

Directed Technical Change and the British Industrial Revolution

David I. Stern, John C. V. Pezzey, Yingying Lu

Abstract: We build a directed technical change model where one intermediate goods sector uses a fixed quantity of biomass energy (“wood”) and another uses coal at a fixed price, matching stylized facts for the British Industrial Revolution. Unlike previous research, we do not assume the level or growth rate of productivity is inherently higher in the coal-using sector. Analytically, greater initial wood scarcity, initial relative knowledge of coal-using technologies, and/or population growth will boost an industrial revolution, while the converse may prevent one forever. An industrial revolution, with eventual dominance by the coal-using sector, is the model’s main dynamic outcome, but not inevitable if inter-good substitutability is high enough. Empirical calibration for 1560–1900 produces historically plausible results for changes in energy-related variables during British industrialization, and through counterfactual simulations confirms that it was the growing relative scarcity of wood caused by population growth that resulted in innovation to develop coal-using machines.

JEL Codes: N13, N73, O33, O41, Q43

Keywords: economic growth, economic history, energy, coal, structural change

DID COAL PLAY A VITAL ROLE in the acceleration of British economic growth known as the Industrial Revolution? Economists and historians are divided on the importance of coal in fueling the increase in the rate of economic growth. Many researchers (e.g., Wilkinson 1973; Wrigley 1988, 2010; Pomeranz 2001; Allen 2009; Kander et al. 2014;

David I. Stern (corresponding author) is in the Crawford School of Public Policy, Australian National University, 132 Lennox Crossing, Acton, ACT 2601, Australia (<https://orcid.org/0000-0001-6595-4268>; david.stern@anu.edu.au). John C. V. Pezzey is in the Fenner School of Environment and Society, Australian National University, 141 Linnaeus Way, Acton, ACT 2601, Australia (<http://orcid.org/0000-0001-6473-9355>). Yingying Lu is in the School of Economics, Shanghai University, China, and CSIRO, Black Mountain, ACT 2601, Australia. We thank Prema-Chandra Athukorala, Ed Barbier, Timothy Considine, Roger Fouquet, Daniel Garcia-Macia, Nick Hanley, Wei Jin, Brooks Kaiser, Astrid Kander, Bruno Lanz, Akshay

Received April 29, 2019; Accepted March 26, 2021; Published online August 27, 2021.

Journal of the Association of Environmental and Resource Economists, volume 8, number 6, November 2021.
© 2021 The Association of Environmental and Resource Economists. All rights reserved. Published by The University of Chicago Press for The Association of Environmental and Resource Economists.
<https://doi.org/10.1086/714602>

Gars and Olovsson 2019) argue that innovations in the use, and growth in the quantity consumed, of coal played a crucial role in driving the Industrial Revolution. By contrast, some economic historians (e.g., Clark and Jacks 2007; Kunnas and Myllyntaus 2009) and economists (e.g., Madsen et al. 2010) either argue that it was not necessary to expand the use of modern energy carriers such as coal or do not give coal a central role (e.g., Clark 2014). This debate matters not just for understanding the history of economic development but also for assessing the prospects for a global energy transition from fossil fuels to renewables in order to avoid dangerous climate change.

We develop a model that shows both analytically and numerically how the scarcity of biomass energy (referred to here as wood, which includes both firewood and charcoal) relative to coal in Britain could have directed technical change toward the development of coal-using technologies, resulting in an increase in the rate of economic growth. Our baseline empirical model reproduces several stylized facts of the British Industrial Revolution during 1560–1900, without assuming that the level or growth rate of productivity is inherently higher in the coal-using sector.¹ Our primary contribution is thus to show for the first time how and why the Industrial Revolution took place in a country with increasingly scarce wood and abundant coal, namely Britain, without in some sense “stacking the deck” in coal’s favor.

Our second contribution is the discovery of a possible development path we call pre-industrial stagnation, where growth is slow and final output becomes ever more concentrated in the wood-using sector. This can happen if the elasticity of substitution between wood-intensive goods and coal-intensive goods is high enough and if initial conditions are sufficiently in wood’s favor.

We use an expanding machine varieties (horizontal innovation) approach to modeling directed technical change, which is appropriate since new types of machines and industrial processes using coal were characteristic of the Industrial Revolution. Though the organization of R&D was undoubtedly different in this period than today, entrepreneurs did carry out deliberate R&D and experiments to develop new products and techniques (Allen 2009, 241) and, Allen argues, responded strongly to price incentives in their choice of innovations to pursue. Our model has two intermediate goods sectors—the “Malthus” and “Solow” sectors—that use wood and coal inputs, respectively, and sectoral goods are then combined into final output via a constant elasticity of

Shanker, Gregor Singer, Olli Tahvonen, Mike Toman, and other participants at seminars at the London School of Economics, the Australian National University, the University of Wyoming, the annual conferences of the Australian Agricultural and Resource Economics Society, AERE, and EAERE and the World Congress of Environmental and Resource Economics. We thank the Australian Research Council for support under Discovery grant DP120101088: “Energy Transitions: Past, Present and Future,” and the National Natural Science Foundation of China for support under grant number 71703096.

1. We do not attempt to model the period after 1900, as from then on additional energy transitions to oil, electricity, etc. come into play.

substitution (CES) production function. Each sector also uses labor- and sector-specific machine inputs. Unlike almost all previous research discussed below, we do not assume that productivity is inherently higher or faster growing in the Solow than in the Malthus sector. Instead, we assume that wood is supplied perfectly inelastically (i.e., with constant quantity), while coal is supplied perfectly elastically (i.e., at a constant price).² In the next section, we show that these key asymmetric assumptions about energy supply are broadly consistent with the available historical data.

As in other directed technical change models, when the elasticity of substitution between the two sectoral goods is greater than unity, relative innovation activity is positively related to the relative abundance of the two sector-specific factors. Thus, an increase in the scarcity of wood relative to coal increases the level of innovation in the coal-using Solow sector relative to that in the wood-using Malthus sector. Kander and Stern (2014) show that the elasticity of substitution between biomass and fossil fuel energy was greater than unity in Sweden in the late nineteenth and early twentieth centuries, and we assume that this was the case for wood-intensive and coal-intensive goods in Britain.³

In our model, growth is forever unbalanced, and there is an underlying tendency toward the relative growth of the coal-using sector, whereas if the quantities of both coal and wood were fixed there would be a balanced growth path (BGP). But when coal is expandable, long-run behavior of the model depends on the degree of substitutability between the two goods. We will show analytically that if wood is initially sufficiently abundant relative to coal, and substitutability between goods produced using wood or coal is high enough, then the incentives for innovation increasingly favor wood-directed technical change, resulting in technical change being wood-biased forever (preindustrial stagnation). For “medium” substitutability, all development paths undergo an industrial revolution—defined as coal-intensive goods output becoming ever more dominant. These paths eventually result in modern economic growth, where coal’s price relative to wood is no longer falling but rising while its relative use is also rising, thus exhibiting the upward-sloping relative demand curve that Acemoglu (2007) calls strong (relative) bias. For “high” substitutability, either an industrial revolution (and modern economic growth) or preindustrial stagnation can happen, depending on the economy’s starting point. Preindustrial stagnation generally also exhibits strong relative bias, where wood’s relative price rises together with its relative use.

Using an empirically calibrated simulation of our model, we then show how growing population could have forced up the price of wood, resulting in the shift of innovative

2. Unlike Hanlon’s (2015) study of the effect of the American civil war on British innovation, supply conditions do not change over time in our model; rather the elasticities of supply of the factor inputs are different.

3. When commenting on relevant literature, we use the various terms that each researcher uses for different energy resources such as fossil fuels, renewable energy, biomass, etc., some of which also differ substantively from the “coal” and “wood” categories used in our model.

activity to the coal-using sector. But one of our counterfactual simulations shows how a combination of abundant initial wood and low population growth would lead to wood-dominant, preindustrial stagnation. So necessity is the mother of invention in our model.

Of course, our model abstracts from other issues such as Allen's (2009) argument that expensive labor was the reason why coal-directed innovation was profitable in Britain long before it was elsewhere. We also implicitly assume that the British institutional environment was appropriate for accelerating growth to occur, for example by having the well-developed patenting system which Madsen et al. (2010) find to be econometrically significant.⁴ Furthermore, we do not make a distinction between the usefulness of different inventions. Authors such as Mokyr (2009a) and Allen (2009) view macro-inventions like the steam engine or coke smelting as having a significant role in the Industrial Revolution (Crafts 2010). However, it can take more than a century of small improvements ("micro-inventions") for technical efficiency to improve enough for a macro-invention to have a significant macroeconomic impact (Clark and Jacks 2007; Allen 2009). Modeling technological change as deterministic and incremental, as we do here, rather than stochastic and sometimes revolutionary, therefore arguably misses no vital feature of the Industrial Revolution. Finally, we abstract from other properties of coal relative to wood such as higher energy density per cubic meter or per hectare of the land used for energy production.

Previous research relevant to our model falls into three areas. First are unified growth models, which explain the take-off from Malthusian stagnation (where any technical progress results in population rather than income growth) but do not model fossil fuels explicitly. Hansen and Prescott (2002) have two sectors, with a fixed land input in the agricultural, "Malthus" sector, no natural resource input to the industrial, "Solow" sector, semi-endogenous population growth, and exogenous technical progress that is assumed a priori to be much faster in the Solow than in the Malthus sector. As a result, the economy transitions from the Malthus to the Solow sector. Other papers in this vein include O'Rourke et al. (2013), who introduce directed technical change in a unified growth model, but with sectors distinguished by high or low labor skills rather than by use of land; and Kögel and Prskawetz (2001), Voigtländer and Voth (2006), and Strulik and Weisdorf (2008), who make assumptions about differences in productivity growth, the capital externality, or the elasticity of consumer demand for the output from agricultural and manufacturing sectors.⁵

The second area of relevant literature comprises papers that do model the effect of fossil fuels on long-run growth (Tahvonen and Salo 2001; Fröling 2011; Eren and

4. But see Mokyr (2009b) on the limitations of the patent system.

5. Lewis (1954) was, of course, the first to develop a two-sector model of the transformation of a preindustrial economy. He assumed an infinitely elastic supply of labor in the traditional, land-based sector, and that capital was only used in the modern sector. But these assumptions about economies in the first stages of industrialization are not necessarily accurate (Gollin 2014).

Garcia-Macia 2013; Gars and Olovsson 2019). These researchers all assume, like Hansen and Prescott, that fossil-fuel-directed technical change is more rapid or less costly to undertake than renewables-directed technical change. Gars and Olovsson (2019) is particularly relevant, as they allow for an equilibrium with persistent slow growth and biofuel-directed technical change. Their model has a CES function with an elasticity of substitution greater than one combining intermediate energy goods made from fossil or biofuel energy sources and sector-specific machines. This energy aggregate then combines with another intermediate good, made from labor and machines in a CES function with elasticity of substitution of less than one to produce the final output. There is Schumpeterian endogenous technical change, no population growth, and the fossil fuel stock grows exogenously. Crucially, they also assume that it is more costly to innovate in the biofuel sector than in the fossil fuel sector. For a single-country model, if fossil fuels are available, then the economy uses only fossil fuels and only develops the fossil fuel technology. Otherwise, it grows much more slowly by developing the biofuel technology. When there are multiple heterogeneous countries and international trade, if developed countries are much more advanced, they bid up the fossil fuel price and make fossil-fuel-using innovation unprofitable in developing economies. Developed countries then follow the fossil fuel growth path, and developing countries follow the biofuel-only path. By contrast, in our model, both fuels are always used, and we allow for an endogenous transition from an economy with predominantly biofuel-directed technical change to one with fossil-fuel-directed change and more rapid economic growth.

The third area of relevant literature is empirical work on the historical role of coal in the Industrial Revolution. Clark and Jacks (2007, 68) argue that an industrial revolution could still have happened in a coal-less Britain with only “modest costs to the productivity growth of the economy” because the value of coal was only a modest share of British GDP, and they argue that Britain’s energy supply could have been greatly expanded, albeit at about twice the cost of coal, by importing wood from the Baltic. Madsen et al. (2010) find that coal production in British coal mines had no econometrically significant effect on per capita output. But both Clark and Jacks (2007) and Madsen et al. (2010) do not allow for the dynamic effects of resource scarcity on the rate of innovation. Finally, Kander and Stern (2014) econometrically estimate a model of the transition from biomass energy (mainly wood) to fossil fuel (mainly coal) in Sweden, which shows the importance of this transition in economic growth there. However, they assume exogenous factor-directed technical change.

1. STYLIZED FACTS

Figure 1a shows the evolution of GDP per capita over 20-year periods from 1560 to 1900.⁶ Up to 1660, GDP per capita was flat or declining, after which it grew at an

6. For 1870 to 1900 we use the Composite GDP (E) measure of real GDP at 2006 prices from Hills et al. (2010). From 1540 to 1870 we used the growth rates from Broadberry et al.’s (2015) estimate of GDP for Great Britain in constant prices of 1700 to project real GDP back to 1540.

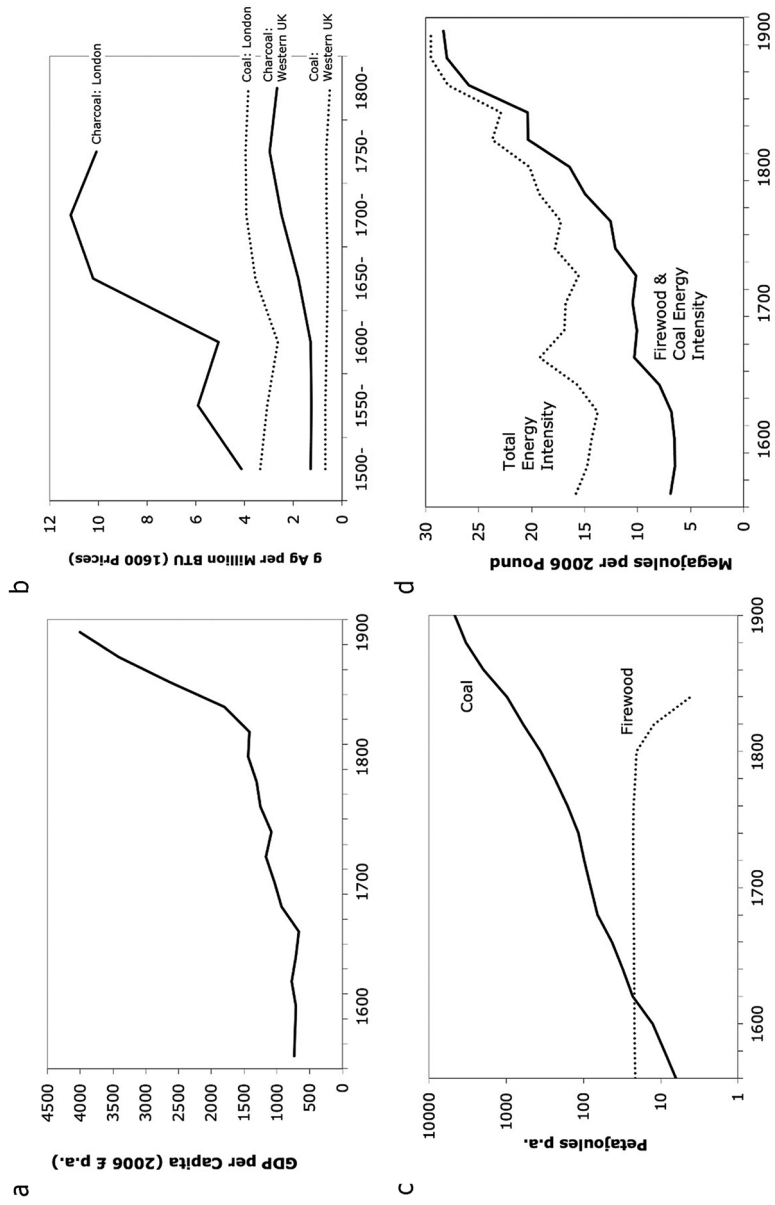


Figure 1. Historical data. *Source:* a, Broadberry et al. (2015). b, Allen (2009, table 4.3); units are grams of silver per million BTU at constant prices of 1550. c, Warde (2007). d, Authors' calculations from data in Warde (2007), Hills et al. (2010), and Broadberry et al. (2015).

accelerating rate, though the growth rate was quite erratic and in the second half of the nineteenth century ranged from 0.8% to 1.9% per annum, which is low by twentieth- or twenty-first-century standards.⁷

Figure 1*b* shows the real prices of coal and charcoal in London and Western Britain (Allen 2009). The price of charcoal rose steeply from the beginning of the seventeenth century to the late eighteenth century after which it appears to level off and fall (Fouquet 2011). The price of coal though is relatively stable over time in both regions. Clark and Jacks (2007) explain that throughout this period innovation in coal extraction overcame the effects of depletion, resulting in the long-run supply of coal being highly elastic.⁸ Figure 1*c* shows the energy content of firewood (including charcoal) and coal consumed in England and Wales (Warde 2007, appendix). Firewood provided about 80% of total fuel in 1560, declining to about 25% by 1700 and to zero by 1850. The absolute quantity of firewood used was fairly constant from about 1560 until 1800. Though timber was increasingly imported to Britain, especially in the nineteenth century (Iriarte-Goñi and Ayuda 2012), there does not seem to have been significant international trade in firewood (Thomas 1986; Warde 2007). Coal use increased 700-fold over the period. Though the quantity of firewood used eventually fell to zero during the nineteenth century, for simplicity our model will assume that wood use for energy (including charcoal) was constant throughout.

Energy intensity in Britain increased till the end of the nineteenth century, after which it declined (Kander et al. 2014). From 1720 to 1900 it roughly doubled, but prior to the mid-eighteenth century it was fairly constant (fig. 1*d*). However, figure 1*d* also shows that if one includes only coal and wood in the energy aggregate—thus eliminating more than half of the energy used in 1560—then intensity also rose since the early seventeenth century and quadrupled by 1900.

Gentvilaite et al. (2015) calculate that Britain's energy cost share declined from around 25% of total costs in 1800 to around 15% in 1900. We do not have sufficient data to estimate the cost share before 1800 consistent with Gentvilaite et al.'s (2015) data. However, using "back of the envelope" calculations based on the data shown above, it seems that the cost of energy relative to the GDP may have been constant or rising till the late seventeenth century before beginning its decline. Therefore, it is reasonably consistent with history that our model will, for the sake of simplicity, assume a constant energy cost share.

7. Though an acceleration of the rate of economic growth was a defining feature of the Industrial Revolution, the time path of income (per capita) over the last millennium is still deeply disputed among economic historians (Fouquet and Broadberry 2015). For example, Clark (2010) estimates English income to have changed very little between preindustrial times and 1800, but Broadberry et al. (2015) estimate that income nearly tripled between 1270 and 1800.

8. We model neither depletion nor innovation in the coal mining sector itself, focusing on innovation in the downstream industrial sectors.

2. THE MODEL

We assume that there are two energy sources—wood and coal—which can both be augmented by technological change. Goods produced using either of the two energy sources are good substitutes for each other. In common with Acemoglu (2002), technical change is modeled as an expansion of machine varieties, but as in Acemoglu et al. (2012), in addition to intermediate machines and labor, natural resources contribute to production. While only one sector has a resource input in Acemoglu et al. (2012), in our model each sector has a resource input—wood or coal. Following our discussion of figure 1b and 1c, we assume that the wood quantity and coal price are exogenously fixed. The consumers supply a unit of wood whose supply is regulated by institutions, and coal is available at a fixed marginal cost in terms of final output. Therefore, we do not consider the nonrenewable nature of coal—or the renewable nature of wood—explicitly, and neither do we model innovation in extraction. Except for some analytic results and in the constant population scenario in section 5, we assume that population, and hence the labor force, grows exogenously up to a finite limit, so that the available wood quantity per worker falls. As in Acemoglu et al. (2012), we use discrete time and assume that a patent for any variety of machine only lasts one period, here 20 years.⁹ We assume that at the beginning of each period, patents for all existing machine varieties are reissued at random, meaning that all varieties (new and existing) are produced by monopolistic firms, which maximize only current period profits.¹⁰ The 20-year period also is a convenient time step for the assumption that all machines depreciate fully within one period. As a result, the consumer plays no active role in our model: profit maximization ensures that consumption is maximized and there is no intertemporal investment decision, which greatly simplifies the model. For our dynamic analytical results and all our

9. If innovators are granted perpetual patents, then they need to consider the net present value of the stream of future profits when deciding how much to invest in innovation activities. As explained by Acemoglu (2002), this decision is then complicated because not only might the interest rate vary over time off a balanced growth path—and in our model a balanced growth path is highly unlikely due to the fixed wood supply—but also the relative prices of the two goods will change over time. This would lead to a complicated dynamic programming problem, which is why Acemoglu (2002) focuses on deviations from balanced growth. Twenty years is the current length of a UK patent. The 1624 Statute of Monopolies set a 14-year period (Khan and Sokoloff 2004).

10. This is similar to the assumption in Acemoglu et al. (2012). We could instead assume that when the patent expires each machine variety is produced competitively in all following periods, so that its price equals marginal cost, and newly developed machine varieties will, therefore, be priced higher than older varieties and used in smaller amounts (see Gancia and Zilibotti [2005] and app. B9 of Acemoglu et al. [2012] for similar models). This is what is seen in the real world, where new technologies are expensive and sold in smaller quantities but later become commodified. However, this assumption complicates our analytical model without changing our qualitative results or adding any useful insights.

numerical simulations, we also assume that the asymptotic rate of growth will be constant, to match the Kaldor stylized facts (Kaldor 1963).

2.1. Production, Prices, and the Allocation of Labor and Energy

Final output, Y , is produced competitively from two intermediate goods, Y_M and Y_S , where the subscripts M and S refer to Malthus and Solow, via a constant elasticity of substitution production function:

$$Y_t = \left[\gamma Y_{M,t}^{\frac{\sigma-1}{\sigma}} + (1-\gamma) Y_{S,t}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad (1)$$

where $\sigma > 1$ is the elasticity of substitution, $0 < \gamma < 1$ is the distribution parameter, and t indicates the (discrete) time period. Various critical values of σ will be shown later to be important in our analysis. The two goods are produced competitively using the following Cobb-Douglas technologies:

$$Y_{M,t} = \frac{1}{\beta} \left(\int_0^{N_{M,t}} x_{M,t}(j)^\beta dj \right) \bar{E}_M^\alpha L_{M,t}^{1-\alpha-\beta}, \quad (2)$$

$$Y_{S,t} = \frac{1}{\beta} \left(\int_0^{N_{S,t}} x_{S,t}(j)^\beta dj \right) E_{S,t}^\alpha L_{S,t}^{1-\alpha-\beta}, \quad (3)$$

where $0 < \alpha, \beta, \alpha + \beta < 1$.¹¹ The Malthus sector uses the fixed wood supply, \bar{E}_M , and a range $N_{M,t}$ of varieties of wood-using machines as inputs, with each variety j used in amount $x_{M,t}(j)$. Asymmetrically, the Solow sector uses an indefinitely expandable coal supply, $E_{S,t}$, and a range $N_{S,t}$ of varieties of coal-using machines as inputs, with each variety used in amount $x_{S,t}(j)$. The initial ranges of machine varieties that can be used with wood and coal, respectively $N_{M,0} > 0$ and $N_{S,0} > 0$, are given as parameters. The numbers of types of machines are also referred to as knowledge in the following.

The labor used in each sector is given by $L_{M,t}$ and $L_{S,t}$. Crucially, labor is mobile between sectors, but the sum, L_t , is assumed to be exogenous and equal to the level of population, which grows toward a finite limit L_∞ :

11. For analytical tractability, given our model's historically realistic asymmetry between wood and coal supply conditions, we use this Cobb-Douglas form, thus departing from another realistic assumption, namely, that the elasticity of substitution between energy and machines is less than 1, as used in previous research (Stern and Kander 2012; Kander and Stern 2014). Lemoine (2017) shows how low substitutability between energy and machines avoids an important source of sectoral lock-in that may prevent long-run energy transitions; so, generalizing (2)–(3) to the non-Cobb-Douglas case remains a worthwhile topic for further research. However, numerical simulations show that an elasticity less than 1 gives results not much different from those in this paper. Note that for most purposes our β corresponds to $1 - \beta$ in Acemoglu (2002).

$$L_{M,t} + L_{S,t} = L_t; \lim_{t \rightarrow \infty} L_t \equiv L_{\infty}. \quad (4)$$

In our baseline simulation in section 5, population L_t is closely matched to British history.

