NBER WORKING PAPER SERIES

WHEN DID GROWTH BEGIN? NEW ESTIMATES OF PRODUCTIVITY GROWTH IN ENGLAND FROM 1250 TO 1870

Paul Bouscasse
Emi Nakamura
Jón Steinsson

Working Paper 28623
http://www.nber.org/papers/w28623

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021, Revised June 2024

We thank Venance Riblier and Caleb Wroblewski for excellent research assistance. We thank Bernard Beaudreau, Gillian Brunet, Gregory Clark, Ernesto Dal Bó, Brad DeLong, Edward Nelson, James Robinson, Jean-Laurent Rosenthal, Jan De Vries, Jonathan Vogel, David Weinstein, Jacob Weisdorf and seminar participants at various institutions for valuable comments and discussions. We thank the Alfred P. Sloan Foundation, the Smith Richardson Foundation, and the Program for Economic Research at Columbia University for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

When Did Growth Begin? New Estimates of Productivity Growth in England from 1250 to 1870
Paul Bouscasse, Emi Nakamura, and Jón Steinsson
NBER Working Paper No. 28623
March 2021, Revised June 2024
JEL No. N13,O11,O47

ABSTRACT

We estimate productivity growth in England from 1250 to 1870. Real wages over this period were heavily influenced by plague-induced swings in the population. Our estimates account for these Malthusian dynamics. We find that productivity growth was zero prior to 1600. Productivity growth began in 1600—almost a century before the Glorious Revolution. Thus, the onset of productivity growth preceded the bourgeois institutional reforms of 17th century England. We estimate productivity growth of 2% per decade between 1600 and 1800, increasing to 5% per decade between 1810 and 1860. Much of the increase in output growth during the Industrial Revolution is explained by structural change—the falling importance of land in production—rather than to faster productivity growth. Stagnant real wages in the 18th and early 19th centuries—“Engel’s Pause”—is explained by rapid population growth putting downward pressure on real wages. Yet, feedback from population growth to real wages is sufficiently weak to permit sustained deviations from the “iron law of wages” prior to the Industrial Revolution.

Paul Bouscasse
Sciences Po
28 Rue des Saints-Peres
Paris 75007
France
paul.bouscasse@sciencespo.fr

Emi Nakamura
Department of Economics
University of California, Berkeley
685 Evans Hall
Berkeley, CA 94720
and NBER
enakamura@berkeley.edu

Jón Steinsson
Department of Economics
University of California, Berkeley
671 Evans Hall
Berkeley, CA 94720
and NBER
jsteinsson@berkeley.edu

1 Introduction

When did economic growth begin? A traditional view holds that economic growth began with the Industrial Revolution around 1800. Recent work has challenged this view, pushing back the date of the onset of growth. Crafts (1983, 1985) and Harley (1982) revised downward previous estimates of growth in Britain during the Industrial Revolution. These new estimates indicate that British output per capita was larger by the mid-18th century than was previously thought, implying that substantial growth must have occurred at an earlier date (see also Crafts and Harley, 1992). Acemoglu, Johnson, and Robinson (2005) argue that a First Great Divergence occurred starting around 1500 with Western Europe growing apart from other areas of the world following the discovery of the Americas and the sea route to India. They support this view with data on urbanization rates. Broadberry et al. (2015) argue that growth began even earlier than this. They present new estimates of GDP per person for Britain back to 1270. These data show slow but steady growth in GDP per person from the beginning of their sample. Finally, Kremer (1993) uses world population estimates to argue for positive but glacially slow growth for hundreds of thousands of years.

An important facet of the debate about when growth began is when *productivity* growth began. We contribute to this debate by constructing a new series for productivity growth (TFP) in England back to 1250. Figure 1 plots our new productivity series (solid black line). Our main finding is that productivity growth in England began in 1600. Between 1250 and 1600, we estimate that productivity growth was zero.[1] We estimate productivity growth of about 2% per decade from 1600 to 1760. Productivity growth then increased to 5% per decade between 1810 and 1860. We attribute much of the increase in output growth during the Industrial Revolution to structural change—a fall in the importance of land in production—rather than to an increase in productivity growth.

Our results help distinguish between different theories of *why* growth began. They suggest that researchers should focus on developments proximate to the 16th and 17th centuries. An important debate regarding the onset of growth centers on the role of institutional change. Our results help sharpen this debate. We find that productivity growth began almost a century before the Glorious Revolution and well before the English Civil War. While the institutional changes associated with these events may have been important for subsequent growth, researchers must look to earlier

[1] The positive but glacially slow productivity growth rate implied by Kremer's (1993) population data for the period 1200 to 1500 lies within our credible set.

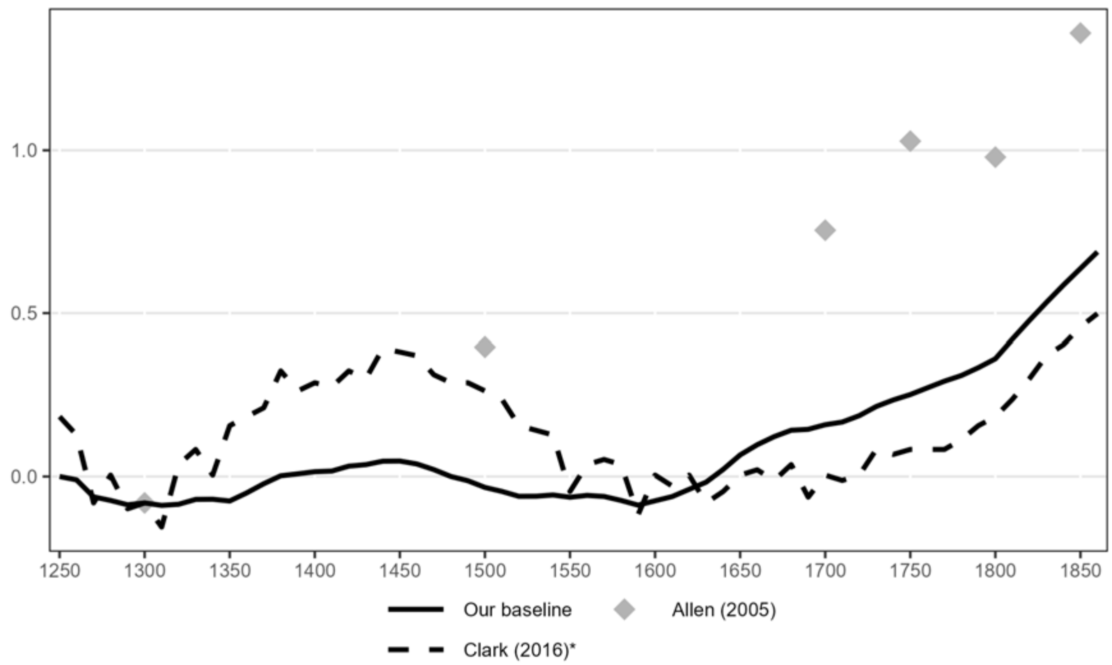

Figure 1: Estimates of Productivity in England

Note: Each series is the natural logarithm of productivity. The series denoted by "Clark (2016)*" is the series from Clark (2016) extended to 1860. We received this series from Clark in private correspondence. Clark's series estimates TFP for the entire economy based on a dual approach. Allen's (2005) estimates are for TFP in the agricultural sector using a primal approach. Our preferred productivity series is normalized to zero in 1250. The other two series are normalized to match our preferred series in 1300.

events for the seeds of modern growth. Plausible candidates include the Reformation, the decline of feudalism, the rise of the yeoman, movable type printing and the associated increase in literacy, and expansion of international trade. We discuss these in more detail below.

The most comprehensive existing productivity series for England was constructed by Clark (2010, 2016). Clark estimated changes in TFP for the entire English economy from 1209 onward using the "dual approach"—i.e., as a weighted average of changes in real factor prices (e.g., Hsieh, 2002). Figure 1 plots Clark's series over our sample period (broken black line).[2] A striking feature of this series is that it implies that productivity in England was no higher in the mid-19th century than in the 15th century. This result does not line up well with other existing (less comprehensive) measures of productivity in England or with less formal assessments of the English economy. For example, Allen (2005) estimates that TFP in agriculture was 162% higher in 1850 than in 1500 (grey diamonds in Figure 1).[3] Clark himself commented that if the fluctuations in his series are not mea-

[2]Clark (2016) published an update of Clark's better-known 2010 series for the shorter time period 1250-1600. The series we plot in Figure 1 is Clark's 2016 series extended to 1860. We received this series from Clark by private correspondence. The 2016 series differs from the 2010 series prior to 1600 due to a new land rent series and because Clark corrected an important error in the 2010 series. We discuss this in more detail in Appendix G.

[3]Allen (2005) employs the familiar "primal approach" of Solow (1957), i.e., subtracts a weighted average of growth

surement error "they imply quite inexplicable fluctuations in the performance of the preindustrial economy."

Our conclusions about productivity in England are quite different from those of Clark (2010, 2016). According to our estimates, productivity in England was 96% higher in 1850 than in 1500 rather than being essentially unchanged. We also estimate smaller fluctuations in productivity prior to 1600. In particular, our productivity series falls much less between 1450 and 1600. These substantial differences arise from differences in the data and methodology we use. We take the labor demand curve as our starting point and estimate changes in productivity as shifts in the labor demand curve. This means that the key data series that inform our estimates are real wages and population. These are arguably among the best measured series of all economic time series over our long sample period.

Our approach is best understood by considering Figure 2. This figure presents a scatter plot of the logarithm of real wages in England (y-axis) against the logarithm of the population in England (x-axis). From 1250 to 1300, the population of England increased and real wages decreased. The period from 1300 to 1450 was a period of frequent plagues—the most famous being the Black Death of 1348. Over this period, the population of England fell by a factor of two and real wages rose substantially. Then, from 1450 to 1600, the population recovered and real wages fell. In 1630, the English economy was back to almost exactly the same point it was at in 1300.

One way to explain these dynamics between 1300 and 1630 is as movements along a stable labor demand curve with no change in productivity. Had productivity grown between 1300 and 1630, the economy could not have returned to essentially the same point in 1630 as it was in 1300 since the labor demand curve would have shifted up and to the right over the intervening period. Then in the 17th century, something important seems to change. The points start moving off the prior labor demand curve. Specifically, they start moving up and to the right relative to the earlier curve. This suggests that productivity started growing in the 17th century in England.

The basic idea behind our approach is to estimate a labor demand curve for England and then back out productivity growth as shifts in this labor demand curve. To get a better sense for how this approach works, consider the following simple labor demand curve for a pre-modern economy

$$W_t = (1-\alpha)A_t\left(\frac{Z}{L_t}\right)^{\alpha},$$

where W_t denotes real wages, A_t denotes productivity (TFP), Z denotes land (which is fixed), and

in factor inputs from output growth, but is only able to do this for agriculture and for a few years.

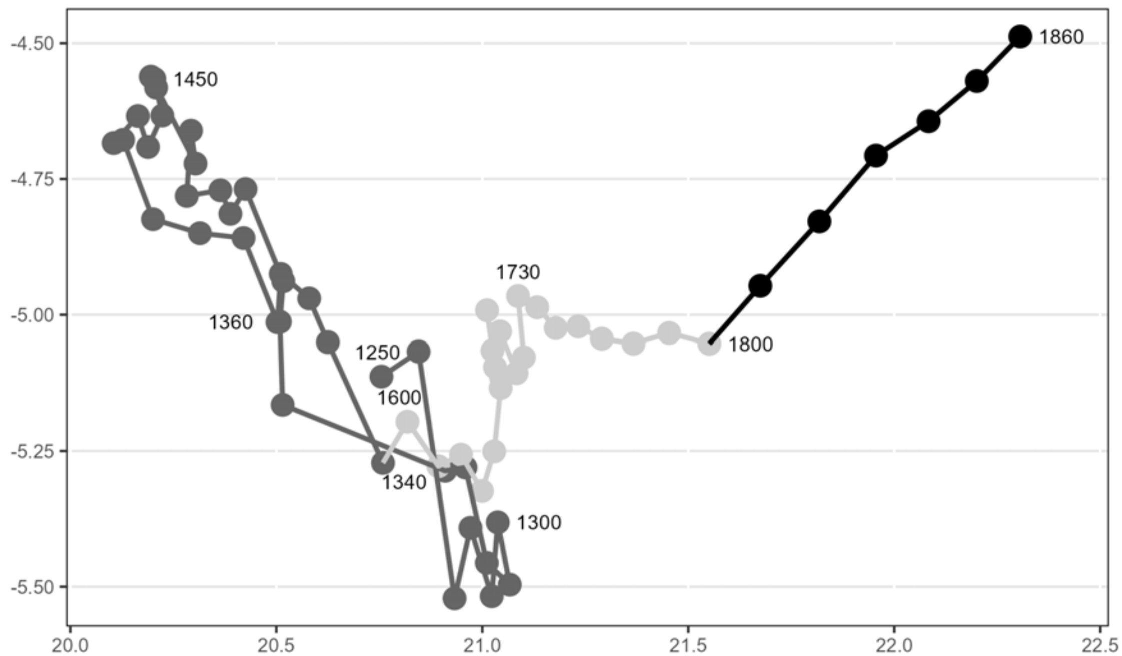

Figure 2: Real Wages and Population

Note: The figure presents a scatter plot of the logarithm of real wages in England against the logarithm of the population in England over the period 1250-1860. The data on real wages are from Clark (2010). Estimates of the population are based on our calculations (baseline case).

L_t denotes labor. We consider more general models later in the paper. But the basic challenge we face can be grasped using this simple model. If we take logarithms, this equation becomes

$$w_t = \phi - \alpha l_t + a_t,$$

where lower case letters denote logarithms of upper case letters. Armed with data on real wages, the population, and an estimate of the slope of the labor demand curve α, one can back out estimates of productivity a_t.

We consider two approaches to estimating the slope of the labor demand curve α. The first is to use the Black Death as a large exogenous shock to the population. In this case, we estimate α simply from data on real wages and population before and after the Black Death. Our second method is to structurally estimate a Malthusian model. In this case, we are modelling the endogenous response of the population to changes in the real wage. These two methods yield similar results.

Since our analysis extends into the early industrial era, we must confront the fact that the importance of land as a factor of production fell rapidly with the spread of steam power, which meant that the production of energy was no longer land intensive (Wrigley, 2010). To capture this crucial development, we allow the output elasticity of land, capital, and labor change over time after the onset of the Industrial Revolution. We use data on land rents after 1760 to pin down how rapidly the importance of land in production fell. The modest increase in land rents that we observe in the face of explosive growth in labor and capital after 1760 leads us to estimate a rapidly falling importance of land in production.

The fact that we allow for this structural transformation implies that the standard way of measuring productivity (a multiplicative A_t in front of a function $F(L_t, K_t, ...)$) is no longer valid. Following Caves, Christensen, and Diewert (1982) we derive a Malmquist productivity index (Malmquist, 1953) for our setting. The Malmquist index reduces to A_t in the familiar setting of constant factor elasticities, but remains valid even when the structure of the production function is changing.

Allowing the importance of land in production to fall after the onset of the Industrial Revolution has important implications for our estimates of productivity. If we don't allow for this change, we estimate a much larger break in productivity in 1760. Productivity growth is, of course, just a measure of our ignorance. Modelling the shift of production away from land intensive technology allows us to explain a larger part of growth during the Industrial Revolution, leaving less for the residual.

Our estimates also shed light on the lack of real wage growth during the latter part of the 18th century, sometimes referred to as "Engels' Pause" (Engels, 1845, Allen, 2009b). Our Malthusian model implies that during this period real wages were held back by very rapid increases in the population, which in a Malthusian world put downward pressure on the marginal product of labor. This explanation contrasts with the common idea that the absence of real wage growth during this period resulted from the lion's share of the fruits of technical change going to capital as opposed to labor. This idea has received attention in the modern context in relation to the development of automation and artificial intelligence (Acemoglu and Restrepo, 2019a).

In addition to estimates of productivity, our methodology yields estimates of the speed of Malthusian population dynamics in pre-modern England. Our estimates imply that these population dynamics were very slow: a doubling of real incomes led to an increase in population growth that was only about 3 to 6 percentage points per decade. Together with our other estimates, this implies that the half-life of a plague-induced drop in the population was more than 100 years prior to the onset of the Industrial Revolution. As the importance of land in production fell after 1760, the Malthusian population dynamics became even slower and weaker. By 1860, our estimate of the half-life of a population shock have risen to several hundred years. Earlier estimates of the speed of Malthusian population dynamics in England also indicate that they were slow. For example, Lee and Anderson (2002) find a half-life of 107 years, while Crafts and Mills (2009) find a half-life of 431 years. Chaney and Hornbeck (2016) document very slow population dynamics in Valencia after the expulsion of the Moriscos in 1609.

The weakness of the Malthusian population dynamics we estimate imply that our model is consistent with sustained deviations from "the iron law of wages" (i.e., that wages in a Malthusian economy are stuck at subsistence). Modest productivity growth over a few centuries can temporarily overwhelm the Malthusian population dynamics in our model and result in sustained periods of real wages several times higher than at other times. Our model can therefore make sense of episodes that historians sometimes refer to as "golden ages" or "efflorescences" (Goldstone, 2002). Once productivity growth falters, real wages will slowly fall back to a lower level. But this will take several centuries.

We are not the first to plot a figure like Figure 2 and argue that it has implications about the evolution of productivity in England. We formalize this intuitive idea and assess what exactly it implies about productivity. Clark (2005, 2007a) discuss informally how shifts in the labor demand curve of a Malthusian model can be informative about the timing of the onset of economic growth. The existing papers most closely related to ours from a methodological point of view are Lee and Anderson (2002) and Crafts and Mills (2009). These papers structurally estimate a Malthusian model of the English economy, as we do. Relative to these papers, we extend the sample period back in time considerably (theirs starts in 1540 while ours start in 1250). This allows us to assess when growth began. We also estimate α differently, incorporate capital, and allow the importance of land to change after the onset of the Industrial Revolution, among other differences.

Our paper is also related to the literature in macroeconomics on the transition from preindustrial stagnation to modern growth—often referred to as the transition "from Malthus to Solow." Important papers in this literature include Galor and Weil (2000), Jones (2001), and Hansen and Prescott (2002). Relative to these papers, our work is more empirical. We contribute detailed estimates of the evolution of productivity, while these papers propose theories of how productivity growth rose. Our work is also related to recent work by Hansen, Ohanian, and Ozturk (2020).

Our paper proceeds as follows. Section 2 presents a simple estimate of productivity growth in England with α estimates from the Black Death. Section 3 presents our full Malthusian model of the economy. Section 4 presents our results on productivity based on the full model. Section 5 presents our results on the strength of the Malthusian population force. Section 6 presents our estimates of the population. Section 7 concludes.

2 A Simple Estimate of Productivity Growth in England

We begin by presenting a very simple estimate of productivity growth in England. Later sections develop a number of extensions and refinements to the basic approach adopted in this section. We model time as discrete and denote it by a subscript t. Since we use decadal data throughout the paper, each time period in the model is meant to represent a decade. Our sample period is from 1250 to 1860. All the data we use are decadal averages. In our figures, a data point listed as 1640 refers to the decadal average from 1640 to 1649. We sometimes refer to a variable at a point in time (say 1640) when we mean the decadal average for that decade. In other words, we use 1640 and "the 1640s" interchangeably.

Consider an economy where output is produced with land and labor according to the following production function:

$$Y_t = A_t Z^\alpha L_t^{1-\alpha},$$

where Y_t denotes output, Z denotes land (which is fixed), L_t denotes labor, and A_t denotes productivity (TFP). We assume that producers hire workers in a competitive labor market taking wages as given. Producer optimization then gives rise to the following labor demand curve:

$$W_t = (1-\alpha) A_t Z^\alpha L_t^{-\alpha},$$

where W_t denotes the real daily wage. Taking logarithms of this equation yields

$$w_t = \phi + a_t - \alpha l_t, \tag{1}$$

where lower case letters denote logarithms of upper case letters and $\phi = \log(1-\alpha) + \alpha \log Z$. Assuming that labor is paid its marginal product in pre-industrial England is a strong assumption that we discuss in greater detail below.

To estimate changes in productivity using equation (1), we need data on real wages, data on labor supply, and an estimate of the slope of the labor demand curve α. We discuss these in turn.

Figure 3: Real Wages in England, 1250-1860

Note: The figure presents estimates of the real wages of unskilled building workers in England from Clark (2010).

2.1 Data on Real Wages and the Population

Our baseline measure of real wages in England is Clark's (2010) series for unskilled building workers. Figure 3 plots this series. The main features of the series are a large and sustained rise between 1300 and 1450, a large and sustained fall between 1450 and 1600, some recovery over the 17th century, stagnation during the 18th century, and finally a sharp increase after 1800. Figure A.1 compares this series with several other series for real wages in England. This comparison shows that the real wage series we use is quite similar to Clark's real wage series for farmers and for craftsmen. We have redone our analysis with these series and discuss this analysis in section 4.3.[4]

Our baseline assumption is that labor supply is proportional to the population of England. Figure 4 presents the population data that we use. For the period from 1540 onward, we use population estimates from Wrigley et al. (1997), which in turn build on the seminal work of Wrigley and Schofield (1981). Sources for population data prior to 1540 are less extensive. Clark (2007b) uses unbalanced panel data on the population of villages and manors from manorial records and

[4]Much controversy has centered on the behavior of real wages in England between 1770 and 1850. This debate revolves around the extent to which laborers shared in the benefits of early industrialization (see, e.g. Feinstein, 1998, Clark, 2005, Allen, 2007, 2009b). In Figure A.1, we also plot Allen's (2007) wage series (which starts in 1770). The figure shows that the differences discussed in the prior literature are modest from our perspective and therefore do not materially affect our analysis.

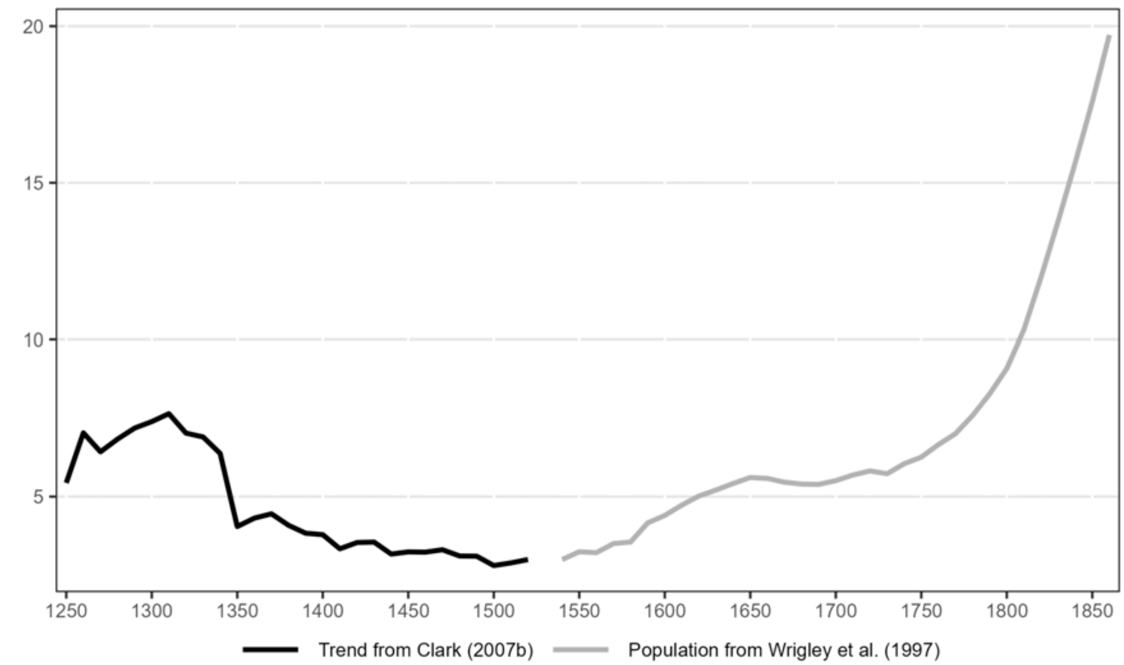

Figure 4: Population Data for England, 1250-1860

Note: The figure presents population estimates for England for the period 1540-1860 from Wrigley et al. (1997) (grey line) and the population trend estimates by Clark (2007b) for the period 1250-1520 (black line). In this figure, the black line is normalized for visual convenience such that its last point is equal to the first point of the grey line.

penny tithing payments to construct estimates of the population prior to 1540. We build on Clark's work to construct an estimate of the population before 1540.

We cannot directly use Clark's pre-1540 population series since Clark's method for constructing his series involves making assumptions about the evolution of productivity.[5] Since we aim to use the population series to make inference about the evolution of productivity in England, we cannot use a population series that already embeds assumptions about productivity growth. However, as an intermediate input into constructing his pre-1540 population series, Clark estimates a regression of his village and manor level population data on time and village/manor fixed effects. Clark refers to the time effects from this regression as a population trend. We plot this population trend in Figure 4 (normalized for visual convenience). We base our estimates of the population of England prior to 1540 on this population trend series. In section 4.3, we discuss how this series compares to (lower frequency) population data reported in Broadberry et al. (2015).

We assume that the true population is measured with error. Specifically, we assume that $n_t = \psi + \tilde{n}_t + \iota_t^n$, where n_t denotes the true unobserved population, $\tilde{n}_t$ denotes our observed population

[5]Appendix A discusses Clark's method in more detail.

series (Clark's population trend series prior to 1530 and the population series from Wrigley et al. (1997) after 1530), $\iota_t^n \sim t_{\nu_n}(0, \sigma_n^2)$ denotes measurement error, and ψ denotes a normalization constant. We normalize ψ to zero after 1530 and estimate its value for the pre-1530 Clark series. We allow for a structural break in the variance of the measurement error σ_n^2 in 1540. Finally, population data are missing for 1530. We view it as unobserved an estimate its value.

2.2 Identification and Estimation

We estimate the slope of the labor demand curve—α in equation (1)—under the assumption that the Black Death is an exogenous shock to the population. Specifically, we estimate α as the ratio of the change in real wages and the change in the population between 1340 and 1360. This is the slope between the point for 1340 and the point for 1360 in Figure 2.[6] The identifying assumption is that the Black Death dominates the change in wages and population between 1340 and 1360. We choose 1360 as opposed to 1350 as our post Black Death point due to attempts by the lords in England to keep wages low in the immediate aftermath of the Black Death. In 1351, the Statute of Laborers was passed which set a maximum wage equal to the prevailing wage prior to the Black Death. These attempts may have had some effect on wages during the 1350s, but wages responded strongly in the 1360s.

This procedure yields a value of $\alpha = 0.70$, which lines up well with the overall slope of the points in Figure 2 prior to the 17th century. This suggests that much of the variation in the population and real wages of England prior to the 17th century was driven by labor supply shocks such as plagues and other disease.[7] Armed with data on w_t and l_t and an estimate of α, we can "back out" values for log productivity, a_t, from equation (1) (up to a constant). Since we are primarily interested in persistent changes in productivity, we filter out high frequency variation in a_t by assuming that a_t is made up of a permanent and transitory component. We assume the permanent component follows a random walk process with drift μ:

$$\tilde{a}_t = \mu + \tilde{a}_{t-1} + \epsilon_{1t}. \tag{2}$$

[6] Figure 2 adjusts for measurement error. To estimate α, we use the raw data for 1340 and 1360. Our estimate of α is, therefore, only approximately the slope between the point for 1340 and the point for 1360 in Figure 2.

[7] In appendix B, we show that the relationship between the land share of production and the slope of the labor demand curve is sensitive to the elasticity of substitution between land and labor. A CES production function implies that the land share is $\sigma\alpha$, where σ is the elasticity of substitution between land and labor. Assuming an elasticity of substitution of 0.66 (Boppart et al., 2023) implies a land share of $0.66 \times 0.70 = 0.46$. More generally, additional parameters are needed to go from the slope of the labor demand curve to factor shares.

