section 6 reports the results from estimating the model for up to 17 OECD countries from 900CE to 1870CE and compares the results with the existing literature. Section 7 summarises the findings and concludes.

## 2 Related Literature

The Malthusian trap is the seminal explanation for the stagnation of living standards prior to the industrial revolution and is also a central premise of unified growth theory (Galor and Weil, 2000; Galor, 2011). Galor and Moav (2002), Clark (2007), and Ashraf and Galor (2011) argue that Malthusian mechanisms are a key reason for the economic stagnation since the Neolithic era.<sup>1</sup> As we document below, there is also a significant

---

<sup>1</sup> Other notable studies that use the Malthusian framework include: Postan (1972), Phelps Brown and Hopkins (1981), Kögel and Prskawetz (2001), Hansen and Prescott (2002), Strulik and Weisdorf

body of econometric research aimed at testing the Malthusian model.

Nearly all studies find some evidence that wages have had a significant positive affect on population growth. Nevertheless, most studies also find that the key parameter estimates are very small and, when compared across different centuries, often become insignificant in the 18<sup>th</sup> or 19<sup>th</sup> centuries, when population and income growth were most rapid. The finding of very small coefficients is problematic for the Malthusian-trap hypothesis as it implies that technological shocks would have a gradual and weak impact on population but an immediate and persistent effect on wages. This is sometimes referred to as “weak homeostatisis” (Lee and Anderson, 2002; Crafts and Mills, 2009). It implies that a positive technology shock will increase population only gradually and hence exert very little downward pressure on wages so that wages may exceed the steady-state “iron-law” level for centuries.

Much of the existing evidence therefore implies that Malthusian mechanisms were not an important cause of economic stagnation because the effect of population on wages was too weak. Thus, in discussing the literature, it is necessary to focus on not only the statistical significance of Malthusian models but also on the economic size of the key parameter estimates to clarify whether wage and population interactions were large enough to be economically important.

## 2.1 Malthusian Econometrics Literature

Building on pioneering studies by Lee (1973, 1980), Weir (1984) and Bailey and Chambers (1993), many studies find evidence of statistically significant Malthusian demographic and wage responses. For pre-industrial England from the 16<sup>th</sup> to the 19<sup>th</sup> centuries, these include Lee and Anderson (2002), Nicolini (2007), Crafts and Mills (2009), Kelly and Ó Gráda (2012) and Møller and Sharp (2014).<sup>2</sup> Similarly Klemp and Møller (2016) find evidence for Denmark (1824-1890), Norway (1776-1853) and Sweden (1777-1873). Lagerlöf (2015) finds evidence of Malthusian demographic effects in Sweden, and Voutilainen (2015) finds significant Malthusian effects in Finland and Sweden. Edvinsson et al. (2015) also present evidence for Sweden, and Chiarini (2010) presents evidence for Italy.

Nevertheless, Lee and Anderson (2002) find that the estimated coefficients are very small and there was, at best, only very slow adjustment to the Malthusian equilibrium, with (2008, 2014), Madsen et al. (2010), Vollrath (2011), and Chaney and Hornbeck (2016).

<sup>2</sup> A seminal study by Lee (1973), estimated a structural equation model for England 1250-1700 but found no relationship between wages and population growth.

implied half-lives of adjustment of population and wages of over a century. Crafts and Mills (2009) similarly find that the elasticity of population with respect to the wage after 1646 is either insignificant or is very small, implying a half-life of adjustment exceeding three to four centuries. Lagerlöf (2015) finds significant but economically small effects for Sweden that are similar in magnitude to Lee and Anderson (2002), and Chiarini (2010) finds mixed results in Italy. Murphy (2010) likewise finds only very limited support for the Malthusian model in France.<sup>3</sup>

There is also a related literature that explores the Malthusian model using vector autoregression (VAR) approaches. Nicolini (2007) finds that there is evidence that higher wages increased birth rates up until 1740 but after that the relationship becomes negative, indicating a breakdown of the Malthusian mechanism well before the industrial revolution.<sup>4</sup> Nicolini (2007) also finds that only 13.9 percent of the variance in the crude birth rate is explained by variations in the log of the wage rate over his whole sample period, 1541-1840, again suggesting that Malthusian effects were not very important overall. Results of very similar magnitude are found by Voutilainen (2015) for Sweden and Finland. Craft and Mills (2009) also revisit Nicolini's (2007) VAR model and confirm his results, also noting the weakening and breakdown of the Malthusian model by the 1800s. Møller and Sharp (2014) use a non-stationary VAR model and, likewise, find significant but small coefficients. Møller and Sharp (2014) cast further doubt on the existence of a Malthusian trap; finding that the assumption of stationarity of their explanatory variables is rejected.<sup>5</sup> As noted by Møller and Sharp (2014), without stationarity, the Malthusian model cannot explain a constant deterministic wage. They conclude that England entered a post-Malthusian phase as early as 1500, where Malthusian responses were outweighed by productivity growth.

With this lack of evidence of large Malthusian effects there has been an imperative in the recent literature to explore alternative approaches and data sources. Perhaps the strongest evidence for the importance of Malthusian responses is Kelly and Ó Gráda (2012). As they note, their results are an exception to the existing literature, which they attribute to the use of parish-level data and, hence, a greater number of observations. They also note the need for further research using alternative data sets.

---

<sup>3</sup> Crafts and Mills (2009) and Lee (2015) note further that a feature of the literature is that the estimated elasticity of wages with respect to population often violates the constraint that factor income shares cannot exceed unity. This points toward biased estimates as a result of feedback effects, identification issues and measurement errors.

<sup>4</sup> Nicolini (2007) also finds weak and inconsistent evidence for the effect of wages on population, which he attributes to a lack of data to model the appropriate lags.

<sup>5</sup> They also note that the finding of stationarity in the data used by Nicolini (2007) and Crafts and Mills (2009) is a fragile result that depends on a few very early observations in their data.

Ashraf and Galor (2011) also find support for the Malthusian model using a novel approach and longer run data. Specifically, they examine population and per capita income levels across a range of countries during 1-1500CE and find that variations in the timing of the Neolithic Revolution in each country affected population density but did not affect per capita incomes. Ashraf and Galor's (2011) results suggest that population levels responded to income over the very long run. Nevertheless, their results do not necessarily imply that demographic responses were large. For example, their results are consistent with a scenario of weak Malthusian responses where technology shocks generate immediate and persistent wage increases, but only generate significant increases in population in the very long run, perhaps after several centuries.

Thus, the literature is typified by very mixed results regarding support for the Malthusian trap and its importance as a key cause of economic stagnation. Specifically, there is an absence of evidence to show that the interaction between population and wages was strong enough to be the key cause of wage stagnation.<sup>6</sup> A typical estimate is that of Crafts and Mills (2009) who conclude that the half-life of adjustment of wages and population to a shock exceeds three centuries. With such strong persistence, or near non-stationary dynamics, even very occasional positive productivity gains could generate persistent wage increases above the steady-state or "iron-law" wage level. Instead, the principal cause of stagnation could simply be a low rate of technological change. Further evidence is therefore needed to understand these historical sources of growth and hence, also, the constraints facing contemporary developing countries.

### 3 Household Fertility Choice

While many versions of Malthusian models exist, the empirical literature has lacked micro-foundations.<sup>7</sup> This means that the coefficients cannot be meaningfully related back to parameters of a behavioural model. In what follows, therefore, we describe a simple model of household fertility choice and extend it to derive the dynamic implications and a parsimonious empirical model for wage and population dynamics. The model makes the relationship between the parameters of the utility and production functions and the

---

<sup>6</sup> It is worth noting that the evidence generally also shows that only the preventative check (the interaction between wages and fertility) is found to be significant, with most of the literature finding the positive check (the interaction between income and mortality) to be unimportant. In what follows we focus on the effect of incomes on population growth implicitly treating mortality as a constant.

<sup>7</sup> Lee and Anderson (2002) derive a time path for population and wages but their model is not based on micro-foundations. Ashraf and Galor (2011) and Galor (2011) do consider micro-foundations and dynamics but focus on steady-state relationships in their empirical applications.

empirical model explicit. It also shows how the Malthusian stagnation hypothesis can readily be tested from the size of the coefficient on initial wages or population.

### 3.1 Production

We assume that aggregate output is given by a production function that exhibits diminishing returns to labour. Assuming technological progress can substitute for land, and other inputs, we can simply aggregate all non-labour inputs into one term,  $X_t$ . For convenience we can therefore think of  $X_t$  as an aggregate of “effective non-labour inputs” including a total factor productivity term.<sup>8</sup>

Given constant returns to all factors, output,  $Y_t$ , for an economy at a point in time, or over a generation, is given by a quasi-concave production function

$$Y_t = F(X_t, L_t), \tag{1}$$

where  $L_t$  is labour,  $F_X > 0$ ,  $F_L > 0$  and  $F(X, L)$  is homogeneous of degree 1. Output per worker is

$$y_t = F(X_t/L_t, 1) = f(x_t), \tag{2}$$

where  $x_t \equiv X_t/L_t$ . Thus, output per worker is increasing in effective land per worker. It will be useful to define the income share of non-labour inputs as  $\alpha(x_t) = x_t f'(x_t)/f(x_t)$ . This completes the description of the technology.

### 3.2 Preferences

We assume there is a representative household that allocates time between consumption and child rearing. A household in generation  $t$  has a strictly quasi-concave utility function given by

$$u_t = u(c_t, m_t), \tag{3}$$

where  $c_t$  is lifetime consumption of goods for generation  $t$  and  $m_t \equiv L_{t+1}/L_t$  is the number of surviving children per adult.

The representative household in each generation is endowed with a unit of time that is employed inelastically to produce output,  $y_t$ . The household’s problem can be characterised as maximizing (3) subject to the constraint  $c_t = y_t - p_t m_t$  where  $p_t$  is the per

---

<sup>8</sup>For convenience, however, we sometimes simply refer to this as the level of technology.

child cost of child rearing activities in terms of units of consumption goods. This implies an indirect utility function given by

$$\nu(p_t, y_t). \tag{4}$$

Roy's identity gives the Marshallian demand for children,

$$m_t = m(p_t, y_t), \tag{5}$$

where the partial derivative of the price of children is negative,  $m_p < 0$ , and the partial derivative of the second term is non-negative,  $m_y \geq 0$ .<sup>9</sup> Thus (5) shows that the demand for children depends on the relative price of children, and incomes.

Next, we consider the supply of children. We assume that a fixed proportion,  $\tau$ , of output is allocated to child rearing for each child.<sup>10</sup> Since the only other use of output is consumption, we have  $c_t = y_t(1 - \tau m_t)$  and so the supply price of child rearing is  $p_t = \tau y_t$ . Thus, we can write the demand for children as a function of income only,  $m(p_t, y_t) = m(\tau y_t, y_t) \equiv \mu(y_t)$ .

### 3.3 Steady State

We define a steady-state equilibrium in our Malthusian model as a situation where per capita incomes are constant.

**Definition 1.** A steady-state is a value of  $y$  such that  $y_{t+1} = y_t$  for all  $t$ .

From (2) this requires  $x_t \equiv X_t/L_t$  is constant. We denote the steady-state ratio of effective non-labour inputs per worker as  $x_t = x^*$ . A steady-state therefore requires that  $L_t$  grows at the same rate as  $X_t$ . Effective non-labour inputs,  $X_t$ , evolve exogenously at the rate  $1 + g = X_{t+1}/X_t$  and, hence, a steady-state also requires  $L_{t+1}/L_t = 1 + g$ . Population is endogenous, however, and must equal the demand for children in equilibrium. Therefore, on a steady-state, the demand for children,  $\mu(y^*) = m_t \equiv L_{t+1}/L_t$ , must satisfy

$$\mu(y^*) = 1 + g. \tag{6}$$

Equation (6) thus determines the unique steady-state equilibrium value of per capita

---

<sup>9</sup> In the interest of parsimony we think of this as a net demand for children, adjusting for expected mortality. This is a useful simplification since the literature suggests that changing birth rates have been the main determinant of population growth.

