Appendix to:

When Did Growth Begin? New Estimates of Productivity Growth in England from 1250 to 1870

A Appendix Figures and Tables

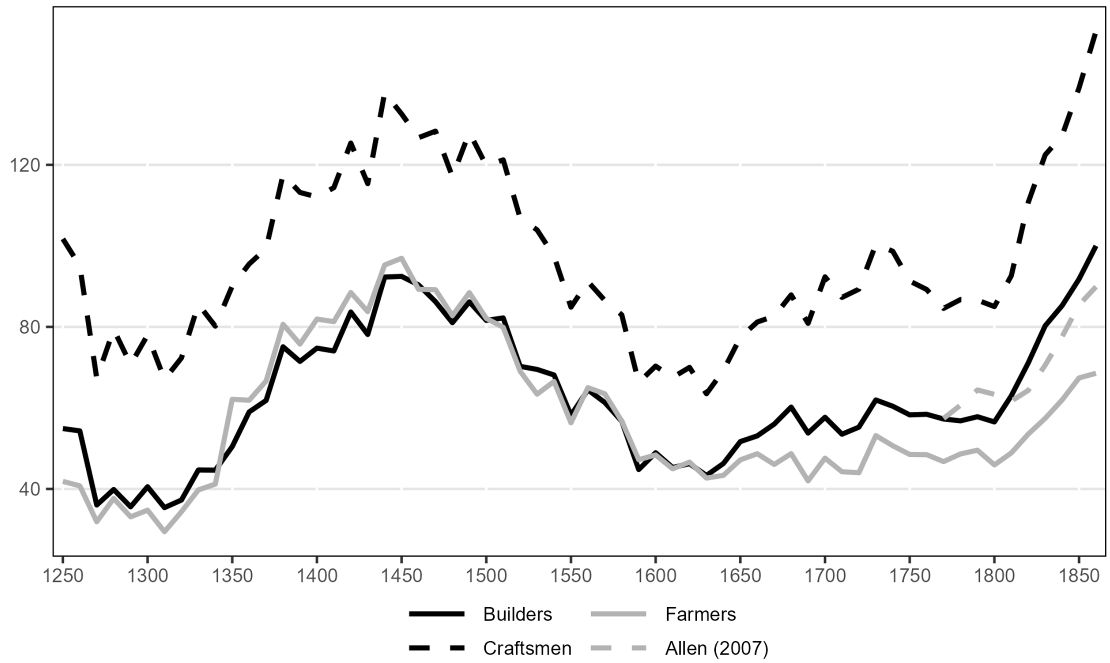

Figure A.1: A Comparison of Real Wage Measures in England, 1250-1860

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.

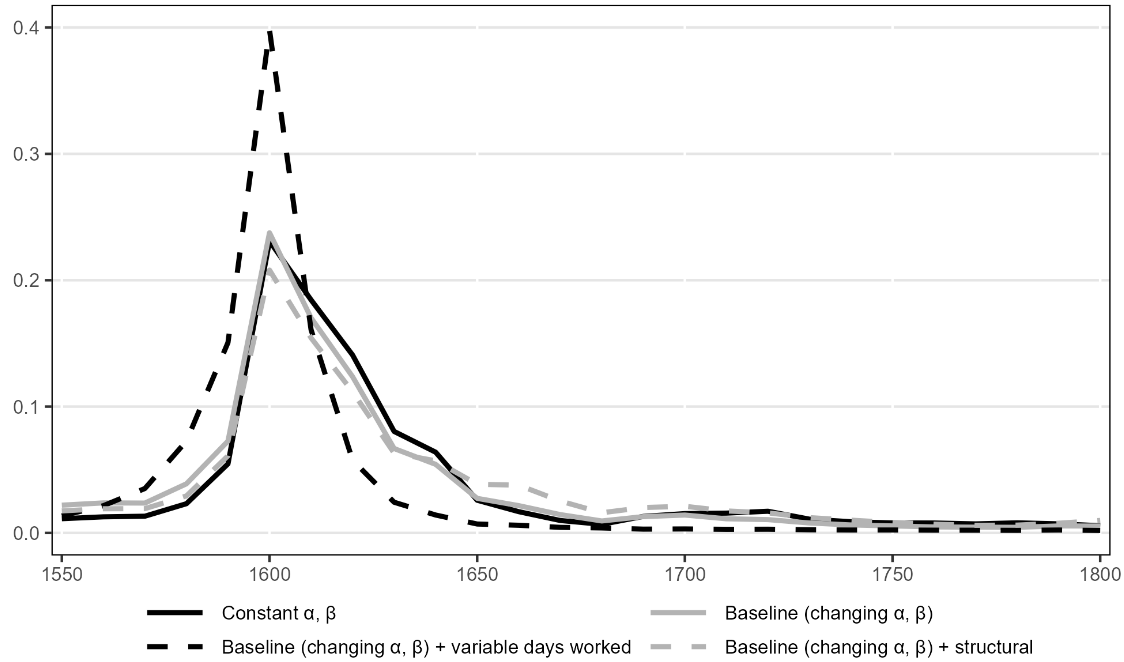

Figure A.2: Probability of Breaks Dates for Different Specifications

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.