We use final output, Y , as the numeraire, normalizing its price to 1. The prices of the two goods inputs are thus related as follows:

$$\gamma^\sigma p_{M,t}^{1-\sigma} + (1-\gamma)^\sigma p_{S,t}^{1-\sigma} = 1, \quad (5)$$

and so, defining $p_t \equiv p_{M,t}/p_{S,t}$ and $\Gamma \equiv \gamma/(1-\gamma)$, (see app. 1; apps. 1–11 are available online):

$$p_{M,t} = (1-\gamma)^{\frac{\sigma}{\sigma-1}} (p_t^{\sigma-1} + \Gamma^\sigma)^{\frac{1}{\sigma-1}} \quad \text{and} \quad p_{S,t} = (1-\gamma)^{\frac{\sigma}{\sigma-1}} (1 + \Gamma^\sigma p_t^{1-\sigma})^{\frac{1}{\sigma-1}}. \quad (6), (7)$$

The goods price ratio, p_t , is given in competitive equilibrium by:

$$p_t = \left(\frac{Y_{M,t}}{Y_{S,t}} \right)^{-\frac{1}{\sigma}} = \Gamma y_t^{-\frac{1}{\sigma}} \quad \text{or} \quad y_t \equiv \frac{Y_{M,t}}{Y_{S,t}} = \Gamma^\sigma p_t^{-\sigma}, \quad (8)$$

which we use later to replace p_t by the sectoral output ratio, y_t , or vice versa. The marginal value products and hence prices of wood and coal are respectively given by:

$$e_{M,t} = p_{M,t} \frac{\partial Y_{M,t}}{\partial \bar{E}_M} = \alpha p_{M,t} \frac{Y_{M,t}}{\bar{E}_M}, \quad (9)$$

$$\bar{e}_S = p_{S,t} \frac{\partial Y_{S,t}}{\partial \bar{E}_{S,t}} = \alpha p_{S,t} \frac{Y_{S,t}}{\bar{E}_{S,t}}, \quad (10)$$

where the coal price, \bar{e}_S , is assumed to be constant, as noted above. We define the energy price ratio, e_t , as the wood price relative to the coal price.

Our model thus adds two features to the model in Acemoglu (2002):

- Energy inputs and prices are asymmetric (\bar{E}_M vs. $E_{S,t}$, $e_{M,t}$ vs. \bar{e}_S). As shown in equation (22) below, optimal coal use depends on the number of Solow varieties, $N_{S,t}$, but wood is fixed in quantity. This means that an increase in N_S has a greater effect on output in that sector than the same increase in N_M has on output in the Malthus sector. With fixed quantities of both coal and wood, we would asymptotically have a BGP and the coal price would rise to help steer the economy back to the BGP. This asymmetry explains why our model has no BGP, and why our results are mostly more complex than those in models with BGPs. The expandability of fixed-price coal means that a transition from wood-based production to coal-based production is the dominant development path. However, we will see that, given a high enough elasticity of substitution, a highly wood-dependent path can

diverge to become ever more wood-dependent, rather than being steered back to a BGP by a falling relative coal price.

- A third production input, labor, L , mobile between sectors and growing exogenously over time that enhances the effect of energy asymmetry. The presence of $1 - \alpha - \beta > 0$ in many of our results reflects the effect of labor mobility, and if one sets $1 - \alpha - \beta = 0$, thus eliminating the labor inputs in (2)–(3), several of our model's early results (those not involving innovation) revert to the corresponding results in Acemoglu (2002). Labor mobility between sectors is easier, the higher is the elasticity of substitution, σ .

Equations (9), (10), and (8) then give this expression for e_t , which along with (8) and (13) will be helpful later:

$$e_t \equiv \frac{e_{M,t}}{\bar{e}_S} = \frac{p_t y_t}{E_t} = \frac{\Gamma y_t^{\frac{\sigma-1}{\sigma}}}{E_t} \text{ where } E_t \equiv \frac{\bar{E}_M}{E_{S,t}}. \quad (11)$$

Finally, labor mobility between sectors results in a common wage rate w_t , equal to the marginal value product of labor in each sector:

$$w_t = p_{M,t}(1 - \alpha - \beta) \frac{Y_{M,t}}{L_{M,t}} = p_{S,t}(1 - \alpha - \beta) \frac{Y_{S,t}}{L_{S,t}}, \quad (12)$$

which, also using (8) and (11), gives several alternative expressions for the relative cost ratios:

$$l_t \equiv \frac{L_{M,t}}{L_{S,t}} = p_t y_t = \Gamma^\sigma p_t^{-(\sigma-1)} = \Gamma y_t^{\frac{\sigma-1}{\sigma}} = e_t E_t, \quad (13)$$

and thus from (4) the labor used in the Malthus and Solow sectors (see app. 1):

$$L_{M,t}(p_t) = \frac{L_t \Gamma^\sigma p_t^{1-\sigma}}{1 + \Gamma^\sigma p_t^{1-\sigma}} \quad \text{and} \quad L_{S,t}(p_t) = \frac{L_t}{1 + \Gamma^\sigma p_t^{1-\sigma}}. \quad (14), (15)$$

2.2. Market for Machines and Incentives for Innovation

Given the above, the first-order conditions for profit maximization by competitive manufacturers of each intermediate good, Y_i , $i = M, S$, imply that the amount of each variety of machine that they demand is:

$$x_{i,t}(j) = \left(\frac{p_{i,t} E_{i,t}^\alpha L_{i,t}^{1-\alpha-\beta}}{\chi_{i,t}(j)} \right)^{\frac{1}{1-\beta}}, \quad (16)$$

where $\chi_{i,t}(j)$ is the price of a machine of type j . Following Acemoglu (2002), we set the marginal cost of manufacturing a machine at a common constant, ψ . Given our assumption

that all machines are produced under a single-period patent, each machine variety is supplied by a monopolist that maximizes profit, which, using (16) to substitute for $\chi_{i,t}(j)$, is given for variety j by:

$$\pi_{i,t}(j) = (\chi_{i,t}(j) - \psi)x_{i,t}(j) = p_{i,t}E_{i,t}^\alpha L_{i,t}^{1-\alpha-\beta} x_{i,t}^\beta(j) - \psi x_{i,t}(j). \quad (17)$$

Maximizing profit then yields a (privately) optimal machine price of $\chi_{i,t}^*(j) = \psi/\beta$, and we set marginal cost $\psi = \beta$ so that $\chi_{i,t}^*(j) = 1$. Then, from (16), the optimal amount of each machine variety sold by each monopolist is given by:

$$x_{i,t}^*(j) = \left(p_{i,t}E_{i,t}^\alpha L_{i,t}^{1-\alpha-\beta}\right)^{\frac{1}{1-\beta}}, \quad (18)$$

which is independent of j . Substituting this for $x_{i,t}(j)$ in the production functions (2) and (3) gives these sectoral outputs (see app. 1 for the functions $p_{M,t}(p_t)$, $L_{M,t}(p_t)$, etc.):

$$Y_{M,t}(p_t, N_{M,t}) = \frac{N_{M,t}}{\beta} p_{M,t}^{\frac{\beta}{1-\beta}}(p_t) \bar{E}_M^{\frac{\alpha}{1-\beta}} L_{M,t}^{\frac{1-\alpha-\beta}{1-\beta}}(p_t), \quad (19)$$

$$Y_{S,t}(p_t, N_{S,t}) = \frac{N_{S,t}}{\beta} p_{S,t}^{\frac{\beta}{1-\beta}}(p_t) \bar{E}_{S,t}^{\frac{\alpha}{1-\beta}} L_{S,t}^{\frac{1-\alpha-\beta}{1-\beta}}(p_t), \quad (20)$$

from which we derive two equations useful for later results.¹² One is found by dividing (19) by (20), substituting $p_t = \Gamma y_t^{-(1/\sigma)}$ (8) and $l_t = \Gamma y_t^{(\sigma-1)/\sigma}$ (13), and rearranging terms to get the relative goods ratio:

$$y_t(N_t, E_t) = \left(\Gamma^{1-\alpha} E_t^\alpha N_t^{1-\beta}\right)^{\frac{\sigma}{\theta}}, \quad (21)$$

where $N_t \equiv N_{M,t}/N_{S,t}$ is the ratio of machine varieties, which we also call the knowledge ratio, and $\theta \equiv 1 + \alpha(\sigma - 1)$ is the elasticity of substitution between the factors with labor and machines adjusting optimally. If the two energy sources were fixed in quantity, then the evolution of y_t would depend (positively and monotonically) on N_t alone.¹³ The expandability of coal in our model means that this is not the case here. The other equation is for optimal coal use, $E_{S,t}(p_t, N_{S,t})$, in terms of the goods price ratio and the number of coal-using varieties, which we find by substituting (20) into the first-order condition for coal use, (10), for \bar{e}_S and rearranging to get:

$$E_{S,t}(p_t, N_{S,t}) = \left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S}\right)^{\frac{1-\beta}{1-\alpha-\beta}} p_{S,t}^{\frac{1}{1-\alpha-\beta}}(p_t) L_{S,t}(p_t), \quad (22)$$

which when inserted into (20) and rearranged gives Solow-sector output:

12. Equations (19) and (20) are analogous to eq. (15) in Acemoglu (2002).

13. From (11) the same will be true for the relative energy price ratio, e_t .

$$Y_{S,t}(p_t, N_{S,t}) = \left(\frac{N_{S,t}}{\beta} \right)^{\frac{1-\beta}{1-\alpha-\beta}} p_{S,t}^{\frac{\alpha+\beta}{1-\alpha-\beta}}(p_t) \left(\frac{\alpha}{\bar{e}_S} \right)^{\frac{\alpha}{1-\alpha-\beta}} L_{S,t}(p_t). \quad (23)$$

Comparing (23) to (19), we see that, as noted above, a given proportional increase in knowledge in both sectors increases output in the S sector more than in the M sector, the knowledge elasticity of the former being $(1-\beta)/(1-\alpha-\beta)$, which is larger than the unitary knowledge elasticity in the M sector. The reason is that any knowledge change in the S sector is matched by an increase in more energy inputs (see [22]), while in the M sector energy input necessarily stays constant.

From (17) and (18), profit per new variety, and the relative profitability of innovating in the two sectors (both also independent of j), are, therefore:

$$\pi_{i,t}(j) = [\chi_i^*(j) - \psi] x_{i,t}^*(j) = (1-\beta) \left(p_{i,t} E_{i,t}^{\alpha} l_t^{1-\alpha-\beta} \right)^{\frac{1}{1-\beta}}, \text{ and} \quad (24)$$

$$\pi_t \equiv \frac{\pi_{M,t}(j)}{\pi_{S,t}(j)} = \left(p_t E_t^{\alpha} l_t^{1-\alpha-\beta} \right)^{\frac{1}{1-\beta}}. \quad (25)$$

Equation (25) shows that the relative profitability of innovation is determined by a (goods) price effect, $p_t^{1/(1-\beta)}$, a (input) market size effect, $E_t^{\alpha/(1-\beta)}$, familiar from Acemoglu (2002), and also a labor mobility effect, $l_t^{(1-\alpha-\beta)/(1-\beta)}$.¹⁴ There is a profit incentive to innovate in the sector with the dearer good, but also in the sector that uses more inputs. As is familiar from Acemoglu (2002), the net effect for $\sigma > 1$ is to favor innovation in the sector whose relative price is falling. As (13) shows that $l_t = p_t y_t = \Gamma^{\sigma} p_t^{-(\sigma-1)}$, labor flows to the sector whose revenue share is rising and (given $\sigma > 1$) whose relative price is falling. This effect further accentuates the incentives in favor of the sector whose relative price is falling. Comparing relative profitability (25) with relative output from the ratio of (19) and (20), we see that $\pi_t N_t = p_t y_t$, which with the equilibrium relationship between the price and output ratios, (8), gives:

$$\pi_t = p_t y_t N_t^{-1} = \Gamma y_t^{\frac{\sigma-1}{\sigma}} N_t^{-1}. \quad (26)$$

The relative incentives for innovation are thus increasing in the ratio of the revenue shares of the intermediate goods, Y_M and Y_S , in final production and decreasing in the ratio of knowledge. So, there are diminishing returns to developing new varieties as the stock of varieties in a sector increases, but an incentive to innovate in the sector with growing revenue.¹⁵ If the energy quantities were fixed, then the relative goods ratio

14. If our model had no mobile labor so that $1-\alpha-\beta = 0$, (25) would revert to Acemoglu's (2002) (17.1).

15. This is related to Hart's (2013) finding that relative R&D investment rates depend on the relative factor shares. We focus on this approach to describing the innovation incentives, as market size and price effects become ambiguous with the addition of the mobile labor input to the standard directed technical change model.

equation, (21), shows that the relative incentives would depend on the knowledge ratio, N_p , alone.

2.3. Technology Innovation Process and Asymptotic Growth

We use a variation on Acemoglu's (2002) lab equipment model, where new machine varieties generated in sector i and period t , $\Delta N_{i,t} \equiv N_{i,t} - N_{i,t-1}$ (with Δ similarly defined for all other time-dependent variables), are a function of R&D expenditure in each sector, $R_{i,t} > 0$:

$$\Delta N_{i,t} = \eta_i R_{i,t}^\nu; \eta_M = \eta_S = \eta > 0, 0 < \nu < 1; \text{ hence } \Delta N_{i,t} > 0. \quad (27)$$

We assume diminishing returns in knowledge production in each sector with respect to research expenditure as more innovating firms enter the sector and spend on R&D.¹⁶ We also assume, as a key model feature mentioned earlier, that there is no difference in the productivity η with which a given amount of research expenditure can produce new knowledge of either type.¹⁷ As will be clear from (33) below, there is a scale effect of population size on the rate of innovation. However, none of our key results depend on population growth. We rearrange (27) to give the total cost of producing new varieties in each sector in a given period:

$$R_{i,t} = \left(\frac{1}{\eta} \Delta N_{i,t} \right)^{\frac{1}{\nu}}. \quad (28)$$

Assuming free entry, the profit from the last variety, $\pi_{i,t}(j)$ from (24), will equal the marginal cost of innovating a new variety in a sector in a given period, $\partial R_{i,t} / \partial (\Delta N_{i,t})$, calculated from (28):¹⁸

$$\pi_{i,t}(j) = \left(\frac{1}{\eta} \right)^{\frac{1}{\nu}} \frac{1}{\nu} (\Delta N_{i,t})^{\frac{1-\nu}{\nu}}. \quad (29)$$

We define the relative knowledge growth rate or direction of technical change, n_t , as:

16. With the time subscripts given—R&D is funded from current production—diminishing returns are needed to obtain an equilibrium. If there were constant returns to spending on R&D (and there are constant returns to knowledge in the production functions) then infinite R&D is optimal and fundable from the resulting production. We assume that new varieties generated are independent of the stock of knowledge; this is the simplest assumption compatible with a constant long-run growth rate in the Solow sector, a rate we derive at the end of sec. 4.

17. We note in sec. 5 that allowing $\eta_M \neq \eta_S$ did little to improve our simulations' goodness of fit.

18. Because of diminishing returns, this is an equality, so there will always be innovation in both sectors as long as both sectoral goods are produced.

$$n_t \equiv \frac{\frac{\Delta N_{M,t}}{N_{M,t}}}{\frac{\Delta N_{S,t}}{N_{S,t}}} = 1 + \frac{\Delta \ln N_t}{\Delta \ln N_{S,t}}; \quad \Delta N_t \gtrless 0 \Leftrightarrow n_t \gtrless 1. \quad (30)$$

Combining the above definitions with $\pi_t = (\Delta N_{M,t}/\Delta N_{S,t})^{(1-\nu)/\nu}$ from (29) then shows that the direction of (proportional) technical change, n_t , is increasing in the relative profitability, π_t , of innovation and decreasing in the knowledge ratio, N_t :

$$n_t = N_t^{-1} \pi_t^{\frac{\nu}{1-\nu}}. \quad (31)$$

Using this and the right-hand side of the relative incentives for innovation equation, (26), we have:

$$n_t(y_t, N_t) = \Gamma^{\frac{\nu}{1-\nu}} y_t^{\left(\frac{\sigma-1}{\sigma}\right)\frac{\nu}{1-\nu}} N_t^{-\left(\frac{1}{1-\nu}\right)}, \quad (32)$$

which will be the key equation for our phase-plane analysis of unbalanced growth in section 4. The direction of technical change defined by n_t , therefore, depends on the same incentives as relative profitability. If $n_t > 1$ we will say that technical change is Malthus directed and if $n_t < 1$, it is Solow directed. Again, with fixed energy quantities, from the relative goods ratio equation, (21), the direction of technical change would depend only on the knowledge ratio, N_t . At $n_t = 1$ there would be a BGP where N , e , and y (and so p) were constant.

Rearranging (29), dividing both sides by $N_{i,t}$ and inserting (24) for $\pi_{i,t}$ gives:

$$\frac{\Delta N_{i,t}}{N_{i,t}} = N_{i,t}^{-1} \left(\eta^{\frac{1}{\nu}} \nu \right)^{\frac{\nu}{1-\nu}} \pi_{i,t}^{\frac{\nu}{1-\nu}} = N_{i,t}^{-1} \left(\eta^{\frac{1}{\nu}} \nu (1-\beta) \right)^{\frac{\nu}{1-\nu}} \left(p_{i,t} E_{i,t}^{\alpha} L_{i,t}^{1-\alpha-\beta} \right)^{\frac{1}{1-\beta} \left(\frac{\nu}{1-\nu} \right)} > 0, \quad (33)$$

showing the scale effect of population $L_{i,t}$ on innovation noted above. Next, substituting optimal coal use, (22), into the Solow-sector version of (33) gives (see app. 2):

$$\frac{\Delta N_{S,t}}{N_{S,t}} = \left(\eta^{\frac{1}{\nu}} \nu (1-\beta) \right)^{\frac{\nu}{1-\nu}} \left(\frac{\alpha}{\beta \bar{e}_S} \right)^{\frac{\alpha \nu}{(1-\nu)(1-\alpha-\beta)}} N_{S,t}^{\frac{\alpha}{1-\alpha-\beta} \left(\frac{\nu}{1-\nu} \right) - 1} p_{S,t}^{\frac{\nu}{(1-\nu)(1-\alpha-\beta)}} L_{S,t}^{\frac{\nu}{1-\nu}}. \quad (34)$$

As mentioned above, for our dynamic results in section 4 and our simulations in section 5 we impose the restriction that the economy's growth rate, and hence the Solow knowledge growth rate, $\Delta N_{S,t}/N_{S,t}$, is asymptotically constant. In an economy where there is an industrial revolution, production becomes ever more concentrated in the Solow, coal-using sector, so the output ratio, y_t , tends to zero. As a result, both $p_{S,t}$ and, given there is an upper bound on population, $L_{S,t}$ will tend to constants; so, a constant asymptotic growth rate requires:

Assumption 1:

$$\nu = \frac{1 - \alpha - \beta}{1 - \beta}, \quad (35)$$

so that the exponent of $N_{S,t}$ in (34), $\frac{\alpha}{1-\alpha-\beta}(\frac{\nu}{1-\nu})-1$, is zero.

Assumption 1 is not actually needed for most of the dynamic results in section 4, but to simplify our presentation, we use it for most of the proofs of those results.

2.4. Household

Each household supplies a unit of labor and \bar{E}_M/L_t units of wood inelastically. Population is set exogenously. Consumers' income consists of wages, wood rents, and profits from the sale of machines. For simplicity, we assume that producing wood does not require labor. Total consumption is $C_t = Y_t - X_t - \sum_i R_{i,t} - \bar{e}_S E_{S,t}$, where $X_t = \beta(\int_0^{N_{M,t}} x_{M,t}(j) dj + \int_0^{N_{S,t}} x_{S,t}(j) dj)$ is total expenditure on producing machines. As already noted, households are only passive consumers of final output, so we need not specify consumption any further than this.

2.5. Equilibrium

The model yields a system of three simultaneous equations for three unknowns in any period t : the sectoral goods price ratio, p_t , and the growth in numbers of Malthus-sector (wood-using) and Solow-sector (coal-using) machine varieties, $\Delta N_{M,t}$ and $\Delta N_{S,t}$. The first equation is the equilibrium between demand and supply for y given by (8) and the ratio of output in each sector, (19) and (23). The remaining two equations are given by (33) after substituting in the relevant functions that determine the rates of technical change as functions of the innovation incentives:

$$\Gamma^\sigma p_t^{-\sigma} = \frac{Y_{M,t}(p_t, N_{M,t})}{Y_{S,t}(p_t, N_{S,t})}, \quad (36)$$

$$\frac{\Delta N_{M,t}}{N_{M,t}} = N_{M,t}^{-1} \left(\eta^\nu \nu (1 - \beta) \right)^{\frac{\nu}{1-\nu}} \left(p_{M,t}(p_t) \bar{E}_M^\alpha L_{M,t}^{1-\alpha-\beta}(p_t) \right)^{\frac{\nu}{(1-\nu)(1-\beta)}}, \quad (37)$$

$$\frac{\Delta N_{S,t}}{N_{S,t}} = N_{S,t}^{-1} \left(\eta^\nu \nu (1 - \beta) \right)^{\frac{\nu}{1-\nu}} \left(p_{S,t}(p_t) E_{S,t}^\alpha(p_t, N_{S,t}) L_{S,t}^{1-\alpha-\beta}(p_t) \right)^{\frac{\nu}{(1-\nu)(1-\beta)}}. \quad (38)$$

Equations (6), (14), (7), and (15) give the explicit functional forms needed here for $p_{M,t}(p_t)$ and $L_{M,t}(p_t)$ (hence $Y_{M,t}(p_t, N_{M,t})$ via [19]), and for $p_{S,t}(p_t)$ and $L_{S,t}(p_t)$ (hence $Y_{S,t}(p_t, N_{S,t})$ via [20] and $E_{S,t}(p_t, N_{S,t})$ via [22]). Given all these functional forms and the model parameters at the start of period t , namely, \bar{E}_M , \bar{e}_S , $N_{M,t-1}$, $N_{S,t-1}$, α , β , γ , σ , η , and L_ν we establish:

Definition 1: An equilibrium is given by the sequences of wages (w_t), intermediate output prices ($p_{M,t}$, $p_{S,t}$), wood prices ($e_{M,t}$), coal demand ($E_{S,t}$), labor demands ($L_{M,t}$, $L_{S,t}$), machine demands ($x_{M,t}$, $x_{S,t}$), and expenditures on innovation ($R_{M,t}$, $R_{S,t}$) such that in each period t : p_t , $N_{M,t}$, and $N_{S,t}$ are simultaneously given by (36)–(38).

3. COMPARATIVE STATICS

If we fix the number of varieties $N_{M,t}$ and $N_{S,t}$ —that is, treat them as exogenous technology parameters—the equation system consists of just the supply-demand equality (36). This section analyzes this static equilibrium with exogenous technological changes, which will make it easier to understand the dynamic results in section 4. In addition to the effects of the number of varieties on the equilibrium, we also look at the effects of the price of coal, the quantity of wood, and population and at how the elasticity of substitution changes these effects.