<sup>10</sup> Ashraf and Galor (2011) and Galor (2011, 2012) also use this assumption.

income,  $y^*$ , as a function of the parameters of the demand function for children, and the growth rate of non-labour inputs,  $g$ . We can summarise the key properties of the model with the following propositions.

**Proposition 1.** In the steady-state equilibrium, the equilibrium per capita income level,  $y^*$ , is independent of the level of non-labour factor inputs,  $X_t$ .

Thus an increase in non-labour resources,  $X_t$ , for example due to land clearing or a new technology, does not affect the steady-state wage or income level. Rather the level of income depends on the rate of growth of effective non-labour inputs. Proposition 1 thus highlights the difference between the neoclassical growth model and the Malthusian model. In the neoclassical model the rate of productivity growth determines the steady-state *growth rate* of per capita income. In a Malthusian model it determines the steady-state *level* of per capita income.<sup>11</sup>

Since  $y^*$  is uniquely determined by (6), the equilibrium has a unique ratio of effective non-labour to labour inputs,  $x^*$ , and it also defines a unique wage rate, since  $w^* = y^* - x^*f'(x^*)$ .<sup>12</sup> Hence, Proposition 1 shows that both wages and incomes are also constant despite continuous productivity growth. In what follows we also refer to this stationary equilibrium level of wages as the “iron-law” wage rate.<sup>13</sup>

Further, it will be useful to note that if non-labour inputs are fixed in supply so that  $g = 0$ , then  $\mu(y^*) = 1$  defines a subsistence level of income in which there is no population growth. This equilibrium where both wages and population are stationary is a special case of the more general Malthusian trap in (6) where wages are stationary but population is growing. It defines the minimum level of income necessary to prevent population collapse. Since there was slow population growth before the industrial revolution, we focus on the more general case where  $g \geq 0$ .

From (6), it can also be seen that different rates of technological change,  $g$ , imply different equilibrium levels of steady-state per capita incomes and wages. Hence, we have:

**Proposition 2.** The steady-state equilibrium levels of per capita income and wage rate are increasing functions of the growth rate of effective non-labour inputs,  $1 + g$ .

---

<sup>11</sup> By “productivity growth” we mean all exogenously growing non-labour factor inputs that affect  $g$ , such as land expansion and technological change.

<sup>12</sup> Note that the constancy of wages and per capita incomes follows from the model and the assumption of homogeneity of degree one in the production function, and does not require the assumption of an isoelastic production function.

<sup>13</sup> Karl Marx and Fredrich Engels popularised the term in their criticism of Malthus’s population theory. See Baumol (1983) for a discussion.

Specifically differentiating (6) gives

$$\frac{dy^*}{y^*} = \frac{1}{\sigma(1+g)} dg, \quad (7)$$

where  $\sigma \equiv \mu'(y)y/\mu(y)$  is the elasticity of demand for children with respect to income. An increase in the growth rate of effective non-labour inputs,  $g$ , will result in transitional growth of per capita incomes as the economy moves to the new steady-state income level. The size of the increase in per capita income in response to a change in  $g$  is inversely proportional to the elasticity of population growth with respect to income,  $\sigma$ .

Proposition 2 thus shows that there can be transitional income and wage growth in the Malthusian model if  $g$  changes. This will be important when discussing the evidence regarding periods of wage increase through history and when European countries entered a post-Malthusian era, as suggested by unified growth theories, such as Galor (2011), or whether the Malthusian model broke down well before the 19<sup>th</sup> century, as suggested by much of the existing econometric literature, as discussed in Section 2. Specifically Proposition 2 suggests that if  $\sigma$  declines, the Malthusian economy starts to behave more like a neoclassical growth model with respect to changes in the rate of technological progress,  $g$ .

### 3.4 Stability

Before proceeding we note the stability properties of the steady-state and the microeconomic factors that determine the rate of convergence which, as we will show, also depend on  $\sigma$ . Stability requires that there is a positive demographic response to rising incomes,  $\mu'(y) > 0$ .<sup>14</sup> Since the demand for children depends on both income and prices, that is,  $\mu(y) \equiv m(p, y)$ , the condition  $\mu'(y) > 0$  requires the sum of the price and income terms to be positive,  $m_p \tau + m_y > 0$ . Thus, a positive income effect,  $m_y > 0$ , is not a sufficient condition for a Malthusian equilibrium.

This condition can be also expressed in terms of  $\sigma$ . Since  $\sigma \equiv \mu'(y) y / \mu(y)$ , the condition  $\mu'(y) > 0$  implies  $\sigma > 0$ . Totally differentiating  $m(p, y) = \mu(y)$  gives  $\sigma = \eta_{mp} \eta_{py} + \eta_{my}$ , where  $\eta_{mp}$  is the partial elasticity of demand for children with respect to price,  $\eta_{py}$  is the partial elasticity of price with respect to income, and  $\eta_{my}$  is the partial elasticity of demand for children with respect to income. Since  $p = \tau y$ , we have  $\eta_{py} = 1$ , so that  $\sigma > 0$  implies  $\eta_{mp} + \eta_{my} > 0$ . We summarise this stability condition as:

---

<sup>14</sup>For convenience time subscripts are suppressed in this sub-section.

**Proposition 3.** The existence of a stable steady-state equilibrium requires  $\sigma > 0$ , which requires that the sum of the partial elasticity of demand for children with respect to income,  $\eta_{my}$ , and the partial elasticity of demand for children with respect to price,  $\eta_{mp}$ , is positive.

Intuitively, this means the income effect on the demand for children needs to be large relative to the absolute value of the Beckerian price effect, or substitution effect, away from child rearing that occurs because an increase in wages also increases the opportunity cost of child rearing. Thus, the existence of a stable “Malthusian trap” requires either a relatively inelastic price elasticity of demand for children or a relatively elastic income elasticity of demand for children or both.<sup>15</sup> There is no theoretical basis for presuming that these conditions will hold. Hence, whether a Malthusian trap existed or not is an empirical matter.

## 4 Convergence in the Malthusian Era

As noted in our discussion of the literature, even though many studies find significant estimates of demographic responses to wages, much of the literature has found estimates that imply very small values of  $\sigma$ . In this section we also show that a small value of  $\sigma$  will imply very slow convergence to the new steady-state in response to a change in any exogenous change and, therefore, that the path of wages and incomes will follow near non-stationary behaviour. This would imply that even a very slow rate of technological progress could generate a positive trend in wages.

To determine the rate of convergence recall that  $y_t = f(x_t)$ , where  $x_t \equiv X_t/L_t$ . For convenience we now assume  $f(x_t) = x_t^\alpha$ . Then, since  $X_{t+1}/X_t = 1 + g$ ,  $L_{t+1}/L_t = m_t$ , we have  $\ln y_{t+1} - \ln y_t = \alpha(x_t) g - \alpha(x_t) \ln m_t$ .<sup>16</sup>

A first-order Taylor-Series expansion of this expression around the steady-state level of per capita income,  $y^*$ , gives,

$$\ln y_{t+1} - \ln y_t = -\alpha\sigma^*(\ln y_t - \ln y^*) \tag{8}$$

---

<sup>15</sup>Note that this means that stability is not satisfied by the Cobb-Douglas utility function since  $\eta_{mp} = -1$ , and  $\eta_{py} = \eta_{my} = 1$  so that  $\sigma = 0$ . In the Cobb-Douglas case the income and price effects directly offset each other so that rising income has no effect on fertility. This example is also noted by Ashraf and Galor (2011).

<sup>16</sup>Following the literature such as Mankiw et al. (1992) we interpret this Cobb-Douglas case as an approximation to a more general form of  $f(x)$ . Specifically, differentiating  $f(x)$  gives  $\partial \ln y_t = \alpha(x_t) \partial \ln x_t$ , where  $\alpha(x_t) = x_t f'(x_t)/f(x_t)$ . Hence we have  $\ln y_{t+1} - \ln y_t \approx \alpha(x_t)(\ln x_{t+1} - \ln x_t) = \alpha(x_t) g - \alpha(x_t) \ln m_t$ . This approximation is exact in the Cobb-Douglas case.

where  $\sigma^* \equiv \mu'(y^*)y^*/\mu(y^*)$  and  $\alpha\sigma^*$  is the Malthusian per capita income convergence rate.<sup>17</sup> Equation (8) shows that the rate of convergence to a Malthusian steady-state equilibrium depends on the income elasticity of demand for children,  $\sigma^* > 0$ , and the non-labour share of income,  $\alpha > 0$ . Larger values of  $\alpha$  and  $\sigma$  mean a faster rate of convergence to a new equilibrium. Conversely if  $\sigma$  small, so that income only has a small effect on population growth, the rate of convergence will be slow. The rate of convergence also depends on the non-labour share of income,  $\alpha$ . A large value of  $\alpha$  implies that the economy will converge quickly to a new equilibrium, whereas if  $\alpha$  is small, convergence will be slow.

It is well known that a first order difference equation of the form taken in (8) has a solution  $\ln y_t = (1 - \alpha\sigma^*)^t(\ln y_0 - \ln y^*) + \ln y^*$ .<sup>18</sup> Over a discrete interval,  $T$ , this gives

$$\frac{\ln y_T - \ln y_0}{T} = \beta (\ln y_0 - \ln y^*), \quad (9)$$

where  $\beta \equiv -(1 - (1 - \alpha\sigma^*)^T)/T$ . Thus, we have a first order difference equation that describes the rate of adjustment to the Malthusian steady-state.<sup>19</sup>

A similar convergence expression describes population growth. Specifically, since  $X_{t+1}/X_t = 1 + g$  and  $L_{t+1}/L_t = m_t$ , the definition  $x_t \equiv X_t/L_t$  implies that  $\ln x_{t+1} - \ln x_t = g - \ln m_t$ . Taking a first order Taylor-Series expansion of this expression around the steady-state and solving gives,  $\ln x_t = (1 - \alpha\sigma^*)^t(\ln x_0 - \ln x^*) + \ln x^*$ . Noting that  $\ln x^* = (1/\alpha) \ln y^*$  and taking a discrete interval  $T$  gives

$$\frac{\ln L_T - \ln L_0}{T} = \beta \ln L_0 - \beta \ln X_0 + \left(\frac{\beta}{\alpha}\right) \ln y^* + g. \quad (10)$$

Thus, as with the per capita income convergence model, a larger value of  $\sigma$  implies faster convergence of population to its steady-state growth path.

## 5 Empirics

Equations (9) and (10) suggest a single equation method for empirically testing the Malthusian model based on the familiar “ $\beta$  – convergence” framework or partial adjust-

<sup>17</sup> For more details see Appendix 1.

<sup>18</sup>The existence of an equilibrium requires  $|1 - \alpha\sigma| < 1$ . Furthermore, we focus on the monotonic convergence case where  $0 < 1 - \alpha\sigma < 1$ .