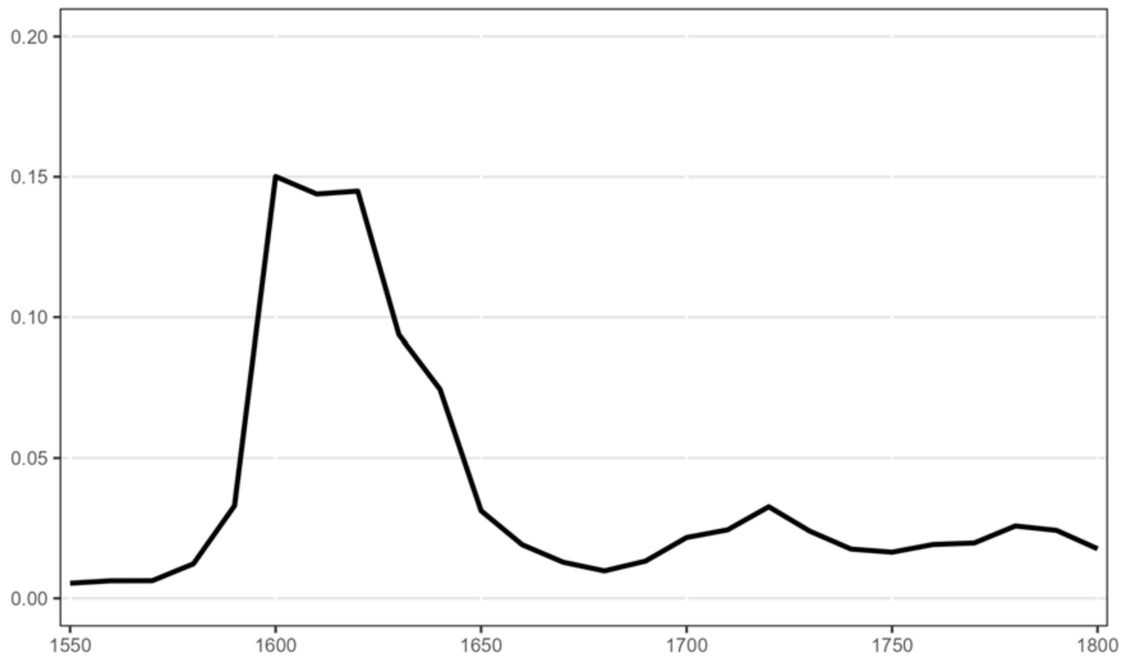

Figure 5: Probability of Different Productivity Growth Break Dates

Note: The figure plots our estimate of the probability that a structural break occurred in the parameters μ, σ_1, and σ_2 in different decades between 1550 and 1800.

The average growth rate of productivity is given by the parameter μ. To capture changes in long-run growth, we allow for structural breaks in μ. Recall that changes in productivity correspond to shifts in the labor demand curve. We allow for a break in 1810 associated with the Industrial Revolution. But Figure 2 suggests that the labor demand curve began shifting out earlier. As a consequence, we allow for one additional break earlier in the sample. This allows for the possibility that there may have been a break in average growth before the Industrial Revolution. We consider a range of possible dates for the first break between 1550 and 1800.[8] The method we use to pin down the timing of this earlier break and to estimate the model more generally is described in detail in section 3.

2.3 When Did Growth Begin?

Figure 5 plots our estimates of the probability that a pre-Industrial Revolution break in productivity growth occurred at different dates between 1550 and 1800. The probability of a break spikes in 1600. It stays high in 1610 and 1620 and then falls off. The probability is very low prior to 1590 and also low after 1650. The probability of a break occurring before 1640—i.e., before the onset of the

[8]The fact that we allow for permanent breaks in productivity growth implies that productivity growth has a unit root component. This allows our model to match the fact that the population is integrated of order two in our sample, which has been emphasized in prior work on this topic (Bailey and Chambers, 1993, Crafts and Mills, 2009).

English Civil War—is estimated to be 60%. The probability of a break occurring before 1680—i.e., before the Glorious Revolution—is 73%. For expositional simplicity, we date the break in 1600. Results from our more complex models in section 3 yield 1600 as the most likely break date (see Figure A.2).

Table 1 presents our estimates of the average growth rate of productivity μ for the three regimes over our sample. We estimate that average productivity growth prior to 1600 was zero. Kremer (1993) used data on the growth rate of the world population to argue that growth has been non-zero and increasing for many millennia. The world population estimates he used indicate that world population growth from 1200 to 1500 was 0.6% per decade. In our Malthusian model (as well as Kremer's model), steady state productivity growth is α times steady state population growth. Using our estimate of α discussed above, this suggests that growth in productivity was 0.4% per decade over the period 1200 to 1500, i.e., positive but glacial. This slow growth rate is well within the credible set of our pre-1600 estimate of μ.

Our results indicate that sustained productivity growth began in 1600 (or around that time). We estimate average productivity growth of 4% per decade over the period 1600 to 1810. In the early 19th century, productivity growth accelerated sharply to 19% per decade. We conclude from these estimates that the period from 1600 to 1810 was a period of transition in England from an era of total stagnation to an era of modern economic growth.

Figure 6 presents our baseline estimates of the time series evolution of the permanent component of productivity. These estimates indicate that the level of productivity in England was very similar in 1600 to what it had been in the late 13th century. In the intervening period, productivity fluctuated a slight bit. After 1600, productivity began a sustained increase, which accelerated sharply in 1810. As in most models in macroeconomics, productivity growth is a residual. In later sections, we consider more complex models with capital and in which we allow the elasticity

Table 1: Productivity Growth

	Mean	St Dev	2.5%	97.5%
$\mu_{a,1}$	0.00	0.01	-0.01	0.02
$\mu_{a,2}$	0.04	0.02	0.02	0.09
$\mu_{a,3}$	0.19	0.01	0.17	0.22

Note: The table presents the mean, standard deviation, 2.5% quantile, and 97.5% quantile of the posterior distribution we estimate for average productivity growth μ in the three regimes, using the simple procedure described in section 2. See table A.1 for the posterior distribution of $\sigma_{\epsilon_{1,t}}$ and $\sigma_{\epsilon_{2,t}}$.

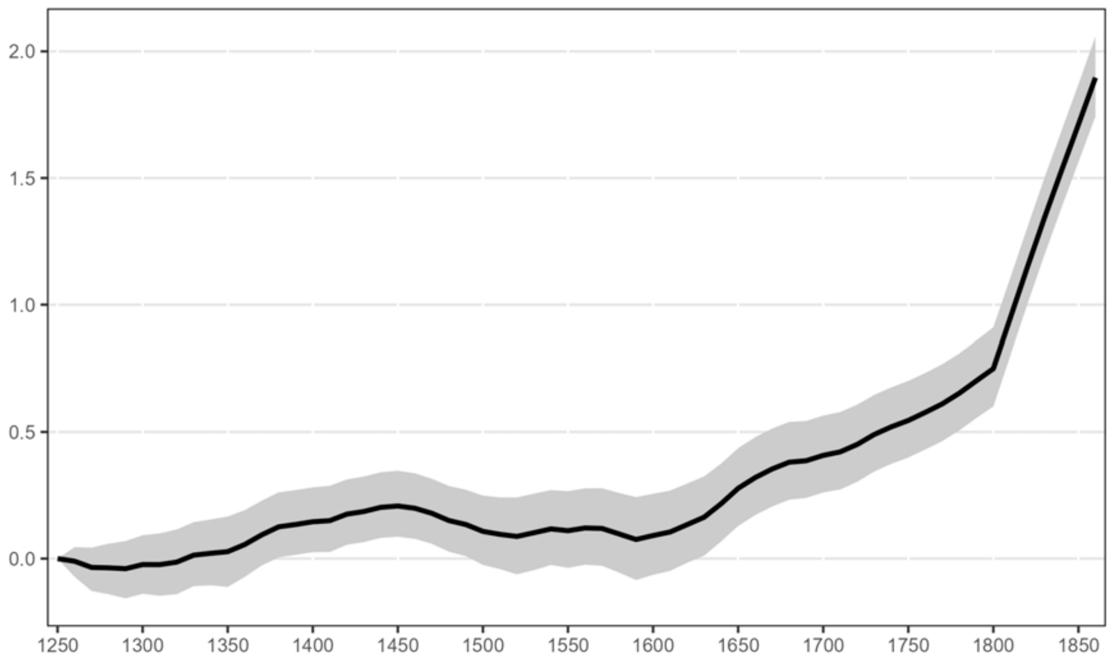

Figure 6: Permanent Component of Productivity

Note: The figure plots our estimates of the evolution of the permanent component of productivity $\tilde{a}_t$ over our sample period (natural logarithm relative to its value in 1250). The black line is the mean of the posterior for each period and the gray shaded area is the 90% central posterior interval.

parameters in the production function to change over time. In these cases, we attribute a substantial fraction of the post-1800 productivity growth in Figure 6 to capital deepening and the falling importance of land in production.[9]

2.4 Real Wages, Productivity, and Engels' Pause

Figure 7 compares our estimate of productivity with the data we use on real wages. This figure illustrates well the importance of accounting for Malthusian population forces when estimating productivity in the pre-industrial era. Our analysis implies that the large changes in real wages prior to 1600 are explained almost entirely by changes in the population and almost not at all by changes in productivity. During this period, the economy moved up and down a relatively stable labor demand curve as suggested by Figure 2. First, plagues reduced the population and this increased wages. Then, the population recovered from these plagues and real wages fell. As a result, changes in productivity were very substantially muted relative to changes in real wages.

[9] Table A.1 presents estimates of the standard deviation of the permanent and transitory productivity shocks. These vary very little across the three regimes. The standard deviation of the permanent productivity shocks is 0.03 prior to 1600 and 0.02 after 1600. The standard deviation of the transitory productivity shocks is 0.05 prior to 1600 and 0.04 after 1600.

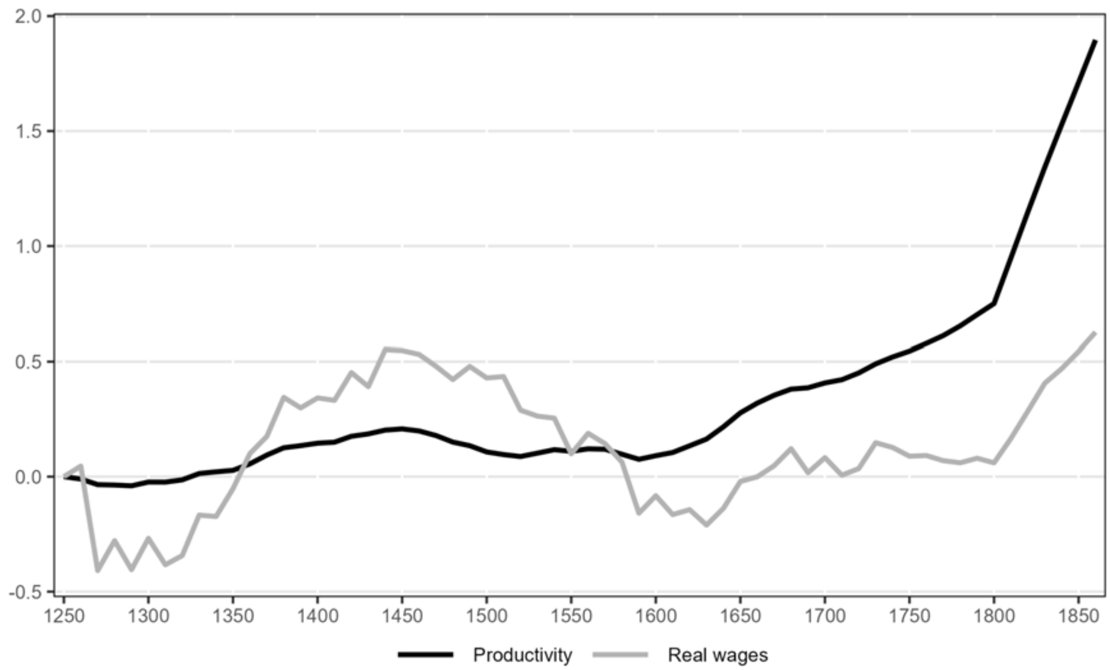

Figure 7: Productivity and Real Wages

Note: The figure plots our estimates of the evolution of the permanent component of productivity $\tilde{a}_t$ along with the real wage series we use.

Real wages are also a poor guide to changes in productivity for the period from 1600 to 1860. Over this period, however, the pattern is reversed: productivity moved substantially *more* than real wages. Productivity started to grow after 1600. But this induced rapid growth in the population. Since the labor demand curve was downward sloping due to land being a fixed factor, growth in the population put downward pressure on wages. Over this period, therefore, real wages lagged behind growth in productivity.

Our analysis provides a simple explanation for "Engels' Pause," the fact that real wages did not rise appreciably during the early decades of industrialization in England. We see in Figure 7 that over the period 1730 and 1800 real wages in England fell slightly despite substantial productivity growth. (The real wage series of Feinstein (1998) and Allen (2007)—which differ somewhat from Clark's real wage series—extend the Pause a few decades into the 19th century.) One explanation for this fact—famously articulated by Engels (1845)—is that the gains from capitalism overwhelmingly accrue to capitalists as opposed to laborers. Our analysis suggests an alternative Malthusian explanation: rapid growth in the population put downward pressure on the marginal product of labor and thus reduced the growth in wages relative to productivity.[10]

[10]Allen (2009b) comes to a similar conclusion using quite different data and methods. He also concludes that population growth was a crucial contributor to stagnant real wages in England during Engels' Pause: "population growth was a necessary condition for stationary real wages: Engels' pause looks like Malthus' dismal science come true" (p.

One reason we chose to use the wage series for unskilled building workers in our analysis is to capture a part of the labor market that involved relatively voluntary labor and therefore a better measure of the marginal product of labor than, e.g., the work of the villein in the countryside. Nevertheless, our assumption of a competitive labor demand curve is clearly a simplification, and it is important to consider how our results might change were we to weaken this assumption. A constant wage markdown due to monopsony power or coercion by employers would not affect our results. It would only affect the estimate of the constant ϕ in equation (1). A wage markdown that was getting smaller over time as the labor market became more free and competitive would show up as an increase in productivity in our analysis. So long as this process was slow—which seems likely—the bias in our results arising from this issue would also be small.

2.5 From When to Why

By dating the onset of productivity growth, our results help discriminate between competing explanations for *why* growth began. We estimate that sustained productivity growth began in England substantially before the Glorious Revolution of 1688. According to our estimates, productivity in England rose by 34% from 1600 to 1680. North and Weingast (1989) argue that the political regime that emerged in England after the Glorious Revolution—characterized by a power sharing arrangement between Parliament, the Crown, and the common law courts—resulted in secure property rights and rule of law and thereby laid the foundation for economic growth. While the institutional changes associated with the Glorious Revolution may well have been important for growth, our results indicate that the seeds of growth in England were sown earlier.

Our results support explanations of the onset of growth that focus attention on developments that occurred in the period surrounding 1600. The Reformation is an obvious candidate. In particular, Henry VIII's confiscation of monastic lands was a big shock to land ownership patterns and the land market in England (Heldring, Robinson, and Vollmer, 2021). England also became a favored destination for skilled immigrants fleeing religious persecution on the continent. This was, furthermore, a period of rapidly increasing urbanization in England. London experienced an explosion of its population around this time—from 55,000 in 1520 to 475,000 in 1670 (Wrigley,

430). Allen's model is more complex than the simple model we analyze in this section. In his model, slow accumulation of capital and a low elasticity of substitution between capital and labor also contribute to low growth in real wages during Engels' Pause. We extend our model to include capital in section 3 below. In Allen's model, growth in capital eventually catches up to growth in the population leading real wage growth to pick up. Our analysis below places a greater emphasis on reductions in the importance of land in production in eliminating the Malthusian character of growth over the course of the 19th century.

2010)—likely due to a rapid increase in international trade. English woolen exports expanded rapidly over this period (new draperies) as did intercontinental trade, colonization, and privateering. The British East India Company was founded in 1600 and the Virginia Company founded its first permanent settlement in North America in 1607.

Our finding that the onset of growth preceded both the Glorious Revolution and the English Civil War (1642-1651) lends support to the Marxist view that economic change propelled history forward and drove political and ideological change. Marx (1867) stressed the transition from feudalism to capitalism. He argued that after the disappearance of serfdom in the 14th century, English peasants were expelled from their land through the enclosure movement. That spoliation inaugurated a new mode of production: one where workers did not own the means of production, and could only subsist on wage labor. This proletariat was ripe for exploitation by a new class of capitalist farmers and industrialists. In that process, political revolutions were a decisive step in securing the rise of the bourgeoisie. To triumph, capitalism needed to break the remaining shackles of feudalism. As the *Communist Manifesto* puts it, "they had to be burst asunder; they were burst asunder" (Marx and Engels, 1848, pp. 40-41). Hill (1940, 1961) offers more recent treatments of the political revolutions in England in the 17th century that stress class conflict and their economic origins.

Acemoglu, Johnson, and Robinson (2005) synthesize the Marxist and institutionalist views. They argue that Atlantic trade enriched a merchant class that then demanded secure property rights and secured these rights through the Civil War and Glorious Revolution. This last narrative lines up well with our result that steady growth—perhaps driven by the Atlantic trade—began about half a century before the Civil War. However, we do not detect a radical increase of growth in the immediate aftermath of either the Civil War or the Glorious Revolution: 3.2% (1600-1640), 4.2% (1640-1680) and 4.5% (1680-1810).

Allen (1992) argues that a long and gradual process of institutional change in England over the 600-year period from the Norman Conquest to the Glorious Revolution resulted in a situation in the 16th century where the yeoman class had acquired a substantial proprietary interest in the land, and thus an incentive to innovate. The timing of Allen's 'rise of the yeoman' lines up reasonably well with our estimate of the onset of growth. According to Allen, property rights, rule of law, and personal freedom gradually expanded, and the social order was gradually transformed from a feudal to a capitalist order. From the 12th century, royal courts helped freeholders gain full ownership over their land. After the Black Death, serfdom collapsed as landlord competed for scarce labor. Early enclosures (15th and early 16th centuries) involved brutal evictions and depopulation of manors. The Crown reacted to this by increasing protection of tenant farmers.

The spread of movable-type printing across Europe after 1450 led to a large increase in literacy in England in the 16th and 17th centuries (Cressy, 1980, Houston, 1982), and a huge drop in the price of books (Clark and Levin, 2011). Dittmar (2011) argues that cities exposed to printing grew substantially faster than otherwise similar cities. Printing and literacy likely had wide ranging effects on culture. Mokyr (2009, 2016) and McCloskey (2006, 2010, 2016) have argued that the crucial change that caused growth to begin was the emergence of a culture of progress based on the idea that mankind can improve its condition through science and rational thought. Others have stressed a Protestant ethic (Weber, 1905) and Puritanism (Tawney, 1926). The timing of these changes lines up reasonably well with our estimates although it is not straightforward to pinpoint precisely what these theories imply about the timing of the onset of growth.

Bogart and Richardson (2011) stress the importance of the post-Glorious Revolution regime's push to reorganize and rationalize property rights through enclosures, statutory authority acts, and estate acts. While our results contradict the notion that growth began with the Glorious Revolution, the fact that England underwent massive institutional change in the 17th and 18th century may have played an important role in sustaining growth during this period.

Allen (2009a) argues that the Industrial Revolution occurred in Britain around 1800 because innovation was uniquely profitable then and there. His theory relies on growth in the 17th century leading to high real wages in England in the 18th century as well as the development of a large coal industry. High wages and cheap coal made it profitable to invent labor saving technologies in textiles such as the spinning jenny, water frame, and mule, as well as coal burning technologies such as the steam engine and coke smelting furnace. While our theory does not point to the Industrial Revolution as the genesis of economic growth, Allen's theory helps explain how growth was sustained and the particular direction in took that led to the huge fall in the importance of land in production that we estimate later in the paper.

3 A Malthusian Model of the Economy

The simple analysis in section 2 makes a number of strong assumptions. In the remainder of the paper, we explore a richer framework that allows us to relax some of these. Most importantly, we model the evolution of the population. Thus, we can present novel estimates of the strength of Malthusian population forces in pre-industrial England. We explicitly incorporate plague shocks into this part of the model, which helps the model distinguish between measurement error and true variation in the population. We also incorporate capital accumulation into the model, explore alternatives to the assumption that labor supply was proportional to the population, allow for a falling importance of land in production after the onset of industrialization, and estimate the slope of the labor demand curve using the entire data set as opposed to identifying it from the Black Death. This richer framework allows us to explore the robustness of our conclusions from section 2 and to present a number of additional interesting results.

While our focus is on the pre-industrial period, our data extends well into the early industrial period. It is therefore important for our model to capture both the character of the pre-industrial economy and the early industrial economy. The role of land in production is particularly important in this regard. Prior to the Industrial Revolution, land was a hugely important factor of production. However, the advent of steam power led to a sharp fall in the role of land in production as fossil fuels substituted for human and animal power in the production of energy (and the role of food production in the economy shrank). To capture this change, we distinguish between the pre-industrial period and the early industrial period and allow the importance of land in production to change after the onset of the Industrial Revolution. Section 3.1 presents our model of the pre-industrial economy, while section 3.2 presents our model of the early industrial economy.

3.1 A Model of the Pre-industrial Economy

Output is produced with land, capital, and labor according to the following production function:

$$Y_t = F_t(Z, K_t, L_t) = A_t Z^{\alpha} K_t^{\beta} L_t^{1-\alpha-\beta},$$

where K_t denotes capital.[11] Producer optimization gives rise to the following labor demand curve:

$$W_t = (1-\alpha-\beta) A_t Z^{\alpha} K_t^{\beta} L_t^{-\alpha-\beta}.$$

Taking logarithms of this equation yields

$$w_t = \tilde{\phi} + a_t + \beta k_t - (\alpha+\beta) l_t, \tag{3}$$

where lower case letters denote logarithms of upper case letters and $\tilde{\phi} = \log(1-\alpha-\beta) + \alpha \log Z$.

[11] Appendix C presents results for a more general production function.

Producers accumulate capital to the point where the marginal product of capital is equal to its user cost. This gives rise to the following capital demand equation:

$$r_t + \delta = \beta A_t Z^\alpha K_t^{\beta-1} L_t^{1-\alpha-\beta}, \tag{4}$$

where r_t is the rental rate for capital and δ is the rate of depreciation of capital. Since we do not have data on capital for the pre-industrial period, we use the capital demand equation—equation (4)—to eliminate K_t from the labor demand equation—equation (3). Taking logs of the resulting equation yields

$$w_t = \tilde{\phi}' + \frac{1}{1-\beta} a_t - \frac{\alpha}{1-\beta} l_t - \frac{\beta}{1-\beta} \log(r_t + \delta), \tag{5}$$

where

$$\tilde{\phi}' = \frac{\beta}{1-\beta} \log \beta + \log(1-\alpha-\beta) + \frac{\alpha}{1-\beta} \log Z.$$

The logarithm of productivity a_t is the sum of a permanent and transitory component:

$$a_t = \tilde{a}_t + \epsilon_{2t}, \tag{6}$$

where $\epsilon_{2t} \sim \mathcal{N}(0, \sigma^2_{\epsilon_2})$ is the transitory component and $\tilde{a}_t$ is the permanent component of productivity, which follows a random walk with drift

$$\tilde{a}_t = \mu + \tilde{a}_{t-1} + \epsilon_{1t}, \tag{7}$$

with $\epsilon_{1t} \sim \mathcal{N}(0, \sigma^2_{\epsilon_1})$. Both ϵ_{1t} and ϵ_{2t} are independently distributed over time. The transitory component of productivity may reflect both measurement error and true transitory variation in productivity (e.g., due to transitory variation in weather). As in most macroeconomic models, productivity is a catch-all variable capturing the influence of all variables that are not explicitly modeled in the production function. Relative to the model in section 2, we now incorporate capital accumulation. Our measure of productivity is narrower as a consequence. Productivity in this model nonetheless captures a number of features of reality in addition to technology, including institutions, the effects of international trade, and colonial exploitation.

The average growth rate of productivity is given by the parameter μ. As in section 2, we allow for two breaks in μ, i.e., two changes in the average growth rate of productivity. (We allow the variances of the transitory and permanent productivity shocks—$\sigma^2_{\epsilon_1}$ and $\sigma^2_{\epsilon_2}$—to break at the same times.) We fix one of these breaks in 1810. This break captures the Industrial Revolution. We allow for another break earlier in the sample between 1550 and 1800.

To pin down the timing of the first break, we estimate a mixture model. Since μ, $\sigma^2_{\epsilon_1}$ and $\sigma^2_{\epsilon_2}$ break twice, they take on three values: one for each regime. We denote these as $\mu(i)$ with $i \in \{1, 2, 3\}$ (with analogous notation for $\sigma^2_{\epsilon_1}$, and $\sigma^2_{\epsilon_2}$). From the beginning of our sample until 1540, $\mu = \mu(1)$. From 1550 until 1800, $\mu = (1 - I)\mu(1) + I\mu(2)$, where I is an indicator variable that switches from zero to one at the time of the first break. Finally, from 1810 until 1860, $\mu = \mu(3)$. The indicator variable I has a multinomial distribution with probabilities of switching from zero to one at each date between 1550 and 1800. We estimate the probabilities of the multinomial distribution for I. The prior for these probabilities is a Dirichlet distribution with concentration vector $c_b \times (1, ..., 1)$. We choose a small value for c_b. This ensures that each draw from the distribution is close to a corner of the distribution, i.e. chooses a specific break date. In particular, we set $c_b = 0.001$.[12] The output from our estimation of these probabilities is a posterior probability distribution over break dates.

We assume that the labor force in the economy is proportional to the population and that each worker works D_t days per year. This implies that

$$L_t = \Lambda D_t N_t,$$

where N_t denotes the population and Λ is a constant. Taking logs of this equation and using the resulting equation to eliminate l_t in equation (5) yields

$$w_t = \phi + \frac{1}{1-\beta} a_t - \frac{\alpha}{1-\beta}(d_t + n_t) - \frac{\beta}{1-\beta} \log\left(r_t + \delta\right), \tag{8}$$

where $\phi = \tilde{\phi}' - \alpha/(1-\beta)\lambda$.

A central aspect of our model is the law of motion for the population. Following Malthus (1798), we assume that population growth is increasing in real income:

$$\frac{N_t}{N_{t-1}} = \Omega(W_{t-1} D_{t-1})^{\gamma} \Xi_t,$$

where Ω is a constant, γ is the elasticity of population growth with respect to real income, and Ξ_t

[12] For a simple exposition of the Dirichlet distribution and the role of the concentration parameter, see Stan Function Reference, section 23: `https://mc-stan.org/docs/functions-reference/dirichlet-distribution.html`.

denotes other (exogenous) factors affecting population growth. Taking logarithms of this equation yields

$$n_t - n_{t-1} = \omega + \gamma(w_{t-1} + d_{t-1}) + \xi_t. \tag{9}$$

Malthus argued that both the birth rate and the death rate varied with real income. He described "preventive checks" on population growth that lowered birth rates. These included contraception, delayed marriage, and regulation of sexual activity during marriage. Malthus also described "positive checks" on population growth that raised death rates. These include disease, war, severe labor, and extreme poverty. In our model, the parameter γ captures the elasticity of both birth rates and death rates with respect to income. This parameter therefore captures any tendency of either preventive or positive checks to lower population growth when income falls.

We allow for two types of exogenous population shocks:

$$\xi_t = \xi_{1t} + \xi_{2t}. \tag{10}$$

First, we allow for "plague" shocks:

$$\exp(\xi_{1t}) \sim \begin{cases} \beta(\beta_1, \beta_2), & \text{with probability } \pi \\ 1, & \text{with probability } 1 - \pi \end{cases} \tag{11}$$

These plague shocks occur infrequently (with probability π) but when they occur they kill a (potentially sizable) fraction of the population. The fraction of the population that survives follows a beta distribution $\beta(\beta_1, \beta_2)$. In addition to the plague shocks, we allow for a second type of population shock: $\xi_{2t} \sim \mathcal{N}(0, \sigma^2_{\xi_2})$. Both of the population shocks are independently distributed over time. Together these population shocks are meant to capture a host of potential influences on population growth, in addition to plagues.

3.2 A Model of the Early Industrial Economy

The last century of our data covers the early industrial period in England. A crucial development over this period is the rapid fall in the importance of land as a factor of production. The primary driving force in this development was the introduction of the steam engine powered by fossil fuels. This technological advance meant that the production of energy was no longer land intensive (Wrigley, 2010). Table 2 presents data on energy consumption from various sources over our sam-

Table 2: Energy Consumption in England and Wales

	1560	1700	1750	1800	1850
Draught Animals	21.1	32.8	33.6	34.3	50.1
People	14.9	27.3	29.7	41.8	67.8
Firewood	21.5	22.5	22.6	18.5	2.2
Wind	0.2	1.4	2.8	12.7	24.4
Water	0.6	1.0	1.3	1.1	1.7
Coal	6.9	84.0	140.8	408.7	1,689.1
Total	65.1	168.9	230.9	517.1	1,835.3
Total Less Coal	58.2	84.9	90.1	108.4	146.2

Note: The figure plots energy consumption in England and Wales from various sources in petajoules. It reproduces a portion of Table 2.1 in Wrigley (2010).

ple period. Prior to the introduction of steam power, the vast majority of energy was derived from draught animals, human power, and firewood. All three of these energy sources were extremely land intensive. Both draught animals and human power rely on the production of food (feed in the case of animals) and firewood requires vast forests. In this environment, the fixed nature of land created a severe bottleneck for economic expansion. Our pre-industrial model captures this by explicitly modelling the reliance of production on land.