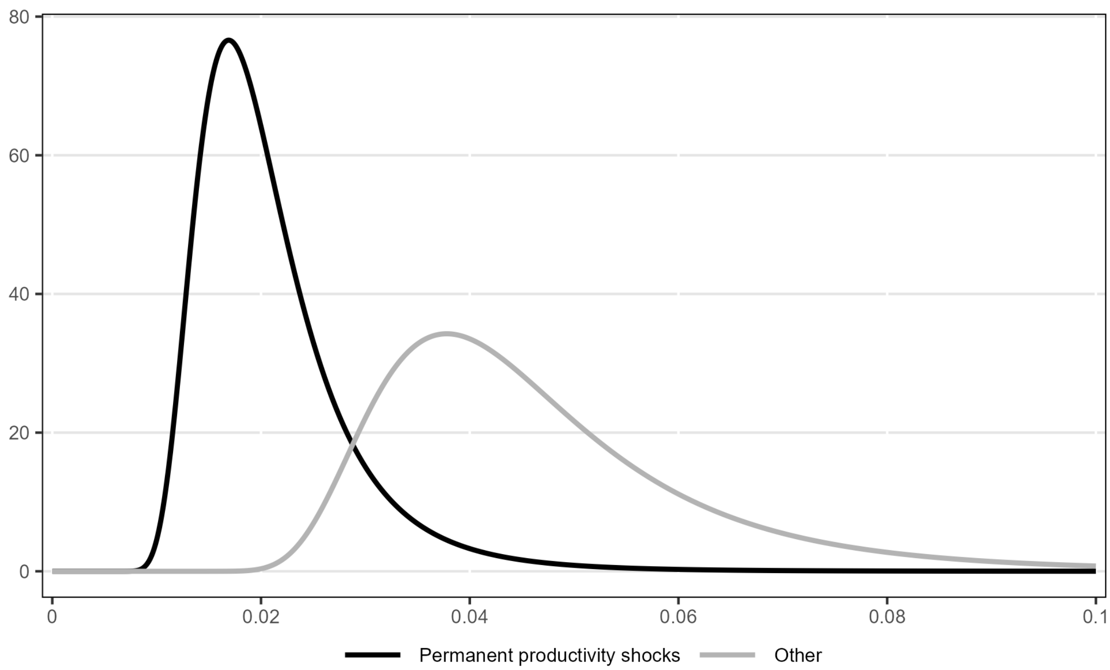

Figure A.3: Prior Densities for Standard Deviations

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.

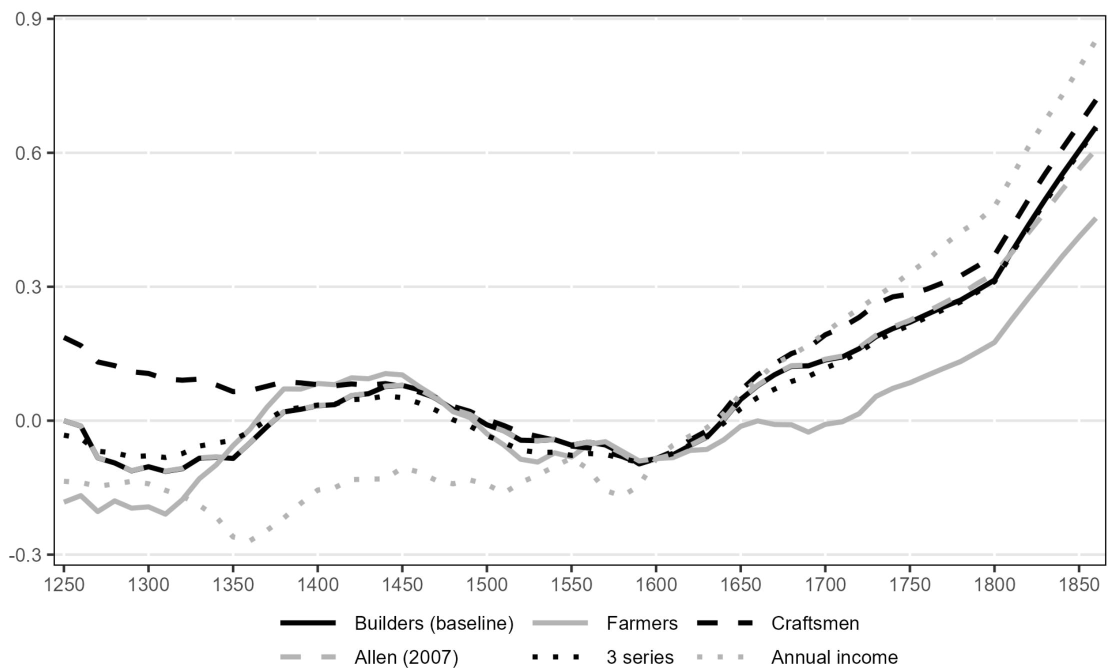

Figure A.5: Productivity using Alternative Wage Series

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".

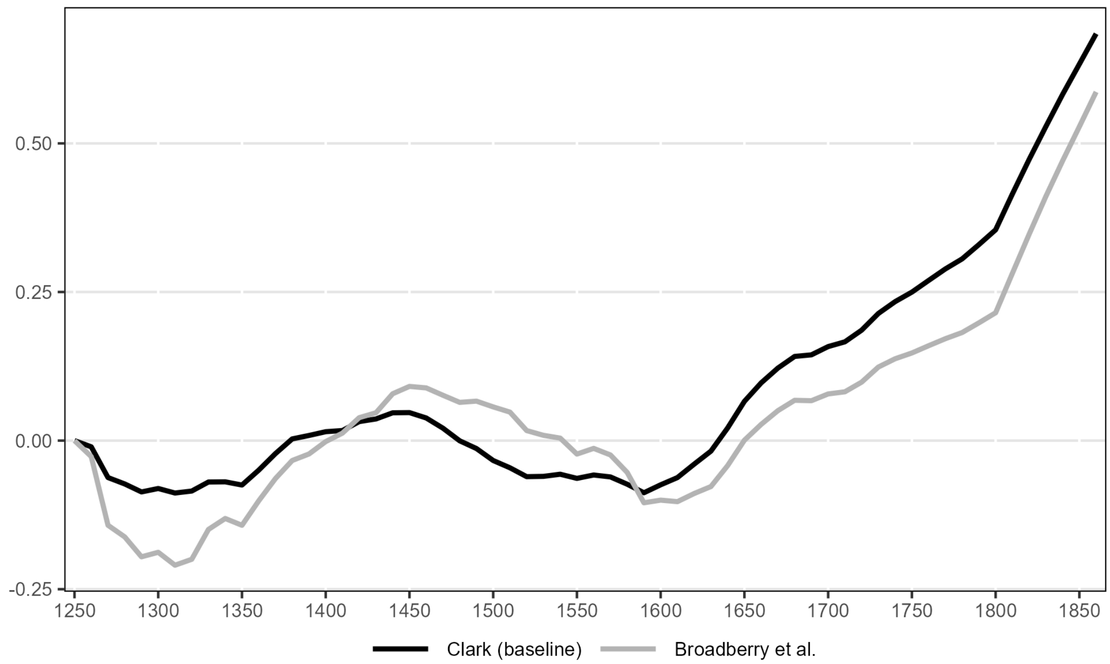

Figure A.6: Productivity using Different Population Data

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).