<sup>19</sup>Clearly the form of (9) is similar to the familiar expression from the extensive neoclassical growth model literature following Barro and Sala-i-Martin (1992) and Mankiw et al. (1992). Naturally however, the determinants of  $y^*$  will differ for Malthusian and neoclassical models.

ment model. Furthermore they suggest an empirical strategy to overcome some challenges identified in the existing literature, such as those discussed by Lee and Anderson (2002), Allen (2008), Crafts and Mills (2009), Guinnane (2011), and Lee (2015), around estimation bias from feedback effects, identification issues and measurement errors in the usual empirical approach to the Malthusian model. Specifically, equations (9) and (10) can be stochastically specified as follows:

$$\Delta \ln y_{it} = \alpha_0 + \beta \ln y_{i,t-30} + Z_{it} \vartheta' + TD + CD + \varepsilon_{1,it}, \quad (11)$$

$$\Delta \ln L_{it} = \phi_0 + \gamma \ln L_{i,t-30} + \lambda \ln X_{i,t-30} + Z_{i,t} \xi' + TD + CD + \varepsilon_{2,it}, \quad (12)$$

where  $\Delta \ln y_{it}$  is the average annualized geometric growth rate of real per capita income over a 30-year interval;  $y_{i,t-30}$  is the initial real per capita income;  $TD$  is time fixed-effects dummies;  $CD$  is country fixed-effects dummies;  $\varepsilon_{1,it}$  and  $\varepsilon_{2,it}$  are stochastic error terms; and  $Z$  is a vector of control variables that governs the steady-state per capita income,  $y^*$ , where  $Z$  includes the number of epidemics,  $epi$ , and the number of great scientists,  $sci$ .  $X_{i,t-30}$  in equation (12) is the initial effective non-labour factor of production and is measured as cultivated land size (sum of crop and arable land),  $land_{i,t-30}$ , and the stock of great scientists,  $S_{it}^{sci}$ , which is a proxy for the initial level of technology. The coefficient of the lagged dependent variable  $\beta$  is thus the per capita income convergence coefficient from which we can obtain estimates of the convergence rate,  $\alpha\sigma^*$  in equations (9) and (10). Likewise,  $\gamma$  is the convergence coefficient in the population equation.<sup>20</sup>

Our estimation strategy has an advantage over the traditional methods in which population is regressed on real wages and their lagged values and vice versa. Under the maintained hypothesis that there is a Malthusian mechanism, parameter estimates in standard population models with contemporaneous wages as explanatory variables will be biased. This is because population shocks create a negative correlation between real wages and population, and neutral technology shocks create a positive correlation. Since slow technological progress prevailed over the pre-industrial period, most shocks would have been population shocks. Thus, in a typical model that regresses population on wages, population shocks will bias the coefficient of real income towards zero and, so, create a bias against the Malthusian hypothesis. Since it is almost impossible to find external instruments for wages and population growth over time, the conventional estimation strategy is unlikely to provide unbiased estimates. The advantage of our empirical

---

<sup>20</sup>In equations (11) and (12) the coefficients in lagged income and population will differ from equations (9) and (10) if they are estimated with annual data, since the elasticity,  $\sigma$ , is defined with respect to demand for children over a lifetime. Specifically, in (11) and (12) with annual data we have  $\beta = \gamma \equiv -(1 - (1 - \alpha\sigma/N)^T)/T$ , where  $N$  is the number of calendar years in a lifetime.

specification is that wages and population are not simultaneously included in the same regression. Specifically, Eqs. (11) and (12) overcome this simultaneity problem because population is not an argument in Eq. (11) and income is not an argument in Eq. (12).

We estimate (11) and (12) in 30-year overlapping intervals to filter out erratic and cyclical movements in the data and to capture the long-run dynamics. The parameters of key interest are the coefficients of initial income and population; both of which must be significantly negative for the Malthusian equilibrium to exist and their absolute size indicates the strength of the Malthusian mechanism. The models are estimated using the cross-sectionally heteroscedastic and time-wise autoregressive estimator of Kmenta (1971), which allows for contemporary cross-country correlation between the error terms following the seemingly unrelated regression (SUR) principle. This estimator yields more efficient parameter estimates than the OLS estimator and caters for unobserved endogenous effects (Egger, 2001). This estimator requires the panel to be balanced and that the number of time periods exceeds the number of cross-sectional units. The estimator corrects for heteroscedasticity and serial correlation.

## 5.1 Data

We estimate equations (11) and (12) for a maximum of 17 OECD countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the UK) over periods of up to nearly a millennium from 900CE to 1870CE. For the UK, the income and population data are for England before 1700 and the wage data are for England throughout the entire period. We refer to “England” and the “UK” interchangeably in our discussion. As discussed in depth below, the baseline regressions are restricted to 9 countries and the periods 1470-1870 and 1600-1870 to ensure that our results are based on data of the best quality.

Income is measured as real wages and per capita GDP. Real wages are calculated as nominal average daily wages of skilled and unskilled building labour in major cities divided by consumer prices. More representative wage data are mostly not available (Allen, 2001; Humphries and Weisdorf, 2017). Real per capita income is measured as real GDP divided by population based on national sources and backdated using urbanization rates, where urbanization is measured as the fraction of the population living in cities with a population size of over 5,000.

A benefit of the wage data is that it is available on an annual basis, whereas the per

capita income data are often heavily interpolated. A potential disadvantage of the real wage data, except for the UK, is that they are measured as daily or hourly wages and, therefore, do not account for secular movements in annual hours worked.<sup>21</sup> This variation in days worked will be captured by the time-dummies in our regressions to the extent that the time-variation of annual hours worked followed the same trend across Europe. The per capita income data do not suffer from this problem of changes in working hours, and these data are available further back in time and for a larger number of countries than the real wage data.

The population data are primarily based on national sources and constitute an improvement over the standard population sources used in the literature, namely McEvedy and Jones (1978) and Maddison (2003). McEvedy and Jones (1978), who provide cross-country population data back to BCE at centennial frequencies before 1500CE and at semi-centennial frequencies thereafter, is the standard source of historical population data used in the literature. Maddison (2003), for example, draws heavily on this source in his 1820-1950 data and, almost entirely on this source in his pre-1850 estimates. However, several new population estimates have emerged since McEvedy and Jones (1978) was published and the long intervals between their observations limits the value of McEvedy and Jones's (1978) data for tests of the Malthusian mechanism. As detailed in the Data Appendix, data have now become available at 1, 5 and 10-year frequencies starting around 1600 for most countries in our sample and go back even further at 10-year frequencies for Germany, the Netherlands, and the UK. Furthermore, McEvedy and Jones (1978) rely on population estimates covering years that often do not coincide with their benchmark years, often with a 10 or 50 year difference (for example, the population data for a country in 1500 may be sourced in 1490 or even in 1450).

The data sources, availability, and the method of construction are discussed in depth for each country in the Data Appendix. Common for all countries is that the quality of the population data improves over time; particularly after the 16<sup>th</sup> century as information on number of deaths, births, population, houses and households become available from church and tax registers. In broad terms, the quality of the data falls into four distinct approximate periods: 950-1400, 1400-1600, 1600-1800, 1800-1870. Over the period 950-1400, the data are not of high quality and are often only available at centennial frequencies; particularly before circa 1100. Before 950, the data are either non-existent or

---

<sup>21</sup> Humphries and Weisdorf (2017) show that the number of working days per year has varied substantially over the past eight centuries in England and increased substantially from the middle ages to the early industrialization. See also Broadberry (2013) who finds that the annual number of holidays in Europe was reduced by around 50 days per year during and after the Protestant reformation and was further reduced by 50 days during the British Industrial Revolution as St. Monday, the practice of tolerating workers not turning up for work on the first day of the week, was gradually abolished.

are of poor quality; particularly during the dark ages (Russell, 1958, 1969; McEvedy and Jones, 1978). Most of the data over the period 950-1300 are from McEvedy and Jones (1978) and although their data are consistent with overlapping data from alternative sources, McEvedy and Jones (1978) provide only limited information on sources or evidence on how the data are constructed. From some of the source material they mention, it appears that, Russell (1958, 1969) is the main source for the medieval data, where Russell estimates population based mostly on information from grave-stone inscriptions and excavated skeletons.

The data frequency increases and the quality of the data improves after approximately 1400 as more frequent information becomes available from tax files, military documents, local registers, and parish registers and the quality of the data improves further during the 16<sup>th</sup> century as documents from parish registers expands rapidly. The period 1600-1800 constitutes a significant improvement over the previous period, 1400-1600, in that parish registers with all the relevant information on births, deaths, marriages, number of households as well as poll tax registers and census surveys become available for almost all countries in our sample. Furthermore, the time-frequency of the data increases substantially after circa 1600 by being mostly available in 1, 5 or 10-year intervals. The introduction of country-wide censuses starting from the second half of the 18<sup>th</sup> century in some countries constitutes a further improvement in the population data in the period 1800-1870 relative to previous periods. The availability of census data and relatively precise population reconstructions for our country sample ensure a high quality of the population data over the period 1800-1870.

The dominant methods for constructing population in the approximate period 1400-1800 are village registers, inverse projection, family reconstruction and hearth tax registers. Inverse projection requires a count of births, age structure, migration and deaths year-by-year and the size of the population in the initial year, and is generally considered reliable provided that the migration rates are low or available (Lee, 1985). Inverse projection is used in the pre-census period for many countries in our study, notably Finland, France, Germany, Italy, Japan, and Sweden. Family reconstructions have been used for England in the pre-census period while hearth tax registers have been used extensively in the population reconstructions, notably for Belgium, Germany, Ireland, and the Netherlands. For other countries, the population estimates are based on information from various documents; mostly parish registers, military surveys of the male adult populations between the ages of 10 and 60, and tax records (typically hearth counts). The largest challenge associated with the construction of population data before country-wide population censuses started being carried out, is scaling. Population estimates based on hearth taxes,

for example, require the average size of the household, and the average number of chimneys per household. Though population estimates based on inverse projection and family reconstruction are quite accurate at parish levels, not all parish archives have survived or years are missing.

Given these data quality considerations for our baseline regressions, we restrict attention to 9 countries (Austria, Belgium, France, Germany, Italy, the Netherlands, Spain, Sweden, and the UK) over the periods 1470-1870 and 1600-1870. The year 1470 is the first year at which real wages are available for these 9 countries in our sample, 1600 signifies an approximate year at which per capita income and population data improve in quality and estimates become available at shorter time-intervals, and 1870 is widely accepted as the start of the modern growth era (Greasley et al., 2013). Though long historical wage data are available for some countries that are not included in our sample, the starting year would, for example, be 61 years later if we included an additional country with the longest wage data that are currently available, noting, as stated, that a balanced panel is required for the SUR estimator used in our estimates.

In some of the regressions below, we expand the estimation period and the country sample, and use different income proxies to ensure that our results are not affected too much by measurement errors and the choice of country sample and estimation period. The consequence of measurement errors is that the coefficients of the key variables, initial income or population, may be biased. However, the bias can be positive as well as negative depending on the serial correlation of the measurement errors and their correlation with other regressors (Temple, 1998). As acknowledged in the literature, however, it is almost impossible to find time-varying instruments for income and population that vary at annual frequencies over the time-period considered here. Since our results are robust to measurement, different estimation periods and country samples, it is likely that our results are not affected significantly by measurement errors.

