

A Theory of Long-run Development with Capital

Nick Pretnar^{12*} Ákos Valentinyi^{3†}

¹*Laboratory for Aggregate Economics and Finance, UCSB*

²*University of Austin Texas (UATX)*

³*University of Manchester*

July 31, 2026[‡]

Abstract

This paper offers a unified framework for understanding long-run growth across the pre-industrial, transitional, and industrial eras. We build a flexible growth model with three factors of production (land, capital, and labor) and two distinct productivities which separately augment land/labor and capital/labor complementarities. The structure generates endogenously variable elasticities of substitution and explains shifts in both factor income shares and aggregate output over an entire time series. Estimated on historical data from England and Wales, the model reconciles multiple growth regimes and provides new insights into the forces that led to industrialization.

*npretnar@ucsb.edu

†valentinyi.a@gmail.com

‡We have benefited from conversations with Arpad Abraham, Pedro Amaral, W.L. Bednar, Adam Blandin, Giancarlo Corsetti, Liuchun Deng, Ying Feng, Paul Jackson, Karen Kopecky, David Lagakos, Oksana Leukhina, Remy Levin, Alexander Ludwig, Kathleen McKiernan, Jeremy Pearce, Thomas Phelan, Franck Portier, B. Ravikumar, José-Víctor Ríos-Rull, Sarah Quincy, Luis Felipe Saenz, James Traina, Guillaume Vandembroucke, Nicola Viegi, Robert Wright, and Donghai Zhang. We also thank two anonymous referees and seminar participants at the University of Barcelona, Vanderbilt University, the Federal Reserve Bank of Saint Louis, the European University Institute, the University of Pretoria, and the National University of Singapore, as well as attendees of the 2026 SED Meetings in Athens, the 2024 MEA Meetings in Chicago, the 2024 SED Winter Meetings in Buenos Aires, the 2024 European Winter Meetings of the Econometric Society in Palma de Mallorca, Spain, and Finn's Skäl at the IE Tower in Madrid, Spain.

1 Introduction

Theories of long-run economic development that include a role for capital in production are decidedly lacking, despite the important role that capital accumulation plays in contributing to aggregate growth. Particularly in the Industrial Revolution literature, only Hansen and Prescott (2002) and Bar and Leukhina (2010a) really think of industrialization as the start of the accumulation of physical, durable capital.¹

This paper extends the exogenous growth model of Solow (1956) and Solow (1957) to a more flexible, single-sector framework capable of reconciling both factor income shares and aggregate output data along the development path. This is because the model, despite its apparently nested CES form, actually generates endogenously variable elasticities of substitution (VES) between inputs. Our key innovation is to allow labor to be augmented by two different kinds of factor productivities — one capturing the efficiency by which labor utilizes land and another capturing the efficiency by which labor utilizes non-landed capital. In this sense we separate land from the capital index, treating it as a standalone production factor in all periods of economic development. This modeling choice yields results that challenge the balanced growth properties of the original Solow (1956) model, namely the independence of GDP per capita from population growth. Yet, as we show, it also can be tractably estimated on time series of input/output data to recover primitive elasticities and each period’s levels of factor-augmenting productivities. Indeed, when applied to input/output data for England + Wales dating from 1200-1860, our model can precisely match factor income shares, while also reconciling the evolution of GDP.

We thus position ourselves squarely within the exogenous growth literature, demonstrating that if such models are built with sufficient parametric flexibility, they can explain long-run growth patterns as summarized both by output per capita and three factor income shares: 1) land income, 2) non-landed capital income, 3) labor income.² Our model yields interpretations regarding how the efficiency of using land versus using non-landed capital changed leading up to early industrialization. Specifically, we can use input/output data to identify, through the model’s lens, when different productivities associated with using land versus non-landed capital change. We actually find that the labor-efficiency of using land started to decline long before the labor-efficiency of non-landed capital utilization rose. This contributed to Malthusian growth stagnation in the centuries preceding industrialization. Improvements to the efficiency of using non-landed capital coincide with early industrialization in England + Wales, a result that runs parallel to theories of industrial development that rely on endogenous human-capital accumulation (Becker, Murphy,

¹In a recent paper Ravikumar and Vandenbroucke (2025) also use a version of the model of Hansen and Prescott (2002) that does not include capital to assess the timing of when industrialization begins across many different countries by exploiting time series data on countries’ agricultural employment share. Their model is similar to Hansen and Prescott (2002) and Bar and Leukhina (2010a) only in so much as land’s relevance for production disappears as the economy enters the Solow growth regime.

²We use the term “non-landed capital” to refer to all capital that is not explicitly land. Non-landed capital includes structures and their rental-equivalent values, but not the rental-equivalent value of the land itself (i.e., the so-called “ground rents”).

and Tamura 1990; Goodfriend and McDermott 1995; Galor and Weil 2000; Jones 2001; Galor and Moav 2002; Tamura 2002).³ This is because in our model human-capital accumulation is mostly embedded within an increase in non-landed capital-utilization efficiency (i.e., structures and machines). Rising capital-utilization efficiency also precedes the reallocation of factor income shares toward capital by 50 to 75 years, allowing the model to replicate the fact that output growth *preceded* input-share reallocation in England + Wales. Further, under a multi-sector re-interpretation of our technology, we show that rising capital-utilization efficiency is not an artifact of labor real-locating out of agriculture, but rather reflects genuine improvements in the efficiency with which workers use capital. Finally, our estimates imply that land, labor, and capital were gross complements in every decade from 1200-1860, with capital and labor becoming more substitutable only as industrialization deepened after 1750.

Our approach to modeling long-run growth can be seen as building on work by Hansen and Prescott (2002), Bar and Leukhina (2010a), and Ravikumar and Vandenbroucke (2025), though our aim is to produce a model that can be inverted to recover an entire time series of observed factor income shares and output without having to exogenously specify or estimate things like an equilibrium transition between two different kinds of growth regimes (i.e., a Malthusian one to a Solow one, as Hansen and Prescott (2002), Bar and Leukhina (2010a), and Ravikumar and Vandenbroucke (2025) do). This is because in our model non-landed capital, land, and labor are all *always* needed in aggregate production. Malthus (1798) argued that pre-industrial economies were characterized by slow (no?) productivity growth, with land being an important fixed factor of production, and rising populations generally leading to declining standards of living. In Solow (1956) and Solow (1957) economies, meanwhile, rising productivity leads to capital accumulation which raises GDP per capita, regardless of how fast or slow the population is growing. Hansen and Prescott (2002), Bar and Leukhina (2010a), and Ravikumar and Vandenbroucke (2025) posit production functions where a Malthus to Solow transition takes place once the productivity associated with using the Solow technology is sufficiently high. In our approach, though, the negative effects of population growth always enter into the growth dynamics, though how much impact they have declines as productivity rises. This flexible approach allows us to replicate and explain all kinds of eras of economic development, not just a Malthus to Solow transition associated with industrialization. For example, we can explain rising GDP per capita accompanying declining population in the Black Death, Malthusian stagnation from 1400-1650, early modern growth in the late 1600s, and full on Industrial Revolution thereafter. While the Hansen and Prescott (2002), Bar and Leukhina (2010a), and Ravikumar and Vandenbroucke (2025) Malthus to Solow technology “fusion” models can explain a crude Malthus to Solow transition, it is not clear that such a framework is flexible enough to explain short- and medium-term oscillations in growth and factor

³Note that many theories of long-run growth and the Industrial Revolution feature some flavor of human capital growth driving the dynamic, economic changes, either via an ‘idea’-accumulation mechanism or mechanisms directly linked to fertility choices and human-capital feedback effects (Becker, Murphy, and Tamura 1990; Goodfriend and McDermott 1995; Galor and Weil 2000; Jones 2001). But, in none of Becker, Murphy, and Tamura (1990), Goodfriend and McDermott (1995), Galor and Weil (2000), and Jones (2001) is physical and *durable* capital featured.

income shares before, during, and after industrialization. This is our innovation.

We proceed by first discussing the kinds of aggregate data series our model is designed to reconcile, focusing on the case of England + Wales from 1200-1860, using data compiled by Clark (2010) and Maddison Project (2020). In Section 2 we present the most generalized form of our aggregate production function. We spend time discussing why it is endogenously VES and characterize its growth properties. We show how factor income shares vary in the growth rates of population and our two different factor-augmenting productivities. We then demonstrate that a classical Solow (1957)-style productivity estimation procedure will be inherently biased downward because it does not explicitly account for land as a fixed factor. After analyzing the model, in Section 3 we estimate the full technology using the Clark (2010) data, backing out productivities and computing elasticities of substitution.

Our findings speak to prominent historical accounts of why growth germinated in seventeenth-century England. The timing and sectoral incidence we measure are consistent with the idea that the institutional transformations of that century — strengthened property rights following the Glorious Revolution (North and Weingast 1989; Acemoglu, Johnson, and Robinson 2005a; Acemoglu, Johnson, and Robinson 2005b), knowledge dissemination through England’s guild and apprenticeship systems (Mokyr 2002), and the emergence of capital-market institutions, such as stock exchanges and insurance markets, out of the commercial sociability of the early coffeehouse era (Jamieson 2001; Withington 2020) — were integral in stimulating early, pre-industrial economic growth. Through the lens of our model, any such stimulus manifests as a sustained rise in the productivity of *using* capital beginning in the mid-to-late 1600s, which *precedes* the emergence of the log-linear GDP per capita growth that characterizes the Industrial Revolution. This timing, which we will discuss in Section 3, does not result from ad hoc assumptions regarding when the Industrial Revolution began: it is estimated from the model structure and input/output data alone. While aggregate decadal data cannot identify which of these candidate stimuli initiated the take-off, our estimates discipline what any candidate explanation must quantitatively deliver, a point we return to in the Conclusion.

1.1 Macroeconomic Development in England + Wales

We will use our theoretical framework to analyze the case of long-run development in the British Isles. Figure 1 presents real GDP per capita in \$2011 for the United Kingdom. The Black Death era from the mid 1300s to the mid 1400s is associated with a rise in GDP per capita, followed by a prolonged period of stagnation from 1400-1650. From 1600-1800 real GDP per capita appeared to then increase linearly before the kind of exponential growth, characteristic of Solow economies, took off after 1800.⁴ Indeed, Figure 1a shows that between 1600 and 1800, output per capita more than doubled.

⁴Note that similar patterns can be gleaned from GDP data for England + Wales only, featured in Clark (2010); this will later be seen in Figure 3. We use Clark (2010) data for structural model estimation, but it is at the decade frequency. Data from Maddison Project (2020) is annual, so we present it here to more vividly illustrate how growth patterns vary over time.