Substituting (6), (7), (14), and (15) for prices $p_{M,t}(p_t)$ and $p_{S,t}(p_t)$ and labor inputs $L_{M,t}(p_t)$ and $L_{S,t}(p_t)$, into the ratio of output in each sector, (19) and (23), we obtain the relative supply of the two outputs on the right-hand side of (36), which for clarity we will label here as y_t^s :

$$y_t^s(\bar{N}_{M,t}, \bar{N}_{S,t}, p_t) = \phi \bar{N}_{M,t} \bar{N}_{S,t}^{-\left(\frac{1-\beta}{1-\alpha-\beta}\right)} \bar{E}_M^{\frac{\alpha}{1-\beta}} \left(\frac{\alpha}{\beta \bar{e}_S} \right)^{-\left(\frac{\alpha}{1-\alpha-\beta}\right)} p_t^{\frac{\beta-(\sigma-1)(1-\alpha-\beta)}{1-\beta}} \left(1 + \Gamma^\sigma p_t^{-(\sigma-1)} \right)^{\left(\frac{\alpha((\sigma-1)(1-\alpha-\beta)-1)}{(\sigma-1)(1-\beta)(1-\alpha-\beta)}\right)} L_t^{-\left(\frac{\alpha}{1-\beta}\right)}, \quad (39)$$

where¹⁹

$$\phi = (1 - \gamma)^{-\frac{\alpha\sigma}{(\sigma-1)(1-\beta)(1-\alpha-\beta)}} \Gamma^{\frac{\sigma(1-\alpha-\beta)}{1-\beta}}.$$

Figure 2 plots the relative supply (39) and relative demand, $y_t^d = \Gamma^\sigma p_t^{-\sigma}$ from (8), using our baseline parameters from section 5 and the values of other variables in 1560 in panel *b*, and with the same parameters but with $\sigma = 1.89$ instead of 4 in panel *a*. As in our simulation, figure 2 uses a normalized production function with the base period of

19. We use the term “relative supply curve” for simplicity even though (39) does not reflect supply conditions alone. This is because it uses the equilibrium labor formulae given by (14) and (15), which apply when the economy is in equilibrium so that $y_t = \Gamma^\sigma p_t^{-\sigma}$ as in (8). Instead, if we used the first-order conditions for labor use, (12), to substitute for $L_{S,t}(p_t)$ in Solow-sector output, (23), this would show that supply in the Solow sector is infinitely elastic in its own price, p_S , because there are constant returns to scale in this sector.

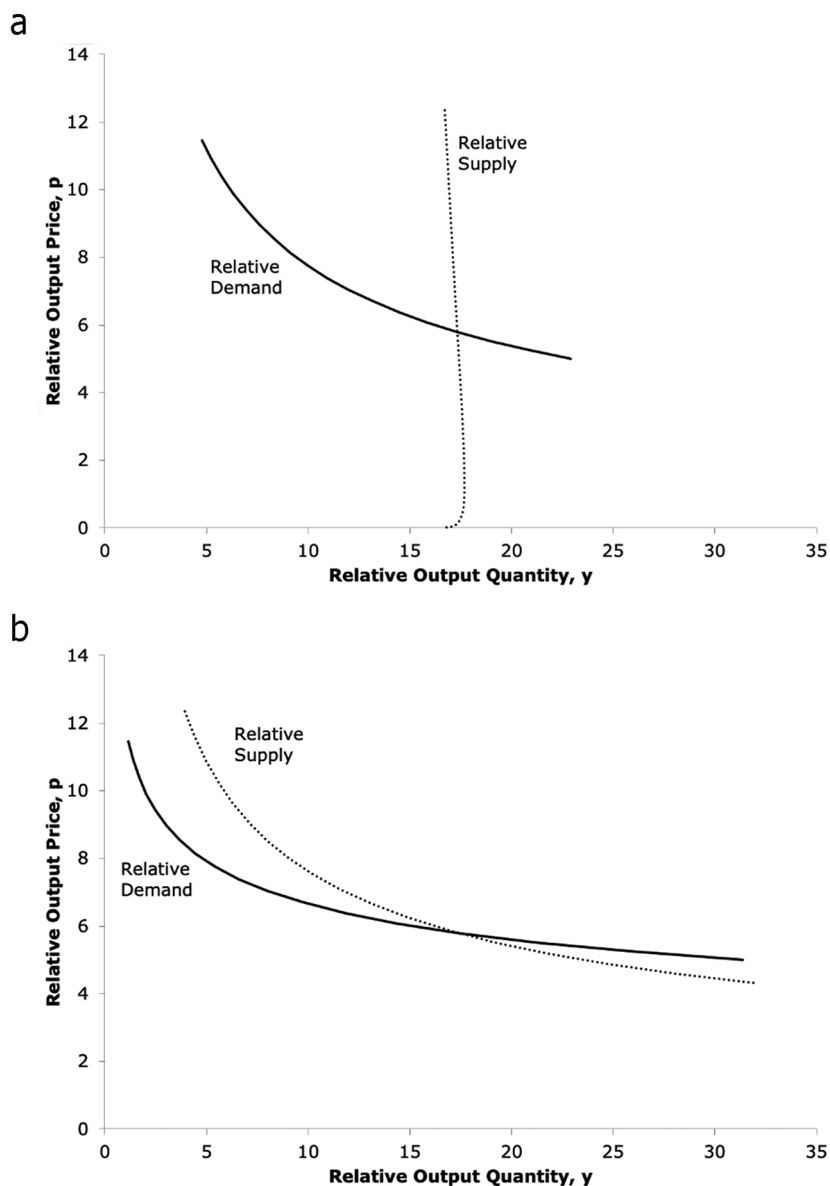


Figure 2. Relative output supply and demand. *a*, $\sigma = 1.89$. *b*, $\sigma = 4$

1560, so changes in σ do not change the equilibrium price and quantity but only the slopes of the supply and demand curves through the equilibrium point.

By computing the supply elasticity, $\rho(\sigma) = \partial \ln y_i^s / \partial \ln p_i$ from (39), which we do in appendix 3, we can show that the relative supply curve is upward sloping for

$\sigma < 1 + \beta/(1 - \alpha - \beta)$ (roughly 1.43 and less for our parameterization in sec. 5), downward sloping for $\sigma > 1/(1 - \alpha - \beta)$ (about 1.9 and greater here) of σ , and backward bending for an intermediate range of σ as in figure 2a. Intuitively, the downward slope of the relative supply curve for higher values of σ is due to the intersectoral labor mobility effect increasing with σ . As equations (19) and (23) show, outputs Y_M and Y_S are rising in their own prices as we would expect. However, sectoral labor allocation depends on the endogenous wage, and equations (14)–(15) show that for $\sigma > 1$, each sector's labor use declines with its own sector output price, and as its sector output price rises, the output price of the other sector falls, strengthening this effect. The greater σ is, the more this labor mobility effect acts against the positive own price effects in (19) and (23); hence, for higher values of σ , outputs are falling in their own prices, and rising p means falling y .

From (39), the parameters shift the relative supply curve as follows:

$$\begin{aligned}
 \frac{\partial \ln y_t^s}{\partial \ln \bar{N}_{M,t}} &= 1, \quad \frac{\partial \ln y_t^s}{\partial \ln \bar{E}_M} = \frac{\alpha}{1 - \beta}, \quad \frac{\partial \ln y_t^s}{\partial \ln \bar{e}_S} = \frac{\alpha}{1 - \alpha - \beta}, \\
 \frac{\partial \ln y_t^s}{\partial \ln \bar{N}_{S,t}} &= -\frac{1 - \beta}{1 - \alpha - \beta}, \quad \frac{\partial \ln y_t^s}{\partial \ln L_t} = -\frac{\alpha}{1 - \beta}.
 \end{aligned} \tag{40}$$

The signs of the first four derivatives are intuitive. For example, one would expect scarcer wood or cheaper coal to spur industrialization. Increased population shifts the relative supply to the left because it increases the optimal amount of coal but obviously does not affect the quantity of wood.

The equilibrium changes in relative output y also depend, of course, on both the demand elasticity, which from (8) is $-\sigma$, and the supply elasticity, $\rho(\sigma)$. Using a standard comparative static result, the equilibrium response of relative output to a shift in the relative supply function is given by

$$d \ln y = \left(\frac{\sigma}{\rho(\sigma) + \sigma} \right) d \ln y^s, \tag{41}$$

where $d \ln y^s$ is an exogenous shift (at constant p) in the relative output supply curve, caused, for example, by a change in N , the knowledge ratio. In appendix 4 we demonstrate the following proposition:

Proposition 1: Any of fewer wood-using varieties, $\bar{N}_{M,t}$, a lower wood quantity, \bar{E}_M , a lower coal price \bar{e}_S , more coal-using varieties, $\bar{N}_{S,t}$ or higher population, L_t , move the economy toward locally lower y_t (i.e., higher industrialization):

$$\partial y_t / \partial \bar{N}_{M,t}, \quad \partial y_t / \partial \bar{E}_M, \quad \partial y_t / \partial \bar{e}_S > 0; \quad \partial y_t / \partial \bar{N}_{S,t}, \quad \partial y_t / \partial L_t < 0. \tag{42}$$

Additionally, an equiproportional increase in $\bar{N}_{M,t}$ and $\bar{N}_{S,t}$, that is $\Delta \ln \bar{N}_{M,t} = \Delta \ln \bar{N}_{S,t} > 0$, hence $\Delta \ln \bar{N}_t = 0$, results in lower y_t , as $\bar{N}_{S,t}$ has a larger relative effect on supply in its sector than $\bar{N}_{M,t}$ does in its, due to the expandability of coal.

Proof. See appendix 4.²⁰

For our dynamic analysis in section 4, we would like to know how the elasticity of substitution, σ , affects these comparative statics. To provide some intuition, we look at how σ affects the price elasticity of the relative supply (39) and demand (8) functions and how, in turn, this affects the comparative statics. It turns out (see app. 5 for details) that for most parameter and price values, including all those relevant to the preindustrial stagnation behavior discussed in section 4.2, $\rho(\sigma)/\sigma$ decreases, hence $\sigma/[\rho(\sigma) + \sigma]$ increases, with increasing σ . This is illustrated by the move from figure 2a to 2b, where the supply elasticity $\rho(\sigma)$ goes from near zero to rather negative as σ increases. Hence from (41), when $\partial \ln y_t / \partial \ln N_t$ is positive, it increases with the elasticity of substitution, σ , a result we return to in section 4.2.

4. DYNAMIC RESULTS

4.1. Introduction

We now analyze the dynamics of the complete system (36)–(38), where growth in the numbers of machine varieties is determined by the incentives for innovation. We describe the paths that the economy can take using the labels industrial revolution, modern economic growth, and preindustrial stagnation, seen in figures 3a–4b (which we introduce shortly), and formally defined as:

Definition 2: A development path of the model undergoes an industrial revolution if $\Delta N_t < 0$ and $\Delta y_t < 0$ forever after some t on the path, so that Solow machine varieties and goods output are rising relative to Malthus varieties and output, with $N_t \rightarrow 0$ and $y_t \rightarrow 0$ as $t \rightarrow \infty$. Modern economic growth occurs on an industrial revolution development path if the energy price ratio, e_v , is falling along the path. Preindustrial stagnation occurs on a path if $\Delta N_t > 0$, $\Delta y_t > 0$ initially and forever, with $N_t \rightarrow \infty$ and $y_t \rightarrow \infty$ as $t \rightarrow \infty$, so it never undergoes an industrial revolution.

We establish three dynamic propositions, arranged in order of importance and in what is effectively the United Kingdom's historical sequence, and dependent on which of these three ranges of values that substitutability σ falls into:

20. Throughout our analysis, we approximate what are formally differences in discrete time (Δy_v , ΔN_v , etc.) as differentials in continuous time (dy , dN , etc). Many of our equals signs ($=$) should, therefore, strictly be replaced by approximately equals signs (\approx), but our many simulations (mostly not reported in sec. 5) have confirmed that the formally approximate analytic results thus found here hold true numerically. So, to avoid complexity that adds no insights, we have used only equals signs.

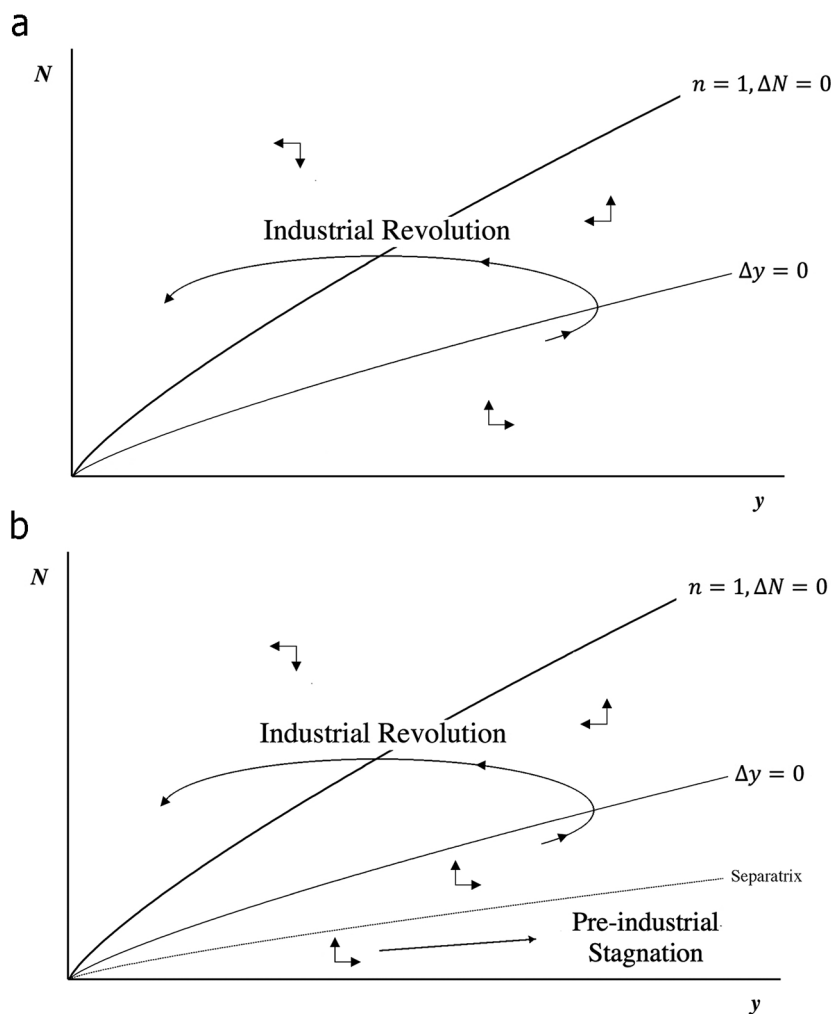


Figure 3. Phase diagrams in Malthus/Solow machine varieties ratio, N , and goods ratio, y , space. *a*, Medium elasticity of substitution ($\tilde{\sigma} < \sigma < \sigma^+$). *b*, High elasticity of substitution ($\sigma > \sigma^+$).

Definition 3: We refer to substitutability as being low if $1 < \sigma < \tilde{\sigma}$, medium if $\tilde{\sigma} < \sigma < \sigma^+$ as in figures 3*a* and 4*a*, and high if $\sigma > \sigma^+$ as in figures 3*b* and 4*b*, where:

$$\tilde{\sigma} \equiv 1 + \frac{1}{1 - \alpha - \beta} < \sigma^+ \equiv 1 + \frac{1 - \beta}{(1 - \alpha - \beta)^2}. \quad (43)$$

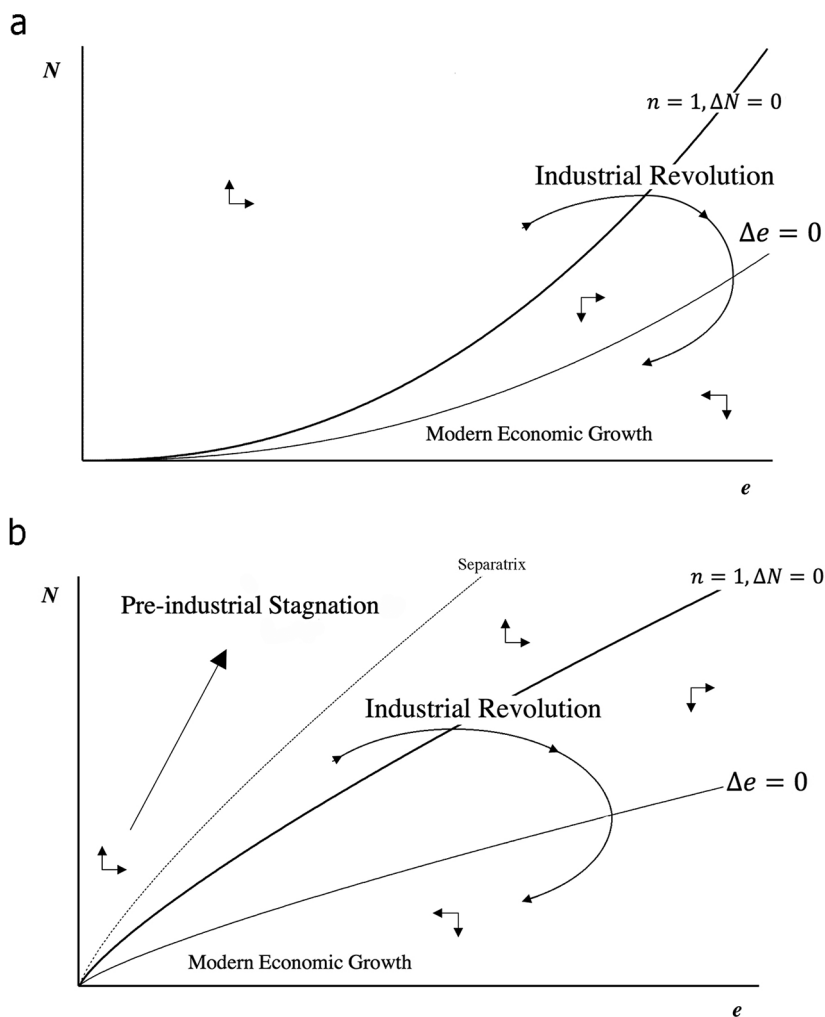


Figure 4. Phase diagrams in Malthus/Solow machine varieties ratio, N , and energy price ratio, e , space. *a*, Medium elasticity of substitution ($\tilde{\sigma} < \sigma < \sigma^\dagger$). *b*, High elasticity of substitution ($\sigma > \sigma^\dagger$).

Proposition 2 (sec. 4.2) will show how, given high substitutability, a zone of preindustrial stagnation will exist, in which the economy will remain if its initial conditions start it in that zone (fig. 3*b*). Proposition 3 (sec. 4.3) will show how, given medium or high substitutability, industrial revolution paths will end up with modern economic growth (fig. 4). Section 4.4 explains why high substitutability is needed for preindustrial stagnation but only medium substitutability is needed for modern economic

growth. Finally, proposition 4 (sec. 4.5) will show how assumption 1 means the economy's asymptotic growth rate (of total output per person) is a positive constant on an industrial revolution path but zero on a preindustrial stagnation path.

4.2. Industrial Revolution or Preindustrial Stagnation?

Figure 3—phase diagrams in (y, N) -space for medium and high substitutability—plots the dynamic evolution of the system as represented by (32), which we repeat here:

$$n_t(y_t, N_t) = \Gamma^{\frac{\sigma}{1-\sigma}} y_t^{\left(\frac{\sigma-1}{\sigma}\right)\frac{\sigma}{1-\sigma}} N_t^{-\left(\frac{1}{1-\sigma}\right)}.$$

Together, these figures illustrate most of our dynamic results and will be explained below.²¹ All development paths are unbalanced, asymptotically approaching either the origin (relative output and machine varieties, $y_t, N_t \rightarrow 0$), or (in fig. 3*b* only) infinity ($y_t, N_t \rightarrow \infty$), with no paths approaching fixed values of the ratios y and N characteristic of balanced growth. Figure 3*a* shows that for medium substitutability, all development paths converge to the origin, with output and machine varieties in the Solow sector growing relative to those in the Malthus sector. Initially, though, the economy follows a path where Malthus output and machine varieties grow faster than Solow ones. Figure 3*b* shows that when the elasticity of substitution is high, if the economy starts out in the preindustrial stagnation zone (whose existence will be demonstrated in proposition 2), where relatively high abundance of wood means that relative output, y , is high given the knowledge ratio, N , then the economy will never cross the separatrix into the industrial revolution zone.

Intuitively, the reason why all development paths in figure 3*a* and most paths in figure 3*b* eventually turn and undergo an industrial revolution (falling y_t) is because coal use rises along these development paths while wood use is fixed. On the arrowed path in figure 3*a*, initially knowledge growth is more rapid in the Malthus sector ($n_t > 1$), so y_t rises. The rise in coal use, however, boosts the profit incentive to research in the Solow sector (lowers π_t in [25]) by enough to lower n_t and hence slow the rise of y_t . In time, coal use rises (E_t falls) fast enough to turn around the direction of development (see [21]), despite the direction of technical change still being in the Malthus direction: the path crosses the $\Delta y_t = 0$ isocline even though N_t is still rising. By contrast, in the preindustrial stagnation zone in figure 3*b*, as we will show, coal use is falling, so development paths never turn around but instead diverge from industrial revolution paths.

21. Unlike phase diagrams for standard dynamic optimization models, where a control variable like consumption can jump in response to a shock, but a state variable like capital is predetermined by history, in our phase diagrams not only is the relative knowledge stock N predetermined, but also relative output y is in some sense predetermined by eq. (36).

From the (y, N) phase-plane equation, (32), the $\Delta N_t = 0$ isocline in figure 3 and the sign of ΔN_t above and below the isocline are given by:

$$\Delta N_t \geq 0 \Leftrightarrow N_t \leq \Gamma^\nu y_t^{\left(\frac{\sigma-1}{\sigma}\right)^\nu}. \quad (44)$$

The isocline is upward sloping because relative market size as measured by the revenue share ratio py ($= y^{(\sigma-1)/\sigma}$) has to be larger for an increased knowledge ratio N to offset the diminishing returns to knowledge. As stated in proposition 1, if knowledge in the Malthus and Solow sectors grows at the same proportional rate ($n_t = 1$) then relative output, y_t , is falling. So, the $\Delta y_t = 0$ isocline must be below the $\Delta N = 0$ isocline, giving rise to unbalanced growth.²² Formally, appendix 6 proves that:

$$\Delta y_t \geq 0 \Leftrightarrow n_t \geq \frac{1-\beta}{1-\alpha-\beta} > 1 \Leftrightarrow N_t \leq \left(\frac{1-\alpha-\beta}{1-\beta}\right)^{1-\nu} \Gamma^\nu y_t^{\left(\frac{\sigma-1}{\sigma}\right)^\nu}. \quad (45)$$

We now prove the key property of the high substitutability case: that the (y, N) phase-space is separated into a lower zone of preindustrial stagnation development paths and an upper zone of industrial revolution paths with a separatrix between them. An analytical proof exists only given the extra, counterfactual assumption of constant population, but we then discuss below the extension by continuity to the historical case of population growth.

Proposition 2: (a) Given a high elasticity of substitution ($\sigma > \sigma^+$) and constant population, there is a monotonic increasing separatrix in (y, N) -space lying strictly below the $\Delta y_t = 0$ isocline (45); all paths below this separatrix exhibit preindustrial stagnation, with eventually $\Delta n_t > 0$ forever, and all paths above it are industrial revolution paths, as in figure 3b. (b) Given a medium or low elasticity of substitution ($\sigma < \sigma^+$), all development paths undergo an industrial revolution (i.e., preindustrial stagnation is not possible).

Proof. See appendix 7.

With population growth ($\Delta \ln L_t > 0$), no analytic proof of part a is possible. But by continuity, proposition 2a holds for at least some small level of population growth,²³ and our numerical simulations found that it does indeed hold for historical British population growth and, moreover, for a wide range of variants on our baseline simulation.

22. Equation (21) shows that with fixed energy quantities, the $\Delta y_t = 0$ and $\Delta N_t = 0$ isoclines would collapse to a single common curve.

23. All key steps in proposition 2a's proof use inequalities (rather than equalities) that would remain true for some small level of population growth, since all our functions are continuously differentiable. For what growth level they would remain true, however, our theory cannot say.