Figure 1 provides a visual summary of the real wage data reported as welfare ratios across countries for comparative purposes, where the welfare ratio is average annual earnings divided by the annual expenditure on a poverty line consumption bundle for a family. A welfare ratio in excess of one indicates a wage above the poverty line, while a ratio less than one means the family lives in poverty (Allen, 2001). The real wage data are converted to be equal to the average welfare ratios computed by Allen (2001) over the period 1700-1750. Deflating each country's wage data by the consumption poverty line effectively provides a purchasing power parity exchange rate that facilitates the visual comparison of the data across countries. Note, however, that potential measurement errors in the welfare ratios have no bearing on the results due to the inclusion of fixed

effects in the panel regression models.

Figure 1 shows that real wages remained low, although mostly well above the poverty line, throughout the period 1440-1870. A secularly decreasing wage trend is noticeable for Austria, Italy, and Spain throughout most of the four centuries considered, where the decline is most pronounced for Italy, while the trends for the other countries are partly disguised by the strong real wage fluctuations. Conversely, England is unique in experiencing an increasing wage trend from around 1600. For the rest, and most of the countries, Belgium, France, Germany, the Netherlands, and Sweden, the real wages tend to fluctuate around a relatively constant level in all four centuries. The cross-country variation in these trends reflects the well-known secular drift in the technological frontier moving north; away from the Mediterranean countries and towards North Western Europe over the second millennium CE (Landes, 2003).

The changes in wages over time have been the subject of much discussion over what they imply for the Malthusian model. Equations (6) and (7) and Propositions 1 and 2 provide some insight. First, the fact that wages exceed the subsistence level, indicated by a welfare ratio above unity in the Netherlands, Belgium and England, is at odds with a strong concept of a Malthusian trap where wages are not only constant but are at a subsistence level. Nevertheless, as shown in Proposition 2, varying levels of wages across countries are consistent with our Malthusian steady-state equilibrium when  $g > 0$ . Likewise, from Eq (6), different preferences and child rearing costs will also affect the steady-state income level and wage rate. Thus, different wage levels across countries and across time may be attributed to different preferences, child rearing costs and rates of technological progress and rates of growth in other non-labour inputs.

Second it can be seen that there is a time-trend in English wages. This is sometimes interpreted as evidence that the Malthusian model had ceased to operate by the 15<sup>th</sup> century in England (for discussion, see e.g., Mokyr, 2008, p.6 and pp.279-280). For example, Clark (2005) and Broadberry et al. (2015) come to different conclusions regarding whether there was a Malthusian trap in England, based on differences in wage trends observed in national income data and wage data. Clark (2005) finds that per capita income was stationary from the 14<sup>th</sup> century, but Broadberry et al. (2015) find modest, but positive, long run growth, which accelerates in the mid-17<sup>th</sup> century. Proposition 2, however, shows that an increase in  $g$  will increase wages as the economy experiences transitional growth between steady-states. The observed trend in English wages is therefore also consistent with the broader concept of a Malthusian steady state where  $g > 0$  and has fluctuated through time.

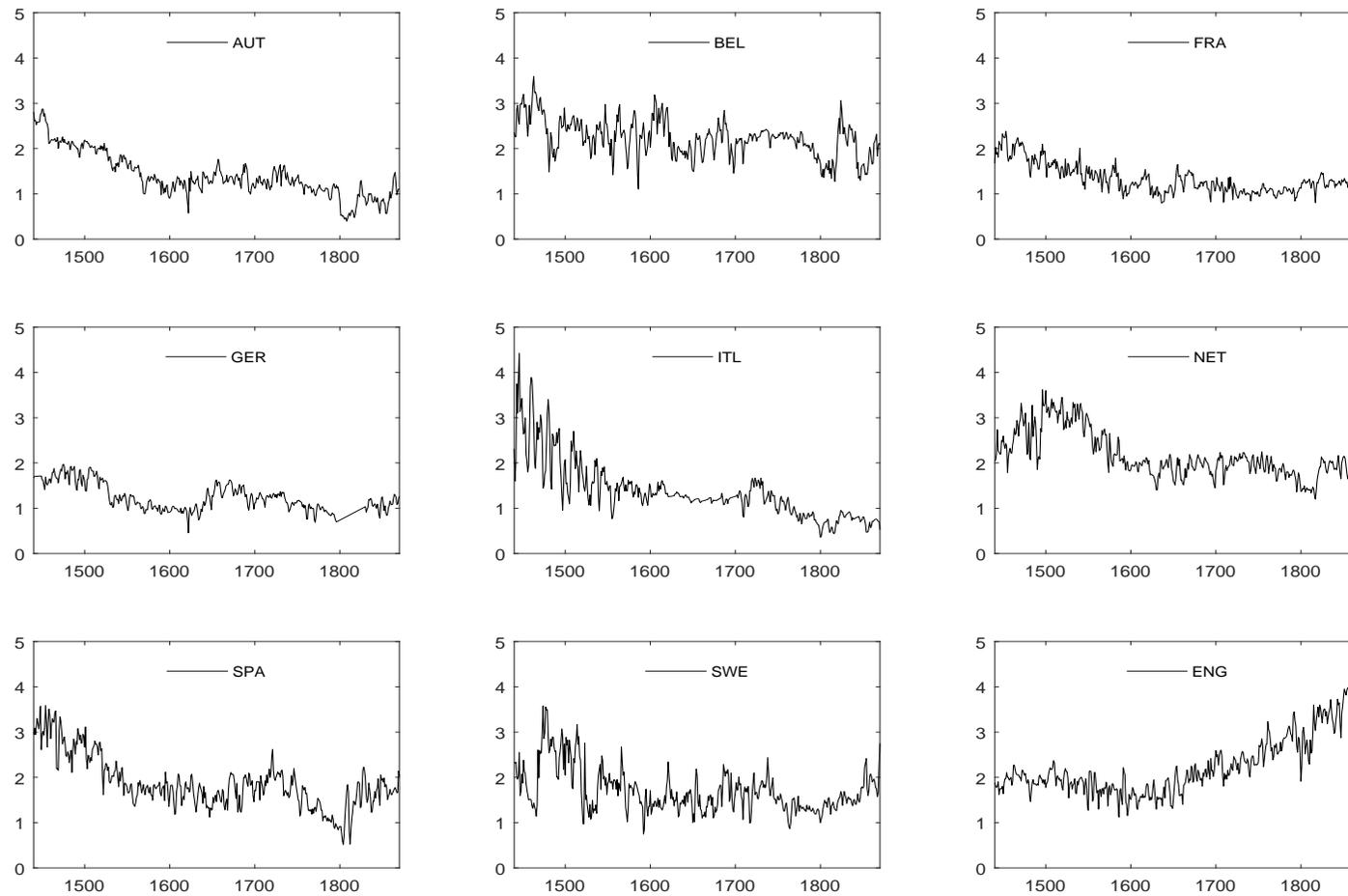


Figure 1: Welfare Ratios for 9 OECD Countries (1440-1870)

**Notes:** The data are real wages converted into welfare ratios by adjusting real wages to be equal to the welfare ratios computed by Allen (2001) over the period 1700-1750. The welfare ratio for Sweden over the period 1700-1750 is the average of that of Germany and the Netherlands.

Thus, a positive trend in the wage data, such as observed in the UK, has two possible explanations. As supposed by Broadberry et al. (2015) it may imply that the Malthusian model is weakening. In our model this corresponds to a fall in  $\alpha$  or  $\sigma$ . Or it may be that  $\alpha$  and  $\sigma$  are relatively stable and that, instead, there has been an acceleration in  $g$ . In particular an acceleration in technological progress marks the beginning of a post-Malthusian era, according to unified growth theory (Galor, 2011). It is therefore an empirical question as to which explanation is the more compelling. A key question of our empirical investigation, therefore, is to see whether there is any evidence of a decline in the convergence coefficients  $\beta$  and  $\lambda$  over time. If these decline over time it would suggest a breakdown of the Malthusian mechanism. But if they are relatively constant it would suggest that the increase in wages was a consequence of an acceleration in productivity growth, or the growth in non-labour inputs, while Malthusian effects still operated, which is consistent with Galor's (2005, 2011) characterization of a post-Malthusian era.

Figure 2 displays the evolution of populations across countries over the period 1440-1870. The population growth, which is generally low, and occasionally negative before 1800, is particularly high over the approximate period 1800-1870. The variation in population growth over time and across countries is consistent with the model, since it may reflect varying rates of technological progress, land acquisition, through clearing and expansion (the Netherlands), epidemics and emigration flows. Considering the pre-1800 period, population growth was particularly slow in Spain (Alvarez-Nogal and Prados de la Escosura, 2007). Similarly, the marked decline in population in Germany over the period 1600-1650 is a result of the Protestant-Catholic Thirty-Year-War, 1618-1648, which caused a very large loss of population in Germany in percentage terms (Greengrass, 2014).

The observed acceleration in population growth toward the end of the 18<sup>th</sup> and 19<sup>th</sup> centuries, also suggests that Malthusian effects were still operating and, hence, the emergence of a post-Malthusian regime. If, alternatively, the Malthusian model had begun to break down by this stage, the acceleration in population would require an alternative explanation.

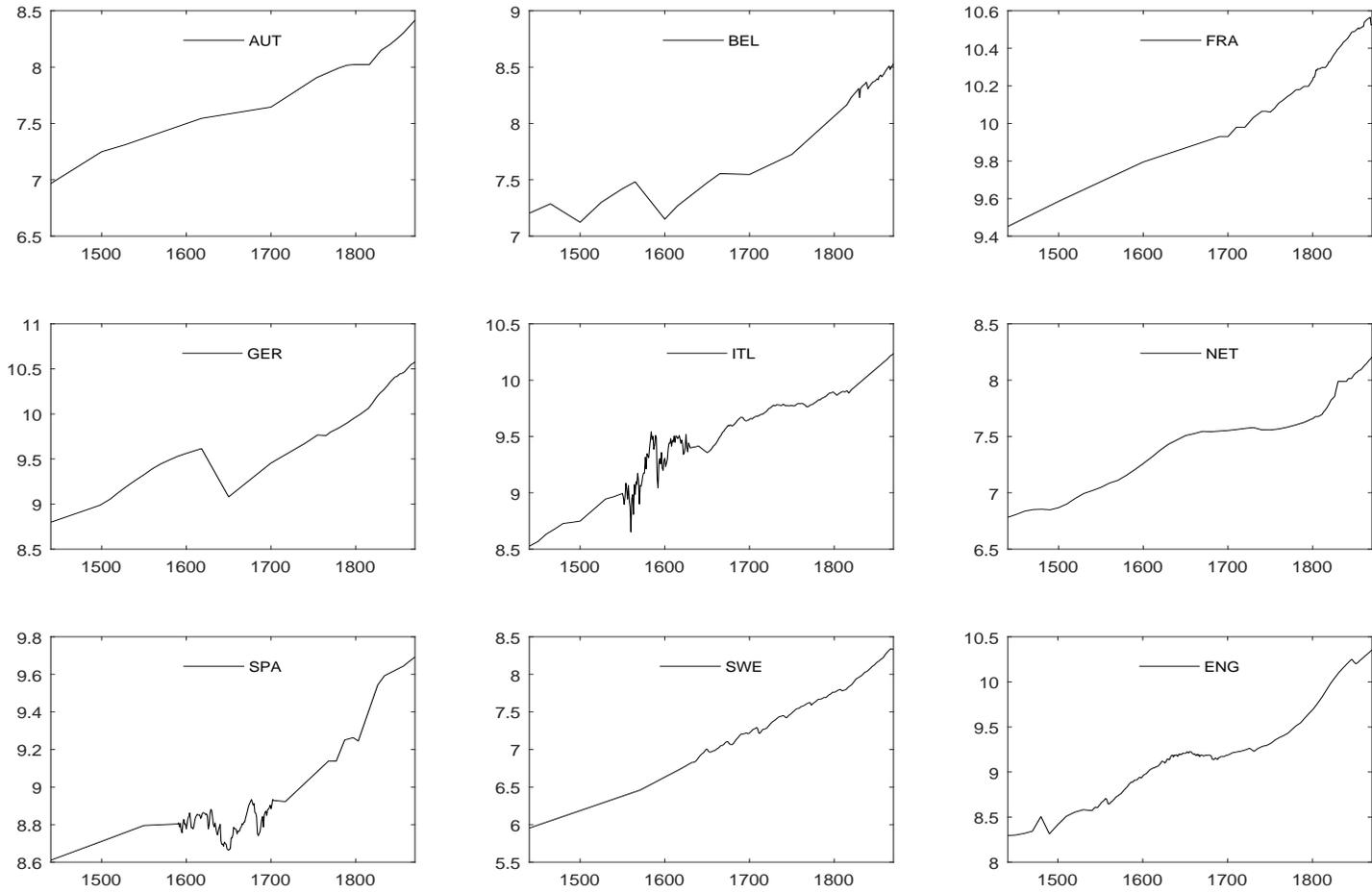


Figure 2: Log Population for 9 OECD Countries (1440-1870)