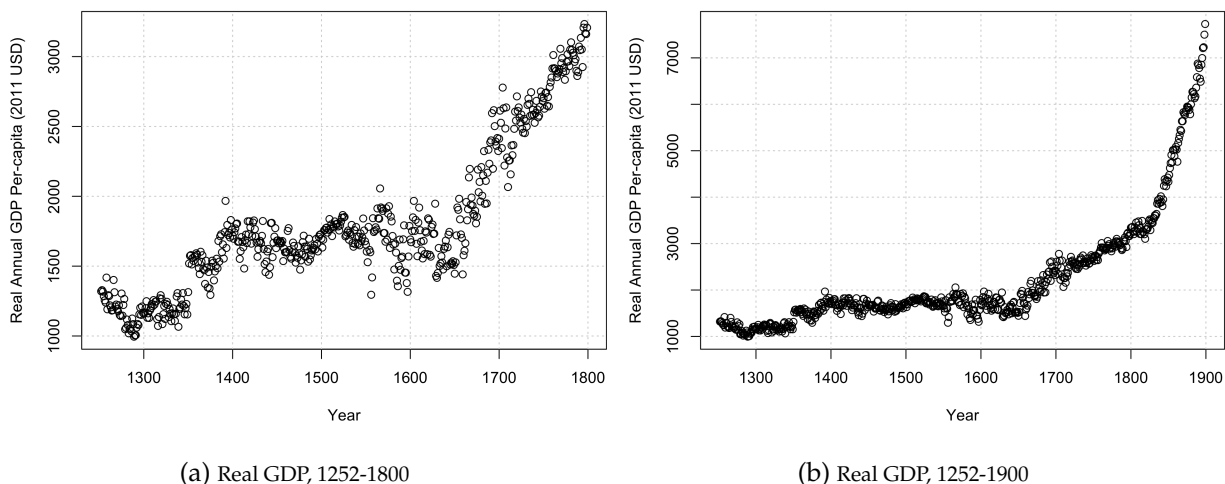


Figure 1: This figure presents annual GDP in real \$2011 for the United Kingdom (Northern Island and Scotland, inclusive) starting in 1252 taken from the Maddison Project (2020). Panel (a) presents data up to 1800, while panel (b) presents data up to 1900.

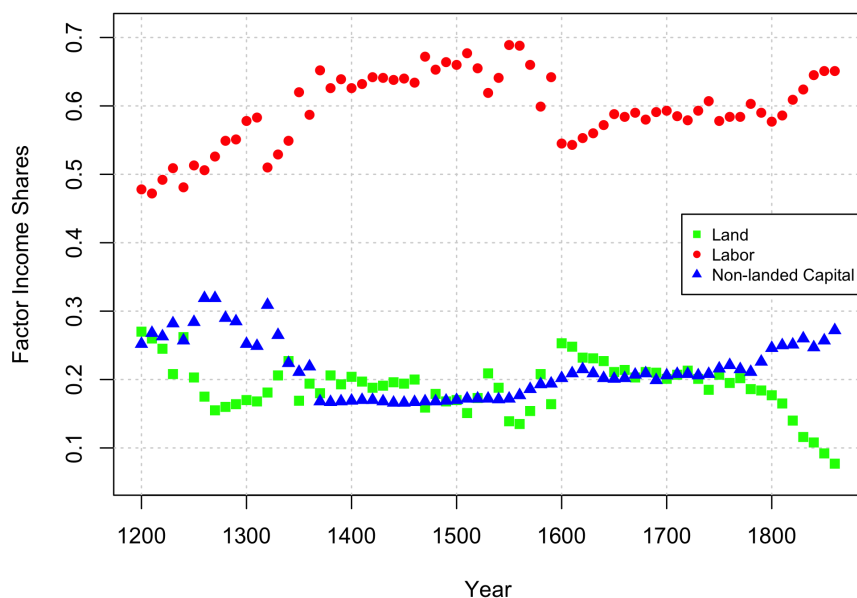


Figure 2: In this figure we present the factor income shares from Clark (2010) for England and Wales. The data start in 1200 with observations by decade up to 1860. Table 13 in Clark (2010) separates out land rents (variable “Land Share in National Income”) from rents associated with other capital (variable “Capital Share in National Income”). The labor share is also included.

Figure 2 presents the evolution of land, labor, and non-landed capital income shares from Clark (2010). Notice that over the entire 1200-1860 period labor oscillates over long cycles between encompassing 50-70% of national income. The mid-1300s Black Death period is preceded

by a gradual rise in the labor share and decline in the non-landed capital share. The period from 1400-1750 (approximately) then sees land and non-landed capital equally at about 20% of national income, while labor encompasses about 60% of national income. The Industrial Revolution period in the mid-1700s is characterized *not* by a decline in labor’s share of income but rather by both the labor and non-landed capital shares of income rising as the land share falls. Still, though, note that even in 1860 — well into a Solow-style growth regime — land’s share of income was 7.7%.⁵ Our structural framework will always allow production to depend on land, even in settings of advanced growth, though land’s relative importance for production may gradually decline (but not disappear) as a result of technological change.

Figure 2 demonstrates that splitting land from capital is essential, as it is clear that variable elasticities of substitution between land and labor versus capital and labor are required for a model to reconcile the full range of major economic transitions.⁶ The transition associated with the Black Death exhibits markedly different patterns in input allocation compared to the transition of the 1700s typically associated with the Industrial Revolution. Further, notice that when comparing Figures 1 and 2 it is clear that output started to grow in the 1600s and 1700s *before* factor-share reallocation occurred. The model should have sufficient flexibility to capture this lagged relationship as well.

While the data in Clark (2010) conveniently allow the econometrician to distinguish between land and non-landed capital, they do not so conveniently allow for inputs to be classified according to their productive use. Such a classification would be necessary for a *meaningful* multi-sector model that allows for all three inputs to be used toward the production of different kinds of outputs. For example, the models of Hansen and Prescott (2002), Bar and Leukhina (2010a), and Ravikumar and Vandenbroucke (2025) require inputs to be classified as to whether or not they are used in service of agricultural versus non-agricultural output;⁷ the Clark (2010) data can only be used to distinguish between labor used in service of agricultural versus non-agricultural production but not land or non-landed capital. Thus, a flexible single-sector model with aggregate land, non-landed capital, and labor as inputs provides for clean identification of technology parameters

⁵Even in modern economies land’s income share remains significantly non-zero. In the United States the Bureau of Economic Analysis reports that net “space rent” (the land component of housing rents) was about 3% of personal income in 2013 (Katz 2017). Using a cross-country accounting approach, Bakker, Niesert, and Seiferling (2023) estimate land rents as a share of GDP for Great Britain at roughly 21% in recent decades.

⁶This is consistent with findings in Juselius (2008), Raurich, Sala, and Sorolla (2012), Karabarbounis and Neiman (2013), and Knoblach, Roessler, and Zwerschke (2019), who all demonstrate that VES is needed to match contemporary variation in factor income shares, though they each take different approaches to show this. For example, in a multi-sector model Karabarbounis and Neiman (2013) show how aggregate factor shares fluctuate as a result of changing sectoral composition, even if each sector is CES but with different production elasticities. Our approach will be of the single-sector variety for several reasons we outline below; thus, our structural efforts should be seen as a way to build a tractable model with sufficient degrees of freedom to permit estimation while *still* generating the variable aggregate elasticities necessary to get factor share fluctuations.

⁷While the models in these papers are not *totally* two-sector models since they feature only one output, that output is produced by the sum of two different technologies. Note that modeling aggregate output in this manner is necessary when sectoral price data is unavailable, while a full multi-sector model in the structural change tradition (see, e.g., Ngai and Pissarides (2007) and Herrendorf et al. (2015)) would leverage variation in sectoral relative prices to understand the drivers of sectoral reallocation.

given data. Indeed, any more granular investigation of long-run growth dynamics with consideration given to different sectors' roles would require multi-sector input/output data that is not widely available prior to World War II.⁸ Building a one-sector quantitative model of industrialization has the advantage that it only requires aggregate data which is much more readily available than sectoral data.⁹

We now move to present such a flexible, one-sector technology that can fit both input and output data for the long-run development profile of England + Wales. For our case study it will need to explain both rising output during the Black Death and factor share reallocation away from non-landed capital toward labor. It will also need to allow for output to grow leading up to the Industrial Revolution *prior* to factor share reallocation. We will quantify the model in Section 3 and assess if, indeed, it can successfully accomplish all of these things. It turns out that, yes, it can.

2 The Technology

In this section we present a flexible, three-factor production technology that is designed to be able to fit shares of income from land, capital, and labor in all epochs of economic development. We can accomplish this with a specific form of a nested technology that appears to be constant elasticity of substitution (CES) but actually generates variable elasticities of substitution (VES) for land/labor, capital/labor, and land/capital tradeoffs in all but knife-edge cases on the technology's primitive, structural parameters. In this manner our production function is constant returns to scale (CRS) with VES in disguise, where it is VES in two different conceptual ways. First, elasticities of substitution may vary over time but *not* because we specify time-varying elasticity *parameters*, but rather as we show, the elasticities of substitution depend on primitive structural elasticities, the *levels* of land and capital per capita, and the technology *levels* which are unobserved but vary over time just like structural residuals a la Solow (1957). Second, the three elasticities of substitution across the three factor-input ratios — land/labor, capital/labor, and land/capital — are also different *within* a period.¹⁰

⁸Sector level input/output data that includes the value of the capital stock is hard to come by even for the post World War II period for the specific set of developing countries which would be suitable for a quantitative study of industrialization. Even for today's *developed* countries, data on the sector-level capital stock is only readily available for the United States (Dennis and Işcan 2009; U.S. Bureau of Economic Analysis 2025). For other modern countries the only multi-sector models one can build are labor-only ones such as in Alvarez-Cuadrado and Poschke (2011), Yang and Zhu (2013), and Huneus and Rogerson (2024).

⁹Still, to assuage concerns that our model does not adequately (theoretically) explain industrialization resulting from an agriculture to manufacturing labor transition, we will show that under a coarse assumption, wherein land utilization is *only* associated with agriculture and capital utilization *only* associated with non-agricultural production, we can make high-level inferences regarding the degree to which productivity variation versus labor reallocation led to industrialization. The formal one-sector/two-sector equivalence underlying this exercise is developed in Technical Appendix D and referenced throughout the main text. We stress that the split is coarse and limited by data availability (cannot split capital and/or land between sectors), since in reality non-landed capital is important in agricultural production (Caunedo and Keller 2021) and land is used in many non-agricultural activities.

¹⁰For illustration, the standard CES function $f(x) = (\sum_j \theta_j x_j^\rho)^{1/\rho}$ has constant elasticities both over time and across input ratios when ρ is fixed. If ρ is time-varying then elasticities vary over time but *not* across input ratios. To vary both across input ratios *and* over time requires a richer structure, which we specify here.