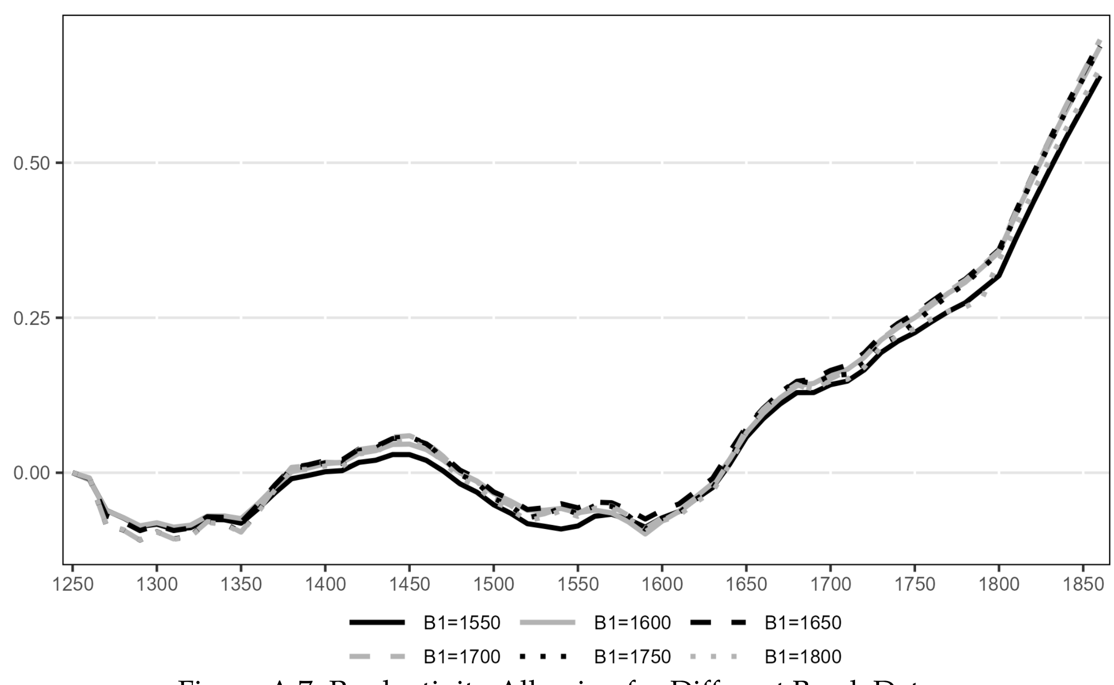

Figure A.7: Productivity Allowing for Different Break Dates

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.

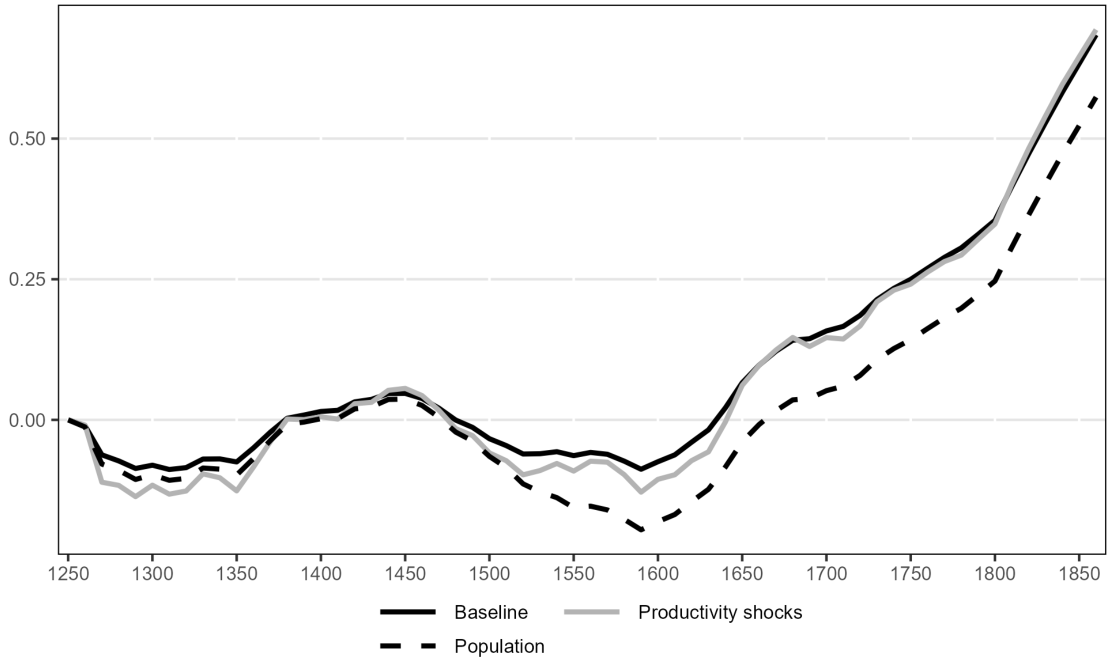

Figure A.8: Productivity using Different Priors

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 $\mathrm{I}\Gamma(3, 0.005)$, 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 $\mathcal{N}(10.86, 10.0)$.