With the introduction of the steam engine, energy production was gradually decoupled from land use. This dramatically reduced the importance of land in production. The last two columns in Table 2 show just how enormous the increase in energy consumption was during even the relatively early phase of the industrial era. By 1850, total energy consumption had risen by over a factor of 10 relative to 1700 with virtually all of this increase coming from coal. It is clear that nothing even remotely like this would have been possible without steam power or some other dramatically less land intensive energy source.

To capture this change parsimoniously, we allow the exponents in the production function to change after the onset of the Industrial Revolution. Acemoglu and Restrepo (2018, 2019b) show how a production function that is explicit about intermediate inputs (such as energy) that are produced by completing various tasks and new technologies (such as the steam engine) change the factor content of production is equivalent to a traditional production function written in terms of primary factors only as long as the exponents on the factors are allowed to vary with technical progress. We take this approach and assume for the early industrial period that

$$Y_t = F_t(Z, K_t, L_t) = A_t Z^{\alpha_t} K^{\beta_t} L_t^{1-\alpha_t-\beta_t}, \tag{12}$$

which is the same as before except that the exponents α_t and β_t are now time varying. A fall in α_t will then capture the fall in the importance of land as a factor of production, and the evolution of β_t will determine to what extent it is capital or labor that increases in importance. We do not model the use of fossil fuels explicitly. Their use is reflected in the changing output elasticities and in higher productivity.

In this case, the labor demand curve—equation (8)—becomes:

$$w_t = \phi_t + \frac{1}{1-\beta_t}a_t - \frac{\alpha_t}{1-\beta_t}(d_t + n_t) - \frac{\beta_t}{1-\beta_t}\log\left(r_t + \delta\right), \tag{13}$$

where:

$$\phi_t = \frac{\beta_t}{1-\beta_t}\log\beta_t + \log\left(1 - \alpha_t - \beta_t\right) + \frac{\alpha_t}{1-\beta_t}\log Z - (\alpha_t + \beta_t)\lambda.$$

The fact that α_t and β_t can change implies that more data is needed to identify the model: an increase in wages for a given level of labor supply could be due to an increase in a_t or a fall in α_t. To address this issue, we make use of the demand curves for land and capital and data on the quantity of capital and the price of land after 1760. We assume that potential renters and owners of land will trade until the rental price of land equals its marginal product:

$$S_t = \alpha_t A_t Z^{\alpha_t - 1} K^{\beta_t} L_t^{1-\alpha_t-\beta_t} \tag{14}$$

where S_t denotes the rental price of land. Capital demand is the same as before—equation (4)—except that α_t and β_t are time varying. Taking logarithms and manipulating the land, capital and labor demand curves yields:

$$s_t = w_t + n_t + d_t - \log Z + \log\alpha_t - \log(1 - \alpha_t - \beta_t) + \lambda \tag{15}$$

$$\log(r_t + \delta) = w_t + n_t + d_t - k_t + \log\beta_t - \log(1 - \alpha_t - \beta_t) + \lambda \tag{16}$$

These two extra equations pin down α_t and β_t. Simple manipulation of the demand curves for labor, land, and capital—which we spell out in detail in Appendix D—establishes three intuitive results. First, an increase in s_t (land rents) holding other variables constant implies an increase in α_t and a decrease in both β_t, and $1 - \alpha_t - \beta_t$. Second, an increase in r_t (the return on capital) holding other variables constant implies an increase in β_t and a decrease in both α_t, and $1 - \alpha_t - \beta_t$. Third, an increase in w_t (wages) holding other variables constant implies and increase in $1 - \alpha_t - \beta_t$ and a decrease in both α_t and β_t. At a mechanical level, we are adding two equations per time period—equations (15) and (16)—and two observable variables per time period—s_t and k_t.

3.3 Measuring Productivity with Structural Transformation

Allowing the exponents in the production function to change raises an important complication. In this case, A_t is no longer a natural measure of productivity.[13] Productivity is meant to capture the rate at which inputs can be converted into outputs. In settings with multiple inputs (or outputs), how to operationalize this concept is ambiguous. In some cases, such as $Y_t = A_t F(X_t)$ where X_t denotes a vector of inputs, all reasonable measures of productivity agree (in this case A_t). But in the more general case of $Y_t = F_t(X_t)$, this is not the case. Caves, Christensen, and Diewert (1982) introduce the notion of a Malmquist productivity index for a quite general case of production technologies, based on ideas in Malmquist (1953). This index uses the notion of the distance between the input-output vector chosen at one point in time and the technological frontier at another point in time. For example, the distance of the input-output vector that is chosen in period $t+1$ from the time t technological frontier is $D_t(X_{t+1}, Y_{t+1}) = F_{t+1}(X_{t+1})/F_t(X_{t+1})$—i.e., actual output at time $t+1$ divided by counterfactual output using the input vector of time $t+1$ but the time t production function. Analogously, the distance of the input-output vector that is chosen at time t from the time $t+1$ technological frontier is $D_{t+1}(X_t, Y_t) = F_t(X_t)/F_{t+1}(X_t)$.

Caves, Christensen, and Diewert (1982) recommend using a geometric average of $D_t(X_{t+1}, Y_{t+1})$ and $D_{t+1}(X_t, Y_t)^{-1}$ as a Malmquist index of productivity:

$$\frac{M_t}{M_{t-1}} = \sqrt{\frac{F_t(Z, K_t, L_t) F_t(Z, K_{t-1}, L_{t-1})}{F_{t-1}(Z, K_t, L_t) F_{t-1}(Z, K_{t-1}, L_{t-1})}}, \qquad M_0 = 1. \tag{17}$$

We adopt this recommendation. With constant exponents α and β, the growth rates of M_t and A_t are the same (i.e., $M_t/M_{t-1} = A_t/A_{t-1}$). More generally, however, the growth rate of M_t will differ from the growth rate of A_t in important ways. See Appendix E for a more detailed discussion of the Malmquist index.

With the production function (12), the logarithm of the Malmquist index is

$$\hat{m}_t = \hat{a}_t + \hat{\alpha}_t \log Z + \hat{\beta}_t \bar{k}_t - (\hat{\alpha}_t + \hat{\beta}_t)(\bar{d}_t + \bar{n}_t + \lambda), \tag{18}$$

[13]One simple way to see this is to consider a change in the units that we use to express labor. Suppose $\ddot{L}_t \equiv \psi L_t$, then: $Y_t = A_t Z^{\alpha_t} K_t^{\beta_t} L_t^{1-\alpha_t-\beta_t} = (A_t/(\psi^{1-\alpha_t-\beta_t})) Z^{\alpha_t} K_t^{\beta_t} \ddot{L}_t^{1-\alpha_t-\beta_t}$. With the new units for labor, it is $A_t/\psi^{1-\alpha_t-\beta_t}$ rather than A_t that multiplies the factors of production. If α_t and β_t change over time, $A_t/\psi^{1-\alpha_t-\beta_t}$ will behave differently from A_t. Clearly, a more general concept of productivity is needed. See Appendix E for a discussion of how the Malmquist index (introduced below) avoids this issue.

where hats denote deviations from the previous period, $\hat{x}_t = x_t - x_{t-1}$, and bars denote the average of period $t-1$ and period t, $\bar{x}_t = (x_{t-1} + x_t)/2$. Once again, if α_t and β_t are constant, this expression just collapses to: $\hat{m}_t = \hat{a}_t$.

As in the pre-industrial era, we assume that the logarithm of productivity (m_t) is subject to permanent and transitory shocks in the early-industrial era:

$$m_t = \tilde{m}_t + \epsilon_{2t}, \tag{19}$$

where

$$\tilde{m}_t = \mu + \tilde{m}_{t-1} + \epsilon_{1t}, \tag{20}$$

$\epsilon_{1t} \sim \mathcal{N}(0, \sigma^2_{\epsilon_1})$, $\epsilon_{2t} \sim \mathcal{N}(0, \sigma^2_{\epsilon_2})$, and ϵ_{1t} and ϵ_{2t} are independently distributed over time. Here, $\tilde{m}_t$ is the permanent component of productivity, which follows a random walk with drift, while ϵ_{2t} is the transitory component of productivity.

3.4 Data on Interest Rates, Land Rents, Capital, and Days Worked

For the pre-industrial period, we use data on rates of return to estimate the evolution of the capital stock. Figure 8 plots our data on rates of return on agricultural land and "rent charges" compiled by Clark (2002, 2010). The rate of return on agricultural land is measured as R/P, where R is the rent and P is the price of a piece of land. "Rent charges" should not be confused with land rents. Rent charges were yields on perpetual nominal debt obligations secured by land or buildings (i.e., a collateralized loan). These are also measured as R/P, where R is the annual payment and P is the price of the obligation (which was usually much smaller than the value of the collateral). See Clark (2010) for more detail.

We view our series on rates of return of agricultural land and rent charges as two noisy measures of the rate of return on capital in England over our sample period. In other words, we assume that $r_t = \tilde{r}_{it} + \iota^r_{it}$, where r_t denotes the true rate of return on capital at time t, $\tilde{r}_{it}$ denotes noisy measure i, and $\iota^r_{it} \sim t_{\nu_{ir}}(0, \tilde{\sigma}^2_{ir})$ denotes measurement error. In periods when neither measure is available, we assume that the interest rate follows a random walk with truncated normal innovations: $r_t \sim \mathcal{N}_{(0,.2)}(r_{t-1}, 0.01^2)$.

We date the shift from the pre-industrial period to the early industrial period to between 1760 and 1770. This is a traditional date for the beginning of the Industrial Revolution. Also, estimating a changing production function—changes in α_t and β_t—requires systematic data on rents and

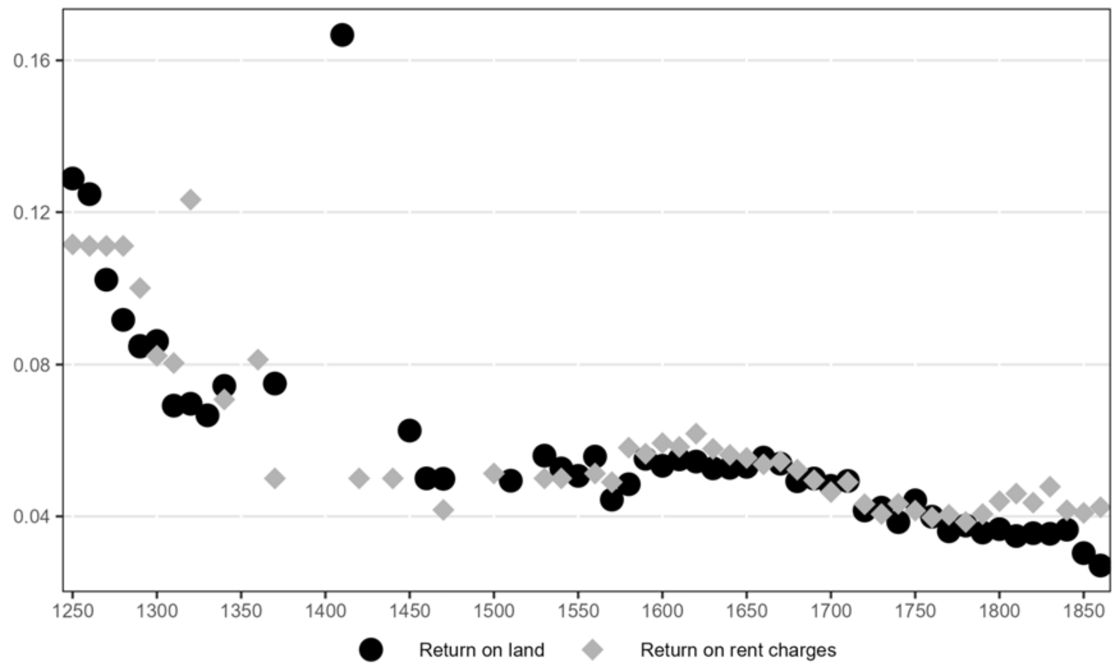

Figure 8: Rates of Return on Land and Rent Charges

Note: The figure plots the data we use on rates of return on land and rent charges. These data are from Clark (2002, 2010).

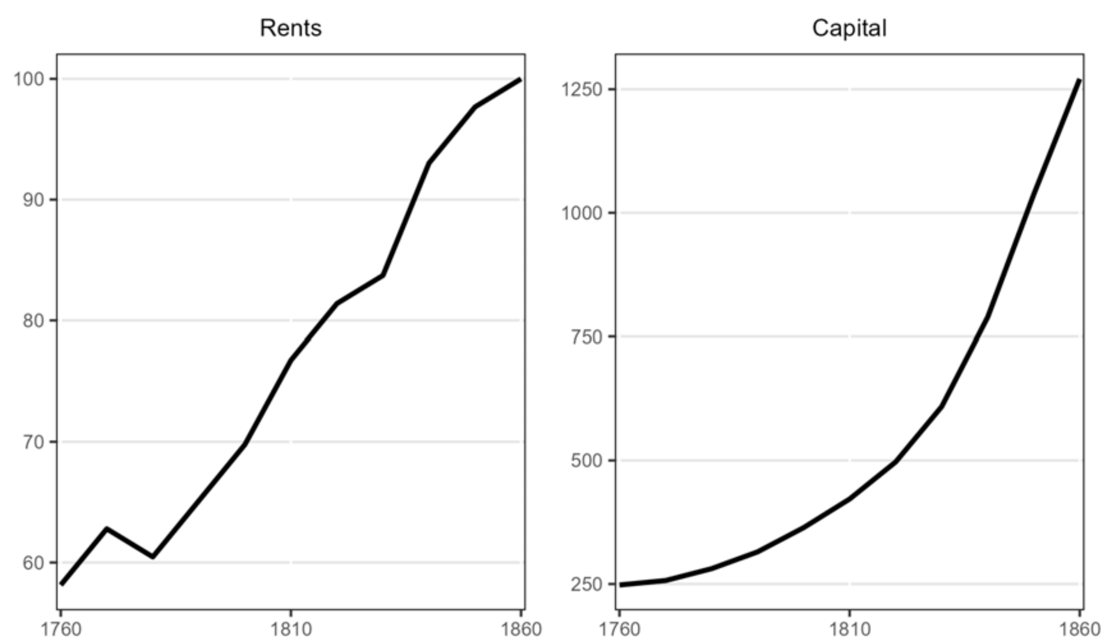

Figure 9: Land Rents and Capital after 1760

Note: The figure plots the data we use on land rents from (Clark, 2002, 2010) and the net capital stock from (Feinstein, 1988, table VIII).

the capital stock. The capital stock series we use is from Feinstein (1988) and is only available after 1760. Figure 9 plots the data on land rents and the capital stock that we use. The land rents series we use is an index from Clark (2002, 2010). Feinstein's series is for the net capital stock and is expressed in millions of pounds in 1851-1860 prices. It reflects both industrial and agricultural investment. We assume that both of these variables are observed with measurement error $s_t = \tilde{s}_t + \iota_t^s$ and $k_t = \tilde{k}_t + \iota_t^k$, where s_t and k_t denote the true land rent and capital stock, respectively, $\tilde{s}_t$ and $\tilde{k}_t$ denote our noisy measures of land rents and the capital stock, respectively, and $\iota_t^s \sim t_{\nu_s}(0, \tilde{\sigma}_s^2)$ and $\iota_t^k \sim t_{\nu_k}(0, \tilde{\sigma}_k^2)$ denote the measurement error in these two variables.

3.5 Priors and Estimation Details

We use Bayesian methods to estimate our model. In particular, we use a Hamiltonian Monte Carlo sampling procedure (Gelman et al., 2013, Betancourt, 2018).[14] Table 3 lists the priors we assume for the model parameters. In all cases, we choose highly dispersed priors. Most of the priors are self-explanatory. But some comments are in order. Our baseline estimation for α is based on movements in real wages and the population between 1340 and 1360 as discussed in section 2. However, we also present results where α is estimated structurally from our full model. The prior listed for α in Table 3 is for this case.

The prior for ψ is set such that the peak population before the Black Death is between 4.5 and 6 million with 95% probability. This range encompasses the estimates of Clark (2007b) and Broadberry et al. (2015). Rather than specifying priors for β_1 and β_2, we specify priors for the mean of ξ_1 which we denote $\mu_{\xi_1} = \beta_1/(\beta_1 + \beta_2)$ and the pseudo sample size of ξ_1 which we denote $\nu_{\xi_1} = \beta_1 + \beta_2$. The priors we choose for these parameters follow the recommendations of Gelman et al. (2013, p. 110) for a flat prior for a beta distribution. Figure A.3 plots the prior densities for the standard deviations of ϵ_1, ϵ_2, and ξ_2. In section 4.3, we discuss how varying our priors affects our main results.

To discipline the behavior of α_t and β_t after 1760, we assume that the simplex $(\alpha_t, \beta_t, 1-\alpha_t-\beta_t)$ follows a Dirichlet distribution with concentration vector $c_s \times (\alpha_{t-1}, \beta_{t-1}, 1 - \alpha_{t-1} - \beta_{t-1})$, where $c_s = 3$. For any value of c_s, this implies that the mean of α_t is α_{t-1} and the mean of β_t is β_{t-1}. The choice $c_s = 3$ implies that, with $\alpha_{t-1} = \beta_{t-1} = 1/3$, the distribution is uniform over simplexes. A smaller value of c_s would concentrate the prior distribution towards the corners of the simplex — draws where one of the coefficients is close to 1 and the others close to 0. A larger value, on

[14] We implement this procedure using Stan (Stan Development Team, 2017).

Table 3: Priors for Model Parameters

Parameter	Prior	Parameter	Prior
α	$\mathcal{U}(0,2)$	γ	$\mathcal{U}(-2,2)$
φ^x	$\mathcal{N}(0,100^2)$	ψ	$\mathcal{N}(10.86,0.07^2)$
ω	$\mathcal{N}(0,1)$	μ	$\mathcal{N}(0,1)$
μ_{ξ_1}	$\mathcal{U}(0.5,0.9)$	ν_{ξ_1}	$\mathcal{P}_I(0.1,1.5)$
π	$\mathcal{U}(0,0.5)$	δ	$\mathcal{N}_{(0,0.2)}(0.1,0.05^2)$
$\sigma^2_{\epsilon_1}$	$\text{I}\Gamma(3,0.001)$	$\sigma^2_{\epsilon_2}$	$\text{I}\Gamma(3,0.005)$
$\sigma^2_{\xi_2}$	$\text{I}\Gamma(3,0.005)$	σ^2_n	$\text{I}\Gamma(3,0.005)$
σ^2_d	$\text{I}\Gamma(3,0.005)$	σ^2_{ri}	$\text{I}\Gamma(3,0.005)$
$\tilde{\sigma}^2_d$	$\text{I}\Gamma(3,0.005)$	$\tilde{\sigma}^2_s$	$\text{I}\Gamma(3,0.005)$
$\tilde{\sigma}^2_k$	$\text{I}\Gamma(3,0.005)$	ν_n^{-1}	$\mathcal{U}(0,1)$
ν_d^{-1}	$\mathcal{U}(0,1)$	ν_s^{-1}	$\mathcal{U}(0,1)$
ν_k^{-1}	$\mathcal{U}(0,1)$	ν_{ir}^{-1}	$\mathcal{U}(0,1)$

the other hand, concentrates the prior towards the mean of the distribution — most draws would be close to $(\alpha_{t-1}, \beta_{t-1}, 1-\alpha_{t-1}-\beta_{t-1})$. Thus, our prior choice is a way to center the distribution around the previous value of the coefficients, while allowing them to change if the likelihood dictates it.

We allow for a structural break in the probability of a plague π in 1680. The timing of this break is chosen to immediately follow the Great London Plague of 1665.[15] This break is meant to capture the fact that plagues are less frequent in the latter part of our sample. The exact timing of this break does not affect our main results in a material way.

Finally, we need to normalize some of the variables in our model. Labor demand—equation (13)—features $\log Z$ and λ, which we normalize to 0. The observed wage, land rents, and our capital series are indices. They are thus not normalized in a theoretically consistent fashion. For wages, for instance, we would want to observe the marginal product of an additional unit of labor expressed in units of aggregate output. What we observe is an index that is normalized to 100 in 1860. Our Bayesian framework allows us to handle this issue in a straightforward fashion. For wages we assume that $w_t = \varphi^w + \tilde{w}_t$, where $\tilde{w}_t$ is the observed wage, w_t is the true wage, and $\varphi^w \sim \mathcal{N}(0,100^2)$ is a normalization constant. For land rents and the capital stock, we add the following normalization constants to their measurement equations: $\varphi^s \sim \mathcal{N}(0,100^2)$ and $\varphi^k \sim$

[15]Notice that the change in the population between the 1660s and the 1670s is affected by the Great London Plague. So, ξ_{1t} for $t = 1670$ will be affected by the Great London Plague. This is why we assume that ξ_{1t} for $t \geq 1680$ is governed by a different π than earlier values of ξ_{1t}.

$\mathcal{N}(0, 100^2)$, respectively. We reproduce the equations and distributional assumptions of our full model in Appendix F for convenience.

4 Estimates of Productivity from the Full Malthusian Model

We present results on the evolution of productivity in England for several variants of the model presented in section 3. We first consider a case in which we maintain our earlier assumptions that the importance of land in production (α) is constant, days worked are constant, and that the slope of the labor demand curve is pinned down by movements in wages and the population at the time of the Black Death. This case differs from the simple model presented in section 2 in that we have added capital and the Malthusian model of population dynamics. We refer to this case as the 'constant α, β' case.

Figure 10 presents results on the evolution of productivity over time for the constant α, β (gray solid line). We also include results for the simple model from section 2 for comparison (black solid line). While the overall pattern is similar, the constant α, β case of the full model implies slower productivity growth than the simple model. The reason for this is that the full model incorporates capital accumulation. As a result, a portion of the increase in labor demand over our sample period is attributed to capital accumulation as opposed to productivity. Less is left in the residual to be attributed to productivity growth in the full model than in the simple model.[16]

4.1 Productivity and the Falling Importance of Land

The models we have discussed above make the standard, but highly unrealistic, assumption that the elasticity of output with respect to land α is constant through the early industrial period up to 1870. We now relax this assumption by allowing the production function parameters α and β to change over time after the onset of the Industrial Revolution. As we describe in section 3, we use data on the quantity of capital and the price of land after 1760 to pin down the evolution of α_t and β_t during the early industrial period.

Figure 11 plots our estimates of the evolution of α_t and β_t over time. Recall that α_t corresponds to the elasticity of output with respect to land, while β_t corresponds to the elasticity of output with respect to capital. We estimate a value for α prior to the Industrial Revolution of 0.54 (0.05), and

[16]Figure A.2 presents our estimate of the timing of the pre-Industrial structural break in productivity growth in the constant α, β case, our baseline case discussed below, as well as several other variants of the model. All these cases strongly favor a break in 1600.

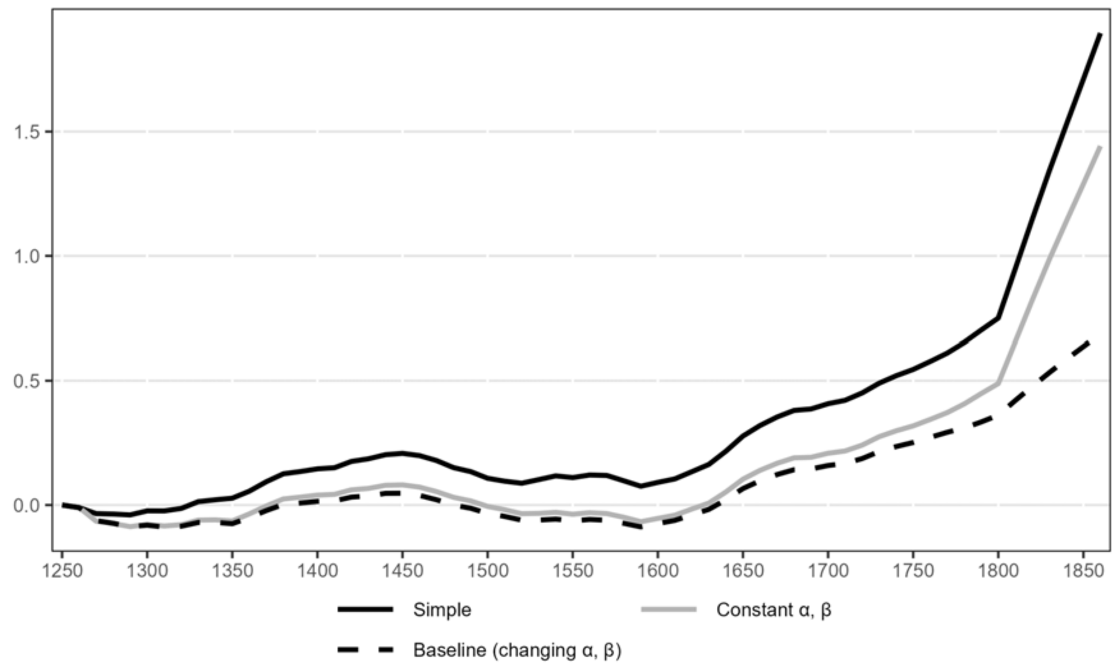

Figure 10: Permanent Component of Productivity

Note: The figure plots our estimates of the evolution of the permanent component of productivity $\tilde{a}_t$ over our sample period for several variants of our full Malthusian model (natural logarithm relative to its value in 1250).

a value for β of 0.23 (0.07).[17] After 1760, we estimate a sharp fall in α_t. By 1860, α_t had fallen by roughly half to a value of 0.29. This sharp fall in α_t reflects the nature of technical change during the Industrial Revolution. As we discuss in section 3, the advent of the steam engine powered by fossil fuels meant that the production of energy was no longer land intensive. This led to a large fall in the importance of land in production. As α_t falls, both β_t and $1 - \alpha_t - \beta_t$ (the elasticity of output with respect to labor) rise. Between 1760 and 1860, β_t increased from 0.23 to 0.35, while $1 - \alpha_t - \beta_t$ increased from 0.23 to 0.36.[18]

Allowing for a fall in the importance of land in production after 1760 has a substantial effect on our estimate of productivity growth after that date. The dashed black line in Figure 10 plots our estimates of productivity growth in this case, which we refer to as our baseline case. Productivity growth up until 1760 is very similar to the constant α, β case, but after 1760 it is much slower. In the constant α, β case, our estimate of average productivity growth between 1800 and 1870 is 16% per decade (Table 4). When we allow α_t and β_t to change in the early industrial period, our

[17] As we discuss in section 2, the land share of production differs from α when the elasticity of substitution between land and labor differs from one. In particular, if land and labor are complements in production the land share is lower than α (see footnote 7). Explicitly, allowing for a production function that is more general than the Cobb-Douglas production function is complicated when the structure of the production function is changing as it is for us after 1760.

[18] Estimate of μ and γ are discussed below. Estimates of other model parameters are presented in Table A.1.