The following explains intuitively why, given enough initial relative wood abundance, high σ gives rise to preindustrial stagnation. Taking logs and first differences of the (y, N) phase-plane equation, (32), and using assumption 1, we have:

$$\Delta \ln n_t = \frac{1 - \alpha - \beta}{\alpha} \left(\frac{\sigma - 1}{\sigma} \right) \Delta \ln y_t - \frac{1 - \beta}{\alpha} \Delta \ln N_t. \quad (46)$$

In order to get preindustrial stagnation, we need $\Delta \ln n_t > 0$ so that technical change is forever increasingly Malthus directed. The higher σ is, then, as from (13) in equilibrium $p_t y_t = \Gamma y_t^{(\sigma-1)/\sigma}$, the more the relative revenue share ratio, $p_t y_t$, increases for a given increase in y_t , increasing the incentive to innovate in Malthus technology rather than Solow technology. Both y and N are increasing in the relevant zone of the phase diagram below the $\Delta y_t = 0$ isocline (fig. 3b). As discussed in section 3, y is driven by changes in N so that here $\partial \ln y_t / \partial \ln N_t > 0$. When the indirect effect of the relative growth of Malthus knowledge, $[(1 - \alpha - \beta)/\alpha][(\sigma - 1)/\sigma](\partial \ln y_t / \partial \ln N_t) \Delta \ln N_t$ in (46), outweighs its direct effect, $-[(1 - \beta)/\alpha] \Delta \ln N_t$, the economy will remain in preindustrial stagnation. In this zone, growth in the ratio of the sectors' revenue shares will have a greater effect on incentivizing Malthus-sector innovation than the effect of diminishing returns to relative knowledge accumulation, creating a self-reinforcing spiral.

As discussed at the end of section 3, $\partial \ln y_t / \partial \ln N_t$ increases with σ for relevant values of the parameters and relative supply, y_t . Hence for values of σ above a minimum bound there will be a zone of preindustrial stagnation in part of the (y, N) -space below the $\Delta y_t = 0$ isocline where n_t is indeed rising. The end of appendix 7 shows how we can derive the bounding value of the elasticity of substitution, σ^\dagger , from further analysis of equation (46).

4.3. Modern Economic Growth and Energy Prices

In this section, we show how for medium and high substitutability the relative price of wood to coal eventually falls even though the relative use of coal is rising, in a phase we call modern economic growth. Using the relative cost ratios equation, (13), to change relative energy quantities, E_t , in the relative goods ratio equation, (21), to relative energy prices, e_t , and rearranging gives the conversion from the (y, N) -plane to the (e, N) -plane:

$$y_t = \Gamma^\sigma e_t^{-\alpha\sigma} N_t^{(1-\beta)\sigma}, \quad (47)$$

which we put into the (y, N) phase-plane equation, (32), and, using assumption 1, obtain the (e, N) phase-plane equation:

$$n_t(e_t, N_t) = \Gamma^{\frac{\sigma(1-\alpha-\beta)}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)} N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}}, \quad (48)$$

which is plotted in figure 4a and b for the medium ($\tilde{\sigma} < \sigma < \sigma^\dagger$) and high ($\sigma > \sigma^\dagger$) substitutability cases. In figure 4a, all development paths eventually converge to the origin as

in figure 3a. In figure 4b, most paths converge to the origin and modern economic growth, but the relative price of wood rises forever along preindustrial stagnation paths. Equation (48) shows that, for $\sigma > 1$, for a given knowledge ratio, N , higher relative wood scarcity discourages wood-directed innovation and encourages coal-directed innovation.

From (48), the isocline $\Delta N_t = 0$ in (e, N) -space and the signs of N_t above and below it are given by:

$$n_t \gtrless 1 \Leftrightarrow \Delta N_t \gtrless 0 \Leftrightarrow N_t \gtrless \Gamma^{\frac{-\sigma}{(1-\beta)(\sigma-\tilde{\sigma})}} e_t^{\frac{\alpha(\sigma-1)}{(1-\beta)(\sigma-\tilde{\sigma})}}. \quad (49)$$

Hence the $\Delta N_t = 0$ isocline is rising in (e, N) -space if $\sigma > \tilde{\sigma}$; and concave (convex) if $\alpha(\sigma - 1)/[(1 - \beta)(\sigma - \tilde{\sigma})] > (<)1$, which from definition 3 implies $\sigma > (<)\sigma^+$.

We now prove formally, and illustrate intuitively, the existence of the $\Delta e_t = 0$ isocline in figure 4, the phase diagrams in (e, N) -space, above which the relative price of wood to coal, e_t , rises, and below which it falls. The falling relative price of wood is the ultimate effect of self-reinforcing technical change. This is an example of strong (relative) biased technical change (Acemoglu 2002, 2007), where the relative price of coal increases alongside increasing coal use, which happens in all of the modern economic growth zone (definition 2) shown in figure 4. Formally, we can state:

Proposition 3: (a) Given medium or high substitutability ($\sigma > \tilde{\sigma}$) and constant population, an upward-sloping isocline $\Delta e_t = 0$ occurs below the $\Delta N_t = 0$ isocline in (e, N) -space, with $\Delta e_t > 0$ above the former isocline and $\Delta e_t < 0$ (modern economic growth) below it, as in figure 4. (b) There is strong relative bias throughout the modern economic growth zone, that is, relative coal use ($E_{S,t}/\bar{E}_M$) is rising even though its relative price ($\bar{e}_S/e_{M,t}$) is also rising.

Proof. See appendix 8.

Strong relative biased technical change occurs when the shift in relative supply of the inputs due to technical change is relatively small compared to the shift in their demand curve, so that the relative price of the inputs rises (falls) as their relative supply rises (falls). The relative inverse demand function for the two energy inputs is given by substituting the relative goods ratio equation, (21), into the energy price ratio, (11), and rearranging:

$$e_t(E_t, N_t) = \Gamma^{\sigma/\theta} E_t^{-1/\theta} N_t^{(1-\beta)(\sigma-1)/\theta}. \quad (50)$$

This demand curve is downward sloping; and given $\sigma > 1$, increases in N —relatively wood-directed technical change—shift the demand curve up and vice versa, so that wood-directed technical change is also wood biased. In other words, for constant E , the relative price of wood to coal, e , increases. To determine where on the relative demand

curve the equilibrium point is, we use the optimal coal quantity, (22), which plays the role of supply here.

In figure 4, below the $\Delta N = 0$ isocline, growth is coal biased and technical change is shifting the relative inverse energy demand function, (50), down over time. From (50), the higher σ is, the greater the exponent of N is,²⁴ hence the more that coal-biased technical change (falling N) shifts the relative energy demand curve down. Intuitively, the greater the flexibility of production, the more the relative marginal products of the factors move due to changes in the ratios of machine varieties and labor (with the quantity of the factors, E , held constant). However, E is falling because coal use is rising over time (according to [22]), which counteracts the effect of falling N .

If σ is great enough, then the shift down of the relative energy demand curve eventually becomes more rapid than the decline in E due to the expansion of coal use (wood use is fixed) so that there is a strong relative bias of technical change toward coal below the $\Delta e = 0$ isocline. This takes time to eventuate because initially as we cross the $\Delta N = 0$ isocline, technical change is still not so biased and does not shift the relative inverse energy demand function, (50), by much compared to the decline in E . So, at first e continues to rise and the $\Delta e = 0$ isocline is below the $\Delta N = 0$ isocline.

Let us take this process to the limit on the far left of figure 4a and b in the late stages of modern economic growth as $\epsilon \rightarrow 0$ and, therefore, $p \rightarrow \infty$. As we explained when introducing assumption 1, both the price of Solow-sector output, $p_{S,t}$ and the quantity of labor in the sector, $L_{S,t}$ will tend to constants, so from optimal coal use, (22), the growth rate of coal use, $\Delta \ln E_{S,t} \rightarrow [(1 - \beta)/(1 - \alpha - \beta)] \Delta \ln N_{S,t}$. Also, $\Delta N_{M,t} \rightarrow 0$, so the main driver of relative knowledge growth, $\Delta \ln N_t$ is $\Delta \ln N_{S,t}$.²⁵ The numerator of the exponent of N_t in the relative inverse energy demand function, (50), is $(1 - \beta)(\sigma - 1)$. If this is larger than $(1 - \beta)/(1 - \alpha - \beta)$, then “supply” of coal relative to wood will increase more slowly than the inverse relative demand curve shifts, resulting in strong relative bias to coal. This is the case for $\sigma > 1 + 1/(1 - \alpha - \beta) \equiv \tilde{\sigma}$, that is, medium substitutability, hence explaining why this is the threshold value needed for proposition 4 to hold.

4.4. Why Does Preindustrial Stagnation Require a Greater Elasticity of Substitution than Modern Economic Growth?

The (e, N) phase-plane equation, (48), readily shows that if $\sigma > \tilde{\sigma}$ it is possible for N , e , and n to all decline together; if so, there is then strong relative bias to coal and increasingly Solow-directed technical change. Proposition 3 showed that, indeed, strong relative bias to coal can occur when $\sigma > \tilde{\sigma}$. Furthermore, if we are not in preindustrial

24. Recall that $\theta \equiv 1 + \alpha(\sigma - 1)$, so $(d/d\sigma)[(\sigma - 1)/\theta] = 1/\theta^2 > 0$.

25. Substituting $p_{M,t}(p_t)$ (7) and $L_{M,t}(p_t)$ (14) into the system equation for the growth rate of Malthus varieties, (37), readily shows that $\Delta N_{M,t}$ declines toward zero as $y \rightarrow 0$, provided $\sigma > \tilde{\sigma}$.

stagnation, n must be falling. Equation (48) also suggests that the reverse—strong relative bias to wood and preindustrial stagnation or rising n —should be possible, but proposition 2 shows that only if $\sigma > \sigma^+$ can preindustrial stagnation actually occur. In appendix 9, we show that strong relative bias to wood where E and e are rising together, and thus coal use is declining, is possible for $\sigma > \tilde{\sigma}$ for some part of the phase plane above the $\Delta e = 0$ isocline in figure 4. However, it turns out that unless $\sigma > \sigma^+$, these paths will all eventually have rising rather than falling coal use and proceed to an industrial revolution. Technical change on these paths is initially Malthus directed but becomes less and less so. This is essentially due to the asymmetric energy supplies that do not allow for the expansion of wood use.

Next, we show that there must be strong relative bias for preindustrial stagnation to occur, using the following equation, derived in appendix 9 by combining the energy demand curve, (50), and the (e, N) phase-plane equation, (48):

$$n_t(e_t, E_t) = \Gamma \frac{\sigma}{\alpha(\sigma-1)} e_t^{\frac{(\sigma-1)(1-2\alpha-\beta)-1}{\alpha(\sigma-1)}} E_t^{\frac{(\sigma-1)(1-\alpha-\beta)-1}{\alpha(\sigma-1)}}. \quad (51)$$

This equation expresses the direction of technical change and, hence, also the incentives for innovation, in terms of an energy price effect, $e_t^{\frac{(\sigma-1)(1-2\alpha-\beta)-1}{\alpha(\sigma-1)}}$, and an energy market size effect, $E_t^{\frac{(\sigma-1)(1-\alpha-\beta)-1}{\alpha(\sigma-1)}}$. For $\sigma > \tilde{\sigma}$ this market size effect is positive. The price effect is negative—a more expensive energy resource disincentivizes innovation in a sector—unless $\sigma > 1 + 1/(1-2\alpha-\beta) > \sigma^+$.²⁶ As the energy price ratio, e , is increasing in the preindustrial stagnation zone, as long as $\sigma < 1 + 1/(1-2\alpha-\beta)$, we have to have strongly wood-biased technical change, where E and not just e is increasing, in order to get preindustrial stagnation. Therefore, the elasticity that allows for preindustrial stagnation must be larger than the one that allows for strongly biased technical change.

4.5. Asymptotic Growth Rates under High Substitutability

We can formally show that the economic growth rate is asymptotically zero under preindustrial stagnation, while, by assumption, it is asymptotically constant on an industrial revolution path. Given assumption 1, the Malthus-sector version of the knowledge growth rate equation, (33), becomes:

$$\frac{\Delta N_{M,t}}{N_{M,t}} = \lambda_M N_{M,t}^{-1} \left(p_{M,t} L_{M,t}^{1-\alpha-\beta} \right)^{\frac{1-\alpha-\beta}{\alpha(1-\beta)}} \text{ where } \lambda_M \equiv \left[\eta^{\frac{1-\beta}{1-\alpha-\beta}} (1-\alpha-\beta) \bar{E}_M^{\frac{\alpha}{1-\beta}} \right]^{\frac{1-\alpha-\beta}{\alpha}}, \quad (52)$$

26. To understand why the sign of the energy price effect flips as σ gets even greater, we see from (50) that constant E means that N must increase to increase e . But as we explained in sec. 4.2, it also increases y , and through the (y, N) phase-plane equation, (32), n . So, these other effects become more powerful than the disincentive of more expensive resources. A similar story explains why the market size effect in (51) is negative for lower values of σ .

while the Solow-sector version, (34), becomes:

$$\frac{\Delta N_{S,t}}{N_{S,t}} = \lambda_S p_{S,t}^{\frac{1}{\alpha}} L_{S,t}^{\frac{1-\alpha-\beta}{\alpha}} \text{ where } \lambda_S \equiv \eta^{\frac{1-\beta}{\alpha}} (1 - \alpha - \beta)^{\frac{1-\alpha-\beta}{\alpha}} \left(\frac{\alpha}{\beta e_S} \right), \quad (53)$$

which then allow us to show:

Proposition 4: Given the upper bound on population, L_∞ , economic growth (i.e., the growth rate of Y_t/L_t , final output per capita) is asymptotically zero under pre-industrial stagnation, when this path exists, and asymptotically constant under an industrial revolution:

$$\lim_{t \rightarrow \infty; PS} \frac{\Delta(Y_t/L_t)}{Y_t/L_t} = 0, \quad (54)$$

$$\lim_{t \rightarrow \infty; IR} \frac{\Delta(Y_t/L_t)}{Y_t/L_t} = \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \lambda_S (1 - \gamma)^{\frac{\alpha}{\sigma-1} \left(\frac{1}{\alpha} \right)} L_\infty^{\frac{1-\alpha-\beta}{\alpha}}. \quad (55)$$

Proof: See appendix 10.

This asymmetry between sectors again stems ultimately from the expandable coal supply, $E_{S,t}$, in the Solow-sector production function, (20), compared to the non-expandable wood supply, \bar{E}_M , in the Malthus-sector version (19) and from our assumption in (4) of an asymptotic limit to population.²⁷ The growth rate in (55) increases with the elasticity of substitution, as predicted by de la Grandville (1989).

5. SIMULATIONS

In this section, we use a Matlab program to find numerical solutions for the dynamic equilibrium defined by definition 1. The simulation uses a normalized production function and exogenous population that approximates historical growth. The parameters are either drawn from the literature or chosen by calibration of the simulation to a set of stylized facts discussed in section 1. Full details are in appendix 11. Critically, in the baseline simulation, we set $\sigma = 4$, which is above $\sigma^+ (= 3.81)$ given our choices of α and β), that is, high.

We first show how this baseline simulation illustrates proposition 3. We then present counterfactual simulations that illustrate the comparative statics presented in section 3, that the Industrial Revolution would have been delayed by any of more abundant wood, a higher coal price, less initial Solow knowledge, or less population growth.

27. It can be shown that with asymptotically growing population, and also a small enough energy share ($\alpha < (1 - \beta)/2$), positive growth would remain possible under preindustrial stagnation. We thank a referee for this point.

We also find that the Industrial Revolution would have been delayed by a higher elasticity of substitution between sectoral goods, consistent with proposition 2a that pre-industrial stagnation is possible only for high enough substitutability.

5.1. Baseline Simulation

Figure 5a graphs our baseline simulation results over time for the relative price of wood and coal, e_t , the log of coal use, $\ln E_{S,t}$, and output per capita, Y_t/L_t . Coal use is converted to logarithms because its overall growth is so great. Also shown are the corresponding data for these three variables based on those presented in figure 1. Simulated results are broadly comparable to the historical data and illustrate both parts of proposition 3, extended to a growing population. However, both the peak in the simulated wood/coal price and the acceleration in economic growth come somewhat later than they do historically. This could be because of factors outside our simple model, such as the growth of trade and the reform of agriculture, which might have contributed to the growth in output per capita in the seventeenth and eighteenth centuries estimated by Broadberry et al. (2015). Our simulated final output better fits Clark's (2010) data, which show only a 22% increase from 1560 to 1800. Though there is technological change, the exogenous growth of population is sufficient to reduce output per capita in this period in our simulation. The Malthus sector's share of labor (not shown in fig. 5) falls from 85% in 1560 to 4% in 1900, which seems reasonable. Coal use increases 250-fold by 1900 in our simulation, which is less than in reality.

5.2. Counterfactual Simulations

In figures 5b–h we simulate the following seven counterfactual scenarios, to highlight the potential effects on economic growth of changing energy resource abundance and scarcity, and other key parameters. We name and define them as follows:

- b. Abundant wood: Wood quantity is 10 times higher than in the baseline scenario, so that $\bar{E}_M = 10$ instead of 1 in the baseline scenario.
- c. Expensive coal: The coal price is doubled, so that $\bar{e}_S = 0.23$ instead of 0.115 in the baseline scenario.
- d. More substitutability: The elasticity of intersectoral substitution σ is increased to 4.2 instead of 4.
- e. Less substitutability: The elasticity of intersectoral substitution σ is reduced to 3.5 instead of 4.
- f. Low Solow knowledge: The initial stock of Solow sector varieties, $N_{S,0}$, is halved to 0.5 instead of 1.
- g. Constant population: $L_t = 1$ always.
- h. Preindustrial stagnation: A combination of the abundant wood, more substitutability, and constant population scenarios, one of many variants that results in a preindustrial stagnation path.

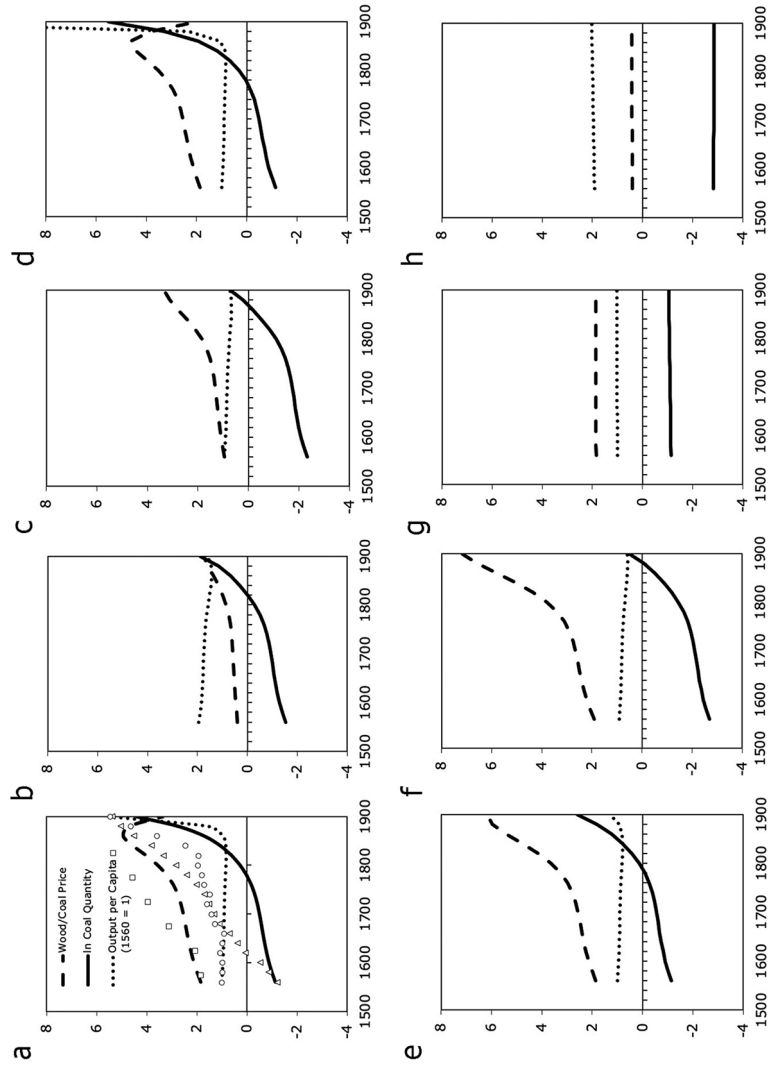


Figure 5. Baseline and counterfactual simulations. *a*, Baseline. *b*, Abundant wood. *c*, Expensive coal. *d*, More substitutability. *e*, Less substitutability. *f*, Low Solow knowledge. *g*, Constant population. *h*, Preindustrial stagnation. In panel *a*, squares are observed e , triangles observed e , and circles observed Y/L .

Except in scenarios *g* and *h*, we assume that population followed its historical path. In all cases apart from *d*, an industrial revolution is delayed, consistent with proposition 1: GDP per capita grows more slowly or declines, less labor shifts to the Solow sector, and coal use and growth are lower.

In the abundant wood scenario (fig. 5*b*) output per capita is nearly twice the baseline level in 1560, and the wood price and (absolute) coal use are both much lower. Output per capita declines slowly as exogenous population grows, and while the price of wood rises steadily, by a factor of four from 1560 to 1900, at the end it does not reach even its initial baseline level. Energy intensity (not shown) declines strongly, and the share of labor in the Malthus sector (also not shown) starts higher (95%) and falls much less, being still 73% in 1900. This scenario clearly illustrates the paradox where an abundance of wood stalls development despite much higher initial output per capita.

The expensive coal scenario (fig. 5*c*) is similar in some ways to the abundant wood scenario, but the initial income level is a little below the baseline scenario and coal use is even lower. The price of wood is about the same as coal initially, so that here both fuels are relatively expensive, whereas in the abundant wood scenario, wood is much cheaper relative to output than in the baseline scenario.

The more substitutability scenario (fig. 5*d*) is an accelerated version of the baseline scenario. Output per capita increases 16-fold, and coal use increases 750-fold overall. The former is far more than occurred in reality, and the latter exactly matches the real-world increase. The wood price falls more steeply after its peak than in the baseline. This might seem counterintuitive, as a high elasticity of substitution can give rise to preindustrial stagnation in our model, but it also increases the asymptotic rate of growth on an industrial revolution path, as predicted by de la Grandville (1989) and in our proposition 4.

By contrast, in the less substitutability scenario (fig. 5*e*), growth is delayed, and the wood price rises more than in the baseline. Assuming the economy had low Solow-sector knowledge in 1560 produces similar results (fig. 5*f*) but with much lower coal use.

The constant population and preindustrial stagnation scenarios (fig. 5*g*, *h*) are very different from the other scenarios. Here there is extremely slow growth in GDP per capita and very little else changes. Constant population starts from the same point as the baseline scenario, while preindustrial stagnation starts from a different point than any other scenario, consistent with proposition 2a. Coal use is very slowly increasing under constant population, and very slowly declining under preindustrial stagnation. We also ran the preindustrial stagnation simulation with a low rate of population growth. Declining coal use is compatible with some amount of population growth. As the results look very similar to figure 5*h*, we have not included them here.

6. DISCUSSION AND CONCLUSIONS

We have shown the potential importance of the differential abundance of renewable and fossil energy resources—wood and coal—in driving the historic transition to modern

economic growth, using a model that both yields analytical insights and reproduces key empirical features of the British Industrial Revolution. We extended and calibrated an increasing machine varieties, directed technical change model, which, unlike previous related research, does not assume productivity or its growth to be inherently higher in the modern, industrial, coal-using, Solow sector than in the traditional, wood-using Malthus sector. Rather, we assume that resource supply conditions differ inherently, so that wood is inelastically and coal elastically supplied, which is a stylized representation of the British historical record.

Comparative static analysis of our model showed the effect of key parameters on the economy's state of development: notably, any of a lower coal price, lower wood quantity, or higher population will further industrialize the economy. Our model's dynamic analytical results show that growth is forever unbalanced, and there is an underlying tendency toward coal-biased growth, stemming from our asymmetric assumptions about the supply of wood and coal. If the elasticity of substitution between wood-intensive and coal-intensive goods is relatively low, then an industrial revolution, where production increasingly uses coal rather than wood, is inevitable. But if the elasticity is high enough and wood is initially sufficiently abundant relative to coal, then thanks to intersectoral labor mobility, the market size effect is strong enough so that the incentives for innovation more and more favor wood-directed innovation, despite diminishing direct innovation returns to accumulated knowledge, resulting in wood-biased technical change forever, which we call preindustrial stagnation. There is then a separatrix between development paths that start with sufficiently abundant wood and those that do not.