Unlike in the setup of Hansen and Prescott (2002), our technology allows for all of land, capital, and labor to *always* be explicitly featured in positive supply in the production technology. Separating the fixed factor (land) from capital ensures that the classic tradeoff between population growth (weighing negatively) and productivity growth (weighing positively) on aggregate output growth is always present in all epochs of economic development.¹¹ Heretofore, the macro-development literature lacks specification of an aggregate technology with all of land, capital, and labor and which can yield any of Malthusian and Solow output dynamics at any time.^{12,13} Traditionally, in the Malthusian economy output per capita falls as the population rises due to the presence of the fixed factor, while under a Solow technology this value is constant in population absent a change in the total factor productivity (i.e., the technological level of production). This explicitly arises *because* the Solow model abstracts from the presence of the fixed factor, simply integrating it into the capital stock. As a result, introducing a fixed factor into the Solow model alters the model's growth properties, challenging some of the empirical regularities identified in the Kaldor (1961) facts.¹⁴

The aggregate technology takes fixed land $L_t = \bar{L}$, durable non-landed capital K_t , and labor N_t (i.e., population) to produce output Y_t according to the nested-CES structure:¹⁵

$$Y_t(L_t, K_t, N_t) = \left(\alpha \left(\omega_L L_t^\gamma + (1 - \omega_L)(z_t^L N_t)^\gamma \right)^{\frac{\rho}{\gamma}} + (1 - \alpha) \left(\omega_K K_t^\kappa + (1 - \omega_K)(z_t^K N_t)^\kappa \right)^{\frac{\rho}{\kappa}} \right)^{\frac{1}{\rho}}. \quad (1)$$

As $\rho \rightarrow 0$, we recover a Cobb-Douglas aggregator:

$$Y_t(L_t, K_t, N_t) = \left(\omega_L L_t^\gamma + (1 - \omega_L)(z_t^L N_t)^\gamma \right)^{\frac{\alpha}{\gamma}} \left(\omega_K K_t^\kappa + (1 - \omega_K)(z_t^K N_t)^\kappa \right)^{\frac{1-\alpha}{\kappa}}. \quad (2)$$

We define the land/labor and capital/labor composites (nests) by $\mathcal{L}$ and $\mathcal{K}$ (script), respectively:

$$\mathcal{L}_t(L_t, N_t) = \left(\omega_L L_t^\gamma + (1 - \omega_L)(z_t^L N_t)^\gamma \right)^{\frac{1}{\gamma}}, \quad (3)$$

$$\mathcal{K}_t(K_t, N_t) = \left(\omega_K K_t^\kappa + (1 - \omega_K)(z_t^K N_t)^\kappa \right)^{\frac{1}{\kappa}}. \quad (4)$$

Note that each composite takes in the *total* population, so workers are not split up according to

¹¹See Jones and Vollrath (2013) for a discussion as to how the presence of the fixed factor leads to the population/productivity horse race.

¹²Even in Hansen and Prescott (2002), the land factor will contribute nothing to production once the Malthus to Solow transition is complete.

¹³We use the term “Malthusian dynamics” to refer to eras where the effects of population growth weighing negatively on output dominate, while we use the term “Solow dynamics” to refer to eras where the effects of productivity growth contributing positively to output dominate.

¹⁴This is not problematic, since Kaldor (1961) regularities are only really valid for economies at already advanced stages of development and, even then, often only over the very long run (Duffy and Papageorgiou 2000).

¹⁵We care about reconciling patterns of long-run growth, so we take an admittedly more crude approach to the labor input than what one might use if examining higher frequency business cycles in a modern economy, for example. Nothing precludes replacing the population level with total hours as the labor input: the analyses are the same, though the units change.

their sector of employment. This is *not* a multi-sector technology, but instead it is a parametric extension of the traditional one-sector growth model with sufficient flexibility to fit input and output data in all stages of development. In this sense our efforts stand in recognition of the spirit of the criticism in Gollin and Kaboski (2023) and Gollin and Rogerson (2025), who advocate for deeper extensions of the one-sector growth model to provide a more granular understanding of the mechanisms underlying growth phenomena.

In our setup labor spent deployed in service to land as an input versus non-landed capital as an input is neither fully separable nor perfectly substitutable.¹⁶ Indeed, combinations of the production parameters and labor-augmenting productivities will dictate the degree to which the land or capital nest dominates in production. This technology thus should not be seen as a fusion of an agricultural technology which has land as an explicit input and an industrial technology which does not. Rather, the mapping from inputs (land, capital, and labor) to output is agnostic as to the composition of those outputs (e.g., agricultural outputs versus physical goods versus intangible services). This is hardly different from the standards in this literature: as mentioned in footnote 12 of Hansen and Prescott (2002), even their model is not one of structural change (i.e., non-agricultural outputs becoming more important than agricultural ones). Rather, the fusion of the two technologies in Hansen and Prescott (2002) allows for aggregate growth dynamics to take on different flavors depending on which of the two technologies dominates (Malthus versus Solow). Our model is similar in spirit but generally more flexible, since (as we will show) it can reconcile both input and output patterns over the many different eras of economic development in England + Wales.¹⁷

One difficulty working with multi-sector models in a structural change kind of approach is that to quantify variations in productivities, one must possess relative price data across sectors or at minimum be able to discuss how relative prices would vary along a transition path. Hansen and Prescott (2002) avoid this problem by interpreting aggregate production as the sum of two technologies (Malthus and Solow). Our approach is the same in spirit though more empirically tractable as a result of its flexibility. Still, one *could* naively interpret the inner nests $\mathcal{L}_t$ and $\mathcal{K}_t$ as a land-intensive sector and a capital-intensive sector, and one could crudely associate $\mathcal{L}_t$ with agricultural production and $\mathcal{K}_t$ with non-agricultural production if one believed that agricultural production could be characterized by a production process lacking capital and vice-versa for all non-agricultural production. This would be wrong though: while agriculture is relatively capital un-intensive prior to industrialization, increasing agricultural capital intensity is associated with the industrialization process (Caunedo and Keller 2021). Nonetheless, in Technical Appendix D

¹⁶This is a very defensible assumption, and in fact, we contend that it is *more* defensible than more coarse assumptions that assume agricultural production *only* uses land and everything else *only* uses non-landed capital. Our defense is that in such models like Hansen and Prescott (2002), land is irrelevant for industrial activity, which surely is unrealistic given factories, mines, etc. sit on (or in the case of mines, “inside of”) land. Mining is a clear example of industrial activity that requires land (where the minerals reside) and capital.

¹⁷Statistically speaking, the nested nature of our technology is useful to introduce enough parametric degrees of freedom to reconcile how factor income shares have changed over time, which will become apparent during quantification in Section 3.

we explore a version of the model where $\mathcal{L}_t$ and $\mathcal{K}_t$ are interpreted as separate sectors and labor is split across them accordingly. We discuss the implications of such a re-interpretation there, though we will continue to refer to various results from that re-interpretive exercise throughout this draft.

Via our preferred single-sector lens, our model is not a standard nested-CES production function, since the *total* supply of one of the inputs (labor) is contained in two separate nests, as long as the weighting parameters, α , ω_L , and ω_K , are all interior to the unit interval. When $\alpha \rightarrow 0$ and/or $\omega_L \rightarrow 0$, land does not matter for production and we recover a Solow economy. A traditional nested-CES where all inputs are restricted to one sub-nest is recovered when either $\omega_L \rightarrow 1$ or $\omega_K \rightarrow 1$. In such a case the model will reduce to something like that in Karabarbounis and Neiman (2013), where elasticities are constant but the factor-augmenting productivity will still allow for income shares to vary over time. Ensuring the weighting parameters are interior while allowing for heterogeneity in the primitive elasticities across nests (i.e., ν , κ , and ρ are not all equal), is what allows the model to generate variable elasticities of substitution across land/labor, capital/labor, and land/capital both within a period and over time. For example, for any given t the land/labor elasticity of substitution is different from the capital/labor elasticity of substitution. Further, for two periods t and t' such that $t \neq t'$, the elasticity of substitution for land/labor (for example) will be different across such periods as long as the level of the labor input (population) and/or the productivity levels are different (i.e., $z_t^L \neq z_{t'}^L$ and/or $z_t^K \neq z_{t'}^K$ for $t \neq t'$). We discuss all of this in detail in Section 2.1 below.

Note that we seek to estimate the model period by period (more specifically, decade by decade) using factor income share data, not just targeting summary statistics. With three factor shares all adding to one, exactly two time-varying parameters are needed to capture idiosyncratic variation in the income shares — hence, z_t^L and z_t^K . Via this setup our specification is both flexible *and* econometrically tractable. Further, we choose to specify these objects as labor-augmenting productivities for interpretation purposes: they encode efficiency with regards to *using* land versus *using* capital, and their variation over time reflects variation in everything from folk knowledge regarding farming practices to capital-adjacent human-capital development.¹⁸

The economy is characterized by a single representative firm, taking the input price of land r_t^L , the input price of non-landed capital r_t^K , and the average wage w_t as given to solve^{19,20}

$$\max_{L_t, N_t, K_t} \left\{ Y_t(L_t, K_t, N_t) - r_t^L L_t - r_t^K K_t - w_t N_t \right\}, \quad \forall t. \quad (5)$$

¹⁸Of course, one could parameterize the model with land- and capital-augmenting productivities or really any combination of two productivities, as long as each inner nest contains one.

¹⁹In practice the representative firm is often constrained to fix $L_t = \bar{L}$, though there are historical exceptions, such as in the case of war, conquest, and annexation. Outside of conflict, though, L_t does sometimes change when lakes and bays are filled in or erosion (e.g., climate change) washes beaches and barrier islands away.

²⁰Hansen and Prescott (2002), Bar and Leukhina (2010a), and Ravikumar and Vandenbroucke (2025) also invoke a representative firm that produces one output. We depart from Hansen and Prescott (2002), Bar and Leukhina (2010a), and Ravikumar and Vandenbroucke (2025) who split the actual *technology* that produces Y_t into the sum of agricultural (or Malthusian) and non-agricultural (or Solow) technologies. We avoid such a split because data limitations in Clark (2010) would keep us from also splitting K_t and L_t into capital and land used in service of agricultural versus non-agricultural production, which is why we seek a flexible, single technology that uses land, capital, and labor aggregates as inputs.

It is assumed that the output price is normalized to unity in each period. Finally, we assume that all resources are fully utilized so that in every period the value of inputs equals outputs: $Y_t = r_t^L L_t + r_t^K K_t + w_t N_t$.

Each composite, (3) and (4), varies over time with labor-augmenting technologies, z_t^L and z_t^K . We call the technologically augmented factors $z_t^L N_t$ and $z_t^K N_t$ “land-efficiency labor” and “capital-efficiency labor,” respectively. As z_t^K rises relative to z_t^L , workers become relatively more efficient at utilizing non-landed capital compared to land. Growth dominated by the transition to capital-intensive industrialization and manufacturing, for example, would be characterized by an increase in the ratio z_t^K/z_t^L (decrease in the inverse). By contrast growth associated simply with improvements to labor’s ability to utilize productive land would be associated with a decline in the ratio z_t^K/z_t^L (increase in the inverse).

So how, exactly, do we interpret z_t^L and z_t^K ? Let us start with z_t^L : one should *not* interpret this object as encoding the *entire* efficiency of agricultural production (i.e., production where land, as an input factor, is relatively important). This is because agricultural production, just like all other production, actually requires utilization of all three inputs — land, labor, and non-landed capital. Using a plow (non-landed capital) improves the efficiency of farming, as do tractors, silos, automated feeding machines, etc., and even more primitive tools like hoes, shovels, and scythes: the value of all of these objects is included in K_t . Rather, z_t^L represents how efficient workers are at utilizing land for output, independent of capital utilization. This value thus encodes knowledge regarding *how* to use the land — when is the best season to plant onions, tomatoes, etc.; when is the best time to harvest, etc. This object can be thought of as thus also capturing “folk knowledge” that is to some extent independent of the complexity of the capital stock. By contrast, augmenting farming practices by introducing new tools and machines is captured by an endogenous increase in K_t . Meanwhile, if workers must gain skills or knowledge with respect to how to use new machines, this would be manifested by a rise in z_t^K , not z_t^L .