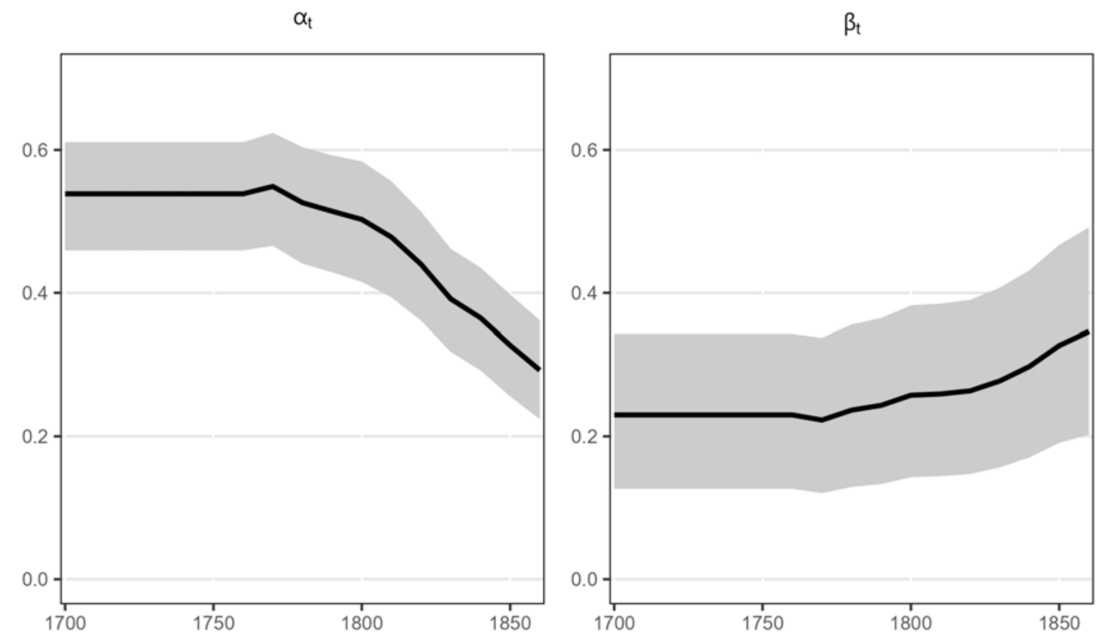

Figure 11: Parameters of the Production Function

Note: The figure plots our estimates of the evolution of the parameters α_t and β_t in our baseline case. They are constant until 1770 by assumption. The black line is the mean of the posterior for each period and the gray shaded area is the 90% central posterior interval.

estimate of average productivity growth falls to a more modest level of 5% per decade over this period. This baseline case is the specification used in Figures 1 and 2 of the introduction. Also, the numerical estimates discussed in the abstract and introduction are for this case.

Intuitively, in the constant α, β case with a large α, the marginal product of labor is sharply downward sloping. After 1760, England experienced explosive growth in its population (Figure 4). This led to strong downward pressure on real wages. Since real wages actually rose over this period, large increases in productivity are needed to fit the data. When α is allowed to fall after 1760, the labor demand curve becomes less downward sloping, which implies that the downward pressure on real wages from population growth is smaller and less productivity growth is therefore needed to explain the increase in real wages.

The model with falling importance of land after 1760 attributes a substantial portion of the explosive economic growth of the early industrial period not to productivity growth but to structural change. Prior to the Industrial Revolution, land was a severe bottleneck for economic growth. The advent of the coal-powered steam engine changed this dramatically, freeing the economy to grow more rapidly. Whether one wants to view this change as an increase in productivity (as we do in the constant α, β case) or as structural change is a matter of how detailed a model one considers.

Table 4: Productivity Growth

	Mean	St Dev	2.5%	97.5%
		Simple		
$\mu_{a,1}$	0.00	0.01	-0.01	0.02
$\mu_{a,2}$	0.04	0.02	0.02	0.09
$\mu_{a,3}$	0.19	0.01	0.17	0.22
		Constant α, β		
$\mu_{a,1}$	-0.00	0.01	-0.01	0.01
$\mu_{a,2}$	0.03	0.01	0.01	0.05
$\mu_{a,3}$	0.16	0.02	0.11	0.20
		Baseline		
$\mu_{a,1}$	-0.00	0.01	-0.01	0.01
$\mu_{a,2}$	0.02	0.01	0.01	0.04
$\mu_{a,3}$	0.05	0.01	0.03	0.08

Note: The table presents the mean, standard deviation, 2.5% quantile, and 97.5% quantile of the posterior distribution we estimate for average productivity growth μ for three cases of our model. See table A.1 for estimate of the posterior distribution of other model parameters.

Productivity captures both forces in the simple model. In the more detailed model of this section, we provide an explanation for a large chunk of this variation in the early industrial period, through the lens of structural change.

The data we use on the capital stock and land rents starts in 1760. We view these variables as being unobserved prior to that time. Figure A.4 plots what our model implies about their evolution over our entire sample period. We infer that capital grew at a modest pace prior to 1600, with faster growth thereafter. We infer relatively stable land rents prior to 1600 with growth thereafter.

4.2 Days Worked and Structural Identification of the Slope of Labor Demand

The results we have presented up to this point have made the assumption that labor supply was proportional to the population of England over our sample. This is a common assumption in the literature. However, recent work by de Vries (2008) and Humphries and Weisdorf (2019) argues that hours worked per worker fluctuated substantially over our sample. Figure 12 plots Humphries and Weisdorf's series. It indicates that days worked dropped sharply after the Black Death and then recovered to its previous level by the early 17th century. After that, days worked

Figure 12: Days Worked per Worker in England, 1260-1840

Note: The figure presents an estimate of the evolution of days worked per worker in England from Humphries and Weisdorf (2019).

kept increasing, rising far above their previous level.[19]

Figure 13 presents results on productivity for a case where we incorporate Humphries and Weisdorf's 2019 estimates of days worked into our analysis. Here, we assume that days worked are exogenous and are measured with error $d_t = \tilde{d}_t + \iota_t^d$, where d_t denotes the true number of days worked per worker, which are unobserved, $\tilde{d}_t$ denotes Humphries and Weisdorf's estimates of days worked, and $\iota_t^d \sim t_{\nu_d}(0, \tilde{\sigma}_d^2)$ denotes the measurement error.[20] Allowing for changing days worked does not affect our conclusion about the timing of the onset of productivity growth. However, we do estimate a somewhat larger increase in average productivity growth after 1600 in this case. Intuitively, more productivity is needed to compensate for the additional labor supply associated with increased days of work. With Humphries and Weisdorf's estimates of variable work hours, we estimate that average productivity growth μ is 3% per decade for the period 1600-

[19]Other researchers have referred to the increase in worker industriousness over this period as an "Industrious Revolution" (de Vries, 1994, 2008). However, the degree to which days worked changed over time in England is controversial. Comparisons of direct estimates by Blanchard (1978) for 1400-1600 and Voth (2000, 2001) for 1760-1830 support the idea that days worked were low in the post-Black Death period and rose sharply in the 17th and 18th centuries. Earlier indirect estimates by Clark and Van Der Werf (1998), however, suggest modest changes in days worked over our sample. Humphries and Weisdorf (2019) argue that their new series on the income of workers on annual contracts represents an important improvement relative to the series used by Clark and Van Der Werf (1998).

[20]Humphries and Weisdorf do not provide estimates for 1250, 1850, and 1860. We extrapolate days worked on these dates assuming that $d_t = d_{t-1} + \eta_t$ where $\eta_t \sim \mathcal{N}(0, \sigma_d^2)$.

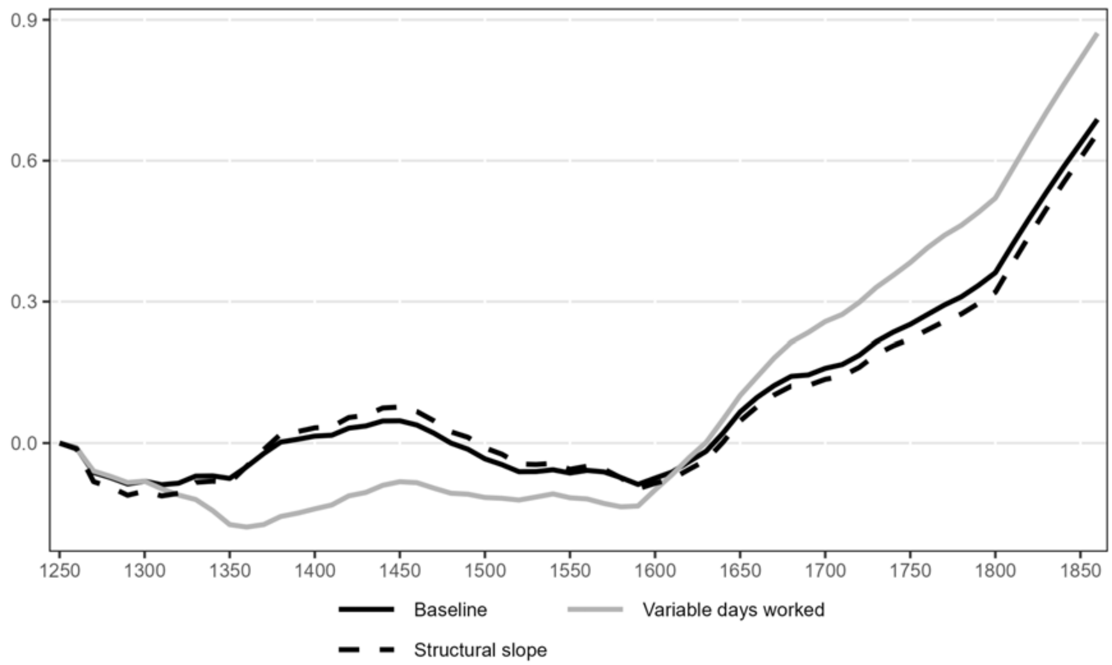

Figure 13: Permanent Component of Productivity

Note: The figure plots our estimates of the evolution of the permanent component of productivity $\tilde{a}_t$ over our sample period for several variants of our full Malthusian model (natural logarithm relative to its value in 1250).

1800 and 6% per decade for the period 1810-1870.

We have used the Black Death as an exogenous shock to identify the slope of the labor demand curve in the pre-industrial era. We also consider the alternative approach of using the structure of the full model to identify this slope. In this case, we specify a relatively diffuse prior for both α and β and use Bayesian updating to calculate a posterior mean for these and for the slope of the labor demand curve conditional on the entire sample, as we do for other parameters. We are then relying on the Malthusian model of population dynamics to account for the endogeneity of the population. Figure 13 presents results on productivity for this case. The results are very similar to the baseline case.

4.3 Robustness

Alternative Real Wage Data: Our results up to this point use real wage data for unskilled builders from Clark (2010). Figure A.5 presents five alternative productivity series where we instead use other wage series. First, we present estimates of productivity using the following day wage series: 1) Clark's (2010) real wages series for farm laborers, 2) Clark's (2010) real wages series for building craftsmen, 3) Allen's (2007) real wage series for the period 1770 onward (with our baseline wage series before that time). We also present estimates of productivity based on the assumption that the builders, farmers, and craftsmen series from Clark (2010) are all noisy signals of the underlying true wage. Finally, we use Humphries and Weisdorf's (2019) annual income series along with the assumption that days worked are constant. We do this robustness analysis for the structural slope case. These alternative productivity series yield similar results to baseline productivity series, although there is some divergence early in the sample period.

Population Data: Our results up to this point use population data prior to 1540 from Clark (2010). Figure A.6 presents estimates of productivity using population data from Broadberry et al. (2015) for the period prior to 1540. Broadberry et al.'s (2015) estimates of the population are infrequent and irregular in their frequency. There are quite a few decades for which Broadberry et al. (2015) have no estimate, e.g., they present no estimate between 1450 and 1522. In this robustness analysis, we view the population as an unobserved variable in decades for which we do not have an estimate from Broadberry et al. (2015). Our results on the evolution of productivity for this case are very similar to our baseline model.

Break Dates: Figure A.7 compares the evolution of the permanent component of productivity for four different assumptions about when the pre-Industrial Revolution break to the productivity component occurred. Recall that Figure 10 presents the evolution of the permanent component of productivity integrating over the probability of breaks occurring at different dates as in Figure A.2. Figure A.7 compares these results with results when we condition on the break occurring at a specific date. We do this for the baseline case and present results for several different dates between 1550 and 1760. These alternative results are very similar.

Priors: Figure A.8 presents estimates of productivity using different prior distributions than we use in our main analysis. First, we present results for a case where we change the prior on σ_{ϵ_1}—the variance of permanent productivity shocks—to be $\text{I}\Gamma(3, 0.005)$, i.e., the same as the prior on the other productivity and population shocks. Second, we present results for a case where we change the prior on ψ—the level of the population prior to 1540—to be $\mathcal{N}(10.86, 10^2)$, i.e., much wider than in our main analysis. In both cases, the resulting productivity series are very similar to our main results. Other priors are quite dispersed.

Comparison with Clark (2010, 2016): Our estimates of productivity differ substantially from those of Clark (2010, 2016). Appendix G presents a detailed decomposition of the factors leading to the differences. Several factors are important. One important contributor to the difference between our series and Clark's (better known) 2010 series is an error in that series that leads to a spurious 25 log point drop between 1540 and 1550. Using the average of factor output elasticities at time t and $t-1$ when calculating changes in productivity between time t and $t-1$ also explains an important part of the difference in our results relative to Clark's, especially for the period prior to 1600. Differences in the factor prices and factor output elasticities implied by our approach, relative to those used by Clark, explain the remaining differences.

5 Liberating the Economy from the Iron Law of Wages

In section 4, we estimate a gradual increase in productivity growth μ and a sharp fall in the importance of land in production α after the onset of industrialization. Here we discuss how both of these developments—as well as our small estimate of the elasticity of population growth to real income γ (see discussion below)—contributed to liberating the economy from the Malthusian "iron law of wages," i.e., the notion that wages tend to a very low (subsistence) level. These estimates also reconcile the Malthusian model with episodes historians have identified prior to the Industrial Revolution when some parts of the world have experienced substantial economic growth over a sustained period of time, sometimes several hundred years. Goldstone (2002) refers to these episodes as efflorescences. They include ancient Greece, ancient Rome, Song China, the Islamic golden age, and the golden age of Holland, to name but a few. These episodes have often been used as evidence against the Malthusian model (e.g., Persson, 2008).

It is important to recognize that steady positive productivity growth in a Malthusian model like ours results in a persistent force pushing wages higher. As wages rise, a counteracting force comes into play pushing wages lower (population growth). The strength of the force pushing wages lower is increasing in the level of the real wage. This implies that as wages rise the downward force gets stronger and stronger and eventually chokes off further increases in wages. In other words, there is a steady state real wage for each level of average productivity growth in a Malthusian model. This steady state is not at subsistence. Rather, the steady state real wage is increasing in average productivity growth.

In appendix H, we show that the steady state wage in our Malthusian model is given by

$$\bar{w} = \frac{\mu}{\alpha\gamma} + \text{constant}, \tag{21}$$

As we discussed above, faster productivity growth μ results in a stronger force pushing wages up and therefore a higher steady state wage. The strength of the counteracting force—which we call

Table 5: Estimates of γ

	Mean	St Dev	2.5%	97.5%
Baseline (constant days worked)	0.03	0.05	-0.06	0.12
Variable days worked	0.09	0.02	0.05	0.14

Note: The table presents the mean, standard deviation, 2.5% quantile, and 97.5% quantile of the posterior distribution we estimate for the elasticity of population growth to income γ. The first row presents results assuming constant days worked (our baseline case), while the second row presents results assuming days worked vary as in Humphries and Weisdorf (2019).

the Malthusian population force—is governed by two parameters in our model: the importance of land in production α and the elasticity of population growth with respect to per capita income γ. Intuitively, α determines the slope of the long-run labor demand curve, i.e., how rapidly real wages fall as population rises, while γ determines how rapidly the population increases when real wages are high.

With zero productivity growth, the steady state real wage is potentially very low. Its level depends on factors outside of the scope of our analysis such as hygiene, the marriage rate, contraceptive technology, and the level of violence in society. With positive productivity growth, the steady state real wage can be much higher. Whether it is depends on the size of α and γ. We discussed our estimates of α in section 4. We next turn to our estimates of γ.

5.1 The Elasticity of Population Growth with Respect to Income

Table 5 presents our estimate of the elasticity of population growth with respect to real income γ for our baseline case and the case with variable days worked. In our baseline case, our estimate of γ is extremely small at 0.03. This means that a 100 log point increase in real wages increases population growth per decade by only 3 log points. Between 1270 and 1440, real wages in England rose by 161% (or 96 log points). So, this increase in real wages stimulated population growth by a mere 3 log points per decade.

Table 5 also presents results on γ for our variable days worked case. In this case, we estimate a somewhat larger γ of 0.09. The lower bound of the credible interval in this case exceeds zero. While this estimate is somewhat larger, it remains very small. Allowing for changes in days worked, real incomes in England rose by 70% between 1270 and 1440. A γ of 0.09 implies that this stimulated population growth by 5 percentage points per decade, while a doubling of real income would stimulate population growth by 6 percentage points per decade.

Another way to gauge the quantitative magnitude of our estimates of γ is to calculate the half-life of population dynamics after an exogenous shock to the level of the population, e.g., due to a plague. Assuming for simplicity that days worked and the return on capital are constant, that all shocks are equal to their average value, and that there's no productivity growth ($\mu = 0$), we show in appendix H that the dynamics of the population after an initial disturbance are given by the following AR(1) process:

$$n_{t+1} = \left(1 - \frac{\gamma\alpha}{1-\beta}\right) n_t + \text{constant}. \tag{22}$$

The speed of population recovery after a plague-induced decrease is, thus, governed by $1 - \gamma\alpha/(1-\beta)$ in this case. In particular, the half-life of the population dynamics, i.e., the time it takes the population to recover half of the way back to steady state after a plague-induced drop, is $\log 0.5/\log(1 - \alpha\gamma/(1-\beta))$. (The half-life of real wage dynamics is the same.)

Plugging our estimates of the parameters γ, α, and β into the formula above, we find that the half-life of population and real wage dynamics prior to the Industrial Revolution for our baseline case is 354 years. For the variable days case, the half-life is 116 years. Since α falls sharply after 1760 and β changed little, the Malthusian population force becomes even weaker after 1760. By 1860, the half-life of population and real wage dynamics have risen to 556 years in the baseline case and 241 years with variable days. These long half-lives imply that the strength of the Malthusian population force was weak in England over our sample period—a result that is sometimes referred to as "weak homeostasis." Prior work has also found weak homeostasis. For example, Lee and Anderson (2002) find a half-life of 107 years, while Crafts and Mills (2009) find a half-life of 431 years.

5.2 Prosperity and the Slope of the Long-Run Labor Demand Curve

We can now use equation (21) and our parameter estimates to assess how changes in the economy after the onset of growth in 1600 affected the long-run steady state real wage the economy was tending towards. Figure 14 plots the steady state wage in our Malthusian economy for different values of productivity growth μ and the slope of the long-run labor demand curve α relative to the steady state wage with zero productivity growth. Each line in the figure gives the steady state wage for a particular value of productivity growth as the value of α varies. We do this analysis for the variable days case, i.e., using $\gamma = 0.09$.

The figure illustrates clearly how important the fall in α is for liberating the economy from the iron law of wages. Consider first how steady state real wages respond to productivity growth when α is equal to our pre-industrial estimate of 0.48. In this case, productivity growth of 3% per decade raises the real wage by a factor of 2 in the long run, while 6% productivity growth

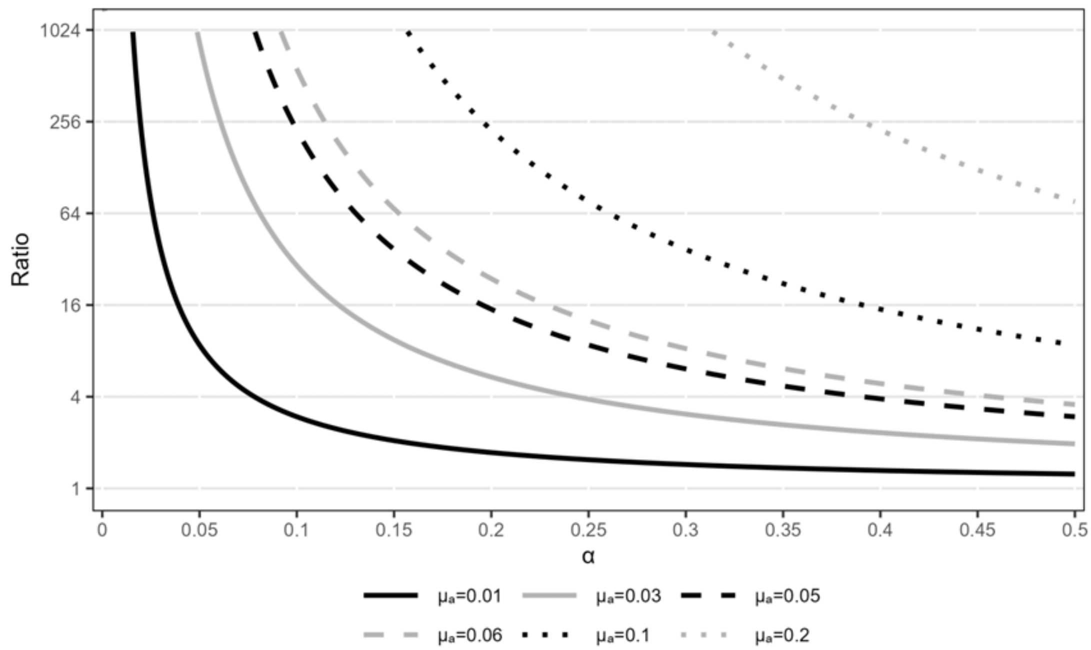

Figure 14: Steady State Wage with Productivity Growth

Note: The figure plots steady state real wage for different average levels of productivity growth (μ) and output elasticity of land (α) relative to the steady state real wage with zero productivity growth. The parameter α varies on the x-axis and each line plots real wages for a given level of average productivity growth. These impulse responses are calculated assuming that all other model parameters are at their posterior mean values. We do this for the variable days case.

raises the real wage by a factor of 4. These results show that our Malthusian model is consistent with substantial, multi-hundred year efflorescences of the kind discussed by Goldstone (2002) if we allow for modest productivity growth. The small value we estimate for γ and the associated weak homeostasis are key to this result.

As α falls, the steady state real wage for any given level of average productivity growth rises sharply. With our estimate of α for 1860 ($\alpha = 0.22$), productivity growth of 6% per decade can raise the real wage by a factor of 19 in the long run. In other words, this level of productivity growth would have eventually led to a 19-fold increase in real wages even if the Demographic Transition had not occurred and the Malthusian population force had continued at its 1860 strength. Clearly, a flattening of the long-run labor demand curve is a powerful force for liberating the economy from the iron law of wages when productivity growth is positive.

Another way to visualize these effects is to plot the steady state wage relative to the actual wage over time. We do this in Figure 15. Before 1600, the ratio of the steady state wage to the actual wage is relatively stable around 1. Once productivity growth begins, however, the steady state wage jumps higher and the actual wage only gradually catches up. After 1760, the steady

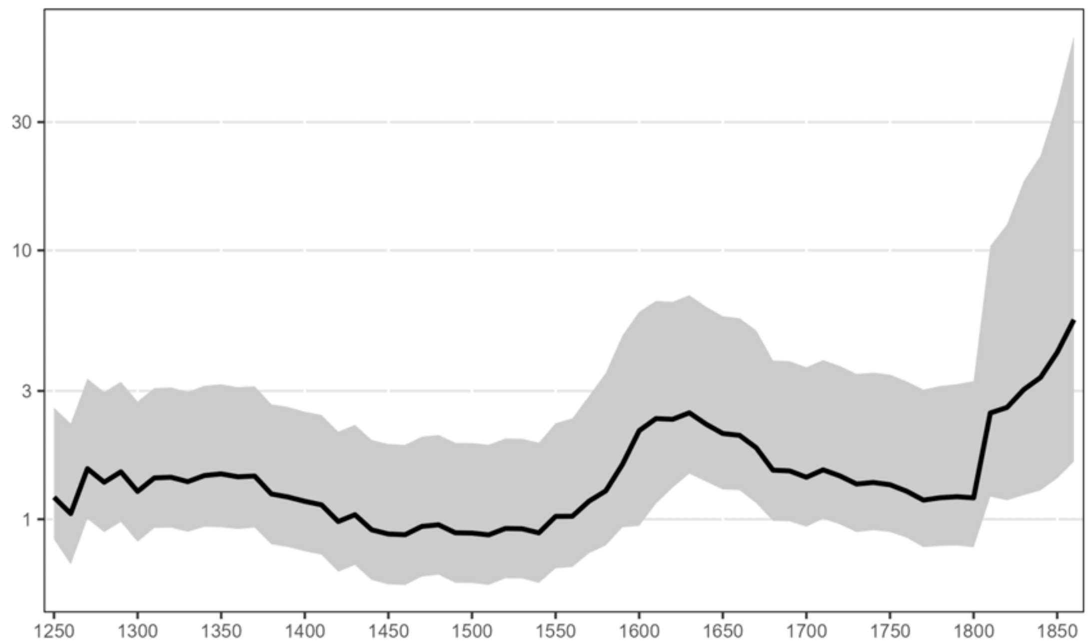

Figure 15: Steady State Wage Relative to Actual Wage

Note: The figure plots the ratio of the steady state wage—given by equation (49)—and the actual wage over time. Note that the steady state wage is a function of days worked. In the figure, we use the days worked at each date as the d^* in equation (49). The black line is the median of the posterior for each period and the gray shaded area is the 90% central posterior interval. We do this for the variable days case.

state wage begins to rise rapidly as the slope of the long-run labor demand curve flattens. By 1860, the steady state wage is more than five times higher than the actual wage.

5.3 Post-1750 Population Explosion

The modest strength of the Malthusian population force in our model begs the question whether our model can explain, with these parameter values, the large increase in the population of England that occurred after 1750 (see Figure 4). In 1740, the population of England was 6 million. By 1860, it had risen to almost 20 million. The population therefore grew at a compound rate of 10.4% per decade over this 120-year period.

Figure 16 compares the evolution of the population in England from 1750 to 1860 with the predicted evolution of the population from the variable days version of our model. We construct the predicted evolution by taking the evolution of real wages and days worked in England as given and simulating the evolution of the population using equation (9) starting from its actual value in 1740 and assuming no population shocks. This analysis shows that in fact our model with variable days worked can explain the vast majority of the rapid increase in the population between 1740 and 1860. This may seem surprising given the weak Malthusian population force

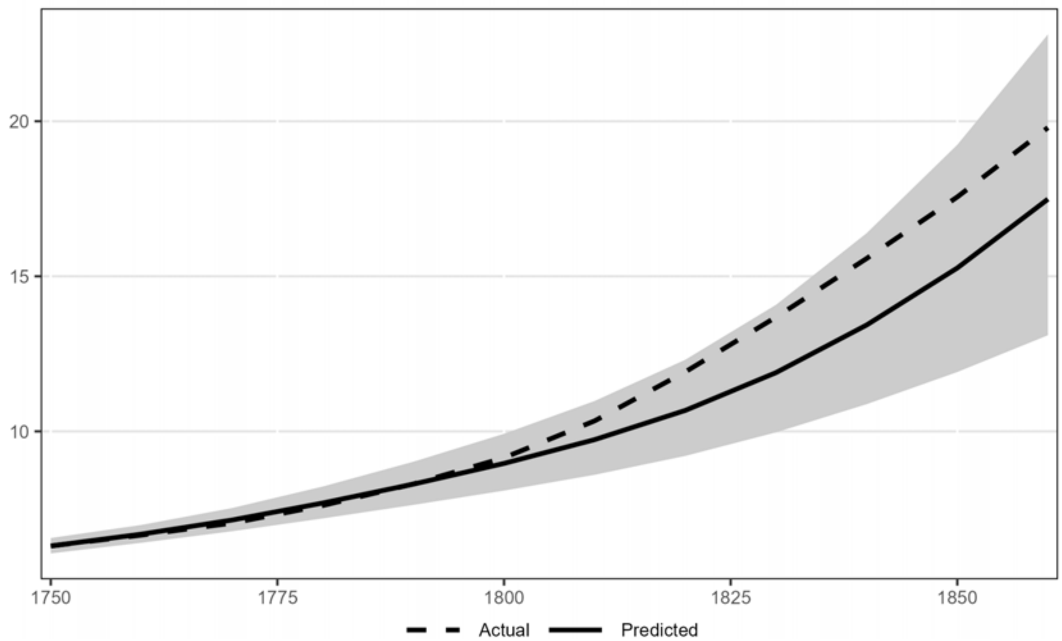

Figure 16: Actual and Predicted Population Dynamics after 1750

Note: The dashed line is the evolution of the population in England. The solid line is the predicted evolution of the population in England from our Malthusian model. In calculating this line, we take the evolution of real wages and days worked in England as given and simulate the evolution of the population using equation (9) starting from its actual value in 1740. The gray shaded area is the 90% central predictive interval given our estimates of α and γ.

and the somewhat modest increase in real wages over this period. However, allowing for variable days worked, per capita income in England over this period rose quite substantially (see Figure 12).

6 Plagues and the Population

Figure 17 plots our baseline estimate of the evolution of the population of England from 1250 to 1550 along with prior estimates from Clark (2007a), Clark (2010), and Broadberry et al. (2015). Our estimates are very similar to Clark's. This implies that our estimation procedure largely validates the assumptions Clark makes regarding the evolution of productivity in constructing his population estimates. The estimates of Broadberry et al. (2015) are substantially lower early in the sample period, but then gradually converge.