If the asymptotic rate of economic growth in an industrial revolution path is constant, then the asymptotic rate under preindustrial stagnation is zero. Preindustrial stagnation is also typically characterized by strong relative bias, defined as wood use relative to coal rising despite the relative price of wood rising. For medium and high levels of substitutability an industrial revolution eventually has strong relative bias, too, with relative coal use and price both rising.

Given some parameter values from the literature, fitting our model to some basic stylized historical facts results in a baseline simulation with sensible values for the free parameters, and a development path that reproduces the key features of the British Industrial Revolution. From the start, the growth rate of coal-using machine varieties exceeds that of wood-using varieties, though its absolute growth is less until 1820. The only exogenous driver in our model is the historical rate of population growth. This should be endogenized in future research, but leaving it exogenous here better highlights the role of natural resource scarcity in driving growth. The rate of economic growth accelerates partly because of the shift to the coal-using sector and partly because increased population increases the rate of innovation.

Compared to the previous literature (see Ashraf and Galor 2011), our model introduces a new reason for why an economy may either remain forever in preindustrial

stagnation or fail to make a timely industrial revolution, since either may be caused by abundant wood, high elasticities of substitution, and/or slow population growth. Our model's counterfactual simulations show that a much higher fixed quantity of wood input or fixed price of coal, and/or slower population growth would have greatly delayed growth of GDP per capita and the rate of innovation. In our model, it is the growing relative scarcity of wood caused by population growth that results in innovation to develop coal-using machines. Necessity is thus indeed the mother of invention: on its own, the unlimited supply of coal does not trigger a transition if wood is not relatively scarce.

Our model thus partly supports views by Allen (2009) and Wrigley (2010) that the Industrial Revolution first happened in Britain mainly because of its cheap, abundant coal. Counter to Clark and Jacks (2007) and Madsen et al. (2010), our model tells a plausible story of how coal could have played a central role in the Industrial Revolution.

However, we stress that our support is partial, because our model does not imply that cheap coal alone would have been sufficient for the Industrial Revolution to happen in Britain in the eighteenth century. Improved institutions or growth in human capital (Clark 2014), differences in demography (Voigtländer and Voth 2006), and even the location of monasteries (Andersen et al. 2017) have all been suggested as key factors that might explain why the Industrial Revolution happened when and where it did. Our results should not be seen as disagreeing with these views. Institutional factors are invisibly assumed in the mathematical structure of most economic growth models, including ours, so we implicitly treat them as also being necessary for growth. If economic analysis can take these factors as well as renewable energy scarcity and fossil fuel availability all into account, then the Industrial Revolution may not “remain one of history's mysteries” (Clark 2014, 260) for much longer.

REFERENCES

- Acemoglu, Daron. 2002. Directed technical change. *Review of Economic Studies* 69:781–810.
- . 2007. Equilibrium bias of technology. *Econometrica* 75 (5): 1371–409.
- Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. 2012. The environment and directed technical change. *American Economic Review* 102 (1): 131–66.
- Allen, Robert C. 2009. *The British Industrial Revolution in global perspective*. Cambridge: Cambridge University Press.
- Andersen, Thomas Barnebeck, Jeanet Bentzen, Carl-Johan Dalgaard, and Paul Sharp. 2017. Pre-reformation roots of the Protestant Ethic. *Economic Journal* 127 (604): 1756–93.
- Ashraf, Quamrul, and Oded Galor. 2011. Dynamics and stagnation in the Malthusian epoch. *American Economic Review* 101:2003–41.
- Broadberry, Stephen, Bruce M. S. Campbell, Alexander Klein, Mark Overton, and Bas van Leeuwen. 2015. *British economic growth, 1270–1870*. Cambridge University Press.
- Clark, Gregory. 2010. The macroeconomic aggregates for England, 1209–2008. *Research in Economic History* 27:51–140.

- . 2014. The Industrial Revolution. In *Handbook of economic growth*, vol. 2A, ed. Philippe Aghion and Steven Durlauf, 217–62. Amsterdam: North Holland.
- Clark, Gregory, and David Jacks. 2007. Coal and the Industrial Revolution, 1700–1869. *European Review of Economic History* 11:39–72.
- Crafts, Nicholas. 2010. Explaining the First Industrial Revolution: Two views. *European Review of Economic History* 15:153–68.
- de La Grandville, O. 1989. In quest of the Slutsky diamond. *American Economic Review* 79 (3): 468–81.
- Eren, Egemen, and Daniel Garcia-Macia. 2013. From wood to coal may well be from Malthus to Solow. Unpublished manuscript, Stanford University, Department of Economics.
- Fouquet, Roger. 2011. Divergences in long run trends in the prices of energy and energy services. *Review of Environmental Economics and Policy* 5 (2): 196–218.
- Fouquet, Roger, and Stephen Broadberry. 2015. Seven centuries of European economic growth and decline. *Journal of Economic Perspectives* 29 (4): 227–44.
- Fröling, Maria. 2011. Energy use, population and growth, 1800–1970. *Journal of Population Economics* 24 (3): 1133–63.
- Gancia, Gino, and Fabrizio Zilibotti. 2005. Horizontal innovation in the theory of growth and development. In *Handbook of economic growth*, vol. 1A, ed. Philippe Aghion and Steven Durlauf, 111–70. Amsterdam: North Holland.
- Gars, Johan, and Conny Olovsson. 2019. Fuel for economic growth? *Journal of Economic Theory* 184:104941.
- Gentvilaite, Ruta, Astrid Kander, and Paul Warde. 2015. The role of energy quality in shaping long-term energy intensity in Europe. *Energies* 8 (1): 133–53.
- Gollin, Douglas. 2014. The Lewis model: A 60-year retrospective. *Journal of Economic Perspectives* 28 (3): 71–88.
- Hanlon, W. Walker. 2015. Necessity is the mother of invention: Input supplies and directed technical change. *Econometrica* 83:67–100.
- Hansen, Gary D., and Edward C. Prescott. 2002. From Malthus to Solow. *American Economic Review* 92 (4): 1205–17.
- Hart, R. 2013. Directed technological change and factor shares. *Economics Letters* 119 (1): 77–80.
- Hills, Sally, Ryland Thomas, and Nicholas Dimsdale. 2010. The UK recession in context: What do three centuries of data tell us? *Bank of England Quarterly Bulletin* 2010, Q4: 277–88.
- Iriarte-Goñi, Iñaki, and María-Isabel Ayuda. 2012. Not only subterranean forests: Wood consumption and economic development in Britain (1850–1938). *Ecological Economics* 77:176–84.
- Kaldor, Nicholas. 1963. Capital accumulation and economic growth. In *The theory of capital: Proceedings of a conference*, ed. F. A. Lutz and D. C. Hague, 177–222. New York: Macmillan.
- Kander, Astrid, Paolo Malanima, and Paul Warde. 2014. *Power to the people: Energy and economic transformation of Europe over four centuries*. Princeton, NJ: Princeton University Press.
- Kander, Astrid, and David I. Stern. 2014. Economic growth and the transition from traditional to modern energy in Sweden. *Energy Economics* 46:56–65.
- Khan, B. Zorina, and Kenneth L. Sokoloff. 2004. Institutions and technological innovation during early economic growth: Evidence from the great inventors of the United States, 1790–1930. NBER Working paper 10966, National Bureau of Economic Research, Cambridge, MA.
- Kögel, Tomas, and Alexia Prskawetz. 2001. Agricultural productivity growth and escape from the Malthusian trap. *Journal of Economic Growth* 6:337–57.

- Kunnas, Jan, and Timo Myllyntaus. 2009. Postponed leap in carbon dioxide emissions: The impact of energy efficiency, fuel choices and industrial structure on the Finnish economy, 1800–2005. *Global Environment* 3:154–89.
- Lemoine, Derek. 2017. Innovation-led transitions in energy supply. NBER Working paper 23420, National Bureau of Economic Research, Cambridge, MA.
- Lewis, W. Arthur. 1954. Economic development with unlimited supplies of labor. *Manchester School* 22 (2): 139–91.
- Madsen, Jakob B., James B. Ang, and Rajabrata Banerjee. 2010. Four centuries of British economic growth: The roles of technology and population. *Journal of Economic Growth* 15 (4): 263–90.
- Mokyr, Joel. 2009a. *The enlightened economy: An economic history of Britain, 1700–1850*. New Haven, CT: Yale University Press.
- . 2009b. Intellectual property rights, the Industrial Revolution, and the beginnings of modern economic growth. *American Economic Review* 99 (2): 349–55.
- O'Rourke, Kevin Hjortshøj, Ahmed S. Rahman, and Alan M. Taylor. 2013. Luddites, the Industrial Revolution, and the demographic transition. *Journal of Economic Growth* 18:373–409.
- Pomeranz, Kenneth L. 2001. *The great divergence: China, Europe and the making of the modern world economy*. Princeton, NJ: Princeton University Press.
- Stern, David I., and Astrid Kander. 2012. The role of energy in the Industrial Revolution and modern economic growth. *Energy Journal* 33 (3): 125–52.
- Strulik, Holger, and Jacob Weisdorf. 2008. Population, food, and knowledge: A simple unified growth theory. *Journal of Economic Growth* 13:105–216.
- Tahvonen, Olli, and Seppo Salo. 2001. Economic growth and transitions between renewable and nonrenewable energy resources. *European Economic Review* 45:1379–98.
- Thomas, Brinley. 1986. Was there an energy crisis in Great Britain in the 17th century? *Explorations in Economic History* 23:124–52.
- Voigtländer, Nico, and Hans-Joachim Voth. 2006. Why England? Demographic factors, structural change and physical capital accumulation during the Industrial Revolution. *Journal of Economic Growth* 11 (4): 319–61.
- Warde, Paul. 2007. *Energy consumption in England and Wales, 1560–2000*. Naples: ISSM-CNR.
- Wilkinson, Richard G. 1973. *Poverty and progress: An ecological model of economic development*. London: Methuen.
- Wrigley, E. Anthony. 1988. *Continuity, chance, and change: The character of the Industrial Revolution in England*. Cambridge: Cambridge University Press.
- . 2010. *Energy and the English Industrial Revolution*. Cambridge: Cambridge University Press.

APPENDIX FOR ONLINE PUBLICATION IN *JAERE***(including Annex starting on p.16) for****Directed Technical Change and the British Industrial Revolution**

David I. Stern, John C. V. Pezzey, and Yingying Lu

Appendix 1: Derivation of Static Equilibrium Equations

Intermediate goods prices and the labor allocation are jointly determined economy-wide because of the labor adding-up condition (4) and the numeraire equation (5). Given goods prices and the labor allocation, all other quantities can then be determined for each sector. First, substitute $p_{M,t} = p_t p_{S,t}$ into the LHS of the numeraire equation (5):

$$\gamma^\sigma (p_t p_{S,t})^{1-\sigma} + (1-\gamma)^\sigma p_{S,t}^{1-\sigma} = 1$$

Dividing both sides by $p_{S,t}^{1-\sigma}$ and raising them to the power of $\frac{1}{\sigma-1}$ gives three forms of $p_{S,t}$ for use in (38) and elsewhere (the second and third using $\Gamma \equiv \frac{\gamma}{1-\gamma}$ and $\Gamma^\sigma p_t^{1-\sigma} = \Gamma y_t^{\frac{\sigma-1}{\sigma}}$ from (8)):

$$p_{S,t} = [\gamma^\sigma p_t^{1-\sigma} + (1-\gamma)^\sigma]^{\frac{1}{\sigma-1}} = (1-\gamma)^{\frac{\sigma}{\sigma-1}} (1 + \Gamma^\sigma p_t^{1-\sigma})^{\frac{1}{\sigma-1}} = (1-\gamma)^{\frac{\sigma}{\sigma-1}} (1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}})^{\frac{1}{\sigma-1}} \quad (7)$$

The price in the Malthus sector for use in (37) is then:

$$\begin{aligned} &= p_t p_{S,t} = (1-\gamma)^{\frac{\sigma}{\sigma-1}} p_t (1 + \Gamma^\sigma p_t^{1-\sigma})^{\frac{1}{\sigma-1}} = (1-\gamma)^{\frac{\sigma}{\sigma-1}} (p_t^{\sigma-1} + \Gamma^\sigma)^{\frac{1}{\sigma-1}} \\ &= (1-\gamma)^{\frac{\sigma}{\sigma-1}} \Gamma \left(y_t^{\frac{\sigma-1}{\sigma}} + \Gamma \right)^{\frac{1}{\sigma-1}} \end{aligned} \quad (6)$$

Next, we find the optimal levels of labor. Using (12) and (8):

$$\frac{Y_{M,t}}{Y_{S,t}} \equiv y_t = \Gamma^\sigma p_t^{-\sigma} = \frac{1}{p_t} \frac{L_{M,t}}{L_{S,t}} \Rightarrow l_t \equiv \frac{L_{M,t}}{L_{S,t}} = \Gamma^\sigma p_t^{1-\sigma} \quad (13)$$

Given (8) and (4) ($L_t = L_{M,t} + L_{S,t}$), $L_{S,t}$ and $L_{M,t}$, for use in (37), (38) and elsewhere, are given by:

$$L_t = (\Gamma^\sigma p_t^{1-\sigma} + 1) L_{S,t} \Rightarrow L_{S,t}(p_t) = \frac{L_t}{1 + \Gamma^\sigma p_t^{1-\sigma}} = L_{S,t}(y_t) = \frac{L_t}{1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}} \quad (15)$$

and

$$L_{M,t}(p_t) = L_t - L_{S,t}(p_t) = \frac{L_t \Gamma^\sigma p_t^{1-\sigma}}{1 + \Gamma^\sigma p_t^{1-\sigma}} = L_{M,t}(y_t) = \frac{L_t \Gamma}{\Gamma + y_t^{\frac{\sigma-1}{\sigma}}} \quad (14)$$

Then we substitute the optimal amount of machines sold, $x_{i,t}^*(j)$ from (18), into the goods production functions (2) and (3). Since $x_{i,t}^*(j)$ does not vary with j , this yields:

$$\begin{aligned} Y_{i,t} &= \frac{1}{\beta} \left(\int_0^{N_{i,t}} \left((p_{i,t} E_{i,t}^\alpha L_{i,t}^{1-\alpha-\beta})^{\frac{1}{1-\beta}} \right)^\beta dj \right) E_{i,t}^\alpha L_{i,t}^{1-\alpha-\beta} \\ &= \frac{1}{\beta} \left(N_{i,t} (p_{i,t} E_{i,t}^\alpha L_{i,t}^{1-\alpha-\beta})^{\frac{\beta}{1-\beta}} \right) E_{i,t}^\alpha L_{i,t}^{1-\alpha-\beta} \end{aligned}$$

hence

$$Y_{M,t}(p_t, N_{M,t}) = \frac{1}{\beta} N_{M,t} p_{M,t}^{\frac{\beta}{1-\beta}}(p_t) \bar{E}_M^{\frac{\alpha}{1-\beta}} L_{M,t}^{\frac{1-\alpha-\beta}{1-\beta}}(p_t) \quad (19)$$

and

$$Y_{S,t}(p_t, N_{S,t}) = \frac{1}{\beta} N_{S,t} p_{S,t}^{\frac{\beta}{1-\beta}}(p_t) E_{S,t}^{\frac{\alpha}{1-\beta}}(p_t, N_{S,t}) L_{S,t}^{\frac{1-\alpha-\beta}{1-\beta}}(p_t) \quad (20)$$

We also need to find $E_{S,t}(p_t, N_{S,t})$, the optimal amount of coal, in terms of p_t and $N_{S,t}$. Substituting (20) into (10) for \bar{e}_S and rearranging yields:

$$\bar{e}_S = \alpha p_{S,t} \frac{Y_{S,t}}{E_{S,t}} = \frac{\alpha}{\beta} p_{S,t}^{\frac{1}{1-\beta}} N_{S,t} E_{S,t}^{\frac{-(1-\alpha-\beta)}{1-\beta}} L_{S,t}^{\frac{1-\alpha-\beta}{1-\beta}}$$

Then solving this for the coal quantity we have:

$$E_{S,t}(p_t, N_{S,t}) = \left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S} \right)^{\frac{1-\beta}{1-\alpha-\beta}} p_{S,t}^{\frac{1}{1-\alpha-\beta}}(p_t) L_{S,t}(p_t) \quad (22)$$

Inserting (22) back into (20) gives:

$$\begin{aligned} Y_{S,t}(p_t, N_{S,t}) &= \frac{1}{\beta} N_{S,t} p_{S,t}^{\frac{\beta}{1-\beta}}(p_t) \left[\left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S} \right)^{\frac{1-\beta}{1-\alpha-\beta}} p_{S,t}^{\frac{1}{1-\alpha-\beta}}(p_t) L_{S,t}(p_t) \right]^{\frac{\alpha}{1-\beta}} L_{S,t}^{\frac{1-\alpha-\beta}{1-\beta}}(p_t) \\ &= \left(\frac{N_{S,t}}{\beta} \right)^{\frac{1-\beta}{1-\alpha-\beta}} \left(\frac{\alpha}{\bar{e}_S} \right)^{\frac{\alpha}{1-\alpha-\beta}} p_{S,t}^{\frac{\alpha+\beta}{1-\alpha-\beta}}(p_t) L_{S,t}(p_t) \end{aligned} \quad (23)$$

Inserting (19) and (23) into (8) then gives the equilibrium output price ratio in the form:

$$\Gamma^\sigma p_t^{-\sigma} = \frac{Y_{M,t}(p_t, N_{M,t})}{Y_{S,t}(p_t, N_{S,t})} \quad (36)$$

Lastly, (6), (7), (14), and (15) are the functional forms used in (37)-(38).

Appendix 2: Derivation of Equation (34) for $n_{S,t}$, the Growth Rate of Solow knowledge

The Solow-sector version of (33) is

$$n_{S,t} \equiv \frac{\Delta N_{S,t}}{N_{S,t}} = N_{S,t}^{-1} \left(\eta^{\frac{1}{\nu}} \nu (1-\beta) \right)^{\frac{\nu}{1-\nu}} (p_{S,t} E_{S,t}^\alpha L_{S,t}^{1-\alpha-\beta})^{\frac{1}{1-\beta} \left(\frac{\nu}{1-\nu} \right)} \quad (A1)$$

and

$$\begin{aligned} E_{S,t}(p_t, N_{S,t}) &= \left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S} \right)^{\frac{1-\beta}{1-\alpha-\beta}} L_{S,t}(p_t) p_{S,t}^{\frac{1}{1-\alpha-\beta}}(p_t) \\ \Rightarrow E_{S,t}^{\frac{\alpha}{1-\beta} \left(\frac{\nu}{1-\nu} \right)} &= \left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S} \right)^{\frac{\alpha}{1-\alpha-\beta} \left(\frac{\nu}{1-\nu} \right)} L_{S,t}^{\frac{\alpha}{1-\beta} \left(\frac{\nu}{1-\nu} \right)} p_{S,t}^{\frac{\alpha}{1-\beta} \left(\frac{\nu}{1-\nu} \right) \frac{1}{1-\alpha-\beta}} \end{aligned} \quad (22)$$

Inserting this into (A1):

$$\begin{aligned} &\Rightarrow n_{S,t} \\ &= N_{S,t}^{-1} \left(\eta^{\frac{1}{\nu}} \nu (1-\beta) \right)^{\frac{\nu}{1-\nu}} p_{S,t}^{\frac{1}{1-\beta} \left(\frac{\nu}{1-\nu} \right)} \left(\frac{\alpha}{\beta \bar{e}_S} \right)^{\frac{\alpha}{1-\alpha-\beta} \left(\frac{\nu}{1-\nu} \right)} N_{S,t}^{\frac{\alpha}{1-\alpha-\beta} \left(\frac{\nu}{1-\nu} \right)} L_{S,t}^{\frac{\alpha}{1-\beta} \left(\frac{\nu}{1-\nu} \right)} p_{S,t}^{\frac{\alpha}{1-\beta} \left(\frac{\nu}{1-\nu} \right) \frac{1}{1-\alpha-\beta}} L_{S,t}^{\frac{1-\alpha-\beta}{(1-\beta) \left(\frac{\nu}{1-\nu} \right)}} \end{aligned}$$

$$= \left(\eta^{\frac{1}{\nu}} \nu (1 - \beta) \right)^{\frac{\nu}{1-\nu}} \left(\frac{\alpha}{\beta \bar{e}_S} \right)^{\frac{\alpha \nu}{(1-\nu)(1-\alpha-\beta)}} N_{S,t}^{\frac{\alpha}{1-\alpha-\beta} \left(\frac{\nu}{1-\nu} \right) - 1} p_{S,t}^{\frac{\nu}{(1-\nu)(1-\alpha-\beta)}} L_{S,t}^{\frac{\nu}{1-\nu}} \quad (34)$$

Appendix 3: Signing the slope of the relative supply curve $y^s(N_M, N_S, p)$

To sign this slope, we take the log of $y^s(N_M, N_S, p)$ (39):

$$\ln y^s = \text{constant} + \frac{\beta - (\sigma - 1)(1 - \alpha - \beta)}{1 - \beta} \ln p + \frac{\alpha((\sigma - 1)(1 - \alpha - \beta) - 1)}{(\sigma - 1)(1 - \beta)(1 - \alpha - \beta)} \ln(1 + \Gamma^\sigma p^{-(\sigma-1)})$$

where *constant* is the terms that do not depend on p . We then take the derivative with respect to p , multiply the result by p , and simplify:

$$\rho(\sigma) = \frac{\partial \ln y^s}{\partial \ln p} = \frac{\beta - (\sigma - 1)(1 - \alpha - \beta)}{1 - \beta} - \frac{\alpha((\sigma - 1)(1 - \alpha - \beta) - 1)}{(1 - \beta)(1 - \alpha - \beta)} \frac{\Gamma^\sigma}{p^{\sigma-1} + \Gamma^\sigma}$$

We then rearrange the first term:

$$\rho(\sigma) = -\frac{\sigma - 1 - \frac{\beta}{1-\alpha-\beta}}{(1 - \beta)/(1 - \alpha - \beta)} - \frac{\alpha((\sigma - 1)(1 - \alpha - \beta) - 1)}{(1 - \beta)(1 - \alpha - \beta)} \frac{\Gamma^\sigma}{p^{\sigma-1} + \Gamma^\sigma} \quad (A2)$$

If $\sigma < 1 + \frac{\beta}{1-\alpha-\beta}$ ($< 1 + \frac{1}{1-\alpha-\beta}$) then the first and second terms are both positive and hence $\rho(\sigma) > 0$, so that the relative supply curve is upward sloping.

If $\sigma > 1 + \frac{\beta}{1-\alpha-\beta}$, as $p \rightarrow \infty$, $\rho(\sigma) \rightarrow -\frac{\sigma - 1 - \frac{\beta}{1-\alpha-\beta}}{(1 - \beta)/(1 - \alpha - \beta)} < 0$, and so the curve slopes down. As $p_t \rightarrow 0$,

$$\rho(\sigma) \rightarrow \frac{\beta - (\sigma - 1)(1 - \alpha - \beta)}{1 - \beta} - \frac{\alpha((\sigma - 1)(1 - \alpha - \beta) - 1)}{(1 - \beta)(1 - \alpha - \beta)} = \frac{1}{1 - \alpha - \beta} - \sigma \quad (A3)$$

So, if $\sigma > \frac{1}{1-\alpha-\beta} = 1.905$ for our baseline parameters, the curve slopes down for all p_t ; but if $1 + \frac{\beta}{1-\alpha-\beta} < \sigma < \frac{1}{1-\alpha-\beta}$, the curve is backward bending, i.e. sloped upward for low p but downward for high p .