## 6 Estimation Results

### 6.1 Wage and Income Convergence

The results from estimating equation (11) for real wages and per capita incomes are reported in Tables 1 and 2, respectively. Column (1) in Table 1 reports the results from the OLS regression for the benchmark sample period 1470-1870 for the 9 countries, where the country sample used in various regressions is listed in the notes to the tables. The coefficient of the initial real wage,  $\hat{\beta}$ , is statistically significant, but economically insignificant with a value of -0.003. The half-life is given by  $h = -\ln(2)/(\ln(1 + \hat{\beta} T)/T)$ .

Given  $\hat{\beta} = 0.003$  and  $T = 30$ , this gives  $h = 220$ . With a half-life of 220 years, it is clear that technology and demographic shocks will have a very persistent effect on wages. This finding of very slow convergence is, therefore, consistent with the existing literature that also finds significant but small Malthusian effects before the industrial revolution (Bailey and Chambers, 1993; Lee and Anderson, 2002; Nicolini, 2007; Crafts and Mills, 2009; Voutilainen, 2015; Lagerlöf, 2015).

Nevertheless, these OLS estimates are likely to be biased towards zero due to unobserved cross-country heterogeneity, endogeneity, and the omission of variables that change at the same rate across countries. To address these issues, we include country and time fixed effects in the regressions and use the SUR estimator to cater for cross-country residual correlation, heteroscedasticity and serial correlation. These results are reported in columns (2)-(4) for the benchmark sample period, 1470-1870, and the two sub-periods, 1470-1600 and 1600-1870, where, as discussed above, the year 1600 signifies the approximate date from which there is a substantial improvement in the quality of the data.

The results of using the SUR estimator and including fixed effects dummies, as shown in column (2), yield a coefficient of the initial real wage of  $\hat{\beta} = -0.029$ , implying a half-life of just 10 years. The coefficients of the initial wage are quite similar in the pre- and post-1600 estimates in columns (3) and (4):  $\hat{\beta} = -0.038$  (1470-1600) and  $\hat{\beta} = -0.031$  (1600-1870). In columns (5)-(7), we extend the observation interval to 40 years to allow for the possibility of even slower adjustment of wages towards the steady state than in the baseline regressions. The coefficient,  $\hat{\beta}$ , in column (5) is only slightly reduced to approximately  $-0.019$ , remains highly significant and has approximately the same magnitude in the pre- and post-1600 regressions in columns (6) and (7). The lower absolute value of  $\hat{\beta}$  in the 40-year estimates, compared to the 30-year interval estimates, is unsurprising since the signal-to-noise ratio is likely to decrease with the length of

the observation interval as low frequency data tend to smooth out medium-term cycles. Conversely, decreasing the observation interval to 20 years, as shown in columns (8)-(10), increases the absolute value of the estimated  $\hat{\beta}$  to the range of  $-0.049$  to  $-0.052$ , depending on the sample time period. Thus, overall, the estimates of  $\hat{\beta}$  using real wage data provides evidence of strong Malthusian effects in the four centuries leading up to the industrial revolution.

Table 2 presents estimates of equation (11) using per capita incomes, rather than wages, for which we have a relatively larger sample of countries and longer time periods. To facilitate comparison with the regressions in Table 1, we first consider results for the benchmark sample of 9 countries over the period 1470-1870. Using the OLS estimator, the convergence coefficient becomes  $\hat{\beta} = -0.002$ , which, again, implies very slow convergence. Including fixed effects dummies and allowing for cross-country residual correlation yields much larger values of  $\hat{\beta}$ . For the period, 1470-1870,  $\hat{\beta} = -0.017$  and remains highly significant in the pre- and post-1600 sub-periods with  $\hat{\beta} = -0.024$  (1470-1600) and  $\hat{\beta} = -0.020$  (1600-1870).

Extending the sample to 17 countries yields a coefficient of initial income of  $\hat{\beta} = -0.020$  (column 5); a result which is quite similar to that of the 9-country sample. Columns (6) and (7) report the results for a sample of 11 countries for which long run per capita income data are available. Column (6) shows the results for the period 1470-1870, and in column (7) we expand the time period to 930-1870. The coefficient estimates remain large, highly significant and relatively consistent across the extended period of nearly a millennium, with,  $\hat{\beta} = -0.018$  (1470-1870) and  $\hat{\beta} = -0.015$  (930-1870). As the data are often only available with long time intervals, and are based on small samples, except for a few census counts (Russell, 1958, 1969), the data quality deteriorates quickly before 1470, suggesting that the results in column (7) are somewhat speculative. The results, nevertheless, point towards a Malthusian mechanism that was operative throughout the second millennium until the 19<sup>th</sup> century. The per capita income results for the 20 and 40-year observation intervals, which are provided in the online Appendix, Table A.1, yield estimates of  $\hat{\beta}$  that are statistically highly significant and with absolute magnitudes ranging from  $-0.022$  to  $-0.017$ , when the interval is widened to 40 years, and  $-0.033$  to  $-0.016$  when the observation interval is narrowed down to 20 years.

Table 1: Wage Convergence (Eq.11)

<i>Dep. Var.</i> $= \Delta \ln w_{it}$	(1) OLS	(2) SUR	(3) SUR	(4) SUR	(5) SUR	(6) SUR	(7) SUR	(8) SUR	(9) SUR	(10) SUR
$epi_{it}$	-20.881 <sup>***</sup> (5.183)	-21.121 <sup>***</sup> (4.475)	27.942 <sup>***</sup> (5.422)	-24.861 <sup>***</sup> (5.862)	0.185 (4.467)	40.191 <sup>***</sup> (5.044)	-6.151 (5.636)	-12.156 <sup>**</sup> (4.817)	36.593 <sup>***</sup> (6.013)	-18.204 <sup>***</sup> (6.343)
$sci_{it}$	-0.873 (1.585)	12.513 <sup>***</sup> (1.699)	0.879 (2.009)	32.128 <sup>***</sup> (3.004)	10.900 <sup>***</sup> (1.751)	2.191 (1.986)	20.812 <sup>***</sup> (3.065)	11.367 <sup>***</sup> (1.852)	-1.045 (2.265)	31.988 <sup>***</sup> (3.394)
$\ln w_{it-T}$	-0.003 <sup>***</sup> (0.0002)	-0.029 <sup>***</sup> (0.0005)	-0.038 <sup>***</sup> (0.0006)	-0.031 <sup>***</sup> (0.0006)	-0.019 <sup>***</sup> (0.0003)	-0.024 <sup>***</sup> (0.0008)	-0.021 <sup>***</sup> (0.0004)	-0.051 <sup>***</sup> (0.0007)	-0.049 <sup>***</sup> (0.0014)	-0.052 <sup>***</sup> (0.0009)
TD	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
CD	N	Y	Y	Y	Y	Y	Y	Y	Y	Y
Est. interval	30-yr	30-yr	30-yr	30-yr	40-yr	40-yr	40-yr	20-yr	20-yr	20-yr
No. of country	9	9	9	9	9	9	9	9	9	9
No. of obs.	3609	3609	1179	2439	3519	1089	2439	3699	1269	2439
Est. period	1470-1870	1470-1870	1470-1600	1600-1870	1480-1870	1480-1600	1600-1870	1460-1870	1460-1600	1600-1870

**Notes:** The panel is estimated in various overlapping intervals (30-years as the default, and 20 and 40-years as alternatives) over various periods as indicated. The 9-country sample consists of: Austria, Belgium, France, Germany, Italy, the Netherlands, Spain, Sweden, and the UK. Country (CD) and time (TD) fixed effects are included. OLS estimates account for cross-section heteroscedasticity and serial correlation. SUR estimates account for cross-section heteroscedasticity, serial correlation, and cross-section correlations. Figures in parentheses are robust standard errors. Note that for control variables, both the estimated coefficients and the standard errors, in parentheses, are multiplied by a scale factor of 100. Significance at \*10%, \*\*5%, \*\*\*1% levels.

Table 2: Per Capita Income Convergence (Eq.11)

<i>Dep. Var.</i> $= \Delta \ln y_{it}$	(1) OLS	(2) SUR	(3) SUR	(4) SUR	(5) SUR	(6) SUR	(7) SUR
$epi_{it}$	5.697*** (2.137)	0.569 (0.824)	-0.608*** (0.193)	4.281*** (1.432)	-0.013 (0.068)	0.025 (0.09)	-0.076 (1.171)
$sci_{it}$	3.896*** (0.78)	0.136 (0.276)	0.619*** (0.09)	0.897 (0.828)	0.194*** (0.05)	0.192*** (0.064)	-0.137 (0.09)
$\ln y_{it-30}$	-0.002*** (0.0002)	-0.017*** (0.0005)	-0.024*** (0.0003)	-0.020*** (0.0007)	-0.020*** (0.0004)	-0.018*** (0.0004)	-0.015*** (0.0002)
TD	N	Y	Y	Y	Y	Y	Y
CD	N	Y	Y	Y	Y	Y	Y
Est. interval	30-yr						
No. of country	9	9	9	9	17	11	11
No. of obs	3609	3609	1179	2439	6817	4411	10351
Est. period	1470-1870	1470-1870	1470-1600	1600-1870	1470-1870	1470-1870	930-1870

**Notes:** The 9-country sample consists of Austria, Belgium, France, Germany, Italy, the Netherlands, Spain, Sweden, and the UK. The 17-country sample consists of Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the UK. The 11-country sample consists of Belgium, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, Portugal, Spain, and the UK. Figures in parentheses are robust standard errors. Note that for control variables, both the estimated coefficients and the standard errors, in parentheses, are multiplied by a scale factor of 100. Significance at \*10%, \*\*5%, \*\*\*1% levels.

Overall, the wage and income regressions provide evidence of strong Malthusian effects over many centuries for up to as many as 17 countries. Specifically, the estimated coefficients of convergence,  $\hat{\beta}$ , range from  $-5\%$  to  $-1.6\%$  per year, with most results ranging between  $-2\%$  and  $-3\%$  per year, implying half-lives of between 10 and 30 years. With these convergence coefficients, technological and demographic shocks would have significantly dissipated over a few decades.