The evolution of the population in England over our sample period is heavily affected by plagues. Our model captures plagues (and other influences on the population other than changes in real income) through the shocks ξ_{1t} and ξ_{2t}. Figure A.9 plots the evolution of the sum of these population shocks over our sample period. The largest population shock by far is the Black Death

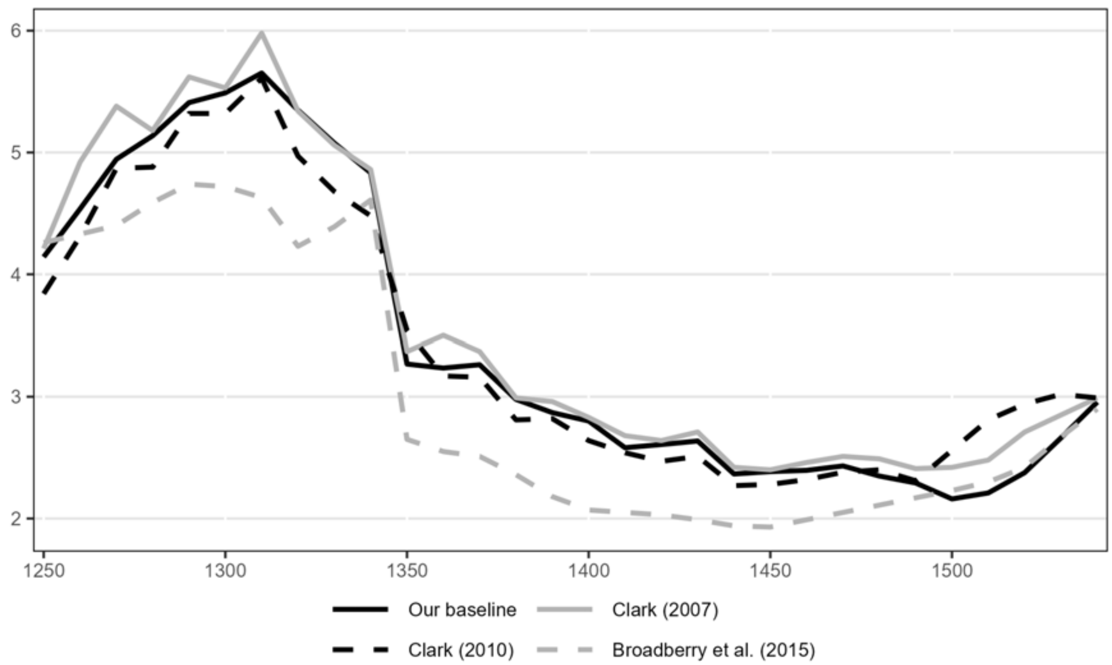

Figure 17: Comparison of Population Estimates for England

Note: The figure plots our estimates of the evolution of the population of England along with estimates from Clark (2007a), Clark (2010), and Broadberry et al. (2015)

of 1348. We estimate that the population shocks associated with the Black Death lead the population of England to shrink by 32%. But Figure A.9 also makes clear that England faced steady population headwinds—i.e., persistent negative population shocks—from the early 14th century until about 1500.

7 Conclusion

In this paper, we estimate the evolution of productivity in England from 1250 to 1870 as shifts in the labor demand curve. Our principal finding is that productivity growth began in 1600. Before 1600, productivity growth was zero. We estimate a growth rate of productivity of 2% per decade between 1600 and 1800 and an increase to 5% per decade between 1810 and 1870. Our results indicate that sustained growth in productivity began well before the Glorious Revolution and Industrial Revolution. We demonstrate that the early 17th century was a crucial turning point for productivity growth in England, a result that helps distinguish between competing lines of thought regarding the ultimate causes of the emergence of growth.

We attribute the high output growth of the Industrial Revolution only partly to productivity growth. A second important factor was the rapidly falling importance of land in production associated with the transition to steam power fueled by coal. We leverage our model to estimate the strength of the Malthusian population force in pre-Industrial England. This force was quite weak. The half-life of the response of real wages after a plague-induced decrease in the population was more than 100 years prior to the onset of the Industrial Revolution and increased to several hundred years by 1860.

A Clark's Population Series

As we discuss in the main text, Clark (2007b) uses unbalanced panel data on the population of villages and manors from manorial records and penny tithing payments to construct estimates of the population prior to 1540. Clark starts by running a regression of this data on time fixed effects and manor/village fixed effects. He refers to the time fixed effects from this regression as as a population trend series.

Clark's population trend series does not provide information on the overall level of the population prior to 1540, only changes in the population (i.e., a normalization is needed). In addition, Clark's microdata is sufficiently unreliable for the 1530s that Clark does not make use of his estimated population trend for that decade. Clark uses the following procedure to surmount these problems. First, he regresses his population trend on real wages from 1250 to 1520, and separately regresses the Wrigley et al. (1997) population series on wages from 1540 to 1610. He observers that the R^2 in both regressions are high and that they yield similar slope coefficients. He concludes from this that (i) the English economy moved along stable labor demand curves during both subsamples and (ii) these two labor demand curves had similar slopes.

Clark next makes the assumption that there was no productivity growth between 1520 and 1540—the labor demand curve did not shift during this time. This allows him to extrapolate the relationship that he finds in the post-1540 data to the earlier sample, and infer both the population in 1530 and the missing normalization from the level of real wages. Clark also uses the fitted values for the population from his labor demand curve as an alternative estimate of the population and averages this with the trend series to get what he calls the "best" estimate of population before 1540.

B CES Production Function

Consider the production function

$$Y_t = A_t \left[\alpha'^{\frac{1}{\sigma}} Z^{\frac{\sigma-1}{\sigma}} + (1-\alpha')^{\frac{1}{\sigma}} (L_t)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}},$$

where σ denotes the elasticity of substitution between land and labor in production. Optimal choice of labor by land owners gives rise to the following labor demand curve

$$W_t = (1-\alpha')^{\frac{1}{\sigma}} A_t \left[\alpha'^{\frac{1}{\sigma}} \left(\frac{Z}{L_t} \right)^{\frac{\sigma-1}{\sigma}} + (1-\alpha')^{\frac{1}{\sigma}} \right]^{\frac{1}{\sigma-1}}.$$

A log-linear approximation of this equation yields

$$w_t = \phi - \alpha l_t + a_t,$$

where

$$\alpha = \left[\sigma \left(1 + \left(\frac{1-\alpha'}{\alpha'} \right)^{\frac{1}{\sigma}} \left(\frac{L}{Z} \right)^{\frac{\sigma-1}{\sigma}} \right) \right]^{-1}$$

and L is the level of labor we linearize around. Notice that $\alpha \to \alpha'$ when $\sigma \to 1$.

It is furthermore easy to show that with the CES production function given above, the labor share of output is

$$\bar{LS} = 1 - \left[1 + \left(\frac{1-\alpha'}{\alpha'} \right)^{\frac{1}{\sigma}} \left(\frac{L}{Z} \right)^{\frac{\sigma-1}{\sigma}} \right]^{-1}.$$

Combining these last two equations, we get that

$$\alpha = \frac{1-\bar{LS}}{\sigma}.$$

This implies that the land share is $\sigma\alpha$ in this case.

C More General Production Function for Pre-Industrial Era

Consider the concave production function

$$Y_t = A_t F(Z, L_t, K_t) \tag{23}$$

The first-order conditions are

$$W_t = A_t F_L(Z, L_t, K_t)$$
$$r_t + \delta = A_t F_K(Z, L_t, K_t)$$

where δ is the depreciation rate of capital.

Taking logs in the FOC

$$w_t = a_t + \log\left(F_L(Z, L_t, K_t)\right) \approx \tilde{\phi}' + a_t + \frac{LF_{LL}}{F_L} l_t + \frac{KF_{LK}}{F_L} k_t \tag{24}$$

$$\log(r_t + \delta) = a_t + \log\left(F_K(Z, L_t, K_t)\right) \approx \tilde{\phi}'' + a_t + \frac{LF_{LK}}{F_K} l_t + \frac{KF_{KK}}{F_K} k_t \tag{25}$$

Solving for k_t in equation (25)

$$k_t = \tilde{\phi}''' + \frac{F_K}{KF_{KK}}\left(\log(r_t + \delta) - a_t\right) - \frac{LF_{LK}}{KF_{KK}} l_t \tag{26}$$

Substituting into equation (24)

$$w_t \approx \tilde{\phi} + \left(1 - \frac{F_K F_{LK}}{F_L F_{KK}}\right) a_t + \frac{L}{F_L F_{KK}}\left(F_{LL}F_{KK} - F_{LK}^2\right) l_t + \frac{F_K F_{LK}}{F_L F_{KK}} \log(r_t + \delta)$$

Which can be rewritten

$$w_t \approx \tilde{\phi} + \left(1 + \tilde{\beta}\right) a_t - \tilde{\alpha} l_t - \tilde{\beta} \log(r_t + \delta) \tag{27}$$

where

$$\tilde{\alpha} = -\frac{L}{F_L F_{KK}}\left(F_{LL}F_{KK} - F_{LK}^2\right)$$
$$\tilde{\beta} = -\frac{F_K F_{LK}}{F_L F_{KK}}$$

Equation (27) shows that a_t is identified up to a first-order approximation. This result does not require a Cobb-Douglas production function, not even constant returns to scale.

D Identification of α_t and β_t

Consider the demand curves for labor, land, and capital in the early-industrial era:

$$W_t = (1 - \alpha_t - \beta_t) A_t Z^{\alpha_t} K_t^{\beta_t} L_t^{-\alpha_t - \beta_t}, \tag{28}$$

$$S_t = \alpha_t A_t Z^{\alpha_t - 1} K^{\beta_t} L_t^{1 - \alpha_t - \beta_t}, \tag{29}$$

$$r_t + \delta = \beta_t A_t Z^{\alpha_t} K_t^{\beta_t - 1} L_t^{1 - \alpha_t - \beta_t}. \tag{30}$$

We begin by dividing land demand and capital demand by labor demand:

$$\frac{S_t}{W_t} = \frac{\alpha_t}{1-\alpha_t-\beta_t}\frac{L_t}{Z}, \tag{31}$$

$$\frac{r_t+\delta}{W_t} = \frac{\beta_t}{1-\alpha_t-\beta_t}\frac{L_t}{K_t}. \tag{32}$$

Manipulating equation (31) yields

$$\alpha_t = X_t - X_t\beta_t, \tag{33}$$

where

$$X_t = \frac{S_t/W_t}{(L_t/Z)+(S_t/W_t)}.$$

Manipulation equation (32) yields

$$\alpha_t = Y_t - Y_t\beta_t, \tag{34}$$

where

$$Y_t = \frac{(r_t+\delta)/W_t}{(L_t/K_t)+((r_t+\delta)/W_t)}.$$

Solving equations (33) and (34) for α_t and β_t yields

$$\alpha_t = X_t\frac{1-Y_t}{1-X_tY_t}, \tag{35}$$

$$\beta_t = Y_t\frac{1-X_t}{1-X_tY_t}, \tag{36}$$

and we then also have that

$$1-\alpha_t-\beta_t = \frac{(1-X_t)(1-Y_t)}{1-X_tY_t}. \tag{37}$$

Consider a case were S_t (land rents) goes up while all other variable remain constant. This increase X_t but leaves Y_t unchanged. As a consequence, α_t increases and both β_t and $1-\alpha_t-\beta_t$ decrease.[21]

Next, consider a case were r_t (rental rate of capital) goes up while all other variables remain constant. This increases Y_t but leaves X_t unchanged. As a consequence, β_t increases and both α_t and $1-\alpha_t-\beta_t$ decrease.

Finally, consider a case where W_t (wage) goes up while all other variables remain constant.

[21] The derivative of $(1-X_t)/(1-X_tY_t)$ with respect to X_t is $-(1-Y_t)/(1-X_tY_t)^2$, which is negative.

This decreases both X_t and Y_t. As a consequence, both α_t and β_t decrease and $1 - \alpha_t - \beta_t$ increases.[22]

E The Malmquist Productivity Index

The concept of productivity is meant to measure the ratio of output to inputs (Diewert and Nakamura, 2007). In situations with more than one inputs (or outputs), the exact way in which this basic concept is operationalized is ambiguous. In some special cases, all reasonable measures of productivity will agree. This is, for example, the case if production is assumed to take the following form $Y_t = A_t F(X_t)$, where Y_t denotes output and X_t denotes a vector of inputs. In this case, A_t is the natural measure of productivity. In the more general case of $Y_t = F_t(X_t)$ the definition of productivity is less clear cut.

Caves, Christensen, and Diewert (1982) introduce the notion of a Malmquist productivity index for a quite general case of production technologies, based on ideas in Malmquist (1953). The discussion below builds on the exposition of these concepts in Färe et al. (1994). Consider a production technology S_t that transforms inputs $X_t \in \mathbb{R}_+^N$ into output $Y_t \in \mathbb{R}_+$: $S_t = \{(X_t, Y_t) : X_t \text{ can produce } Y_t\}$. Written in terms of a production function $Y_t = F_t(X_t)$, we have $S_t = \{(X_t, Y_t) : Y_t \leq F_t(X_t)\}$. In other words, S_t defines the set of all feasible input-output vectors.

Caves, Christensen, and Diewert (1982) define the Malmquist productivity index in terms of the distance function $D_t(X_s, Y_s) = \inf\{\theta : (X_s, Y_s/\theta) \in S_t\}$. The distance $D_t(X_s, Y_s)$ is then the minimum multiplicative proportion by which Y_s needs to be scaled down for the input-output vector (X_s, Y_s) to be feasible with time t technology. For example, if period s is a later period than period t and technology is "more advanced" at this later period, (X_s, Y_s) may be feasible using technology S_s, but $Y_s/D_t(X_s, Y_s)$ with $D_t(X_s, Y_s) > 1$ may be the largest output that is feasible given input use X_s and the inferior technology S_t.

Given the definition of S_t, the distance is the smallest θ such that $Y_s/\theta \leq F_t(X_s)$, which means $D_t(X_s, Y_s) = Y_s/F_t(X_s)$. Under our maintained assumptions in this paper, $D_t(X_t, Y_t) = 1$, i.e., the output actually produced at time t with inputs X_t is exactly feasible. (More generally, one can imagine production at time t being inside the technical frontier at time t. In this case, $D_t(X_t, Y_t) < 1$.)

[22] The total derivative of $1 - \alpha_t - \beta_t$ with respect to W_t is: $-\left((1 - Y_t)^2 \times \partial X_t/\partial W_t + (1 - X_t)^2 \times \partial Y_t/\partial W_t\right)/(1 - X_t Y_t)^2$. Since X_t and Y_t are both decreasing in W_t, this derivative is positive. For $1 - \alpha_t - \beta_t$ to increase, α_t or β_t must decrease. Manipulating equations (31) and (32), we have: $\alpha_t/\beta_t = S_t/(r_t + \delta) \times Z/K_t$. Since the ratio of α_t over β_t is constant and at least one of them decreases, both must decrease.

Next consider $D_t(X_{t+1}, Y_{t+1})$, i.e., the distance of the input-output vector at time $t+1$ from the technical frontier at time t. Applying the definition of the distance function we have that $Y_{t+1}/D_t(X_{t+1}, Y_{t+1}) = F_t(X_{t+1})$, which implies

$$D_t(X_{t+1}, Y_{t+1}) = \frac{Y_{t+1}}{F_t(X_{t+1})} = \frac{F_{t+1}(X_{t+1})}{F_t(X_{t+1})}.$$

This is very intuitive: The distance of the time $t+1$ technology from the time t technology evaluated at the time $t+1$ input-output vector is simply the output at time $t+1$, i.e., $F_{t+1}(X_{t+1})$, divided by what output would be if the input vector at time $t+1$ were used with the time t technology, i.e., $F_t(X_{t+1})$.

A Malmquist index for productivity growth between periods t and $t+1$ that uses the production technology of time t as a reference technology is then defined as

$$M^t_{t,t+1} \equiv \frac{D_t(X_{t+1}, Y_{t+1})}{D_t(X_t, Y_t)} = \frac{F_{t+1}(X_{t+1})/F_t(X_{t+1})}{1} = \frac{F_{t+1}(X_{t+1})}{F_t(X_{t+1})}.$$

We can also consider $D_{t+1}(X_t, Y_t)$, i.e., the distance of the input-output vector at time t from the technical frontier at time $t+1$. Applying the definition of the distance function, we have that $Y_t/D_{t+1}(X_t, Y_t) = F_{t+1}(X_t)$, which implies

$$D_{t+1}(X_t, Y_t) = \frac{Y_t}{F_{t+1}(X_t)} = \frac{F_t(X_t)}{F_{t+1}(X_t)}.$$

A Malmquist index for productivity growth between periods t and $t+1$ that uses the production technology of time $t+1$ as a reference technology is then defined as

$$M^{t+1}_{t,t+1} \equiv \frac{D_{t+1}(X_{t+1}, Y_{t+1})}{D_{t+1}(X_t, Y_t)} = \frac{1}{F_t(X_t)/F_{t+1}(X_t)} = \frac{F_{t+1}(X_t)}{F_t(X_t)}.$$

Caves, Christensen, and Diewert (1982) recommend defining the Malmquist index as the geometric average of $M^t_{t,t+1}$ and $M^{t+1}_{t,t+1}$. In this case the Malmquist index becomes

$$M_{t,t+1} \equiv \left(\frac{D_t(X_{t+1}, Y_{t+1})}{D_t(X_t, Y_t)} \frac{D_{t+1}(X_{t+1}, Y_{t+1})}{D_{t+1}(X_t, Y_t)} \right)^{1/2} = \left(\frac{F_{t+1}(X_{t+1})}{F_t(X_{t+1})} \frac{F_{t+1}(X_t)}{F_t(X_t)} \right)^{1/2}.$$

This definition avoids favoring the technology in one of the two periods over the other.

E.1 Normalization and the Malmquist Index

As we discuss in footnote 13 in the body of the paper, one symptom of A_t not being a good measure of productivity in the case were the functional form of the production function changes over time is that the growth rate of A_t will be sensitive to the choice of normalization of the inputs to production. This is not the case for the Malmquist index.

To illustrate this, consider again the change in the unit in which labor is expressed that we discussed in footnote 13: $\ddot{L}_t \equiv \psi L_t$. In this case we have that

$$F_t(Z, K_t, L_t) \equiv A_t Z^{\alpha_t} K_t^{\beta_t} L_t^{1-\alpha_t-\beta_t} = \ddot{A}_t Z^{\alpha_t} K_t^{\beta_t} \ddot{L}_t^{1-\alpha_t-\beta_t} \equiv \ddot{F}_t(Z, K_t, \tilde{L}_t), \tag{38}$$

where

$$\ddot{A}_t \equiv \frac{A_t}{\psi^{1-\alpha_t-\beta_t}}$$

Clearly, if α_t or β_t vary over time, the growth rates of A_t and $\ddot{A}_t$ will not be the same.

The Malmquist index, however, suffers no such issue. Since, by equation (38), $F_t(Z, K_t, L_t) = \ddot{F}_t(Z, K_t, \tilde{L}_t)$, this equation immediately implies that the Malmquist index remains the same. In fact, any rewriting of the production function that leaves the mapping from input to output unchanged, i.e. that does not change the production possibility frontier, implies the same Malmquist index because the formula for the Malmquist index only depends on output for some quantities of inputs.

We can illustrate this point by deriving an expression for the Malmquist index in terms of the observables in our model—equation (18)—for both F_t and $\ddot{F}_t$ and denoting the associated indices as m_t and $\ddot{m}_t$:

$$\begin{aligned}
\hat{m}_t &= \hat{a}_t + \hat{\alpha}_t \log Z + \hat{\beta}_t \bar{k}_t - (\hat{\alpha}_t + \hat{\beta}_t)\bar{l}_t \\
&= \hat{a}_t + (\hat{\alpha}_t + \hat{\beta}_t)\log\psi + \hat{\alpha}_t \log Z + \hat{\beta}_t \bar{k}_t - (\hat{\alpha}_t + \hat{\beta}_t)(\bar{l}_t + \log\psi) \\
&= \hat{\ddot{a}}_t + \hat{\alpha}_t \log Z + \hat{\beta}_t \bar{k}_t - (\hat{\alpha}_t + \hat{\beta}_t)\bar{\ddot{l}}_t \\
&= \hat{\ddot{m}}_t.
\end{aligned}$$

Recall that hats denote deviations from the previous period, $\hat{x}_t = x_t - x_{t-1}$, and bars denote the average of period $t-1$ and period t, $\bar{x}_t = (x_{t-1} + x_t)/2$. To go from the first to the second line, we added and subtracted the normalization that transforms l_t into $\ddot{l}_t$: $(\hat{\alpha}_t + \hat{\beta}_t)\log\psi$. In the third line, this time-varying normalization is absorbed by the A residual, $\hat{\ddot{a}}_t$, and $\bar{l}_t + \log\psi$ is converted to $\bar{\bar{l}}_t$. From this we see that while the A residual is normalization-dependent the Malmquist index is not.

F Model Equations

We reproduce the equations and distributional assumptions of our full model here for convenience:

$$\begin{aligned}
w_t &= \phi_t + \tfrac{1}{1-\beta_t} a_t - \tfrac{\alpha_t}{1-\beta_t}(d_t + n_t) - \tfrac{\beta_t}{1-\beta_t}\log\left(r_t + \delta\right) \\
\phi_t &= \log\beta_t + \log\left(1 - \alpha_t - \beta_t\right) + \tfrac{\alpha_t}{1-\beta_t} z - (\alpha_t + \beta_t)\lambda \\
s_t &= w_t + n_t + d_t - z + \log\alpha_t - \log(1 - \alpha_t - \beta_t) \\
k_t &= w_t + n_t + d_t - \log(r_t + \delta) + \log\beta_t - \log(1 - \alpha_t - \beta_t) \\
n_t &= n_{t-1} + \omega + \gamma(w_{t-1} + d_{t-1}) + \xi_{1t} + \xi_{2t} \\
\hat{m}_t &= \hat{a}_t + \hat{\alpha}_t z + \hat{\beta}_t \bar{k}_t - \left(\hat{\alpha}_t + \hat{\beta}_t\right)(\bar{d}_t + \bar{n}_t) \\
m_t &= \tilde{m}_t + \epsilon_{2t} \\
\tilde{m}_t &= \mu + \tilde{m}_{t-1} + \epsilon_{1t}
\end{aligned}$$

$$\exp(\xi_{1t}) \sim \begin{cases} \beta(\beta_1, \beta_2), & \text{with probability } \pi \\ 1, & \text{with probability } 1 - \pi \end{cases}$$

$$\epsilon_{1t} \sim \mathcal{N}(0, \sigma^2_{\epsilon_1}), \quad \epsilon_{2t} \sim \mathcal{N}(0, \sigma^2_{\epsilon_2}), \quad \xi_{2t} \sim \mathcal{N}(0, \sigma^2_{\xi_2})$$

Before 1760, α_t and β_t are assumed to be constant. This implies that the sixth equation collapses to $\hat{a}_t = \hat{m}_t$ before 1760. As a result, rents s_t and the capital stock k_t only appear in the equations that define them (the third and fourth equations). This is also the period for which we do not have data on s_t and k_t. For this period, we therefore use the third and fourth equations to estimate s_t and k_t.

Below we reproduce the assumptions we make about measurement error and normalizations in our data:

$$
\begin{aligned}
w_t &= \varphi^w + \tilde{w}_t \\
n_t &= \psi + \tilde{n}_t + \iota_t^n \\
d_t &= \tilde{d}_t + \iota_t^d \\
r_t &= \tilde{r}_{it} + \iota_{it}^r \\
s_t &= \varphi^s + \tilde{s}_t + \iota_t^s, \\
k_t &= \varphi^k + \tilde{k}_t + \iota_t^k,
\end{aligned}
$$

Here, the variables with tilde's are the measured variables, while the variables without tilde's are the true variables, $\varphi^w \sim \mathcal{N}(0, 100^2)$, $\varphi^s \sim \mathcal{N}(0, 100^2)$, and $\varphi^k \sim \mathcal{N}(0, 100^2)$ are normalization constants, and $\iota_t^n \sim t_{\nu_n}(0, \sigma_n^2)$, $\iota_t^d \sim t_{\nu_d}(0, \tilde{\sigma}_d^2)$, $\iota_{it}^r \sim t_{\nu_{ir}}(0, \tilde{\sigma}_{ir}^2)$, $\iota_t^s \sim t_{\nu_s}(0, \tilde{\sigma}_s^2)$, and $\iota_t^k \sim t_{\nu_k}(0, \tilde{\sigma}_k^2)$ capture measurement error. A few additional details regarding missing observations are given in the main text.

G A Comparison with Clark (2010, 2016)

Our approach to estimating productivity in England from the 13th to 19th centuries yields quite different results than the most comprehensive existing estimates by Clark (2010, 2016). Here, we consider from where the differences arise. We break this discussion into three parts. First, we discuss Clark's dual approach and differences between his 2010 series and his 2016 series. Second, we discuss how Clark's dual approach relates to our Malmquist approach. Third, we discuss differences that arise from the fact that our approach has different implications for the evolution of factor prices and factor output elasticities than Clark's approach.

A summary of our conclusions is as follows. First, Clark (2010) made an error in calculating the growth rate of his index from 1540 to 1550 which contributes to the difference between this series and our series. Clark (2016) corrects this error. Second, using the average of factor output elasticites at time t and $t-1$ when calculating changes in productivity between time t and $t-1$ explains an important part of difference in our results, especially prior to 1600. Conditional on doing this Clark's dual approach is approximately equal to our Malmquist approach. Third, differences in the factor prices and factor output elasticites implied by our approach, relative to those used by Clark, explain the remaining differences in the evolution of productivity.

G.1 Clark's Dual Approach and Differences Between Clark (2010) and Clark (2016)

Clark (2010, 2016) employs a "dual approach" to estimating productivity. Specifically, his estimate of the growth rate of productivity is

$$\frac{E_t}{E_{t-1}} = \left(\frac{S_t}{S_{t-1}}\right)^{s_{Z,t-1}} \left(\frac{r_t+\lambda}{r_{t-1}+\lambda}\right)^{s_{K,t-1}} \left(\frac{W_t}{W_{t-1}}\right)^{s_{L,t-1}} \frac{1-\tau_{t-1}}{1-\tau_t}. \tag{39}$$

where we use E_t (for efficiency) to denote the dual estimate of productivity, λ is a risk premium, τ_t is the share of national income paid in indirect taxes, and $s_{Z,t-1}$, $s_{K,t-1}$, and $s_{L,t-1}$ are time-varying estimates of the elasticity of output with respect to land, capital, and labor, respectively.[23]

Clark's 2016 productivity series is an updated version of his better known 2010 productivity series for the sample period 1250-1600. Clark has shared with us the file he used to construct his 2016 series by private correspondence. This file extends his 2016 series from 1600 to 1860 and contains the component series Clark uses to construct this series. Our discussion here is based on these series. For the period after 1600, the new productivity series coincides with Clark's 2010 series.

Figure G.1 plots Clark's 2010 productivity series (solid gray line) and Clark's 2016 productivity series extended to 1860 using the file Clark shared with us (broken black line). We refer to the extended 2016 series as "Clark (2016)*". These series differ for two reasons. First, Clark's 2010 series contains an error in the growth rate from 1540 to 1550. This error creates a 25 log point spurious drop in the 2010 series. Clark's 2016 series corrects this error. Second, Clark´s 2016 series incorporates a new land rent series for the period 1250-1600. Both of these changes make Clark's 2016 series more similar to our baseline productivity estimate (solid black line in Figure G.1) than his 2010 series.

The Malmquist index we use for our baseline estimates uses average factor output elasticities rather than lagged factor output elasticities. Using average factor output elasticities is also recommended by Barro and Sala-i-Martin (2004, p. 435). We can modify Clark's dual approach—equation (39)—to use average factor output elasticities as follows:

$$\frac{E_t}{E_{t-1}} = \left(\frac{S_t}{S_{t-1}}\right)^{\bar{s}_{Z,t}} \left(\frac{r_t+\lambda}{r_{t-1}+\lambda}\right)^{\bar{s}_{K,t}} \left(\frac{W_t}{W_{t-1}}\right)^{\bar{s}_{L,t}} \frac{1-\tau_{t-1}}{1-\tau_t}. \tag{40}$$

[23]The discussion in Clark (2010, 2016) suggests that Clark estimates the level of productivity rather than its growth rate. However, data Clark has shared with us (discussed below) makes clear that he, in fact, estimates growth rates of productivity. This distinction is important as the level formula Clark discusses in his 2010 and 2016 papers does not provide a valid measure of productivity when factor shares are allowed to vary over time. See footnote 13 in the main text for more detail on this point.