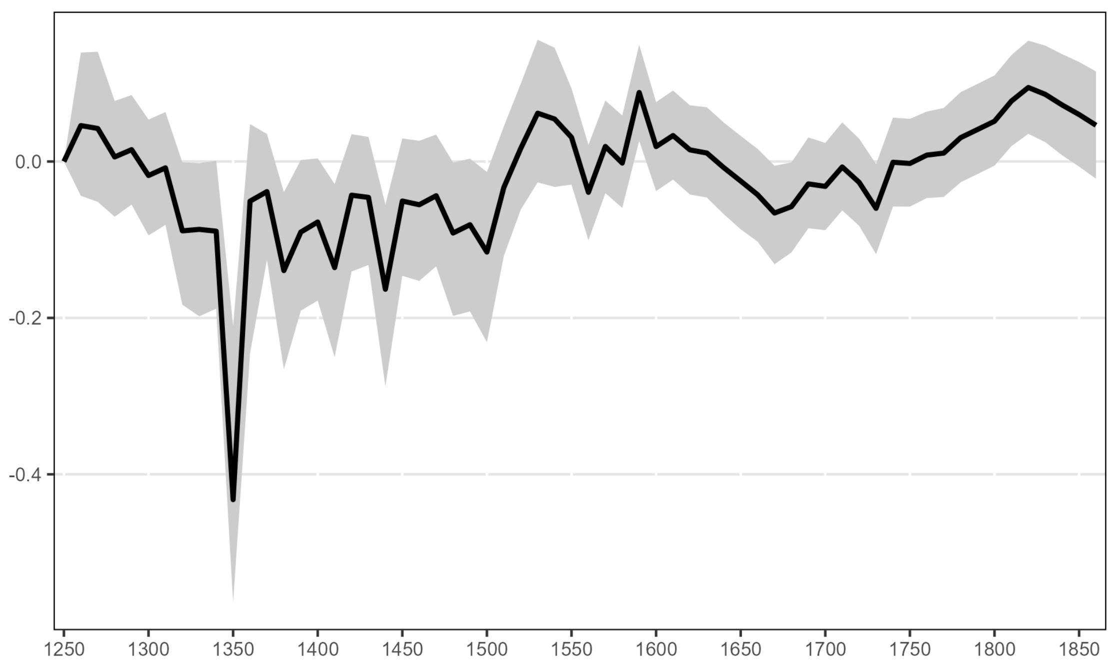

Figure A.9: Population Shocks

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.

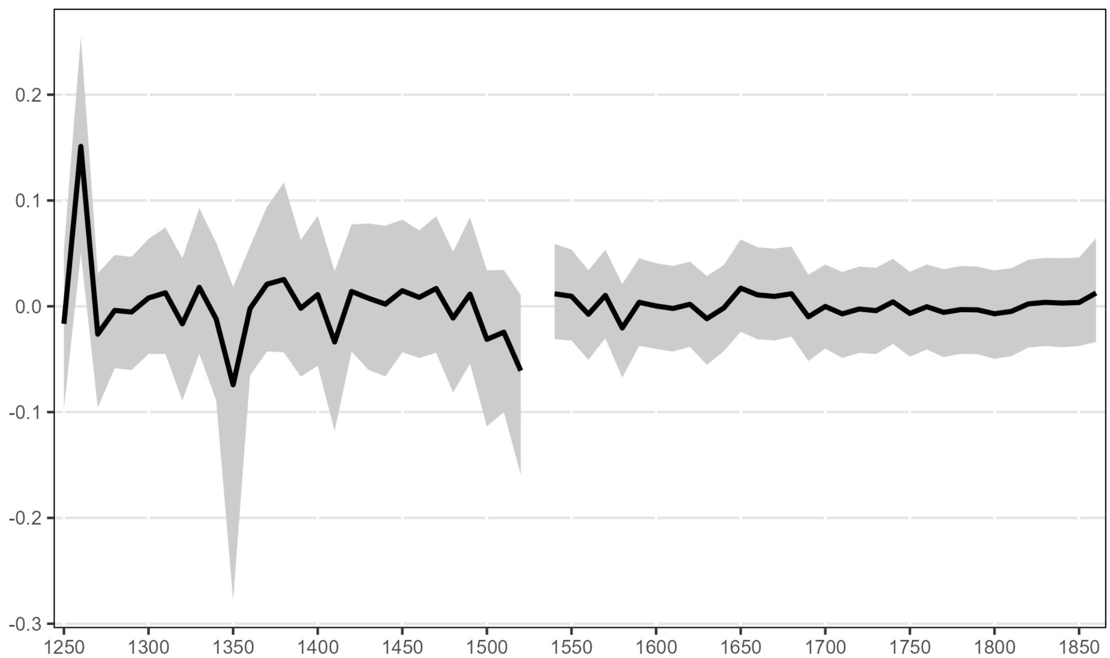

Figure A.10: Measurement Error in Population Data

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
$\pi_{t\geq1680}$	0.10	0.09	0.00	0.35
μ_{ξ_1}	0.82	0.08	0.59	0.90
ν_{ξ_1}	7.18	29.07	1.04	36.32
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
$\sigma_{n,t\geq1540}$	0.03	0.00	0.02	0.04
$\nu_{n,t<1540}$	18.45	1127.28	1.14	42.78
$\nu_{n,t\geq1540}$	74.50	1984.14	2.13	258.69

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.

B 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 he 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.

C CES Production Function

Consider the production function

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

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.

D 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.

E 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

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

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

and we then also have that

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

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 $(1 - X_t)/(1 - X_t Y_t)$ with respect to X_t is $-(1 - Y_t)/(1 - X_t Y_t)^2$, 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]

F 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, $D_t(X_t, Y_t) < 1$.)

Next consider $D_t(X_{t+1}, Y_{t+1})$, i.e., the distance of the input-output vector at time $t+1$ from

[22] The total derivative of $1-\alpha_t-\beta_t$ with respect to W_t is: $-\left((1-Y_t)^2 \times \partial X_t/\partial W_t + (1-X_t)^2 \times \partial Y_t/\partial W_t\right)/(1-X_tY_t)^2$. 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.

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.

F.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 where 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.

G Model Equations

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

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

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

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

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.

H 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.

H.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 H.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 H.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.