Appendix 4: Proof of Proposition 1 on Comparative Statics

Combining Equations (40) and (41), the derivatives in Proposition 1 in elasticity form are:

$$\begin{aligned} \partial \ln y_t / \partial \ln \bar{N}_{M,t} &= \frac{1}{\rho(\sigma)/\sigma + 1} \\ \partial \ln y_t / \partial \ln \bar{E}_M &= \frac{\alpha}{1 - \beta} \left(\frac{1}{\rho(\sigma)/\sigma + 1} \right) \end{aligned}$$

$$\begin{aligned}\partial \ln y_t / \partial \ln \bar{e}_S &= \frac{\alpha}{1 - \alpha - \beta} \left(\frac{1}{\rho(\sigma)/\sigma + 1} \right) \\ \partial \ln y_t / \partial \ln \bar{N}_{S,t} &= -\frac{1 - \beta}{1 - \alpha - \beta} \left(\frac{1}{\rho(\sigma)/\sigma + 1} \right) \\ \partial \ln y_t / \partial \ln L_t &= -\frac{\alpha}{1 - \beta} \left(\frac{1}{\rho(\sigma)/\sigma + 1} \right)\end{aligned}$$

As we show in the following, $\frac{1}{\frac{\rho(\sigma)}{\sigma} + 1} > 0$, hence the signs of the derivatives in Proposition 1 are as shown there. Also, since $1 - \frac{1 - \beta}{1 - \alpha - \beta} < 0$, the above shows that an equiproportional increase in $\bar{N}_{M,t}$ and $\bar{N}_{S,t}$, i.e. $\Delta \ln(\bar{N}_{M,t}) = \Delta \ln(\bar{N}_{S,t}) > 0$, hence $\Delta \ln(\bar{N}_t) = 0$, results in lower y_t .

To show $\frac{1}{\frac{\rho(\sigma)}{\sigma} + 1} > 0$ we need to have $\frac{\rho(\sigma)}{\sigma} > -1$. Dividing $\rho(\sigma)$ from (A2) by σ yields:

$$\frac{\rho(\sigma)}{\sigma} = \frac{\beta - (\sigma - 1)(1 - \alpha - \beta)}{\sigma(1 - \beta)} - \frac{\alpha((\sigma - 1)(1 - \alpha - \beta) - 1)}{\sigma(1 - \beta)(1 - \alpha - \beta)} \frac{\Gamma^\sigma}{p^{\sigma-1} + \Gamma^\sigma} \quad (\text{A4})$$

We have $1 > \frac{\Gamma^\sigma}{p^{\sigma-1} + \Gamma^\sigma} > 0$. Evaluating (A4) using the limiting values of $\frac{\Gamma^\sigma}{p^{\sigma-1} + \Gamma^\sigma}$ as p goes to zero or infinity, we have for $\frac{\Gamma^\sigma}{p^{\sigma-1} + \Gamma^\sigma} = 1$:

$$\frac{(\beta - (\sigma - 1)(1 - \alpha - \beta))(1 - \alpha - \beta) - \alpha((\sigma - 1)(1 - \alpha - \beta) - 1)}{\sigma(1 - \beta)(1 - \alpha - \beta)} = \frac{1}{\sigma(1 - \alpha - \beta)} - 1$$

for $\frac{\Gamma^\sigma}{p^{\sigma-1} + \Gamma^\sigma} = 0$:

$$\frac{\beta - (\sigma - 1)(1 - \alpha - \beta)}{\sigma(1 - \beta)} = \frac{1 - \alpha}{\sigma(1 - \beta)} - \frac{(1 - \alpha - \beta)}{(1 - \beta)}$$

So, in both cases, and therefore, in intermediate cases as well, $\frac{\rho(\sigma)}{\sigma} > -1$.

Appendix 5: Exploring when an increased elasticity of substitution, σ , decreases $\frac{\rho(\sigma)}{\sigma}$

Here we explore the relevant values of σ , mentioned at the end of Section 3, for which $\frac{\partial(\rho(\sigma)/\sigma)}{\partial \sigma} < 0$. Since $\frac{\partial(\rho(\sigma)/\sigma)}{\partial \sigma} = \frac{\partial \rho}{\partial \sigma} - \frac{\rho}{\sigma^2} < 0$ if $\frac{\partial \rho}{\partial \sigma} < 0$, we investigate whether $\frac{\partial \rho}{\partial \sigma} < 0$ for the parameter values and range of p needed for the pre-industrial stagnation analysis in Section 4.2, using (A2) for the supply elasticity $\rho(\sigma) \equiv \frac{\partial \ln y}{\partial \ln p}$, whence:

$$\frac{\partial \rho(\sigma)}{\partial \sigma} = \frac{-(1 - \alpha - \beta)}{1 - \beta} - \frac{\alpha}{(1 - \beta)} \frac{\Gamma^\sigma}{p^{\sigma-1} + \Gamma^\sigma} - \frac{\alpha((\sigma - 1)(1 - \alpha - \beta) - 1)}{(1 - \beta)(1 - \alpha - \beta)} \frac{\Gamma^\sigma p^{\sigma-1} (\ln \Gamma - \ln p)}{(p^{\sigma-1} + \Gamma^\sigma)^2} \quad (\text{A5})$$

$\frac{\partial \rho(\sigma)}{\partial \sigma}$ is definitely negative for $p < \Gamma$ and $\sigma > 1 + \frac{1}{1 - \alpha - \beta}$ and *vice versa*. These are the relevant conditions for our discussion in Section 4.2 of Pre-industrial Stagnation, hence our conclusion

in Section 3 that "for...parameter and price values...relevant to the Pre-industrial Stagnation behavior discussed in Section 4.2... $\frac{\rho(\sigma)}{\sigma}$ decreases...with increasing σ ."

We also note that it is negative for a much broader range of values than this as shown by the supply curve shifting from positively sloped to negatively sloped in the previous section. Applying L'Hôpital's rule to (A5) we find:

$$\lim_{p \rightarrow \infty} \frac{\partial \rho}{\partial \sigma} = -\frac{1 - \alpha - \beta}{1 - \beta}, \quad \lim_{p \rightarrow 0} \frac{\partial \rho}{\partial \sigma} = -1$$

As $\rho(\sigma)$ is monotonic in p (see (A2)) if $\rho(\sigma)$ declines sufficiently as σ increases, for example from positive to negative at both extreme values of p , it must also do so for all intermediate values of p . Such changes must also preserve the relevant concavity or convexity properties of (A2). However, this does not preclude $\rho(\sigma)$ locally and temporarily with increasing σ for some intermediate values of p . Evaluating (A5) numerically for different parameter values shows that it is usually negative, though it is possible for σ close to one to get a positive derivative for some range of low values of p . So, there are minor exceptions to our statement above that $\rho(\sigma)/\sigma$ declines with increasing σ .

Appendix 6: Derivation of Equation (45) for $\Delta y_t = 0$ Isoclines in Figures 6a and 6b

Inserting $L_{S,t}(y_t)$ from (15) and $p_{S,t}(y_t)$ from (7) into (22) for coal use:

$$\begin{aligned} \Rightarrow E_{S,t} &= \left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S} \right)^{\frac{1-\beta}{1-\alpha-\beta}} \left[(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}) (1 - \gamma)^\sigma \right]^{\frac{1}{\sigma-1} \left(\frac{1}{1-\alpha-\beta} \right)} \frac{L_t}{1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}} \\ \Rightarrow E_t &= \frac{\bar{E}_M}{E_{S,t}} = \left(\frac{\beta \bar{e}_S}{\alpha N_{S,t}} \right)^{\frac{1-\beta}{1-\alpha-\beta}} \frac{\bar{E}_M}{L_t} \frac{(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}})^{1 - \frac{1}{\sigma-1} \left(\frac{1}{1-\alpha-\beta} \right)}}{(1 - \gamma)^{\frac{\sigma}{\sigma-1} \left(\frac{1}{1-\alpha-\beta} \right)}} \quad (\text{B6a}) \end{aligned}$$

Substituting this into (21) we have:

$$y_t^{1+\alpha(\sigma-1)} = \Gamma^{(1-\alpha)\sigma} N_t^{(1-\beta)\sigma} \left(\frac{\beta \bar{e}_S}{\alpha N_{S,t}} \right)^{\frac{\alpha\sigma(1-\beta)}{1-\alpha-\beta}} \left(\frac{\bar{E}_M}{L_t} \right)^{\alpha\sigma} \frac{(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}})^{\alpha\sigma \left(1 - \frac{1}{\sigma-1} \left(\frac{1}{1-\alpha-\beta} \right) \right)}}{(1 - \gamma)^{\frac{\alpha\sigma^2}{\sigma-1} \left(\frac{1}{1-\alpha-\beta} \right)}} \quad (\text{A6})$$

Taking logs then differences, and substituting $\Delta y_t/y_t = \Delta \ln(y_t)$ gives (see the Annex at the end):

$$\begin{aligned} \frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}}}{1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}} \Delta \ln(y_t) &= \sigma(1 - \beta) \Delta \ln(N_{M,t}) - \sigma \frac{(1 - \beta)^2}{1 - \alpha - \beta} \Delta \ln(N_{S,t}) \\ &\quad - \alpha\sigma \Delta \ln(L_t) \end{aligned} \quad (\text{A7})$$

Using $\Delta \ln(N_{M,t}) = n_t \Delta \ln(N_{S,t}) = n_t \Delta \ln(N_t)/(n_t - 1)$ then gives, after further algebra (again see the Annex):

$$\begin{aligned} & \frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta}\right)\Gamma y_t^{\frac{\sigma-1}{\sigma}}}{1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}}(n_t - 1) \Delta \ln(y_t) \\ &= \sigma(1 - \beta) \left(n_t - \frac{1 - \beta}{1 - \alpha - \beta} \right) \Delta \ln(N_t) - (n_t - 1)\alpha\sigma \Delta \ln(L_t) \end{aligned} \quad (\text{A8})$$

With *constant population*, $\Delta \ln(L_t) = 0$, we have from (A8) and (32):

$$\Delta y_t \gtrless 0 \iff n_t \gtrless \frac{1 - \beta}{1 - \alpha - \beta} > 1 \iff N_t \lesseqgtr \left(\frac{1 - \alpha - \beta}{1 - \beta} \right)^{1-\nu} \Gamma^\nu y_t^{\left(\frac{\sigma-1}{\sigma}\right)\nu} \quad (\text{A9})$$

with $\Delta y_t = 0$ being below the $\Delta N_t = 0$ isocline as shown in the figures; and since $\frac{1-\beta}{1-\alpha-\beta} > 1$, $\Delta y_t > 0$ below the $\Delta y_t = 0$ isocline and < 0 above it, also as shown. With *population growth*, the $\Delta y_t = 0$ isocline is given by $\sigma(1 - \beta)(n_t - \frac{1-\beta}{1-\alpha-\beta})\Delta \ln(N_t) - (n_t - 1)\alpha\sigma\Delta \ln(L_t) = 0$, so that:

$$\Delta y_t = 0 \text{ where } n_t > \frac{1 - \beta}{1 - \alpha - \beta} \text{ and thus } N_t < \left(\frac{1 - \alpha - \beta}{1 - \beta} \right)^{1-\nu} \Gamma^\nu y_t^{\left(\frac{\sigma-1}{\sigma}\right)\nu} \quad (\text{A9})$$

Appendix 7: Proof of Proposition 2 on Existence of Pre-Industrial Stagnation

To prove Proposition 2(a), we first need the following Lemma:

LEMMA. *Given High substitutability, constant population and Assumption 1:*

(i) *the locus of all points in (y, N) -space where $\Delta n_t = 0$ is*

$$\Gamma^{\frac{1-\alpha-\beta}{\alpha}} y_t^{\left(\frac{\sigma-1}{\sigma}\right)\frac{1-\alpha-\beta}{\alpha}} N_t^{-\left(\frac{1-\beta}{\alpha}\right)} = \frac{(\sigma - 1)(1 - \alpha - \beta) - 1 + [(\sigma - 1)(1 - \beta) - \left(\frac{1-\beta}{1-\alpha-\beta}\right)]\Gamma y_t^{\frac{\sigma-1}{\sigma}}}{(\sigma - 1)(1 - 2\alpha - \beta) - 1 + [(\sigma - 1)(1 - \alpha - \beta) - \left(\frac{1-\beta}{1-\alpha-\beta}\right)]\Gamma y_t^{\frac{\sigma-1}{\sigma}}} \quad (\text{A10})$$

(ii) *n falls along the $\Delta n_t = 0$ locus (A10) as it rises in (y, N) -space, and (A10) lies strictly below, and asymptotically (as $y \rightarrow \infty$) approaches, the locus defined by*

$$\Gamma^{\frac{1-\alpha-\beta}{\alpha}} y^{\left(\frac{\sigma-1}{\sigma}\right)\frac{1-\alpha-\beta}{\alpha}} N^{-\left(\frac{1-\beta}{\alpha}\right)} = \frac{(1 - \beta)(\sigma - \tilde{\sigma})}{(1 - \alpha - \beta)(\sigma - \sigma^\dagger)} \equiv n_\infty \quad (\text{A11})$$

Proof of Lemma: (i) We find the $\Delta n_t = 0$ locus in (y, N) -space by taking logs then the differences of (32) with Assumption 1 inserted, $n_t = \Gamma^{\frac{1-\alpha-\beta}{\alpha}} y_t^{\left(\frac{\sigma-1}{\sigma}\right)\frac{1-\alpha-\beta}{\alpha}} N_t^{-\left(\frac{1-\beta}{\alpha}\right)}$, and then setting $\Delta \ln(n_t) = 0$:

$$0 = \left(\frac{\sigma - 1}{\sigma} \right) \frac{1 - \alpha - \beta}{\alpha} \Delta \ln(y_t) - \frac{1 - \beta}{\alpha} \Delta \ln(N_t) \implies \frac{\Delta \ln(N_t)}{\Delta \ln(y_t)} = \left(\frac{\sigma - 1}{\sigma} \right) \frac{1 - \alpha - \beta}{1 - \beta} \quad (\text{A12})$$

Substitute this in (A8) with $\Delta \ln(L_t) = 0$, which relates the growth rate of y and N given constant population, and multiply by $\left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right)$:

$$\begin{aligned}
 & \left[1 + \alpha(\sigma - 1) + \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right] (n_t - 1) \Delta \ln(y_t) \\
 & = \left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right) \sigma(1 - \beta) \left(n_t - \frac{1 - \beta}{1 - \alpha - \beta} \right) \Delta \ln(N_t) \\
 & = \left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right) \sigma(1 - \beta) \left(n_t - \frac{1 - \beta}{1 - \alpha - \beta} \right) \left(\frac{\sigma - 1}{\sigma} \right) \frac{1 - \alpha - \beta}{1 - \beta} \Delta \ln(y_t)
 \end{aligned}$$

Divide by $\Delta \ln(y_t)$ and rearrange:

$$\begin{aligned}
 \Rightarrow & \left[1 + \alpha(\sigma - 1) + \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right] (n_t - 1) = \left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right) (1 - \alpha - \beta)(\sigma - 1) \left(n_t - \frac{1 - \beta}{1 - \alpha - \beta} \right) \\
 \Rightarrow & \left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right) (1 - \alpha - \beta)(\sigma - 1) \frac{1 - \beta}{1 - \alpha - \beta} - \left[1 + \alpha(\sigma - 1) + \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right] = \left[\left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right) (1 - \alpha - \beta)(\sigma - 1) - \left\{ 1 + \alpha(\sigma - 1) + \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right\} \right] n_t
 \end{aligned}$$

The $\Delta n_t = 0$ locus in (y, N) -space is thus:

$$\begin{aligned}
 n(y, N) & = \Gamma^{\frac{1-\alpha-\beta}{\alpha}} y_t^{\frac{(\sigma-1)(1-\alpha-\beta)}{\alpha}} N_t^{-\left(\frac{1-\beta}{\alpha}\right)} \\
 & = \frac{(\sigma - 1)(1 - \beta) \left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right) - \left[1 + \alpha(\sigma - 1) + \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right]}{(\sigma - 1)(1 - \alpha - \beta) \left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right) - \left[1 + \alpha(\sigma - 1) + \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right]} \\
 & = \frac{(\sigma - 1)(1 - \alpha - \beta) - 1 + \left[(\sigma - 1)(1 - \beta) - \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \right] \Gamma y_t^{\frac{\sigma-1}{\sigma}}}{(\sigma - 1)(1 - 2\alpha - \beta) - 1 + \left[(\sigma - 1)(1 - \alpha - \beta) - \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \right] \Gamma y_t^{\frac{\sigma-1}{\sigma}}} \quad (A10)
 \end{aligned}$$

(ii) Now by straightforward algebra (see Annex at the end), provided $\sigma > \sigma^\dagger$, n falls as y rises on locus (A10); and from (A10) and (43), the asymptotic lower bound of n as $y \rightarrow \infty$ is

$$\lim_{y \rightarrow \infty} n(y, N) = \frac{(\sigma-1)(1-\beta) - \left(\frac{1-\beta}{1-\alpha-\beta}\right)}{(\sigma-1)(1-\alpha-\beta) - \left(\frac{1-\beta}{1-\alpha-\beta}\right)} = \frac{(1-\beta)(\sigma-\tilde{\sigma})}{(1-\alpha-\beta)(\sigma-\sigma^\dagger)} \equiv n_\infty \quad (A11) \blacksquare$$

So, the $\Delta n_t = 0$ locus lies below the curve $n(y, N) = n_\infty$.

Proof of Proposition 2(a): Since $\tilde{\sigma} < \sigma^\dagger$, $n_\infty > \frac{(1-\beta)}{(1-\alpha-\beta)}$, hence the curve $n(y, N) = n_\infty$ lies below the $\Delta y_t = 0$ locus $n(y, N) = \frac{1-\beta}{1-\alpha-\beta}$; so all development paths below $n(y, N) = n_\infty$ have rising y_t . And by Lemma result (ii), n falls along the $\Delta n_t = 0$ locus as it rises from left to right in (y, N) -space, hence the locus crosses upwards over curves defined by $n(y, N) = \text{constant}$. By the definition of the locus, all development paths that cross it are locally tangent to the curve with $n(y, N) = \text{constant}$ at the point of crossing; and because the locus lies below $n(y, N) = n_\infty$, y_t is increasing along those paths. So, all development paths that cross the rising $\Delta n_t = 0$ locus do so from the left, above the locus, to the right, below it, which means they can never rise above the locus later. Such paths therefore have permanently rising y_t , the definition of Pre-industrial Stagnation, and moreover $\Delta n_t > 0$ forever once they cross the $\Delta n_t = 0$ locus. So there is a separatrix in (y, N) -space above the $\Delta n_t = 0$ locus but below the $\Delta y_t = 0$ locus, such that all development paths below the separatrix have forever rising y_t and eventually forever rising n_t , and all development paths above the separatrix eventually cross the $\Delta y_t = 0$ locus, with $\Delta y_t < 0$ thereafter. ■

Proof of Proposition 2(b): We first prove that $\sigma < \sigma^\dagger$ (Medium or Low substitutability) means all paths under the $\Delta y_t = 0$ locus in (y, N) -space eventually rise to cross the locus upwards. For this, we need to show that at any point under this locus, the slope of the path through that point is steeper than the curve $n(y, N) \equiv \Gamma^{\frac{1-\alpha-\beta}{\alpha}} y_t^{\frac{(\sigma-1)}{\sigma} \frac{1-\alpha-\beta}{\alpha}} N_t^{-\frac{(1-\beta)}{\alpha}} = \bar{n}$, where \bar{n} is a constant, through that point. That is, from (A9) and (A8), we need to show that:

$$\left\{ \bar{n} > \frac{1-\beta}{1-\alpha-\beta} (> 1) \text{ and } \sigma - 1 < \frac{1-\beta}{(1-\alpha-\beta)^2} \right\}$$

$$\Rightarrow \frac{\Delta \ln(N_t)}{\Delta \ln(y_t)} = \frac{\left[1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right] (\bar{n} - 1)}{\sigma(1-\beta) \left(\bar{n} - \frac{1-\beta}{1-\alpha-\beta} \right) \left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right)} + \frac{(\bar{n} - 1) \alpha \sigma \Delta \ln(L_t)}{\sigma(1-\beta) \left(\bar{n} - \frac{1-\beta}{1-\alpha-\beta} \right) \Delta \ln(y_t)}$$

$$> \frac{\sigma - 1}{\sigma \left(\frac{1-\beta}{1-\alpha-\beta} \right)}$$

Since $\bar{n} > \frac{1-\beta}{1-\alpha-\beta}$, $\Delta \ln(L_t) > 0$ always, and $\Delta \ln(y_t) > 0$ below the $\Delta \ln(y_t) = 0$ locus, the second term on the LHS is > 0 . So it will be enough just to prove that

$$\frac{\left[1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right] (\bar{n} - 1)}{\left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right) \left(\bar{n} - \frac{1-\beta}{1-\alpha-\beta} \right)} > (\sigma - 1)(1 - \alpha - \beta) \quad (\text{A13})$$

The proof of this by straightforward but tedious algebra is given in the Annex.

Then from (45) and (44), all paths above that isocline eventually cross the $\Delta N_t = 0$ isocline leftwards into the region where $\Delta N_t < 0$, $\Delta y_t < 0$ forever, as in figure 6a. ■

Bounding elasticity value, σ^\dagger . Finally, we show how from (46) we can derive σ^\dagger , the threshold value of σ for which the growth of the cost share ratio will first be able to outweigh the diminishing returns to knowledge. Expanding and rearranging (46) and using derivatives from Proposition 1:

$$\Delta \ln n_t = \left(\frac{1-\alpha-\beta}{\alpha} \frac{\sigma-1}{\sigma} \frac{\partial \ln y_t}{\partial \ln N_{M,t}} - \frac{1-\beta}{\alpha} \right) \Delta \ln N_{M,t}$$

$$- \left(\frac{1-\alpha-\beta}{\alpha} \frac{\sigma-1}{\sigma} \frac{1-\beta}{1-\alpha-\beta} \frac{\partial \ln y_t}{\partial \ln N_{M,t}} - \frac{1-\beta}{\alpha} \right) \Delta \ln N_{S,t} \quad (\text{A14})$$

Substituting in $\frac{\partial \ln y_t}{\partial \ln N_{M,t}} = \frac{\sigma}{\rho + \sigma}$ from (40) and (41):

$$\Rightarrow \Delta \ln n_t = \left(\frac{1-\alpha-\beta}{\alpha} \frac{\sigma-1}{\rho + \sigma} - \frac{1-\beta}{\alpha} \right) \Delta \ln N_{M,t}$$

$$- \left(\frac{1-\alpha-\beta}{\alpha} \frac{\sigma-1}{\sigma} \frac{1-\beta}{1-\alpha-\beta} \frac{\sigma}{\rho + \sigma} - \frac{1-\beta}{\alpha} \right) \Delta \ln N_{S,t} \quad (\text{A15})$$

The necessary conditions for $\Delta \ln n_t > 0$ is that the coefficient of $\Delta \ln N_{M,t}$ here is positive, and if it is close to zero then $\Delta \ln N_{S,t} \rightarrow 0$. In the bounding case where we set the coefficient to zero and cross-multiplying and simplifying, we have the condition:

$$(1 - \alpha - \beta)(\sigma^\dagger - 1) - (1 - \beta)(\rho + \sigma^\dagger) = 0$$

Now substituting in the formula for ρ for the case where $p = 0$ from (A3):

$$(1 - \alpha - \beta)(\sigma^\dagger - 1) - (1 - \beta) \frac{1}{1 - \alpha - \beta} = 0$$

$$\Rightarrow \sigma^\dagger = 1 + \frac{(1-\beta)}{(1-\alpha-\beta)^2}. \quad \blacksquare$$

As σ increases, at first the zone of preindustrial stagnation will appear in the bottom right of figure 6b where very relatively abundant wood results in relative output, y , being very high given relative knowledge stocks, N .