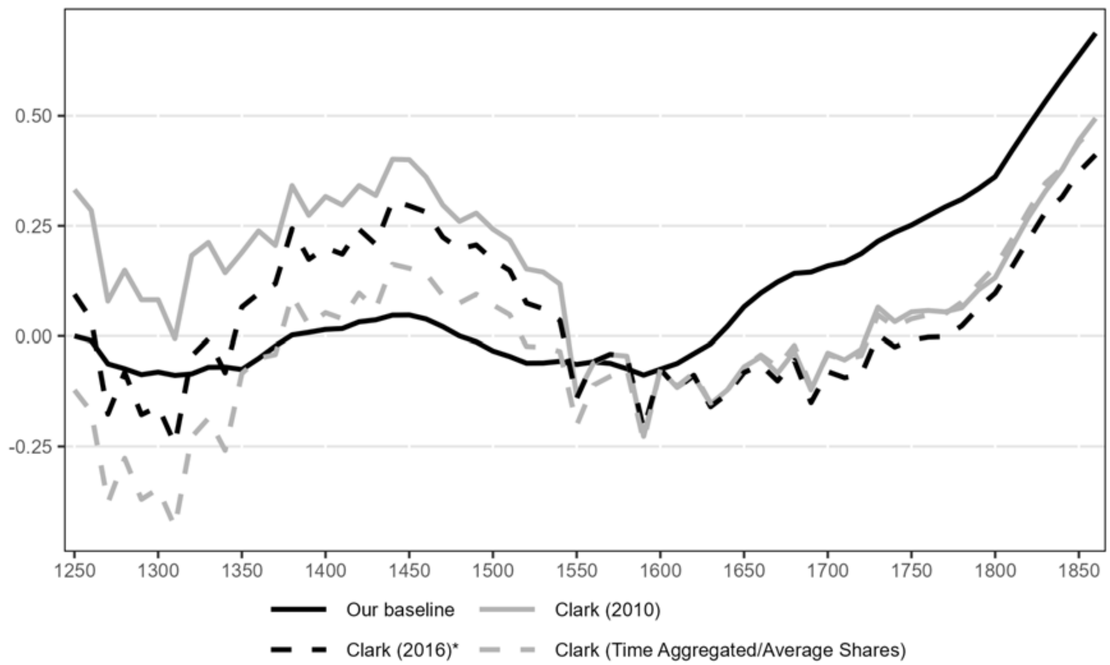

Figure G.1: A Comparison between Clark (2010) and Clark (2016)

Note: The Figure plots four productivity series. The solid black line is our baseline Malmquist index. The solid gray line is Clark's (2010) original productivity series. The broken black line—labeled "Clark (2016)*"—is Clark's (2016) productivity series extended to 1860. We obtained this series from Clark in private correspondence. The broken gray line is an estimate of productivity using equation (39) with decadal data, i.e., this series moves to average output elasticities and time aggregates relative to the Clark (2016)* series. The latter three series are normalized to be equal to the Malmquist index in 1600.

As in the main text, the bar on top of each s signifies an average between $t-1$ and t: $\bar{s}_{Z,t} = (s_{Z,t-1} + s_{Z,t})/2$ and similarly for $\bar{s}_{K,t}$ and $\bar{s}_{L,t}$.

The fourth line plotted in Figure G.1 is productivity growth estimated using equation (40) and Clark's data series for factor prices and factor output elasticities (broken gray line). This line also differs from the two Clark series because of time aggregation. Clark estimates productivity using equation (39) at an annual frequency and then averages over decades. To be consistent with our approach in the rest of the paper, we average the data over each decade and then use equation (39) to estimate productivity at a decadal frequency. We see that moving from lagged to average factor output elasticities and decadal time aggregation results in estimates of productivity that are lower early in the sample. This difference is mostly due to the switch to average factor output elasticities—time aggregation only makes a small difference. These changes result in a productivity series that is closer to ours between 1350 and 1600.

G.2 The Dual Approach versus the Malmquist Approach

We next show that our Malmquist index and the dual approach are equivalent up to a first-order approximation. To see this, we go back to equation (18), which we reproduce here for convenience:

$$\hat{m}_t = \hat{a}_t + \hat{\alpha}_t \log Z + \hat{\beta}_t \bar{k}_t - (\hat{\alpha}_t + \hat{\beta}_t) l_t.$$

In this equation, bars denote arithmetic averages across period $t-1$ and t and hats denote differences between the two periods. Rearranging this equation yields[24]

$$\hat{m}_t = \hat{y}_t - \bar{\beta}_t \hat{k}_t - (1 - \bar{\alpha}_t - \bar{\beta}_t)\hat{l}_t. \tag{41}$$

The right-hand side is the primal measure of the growth rate of productivity, i.e., the Solow residual (Solow, 1957). Here, weights are given by the arithmetic average of the factor output elasticities across the two periods. To go from the primal measure to the dual measure, we can follow Hsieh (2002) and start from the fact that the value of output must equal payments to factors: $Y_t = S_t Z + (r_t + \delta) K_t + W_t L_t$. Taking a log-linear approximation of this expression at times $t-1$ and t around a situation where factor output elasticities are the averages of the two periods yields the following expression:

$$\hat{y}_t = \bar{\alpha}_t \hat{s}_t + \bar{\beta}_t \left(\log \left(\frac{r_t + \delta}{r_{t-1} + \delta} \right) + \hat{k}_t \right) + (1 - \bar{\alpha}_t - \bar{\beta}_t) \left(\hat{w}_t + \hat{l}_t \right),$$

[24] The derivation is

$$\begin{aligned}
\hat{m}_t =& a_t - a_{t-1} + \frac{1}{2} \left(\alpha_t \log Z + \beta_t k_t + (1 - \alpha_t - \beta_t) l_t - (\alpha_{t-1} \log Z + \beta_{t-1} k_{t-1} + (1 - \alpha_{t-1} - \beta_{t-1}) l_{t-1}) \right) \\
&- \frac{1}{2} \left(\alpha_{t-1} \log Z + \beta_{t-1} k_t + (1 - \alpha_{t-1} - \beta_{t-1}) l_t - (\alpha_t \log Z + \beta_t k_{t-1} + (1 - \alpha_t - \beta_t) l_{t-1}) \right) \\
=& a_t + \alpha_t \log Z + \beta_t k_t + (1 - \alpha_t - \beta_t) l_t - (a_{t-1} + \alpha_{t-1} \log Z + \beta_{t-1} k_{t-1} + (1 - \alpha_{t-1} - \beta_{t-1}) l_{t-1}) \\
&- \frac{1}{2} \left(\alpha_t \log Z + \beta_t k_t + (1 - \alpha_t - \beta_t) l_t - (\alpha_{t-1} \log Z + \beta_{t-1} k_{t-1} + (1 - \alpha_{t-1} - \beta_{t-1}) l_{t-1}) \right) \\
&- \frac{1}{2} \left(\alpha_{t-1} \log Z + \beta_{t-1} k_t + (1 - \alpha_{t-1} - \beta_{t-1}) l_t - (\alpha_t \log Z + \beta_t k_{t-1} + (1 - \alpha_t - \beta_t) l_{t-1}) \right) \\
=& \hat{y}_t - \bar{\beta}_t \hat{k}_t - (1 - \bar{\alpha}_t - \bar{\beta}_t) \hat{l}_t.
\end{aligned}$$

For the first equality, we just use the definition of the bar and hat symbols. For the second equality, we add and subtract the expression contained in the line that follows the second equal sign. The third equality is again a straightforward use of the bar and hat symbols.

where we have dropped higher order terms. Combining this equation and equation (41), we obtain

$$\hat{m}_t = \bar{\alpha}_t \hat{s}_t + \bar{\beta}_t \log\left(\frac{r_t + \delta}{r_{t-1} + \delta}\right) + (1 - \bar{\alpha}_t - \bar{\beta}_t)\hat{w}_t. \tag{42}$$

This equation shows that the log-change in the Malmquist index is equal to the dual measure of productivity growth up to a first order approximation.

The productivity measures in equation (42) differ in some details from the ones plotted in Figure G.1. First, the left-hand-side of equation (42) is $\hat{m}_t$, the change in m_t. The productivity measure plotted as our baseline estimate in Figure G.1 (solid black line) is $\tilde{m}_t$ rather than m_t. Recall that $\tilde{m}_t$ is the permanent component of productivity (see equations (19)–(20)). Our baseline estimate in Figure G.1, thus, filters out some high frequency variation in productivity, which makes it smoother than estimates based on the dual approach.

Clark's dual approach also differs in a few details from the right-hand side of equation (42). Clark's dual approach does not incorporate capital depreciation (δ), but it includes a risk premium (λ) and taxes (τ_t) that are not incorporated into equation (42). The similarity (and difference in details) between the right-hand side of (42) and Clark's dual approach can be more easily seen by taking logarithms of equation (40):

$$\hat{e}_t = \bar{s}_{Z,t}\hat{s}_t + \bar{s}_{K,t} \log\left(\frac{r_t + \lambda}{r_{t-1} + \lambda}\right) + \bar{s}_{L,t}\hat{w}_t - \log\left(\frac{1 - \tau_t}{1 - \tau_{t-1}}\right) \tag{43}$$

Comparing this equation to equation (42), notice that in our model, α_t, β_t, and $1 - \alpha_t - \beta_t$ are the land, capital, and labor output elasticicities, while in equation (43) these are denoted by $s_{Z,t}$, $s_{K,t}$ and $s_{L,t}$, respectively. The two formulas are, thus, the same apart from δ being replaced by λ, and the presence of τ_t in equation (43).

These details turn out not to make much of a difference. To see this, Figure G.2 plots a dual measure of productivity using the formula in equation (40) but with our factor price and factor output elasticities series (broken black line). In other words, this measure of productivity, differs from ours only in terms of method, not data. We see that the resulting productivity series tracks our baseline productivity series very closely. The only difference is that our measure is smoother at high frequency reflecting the fact that it filters out high-frequency movements in productivity.

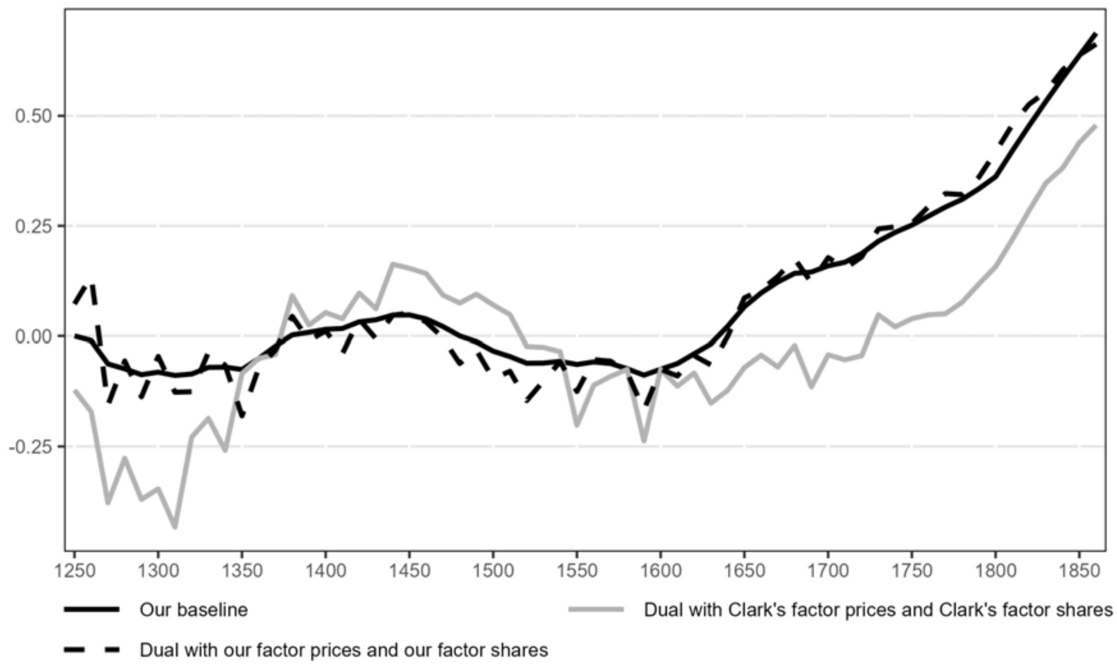

Figure G.2: Our Productivity Measure Compared with the Dual Approach

Note: The Figure plots three productivity series. The solid black line is our baseline Malmquist index. The solid gray line is the index constructed with Clark's factor prices and factor shares using equation (40). This is the same line as the one we label "Clark (Time Aggregation/Average Shares)" in Figure G.1. The dashed black line is the index constructed with our factor prices and factor shares using equation (40). The latter two series are normalized to be equal to the Malmquist index in 1600.

G.3 Factor Prices and Factor Output Elasticities

We now turn to the role played by differences in the factor price and factor output elasticity series used by Clark relative to those implied by our analysis. Since we have shown above that the dual approach and the Malmquist approach are virtually equivalent, we will carry out the rest of the analysis using the dual approach for concreteness. In particular, we will calculate productivity using equation (40) with different combinations of Clark's and our factor price and factor output elastiticy series. (Clark refers to factor output elasticities as factor shares.) In the case of Clark's series, we will use Clark's 2016 series extended to 1860. We have already plotted two such cases in Figure G.2. The solid gray line uses Clark's factor price and factor output elasticity series, while the broken black line uses our factor price and factor output elasticity series. Next, we consider intermediate cases.

A complication that arises if we seek a decomposition of the remaining difference between our productivity index and Clark's—the solid gray line and the broken black line in Figure G.2—into the share explained by factor prices and the share explained by factor output elasticities is that the productivity indexes we are considering are non-linear. This implies that the difference in

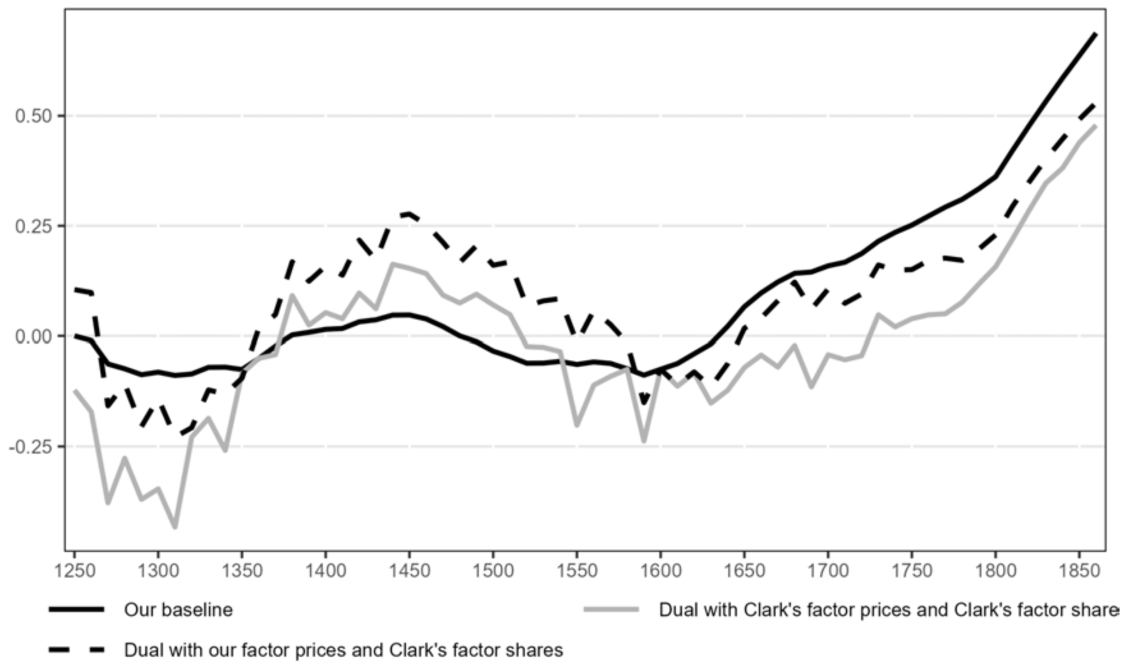

Figure G.3: Contribution of Factor Prices to Differences in Productivity Estimates

Note: The Figure plots three productivity series. The solid black line is our baseline Malmquist index. The solid gray line is the index constructed with Clark's factor prices and factor shares using equation (40). The dashed gray line is the index constructed with Clark's factor prices and Clark factor shares using equation (40). The latter two series are normalized to be equal to the Malmquist index in 1600.

question is not simply the sum of the effect of changing the factor prices, on the one hand, and the effect of changing the factor output elasticities, on the other hand. Rather, there is also an interaction term, which is non-trivial.

With this in mind, we begin by considering how changing the factor price series alone affects the productivity series. Figure G.3 plots a dual estimate of productivity using Clark's factor output elasticity but our factor price series (broken black line). The difference between the solid gray line and the broken black line in Figure G.3 is thus due to moving from Clark's factor price series to our factor price series. Focusing on the period after 1600, we see that this change explains a sizable portion of the difference between our results and the series using Clark's factor prices and factor output elasticities, especially during the 17th and early 18th centuries. Prior to 1600, moving to Clark's factor price series raises productivity which helps explain the difference between our results and Clark's early in the sample, but makes this difference larger between 1350 and 1600.

Figure G.4 plots Clark's factor price series (solid gray lines) and our factor price series (solid black lines). In the case of land rents, we also plot the series used in Clark (2010) (broken gray line). Our real wage series looks similar to Clark's. The raw real interest rate date we use is also similar to that used by Clark. However, we allow for measurement error in real interest rates and

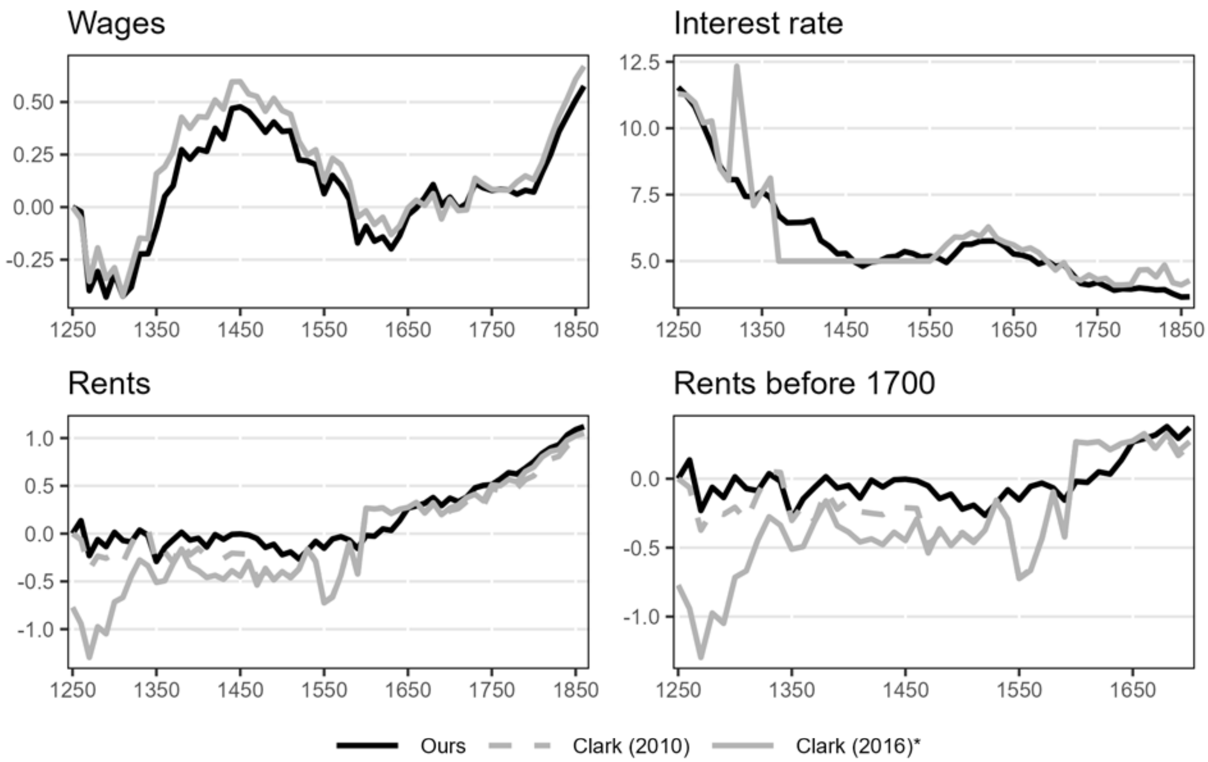

Figure G.4: A Comparison of Price Series

Note: The top two panels plots the wage and interest rate series used in our analysis and used by Clark (2016). The bottom two panels plot the land rent series used by Clark (2010, 2016) and the land rents that are implied by our analysis.

make use of two return series (rates of return on land and rent charges). This implies that our real interest rate series is substantially smoother in the early part of our sample and around 1600. In particular, Clark's interest rate series is constant between 1370 and 1540, reflecting Clark's choice of how to interpolate over a period of relatively sparse data, while our series falls more gradually over the early part of this period.

For land rents, we use the same data as Clark after 1760 but choose to infer land rents from the model prior to 1760. Our inferred series differs quite a bit from Clark's series, especially early in the sample. Clark's data is quite noisy over this early period. But measuring land rents prior to 1650 is difficult due to the complexity of the relationship between landlords and tenants in a feudal era. It is also notable that Clark's 2016 series for land rents differs quite substantially from his earlier 2010 land rent series for the period prior to 1500. From 1250 to 1500, the 2016 series increases by 47%, while the 2010 series falls by 68%.

Turning to factor output elasticities, Figure G.5 plots a dual estimate of productivity using Clark's factor prices but our factor output elasticity series (broken black line). The difference

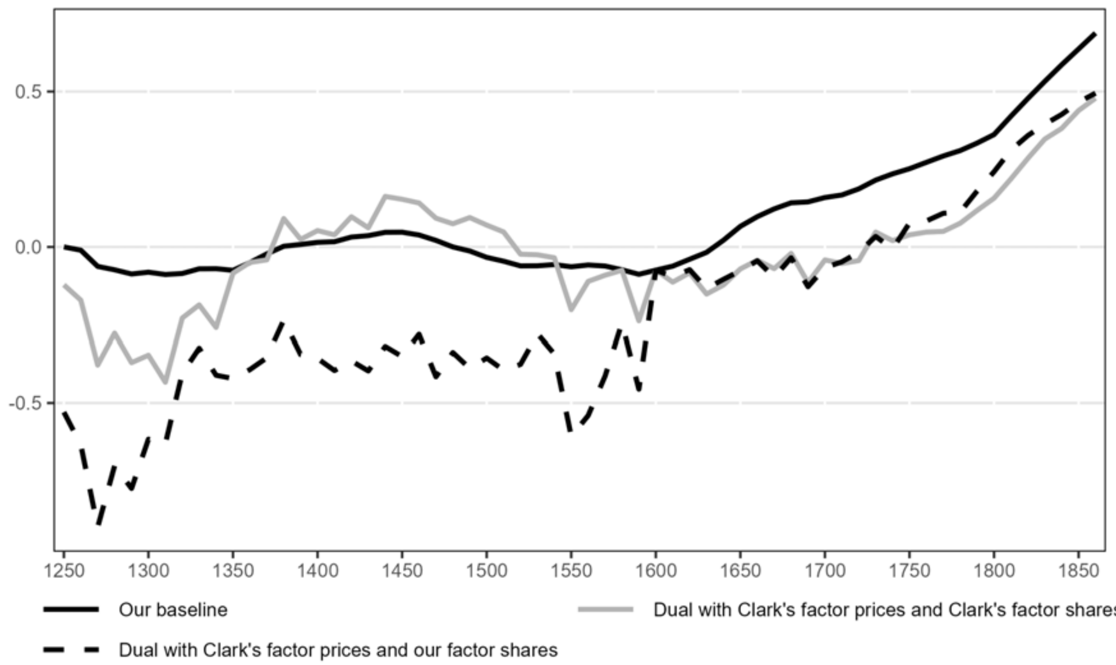

Figure G.5: Contribution of Factor Shares to Differences in Productivity Estimates

Note: The Figure plots three productivity series. The solid black line is our baseline Malmquist index. The solid gray line is the index constructed with Clark's factor prices and factor shares using equation (40). The dashed gray line is the index constructed with Clark's factor prices and our factor shares using equation (40). The latter two series are normalized to be equal to the Malmquist index in 1600.

between the solid gray line and the broken black line in Figure G.5 is thus due to moving from Clark's factor output elasticity series to our factor output elasticitiy series. For the period after 1600, this change has a minimal effect. Prior to 1600, the differences are larger. Shifting to our factor output elasticity results in a sharp rise of the productivity series from 1250 to 1400. This reflects the increase in Clark's 2016 rent series (which both the solid gray and broken black lines are using). It also results in high volatility and a substantial increase in the 16th century. Figures G.1, G.2, G.3, and G.5 taken together indicate that the difference between Clark's series and our series before 1600 is a complicated combination of the effects of factor prices, factor output elasticities, their interaction, time aggregation, and average versus lagged factor output elasticities.

Figure G.6 compares our estimates of factor output elasticities (black lines) with Clark's (gray lines). The largest difference is for the output elasticity of land. We estimate a substantially larger output elasticity of land than Clark. Recall that our estimate of the output elasticity of land is derived from our estimate of the slope of the labor demand curve. Clark's constructs his estimate from estimates of factor shares. His basic approach is to calculate payments to factors by multiplying factor prices with the quantity of those factors. A challenge with this approach is that Clark does not have much data on factor quantities. This forces him to make strong assumptions

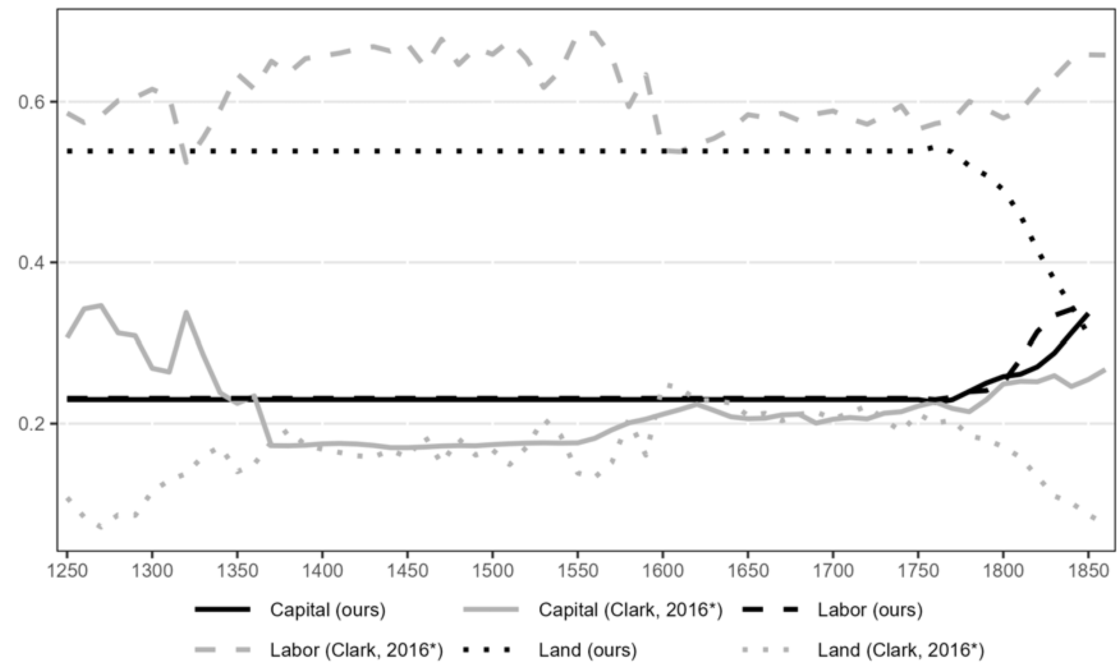

Figure G.6: Factor Shares

Note: The figure presents the factor shares implies by our analysis (black lines) and those used by Clark (2016) extended to 1860 (gray lines). We obtained the latter series from Clark in private correspondence.

(educated guesses) about the factor quantities.

For instance, Clark's estimate of payments to labor is: $W_t \times 300 \times \nu N_t$, where W_t is the average daily wage, 300 is the assumed number of days worked, N_t is population, and ν is the fraction of the population that is economically active, which he assumes to be 34%. Clark's assumption that days worked are constant over the entire sample period contrasts sharply with the estimates of Humphries and Weisdorf (2019). Also, it is not clear why he choses 300 days. Earlier work often chose 250. Finally, the notion that the fraction of the population that was economically active was constant over our sample is also a strong assumption. In particular, an important literature has highlighted variation in marriage patterns over our sample and associated variation in the employment of women (De Moor and van Zanden, 2010, Voigtländer and Voth, 2013).