Turning to the control variables, the coefficients of scientists,  $\Delta \ln(sci)_{it}$ , are positive and significant in 10 of the 17 cases in Tables 1 and 2 and in no cases are they significantly negative. Thus, there is a tendency for the knowledge stock to increase per capita income; however, the Malthusian feedback mechanism ensures that most of the positive growth effects are temporary. The income effect of epidemics is mixed, which may reflect the fact that epidemics simultaneously affect productivity as well as overall population levels. For example, following an epidemic outbreak, the reduction in the population growth drag gives a temporary positive effect on wages and incomes. Conversely the health status of the population surviving the epidemics may reduce productivity and also may affect dependency ratios. Thus the sign of the net effect of these changes is likely to vary over time and across countries.

## 6.2 Population Convergence

The distinguishing feature of the Malthusian growth model is that population growth is endogenous. Hence the Malthusian model yields a convergence equation for population as well as income, as given by equation (12). Specifically, the convergence parameter,  $\hat{\gamma}$ , should be significantly negative and not too small in absolute value for the Malthusian mechanism to be effective.

Table 3 reports the results of estimating equation (12) for the 9-country sample over the baseline period 1470-1870. The OLS estimate in column (1) shows that the estimated coefficient of the initial population has a very small value of only  $\hat{\gamma} = -0.002$ . This implies a half-life of  $h = 336$  years, where  $h = -\ln(2) / (\ln(1 + \hat{\gamma}T)/T)$ . Again, the coefficient of initial population is likely to be biased towards zero due to the correlation between initial population, omitted fixed effects, and the presence of common cross-country shocks. We therefore include country and time fixed effects in the models and use the SUR estimator, and find evidence of strong population convergence. Specifically, we find  $\hat{\gamma} = -0.028$  over the full period (column 2), which implies much faster convergence, with a half-life of 11 years. Likewise, we have similarly large coefficients over the subperiods with  $\hat{\gamma} = -0.034$  over 1470-1600 (column 3) and  $\hat{\gamma} = -0.032$  over the period 1600-1870 (column 4).

Table 3: Population Convergence (Eq.12)

<i>Dep. Var.</i> $= \Delta \ln L_{it}$	(1) OLS	(2) SUR	(3) SUR	(4) SUR	(5) SUR	(6) SUR	(7) SUR
$epi_{it}$	-7.165*** (1.040)	-0.325 (0.325)	-0.493** (0.211)	-0.188 (0.373)	-0.229*** (0.077)	-0.161 (0.133)	0.687 (0.512)
$\ln land_{it-T}$	-0.061*** (0.008)	1.526*** (0.046)	0.391*** (0.018)	2.335*** (0.082)	1.489*** (0.042)	2.390*** (0.036)	1.782*** (0.055)
$\ln S_{it-T}^{sci}$	0.116*** (0.005)	-0.006 (0.011)	0.015** (0.006)	0.027 (0.018)	0.003 (0.004)	0.012** (0.005)	0.151*** (0.02)
$sci_{it}$	9.184*** (0.334)	0.082 (0.259)	0.453*** (0.156)	0.469 (0.427)	0.037 (0.085)	0.203* (0.122)	2.556*** (0.51)
$\ln L_{it-T}$	-0.002*** (0.0001)	-0.028*** (0.0006)	-0.034*** (0.0004)	-0.032*** (0.0006)	-0.029*** (0.0003)	-0.029*** (0.0004)	-0.021*** (0.0006)
TD	N	Y	Y	Y	Y	Y	Y
CD	N	Y	Y	Y	Y	Y	Y
Est. interval	30-yr						
No. of country	3609	3609	1179	2439	15997	7528	2823
No. of obs	9	9	9	9	17	8	3
Est. period	1470-1870	1470-1870	1470-1600	1600-1870	930-1870	930-1870	930-1870

*Notes:* The 9-country sample consists of: Austria, Belgium, France, Germany, Italy, the Netherlands, Spain, Sweden, and the UK. The 17-country sample consists of: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the UK. The 8-country sample consists of: Japan, Belgium, France, Germany, Italy, the Netherlands, Sweden and the UK. The 3-country sample consists of: Italy, the Netherlands and the UK. Figures in parentheses are robust standard errors. Note that for control variables, both the estimated coefficients and the standard errors, in parentheses, are multiplied by a scale factor of 100. Significance at \*10%, \*\*5%, \*\*\*1% levels.

Expanding the sample to 17 countries (column 5) yields  $\hat{\gamma} = -0.029$ , which is consistent with the 9-country sample estimates. Likewise, when we expand the sample period to 930-1870 for 8 countries (column 6) the coefficient remains at  $\hat{\gamma} = -0.029$ . Restricting the sample to the countries with the best data coverage and the most reliable historical population data, namely Italy, the Netherlands and the UK, back to 930, yields  $\hat{\gamma} = -0.021$ , which is, again, consistent with the Malthusian stagnation hypothesis. Thus, the results are robust to variations in estimation periods and country sample.

The estimates of the population convergence for 40 and 20-year observation intervals are given in the online Appendix, Table A.2. For the 40-year interval estimates we find  $\hat{\gamma} = -0.024$  for the full period, 1480-1870, and  $\hat{\gamma} = -0.026$  and  $\hat{\gamma} = -0.023$  over the periods 1480-1600 and 1600-1870 (columns 1-3). The coefficients of the initial population range from -0.042 to -0.024 in the 20-year interval estimates. Thus, the results are robust to the choice of 20, 30, and 40-year time-intervals. Turning to the control variables in Table 3, except for the OLS estimates the coefficients of land size are all positive and significant, which is consistent with our model. Also consistent with the Malthusian hypothesis, the coefficients of the knowledge stock are positive in 8 of the 14 cases– the remainder are insignificant. Finally, as with the previous results for wages, the coefficients of epidemics are either significantly negative, as might be expected, or insignificant.

### 6.3 The Post-Malthusian Era

Thus, our results reveal a clear pattern of population and wage convergence. Specifically, we find coefficients on initial income, wages and population that are typically in the range of 2% - 3% per year, and which imply half-lives in the range of 10-30 years. These suggest that any increase or fall in wages due to a shock, would have quickly been driven back to the stationary steady-state level, consistent with the notion of an “iron-law of wages”.

A key debate in the literature refers to whether the Malthusian mechanism began to weaken during or before the industrial revolution. According to unified growth theory, the post-Malthusian period is defined as the period where the Malthusian mechanism still operates but there is also accelerating population growth and a positive trend in wages and per capita incomes. Specifically, Galor and Weil (2000) show that population growth rates across Europe approximately doubled in the 18<sup>th</sup> century relative to the 17<sup>th</sup> century and the population growth rate reaches a peak in the 19<sup>th</sup> century, 1820-1870, when there is also a significant rise in per capita GDP growth (Galor and Weil, 2000; Galor, 2011).

Our results show that the Malthusian convergence dynamics are relatively constant throughout the sample periods from the 10<sup>th</sup> century up until the 19<sup>th</sup> century. For example, in the wage and per capita income regressions, the coefficient on the lagged dependent variable in the 15<sup>th</sup> to 17<sup>th</sup> century sample barely changes from the 17<sup>th</sup> to 19<sup>th</sup> century sample. For the population results, the coefficient for the latter period, 1470-1870 is, if anything, slightly larger than the coefficient estimate for the preceding period 1470-1600.

This consistency contrasts with Nicolini (2007) who finds insignificant Malthusian responses after 1740; Crafts and Mills (2009) who find economically insignificant effects from 1650 onwards. The conclusion that Malthusian demographic responses remained large across the millennium supports the unified growth theory perspective that a post-Malthusian regime emerged toward the end of the sample period in the 19<sup>th</sup> century, where Malthusian demographic effects were still strong but rapid technological progress nevertheless resulted in rising wages. Thus, our results suggest that across countries, and up until the industrial revolution, there were strong Malthusian responses. These strong responses meant that there was rapid convergence back to the steady-state wage following a shock, so that wage gains from technological progress, or land expansion, would have been substantially eroded by population growth in the course of a generation. We find no evidence of the Malthusian mechanism breaking down in the latter centuries. Rather, our results support the idea of a post-Malthusian era where technological progress accelerated and populations boomed.

## 6.4 Country Sub-Samples

So far all of our regressions have been based on panel data in order to gain efficiency through a larger sample size and by allowing for the correlation between the residuals across countries. Furthermore, the inclusion of time dummies in panels reduces the bias from omitted variables that are correlated with the regressors to the extent that the omitted variables have a high covariation across countries. The downside of pooling is that the coefficients may be biased if they vary significantly across countries. To investigate this possibility, we regress equations (11) and (12) for each individual country and for three country groups distinguished by their religion. As argued above, the productivity effects of the reduction in holidays during the Protestant Reformation would be picked up by the time-dummies to the extent that the change occurred across all countries and at the same time. It is, therefore, natural to split the countries up in three groups: Catholic countries, Protestant countries and Other countries that have mixed Catholic, Protes-

tant or other religions, including Japan, which is predominantly Buddhist. Furthermore, as discussed by Lynch (2006), there is a large literature advocating different fertility patterns across religions; particularly contrasting Catholicism and Protestantism. If the fertility responses differ across religions, we would expect a differentiated response to shocks and a different speed of adjustment towards the steady state.

First consider the regressions for individual countries in Table 4 over the periods 1470-1870, 1470-1600 and 1600-1870, where only the parameter estimates of the convergence coefficients are reported for brevity. For wages and income, the convergence coefficients are almost all highly significantly negative across countries and across time-periods and the size of the coefficients are quite close to those obtained in our panel regressions. In the population growth equations, the convergence coefficients are significantly negative in 8 of the 9 cases in the full-period estimates and in 15 of the 18 cases in the sub-period regressions. While there are a few instances where the convergence coefficient is significantly positive in the population growth model, this is likely to be an outcome of small-sample bias, measurement errors, and other omitted fixed effects that we can control for in panel data, such as technology that is transmitted across countries and pandemics.

Turning to the estimates for country groups with different religious beliefs in Table 5, the convergence coefficients are generally very close to those of the baseline regression for wages, income and population, regardless of religious adherence. These results reinforce the results in Table 4 that the convergence coefficients are quite similar across countries regardless of religious adherence.