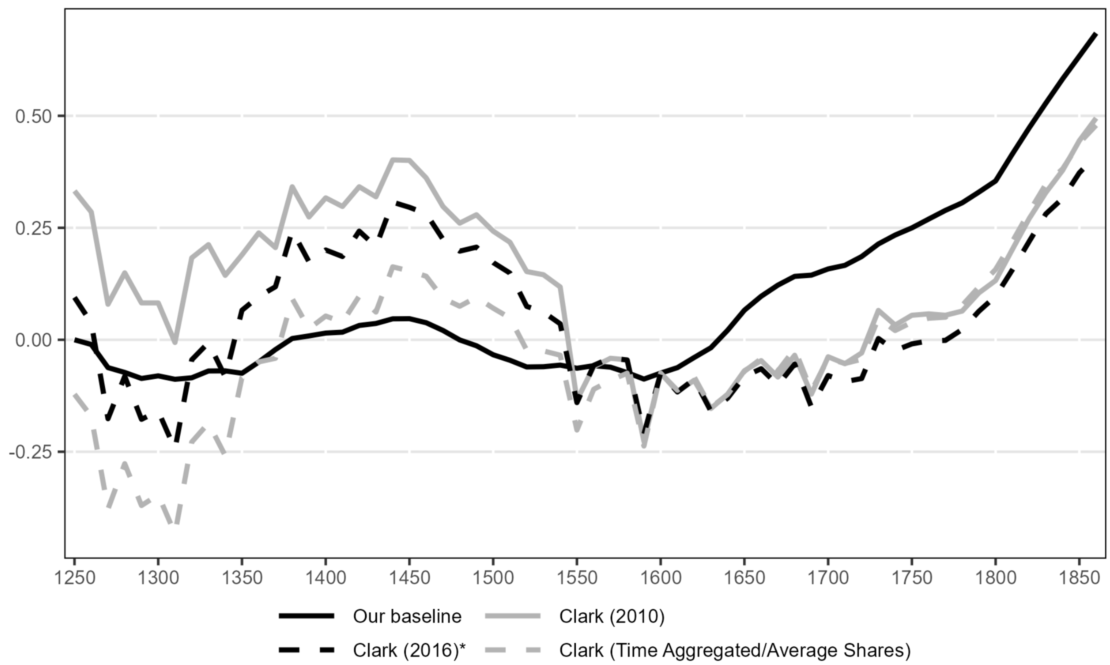

Figure H.1: A Comparison between Clark (2010) and Clark (2016)

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 H.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.

H.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:

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

[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 H.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 H.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 H.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 elasticities, 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 H.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.

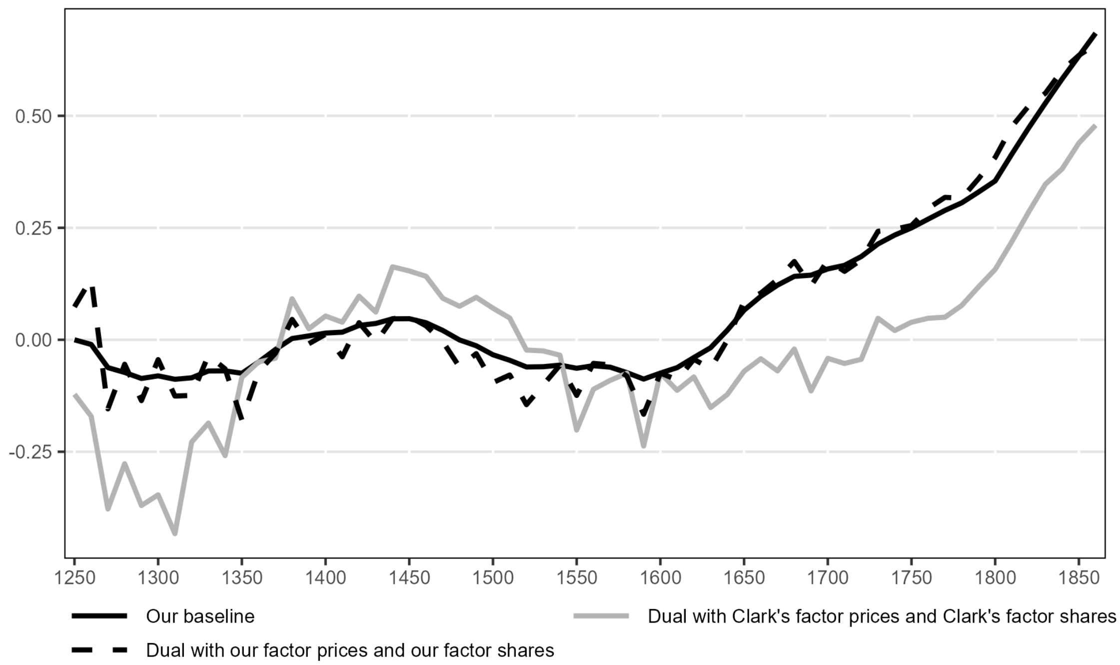

Figure H.2: Our Productivity Measure Compared with the Dual Approach

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 H.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.

H.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 H.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 H.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

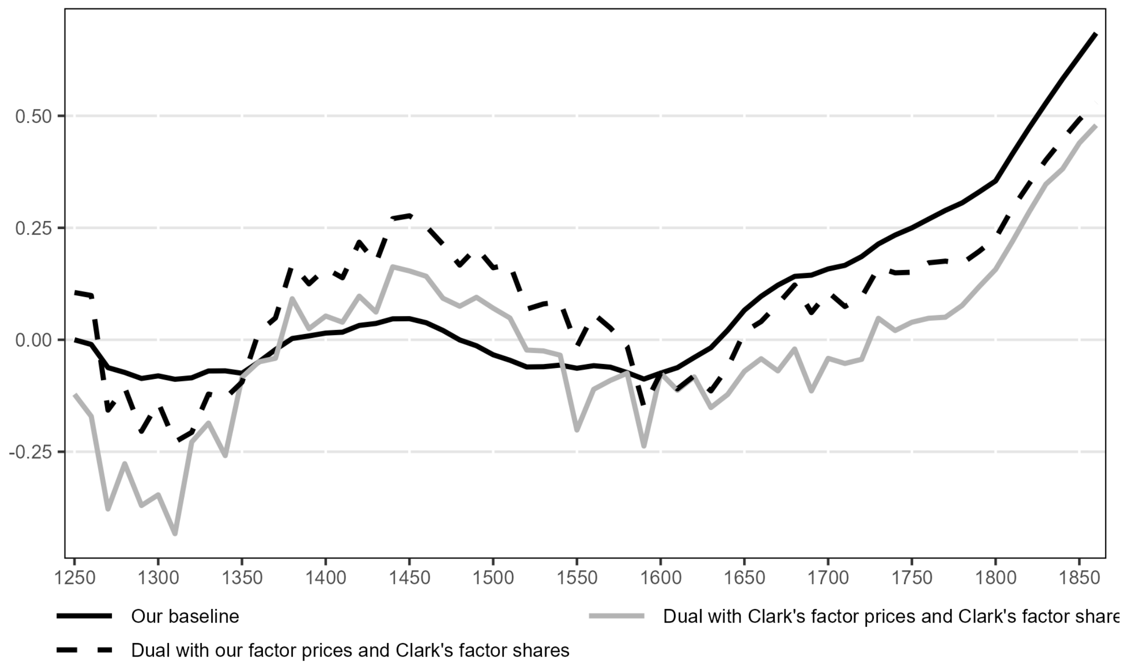

Figure H.3: Contribution of Factor Prices to Differences in Productivity Estimates

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 H.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 H.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 H.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