Similarly, to construct payments to capital, Clark makes educated guesses on the stock of housing, improvements to land, livestock, etc. He estimates payments to land by multiplying the rent index with a fixed stock of land (28.24 million acres) before the 1840s and direct estimates from tax returns after this date. With factor payments estimated in this manner, each factor's share can be obtained by dividing payments accruing to that factor by payments accruing to all factors. Clark's estimates of factor share are, thus, based on a number of strong empirical conjectures. A curious aspect of his estimates is that he estimates a large capital share in the 13th century that then falls by about half in the 14th century (mostly before the Black Death).

H Impulse Response Functions

H.1 Dynamics After Change in Productivity Growth

Our Malthusian model implies that an increase in productivity growth will result in higher steady state wages. To see this, we first abstract for notational simplicity from all the shocks in our model. More precisely, we set the value of all shocks equal to their mean. The mean value of ϵ_{1t}, ϵ_{2t}, and ξ_{2t} is zero. The mean value of ξ_{1t}, however, is $E\xi_{1t} = \pi(\psi\beta_1) - \psi(\beta_1 + \beta_2))$, where $\psi(\cdot)$ is the digamma function. We furthermore, assume that days worked and the interest rate are constant at d^* and r^*.

Given these assumptions, our model simplifies to:

$$w_t = \phi + \frac{1}{1-\beta}\tilde{a}_t - \frac{\alpha}{1-\beta}(n_t + d^*) - \frac{\beta}{1-\beta}\log(r^* + \delta) \tag{44}$$

$$n_t - n_{t-1} = \omega + \gamma(w_{t-1} + d^*) + E\xi_{1t} \tag{45}$$

$$\tilde{a}_t = \mu + \tilde{a}_{t-1}. \tag{46}$$

We can use equation (44) to eliminate w_t in equation (45). This yields:

$$n_t - n_{t-1} = \omega + \gamma\phi + \frac{\gamma}{1-\beta}\tilde{a}_{t-1} - \frac{\alpha\gamma}{1-\beta}n_{t-1} - \frac{\beta\gamma}{1-\beta}\log(r^* + \delta) + \gamma\frac{1-\alpha-\beta}{1-\beta}d^* + E\xi_{1t}.$$

This equation can be rewritten as

$$n_{t+1} = \left(1 - \frac{\gamma\alpha}{1-\beta}\right)n_t + \frac{\gamma}{1-\beta}\tilde{a}_{t-1} + \text{constant}. \tag{47}$$

Next, we subtract α times the second-to-last equation from equation (46) and rearrange. This yields:

$$\tilde{a}_t - \alpha n_t = \mu - \kappa + \frac{1-\alpha\gamma-\beta}{1-\beta}\left(\tilde{a}_{t-1} - \alpha n_{t-1}\right), \tag{48}$$

where

$$\kappa = \alpha\left(\omega + \gamma\phi + \gamma\frac{1-\alpha-\beta}{1-\beta}d^* - \frac{\beta\gamma}{1-\beta}\log(r^* + \delta) + E\xi_{1t}\right).$$

This shows that $\tilde{a}_t - \alpha \times n_t$ follows an $AR(1)$ and therefore settles down to a steady state in the long

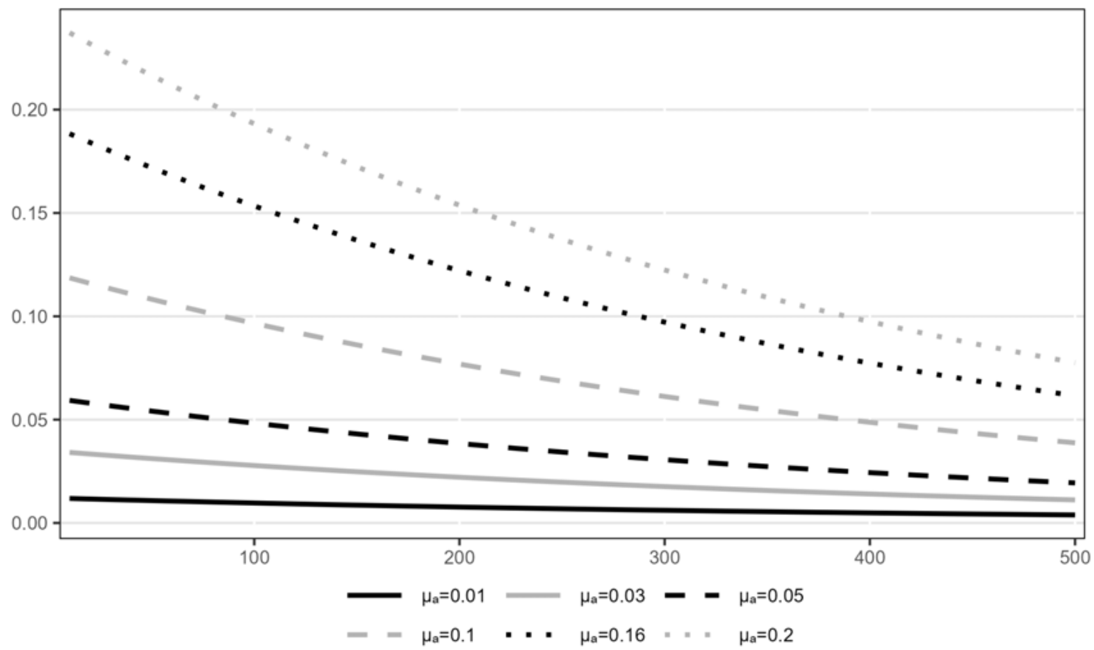

Figure H.1: Real Wage Growth After an Increase in Productivity Growth

Note: Each line plots the growth rate of real wages over time after an increase in productivity growth from $\mu = 0$ to a higher value. These impulse responses are calculated assuming that all model parameters are at their posterior mean values and α is equal to our pre-Industrial estimate of 0.38.

run as long as $|(1 - \alpha\gamma - \beta)/(1 - \beta)| < 1$. The steady state value of $\tilde{a}_t - \alpha n_t$ is $(\mu - \kappa)(1 - \beta)/(\alpha\gamma)$ and (using equation (44)) the steady state real wage is

$$w^* = \frac{\mu}{\alpha\gamma} - d^* - \frac{\omega}{\gamma} - \frac{E\xi_{1t}}{\gamma} \tag{49}$$

We see from this that the steady state real wage in our Malthusian economy is increasing in the productivity growth rate μ and the extent to which this is the case is influenced by the strength of the Malthusian population force as summarized by $\alpha\gamma$.

Figures H.1 and H.2 present impulse responses to a change in productivity growth that show quantitatively how much changes in productivity growth increase wages over time according to our model when α is set to our pre-Industrial estimate ($\alpha = 0.38$). For each impulse response, we start the economy off in a steady state with zero productivity growth ($\mu = 0$). At time zero in the figures, productivity growth increases. In Figure H.1, we show the evolution of the growth rate of wages (log change) over the subsequent 500 years. In Figure H.2, we show the evolution of the level of wages relative to its earlier steady state level over the subsequent 1000 years. In both figures, we assume that all other shocks are constant at their mean values.

In Figure H.1, we see that the growth rate of wages is initially equal to the change in produc-

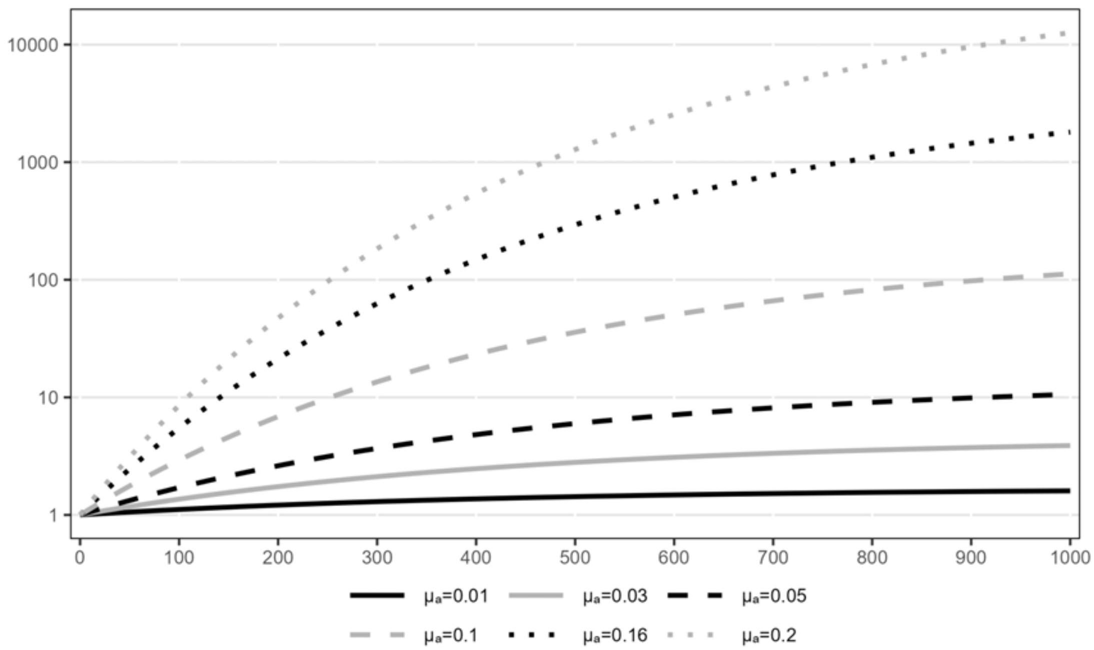

Figure H.2: Evolution of Real Wages After an Increase in Productivity Growth

Note: Each line plots the evolution of real wages over time after an increase in productivity growth from $\mu = 0$ to a higher value. These impulse responses are calculated assuming that all model parameters are at their posterior mean values and α is equal to our pre-Industrial estimate of 0.38.

tivity. As wages rise and the Malthusian population force kicks in, the growth rate of wages falls. This process takes a very long time due to the weakness of the Malthusian population force. As we discussed above, the half-life of wage growth is roughly 170 years when the land share is at its pre-1760 value. The fact that wage growth continues for hundreds of years after a change in productivity implies that the cumulative increase in wages is substantial. In Figure H.2, we can read off the long-run effect of higher productivity growth on wages. For a "modern" productivity growth rate of $\mu = 0.1$, we find that the long-run effect on the level of wages is an increase of a factor of 20.

H.2 Dymanics after Change in α

We now study the impulse response function of our Malthusian economy to a change in α. The thought experiment is the following: before time 0, the economy is on a balanced growth path with α constant and equal to α^H. At time 0, the value of α falls to α^L. β is constant at all times. Like before, we shut shocks down by setting them equal to their expected value. The permanent component of the Malmquist index follows its law of motion throughout the experiment:

$$\tilde{m}_t = \mu + \tilde{m}_{t-1}. \tag{50}$$

Since the economy is on the balanced growth path before time 0, the derivations of section H.1 apply and we have for all $t < 0$:

$$w_t = \frac{\mu}{\alpha^H \gamma} - d^* - \frac{\omega}{\gamma} - \frac{E\xi_{1t}}{\gamma}$$

$$\tilde{a}_t - \alpha^H n_t = \frac{(\mu - \kappa^H)(1-\beta)}{\alpha^H \gamma},$$

where κ^H is the value of κ when $\alpha = \alpha^H$. Similarly, since α is constant for $t \geq 1$, equations (44)–(46) hold and so does equation (48). Therefore, the convergence result for $a_t - \alpha^L n_t$ and w_t, $t > 0$, apply with $\alpha = \alpha^L$.

At time 0, things are more subtle as the change in α implies that equation (46) is replaced by equation (50). Combining equation (45) at time 0 and the formula for w_t with $t < 0$, we know n_0:

$$n_0 = n_{-1} + \frac{\mu}{\alpha^H}.$$

Invoking equation (18), we can solve for a_0:

$$\tilde{a}_0 = \tilde{a}_{-1} + \mu + (\alpha^H - \alpha^L)(\log Z - d^* - \bar{n}_0 - \lambda), \tag{51}$$

where we have used the fact that β is constant, $a_t = \tilde{a}_t$, $m_t = \tilde{m}_t$, and $\hat{\tilde{m}}_t = \mu$. Finally, w_0 is given by equation (44) with $\alpha = \alpha^L$.

We show the impulse response functions of W_t and N_t in Figures H.3 and H.4. We set α^H, the value of α before time 0, to 0.38, which is the posterior mean before 1770 and show the results for various values of α^L. The lowest one, 0.15, is the posterior mean for α_t in the last decade of the sample. For simplicity, we set $\mu = 0$ so that population has a well-defined steady state. Both variables are expressed as a multiple of their steady state value with $\alpha = \alpha^H$. Note that, by assumption, the variables are in the latter steady state before time 0.

Real wages jump on impact. Since productivity, defined as the Malmquist index, is held constant throughout, there is no change in output at time 0 and this jump is entirely explained by the

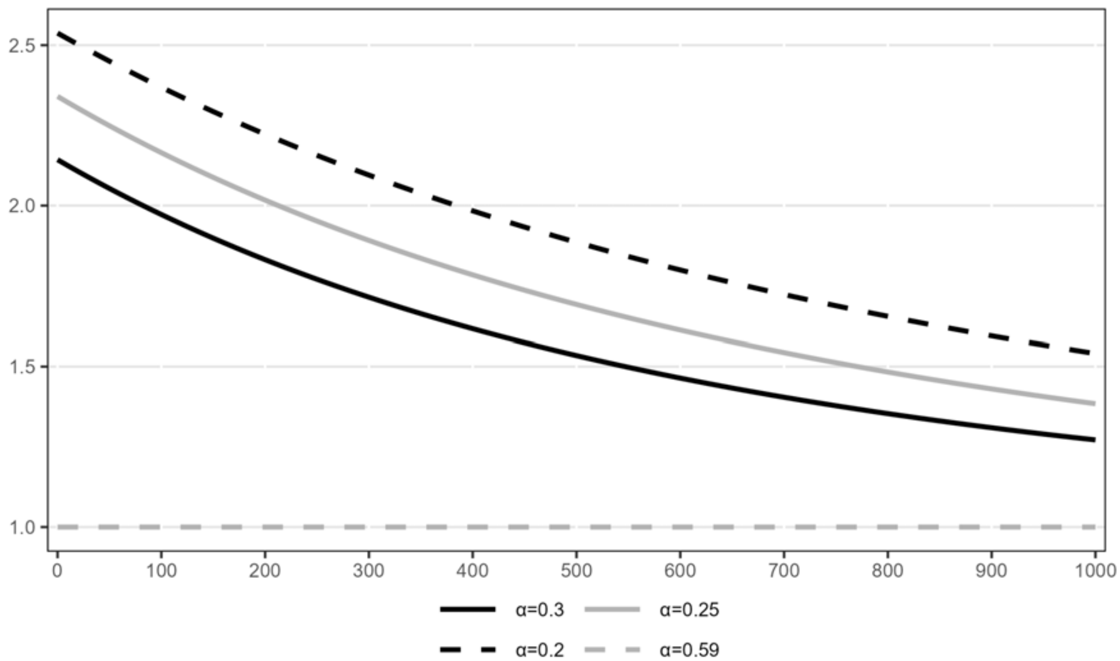

Figure H.3: Response of W_t to a Change in α

Note: The figure plots the response of the real wage (W_t) to a drop in α from its posterior mean before 1770 (0.38) to the value in the legend. W_t is expressed in multiple of its steady state value before the drop.

increase in the labor share.[25] Indeed, a drop in α from 0.38 to 0.15 implies a 51% increase in the labor share, which is exactly the increase in W_t on impact. From time 1 onward, population increases which pushes the wage down to the old steady state—without growth ($\mu = 0$), the steady state wage doesn't depend on α.

Population is predetermined at time 0, so it does not change on impact. As income rose in period 0, however, it starts increasing in period 1 and slowly converges to a permanently higher level. With a larger labor share, a bigger population can be sustained in steady state.

[25]Formally, the change in output is:

$$\hat{y}_0 = \hat{\hat{a}}_0 + \hat{\alpha}_0 \log Z + \beta \hat{k}_0 - \hat{\alpha}_0 (n_{-1} + d^* + \lambda) = \frac{1}{1-\beta} \left(\hat{\hat{a}}_0 + \hat{\alpha}_0 \left(\log Z - (n_{-1} + d^* + \lambda) \right) \right) = 0,$$

where we used the fact that $n_0 = n_{-1}$ when $\mu = 0$ in the first equality, the capital demand in the second one, and equation (51) in the third one.

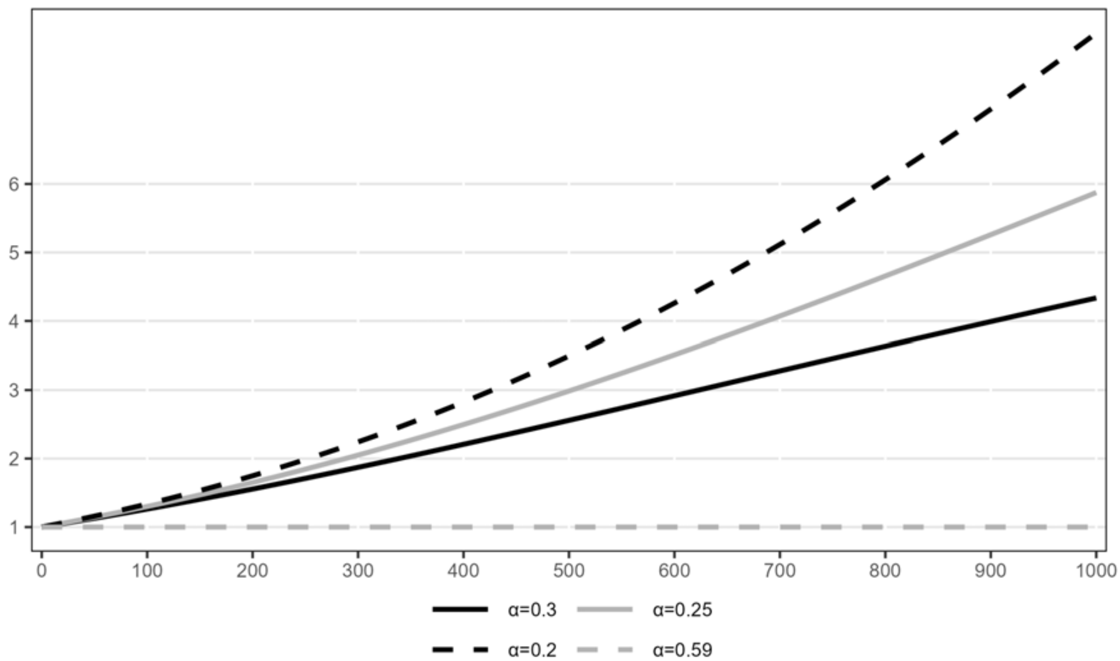

Figure H.4: Response of N_t to a Change in α

Note: The figure plots the response of population (N_t) to a drop in α from its posterior mean before 1770 (0.38) to the value in the legend. N_t is expressed in multiple of its steady state value before the drop.

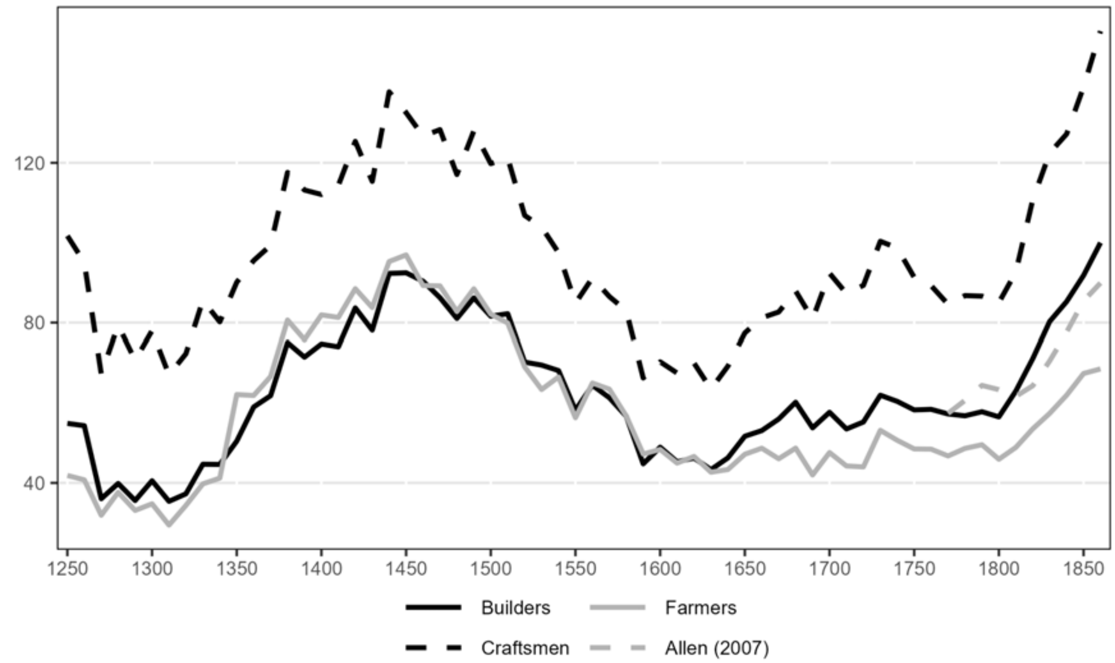

Figure A.1: A Comparison of Real Wage Measures in England, 1250-1860

Note: The figure presents four estimates of the real wages in England. Three are from Clark (2010): builders, farmers, and craftsmen. The remaining series is from Allen (2007). The builders series is the series we use in our main analysis. The builders series is normalized to 100 in 1860. The levels of the farmers and craftsmen series indicate differences in real earnings relative to builders. The Allen (2007) series is normalized to equal the builders series in 1770.

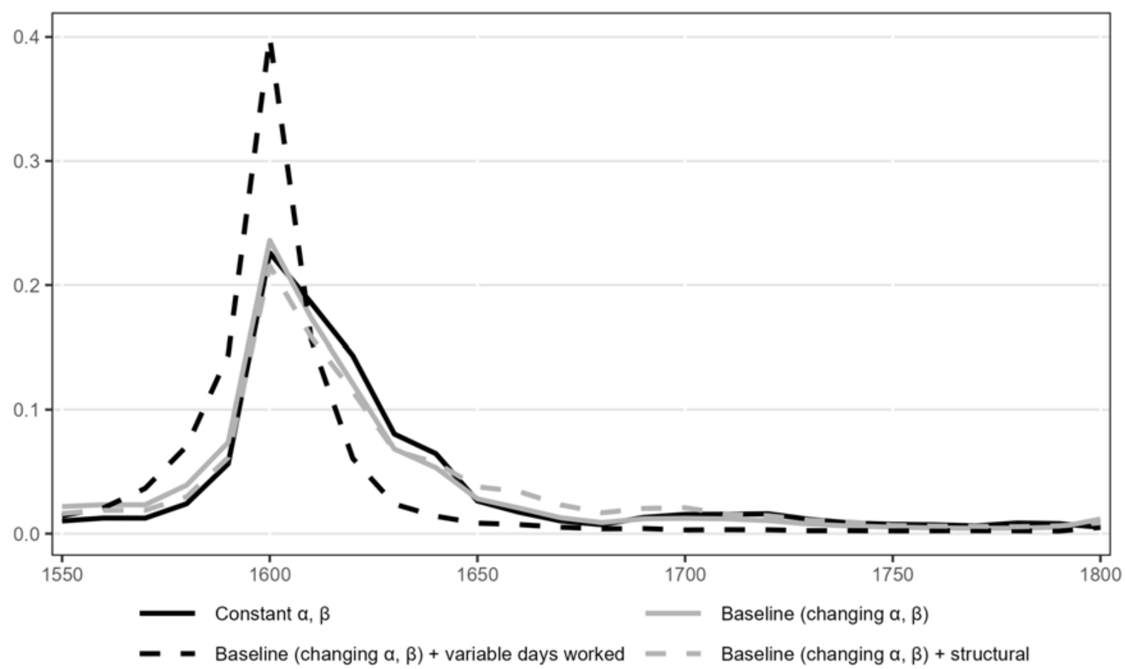

Figure A.2: Probability of Breaks Dates for Different Specifications

Note: The figure plots our estimate of the probability that a structural break occurred in the parameters μ, σ_1, and σ_2 in different decades between 1550 and 1800 for various specifications of our model.

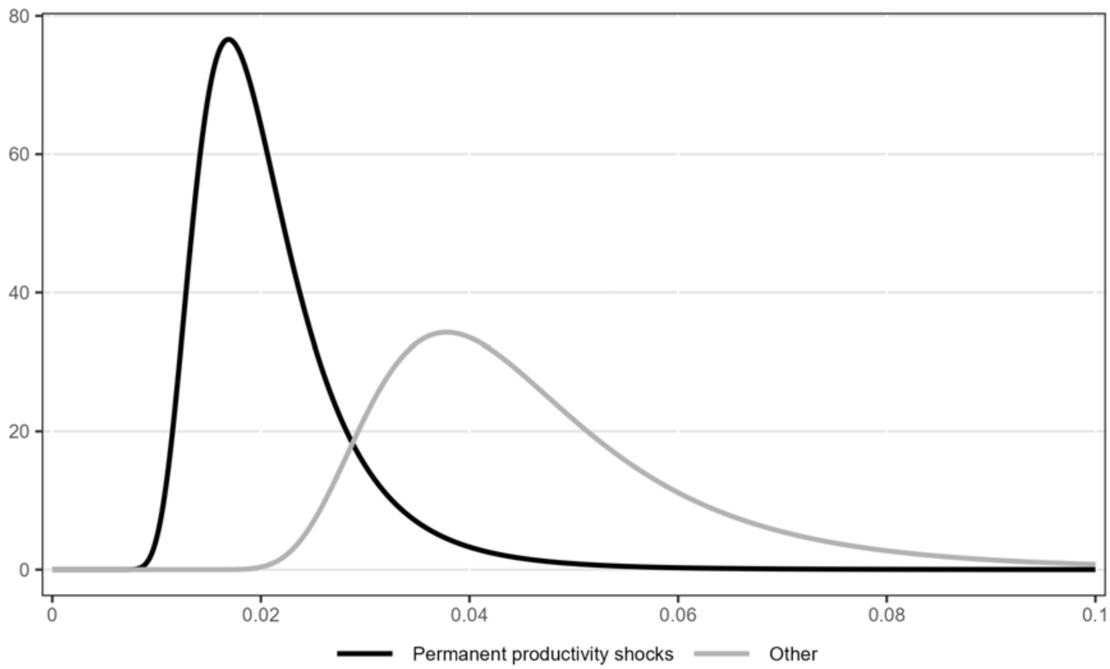

Figure A.3: Prior Densities for Standard Deviations

Figure A.4: Capital and Rents

Note: The figure plots our estimates of the evolution of the logarithm of capital, k_t, and rents, s_t. They are normalized to 0 in 1250. The black line is the mean of the posterior for each period and the gray shaded area is the 90% central posterior interval.