Table 4: Convergence by Country

	Wage			Income			Population		
	1470-1870	1470-1600	1600-1870	1470-1870	1470-1600	1600-1870	1470-1870	1470-1600	1600-1870
Austria	-0.027 <sup>***</sup> (0.0017)	-0.014 <sup>****</sup> (0.0031)	-0.031 <sup>***</sup> (0.0021)	-0.010 <sup>***</sup> (0.0012)	-0.023 <sup>***</sup> (0.0020)	-0.001 (0.0010)	-0.001 (0.0021)	-0.014 <sup>***</sup> (0.0011)	-0.004 (0.0033)
Belgium	-0.029 <sup>***</sup> (0.0017)	-0.028 <sup>***</sup> (0.0033)	-0.031 <sup>***</sup> (0.0019)	0.006 (0.0044)	-0.031 <sup>***</sup> (0.0019)	0.060 <sup>***</sup> (0.0071)	-0.006 <sup>***</sup> (0.0012)	-0.051 <sup>***</sup> (0.0052)	-0.027 <sup>***</sup> (0.0014)
France	-0.030 <sup>***</sup> (0.0016)	-0.025 <sup>***</sup> (0.0033)	-0.035 <sup>***</sup> (0.0019)	-0.006 <sup>***</sup> (0.0022)	-0.031 <sup>***</sup> (0.0021)	-0.008 <sup>***</sup> (0.0027)	-0.010 <sup>***</sup> (0.0018)	-0.001 <sup>***</sup> (0.0001)	-0.014 <sup>***</sup> (0.0025)
Germany	-0.031 <sup>***</sup> (0.0017)	-0.029 <sup>***</sup> (0.0030)	-0.033 <sup>***</sup> (0.0021)	0.001 (0.0028)	-0.039 <sup>***</sup> (0.0035)	0.006 <sup>*</sup> (0.0030)	-0.029 <sup>***</sup> (0.0019)	0.043 <sup>***</sup> (0.0048)	-0.013 <sup>***</sup> (0.0016)
Italy	-0.030 <sup>***</sup> (0.0016)	-0.033 <sup>***</sup> (0.0027)	-0.032 <sup>***</sup> (0.0021)	-0.036 <sup>**</sup> (0.0015)	-0.032 <sup>***</sup> (0.0025)	-0.037 <sup>***</sup> (0.0020)	-0.037 <sup>***</sup> (0.0017)	-0.045 <sup>***</sup> (0.0042)	-0.032 <sup>***</sup> (0.0010)
Netherlands	-0.029 <sup>***</sup> (0.0016)	-0.032 <sup>***</sup> (0.0031)	-0.033 <sup>***</sup> (0.0019)	-0.020 <sup>***</sup> (0.0015)	-0.020 <sup>***</sup> (0.0033)	-0.032 <sup>***</sup> (0.0018)	-0.005 <sup>***</sup> (0.0016)	-0.067 <sup>***</sup> (0.0038)	-0.017 <sup>***</sup> (0.0031)
Spain	-0.031 <sup>***</sup> (0.0016)	-0.026 <sup>***</sup> (0.0029)	-0.034 <sup>***</sup> (0.0020)	-0.018 <sup>***</sup> (0.0025)	-0.000 (0.0006)	-0.019 <sup>***</sup> (0.0020)	-0.024 <sup>***</sup> (0.0017)	-0.004 (0.0053)	-0.032 <sup>***</sup> (0.0020)
Sweden	-0.034 <sup>***</sup> (0.0016)	-0.037 <sup>***</sup> (0.0030)	-0.034 <sup>***</sup> (0.0019)	-0.028 <sup>***</sup> (0.0018)	-0.015 <sup>***</sup> (0.0017)	-0.031 <sup>***</sup> (0.0023)	-0.010 <sup>***</sup> (0.0013)	-0.001 (0.0011)	-0.020 <sup>***</sup> (0.0020)
UK	-0.017 <sup>***</sup> (0.0016)	-0.027 <sup>***</sup> (0.0037)	-0.016 <sup>***</sup> (0.0018)	-0.001 (0.0012)	-0.017 <sup>***</sup> (0.0015)	-0.003 <sup>*</sup> (0.0015)	-0.027 <sup>***</sup> (0.0019)	-0.032 <sup>***</sup> (0.0031)	-0.021 <sup>***</sup> (0.0024)

*Notes:* Figures in parentheses are robust standard errors. Significance at \*10%, \*\*5%, \*\*\*1% levels.

Table 5: Convergence by Country Group (Eqs. 11-12)

	Baseline	Catholic	Protestant	Mix
Wage				
1470-1870	-0.030 <sup>***</sup> (0.0005)	-0.030 <sup>***</sup> (0.0008)	-0.027 <sup>***</sup> (0.0016)	-0.022 <sup>***</sup> (0.0014)
1470-1600	-0.037 <sup>***</sup> (0.0007)	-0.028 <sup>***</sup> (0.001)	-0.054 <sup>***</sup> (0.002)	-0.039 <sup>***</sup> (0.0012)
1600-1870	-0.031 <sup>***</sup> (0.0006)	-0.030 <sup>***</sup> (0.001)	-0.027 <sup>***</sup> (0.0017)	-0.024 <sup>***</sup> (0.0016)
Income				
1470-1870	-0.017 <sup>***</sup> (0.0005)	-0.024 <sup>***</sup> (0.0001)	-0.019 <sup>***</sup> (0.0008)	-0.034 <sup>***</sup> (0.001)
1470-1600	-0.024 <sup>***</sup> (0.0003)	-0.027 <sup>***</sup> (0.0001)	-0.007 <sup>**</sup> (0.0005)	-0.019 <sup>***</sup> (0.0012)
1600-1870	-0.020 <sup>***</sup> (0.0007)	-0.024 <sup>***</sup> (0.0002)	-0.018 <sup>***</sup> (0.001)	-0.041 <sup>***</sup> (0.0011)
Population				
1470-1870	-0.028 <sup>***</sup> (0.0006)	-0.029 <sup>***</sup> (0.0007)	-0.023 <sup>***</sup> (0.0007)	-0.028 <sup>***</sup> (0.0009)
1470-1600	-0.017 <sup>***</sup> (0.0006)	-0.031 <sup>***</sup> (0.0006)	-0.018 <sup>***</sup> (0.0013)	-0.032 <sup>***</sup> (0.0013)
1600-1870	-0.030 <sup>***</sup> (0.0007)	-0.031 <sup>***</sup> (0.0006)	-0.027 <sup>***</sup> (0.0009)	-0.025 <sup>***</sup> (0.001)

*Notes:* The SUR estimator is used and the interval  $T$  is 30 years. The Catholic group consists of Austria, Belgium, France, Greece, Ireland, Italy, Portugal and Spain in the population equation, and Austria, Belgium, France, Italy and Spain in the wage equation. The Protestant group consists of Denmark, Finland, Norway, Sweden, the UK in the population equation, and Sweden and the UK in the wage equation. Finally, the Mix group consists of Japan, Germany, Netherlands, Switzerland (population equation), and Germany, Netherlands (wage equation). Figures in parentheses are robust standard errors. Significance at \*10%, \*\*5%, \*\*\*1% levels.

## 6.5 Comparing The Speed of Convergence Across Studies

Our finding of strong Malthusian effects contrasts with much of the literature. Specifically, the large coefficients we find imply a very strong Malthusian trap, so that wages would have quickly reverted to the steady-state level following a population or technology shock. To see how our estimates differ in their implications for the speed of convergence from existing studies, Table 6 reports the coefficients from several other prominent studies and the implied half-lives. We compare these with those implied by the present study's estimates of  $\hat{\beta}$  and  $\hat{\gamma}$ , reported in the last three columns. We recover the implied convergence rate from existing studies from the reported values of the partial elasticity of population growth with respect to log wages, which corresponds to an annualized value of  $\sigma$ , and the estimated non-labour share,  $\alpha$ . When an estimate of  $\alpha$  is not available, we assume  $\alpha = 0.4$  in order to construct an indicative estimate the convergence rate based on the reported partial elasticity of population growth with respect to log wages only.

It can be seen that our estimated coefficients imply substantially faster convergence than those in the existing literature with half-lives of 1-3 decades as opposed to half-lives that are typically in the range of 1-2 centuries, or in some cases much longer.<sup>22</sup> A possible exception is found in Kelly and Ó Gráda (2012) for the sub period 1700-1750. Assuming that  $\alpha = 0.4$ , their coefficient estimates imply a half-life of 100 years. In their other sub-periods, however, it can be seen that the implied half-lives exceed a century under the same assumption.

---

<sup>22</sup> The only cases in the literature where our estimated speeds of convergence coincides with that of the literature are in Crafts and Mills (2009) for the sub period 1540-1645. As they acknowledge, however, this particular estimate depends on an estimated factor income share that exceeds unity.

Table 6: Comparison of Results with Existing Studies

Authors	Model	Period	$\partial n / \partial \ln w$	Non-labour Share $\alpha$	Implied Convergence Rate	Coefficient $-\beta$ or $-\gamma$	half-life years
(1)	(2)	(3)	(4)	(5)	(6)=(4)*(5)	(7)	(8)
Bailey and Chambers (1993)	Cointegration	1539-1800	0.0084	0.400	0.003	-	206
Lee and Anderson (2002)	Structural Model	1548-1834	0.0062	1.045	0.006	-	107
Nicolini (2007)	Stationary VAR	1541-1840	0.0005	0.400	0.000	-	3466
Crafts and Mills (2009)	Structural Model	1540-1870	0.0063	0.920	0.006	-	120
Crafts and Mills (2009)	Stationary VAR	1540-1740	0.0050	0.400	0.002	-	347
Møller and Sharp (2014)	CVAR	1564-1802	0.0073	0.400	0.003	-	237
Kelly and Ó Gráda (2012)	Regression	1539-1600	0.0112	0.400	0.004	-	155
Kelly and Ó Gráda (2012)	Regression	1600-1650	0.0088	0.400	0.004	-	197
Kelly and Ó Gráda (2012)	Regression	1650-1700	0.0109	0.400	0.004	-	159
Kelly and Ó Gráda (2012)	Regression	1700-1750	0.0173	0.400	0.007	-	100
Kelly and Ó Gráda (2012)	Regression	1750-1800	0.0055	0.400	0.002	-	315
Present Study	Wage Regression	1470-1870	-	-	-	0.029	11
Present Study	Income Regression	1470-1870	-	-	-	0.017	29
Present Study	Population Regression	1470-1870	-	-	-	0.028	12

**Notes:** Notes: The implied convergence rates from existing studies are estimated from reported coefficients  $\partial n / \partial \ln w$  in population regressions, e.g. the response of population growth rates to relative changes in real wages, and assuming  $\alpha = 0.4$ , except for Lee and Anderson (2002) and Crafts and Mills (2009) where estimates of  $\alpha$  are provided. The implied half-life is calculated by dividing  $\ln(2)$  by column (6). For the present study the formula used is described in the text.

## 7 Conclusion

The industrial revolution represents the turning point between economic stagnation and growth. Understanding what forces preserved economic stagnation in the pre-industrial era is a quintessential issue in Economics. Many economists place the Malthusian trap at the centre of their explanations (Hansen and Prescott, 2002; Clark, 2007; Galor, 2011). Nevertheless most of the existing econometric literature finds that Malthusian responses were very small. This raises doubts as to whether Malthusian demographic responses were strong enough to keep wages and incomes constant at the “iron-law” wage level in the face of land cultivation and other productivity improvements or technology shocks. Instead, it may simply be that economic stagnation was a consequence of low productivity growth, and that Malthusian demographics did not play an important role.

In this paper we provide new evidence that strongly supports the Malthusian view that economic stagnation was the result of the Malthusian trap, and that overcoming the Malthusian trap is, therefore, a key ingredient in understanding the sources of modern growth and underdevelopment. Our model and results advance the existing literature in several ways. First, we derive a dynamic representation of the Malthusian model based on a model of household fertility choice. The model shows that the relevance of the Malthusian mechanisms in generating stagnation is an empirical issue that depends on the sign and size of total income elasticity of demand for children,  $\sigma$ .

Second, we derive a parsimonious empirical strategy for testing the Malthusian trap hypothesis by linearizing around the steady state to obtain a familiar  $\beta$  – convergence or error correction model. The key parameters of interest can be directly estimated from the initial values of income, wages and population in the respective regression equations. Moreover, the estimation strategy overcomes identification issues that have been discussed in the literature and that may be responsible for inconsistent parameter estimates and very small convergence coefficients in previous studies. We test the Malthusian model using panel data with up to 17 OECD countries from 900CE to 1870CE, thus using a far larger data set than has been used before.