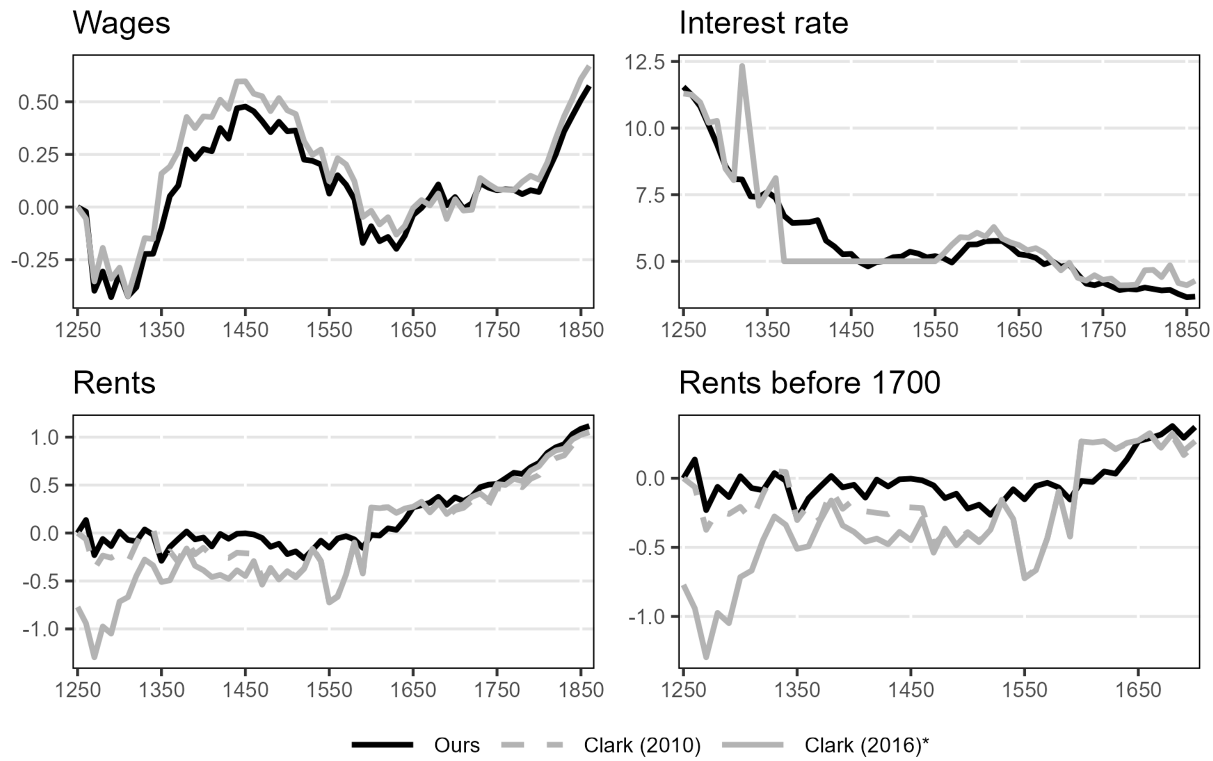

Figure H.4: A Comparison of Price Series

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 45%, while the 2010 series falls by 32%.

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

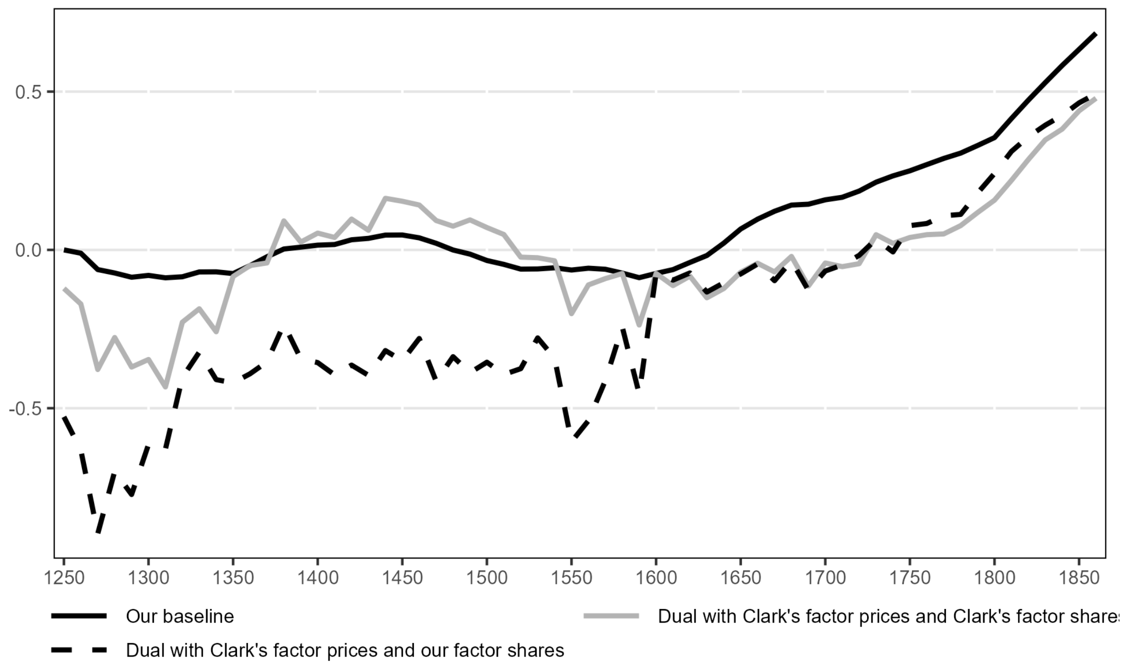

Figure H.5: Contribution of Factor Shares to Differences in Productivity Estimates

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 H.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 H.1, H.2, H.3, and H.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 H.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 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 (educated