Our contention is that industrialization results from an increase in the absolute efficiency of utilizing capital. A rise in z_t^K would primarily embed the kinds of human-capital and idea-generation stories featured in Romer (1986), Lucas Jr (1988), Stokey (1988), Romer (1990), Grossman and Helpman (1991), Stokey (1991), Tamura (1991), Goodfriend and McDermott (1995), Jones (1995), Eaton and Kortum (1996), Tamura (1998), Galor and Weil (2000), Jones (2001), Galor and Moav (2002), Jones (2002), and Tamura (2002) under the assumption that knowledge is generally more capital adjacent than land adjacent. High growth in the non-landed capital stock is associated with high growth in machine adoption, which would require a relatively higher worker skill base to operate. Further, z_t^K encodes improvements in capital-utilization efficiency (i.e., mechanical quality and complexity), which means that fewer workers per machine are required as z_t^K rises. Think, again, of agricultural production before and after the invention of the tractor. Adoption of tractors represents an increase in the quantity of K_t (as would increased adoption of all kinds of other, non-agricultural machines). But, a farmer who was tilling his field using a plow hitched to a team of oxen has to *learn* how to drive and maintain a tractor — an increase in capital-utilization

knowledge, z_t^K .

Rounding out our discussion of the technology's ingredients, note that the parameters ν and κ govern the elasticity of substitution between land and land-efficiency labor and capital and capital-efficiency labor respectively, while ρ governs the elasticity of substitution between the land and capital *nests*. The elasticities of substitution between land, *total* labor, and capital are *not* generally constant. That is, ν , for example, does *not* fully encode the elasticity of substitution between land and labor, because total labor is used in both the land- and capital-efficiency composites; in Section 2.1 we derive the model's elasticities of substitution for inputs L_t , N_t , and K_t and show this. Generally, such elasticities are functions of structural parameters and are themselves variable in both ratios and levels. Indeed, the elasticities of substitution for land/labor, capital/labor, and land/capital will depend on ν , κ , and ρ , as well as ω_L and ω_K (i.e., the land and capital weights in their respective production composites) and α (i.e., the average contribution of the land composite to aggregate production), outside of knife-edge cases.

Finally, note that (1) is homogeneous of degree one in factor inputs. We can rewrite output as a two-factor technology, with output per person as a function of land per person, $\ell_t = L_t/N_t$, and non-landed capital per person, $k_t = K_t/N_t$:

$$y_t(\ell_t, k_t) = \left(\alpha \left(\omega_L \ell_t^\nu + (1 - \omega_L)(z_t^L)^\nu \right)^{\frac{\rho}{\nu}} + (1 - \alpha) \left(\omega_K k_t^\kappa + (1 - \omega_K)(z_t^K)^\kappa \right)^{\frac{\rho}{\kappa}} \right)^{\frac{1}{\rho}}. \quad (6)$$

Lower-case letters will denote per capita variables.

2.1 Elasticities of Substitution

We previously mentioned that, despite appearing to be of the nested-CES form, outside of specific knife-edge cases our model is *endogenously* VES both across different inputs within a period and over time. That is, elasticities of substitution depend on the representative firm's chosen level of inputs and the technology levels. A VES structure is generally necessary if we want the model to be able to generate factor income shares that vary appropriately along the development path according to findings in (Juselius 2008; Raurich, Sala, and Sorolla 2012; Knoblach, Roessler, and Zwerschke 2019). Because we have three factor income shares to match, we desire a more flexible function that can both reconcile the patterns in Figure 2 and the growth of GDP per-capita seen in Figure 1. With our function, taken to data, in Section 3 we empirically assess the degree to which the elasticities of substitution actually *do* vary under the fitted model.

Different modeling strategies are used in the literature to achieve time-varying factor income shares. To get time-varying labor-income shares, Karabarbounis and Neiman (2013) posit a nested-CES structure with intermediate producers whose products are combined into final consumption according to a technology that features a time-varying parameter which simultaneously encodes the intermediate producers' markup over marginal cost and the elasticity of substitution for varieties of inputs. By further assuming intermediate inputs are produced using a CES technology combining capital and labor, this strategy buys them the flexibility to fit time-varying labor shares

both when the intermediate production technology is Cobb-Douglas and when it is a generalized CES: when it is Cobb-Douglas they *need* the time-varying markup/elasticity parameter to match a time-varying labor share, but when it is a general CES, the labor share also depends on the technology level which varies over time. For two-factor Cobb-Douglas technologies a simpler strategy is to let the input/output elasticity, commonly referred to as α_t , be time-varying (Zhang et al. 2022). For a survey of alternative production function structures that yield both CES and VES in different forms, see Growiec and Mućk (2020).

To see how and why our model generates VES between inputs, recall the Hicks (1932) definition of the elasticity of substitution between two arbitrary variables x and y with input prices p^x and p^y and marginal products MP_x and MP_y : $\mathcal{E}^{x,y} = -\frac{d \ln(x/y)}{d \ln(p^x/p^y)} \equiv -\frac{d \ln(x/y)}{d \ln(MP_x/MP_y)}$. Looking at (1), it is clear that the elasticity of substitution between L_t and $z_t^L N_t$ (i.e., land and land-efficiency labor) is constant: $\mathcal{E}^{L,z^L N} = 1/(1 - \nu)$. The same is true for capital and capital-efficiency labor, whose elasticity is $\mathcal{E}^{K,z^K N} = 1/(1 - \kappa)$. But the elasticities of substitution for land/labor, capital/labor, and land/capital depend separately on the marginal products of land, capital, and labor, which are each governed by different primitive parameters.

Consider the first-order conditions of the representative firm from (5):

$$\alpha \omega_L Y_t^{1-\rho} \mathcal{L}_t^{\rho-\nu} L_t^{\nu-1} = r_t^L, \quad (7)$$

$$(1 - \alpha) \omega_K Y_t^{1-\rho} \mathcal{K}_t^{\rho-\kappa} K_t^{\kappa-1} = r_t^K, \quad (8)$$

$$\text{and } \alpha(1 - \omega_L) Y_t^{1-\rho} \mathcal{L}_t^{\rho-\nu} (z_t^L)^\nu N_t^{\nu-1} + (1 - \alpha)(1 - \omega_K) Y_t^{1-\rho} \mathcal{K}_t^{\rho-\kappa} (z_t^K)^\kappa N_t^{\kappa-1} = w_t. \quad (9)$$

The left-hand sides of (7), (8), and (9) are the marginal products of land, capital, and labor, respectively. To get the different substitution elasticities requires computing the marginal rates of technical substitution between each input then differentiating in the input price ratios. These objects are non-trivial *except* when the primitive, structural elasticities satisfy very specific conditions. Part of the reason for this is that the marginal product of labor (i.e., the left-hand side of (9)) is just the productivity-weighted sum of the marginal products of land-efficiency and capital-efficiency labor:

$$MP_N = z_t^L MP_{z^L N} + z_t^K MP_{z^K N}. \quad (10)$$

Let $\Upsilon_t^{x,y}$ be a function that describes the marginal rate of technical substitution between inputs x and y , $\Upsilon_t^{x,y} = MP_y/MP_x$. The marginal rates of technical substitution for land/labor and capital/labor are then

$$\Upsilon_t^{L,N} := \frac{1 - \omega_L}{\omega_L} (z_t^L)^\nu \left(\frac{N_t}{L_t} \right)^{\nu-1} + \frac{(1 - \alpha)(1 - \omega_K)}{\alpha \omega_L} (z_t^K)^\kappa \frac{\mathcal{K}_t(K_t, N_t)^{\rho-\kappa}}{\mathcal{L}_t(L_t, N_t)^{\rho-\nu}} \frac{N_t^{\kappa-1}}{L_t^{\nu-1}}, \quad (11)$$

$$\Upsilon_t^{K,N} := \frac{\alpha(1 - \omega_L)}{(1 - \alpha)\omega_K} (z_t^L)^\nu \frac{\mathcal{L}_t(L_t, N_t)^{\rho-\nu}}{\mathcal{K}_t(K_t, N_t)^{\rho-\kappa}} \frac{N_t^{\nu-1}}{K_t^{\kappa-1}} + \frac{1 - \omega_K}{\omega_K} (z_t^K)^\kappa \left(\frac{N_t}{K_t} \right)^{\kappa-1}. \quad (12)$$

Equating (11) and (12) with w_t/r_t^L and w_t/r_t^K respectively traces out the isoquants describing the factor tradeoffs as functions of relative prices. Since total labor is featured in *both* the land and capital nests, (11) and (12) also take on an additive form. To understand why, consider how aggregate production would respond to an increase in the labor supply in this economy. Increasing N_t would affect the value of land-owner rents relative to wages, but this effect is the sum of two secondary effects. The first term in (11) encodes a direct effect on the land/labor tradeoff, while the second term in (11) captures the indirect effect due to imperfect complementarities/substitutabilities between land and capital. The same logic holds for the two terms in the sum in (12). Finally, the marginal rate of technical substitution between capital and land is

$$\Upsilon_t^{K,L} := \left(\frac{\alpha \omega_L}{(1-\alpha) \omega_K} \right) \frac{\mathcal{L}_t(L_t, N_t)^{\rho-\nu} L_t^{\nu-1}}{\mathcal{K}_t(K_t, N_t)^{\rho-\kappa} K_t^{\kappa-1}}. \quad (13)$$

While (13) is simpler than (11) and (12), we show below that the elasticity of substitution between land and capital is still non-trivial.

We now require an explicit definition that will help us better illustrate why the model is VES.

Proposition 1. *Let $\mathcal{E}(z, p)$ be the elasticity of z w.r.t. p , let $\vartheta(x, y) = x/y$ be the ratio of two production inputs, and let $\tilde{r}(r^x, r^y) = r^x/r^y$ be the ratio of two rates of return. Then, the Hicks (1932) elasticity of substitution is*

$$\mathcal{E}^{x,y} := \mathcal{E}(\vartheta(x, y), \tilde{r}(r^x, r^y)) = \left| \left(\frac{\partial \Upsilon^{x,y}(\vartheta)}{\partial \vartheta(x, y)} \right)^{-1} \frac{\tilde{r}(r^x, r^y)}{\vartheta(x, y)} \right|. \quad (14)$$

Proof. By definition the elasticity of substitution is $\frac{d\vartheta}{d\tilde{r}} \frac{\tilde{r}}{\vartheta}$. Since $\Upsilon^{x,y}(\vartheta_t) = \tilde{r}_t$, differentiating both sides in $\tilde{r}_t$ gives $\frac{\partial \Upsilon^{x,y}}{\partial \vartheta_t} \frac{d\vartheta_t}{d\tilde{r}_t} = 1$. Rearranging and multiplying by the ratio of relative prices to relative inputs yields the result. **QED**

Proposition 1 says that the Hicks (1932) elasticity of substitution is a function of the derivative of the marginal rate of technical substitution in the input ratio. How $\vartheta_t(x, y)$ (the ratio function) enters $\Upsilon_t^{x,y}$ will determine whether or not the model is VES. It can be directly seen that, indeed, this object is non-constant for all but knife-edge restrictions on the primitive, structural elasticities in our model. We now walk through how we re-write the MRTS conditions in (11) through (13) to depend on ratios, noting that the capital, land, and population levels will still enter these functions except for very specific cases. This is because, generally speaking, we must take the derivative of $\Upsilon_t^{x,y}$ with respect to ϑ_t around fixed levels, assuming only one of the inputs varies.