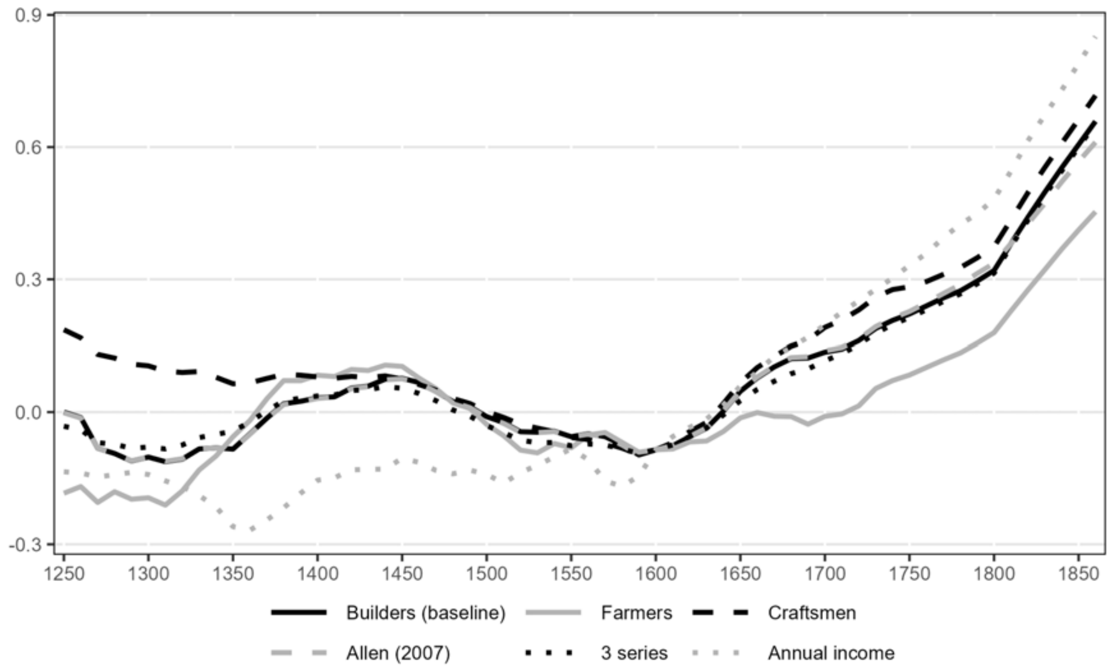

Figure A.5: Productivity using Alternative Wage Series

Note: The figure compares our baseline estimates of the evolution of the permanent component of productivity $\tilde{m}_t$ with estimates using different wage series. The "Farmers" series is the farm worker series from Clark (2010), the "Craftsmen" series is the building craftsmen series from Clark (2010), the "Allen (2007)" series uses Allen's (2007) series from 1770 onward (but our baseline wage series before that). Finally, we present estimates of productivity based on the assumption that the builders, farmers, and craftsmen series are all noisy signals of the true underlying wage. These estimates are labeled "3 series".

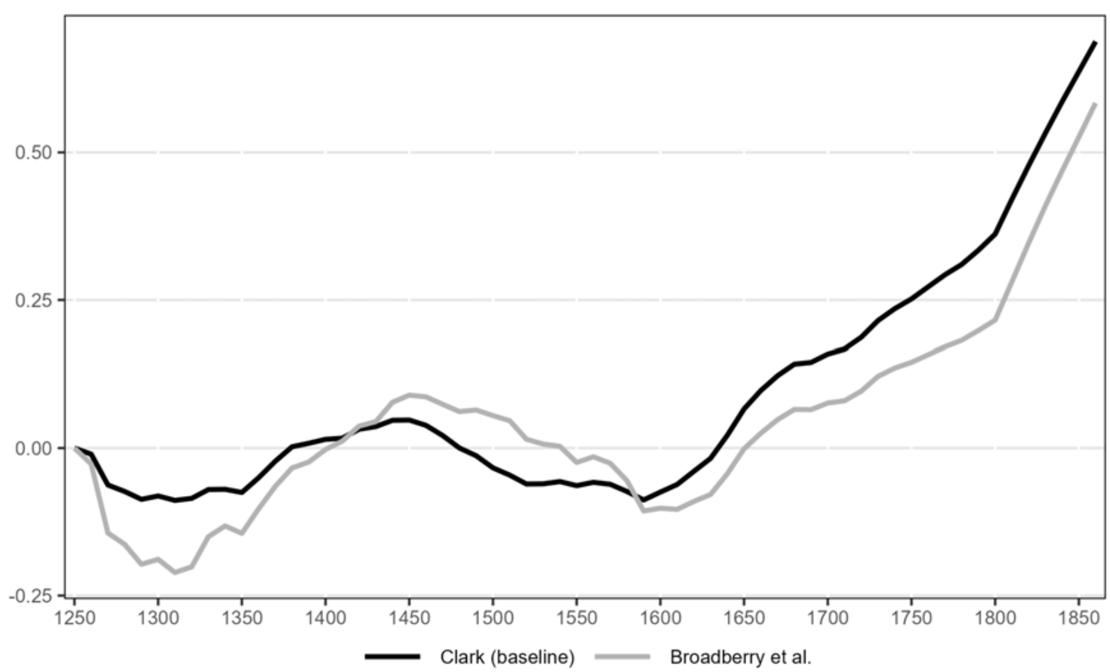

Figure A.6: Productivity using Different Population Data

Note: The figure compares our baseline estimates of the evolution of the permanent component of productivity $\tilde{m}_t$ with estimates using data on the population of England prior to 1540 from Broadberry et al. (2015).

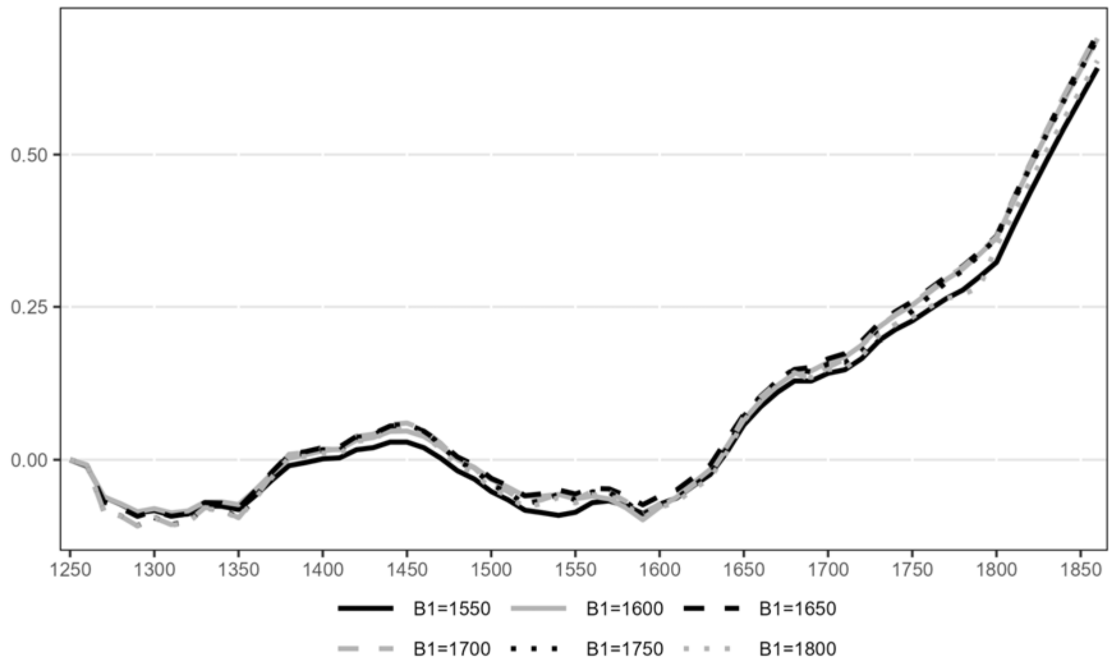

Figure A.7: Productivity Allowing for Different Break Dates

Note: The figure compares estimates of the evolution of the permanent component of productivity $\tilde{m}_t$ when we allow for different dates for the first productivity break. B1 and B2 stand for break 1.

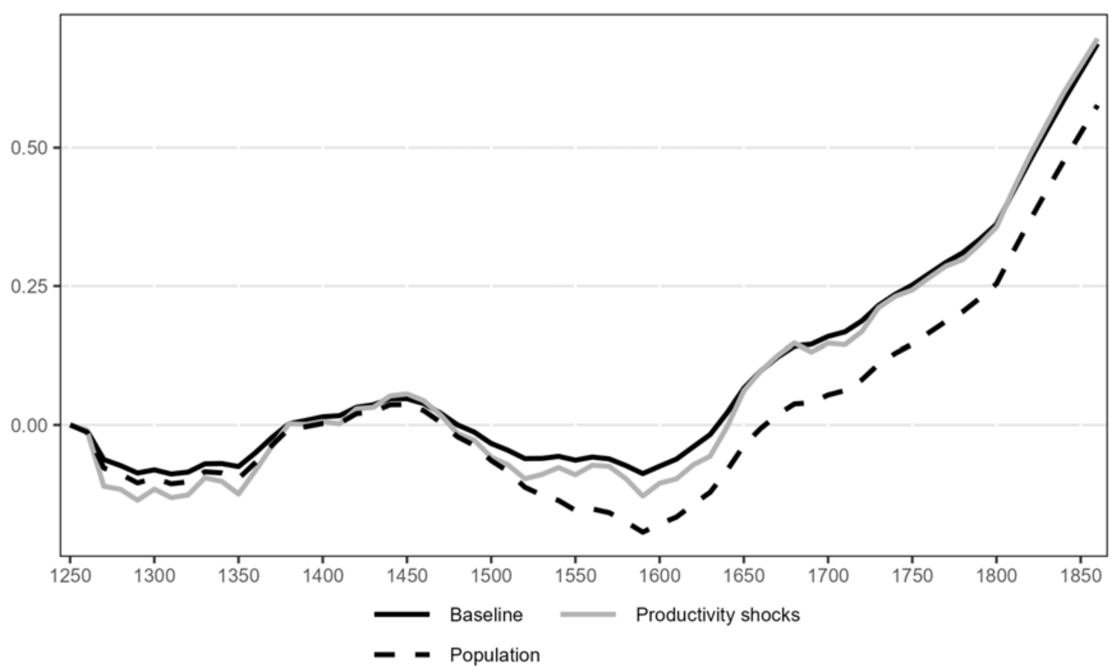

Figure A.8: Productivity using Different Priors

Note: The figure compares our baseline estimates of the evolution of the permanent component of productivity $\tilde{m}_t$ with estimates using different prior distributions. The "Productivity shocks" series changes the prior on σ_{ϵ_1} to be $\text{I}\Gamma(3, 0.005)$, i.e., the same as the prior on the other productivity and population shocks. The "Population level" series changes the prior on ψ to be $\mathcal{N}(10.86, 10.0)$.

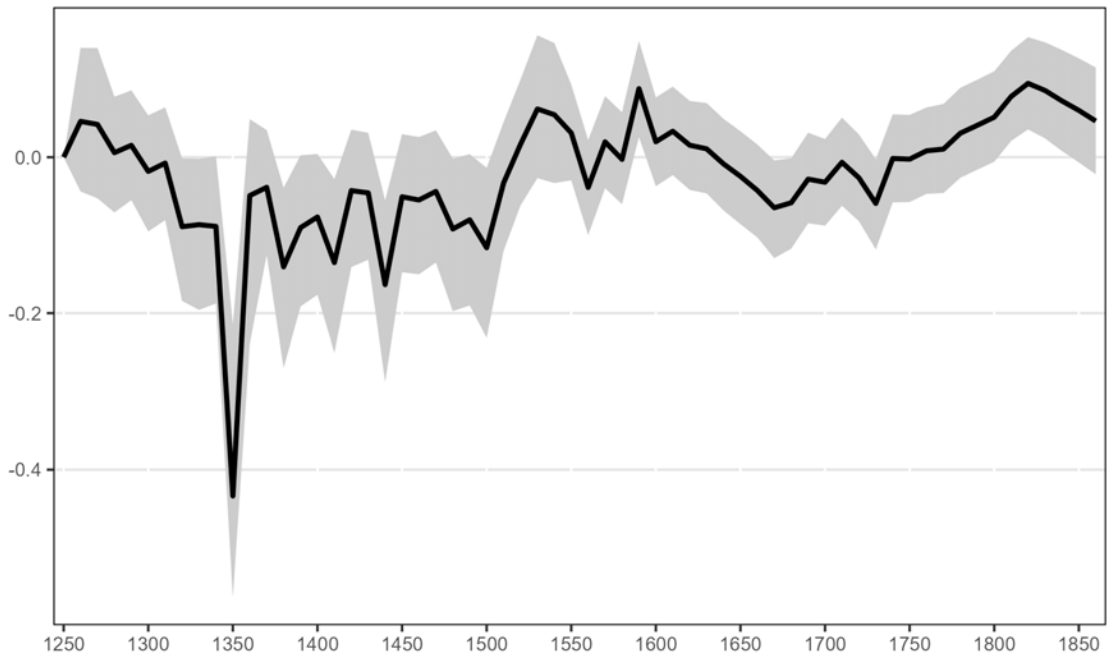

Figure A.9: Population Shocks

Note: The figure plots our estimates of the population shocks hitting the English economy over our sample period, i.e., $\xi_{1t} + \xi_{2t}$. The black line is the mean of the posterior for each period and the gray shaded area is the 90% central posterior interval.

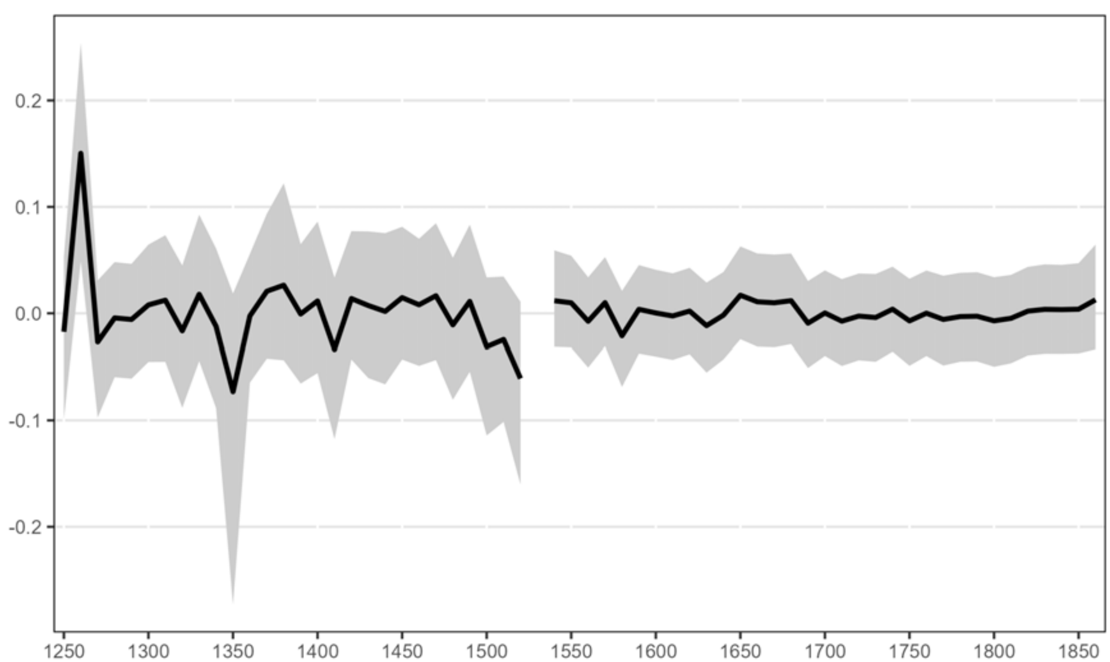

Figure A.10: Measurement Error in Population Data

Note: The figure plots our estimate of the measurement error in our population data ι_t^n.

Table A.1: Parameter Estimates—changing α, β

	Mean	St Dev	2.5%	97.5%
Main Parameters				
α	0.54	0.05	0.44	0.62
β	0.23	0.07	0.11	0.37
γ	0.03	0.05	-0.06	0.12
ω	0.03	0.02	-0.01	0.08
Productivity Shock Parameters				
$\sigma_{\epsilon_1,1}$	0.03	0.01	0.02	0.05
$\sigma_{\epsilon_1,2}$	0.02	0.01	0.01	0.03
$\sigma_{\epsilon_1,3}$	0.02	0.01	0.01	0.04
$\sigma_{\epsilon_2,1}$	0.06	0.01	0.04	0.08
$\sigma_{\epsilon_2,2}$	0.04	0.01	0.02	0.05
$\sigma_{\epsilon_2,3}$	0.04	0.01	0.02	0.06
Population Parameters				
$\pi_{t<1680}$	0.26	0.14	0.03	0.49
$\pi_{t\geq 1680}$	0.10	0.09	0.00	0.34
μ_{ξ_1}	0.82	0.08	0.58	0.90
ν_{ξ_1}	7.10	29.75	1.04	35.38
σ_{ξ_2}	0.06	0.01	0.04	0.08
Population Measurement Error Parameters				
$\sigma_{n,t<1540}$	0.04	0.01	0.02	0.06
$\sigma_{n,t\geq 1540}$	0.03	0.00	0.02	0.04
$\nu_{n,t<1540}$	11.98	265.43	1.14	43.50
$\nu_{n,t\geq 1540}$	59.63	1392.08	2.13	264.34

Note: The table presents the mean, standard deviation, 2.5% quantile, and 97.5% quantile of the posterior distribution we estimate for , using the three procedures described in sections 2–3. *Note:* The table presents the mean, standard deviation, 2.5% quantile, and 97.5% quantile of the posterior distribution we estimate for the parameters of the production function α, β, the elasticity of population growth to income γ, the subsistence wage parameter ω, the standard deviation of the permanent and transitory productivity shocks ϵ_{1t} and ϵ_{2t} in the three regimes, the probability of a plague shock π, the mean of the plague shock μ_{ξ_1}, the pseudo sample size of the plague shocks ν_{ξ_1}, the standard deviation of the normal population shock σ_{ξ_2}, the scale and degrees of freedom parameters of the population measurement error shocks, σ_n and ν_n, respectively.

References

ACEMOGLU, D., S. JOHNSON, AND J. ROBINSON (2005): "The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth," *American Economic Review*, 95, 546–579.

ACEMOGLU, D. AND P. RESTREPO (2018): "The Race Between Man and Macine: Implications of Technology for Growth, Factor Shares, and Employment," *American Economic Review*, 108, 1488–1542.

——— (2019a): "Artificial Intelligence, Automation, and Work," in *The Economics of Artificial Intelligence: An Agenda*, ed. by A. Agrawal, J. Gans, and A. Goldfarb, Chicago, IL: University of Chicago Press, 197–236.

——— (2019b): "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 33, 3–30.

ALLEN, R. C. (1992): *Enclosure and the Yeoman*, Oxford, UK: Oxford University Press.

——— (2005): "English and Welsh Agriculture, 1300-1850: Output, Inputs, and Income," Working Paper, Oxford University.

——— (2007): "Pessimism Preserved: Real Wages in the British Industrial Revolution," Oxford University Department of Economics Working Paper 314.

——— (2009a): *The British Industrial Revolution in Global Perspective*, Cambridge, UK: Cambridge University Press.

——— (2009b): "Engels' Pause: Technical Change, Capital Accumulation, and Inequality in the British Industrial Revolution," *Explorations in Economic History*, 46, 418–435.

BAILEY, R. E. AND M. J. CHAMBERS (1993): "Long-Term Demographic Interactions in Precensus England," *Journal of the Royal Statistical Society, Series A*, 156, 339–362.

BARRO, R. J. AND X. SALA-I-MARTIN (2004): *Economic Growth*, Cambridge: MIT Press, second ed.

BETANCOURT, M. (2018): "A Conceptual Introduction to Hamiltonian Monte Carlo," Working Paper, arXiv:1701.02434v2.

BLANCHARD, I. (1978): "Labor Productivity and Worker Psychology in the English Mining Industry, 1400-1600," *Economic History Review*, 31, 1–24.

BOGART, D. AND G. RICHARDSON (2011): "Property Rights and Parliament in Industrializing Britain," *Journal of Law and Economics*, 54, 241–274.

BOPPART, T., P. KIERNAN, P. KRUSELL, AND H. MALMBERG (2023): "The Macroeconomics of Intensive Agriculture," Working Paper, IIES, Stockholm University.

BROADBERRY, S., B. M. CAMPBELL, A. KLEIN, M. OVERTON, AND B. VAN LEEUWEN (2015): *British Economic Growth, 1270-1870*, Cambridge, UK: Cambridge University Press.

CAVES, D. W., L. R. CHRISTENSEN, AND W. E. DIEWERT (1982): "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity," *Econometrica*, 50, 1393–1414.

CHANEY, E. AND R. HORNBECK (2016): "Economic Dynamics in the Malthusian Era: Evidence from the 1609 Spanish Expulsion of the Moriscos," *Economic Journal*, 126, 1404–1440.

CLARK, G. (2002): "Land Rental Values and the Agrarian Economy: England and Wales, 1500-

1914," *European Review of Economic History*, 6, 281–308.

——— (2005): "The Condition of the Working Class in England, 1209-2004," *Journal of Political Economy*, 113, 1307–1340.

——— (2007a): *A Farewell to Alms*, Princeton, NJ: Princeton University Press.

——— (2007b): "The Long March of History: Farm Wages, Population, and Economic Growth, England 1209-1869," *Economic History Review*, 60, 97–135.

——— (2010): "The Macroeconomic Aggregates for England 1209-1869," *Research in Economic History*, 27, 51–140.

——— (2016): "Microbes and Markets: Was the Black Death an Economic Revolution?" *Journal of Demographic Economics*, 82, 139–165.

CLARK, G. AND P. A. LEVIN (2011): "How Different Was the Industrial Revolution? The Revolution in Printing, 1350-1835," Unpublished Manuscript, University of California, Davis.

CLARK, G. AND Y. VAN DER WERF (1998): "Work in Progress? The Industrious Revolution," *Journal of Economic History*, 58, 830–843.

CRAFTS, N. AND T. C. MILLS (2009): "From Malthus to Solow: How Did the Malthusian Economy Really Evolve?" *Journal of Macroeconomics*, 31, 68–93.

CRAFTS, N. F. (1983): "British Economic Growth, 1700-1831: A Review of the Evidence," *Economic History Review*, 36, 177–199.

——— (1985): *British Economic Growth During the Industrial Revolution*, Oxford, UK: Clarendon Press.

CRAFTS, N. F. AND C. K. HARLEY (1992): "Output Growth and the British Industrial Revolution: A Restatement of the Crafts-Harley View," *Economic History Review*, 45, 703–730.

CRESSY, D. (1980): *Literacy and the Social Order*, Cambridge, UK: Cambridge University Press.

DE MOOR, T. AND J. L. VAN ZANDEN (2010): "Girl Power: The European Marriage Pattern and Labor Markets in the North Sea Region in the Labe Medieval and Early Modern Period," *Economic History Review*, 63, 1–33.

DE VRIES, J. (1994): "The Industrial Revolution and the Industrious Revolution," *Journal of Economic History*, 54, 249–270.

——— (2008): *The Industrious Revolution: Consumer Behavior and the Household Economy, 1650 to the Present*, Cambridge, UK: Cambridge University Press.

DIEWERT, W. E. AND A. O. NAKAMURA (2007): "The Measurement of Productivity for Nations," in *Handbook of Econometrics, Volume 6, Part A*, ed. by J. J. Heckman and E. E. Leamer, Amsterdam, Holland: Elsevier, 4501–4586.

DITTMAR, J. E. (2011): "Information Technology and Economic Change: The Impact of the Printing Press," *Quarterly Journal of Economics*, 126, 1133–1172.

ENGELS, F. (1845): *The Condition of the Working Class in England*, Oxford, UK: Oxford University Press, 2009, (Originally published in German by Otto Wigand in Leigzig).

FÄRE, R., S. GROSSKOPF, M. NORRIS, AND Z. ZHANG (1994): "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries," *American Economic Review*, 84, 66–83.

FEINSTEIN, C. H. (1988): "National Statistics, 1760-1920," in *Studies in Capital Formation in the United Kingdom, 1750-1920*, ed. by C. H. Feinstein and S. Pollard, Oxford, UK: Oxford University Press, 259–471.

——— (1998): "Pessimism Perpetuated: Real Wages and the Standard of Living in Britain during and after the Industrial Revolution," *Journal of Economic History*, 58, 625–658.

GALOR, O. AND D. N. WEIL (2000): "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review*, 90, 806–828.

GELMAN, A., J. B. CARLIN, H. S. STERN, D. B. DUNSON, A. VEHTARI, AND D. B. RUBIN (2013): *Bayesian Data Analysis, Third Edition*, CRC Press.

GOLDSTONE, J. A. (2002): "Efflorescences and Economic Growth in World History: Rethinking the "Rise of the West" and the Industrial Revolution," *Journal of World History*, 13, 323–389.

HANSEN, G. D., L. E. OHANIAN, AND F. OZTURK (2020): "Dynamic General Equilibrium Modelling of Long and Short-Run Historical Events," NBER Working Paper No. 28090.

HANSEN, G. D. AND E. C. PRESCOTT (2002): "Malthus to Solow," *American Economic Review*, 92, 1205–1217.

HARLEY, C. K. (1982): "British Industrialization Before 1841: Evidence of Slower Growth During the Industrial Revolution," *Journal of Economic History*, 42, 267–289.

HELDRING, L., J. A. ROBINSON, AND S. VOLLMER (2021): "The Long-Run Impact of the Dissolution of the English Monasteries," *Quarterly Journal of Economics*, 136, 2093–2145.

HILL, C. (1940): *The English Revolution 1640*, London, UK: Lawrence and Wishart Ltd.

——— (1961): *The Century of Revolution*, New York, NY: Van Nostrand Reinhold (International) Co. Ltd.

HOUSTON, R. A. (1982): "The Development of Literacy: Northern England, 1640-1750," *Economic History Review*, 35, 199–216.

HSIEH, C.-T. (2002): "What Explains the Industrial Revolution in East Asia? Evidence From the Factor Markets," *American Economic Review*, 92, 502–525.

HUMPHRIES, J. AND J. WEISDORF (2019): "Unreal Wages? Real Income and Economic Growth in England, 1260-1850," *Economic Journal*, forthcoming.

JONES, C. I. (2001): "Was an Industrial Revolution Inevitable? Economic Growth Over the Very Long Run," *Advances in Macroeconomics*, 1, 1–43.

KREMER, M. (1993): "Population Growth and Technological Change: One Million B.C. to 1990," *Quarterly Journal of Economics*, 108, 681–716.

LEE, R. AND M. ANDERSON (2002): "Malthus in State Space: Macro Economic-Demographic Relationships in English History, 1540-1870," *Journal of Population Economics*, 15, 195–220.

MALMQUIST, S. (1953): "Index Numbers and Indifference Surfaces," *Trabajos de estadística*, 4, 209–242.

MALTHUS, T. R. (1798): *An Essay on the Principle of Population*, London, UK: J. Johnson in St. Paul's Church-yard, (Originally published anonymously).

MARX, K. (1867): *Capital. A Critique of Political Economy*, vol. 1, Chicago, IL: Charles H. Kerr & Company, 1932, English translation by Samuel Moore and Edward Aveling.

MARX, K. AND F. ENGELS (1848): *The Communist Manifesto*, London, UK: Pluto Press, 2008, English translation by Samuel Moore in cooperation with Friedrich Engels, 1888.

MCCLOSKEY, D. (2006): *The Bourgeois Virtues: Ethics for an Age of Commerce*, Chicago, IL: University of Chicago Press.

——— (2010): *Bourgeois Dignity: Why Economics Can't Explain the Modern World*, Chicago, IL: University of Chicago Press.

——— (2016): *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*, Chicago, IL: University of Chicago Press.

MOKYR, J. (2009): *The Enlightened Economy: The Economic History of Britain, 1700-1850*, New Haven, CT: Yale University Press.

——— (2016): *A Culture of Growth: The Origins of the Modern Economy*, Princeton, NJ: Princeton University Press.

NORTH, D. C. AND B. R. WEINGAST (1989): "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth Century England," *Journal of Economic History*, 49, 803–832.

PERSSON, K. G. (2008): "The Malthus Delusion," *European Review of Economic History*, 12, 165–173.

SOLOW, R. M. (1957): "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, 39, 312–330.

STAN DEVELOPMENT TEAM (2017): "Stan Modeling Language Users Guide and Reference Manual," Version 2.17.0, https://mc-stan.org.

TAWNEY, R. H. (1926): *Religion and the Rise of Capitalism*, New York, NY: Routledge, 2017.

VOIGTLÄNDER, N. AND H.-J. VOTH (2013): "How the West "Invented" Fertility Restriction," *American Economic Review*, 103, 2227–2264.

VOTH, H.-J. (2000): *Time and Work in England, 1750-1830*, Oxford, UK: Oxford University Press.

——— (2001): "The Longest Years: New Estimates of Labor Input in England, 1760-1830," *Journal of Economic History*, 61, 1065–1082.

WEBER, M. (1905): *Protestant Ethic and the Spirit of Capitalism*, New York, NY: Penguin Books, 2002, English translation by Peter Baehr and Gordon C. Wells.

WRIGLEY, E. A. (2010): *Energy and the English Industrial Revolution*, Cambridge, UK: Cambridge University Press.

WRIGLEY, E. A., R. S. DAVIES, J. E. OEPPEN, AND R. S. SCHOFIELD (1997): *English Population History from Family Reconstruction, 1580-1837*, Cambridge, UK: Cambridge University Press.

WRIGLEY, E. A. AND R. S. SCHOFIELD (1981): *The Population History of England 1541-1871*, Cambridge, UK: Cambridge University Press.