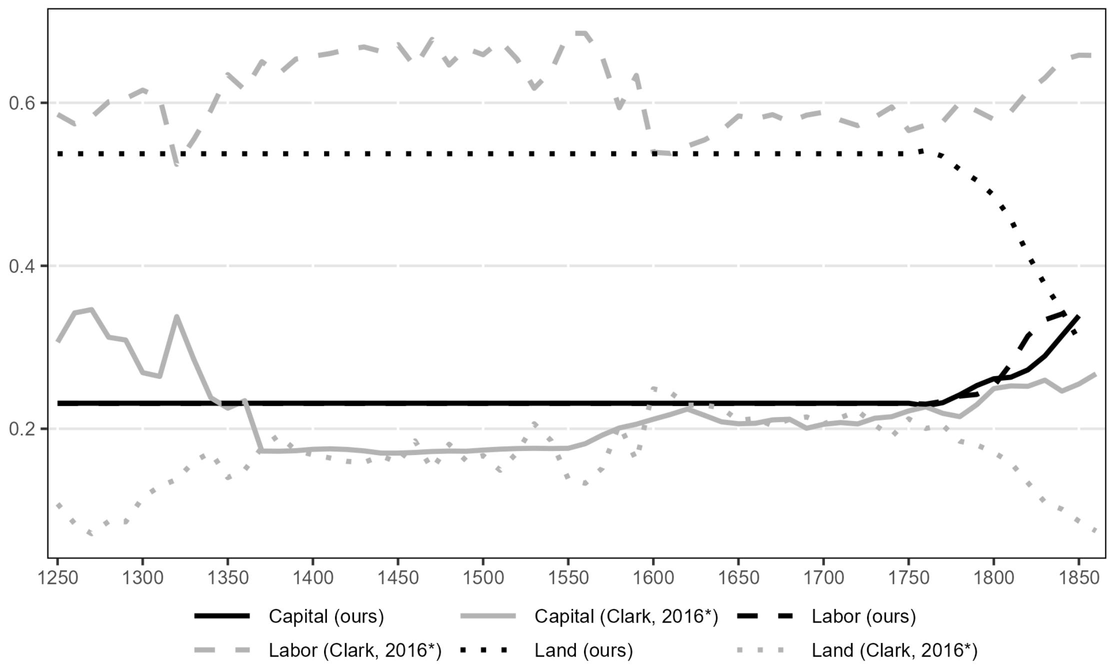

Figure H.6: Factor Shares

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.

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).

I Impulse Response Functions

I.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

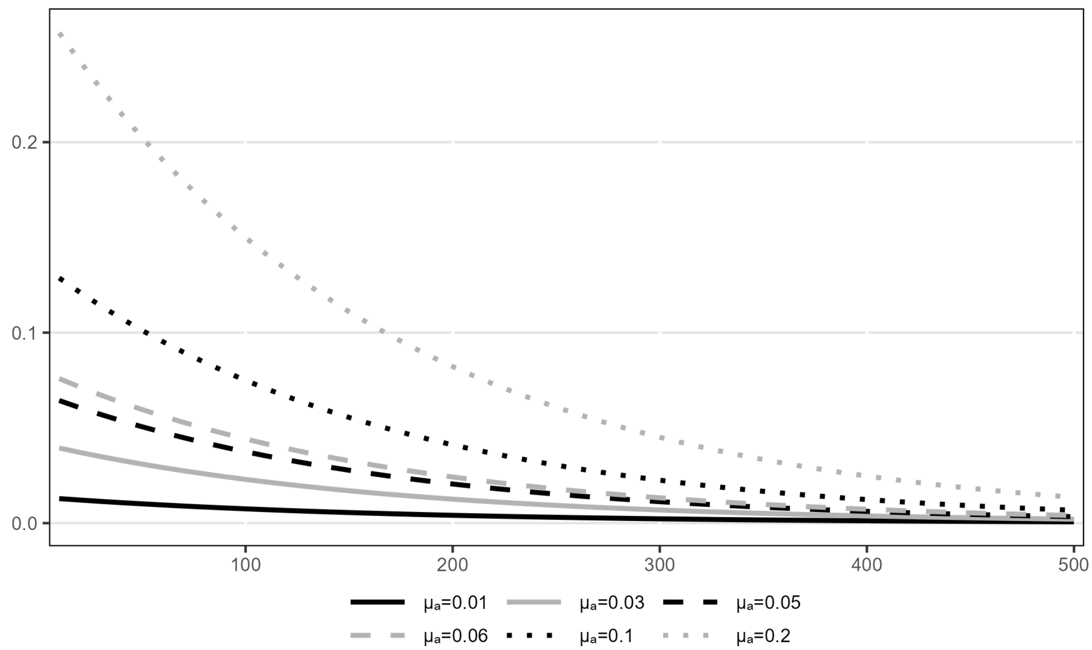

Figure I.1: Real Wage Growth After an Increase in Productivity Growth

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 0.49.

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 I.1 and I.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 ($\alpha = 0.49$). 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 I.1, we show the evolution of the growth rate of wages (log change) over the subsequent 500 years. In Figure I.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 I.1, we see that the growth rate of wages is initially equal to the change in produc-

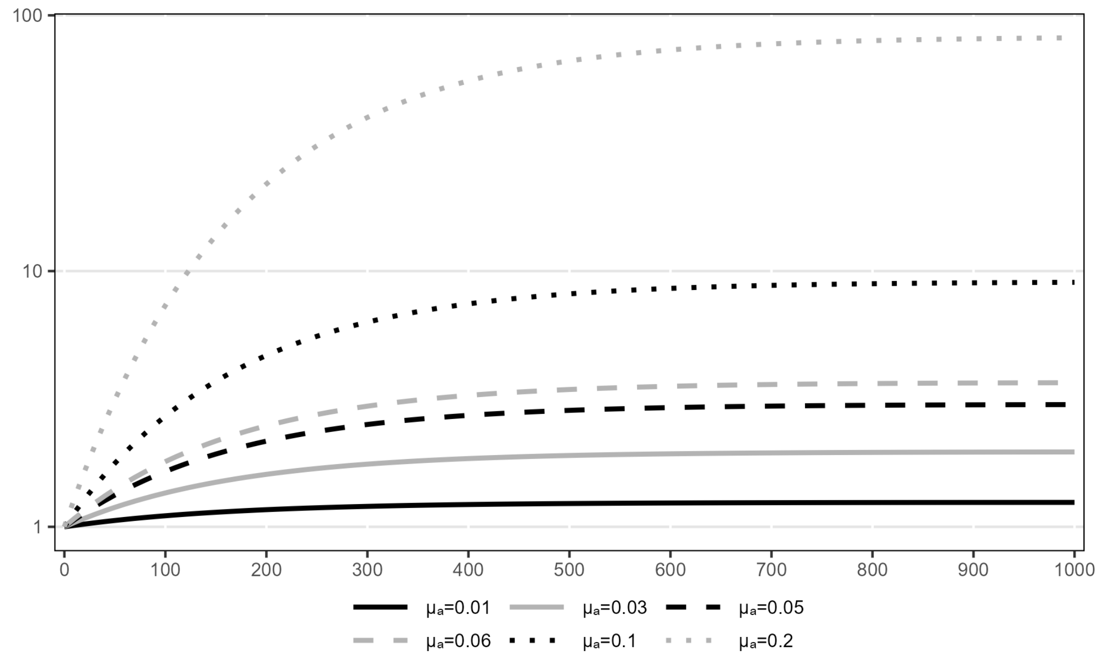

Figure I.2: Evolution of Real Wages After an Increase in Productivity Growth

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 0.49.