Starting with (11) and (12), consider first the ratios of land to labor $\vartheta_t(L_t, N_t) = L_t/N_t$ and capital to labor $\vartheta_t(K_t, N_t) = K_t/N_t$. In each of (11) and (12), we substitute out land $L_t = \vartheta_t(L_t, N_t)N_t$ and capital $K_t = \vartheta_t(K_t, N_t)N_t$ respectively to write each of the MRTS conditions as functions of

both ratios and levels:

$$\begin{aligned} \Upsilon_t^{L,N}(\vartheta_t(L_t, N_t); K_t, N_t) &= \frac{1 - \omega_L}{\omega_L} (z_t^L)^\nu \vartheta_t(L_t, N_t)^{1-\nu} \\ &+ \frac{(1 - \alpha)(1 - \omega_K)}{\alpha \omega_L} (z_t^K)^\kappa \frac{\mathcal{K}_t(K_t, N_t)^{\rho-\kappa}}{\mathcal{L}_t(\vartheta_t(L_t, N_t))^{\rho-\nu}} \vartheta_t(L_t, N_t)^{1-\nu} N_t^{\kappa-\rho}, \end{aligned} \quad (15)$$

$$\begin{aligned} \text{and } \Upsilon_t^{K,N}(\vartheta_t(K_t, N_t); L_t, N_t) &= \frac{\alpha(1 - \omega_L)}{(1 - \alpha)\omega_K} (z_t^L)^\nu \frac{\mathcal{L}_t(L_t, N_t)^{\rho-\nu}}{\mathcal{K}_t(\vartheta_t(K_t, N_t))^{\rho-\kappa}} \vartheta_t(K_t, N_t)^{1-\kappa} N_t^{\nu-\rho} \\ &+ \frac{1 - \omega_K}{\omega_K} (z_t^K)^\kappa \vartheta_t(K_t, N_t)^{1-\kappa}. \end{aligned} \quad (16)$$

Notice that in both (15) and (16) the population level N_t and the level of the *opposite* input (i.e., capital in the land/labor MRTS and land in the capital/labor MRTS) do not fall out of the expression. This is because in (15), for example, the non-landed capital nest does not contain land, and so we cannot substitute out $L_t = \vartheta_t(L_t, N_t)N_t$; the same is true for the land nest with respect to capital in (16). Meanwhile, the own-nest (direct effect) can be written strictly in terms of $\vartheta_t(\cdot, N_t)$ because homogeneity can be exploited. For instance, we can substitute $L_t = \vartheta_t(L_t, N_t)N_t$ into $\mathcal{L}_t(L_t, N_t)^{\rho-\nu}$ to get $N_t^{\rho-\nu} \mathcal{L}_t(\vartheta_t(L_t, N_t))^{\rho-\nu}$. Since an expression for the elasticities of substitution requires differentiating the quotient of two functions that depend non-linearly on the input ratios, it should be clear by simply inspecting (15) and (16) that these objects are *not* associated with constant elasticities except in the most knife-edge parameterizations. Thus, the sensitivity of labor per hectare to the wage/land-rental rate ratio depends on how capital-intensive production is (i.e., K_t). Also, the sensitivity of capital per worker to the wage/capital-rental-rate ratio depends on how geographically big the country is (i.e., L_t).

While the land/capital tradeoff in (13) is less complicated, its expression in ratios is still non-trivial and level-dependent:

$$\Upsilon_t^{K,L}(\vartheta_t(K_t, L_t); L_t, N_t) = \left(\frac{\alpha \omega_L}{(1 - \alpha)\omega_K} \right) \frac{\mathcal{L}_t(L_t, N_t)^{\rho-\nu}}{\mathcal{K}_t(\vartheta_t(K_t, L_t)L_t, N_t)^{\rho-\kappa}} \vartheta_t(K_t, L_t)^{1-\kappa} L_t^{\nu-\kappa}. \quad (17)$$

In (17) we substituted $K_t = \vartheta_t(K_t, L_t)L_t$ for K_t which yields an expression for the MRTS of K_t and L_t as a function of the ratio $\vartheta_t(K_t, L_t)$ around the population and land levels. The way the capital per hectare thus responds to changes in the relative rate of return of capital to land depends on the size of the country, both in terms of its land mass and population.

Proposition 2 makes clear that our model is VES except in one knife-edge case. This is because the objects $\frac{\partial \Upsilon^{x,y}(\vartheta)}{\partial \vartheta(x,y)}$ will depend on the levels of L_t , K_t , and N_t and/or the ratios ϑ_t 's, for most parameterizations. Relative allocations thus vary around relative prices for specific values of L_t , K_t , and N_t along the associated, pair-wise isoquants.

Proposition 2. (*Variable Elasticities of Substitution*) For any given pair (z_t^L, z_t^K) and interior values of ω_L , ω_K , and α , the model contains constant elasticities of substitution if and only if $\rho = \nu = \kappa$.

Proof. Let $\sigma \in \{\nu, \kappa\}$. For sufficiency note that when $\rho = \nu = \kappa$ the levels fall out of each of (15),

(16), and (17), and $\Upsilon_t^{x,y} = \bar{C}_t \vartheta_t^{1-\sigma}$, where $\bar{C}_t$ is some period-specific constant that may depend on the productivities. Using $\Upsilon_t^{x,y} = \tilde{r}_t$, it is readily apparent that $\mathcal{E}^{x,y} = (1 - \sigma)^{-1}$ for all (x, y) pairs, and so the model is CES. For necessity suppose the model is CES but at least one of ρ , ν , and κ are not equal to the others. Then, either $(\rho - \nu)$ or $(\rho - \kappa)$ is non-zero, implying that at least one of $\mathcal{L}_t(\vartheta_t)^{\rho-\nu}$, $\mathcal{K}_t(\vartheta_t)^{\rho-\kappa}$, and/or $\mathcal{K}_t(\vartheta_t L_t, N_t)^{\rho-\kappa}$ will enter Υ_t 's non-trivially, and for some pair (x, y) there will exist $\frac{\partial \Upsilon_t^{x,y}}{\partial \vartheta_t}$ that depends on input levels, contradicting the supposition that the model is CES. QED

2.2 Factor Income Shares

Let θ_t^x be the share of aggregate income that goes to factor x . We can exploit homogeneity to rewrite (7) and (8) in terms of land per person and capital per person, respectively. We can then divide both sides by GDP per capita to get the land and capital income shares as functions of land and capital per person and productivities:

$$\alpha \omega_L y_t^{-\rho} \mathcal{L}_t(\ell_t; z_t^L)^{\rho-\nu} \ell_t^\nu = \theta_t^L \quad (18)$$

$$\text{and} \quad (1 - \alpha) \omega_K y_t^{-\rho} \mathcal{K}_t(k_t; z_t^K)^{\rho-\kappa} k_t^\kappa = \theta_t^K. \quad (19)$$

Clearly when $\rho = 0$ these objects become

$$\alpha \omega_L \mathcal{L}_t(\ell_t; z_t^L)^{-\nu} \ell_t^\nu = \theta_t^L, \quad (20)$$

$$\text{and} \quad (1 - \alpha) \omega_K \mathcal{K}_t(k_t; z_t^K)^{-\kappa} k_t^\kappa = \theta_t^K, \quad (21)$$

where θ_t^L is the land share of income, θ_t^K is the capital share of income, and the labor share of income is $\theta_t^N = 1 - \theta_t^L - \theta_t^K$. By inspecting (18) through (21) it should be apparent that factor shares vary non-trivially in both productivities and levels of land and capital per capita. Further, if we possess data on ℓ_t , k_t , y_t (in the full CES case), θ_t^L , and θ_t^K , then we can estimate the model parameters directly from either of the above equation pairs, treating z_t^L and z_t^K as residuals.

Comparing (18) through (21) to the factor income share expressions in the original model of Hansen and Prescott (2002), where the technology and its inputs are split between agricultural versus non-agricultural production, it can readily be seen that our formulation can more flexibly match factor income shares, period by period, in all phases of development. This is because in Hansen and Prescott (2002) factor income shares are constant *except* during the brief Malthus to Solow transition phase. We show this in Appendix A.2. Thus, even if we had the data to be able to split capital and labor utilization between productive purposes (i.e., agriculture and non-agriculture), the Hansen and Prescott (2002) model cannot reconcile the significant fluctuations in the factor shares we observe in Figure 2 both *before* and *after* the Industrial Revolution.

2.3 Growth Properties — Income Shares ($\rho = 0$)

We now invoke a steady state to study the behavior of model objects along the growth path.²¹ We ask, how do factor income shares vary in population and productivity growth rates? Further, what are the implications for the drivers of growth in output per capita, y ? Throughout this section assume that $\rho = 0$, so that the outer nest is Cobb-Douglas and $\nu < 0$.^{22,23} Further, let $\kappa < 1$ be finite and free. Define g_x as the gross growth rate of variable x and $\tilde{g}_x = \ln g_x$.²⁴

Now, note that $\ell = \bar{L}/N$, and so it only varies as N varies. Log growth in the land share of income can thus be written:

$$\tilde{g}_{\theta^L} = -\nu(\tilde{g}_N + \tilde{g}_L), \quad (22)$$

where $\tilde{g}_L$ describes growth in the land composite, which depends on both g_N and g_{z^L} . The land share of income will vary along the development path *only* in response to variation in population and/or the labor efficiency of using land.

Meanwhile, under the assumption that $\rho = 0$, we can write per capita income as $y = \mathcal{L}^\alpha \mathcal{K}^{1-\alpha}$, so that $\tilde{g}_y = \alpha\tilde{g}_L + (1-\alpha)\tilde{g}_K$. This brings us to Assumption 1, which will help pin down the growth properties.

Assumption 1. (Constant Capital Investment Rate) *The non-landed capital investment rate is constant, s , and households maintain a constant non-landed capital/output ratio along the growth path, so that*

$$k/y = s/\hat{\delta}, \quad \text{where} \quad \hat{\delta} \equiv \delta + g_y g_N - 1,$$

δ is the non-landed capital depreciation rate and $g_y g_N$ is the gross growth rate of aggregate output. Thus, $\tilde{g}_k = \tilde{g}_y$.²⁵

Now, with Assumption 1 in hand we can take logged first differences of (21) and replace $\tilde{g}_k$ with $\tilde{g}_y$ to get an expression for the growth rate of the capital share of income:

$$\tilde{g}_{\theta^K} = \kappa\alpha(\tilde{g}_L - \tilde{g}_K). \quad (23)$$

Using expressions (22) and (23), we can engage in several partial decompositions to understand how income shares respond to variation in population and productivities. We examine the following cases:

1. Stagnation ($g_N > 1$ but $g_y = 1$ with $g_{z^L} = g_{z^K} = 1$);

²¹Throughout this section and the next we drop t subscripts.