Appendix 8: Proof of Proposition 3

Proof of Proposition 3(a), Existence of Modern Economic Growth zone

Taking differences of the log of (47) gives $\Delta \ln(y_t) = \sigma[(1 - \beta)\Delta \ln(N_t) - \alpha \Delta \ln(e_t)]$, and substituting this, and $\Gamma y_t^{\frac{\sigma-1}{\sigma}} = N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}$ from (32), into (A8) gives:

$$\sigma \left[\frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta}\right) N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right] (n_t - 1) [(1 - \beta)\Delta \ln(N_t) - \alpha \Delta \ln(e_t)]$$

$$\approx \sigma \left[(1 - \beta) \left(n_t - \frac{1 - \beta}{1 - \alpha - \beta} \right) \Delta \ln(N_t) - (n_t - 1) \alpha \Delta \ln(L_t) \right] \quad (\text{A16})$$

After much further algebra in the Annex at the end, this yields:

$$\left(\frac{1 + \alpha(\sigma - 1)}{1 - \beta} + \frac{N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 - \alpha - \beta} \right) (n_t - 1) \Delta \ln(e_t)$$

$$= \left[\left\{ \sigma - \tilde{\sigma} + (\tilde{\sigma} - 1) \left(1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}} \right) \right\} n_t - (\sigma - \tilde{\sigma}) \right] \Delta \ln(N_t) + \left(\frac{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 - \beta} \right) (n_t - 1) \Delta \ln(L_t) \quad (\text{A17})$$

From rearranging (A16) with $\Delta \ln(L_t) = 0$, $\Delta e_t = 0$ when:

$$n_t = \frac{\sigma - \tilde{\sigma}}{\sigma - \tilde{\sigma} + (\tilde{\sigma} - 1) \left(1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}} \right)}, < 1 \text{ when } \sigma > \tilde{\sigma} \quad (\text{A18})$$

and given $\sigma > \tilde{\sigma}$, this does have a solution with $0 < n_t < 1$ ($\Delta N_t < 0$) for any permitted parameter values. From (A18), Assumption 1 (35), and (48) (which uses Assumption 1):

$$n_t = \frac{\sigma - \tilde{\sigma}}{\sigma - \tilde{\sigma} + (\tilde{\sigma} - 1) \left(1 + N_t^{\frac{1-\beta}{1-\alpha-\beta}} n_t^{\frac{\alpha}{1-\alpha-\beta}} \right)} = \Gamma^{\frac{\sigma(1-\alpha-\beta)}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)} N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}}$$

$$\Rightarrow \frac{\sigma - \tilde{\sigma}}{\Gamma^{\frac{\sigma(1-\alpha-\beta)}{\alpha}}}$$

$$= N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)} \left[\sigma - 1 + (\tilde{\sigma} - 1) \Gamma^{\sigma} e_t^{-\alpha(\sigma-1)} N_t^{(1-\beta)(\sigma-\tilde{\sigma}) + \frac{1-\beta}{1-\alpha-\beta}} \right]$$

Now take total differences, using $(1 - \beta)(\sigma - \tilde{\sigma}) + \frac{1-\beta}{1-\alpha-\beta} = (\sigma - 1)(1 - \beta)$:

$\Rightarrow 0 =$

$$\begin{aligned}
& \frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha} N_t^{\left[\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}\right]-1} e_t^{-(\sigma-1)(1-\alpha-\beta)} \left[\sigma - 1 \right. \\
& \quad \left. + (\tilde{\sigma} - 1) \Gamma^\sigma N_t^{(\sigma-1)(1-\beta)} e_t^{-\alpha(\sigma-1)} \right] \Delta N_t \\
& - (\sigma - 1)(1 - \alpha - \beta) N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)-1} \left[\sigma - 1 \right. \\
& \quad \left. + (\tilde{\sigma} - 1) \Gamma^\sigma N_t^{(\sigma-1)(1-\beta)} e_t^{-\alpha(\sigma-1)} \right] \Delta e_t \\
& + N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)} (\tilde{\sigma} - 1) \Gamma^\sigma [(\sigma - 1)(1 - \beta)] N_t^{(\sigma-1)(1-\beta)-1} e_t^{-\alpha(\sigma-1)} \Delta N_t \\
& - N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)} (\tilde{\sigma} - 1) \Gamma^\sigma \alpha (\sigma - 1) N_t^{(\sigma-1)(1-\beta)} e_t^{-\alpha(\sigma-1)-1} \Delta e_t \\
& \Rightarrow \left\{ \frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha} N_t^{\left[\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}\right]-1} e_t^{-(\sigma-1)(1-\alpha-\beta)} \left[\sigma - 1 + (\tilde{\sigma} - \right. \right. \\
& \left. \left. 1) \Gamma^\sigma N_t^{(\sigma-1)(1-\beta)} e_t^{-\alpha(\sigma-1)} \right] + N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)} (\tilde{\sigma} - 1) \Gamma^\sigma [(\sigma - 1)(1 - \right. \\
& \left. \beta)] N_t^{(\sigma-1)(1-\beta)-1} e_t^{-\alpha(\sigma-1)} \right\} \Delta N_t \\
& = \left\{ \frac{\alpha(\sigma-1)}{h} N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)-1} \left[\sigma - 1 + (\tilde{\sigma} - \right. \right. \\
& \left. \left. 1) \Gamma^\sigma N_t^{(\sigma-1)(1-\beta)} e_t^{-\alpha(\sigma-1)} \right] + N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)} (\tilde{\sigma} - 1) \Gamma^\sigma \alpha (\sigma - \right. \\
& \left. 1) N_t^{(\sigma-1)(1-\beta)} e_t^{-\alpha(\sigma-1)-1} \right\} \Delta e_t \quad (A19)
\end{aligned}$$

The bracketed expressions multiplying ΔN_t and Δe_t are both unambiguously positive, so $\Delta N_t / \Delta e_t > 0$, i.e. the isocline is upward sloping. That $\Delta e_t > 0$ above the isocline and < 0 below it then follows from the signs in (A17). ■

Proof of Proposition 3(b) Strong Relative Bias in Modern Economic Growth zone

By part (a), the Modern Economic Growth zone will lie strictly below the $\Delta N_t = 0$ isocline in (e, N) -space, so y_t is falling throughout the zone. Now consider coal use (22) expressed as a function of $(y_t, N_{S,t})$:

$$E_{S,t}(y_t, N_{S,t}) = \left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S} \right)^{\frac{1-\beta}{1-\alpha-\beta}} p_{S,t}^{\frac{1}{1-\alpha-\beta}}(y_t) L_{S,t}(y_t)$$

Substituting $p_{S,t}(y_t)$ from (7) and $L_{S,t}(y_t)$ from (15) converts this to

$$E_{S,t}(y_t, N_{S,t}) = \left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S} \right)^{\frac{1-\beta}{1-\alpha-\beta}} (1 - \gamma)^{\frac{\sigma}{\sigma-1} \left(\frac{1}{1-\alpha-\beta} \right)} \left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right)^{-\left(\frac{\sigma-\tilde{\sigma}}{\sigma-1} \right)} L_t, \quad (A20)$$

which is rising, since $N_{S,t}$ always rises, y_t is falling, and $\sigma > \tilde{\sigma} > 1$. So we have rising relative coal use $E_{S,t} / \bar{E}_M$ despite a rising relative coal price $\bar{e}_S / e_{M,t}$ (since $e_t \equiv e_{M,t} / \bar{e}_S$ is falling, by definition of a MEG zone).

Appendix 9: Result on Strong Relative Bias in Pre-industrial Stagnation Zone

Because we already know that e is increasing unless we are below the $\Delta e = 0$ isocline in figures 7a and 7b, in order to determine where strong bias to wood can happen, we just need to check when coal will decline as p declines (and so y is increasing). Looking at the optimal use of coal again:

$$E_{S,t}(p_t, N_{S,t}) = \left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S} \right)^{\frac{1-\beta}{1-\alpha-\beta}} p_{S,t}^{\frac{1}{1-\alpha-\beta}}(p_t) L_{S,t}(p_t), \quad (22)$$

any positive $\Delta N_{S,t}$ increases optimal coal use, *ceteris paribus*. The increasing price of Solow goods, $p_{S,t}$, as the economy becomes more Malthus specialized also will raise coal use, *ceteris paribus*. We therefore need to have a strong enough decline in labor used in the Solow sector in order to overcome these positive effects and allow coal to decline. For this to happen the elasticity of substitution between the goods must be high enough to allow enough substitution from the Solow good to the Malthus good to allow enough flow of labor between the sectors. Using the formulae for $p_{S,t}(p_t)$ and $L_{S,t}(p_t)$, (7) and (15), we have:

$$\begin{aligned} p_{S,t}^{\frac{1}{1-\alpha-\beta}}(p_t) L_{S,t}(p_t) &= (1-\gamma)^{\frac{\sigma}{\sigma-1}} L_t \frac{(1 + \Gamma^\sigma p_t^{1-\sigma})^{\frac{1}{(\sigma-1)(1-\alpha-\beta)}}}{1 + \Gamma^\sigma p_t^{1-\sigma}} \\ &= \phi(1 + \Gamma^\sigma p_t^{1-\sigma})^{\frac{1-(\sigma-1)(1-\alpha-\beta)}{(\sigma-1)(1-\alpha-\beta)}} \equiv \phi z \end{aligned}$$

Next determine $\frac{dz}{dp}$:

$$\frac{dz}{dp} = \frac{1 - (\sigma - 1)(1 - \alpha - \beta)}{(\sigma - 1)(1 - \alpha - \beta)} (1 - \sigma) (1 + \Gamma^\sigma p_t^{1-\sigma})^{\frac{1-2(\sigma-1)(1-\alpha-\beta)}{(\sigma-1)(1-\alpha-\beta)}} \Gamma^\sigma p_t^{-\sigma}$$

If $\sigma > \tilde{\sigma} \equiv 1 + \frac{1}{1-\alpha-\beta}$, then $\frac{dz}{dp} > 0$. So, these terms are declining with declining p (and, therefore rising y) as long as $\sigma > \tilde{\sigma}$. A minor rearrangement of equation (34) (using assumption 1, so that $N_{S,t}^{\frac{\alpha}{1-\alpha-\beta}(\frac{\nu}{1-\nu})-1} = 1$):

$$\frac{\Delta N_{S,t}}{N_{S,t}} = \left(\eta^{\frac{1}{\nu}} \nu (1 - \beta) \right)^{\frac{\nu}{1-\nu}} \left(\frac{\alpha}{\beta \bar{e}_S} \right)^{\frac{\alpha \nu}{(1-\nu)(1-\alpha-\beta)}} \left(p_{S,t}^{\frac{1}{1-\alpha-\beta}} L_{S,t} \right)^{\frac{1-\nu}{\nu}} \quad (A21)$$

shows that exactly the same term, z , drives $\Delta N_{S,t}$. Therefore, if $\sigma > \tilde{\sigma}$ and p is declining, $N_{S,t}$ will eventually converge to a constant and coal use can decline.

It is harder to determine over what zone of the phase plane there is strong bias to wood. A simple comparison of the direct and indirect effects of z on $E_{S,t}$ in (22) is difficult because the level of z affects $E_{S,t}$ both directly and via the change in $N_{S,t}$ through (34). Using the comparative statics framework in Section 3 to find the effect of an exogenous change in $N_{S,t}$ on $E_{S,t}$ will also need to take into account how much $N_{M,t}$ changes, because that also will change the price ratio, p .

However, unless $\sigma > \sigma^\dagger$, all paths which initially have falling coal use eventually end up with rising coal use. This is because, initially, the price ratio, p , was falling strongly enough to reduce

$E_{S,t}$ by more than growing $N_{S,t}$ increases it. But the rate of decline in p slows and the level of z remains sufficiently high that through (34) technical change can dominate and increase $E_{S,t}$ again.

We derive equation (51) by first rearranging $e_t(E_t, N_t) = \Gamma^{\sigma/\theta} E_t^{-1/\theta} N_t^{(1-\beta)(\sigma-1)/\theta}$ (50) into $N_t(e_t, E_t) = \Gamma^{\frac{\sigma}{(1-\beta)(\sigma-1)}} e_t^{\frac{\theta}{(1-\beta)(\sigma-1)}} E_t^{\frac{1}{(1-\beta)(\sigma-1)}}$. Inserting this into $n_t(e_t, N_t) = \Gamma^{\frac{\sigma(1-\alpha-\beta)}{\alpha}} e_t^{-(\sigma-1)(1-\alpha-\beta)} N_t^{\frac{(1-\beta)(1-\alpha-\beta)(\sigma-\tilde{\sigma})}{\alpha}}$ (48), and rearranging, gives

$$n_t(e_t, E_t) = \Gamma^{\frac{\sigma}{\alpha(\sigma-1)}} e_t^{\frac{(\sigma-1)(1-2\alpha-\beta)-1}{\alpha(\sigma-1)}} E_t^{\frac{(\sigma-1)(1-\alpha-\beta)-1}{\alpha(\sigma-1)}} \quad (51)$$

Appendix 10: Proof of Proposition 4 on Asymptotic Growth Rates

Here, we denote growth rates and asymptotic growth rates for variable X_t thus:

$$\frac{\Delta X_t}{X_t} \equiv g(X_t) \text{ and } \lim_{t \rightarrow \infty} \frac{\Delta X_t}{X_t} \equiv g_{\infty}(X_t),$$

with *PS* and *IR* subscripts added as needed.

By definition $y_t \rightarrow \infty$ under PS; and by (6) and (14):

$$\lim_{y_t \rightarrow \infty} p_{M,t} = \lim_{y_t \rightarrow \infty} (1 - \gamma)^{\frac{\sigma}{\sigma-1}} \Gamma \left(y_t^{-\frac{\sigma-1}{\sigma}} + \Gamma \right)^{\frac{1}{\sigma-1}} = (1 - \gamma)^{\frac{\sigma}{\sigma-1}} \Gamma^{\frac{\sigma}{\sigma-1}} \text{ and}$$

$$\lim_{y_t \rightarrow \infty} L_{M,t} = \lim_{y_t \rightarrow \infty} \frac{L_t \Gamma}{\Gamma + y_t^{-\frac{\sigma-1}{\sigma}}} = L_{\infty}$$

and inserting these limits into (52) gives:

$$\begin{aligned} g_{\infty PS}(N_{M,t}) &\equiv \lim_{t \rightarrow \infty, PS} \frac{\Delta N_{M,t}}{N_{M,t}} = \lim_{t \rightarrow \infty, PS} \lambda_M N_{M,t}^{-1} (p_{M,t} \bar{E}_M^{\alpha} L_{M,t}^{1-\alpha-\beta})^{\frac{1-\alpha-\beta}{\alpha(1-\beta)}} \\ &= \left(\lim_{t \rightarrow \infty, PS} N_{M,t}^{-1} \right) \lambda_M^{\frac{1-\alpha-\beta}{\alpha}} [(1 - \gamma)^{\frac{\sigma}{\sigma-1}} \Gamma^{\frac{\sigma}{\sigma-1}}]^{\frac{(1-\alpha-\beta)}{\alpha(1-\beta)}} L_{\infty}^{\frac{(1-\alpha-\beta)^2}{\alpha(1-\beta)}} = 0 \end{aligned} \quad (A22)$$

By definition $y_t \rightarrow 0$ under IR, and by (7) and (15),

$$\lim_{y_t \rightarrow 0} p_{S,t} = \lim_{y_t \rightarrow 0} (1 - \gamma)^{\frac{\sigma}{\sigma-1}} \Gamma (1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}})^{\frac{1}{\sigma-1}} = (1 - \gamma)^{\frac{\sigma}{\sigma-1}} \implies g_{\infty IR}(p_{S,t}) = 0 \quad (A23)$$

$$\lim_{y_t \rightarrow 0} L_{S,t} = \lim_{y_t \rightarrow 0} \frac{L_t}{\Gamma y_t^{\frac{\sigma-1}{\sigma}} + 1} = L_{\infty} \quad (A24)$$

and inserting these limits into (53) gives:

$$g_{\infty IR}(N_{S,t}) \equiv \lim_{t \rightarrow \infty, IR} \frac{\Delta N_{S,t}}{N_{S,t}} \equiv \lim_{t \rightarrow \infty, IR} n_{S,t} \equiv n_{S \infty IR} = \lambda_S (1 - \gamma)^{\frac{\sigma}{\sigma-1}(\frac{1}{\alpha})} L_{\infty}^{\frac{1-\alpha-\beta}{\alpha}} \quad (A25)$$

Next, we find the growth rates of labor productivity (output per capita) for the Malthus and Solow sectors. For the Malthus sector, substituting (6) for $p_{M,t}(y_t)$ and (14) for $L_{M,t}(y_t)$ into (19) for $Y_{M,t}$ and then rearranging, gives for some constant $\phi_M > 0$ (see the Annex at the end):

$$\frac{Y_{M,t}}{L_{M,t}} = \phi_M N_{M,t} (y_t^{-\frac{\sigma-1}{\sigma}} + \Gamma)^{\frac{(\sigma-1)\alpha+\beta}{(\sigma-1)(1-\beta)}} L_t^{-\frac{\alpha}{1-\beta}} \quad (\text{A26})$$

which, using the limits noted above

$$\Rightarrow g_{\infty PS} \left(\frac{Y_{M,t}}{L_{M,t}} \right) = g_{\infty PS}(N_{M,t}) + \frac{(\sigma-1)\alpha+\beta}{(\sigma-1)(1-\beta)} 0 - \frac{\alpha}{1-\beta} 0 = 0 \quad (\text{A27})$$

For the Solow sector, substituting (7) for $p_{S,t}(y_t)$ and (22) for $E_{S,t}$ into (20) for $Y_{S,t}$ and rearranging yields, for some constant $\phi_S > 0$ (again see the Annex):

$$\frac{Y_{S,t}}{L_{S,t}} = \phi_S N_{S,t}^{\frac{1-\beta}{1-\alpha-\beta}} (1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}})^{\frac{\alpha+\beta}{(\sigma-1)(1-\alpha-\beta)}} \quad (\text{A28})$$

which by (A25) and other limits

$$\Rightarrow g_{\infty IR} \left(\frac{Y_{S,t}}{L_{S,t}} \right) = \frac{1-\beta}{1-\alpha-\beta} n_{S\infty IR} = \left(\frac{1-\beta}{1-\alpha-\beta} \right) \lambda_S (1-\gamma)^{\frac{\sigma-1}{\sigma-1}(\frac{1}{\alpha})} L_{\infty}^{\frac{1-\alpha-\beta}{\alpha}} \quad (\text{55})$$

And since $y_t \rightarrow 0$ on an IR path, $L_{S,t} \rightarrow L_t$ and (by (1)) $Y_t \rightarrow (1-\gamma)^{\frac{\sigma-1}{\sigma}} Y_{S,t}$, hence $g_{\infty IR} \left(\frac{Y_{S,t}}{L_{S,t}} \right) = g_{\infty IR} \left(\frac{Y_t}{L_t} \right)$, the economy's "growth rate" (i.e. of final output per capita), which by (54) is asymptotically positive. Similar algebra shows that since $y_t \rightarrow \infty$ on an PS path, economic growth $g_{\infty PS} \left(\frac{Y_t}{L_t} \right) = g_{\infty PS} \left(\frac{Y_{M,t}}{L_{M,t}} \right) = 0$ by (A27). ■

Appendix 11: Details of Simulation

For our historical baseline and counterfactual simulation scenarios we provide the exogenous population input parametrically. We refitted Marchetti *et al.*'s (1996) bilogistic function model using Broadberry *et al.*'s (2015) data for the population of the United Kingdom at 20-year intervals through 2000, resulting in the following fit:

$$S_{\tau} = \frac{9.7}{1 + \exp \left(-\frac{\ln(81)}{267} (\tau - 1530) \right)} + \frac{47.4}{1 + \exp \left(-\frac{\ln(81)}{171} (\tau - 1870) \right)} \rightarrow 57.1 \equiv S_{\infty} \text{ as } \tau \rightarrow \infty \quad (\text{A29})$$

where τ is the calendar year and population S_{τ} is measured in millions. Then we assume that the total (normalized) labor force is given by $L_t \equiv S_t/S_1$ (so $L_{\infty} \equiv S_{\infty}/S_1$), where time t counts 20-year periods from $t = 1$ in 1560, the first year of Warde's (2007) energy data, to $t = 18$ in 1900, so that $t = (\tau - 1540)/20$.

The simulations use a normalized CES production function:

$$\frac{Y_t}{Y_b} = \left[\gamma \left(\frac{Y_{M,t}}{Y_{M,b}} \right)^{\frac{\sigma-1}{\sigma}} + (1-\gamma) \left(\frac{Y_{S,t}}{Y_{S,b}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \quad (\text{A30})$$

where Y_b is final output and $Y_{M,b}$ and $Y_{S,b}$ are the sectoral outputs in the first period of the Baseline simulation, γ is the share parameter corresponding to normalized sectoral outputs $Y_{M,t}/Y_{M,b}$ and $Y_{S,t}/Y_{S,b}$,¹ whose prices are now defined in terms of normalized final output Y_t/Y_b , and other equations are modified accordingly. This ensures that comparisons involving changes in the elasticity of substitution are meaningful (Klump *et al.*, 2012). As under

¹ In general, normalization results in γ in (1) and γ in (A30) having different numerical values.

normalization $p_1 = \Gamma$, the equations for the first period of the Baseline simulation require solving for only $N_{M,1}$ and $N_{S,1}$. The values of $Y_{M,1}$ and $Y_{S,1}$ are then treated as parameters in the remaining periods of the Baseline and for all periods of the counterfactual simulations.

Our full set of Baseline parameters is shown in table 1. We take the cost share of energy in 1800 in Britain to be around 25% not including human and animal power (Gentvilaite *et al.*, 2015), so we set the energy output elasticity to $\alpha = 0.25$. Gentvilaite *et al.* (2015) show that the energy cost share declined to about 15% in 1900. The Cobb-Douglas production functions in each of the two sectors abstract from this reality. We set the output elasticity of machines to $\beta = 0.225$ based on table 13 in Clark (2010). We normalize the quantity of wood, \bar{E}_M , and the stock of machine varieties in the Solow sector in 1540 ($t = 0$), $N_{S,0}$, to 1. We set $\sigma = 4$, close to Kander and Stern's (2014) estimate and above $\sigma^\dagger (= 3.81$ given our choices of α and β), that is, 'High'.

The remaining parameters are $N_{M,0}$, η , \bar{e}_S , and γ . We optimize these by minimizing the sum of squared proportional deviations from six Stylized Facts: two based on the initial state in Britain in 1560, and four based on the change in the variables over its Industrial Revolution.² Using calendar year time subscripts, the chosen stylized facts and proportional deviations are respectively:

1. The price of wood is double the price of coal in 1560 (Allen, 2009): $\ln\left(\frac{e_{M,1560}}{e_{S,1560}}\right) - \ln(2)$.
2. Coal use is 30% of wood use in 1560 (Warde, 2007): $\ln\left(\frac{E_{S,1560}}{E_M}\right) - \ln(0.3)$.
3. The price of wood doubles from 1560 to its peak (Allen, 2009): $\ln\left(\frac{\max(e_{M,t})}{e_{M,1560}}\right) - \ln(2)$.
4. Energy intensity doubles from 1560 to 1900 (Warde, 2007; Broadberry *et al.*, 2015): $\ln\left(\frac{(E/Y)_{1900}}{(E/Y)_{1560}}\right) - \ln(2)$. This reflects the increase in total energy intensity in figure 4. We tried instead using a ratio of 4 in our optimization, to reflect the increase in firewood and coal energy intensity, but this resulted in a much poorer fit to the other stylized facts.
5. Output per capita rises 5.4-fold from 1560 to 1900 (Broadberry *et al.*, 2015): $\ln\left(\frac{(Y/L)_{1900}}{(Y/L)_{1560}}\right) - \ln(5.4)$.
6. Output per capita doubles from 1560 to 1800 (Broadberry *et al.*, 2015): $\ln\left(\frac{(Y/L)_{1800}}{(Y/L)_{1560}}\right) - \ln(2)$.

The model fits Stylized Facts 1 to 5 very well, but Fact 6 less well. The latter is perhaps not surprising given how simplified the model is – for example assuming equal parameter values in each sector – and how much uncertainty there is about the values of those parameters. We can fit some of the stylized facts better only at the expense of fitting others more poorly. For example, by reducing the initial number of varieties in the Malthus sector and therefore increasing the rate of growth in the Malthus sector we can obtain a more realistic trajectory for economic growth with more rapid early growth and slower later growth. However, because demand for wood remains stronger the price of wood does not peak in these scenarios. By contrast to obtain earlier more rapid growth in coal use we end up with even more growth

² As part of these optimizations, at first, we allowed different innovation productivities, η_M and η_S , but found that doing so gave very little reduction in the sum of squared deviations. This strengthens our case for using a common η , to avoid unhelpful theoretical complexity.

concentrated into the 19th Century. Of course, if we allowed more parameters to vary freely rather than be set a priori, we could obtain a better fit to the data.