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 115 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 I.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.

I.2 Dynamics 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 I.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:

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

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 I.3 and I.4. We set α^H, the value of α before time 0, to 0.49, which is the posterior mean before 1770 and show the results for various values of α^L. The lowest one, 0.22, 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

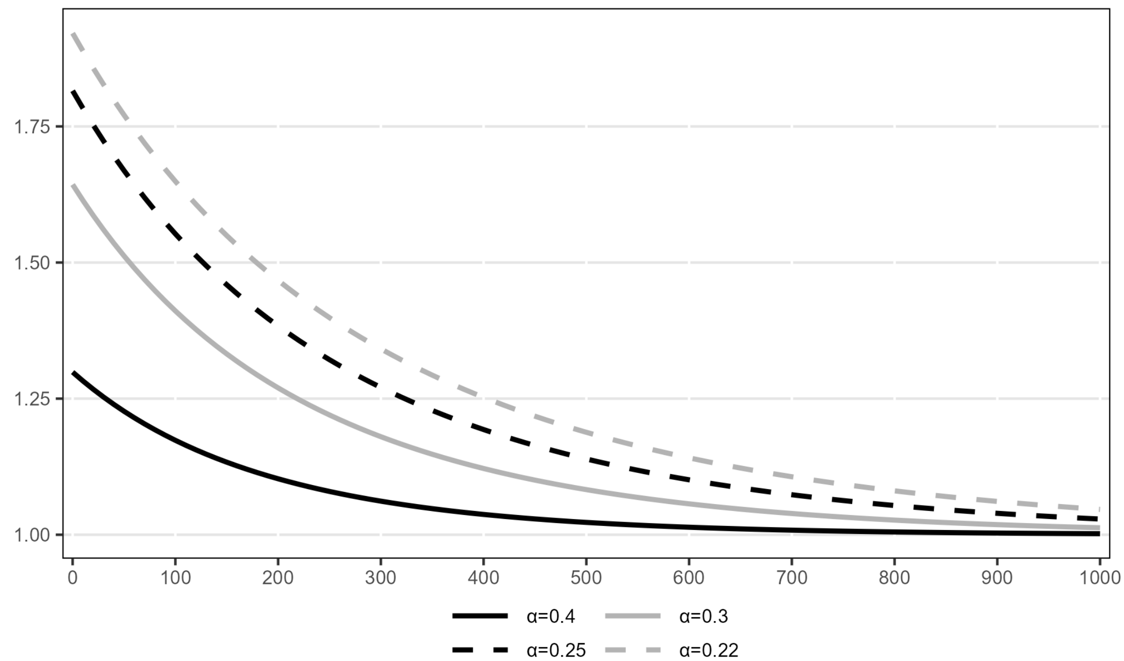

Figure I.3: Response of W_t to a Change in α

Note: The figure plots the response of the real wage (W_t) to a drop in α from its posterior mean before 1770 (0.49) 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 0.49 to 0.22 implies a 92% 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:

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

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.

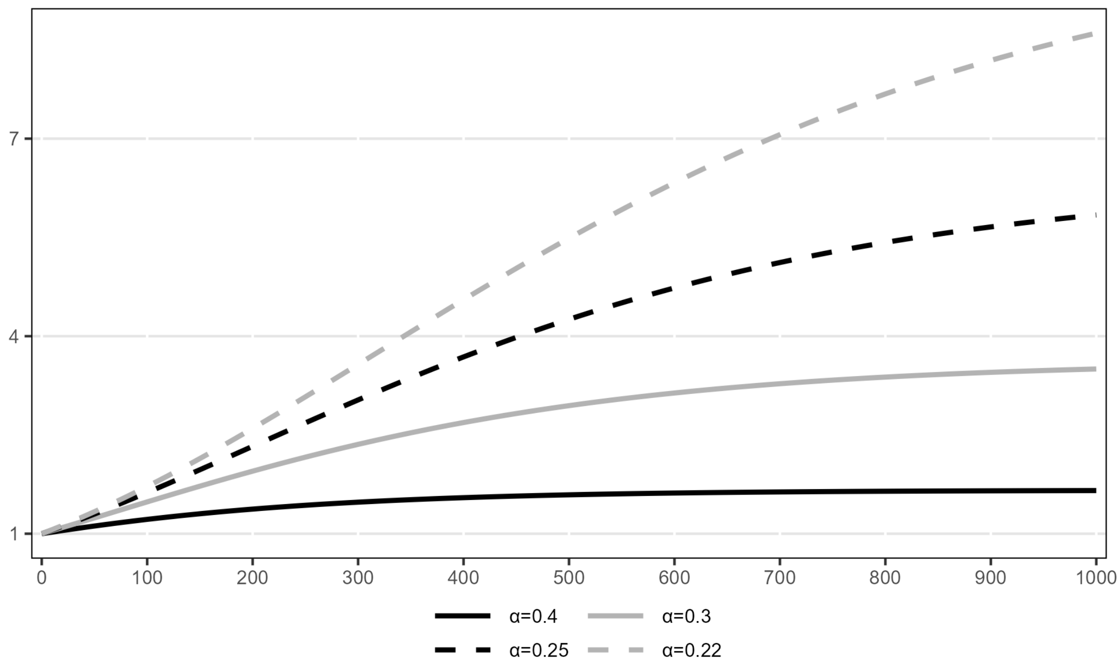

Figure I.4: Response of N_t to a Change in α

Note: The figure plots the response of population (N_t) to a drop in α from its posterior mean before 1770 (0.49) to the value in the legend. N_t is expressed in multiple of its steady state value before the drop.