²²It will become apparent in Section 3 why these are reasonable assumptions.

²³In Section 2.4 we briefly discuss what happens when the $\rho = 0$ assumption is relaxed.

²⁴Thus, if $g_x > 1$ then $\tilde{g}_x > 0$, and x is growing.

²⁵This follows from the capital accumulation identity alone, with no appeal to the structure of the technology: $K_{t+1} = sY_t + (1-\delta)K_t$ implies $K_{t+1}/K_t = s(Y_t/K_t) + 1 - \delta$, and a constant capital/output ratio requires K and Y to grow at the common gross rate $g_y g_N$, giving $K/Y = s/(\delta + g_y g_N - 1)$. Setting $g_y = g_N = 1$ recovers the familiar s/δ , which holds only when aggregate output is static.

2. Pure Malthus ($g_N > 1$ but $g_y < 1$ with $g_{z^L} = g_{z^K} = 1$);
3. Malthusian Productivity Growth ($g_{z^L} > 1$ but $g_y = 1$ with $g_N = g_{z^K} = 1$);
4. Solow Growth ($g_{z^K} > 1$ leading to $g_y > 1$ with $g_N = g_{z^L} = 1$).

Case 1 describes an economy experiencing population growth without the requisite GDP per capita growth.²⁶ Case 1 could be interpreted as a partial decomposition, holding g_y fixed by assumption. Case 2 is pure Malthus (1798): population growth leads to over-crowding of agricultural land and thus declining GDP per capita, absent productivity growth. Case 3 describes what happens when the labor efficiency of using land increases, absent any additional changes. Finally, Case 4 isolates the Solow growth channel, whereby increasing capital-utilization efficiency leads to capital deepening and thus growth in GDP per capita.

2.3.1 Stagnation

Lemma 1. *Under our parametric assumptions population growth absent growth in GDP per capita affects income shares as follows:*

- i. θ^L increases always;
- ii. θ^K increases if $\kappa < 0$ and decreases if $\kappa \in (0, 1)$;
- iii. θ^N decreases if $\kappa < 0$ and increases if $\kappa \in (0, 1)$ and $\kappa\alpha \gg |\nu|$ and/or $|\tilde{g}_{\mathcal{L}}| \gg 0$.

Proof. See Appendix A.1.

QED

Lemma 1 describes how the factor income distribution varies in response to population growth without corresponding growth in GDP per capita, all else equal. In most cases the land and capital shares rise as the population increases, and the labor share falls. Only when capital and capital-efficiency labor are substitutes and land and land-efficiency labor are very weak complements will the labor share rise as the population rises. The capital share rises and falls depending on the sign of κ directly, but since labor is featured in both the land and capital nests, whether the labor share rises in response to population growth depends on the *relative* elasticities, κ and ν . There is also a case where the nest itself, $|\tilde{g}_{\mathcal{L}}|$, grows a lot in absolute value in response to population growth. This occurs when ω_L is close to one, so that the land nest is very land intensive *and* land and land-efficiency labor are weak complements (i.e., $|\nu|$ is very small). In an economy characterized by stagnation, though, land owners and capital owners get bigger pieces of the pie as the population grows *if* such growth cannot be converted into growth in GDP per capita. Thus, Lemma 1 suggests that under most parameterizations rates of return on land and capital rise in response to population growth that does not also translate into GDP per capita growth.

²⁶Such a situation would characterize an under-developed modern economy, like Zambia, for example.

2.3.2 Pure Malthus

Malthus' theorized that population increases led to long-term declines in GDP per capita simply because of over crowding: people would be forced to farm less productive agricultural land thereby reducing per capita food stuffs in spite of aggregate output rising (Malthus 1798). Thus, a purely Malthusian growth regime sees $g_N > 1$ and assumes $g_y < 1$, all else equal. Lemma 2 characterizes how aggregate income is distributed in a purely Malthusian growth regime.

Lemma 2. *Population growth associated with declining GDP per capita affects income shares as follows:*

- i. θ^L increases always;
- ii. θ^K increases if a) both $\kappa < 0$ and $|\nu| < |\kappa|$, or b) $\kappa \in (0, 1)$ and $|\nu| > \kappa$;
- iii. θ^N increases if θ^K falls and population growth is sufficiently slow.

Proof. See Appendix A.1.

QED

Lemma 2 says that if rising population leads to declining GDP per capita because aggregate K is not rising fast enough, any additional income generated from an increase to *total* GDP will usually benefit land owners and capital owners at the expense of workers. When both land and land-efficiency labor and capital and capital-efficiency labor are complementary, as long as capital/labor are relatively *more* complementary, non-landed capital owners benefit at the expense of workers. However, if capital/labor are substitutes within the $\mathcal{K}$ nest, the opposite mechanism is at work, driving up θ^K : if land/labor are strong complements, non-landed capital owners benefit at the expense of workers. Workers only see their share of income grow if the capital share falls faster than the land share rises. This happens when population growth is sufficiently slow *and* land and land-efficiency labor complementarities dominate in a regime in which capital and capital-efficiency labor are still complementary *or* when K and $z^K N$ are substitutes and L and $z^L N$ are only very weak complements.

2.3.3 Malthusian Productivity Growth

What happens when z^L grows, all else equal?²⁷ The premise behind this exercise is to understand how income is redistributed along a development path as labor becomes relatively more efficient at utilizing land, despite population and GDP per capita remaining fixed. Note that the non-landed capital nest, $\mathcal{K}$, is independent of z^L and thus unaffected by growth in z^L , assuming no additional growth in k leading to rising y . Lemma 3 characterizes how income shares respond to land-utilization efficiency changes.

Lemma 3. *Growth in z^L , all else equal, affects income shares as follows:*

- i. θ^L increases always;

²⁷Like in the “stagnation” exercise in Section 2.3.1, we hold y fixed.

ii. θ^K decreases if $\kappa < 0$ and increases if $\kappa \in (0, 1)$;

iii. θ^N increases if $\kappa < 0$ and $\theta^K|\kappa|\alpha > \theta^L|\nu|$.

Proof. See Appendix A.1.

QED

Recall that z^L encodes the labor efficiency of using land. This object embeds knowledge associated with best practices pertaining to agricultural production as well as access to arable land, among other things. In the event of an increase in this productivity, land owners always benefit, but workers may also benefit if non-landed capital and labor are sufficiently *more* complementary than land and the non-landed capital income share is already sufficiently high. When this happens, an increase in z^L frees up labor to be used in more (relatively) capital-intensive production, benefiting workers at the expense of non-landed capital owners. The opposite occurs when capital and capital-efficiency labor are substitutes: an increase in the labor efficiency of using land benefits land owners, but workers are unable to bargain for more capital-intensive work, since capital and labor are substitutes. This drives down wages, and capital owners benefit from cheaper labor. Finally, θ^N may fall even when efficiency labor is complementary with both land and capital if land's share of income is too high.

2.3.4 Solow Growth

A growth regime a la Solow (1956) and Solow (1957) occurs when $g_{z^K} > 1$ leading to $g_y > 1$. That is, the labor-efficiency with respect to utilizing capital drives capital accumulation and thus growth in GDP per capita. Recall that under our assumptions only in this section, GDP per capita grows in constant proportion to capital per capita, which is consistent with the conventions governing exogenous growth models with constant returns to scale production functions.

Lemma 4. *Growth in z^K leading to growth in y affects income shares as follows:*

i. θ^L does not change;

ii. θ^K increases if $\kappa < 0$ and decreases if $\kappa \in (0, 1)$;

iii. θ^N decreases if $\kappa < 0$ and increases if $\kappa \in (0, 1)$.

Proof. See Appendix A.1.

QED

Lemma 4 describes how income is distributed along a Solow-style growth path. The results are remarkably simple. When capital and capital-efficiency labor are complements, capital owners get a bigger share of the pie at the expense of workers, while the opposite is true when capital and labor are substitutes. One implication of Lemma 4 is that land's share of income does not change in response to industrialization driven by purely capitalistic technological improvements. Thus, to replicate the input-share patterns we see in Figure 2 whereby land's share falls as capital's share rises, we need something else to occur. Specifically, we need either z^L or N to fall. Clearly, given

industrialization is almost always associated with increasing population, the former (z^L declining) generates results that are more empirically consistent.²⁸

But, does it make sense that industrialization would also be simultaneously associated with a falling land-efficiency of labor? Let us recall what z^L represents — knowledge with respect to how best to utilize land to achieve productive output, as well as accessibility to arable land by a large enough mass of laborers. Enclosure movements (both formal and informal) gradually limited the peasant class' access to common lands, perhaps adversely affecting the labor-efficiency of land utilization (North 1981; Ostrom 1990). But what about land-utilization knowledge embedded in z^L ? Absent access, such folk knowledge could also atrophy.

Further, consider the following thought experiment: an industrial (or post-industrial) economy suddenly is thrust into apocalyptic nuclear armageddon, killing 80% of the population. All machines are rendered useless due to shocks to electromagnetic fields, so survivors must learn how to engage in manual, subsistence agriculture in order to feed themselves. Do we honestly believe that such survivors, having spent their lives (and descended from multiple generations of people) who have worked in factories and offices, would be just as good (if not better) than their forebears from hundreds of years prior at tending the land without advanced technologies to do so? Of course not. The Medieval peasant was likely far more knowledgeable about best agricultural practices when faced with an input menu devoid of advanced technologies (e.g., tractors, fertilizers, harvesters, etc.). Indeed, under such a thought experiment, it would actually be rather unreasonable to think that z^L grows as economies industrialize and workers stop engaging in subsistence agricultural practices but instead learn to use more capital-intensive technologies (i.e., z^K rises). This theoretical result thus gives us an idea of what to expect from a well-posed quantitative model: rising z^K and falling z^L around the onset of Industrial Revolution. Indeed, our quantitative approach will allow us to test whether or not this is what actually happened.

2.4 Growth Properties — Income Shares ($\rho \neq 0$)

We now briefly discuss what happens when the $\rho = 0$ assumption is relaxed and asymptotic balanced growth properties under a full CES outer nest are considered. This section is primarily just discussion of more detailed results that we reserve for Technical Appendix B.