In contrast to the existing econometric evidence, which shows a surprising lack of support for the Malthusian trap, we find very strong evidence of Malthusian dynamics across the entire sample. Our results thus provide the first time-series econometric evidence for a Malthusian trap across a range of countries. We find very large negative convergence coefficients, which, as we show, are necessary for the existence of a Malthusian trap that can account for the observed stagnation in wages and incomes. Specifically, we find a

convergence coefficient of around 2-3 percent per year, implying half-lives of 10-30 years. This contrasts with the existing literature which typically finds near non-stationarity of wages and extremely weak convergence with half-lives that exceed several centuries.

We find that the evidence of strong Malthusian effects is consistent across a period of almost thousand years leading up to the 19<sup>th</sup> century industrial revolution. The results therefore also support the idea that the Malthusian era lasted until the 19<sup>th</sup> century, when accelerating technological change led to rapid population growth. This transition from a Malthusian to a post-Malthusian era is also a key component of the unified growth theory literature (Galor and Weil, 2000; Galor, 2011). Overall, therefore, the results suggest that fertility and demographic factors were an important source of past economic stagnation and, therefore, may also remain an important source of contemporary underdevelopment.

## References

- Acemoglu, D., Johnson, S. and Robinson, J. (2005), "Institutions as the fundamental cause of long-run growth," in Aghion, P. and Durlauf, S.N. (eds.), *Handbook of Economic Growth*, Vol 1A, pp.385-472, Amsterdam and San Diego: Elsevier, North-Holland.
- Allen, R.C. (2001), "The great divergence in European wages and prices from the Middle Ages to the First World War," *Explorations in Economic History*, 38(4), pp.411-447.
- Allen, R.C. (2008), "A review of Gregory Clark's A Farewell to Alms: A brief economic history of the world," *Journal of Economic Literature*, 46(4), pp.946-973.
- Alvarez-Nogal, C. and De La Escosura, L.P. (2007), "The decline of Spain (1500–1850): Conjectural estimates," *European Review of Economic History*, 11(3), pp.319-366.
- Ashraf, Q. and Galor, O. (2011), "Dynamics and stagnation in the Malthusian epoch," *The American Economic Review*, 101(5), pp.2003-2041.
- Bailey, R.E. and Chambers, M.J. (1993), "Long-term demographic interactions in precensus England," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 156(3), pp.339-362.
- Barro, R.J. and Sala-i-Martin, X. (1992), "Convergence," *Journal of Political Economy*, 100(2), pp.223-251.
- Baumol, William J. (1983) "Marx and the iron law of wages" *The American Economic Review*, 73(2), pp.303-308.
- Broadberry, S. (2013), "Accounting for the Great Divergence," Economic History Working Papers, 184/13, London School of Economics and Political Science.
- Broadberry, S., Campbell, B.M., Klein, A., Overton, M., and Van Leeuwen, B. (2015), *British Economic Growth, 1270-1870*, Cambridge, UK: Cambridge University Press.
- Chaney, E. and Hornbeck, R. (2016), "Economic dynamics in the Malthusian era: Evidence from the 1609 Spanish expulsion of the Moriscos," *The Economic Journal*, 126(594), pp.1404-1440.

- Chiarini, B., 2010. "Was Malthus right? The relationship between population and real wages in Italian history, 1320 to 1870," *Explorations in Economic History*, 47(4), pp.460-475.
- Clark, G. (2005), "The condition of the working class in England, 1209-2004," *Journal of Political Economy*, 113(6), pp.1307-1340.
- Clark, G. (2007), *A Farewell to Alms*, Princeton, NJ: Princeton University Press.
- Crafts, N. and Mills, T.C. (2009), "From Malthus to Solow: How did the Malthusian economy really evolve?" *Journal of Macroeconomics*, 31(1), pp.68-93.
- Edvinsson, R. (2015), "Pre-industrial population and economic growth: Was there a Malthusian mechanism in Sweden?" Stockholm Papers in Economic History No. 17, Department of Economic History, Stockholm University.
- Egger, P. (2001), "SUR estimation of error components models with AR(1) disturbances and unobserved endogenous effects," WIFO Working Papers No. 171.
- Galor, O. (2005), "From stagnation to growth: Unified Growth Theory," in Aghion, P. and Durlauf, S.N (eds.), *Handbook of Economic Growth*, Vol 1A, pp.171-293, North Holland.
- Galor, O. (2011), *Unified Growth Theory*, Princeton, NJ: Princeton University Press.
- Galor, O. (2012), "The demographic transition: Causes and consequences," *Cliometrica*, 6(1), pp.1-28.
- Galor, O. and Moav, O. (2002), "Natural selection and the origin of economic growth," *The Quarterly Journal of Economics*, 117(4), pp.1133-1191.
- Galor, O. and Weil, D.N. (2000), "Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond," *American Economic Review*, 90(4), pp.806-828.
- Greasley, D., Madsen, J.B., and Wohar, M.E. (2013), "Long-run growth empirics and new challenges for unified theory," *Applied Economics*, 45(28), pp.3973-3987.
- Greengrass, M. (2014), *Christendom Destroyed: Europe 1517-1648*, London: Allen Lane.
- Greif, A. (2006), *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*, Cambridge UK: Cambridge University Press.
- Guinnane, T.W. (2011), "The historical fertility transition: A guide for economists," *Journal of Economic Literature*, 49(3), pp.589-614.
- Hajnal, J. (1965), "European marriage patterns in perspective," In Glass, D.V. and Eversley, D.E.C. (eds), *Population in History*, pp.101-143, London, UK: Arnold.
- Hansen, G.D. and Prescott, E.C. (2002), "Malthus to Solow," *The American Economic Review*, 92(4), pp.1205-1217.
- Humphries, J. and Weisdorf, J.L. (2017), "Unreal wages? Real income and economic growth in England, 1260-1850," CEPR Discussion Paper No. DP11999.
- Kelly, M. and Ó Gráda, C. (2012), "The preventive check in medieval and preindustrial England," *The Journal of Economic History*, 72(4), pp.1015-1035.
- Klemp, M. and Møller, N.F. (2016), "Post-Malthusian dynamics in pre-industrial Scandinavia," *The Scandinavian Journal of Economics*, 118(4), pp.841-867.

- Kmenta, J. (1971), *Elements of Econometrics*, New York: Macmillan.
- Kögel, T. and Prskawetz, A. (2001), “Agricultural productivity growth and escape from the Malthusian trap,” *Journal of Economic Growth*, 6(4), pp.337-357.
- Kremer, M. (1993), “Population growth and technological change: One million B.C. to 1990,” *Quarterly Journal of Economics*, 108(3), pp.681-716.
- Lagerlöf, N.P. (2015) “Malthus in Sweden,” *The Scandinavian Journal of Economics*, 117(4), pp.1091-1133.
- Landes, D.S. (2003), *The Unbound Prometheus: Technological Change and Industrial Development in Western Europe from 1750 to the Present*, Cambridge, UK: Cambridge University Press.
- Lee, R. (1973), “Population in preindustrial England: An econometric analysis,” *The Quarterly Journal of Economics*, 87(4), pp.581-607.
- Lee, R. (1980), “A Historical Perspective on Economic Aspects of the Population Explosion: The Case of Preindustrial England, in Easterlin RA (ed), *Population and Economic Change in Developing Countries*, pp.517–566, University of Chicago Press, Chicago.
- Lee, R. (1985) “Inverse projection and back projection: A critical appraisal and comparative results for England, 1539-1871,” *Population Studies*, 39, pp.233-248.
- Lee, R. (2003), “The demographic transition: Three centuries of fundamental change,” *Journal of Economic Perspectives*, 17 (4), pp.167-190.
- Lee, R. (2015), “Becker and the demographic transition,” *Journal of Demographic Economics*, 81(1), pp.67-74.
- Lee, R. and Anderson, M. (2002), “Malthus in state space: Macro economic-demographic relations in English history, 1540 to 1870,” *Journal of Population Economics*, 15(2), pp.195-220.
- Lynch, K. A. (2006), “Theoretical and analytical approaches to religious beliefs, values, and identities during the modern fertility transition,” in Derosas, R. and van Poppel, F. (eds.), *Religion and the Decline of Fertility in the Western World*, pp.21-39, Dordrecht: Springer.
- Maddison, A. (2003), *The World Economy: Historical Statistics*, Paris: OECD.
- Madsen, J.B., Ang, J.B. and Banerjee, R. (2010), “Four centuries of British economic growth: The roles of technology and population,” *Journal of Economic Growth*, 15(4), pp.263-290.
- Malthus, T.R., (1798), *An Essay on the Principle of Population*. New York: Modern Library.
- Mankiw, N.G, Romer, D., and Weil, D.N. (1992), “A contribution to the empirics of economic growth,” *The Quarterly Journal of Economics*, 107(2), pp.407-437.
- McEvedy, C. and Jones, R. (1978), *Atlas of World Population History*, London: Penguin.
- Mokyr, J. (2008), “The institutional origins of the industrial revolution,” in Helpman, E. (ed), *Institutions and Economic Performance*, pp.64-119, Cambridge: Harvard University Press.
- Mokyr, J. (2011), *The Enlightened Economy: Britain and the Industrial Revolution 1700-1850*, London: Penguin Books.
- Møller, N.F. and Sharp, P. (2014), “Malthus in cointegration space: Evidence of a post-Malthusian pre-industrial England,” *Journal of Economic Growth*, 19(1), pp.105-140.

- Murphy, T. (2010), "Persistence of Malthus or persistence in Malthus? Mortality, income, and marriage in the French fertility decline of the long nineteenth century," Working Paper 363, IGER, Bocconi University.
- Nicolini, E.A. (2007), "Was Malthus Right? A VAR analysis of economic and demographic interactions in pre-industrial England," *European Review of Economic History*, 11(1), pp.99-121.
- Persson, K.G. (2008), "The Malthus delusion," *European Review of Economic History*, 12(2), pp.165-173.
- Phelps Brown, H. and Hopkins, S. V. (1981), *A Perspective of Wages and Prices*, London: Methuen.
- Postan, M.M. (1972), *The Medieval Economy and Society: An Economic History of Britain in the Middle Ages*, Berkeley and Los Angeles CA: University of California Press.
- Rosenberg, N. and Birdzell, L.E. Jr. (1987), *How the West Grew Rich: The Economic Transformation of the Industrial World*, New York: Basic Books.
- Russell, J. C. (1958), "Late ancient and medieval population," *Transactions of the American Philosophical Society*, 48(3), pp.1-152.
- Russell, J.C. (1969), "Population in Europe 500-1500", in Cipolla, C. (ed.) *Fontana Economic History of Europe: The Middle Ages*, Vol.1, Sect.1, UK: Collins.
- Strulik, H. and Weisdorf, J. (2008), "Population, food, and knowledge: A simple unified growth theory," *Journal of Economic Growth*, 13(3), pp.169-194.
- Strulik, H. and Weisdorf, J. (2014), "How child costs and survival shaped the industrial revolution and the demographic transition," *Macroeconomic Dynamics*, 18(1), pp.114-144.
- Tabellini, G. (2010), "Culture and institutions: Economic development in the regions of Europe," *Journal of the European Economic Association*, 8(4), pp.677-716.
- Temple, J. (1998). "Robustness tests of the augmented Solow model," *Journal of Applied Econometrics* 13(4), pp.361-375.
- Vollrath, D. (2011), "The agricultural basis of comparative development," *Journal of Economic Growth*, 16(4), pp.343-70.
- Voutilainen, M. (2015), "Malthusian checks in pre-industrial Sweden and Finland: A comparative analysis of the demographic regimes," *Scandinavian Economic History Review*, 63(3), pp.235-259.
- Weir, D. (1984), "Life under pressure: France and England 1670-1870," *Journal of Economic History*, 44(1), pp.27-48.