Table 1. Baseline Parameters

Parameter	Symbol	Value	Sources
CES elasticity in final production	σ	4	Greater than σ^+
Distribution parameter in CES final production	γ	0.853	Optimized
Energy output elasticity	α	0.25	Energy cost share in 1800 in the UK was about 25% not counting animal and human power (Gentvilaite et al., 2015).
Capital (machine) output elasticity	β	0.225	This is based on a share of capital that fluctuates between about 0.2 and 0.25 in Clark (2010).
Productivity innovation in M sector	η_M	2.26	Optimized
Productivity innovation in S sector	η_S	2.26	Optimized
Initial idea stock in M sector	$N_{M,0}$	5	Optimized
Initial idea stock in S sector	$N_{S,0}$	1	Normalized
Constant price of coal	\bar{e}_S	0.115	Optimized
Constant consumption of wood	\bar{E}_M	1	Normalized

Additional References

- Klump, Rainer, Peter McAdam, and Alpo Willman. 2012. The normalized CES production function: theory and empirics. *Journal of Economic Surveys* 26(5): 769–99.
- Marchetti, Cesare, Perrin S. Meyer, and Jesse H. Ausubel. 1996. Human population dynamics revisited with the logistic model: How much can be modeled and predicted? *Technological Forecasting and Social Change* 52: 1–30.

ANNEX (FOR ONLINE PUBLICATION) of Directed Technical Change and the British Industrial Revolution

Derivations in Appendix 6 on Derivation of $\Delta y_t = 0$ Isoclines

Steps from (A6) to (A7)

Take logs then differences of (A6):

$$\begin{aligned}
 [1 + \alpha(\sigma - 1)]\Delta \ln(y_t) &= \sigma(1 - \beta) \Delta (\ln(N_{M,t}) - \ln(N_{S,t})) - \frac{\alpha\sigma(1 - \beta)}{1 - \alpha - \beta} \Delta \ln(N_{S,t}) \\
 &\quad - \alpha\sigma \Delta \ln(L_t) + \alpha \left(1 - \frac{1}{\sigma - 1} \left(\frac{1}{1 - \alpha - \beta} \right) \right) \frac{\Gamma(\sigma - 1)y_t^{\frac{\sigma-1}{\sigma}} \Delta y_t}{\left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right)} \\
 [1 + \alpha(\sigma - 1)]\Delta \ln(y_t) &= \sigma(1 - \beta) \Delta (\ln N_{M,t} - \ln N_{S,t}) - \frac{\alpha\sigma(1 - \beta)}{1 - \alpha - \beta} \Delta \ln(N_{S,t}) \\
 &\quad - \alpha\sigma \Delta \ln(L_t) + \alpha \left[\sigma - 1 - \left(\frac{1}{1 - \alpha - \beta} \right) \right] \frac{\Gamma y_t^{\frac{\sigma-1}{\sigma}} \Delta y_t / y_t}{\left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right)}
 \end{aligned}$$

Substituting $\Delta y_t / y_t = \Delta \ln(y_t)$ and rearranging:

$$\begin{aligned}
 [1 + \alpha(\sigma - 1)] \frac{\left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right)}{1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}} \Delta \ln(y_t) &= \sigma(1 - \beta) \Delta \ln(N_{M,t}) - \left[\sigma(1 - \beta) + \frac{\alpha\sigma(1 - \beta)}{1 - \alpha - \beta} \right] \Delta \ln(N_{S,t}) \\
 &\quad - \alpha\sigma \Delta \ln(L_t) + \alpha \left[\sigma - 1 - \left(\frac{1}{1 - \alpha - \beta} \right) \right] \frac{\Gamma y_t^{\frac{\sigma-1}{\sigma}} \Delta \ln(y_t)}{\left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}} \right)} \\
 \Rightarrow \frac{1 + \alpha(\sigma - 1) + [1 + \alpha(\sigma - 1) - \alpha(\sigma - 1) + \frac{\alpha}{1 - \alpha - \beta}] \Gamma y_t^{\frac{\sigma-1}{\sigma}}}{1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}} \Delta \ln(y_t) &= \sigma(1 - \beta) \Delta \ln(N_{M,t}) - \sigma(1 - \beta) \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \Delta \ln(N_{S,t}) - \alpha\sigma \Delta \ln(L_t) \\
 \Rightarrow \frac{1 + \alpha(\sigma - 1) + \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \Gamma y_t^{\frac{\sigma-1}{\sigma}}}{1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}} \Delta \ln(y_t) &= \sigma(1 - \beta) \Delta \ln(N_{M,t}) - \sigma \frac{(1 - \beta)^2}{1 - \alpha - \beta} \Delta \ln(N_{S,t}) - \alpha\sigma \Delta \ln(L_t)
 \end{aligned} \tag{A7}$$

Steps from (A7) to (A8)

To progress from (A7), we need to replace $\Delta \ln(N_{S,t})$, using this:

$$\begin{aligned}
 n_t = \frac{\Delta N_{M,t} / N_{M,t}}{\Delta N_{S,t} / N_{S,t}} \Rightarrow n_t \Delta \ln(N_{S,t}) &= \Delta \ln(N_{M,t}) = \Delta \ln(N_t N_{S,t}) \\
 &= \Delta \ln(N_t) + \Delta \ln(N_{S,t}) \\
 \Rightarrow \Delta \ln(N_{S,t}) = \frac{\Delta \ln(N_t)}{n_t - 1} \Rightarrow \Delta \ln(N_{M,t}) &= n_t \Delta \ln(N_{S,t}) = \frac{n_t \Delta \ln(N_t)}{n_t - 1}
 \end{aligned}$$

So (A7) becomes:

$$\begin{aligned} & \frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta}\right)\Gamma y_t^{\frac{\sigma-1}{\sigma}}}{1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}} \Delta \ln(y_t) \\ &= \sigma(1 - \beta) \left(\frac{n_t \Delta \ln(N_t)}{n_t - 1} \right) - \sigma \frac{(1 - \beta)^2}{1 - \alpha - \beta} \left(\frac{\Delta \ln(N_t)}{n_t - 1} \right) - \alpha \sigma \Delta \ln(L_t) \\ &= \frac{\sigma(1 - \beta)}{n_t - 1} \left(n_t - \frac{1 - \beta}{1 - \alpha - \beta} \right) \Delta \ln(N_t) - \alpha \sigma \Delta \ln(L_t) \end{aligned}$$

Multiplying both sides by $n_t - 1$:

$$\begin{aligned} & \Rightarrow \frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta}\right)\Gamma y_t^{\frac{\sigma-1}{\sigma}}}{1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}} (n_t - 1) \Delta \ln(y_t) \\ &= \sigma(1 - \beta) \left(n_t - \frac{1 - \beta}{1 - \alpha - \beta} \right) \Delta \ln(N_t) - (n_t - 1) \alpha \sigma \Delta \ln(L_t) \end{aligned} \quad (\text{A8})$$

Derivations in Appendix 7 on Existence of Pre-industrial Stagnation

For proof of Lemma part (ii)

To show n falls as y rises on the $\Delta n_t = 0$ locus (A10), write (A10) as

$$n(y, N) = \frac{A + B\Gamma y^{\frac{\sigma-1}{\sigma}}}{C + D\Gamma y^{\frac{\sigma-1}{\sigma}}} \equiv f(\Gamma y^{\frac{\sigma-1}{\sigma}})$$

where $A \equiv (\sigma - 1)(1 - \alpha - \beta) - 1$, $B \equiv (\sigma - 1)(1 - \beta) - \left(\frac{1-\beta}{1-\alpha-\beta}\right)$, $C \equiv (\sigma - 1)(1 - 2\alpha - \beta) - 1$ and $D \equiv (\sigma - 1)(1 - \alpha - \beta) - \left(\frac{1-\beta}{1-\alpha-\beta}\right)$. We need to prove that

$$f'(\Gamma y^{\frac{\sigma-1}{\sigma}}) = \frac{(C + D\Gamma y^{\frac{\sigma-1}{\sigma}})B - (A + B\Gamma y^{\frac{\sigma-1}{\sigma}})D}{(C + D\Gamma y^{\frac{\sigma-1}{\sigma}})^2} = \frac{BC - AD}{(C + D\Gamma y^{\frac{\sigma-1}{\sigma}})^2} < 0$$

This is true because we have $\sigma > \sigma^\dagger > \tilde{\sigma} = \frac{1}{1-\alpha-\beta} + 1$, and hence

$$BC - DA$$

$$\begin{aligned} &= \left[(\sigma - 1)(1 - \beta) - \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \right] [(\sigma - 1)(1 - 2\alpha - \beta) - 1] \\ &\quad - \left[(\sigma - 1)(1 - \alpha - \beta) - \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \right] [(\sigma - 1)(1 - \alpha - \beta) - 1] \\ &= -\alpha \left[(\sigma - 1)(1 - \beta) - \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \right] + \alpha [(\sigma - 1)(1 - \alpha - \beta) - 1] \\ &= \alpha \left[-(\sigma - 1)(1 - \beta) + \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) + (\sigma - 1)(1 - \beta) - \alpha(\sigma - 1) - 1 \right] \\ &= \alpha \left[\frac{\alpha}{1 - \alpha - \beta} - \alpha(\sigma - 1) \right] = \alpha^2(\tilde{\sigma} - \sigma) < 0 \end{aligned}$$

For proof of Proposition 2(b)

Showing

$$\frac{\left[1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta}\right)\Gamma y_t^{\frac{\sigma-1}{\sigma}}\right](\bar{n} - 1)}{\left(1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}}\right)\left(\bar{n} - \frac{1-\beta}{1-\alpha-\beta}\right)} > (\sigma - 1)(1 - \alpha - \beta) \quad (\text{A13})$$

is the same as showing:

$$\begin{aligned} & \left[\left(\frac{1-\beta}{1-\alpha-\beta} \right) (\bar{n} - 1) - \left(\bar{n} - \frac{1-\beta}{1-\alpha-\beta} \right) (\sigma - 1)(1 - \alpha - \beta) \right] \Gamma y_t^{\frac{\sigma-1}{\sigma}} \\ & > \left(\bar{n} - \frac{1-\beta}{1-\alpha-\beta} \right) (\sigma - 1)(1 - \alpha - \beta) - [1 + \alpha(\sigma - 1)](\bar{n} - 1). \end{aligned}$$

We prove this inequality is true by showing the [LHS] > 0 and the RHS < 0 as follows:

$$\begin{aligned} \frac{[\text{LHS}]}{(1 - \alpha - \beta)} &= \left[\frac{1 - \beta}{(1 - \alpha - \beta)^2} (\bar{n} - 1) - (\sigma - 1) \left(\bar{n} - \frac{1 - \beta}{1 - \alpha - \beta} \right) \right] \\ &= \left(\frac{1 - \beta}{(1 - \alpha - \beta)^2} - (\sigma - 1) \right) \bar{n} - \frac{1 - \beta}{(1 - \alpha - \beta)^2} + (\sigma - 1) \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \end{aligned}$$

which (because $\sigma - 1 < \frac{1-\beta}{(1-\alpha-\beta)^2}$ and $\bar{n} > \frac{1-\beta}{1-\alpha-\beta}$)

$$\begin{aligned} & > \left[\left(\frac{1 - \beta}{(1 - \alpha - \beta)^2} - (\sigma - 1) \right) \frac{1 - \beta}{1 - \alpha - \beta} + \left(\sigma - 1 - \frac{1}{1 - \alpha - \beta} \right) \left(\frac{1 - \beta}{1 - \alpha - \beta} \right) \right] \\ &= \left(\frac{1 - \beta}{(1 - \alpha - \beta)^2} - \frac{1}{1 - \alpha - \beta} \right) \frac{1 - \beta}{1 - \alpha - \beta} > 0. \end{aligned}$$

$$\begin{aligned} \text{RHS} &= \left(\bar{n} - \frac{1 - \beta}{1 - \alpha - \beta} \right) (1 - \alpha - \beta)(\sigma - 1) - [1 + \alpha(\sigma - 1)](\bar{n} - 1) \\ &= (1 - \alpha - \beta)(\sigma - 1)\bar{n} - \bar{n} - \alpha(\sigma - 1)\bar{n} - (1 - \beta)(\sigma - 1) + 1 + \alpha(\sigma - 1) \\ &= [(1 - \alpha - \beta)(\sigma - 1) - \alpha(\sigma - 1) - 1]\bar{n} - (1 - \alpha - \beta)(\sigma - 1) + 1 \end{aligned}$$

which (again because $\bar{n} > \frac{1-\beta}{1-\alpha-\beta}$)

$$\begin{aligned} & < [(1 - \alpha - \beta)(\sigma - 1) - \alpha(\sigma - 1) - 1] \frac{1 - \beta}{1 - \alpha - \beta} - (1 - \alpha - \beta)(\sigma - 1) + 1 \\ &= \left[1 - \beta - \alpha \frac{1 - \beta}{1 - \alpha - \beta} - (1 - \alpha - \beta) \right] (\sigma - 1) - \frac{1 - \beta}{1 - \alpha - \beta} + 1 \\ &= - \left(\frac{1 - \beta}{1 - \alpha - \beta} - 1 \right) [\alpha(\sigma - 1) + 1] < 0. \end{aligned}$$

Derivation in Appendix 8 on Existence of Modern Economic Growth Zone

$$\begin{aligned} & \sigma \left[\frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta}\right)N_t^{\frac{1}{\nu}}n_t^{\frac{1-\nu}{\nu}}}{1 + N_t^{\frac{1}{\nu}}n_t^{\frac{1-\nu}{\nu}}} \right] (n_t - 1)[(1 - \beta)\Delta \ln(N_t) - \alpha\Delta \ln(e_t)] \\ &= \sigma \left[(1 - \beta) \left(n_t - \frac{1 - \beta}{1 - \alpha - \beta} \right) \Delta \ln(N_t) - (n_t - 1)\alpha\Delta \ln(L_t) \right] \quad (\text{A16}) \\ \Rightarrow (1 - \beta) & \left[\frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta}\right)N_t^{\frac{1}{\nu}}n_t^{\frac{1-\nu}{\nu}}}{1 + N_t^{\frac{1}{\nu}}n_t^{\frac{1-\nu}{\nu}}} (n_t - 1) - \left(n_t - \frac{1 - \beta}{1 - \alpha - \beta} \right) \right] \Delta \ln(N_t) \\ &+ (n_t - 1)\alpha\Delta \ln(L_t) = \frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta}\right)N_t^{\frac{1}{\nu}}n_t^{\frac{1-\nu}{\nu}}}{1 + N_t^{\frac{1}{\nu}}n_t^{\frac{1-\nu}{\nu}}} (n_t - 1)\alpha \Delta \ln(e_t) \end{aligned}$$

$$\begin{aligned}
&\Rightarrow \left(\frac{\frac{1+\alpha(\sigma-1)}{1-\beta} + \frac{N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1-\alpha-\beta}}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right) (n_t - 1) \alpha \Delta \ln(e_t) \\
&= \left[\left(\frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta} \right) N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} - 1 \right) n_t + \frac{1 - \beta}{1 - \alpha - \beta} \right. \\
&\quad \left. - \frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta} \right) N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right] \Delta \ln(N_t) + \frac{(n_t - 1) \alpha}{(1 - \beta)} \Delta \ln(L_t) \\
&= \left[\left(\frac{1 + \alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta} \right) N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}} - 1 - N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right) n_t \right. \\
&\quad \left. + \frac{\left(\frac{1-\beta}{1-\alpha-\beta} \right) \left(1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}} - N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}} \right) - 1 - \alpha(\sigma - 1)}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right] \Delta \ln(N_t) + \frac{(n_t - 1) \alpha}{(1 - \beta)} \Delta \ln(L_t) \\
&= \left[\left(\frac{\alpha(\sigma - 1) + \left(\frac{1-\beta}{1-\alpha-\beta} \right) N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}} - N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right) n_t + \frac{\left(\frac{1-\beta}{1-\alpha-\beta} \right) - 1 - \alpha(\sigma - 1)}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right] \Delta \ln(N_t) \\
&\quad + \frac{(n_t - 1) \alpha}{(1 - \beta)} \Delta \ln(L_t)
\end{aligned}$$

Now substitute $\frac{1-\beta}{1-\alpha-\beta} - 1 = \frac{\alpha}{1-\alpha-\beta} = \alpha(\tilde{\sigma} - 1)$, which makes this expression:

$$\begin{aligned}
&= \left[\left(\frac{\alpha(\sigma - 1) + \alpha(\tilde{\sigma} - 1) N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right) n_t + \frac{\alpha(\tilde{\sigma} - 1) - \alpha(\sigma - 1)}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right] \Delta \ln(N_t) \\
&\quad + \frac{(n_t - 1) \alpha}{(1 - \beta)} \Delta \ln(L_t) \\
&\Rightarrow \left(\frac{\frac{1+\alpha(\sigma-1)}{1-\beta} + \frac{N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1-\alpha-\beta}}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} \right) (n_t - 1) \alpha \Delta \ln(e_t) \\
&= \left[\left\{ \sigma - 1 + (\tilde{\sigma} - 1) N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}} \right\} n_t - (\sigma - \tilde{\sigma}) \right] \frac{\alpha \Delta \ln(N_t)}{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}} + \frac{(n_t - 1) \alpha}{(1 - \beta)} \Delta \ln(L_t) \\
&\Rightarrow \left(\frac{1 + \alpha(\sigma - 1)}{1 - \beta} + \frac{N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 - \alpha - \beta} \right) (n_t - 1) \Delta \ln(e_t) \\
&= \left[\left\{ \sigma - \tilde{\sigma} + (\tilde{\sigma} - 1) (1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}) \right\} n_t - (\sigma - \tilde{\sigma}) \right] \Delta \ln(N_t) \quad (A17) \\
&\quad + \left(\frac{1 + N_t^{\frac{1}{\nu}} n_t^{\frac{1-\nu}{\nu}}}{1 - \beta} \right) (n_t - 1) \Delta \ln(L_t)
\end{aligned}$$

Derivations in Appendix 10 on Asymptotic Growth Rates

Derivation of (A26)

$$(19) \ Y_{M,t} = \frac{1}{\beta} N_{M,t} p_{M,t}^{\frac{\beta}{1-\beta}} \bar{E}_M^{\frac{\alpha}{1-\beta}} L_{M,t}^{\frac{1-\alpha-\beta}{1-\beta}} \Rightarrow \frac{Y_{M,t}}{L_{M,t}} = \frac{1}{\beta} N_{M,t} p_{M,t}^{\frac{\beta}{1-\beta}} \bar{E}_M^{\frac{\alpha}{1-\beta}} L_{M,t}^{\frac{-\alpha}{1-\beta}}$$

which, substituting (6) for $p_{M,t}$ and (14) for $L_{M,t}$,

$$= \frac{1}{\beta} N_{M,t} \left[(1 - \gamma)^{\frac{\sigma-1}{\sigma-1}} \Gamma \left(y_t^{-\frac{\sigma-1}{\sigma}} + \Gamma \right)^{\frac{1}{\sigma-1}} \right]^{\frac{\beta}{1-\beta}} \bar{E}_M^{\frac{\alpha}{1-\beta}} \left(\frac{L_t \Gamma}{y_t^{-\frac{\sigma-1}{\sigma}} + \Gamma} \right)^{\frac{-\alpha}{1-\beta}}$$

$$\begin{aligned}
&= \frac{1}{\beta} [(1-\gamma)^{\frac{\sigma}{\sigma-1}} \Gamma]^{\frac{\beta}{1-\beta}} \bar{E}_M^{\frac{\alpha}{1-\beta}} \Gamma^{\frac{-\alpha}{1-\beta}} N_{M,t} (y_t^{-\frac{\sigma-1}{\sigma}} + \Gamma)^{\frac{1}{\sigma-1}(\frac{\beta}{1-\beta}) + \frac{\alpha}{1-\beta}} L_t^{\frac{-\alpha}{1-\beta}} \\
&= \phi_M N_{M,t} (y_t^{-\frac{\sigma-1}{\sigma}} + \Gamma)^{\frac{(\sigma-1)\alpha+\beta}{(\sigma-1)(1-\beta)}} L_t^{\frac{-\alpha}{1-\beta}} \text{ where } \phi_M \equiv \frac{1}{\beta} (1-\gamma)^{\frac{\sigma}{\sigma-1}(\frac{\beta}{1-\beta})} \bar{E}_M^{\frac{\alpha}{1-\beta}} \Gamma^{\frac{\beta-\alpha}{1-\beta}} > 0 \quad (\text{A26})
\end{aligned}$$

Derivation of (A28)

Substituting $E_{S,t} = \left(\frac{\alpha N_{S,t}}{\beta \bar{e}_S}\right)^{\frac{1-\beta}{1-\alpha-\beta}} L_{S,t} p_{S,t}^{\frac{1}{1-\alpha-\beta}}$ (22) into $Y_{S,t} = \frac{1}{\beta} N_{S,t} p_{S,t}^{\frac{\beta}{1-\beta}} E_{S,t}^{\frac{\alpha}{1-\beta}} L_{S,t}^{\frac{1-\alpha-\beta}{1-\beta}}$ (20)

and rearranging: $\Rightarrow Y_{S,t} = \frac{1}{\beta} N_{S,t} p_{S,t}^{\frac{\beta}{1-\beta}} \left[\left(\frac{N_{S,t}}{\bar{e}_S}\right)^{\frac{1-\beta}{1-\alpha-\beta}} L_{S,t} p_{S,t}^{\frac{1}{1-\alpha-\beta}} \right]^{\frac{\alpha}{1-\beta}} L_{S,t}^{\frac{1-\alpha-\beta}{1-\beta}}$

(Powers: $N_{S,t}: \frac{1-\alpha-\beta}{1-\alpha-\beta} + \frac{1-\beta}{1-\alpha-\beta} \frac{\alpha}{1-\beta} = \frac{1-\beta}{1-\alpha-\beta}$; $p_{S,t}: \frac{\beta(1-\alpha-\beta)+\alpha}{(1-\beta)(1-\alpha-\beta)} = \frac{\beta(1-\beta)+\alpha(1-\beta)}{(1-\beta)(1-\alpha-\beta)} = \frac{\alpha+\beta}{1-\alpha-\beta}$)

$$= \frac{1}{\beta} \left(\frac{1}{\bar{e}_S}\right)^{\frac{\alpha}{1-\alpha-\beta}} N_{S,t}^{\frac{1-\beta}{1-\alpha-\beta}} p_{S,t}^{\frac{\alpha+\beta}{1-\alpha-\beta}} L_{S,t}$$

and then using (7) for $p_{S,t}$

$$\begin{aligned}
&\Rightarrow \frac{Y_{S,t}}{L_{S,t}} = \frac{1}{\beta} \left(\frac{1}{\bar{e}_S}\right)^{\frac{\alpha}{1-\alpha-\beta}} N_{S,t}^{\frac{1-\beta}{1-\alpha-\beta}} \left[(1-\gamma)^{\frac{\sigma}{\sigma-1}} (1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}})^{\frac{1}{\sigma-1}} \right]^{\frac{\alpha+\beta}{1-\alpha-\beta}} \\
&= \phi_S N_{S,t}^{\frac{1-\beta}{1-\alpha-\beta}} (1 + \Gamma y_t^{\frac{\sigma-1}{\sigma}})^{\frac{\alpha+\beta}{(\sigma-1)(1-\alpha-\beta)}} \text{ where } \phi_S \\
&\quad \equiv \frac{1}{\beta} \left(\frac{1}{\bar{e}_S}\right)^{\frac{\alpha}{1-\alpha-\beta}} (1-\gamma)^{\frac{\sigma}{\sigma-1}(\frac{\alpha+\beta}{1-\alpha-\beta})} > 0 \quad (\text{A28})
\end{aligned}$$