Several key differences distinguish the reallocative properties under a generalized CES framework from the Cobb-Douglas outer nest. First, the independence of θ^L from variation in z^K no longer holds when $\rho \neq 0$, so that mechanical declines in θ^L can arise under a Solow growth regime, absent any changes to the land-utilization productivity, z^L . In particular if the land and capital nests are themselves weakly gross substitutes ($0 < \rho < 1$ with $\rho < 1$) while the inner nests feature gross complementarities ($\nu, \kappa < 0$), the land share will fall as the capital share rises *mechanically* in response to a simple increase in z^K , independent of any variation in z^L . Mean-

²⁸Some have even posited that the Industrial Revolution is partially the *result* of demographic transition, namely a reduction in mortality rates for the young, leading to population increases overall (Bar and Leukhina 2010a; Bar and Leukhina 2010b).

while, if we flip the signs of the elasticities such that $\rho < 0$ and $\nu, \kappa \in (0, 1)$, the opposite occurs: θ^K falls always in response to increasing capital-utilization efficiency, while θ^L rises — a sort of factor-input cost disease phenomenon.

When quantifying the model, we generally prefer a $\rho = 0$ specification for more technical reasons pertaining to model fit as we discuss in our estimation below. While the case of weak gross substitutes with inner-nest complementarities *does* indeed yield reallocative effects in the Solow growth regime that *exactly mimic* those we find in the data, independent of variation in land-utilization efficiency, model comparison assessments we engage in below suggest that a model where $\rho \in (0, 1)$ and $\nu, \kappa < 0$ is not the best parameterization for the data.

2.5 Growth Properties — TFP vs. Population Growth

While thus far we have analyzed a VES version of the model, in many aggregate economics applications it is common to use a restrictive CES production function, and still then an often more restrictive Cobb-Douglas model. In his original endeavors estimating total factor productivities, Solow (1956) discusses that a Cobb-Douglas version of the aggregate production function should be preferred because it best fits his limited empirical regularities. In this section we consider the growth properties of a simplified, fully Cobb-Douglas version of the per capita production function in (6) (i.e., $\rho = \nu = \kappa = 0$) with both land per capita and capital per capita as inputs to show that Solow-style TFP measurements are actually biased downward when failing to account for land's role in aggregate production. Note that we only engage in this exercise to demonstrate the importance for measurement of including the fixed-factor (land) in the production function: the Cobb-Douglas version has constant income shares, so it would not be able to replicate the patterns we observe from the Clark (2010) data.

The fixed factor (land) ensures that population growth and TFP growth compete against each other, as described in Jones and Vollrath (2013), such that relative growth rates between the two will determine whether or not the economy appears Malthusian or Solow in its properties. The primary innovation we add is to show that with our purely Cobb-Douglas setup when TFP and population grow at the same rate, the positive effects of TFP growth always outweigh the negative effects of population growth.

The Cobb-Douglas (C-D) technology comes from sending ρ , ν , and $\kappa \rightarrow 0$ to yield

$$Y_t = A_t(z_t^L, z_t^K) L_t^{\alpha\omega_L} K_t^{(1-\alpha)\omega_K} N_t^{\alpha(1-\omega_L)+(1-\alpha)(1-\omega_K)}, \quad (24)$$

where the function $A_t(z_t^L, z_t^K) = (z_t^L)^{\alpha(1-\omega_L)} (z_t^K)^{(1-\alpha)(1-\omega_K)}$ is the TFP for this economy, which depends on the different land/labor and capital/labor productivities. In per capita terms (24) becomes:

$$y_t = A_t \ell_t^{\alpha\omega_L} k_t^{(1-\alpha)\omega_K}. \quad (25)$$

Now, consider dynamics along a balanced growth path and drop time subscripts. By exploit-

ing Assumption 1 and substituting $y = (\widehat{\delta}/s)k$ into (25), balanced-growth capital per capita can be written

$$k = \left[\frac{sA}{\widehat{\delta}} \ell^{\alpha\omega_L} \right]^{\frac{1}{1-(1-\alpha)\omega_K}}. \quad (26)$$

The Solow economy, wherein capital per capita is independent of the population level, is recovered as land's importance for production vanishes (either $\omega_L \rightarrow_+ 0$ or $\alpha \rightarrow_+ 0$). As long as production requires land, capital per capita is declining in population, since $\ell = \bar{L}/N$ and $\bar{L}$ is fixed. In the classic Solow setting capital per capita only changes in response to variation in capital accumulation rates (as captured by s), the depreciation rate, or TFP. Solow dynamics as they pertain to the relationship between capital growth and the population are thus broken when a third, fixed production factor (i.e., land) is included.

Because balanced-growth capital per capita depends on the population level, so does balanced-growth output per capita which is

$$y = \left[\frac{s}{\widehat{\delta}} \right]^{\frac{(1-\alpha)\omega_K}{1-(1-\alpha)\omega_K}} A^{\frac{1}{1-(1-\alpha)\omega_K}} \left[\frac{\bar{L}}{N} \right]^{\frac{\alpha\omega_L}{1-(1-\alpha)\omega_K}}, \quad (27)$$

where we write $\ell = \bar{L}/N$ to show that N still enters the expression for y because $\bar{L}$ is fixed. A key feature of a Malthusian economy is that output per capita declines as the population level rises. This is still the case in this model as long as land is positively supplied in production. In the Solow world the primary driver of growth in y is variation in A , however in this economy (27) shows that a population decline *could also* lead to growth in y , even if A were held fixed.

Indeed, the degree to which output per capita grows in response to growth in A versus a decline in N depends on the capital intensity of production, α , and both land/labor and capital/labor input complementarities, ω_L and ω_K . Proposition 3 states that the positive effects of TFP growth will always outweigh the negative effects of population growth if each grow proportionally. Thus, despite the fact that y is decreasing in N , if A grows fast enough, the economy will appear to be characterized by a Solow-style growth regime as opposed to a Malthusian one.

Proposition 3. (*Solow vs. Malthusian Growth Dynamics*) Let α , ω_L , and $\omega_K \in (0, 1)$. Suppose A and N grow at the same rates. Then, the positive effects of TFP growth will always dominate the negative effects of population growth, and output per capita will grow.

Proof. See Appendix A.1. QED

As Proposition 3 explains, A does not even need to grow as fast as N in order to ensure output per capita grows. That is, output growth is less sensitive to the negative effects of population growth than it is to the positive effects of productivity growth. Expressed in terms of elasticities, Proposition 3 simply says that the elasticity of GDP per capita with respect to population is less than the elasticity of GDP per capita with respect to TFP, as long as all factors (land, capital, and

labor) matter for aggregate production (i.e., α , ω_L , and ω_K are all interior). This is self-evidently true for modern, advanced economies which have grown *despite* experiencing population growth. Proposition 3 thus suggests simple, two-factor technologies used to describe growth today may be missing something. In Technical Appendix F we provide a visual illustration of Proposition 3, showing how under different parameter combinations TFP growth more or less significantly dominates population growth. Because $g_y > 0$ requires only $\tilde{g}_A > \alpha\omega_L\tilde{g}_N$, land-intensive economies tip into negative per-capita growth more readily, which is the Malthusian point those panels exist to make.

2.6 Single-sector vs. Multi-sector Interpretation

Our economy is single sector by definition for several reasons that we reiterate: 1) long-run data limitations with regards to splitting up factor income across sectors; 2) much of the long-run development literature actually mathematically operates in quasi single sector economies, despite using language presenting models as sums of two sectors (i.e., agriculture or a “Malthusian” sector and non-agriculture or a “Solow” sector);²⁹ This choice, though, comes with one shortcoming: absent splitting up activity (inputs or outputs) we cannot relate what our model has to say about growth and development to things like the coinciding decline in the share of aggregate labor used in agricultural production and corresponding improvements to agricultural productivity.³⁰ To address this multi-sector critique, in Technical Appendix D we consider a re-interpretation of our aggregate production function as a multi-sector one, where $\mathcal{L}_t$ represents a land-intensive sector (e.g., pre-industrial agriculture) and $\mathcal{K}_t$ represents a capital-intensive one. There, we show that under some additional (admittedly strong) assumptions our one-sector production function is *exactly* the reduced form of a two-sector economy with explicit labor allocation across sectors, the effects of which are embedded in the utilization productivities, z_t^L and z_t^K .

To understand this, let $n_t^L \in (0, 1)$ (lower-case) be the share of aggregate labor in the land-intensive sector and $n_t^K \in (0, 1)$ the share of aggregate labor in the capital-intensive sector (i.e., $n_t^L + n_t^K = 1$). Let Y_t be a final good comprised of the weighted shares of output from the land- and capital-intensive sectors and whose aggregation is fully characterized by either the generalized CES structure in (1) or the Cobb-Douglas structure in (2). Then, land-intensive sector efficiency labor can be written $z_t^L N_t = A_t^L n_t^L N_t$, and capital-intensive sector efficiency labor can be written $z_t^K N_t = A_t^K n_t^K N_t$ where the variables A_t^L and A_t^K can be interpreted as sectoral labor-augmenting productivities. Thus, if the land-intensive sector is *assumed* to not utilize capital and the capital-intensive sector is *assumed* to not utilize land, then our one-sector model is isomorphic to a two-sector model up to a re-interpretation of the exogenous productivity components. In this interpretation the aggregate productivities z_t^L and z_t^K are composites of sector-level labor-

²⁹See Hansen and Prescott (2002), Bar and Leukhina (2010a), and Ravikumar and Vandenbroucke (2025).

³⁰Labor’s share in agricultural employment generally declines while agricultural productivity generally rises as industrialization occurs due to a rise in manufacturing (Gollin, Parente, and Rogerson 2002; Gollin, Parente, and Rogerson 2004; Gollin, Parente, and Rogerson 2007; Restuccia, Yang, and Zhu 2008; Gollin, Lagakos, and Waugh 2014; Gollin and Rogerson 2025).

augmenting productivities and the labor allocation itself.³¹ Each z_t term absorbs both “how productive” a sector’s workers are and “how many” workers the sector employs.

3 Estimating the Full Technology

Continuing with our case study of the macroeconomic development of England + Wales, we now seek to use input/output data to estimate the model’s primitive, structural elasticities and back out time series estimates for the two different labor-augmenting productivities, z_t^L and z_t^K . With elasticities and productivities in hand, we can then examine how elasticities of substitution change over different epochs of economic growth and assess what specifically was the catalyst for industrialization. What was the timing of Solow-style productivity growth? Specifically, we use the model estimates to understand when exactly growth in z_t^K started to really take off and place a timing window on the Industrial Revolution’s germination. In backing out productivities along a growth path, our efforts are analogous to those of Solow (1957), just extended to a longer time horizon with a richer production structure.

Our primary estimators operate on a system of two equations derived by re-arranging the factor-share conditions in (20) and (21) from the $\rho = 0$ model where the outer nest is assumed to be Cobb-Douglas.³² We take logs in order to isolate $\ln z_t^L$ and $\ln z_t^K$ on one side of each equation with parameters and data on the other side:

$$\ln z_t^L = \ln \ell_t + \frac{1}{\nu} \ln \Gamma^L(\ell_t, \theta_t^L) - \frac{1}{\nu} \ln(1 - \omega_L), \quad (28)$$

$$\ln z_t^K = \ln k_t + \frac{1}{\kappa} \ln \Gamma^K(k_t, \theta_t^K) - \frac{1}{\kappa} \ln(1 - \omega_K). \quad (29)$$

The functions Γ^L and Γ^K are³³

$$\Gamma^L(\ell_t, \theta_t^L) = \frac{\alpha \omega_L}{\theta_t^L} - \omega_L, \quad (32)$$

$$\Gamma^K(k_t, \theta_t^K) = \frac{(1 - \alpha) \omega_K}{\theta_t^K} - \omega_K. \quad (33)$$

³¹Such a two-sector interpretation, though, is admittedly coarse especially in light of the importance of capital accumulation in driving measurable productivity growth in agriculture — ostensibly, the Malthusian or land-intensive sector (Caunedo and Keller 2021).

³²In the main text, here, we only include estimates when $\rho = 0$, since, as will be seen, such models have the best fitness properties. In Technical Appendix C.1 we estimate additional models with $\rho \neq 0$, which generally do not fit as well as the best-fitting $\rho = 0$ model, as we discuss below.

³³Note that when $\rho \neq 0$ (i.e., the outer nest is not Cobb-Douglas), these functions also depend on per-capita income, y_t :

$$\Gamma^L(y_t, \ell_t, \theta_t^L) = \left[(\theta_t^L)^{\frac{\nu}{\rho-\nu}} (\alpha \omega_L)^{\frac{\nu}{\rho-\nu}} \ell_t^{\frac{\rho\nu}{\rho-\nu}} y_t^{\frac{\rho\nu}{\rho-\nu}} - \omega_L \right], \quad (30)$$

$$\Gamma^K(y_t, k_t, \theta_t^K) = \left[(\theta_t^K)^{\frac{\kappa}{\rho-\kappa}} ((1 - \alpha) \omega_K)^{\frac{\kappa}{\rho-\kappa}} k_t^{\frac{\rho\kappa}{\rho-\kappa}} y_t^{\frac{\rho\kappa}{\rho-\kappa}} - \omega_K \right]. \quad (31)$$

In fact when $\rho \neq 0$, the Γ^j functions become slightly more complicated, but the general structures of (28) and (29), because all additional complexity is embedded in the Γ^j ’s.

For the full derivation of these estimating equations, see Appendix A.3. To the econometrician z_t^L and z_t^K are unknown time-dependent parameters, so that, just because we place them on the left-hand sides of (28) and (29), they should *not* be interpreted as dependent variables. We write the model this way because *all variables* on the right-hand side are endogenous and it will help us illustrate, below, how the likelihood function is constructed around realizations of unobserved log productivities, $\ln z_t^j$, $j \in \{L, K\}$, which are totally summarized by the structures on the right-hand sides of (28) and (29).

Given data on GDP per capita, land (hectares) per capita, non-landed capital per capita, the income share of land, and the income share of non-landed capital, we can estimate the above system's structural elasticities, letting the labor-efficiency productivities be residuals. For data we turn to Clark (2010). British land is measured as hectares of land in England and Wales and is assumed constant at 15,357,900 hectares. The population (in millions) is taken from Table 7 of Clark (2010). Nominal capital is computed by multiplying the capital share of income from Table 13 against national income, divided by the rate of return on capital from Table 7. For output we use net national income from Table 13. To get real variables we divide all nominal series (in this case, just capital and output) by Clark's national product deflator (i.e., GDP deflator) from Table 27. This yields a dataset with 67 decades of observations (periods) and five variables $\{y_t, \ell_t, k_t, \theta_t^L, \theta_t^K\}_{t=1200}^{1860}$, which we must use to estimate a model with six fixed parameters, $\{\alpha, \rho, \omega_L, \nu, \omega_K, \kappa\}$, and two sets of time varying productivities, $\{z_t^L\}_{t=1200}^{1860}$ and $\{z_t^K\}_{t=1200}^{1860}$.

As Herrendorf et al. (2015) point out, to identify structural elasticities, one must either make additional assumptions on the values that productivities take or normalize the output elasticities, ω_L and ω_K . In this paper we address the identification problem with a Bayesian approach, specifying prior distributions on the structural parameters that satisfy theory. The Bayesian approach also adds computational complexity for the computer at the benefit of reducing coding time for the econometrician. Further, Bayesian Monte Carlo methods which sample long enough from their underlying distribution are guaranteed to converge, so they are robust to things like starting values and the selection of tuning parameters. One issue we face with a Bayesian approach, which will become more apparent below, is that we must impose prior assumptions on the bounds of the elasticity parameters to ensure the model is well posed without the possibility for singularities. In practice this means restricting ν and/or κ such that the various nests imply their components are either substitutes or complements within the specific estimation routine. Ensuring the parameter space being sampled from is smooth is necessary for the particular type of integration scheme we deploy (Hamiltonian Monte Carlo a la Neal (2011)) to recover the posterior distribution of underlying parameters conditional upon data and priors.

We consider six different estimations over our full 1200-1860 sample:³⁴ four estimators restrict $\rho = 0$ (i.e., Cobb-Douglas upper tier) which we present in the main text; in Technical Appendix C.1 one puts $\rho \in (0, 1)$ (i.e., land nest and capital nest are imperfect substitutes), and another puts $\rho < 0$ (i.e., land nest and capital nest are imperfect complements). Of the four estimators that restrict the upper tier nest to Cobb-Douglas, we restrict ν and κ in different ways so that the $\mathcal{L}$ and $\mathcal{K}$ nests are either complements, substitutes, or a combination of the two. Specifically, we consider estimators with $\nu, \kappa < 0$ (both complements), $\nu < 0$ and $\kappa \in (0, 1)$ (labor nest complements; capital nest substitutes), $\nu \in (0, 1)$ and $\kappa < 0$ (labor nest substitutes; capital nest complements); $\nu, \kappa \in (0, 1)$ (both substitutes). In Technical Appendix C.1 when we assume $\rho \in (0, 1)$, we restrict $\nu, \kappa < 0$, while when $\rho < 0$, we restrict $\nu, \kappa \in (0, 1)$. Note that we must restrict the estimation parameter region by parameter region because of discontinuities in (28) and (29) when $\nu, \kappa = 0$.

The inner-nest structural elasticity priors are four combinations of the following:

$$\begin{aligned} -\nu &\sim \mathcal{LN}(0, 2^2) \quad \text{or} \quad \nu \sim \mathcal{U}[0, 1], \\ -\kappa &\sim \mathcal{LN}(0, 2^2) \quad \text{or} \quad \kappa \sim \mathcal{U}[0, 1]. \end{aligned}$$

That is, when $\nu, \kappa < 0$ we assume their absolute values take a log-normal prior ($\mathcal{LN}$), and when they are positive, we give them uniform priors ($\mathcal{U}$) on the unit interval. We let $\omega_j \sim \mathcal{U}[0, 1]$ always for each $j \in \{L, K\}$. We restrict $\alpha \sim \mathcal{U}[\underline{\alpha}, \bar{\alpha}]$ where $\underline{\alpha} = \max\{\theta_t^L\}$ and $\bar{\alpha} = 1 - \max\{\theta_t^K\}$. Using the Clark (2010) data, $\underline{\alpha} = 0.270$ and $\bar{\alpha} = 0.681$.³⁵ Finally, we form the likelihood around the vector $\ln \mathbf{z}_t$ (vectors are denoted with bold face font), which we give a strong, hierarchical prior. A strong hierarchical prior is the Bayesian equivalent of placing additional structure on the productivity process in order to be able to better identify the elasticities a la Herrendorf et al. (2015). Specifically, we let $\ln \mathbf{z}_t$ have an unconditional (i.e., not conditional on other parameters' distributions or the data distribution) marginal distribution with a hierarchical structure

$$\begin{aligned} \ln \mathbf{z}_t &\sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Omega}) \quad \text{with} \quad \boldsymbol{\mu} \sim \mathcal{N}(\mathbf{0}, 100 \cdot \mathbf{I}), \\ \text{where} \quad \boldsymbol{\Omega} &= \boldsymbol{\Xi}(\sigma^2 \mathbf{I})\boldsymbol{\Xi}^\top, \quad \boldsymbol{\Xi} \sim \text{LKJ}(2), \quad \text{and} \quad \sigma_j^2 \sim \mathcal{N}_{(0, \infty)}(0, 2^2), \quad \forall j \in \{L, K\}. \end{aligned}$$

We thus let $\mathbf{z}$ be multivariate log-normal with a flat hierarchical prior, where the variance/covariance matrix $\boldsymbol{\Omega}$ has its Cholesky-factored lower-triangular matrix sampled using an “LKJ” prior for reasons of computational efficiency (see Lewandowski, Kurowicka, and Joe (2009)), and the diagonal

³⁴Data from Clark (2010) may feature biased estimates of the labor-income share, especially early in the sample, for several reasons (we thank a referee for pointing this out). First, Clark (2010) data tend to lack direct measurements as to the earnings of agricultural workers so that the labor-share measures are skewed, especially early in the sample, toward more urban craft/trade workers. Second, there are institutional reasons pre-1500, which may make measurement difficult overall (e.g., feudal serfdom practices making it difficult to measure labor income, etc.). For robustness purposes we do two things. First, we also estimate the model on a shortened time series of data that runs from 1500-1860, featuring the results in Technical Appendix C.7. Second, in Technical Appendix E we bring in micro data from Allen (1992) and Allen (2007) to recompute the labor share from scratch and compare the computation to the labor shares reported in Clark (2010). Nothing qualitatively serious changes under either robustness check.

³⁵These bounds ensure that Γ^L and Γ^K are always positive.

terms of $\mathbf{\Omega}$ are each given half-normal priors. In practice this means that the likelihood of realizing the fixed parameters given data is the probability density function, $p(\cdot)$, of the two-dimensional normal distribution:

$$p(\mathbf{f}_t(\theta_t, \ell_t, k_t; \boldsymbol{\omega}, \nu, \kappa, \alpha); \boldsymbol{\mu}, \mathbf{\Omega}) = \frac{\exp \left\{ -\frac{1}{2}(\mathbf{f}_t - \boldsymbol{\mu})^\top \mathbf{\Omega}^{-1}(\mathbf{f}_t - \boldsymbol{\mu}) \right\}}{2\pi \sqrt{\det(\mathbf{\Omega})}}, \quad (34)$$

where $\mathbf{f}_t(\theta_t, \ell_t, k_t; \boldsymbol{\omega}, \nu, \kappa, \alpha)$ is a vector-valued function stacking the right-hand sides of (28) and (29).^{36,37} Note that $\ln z_t = \mathbf{f}_t(\theta_t, \ell_t, k_t; \boldsymbol{\omega}, \nu, \kappa, \alpha)$ in each period along the development path. The log productivities, $\ln z_t$, are thus realizations from the distribution characterized by (34), the parameters of which (i.e., $\boldsymbol{\omega}$, ν , κ , α , $\boldsymbol{\mu}$, and $\mathbf{\Omega}$) must be estimated given the available data from Clark (2010). More specifically, $\ln z_t$ is the vector of residuals which ensure (28) and (29) hold. The period-by-period log productivities are thus recovered so as to exactly reconcile each period's factor income shares, conditional upon the assumed model structure. With productivities in hand, the model thus fits the factor income shares exactly.

As mentioned we take a Bayesian approach to our estimation strategy, which involves recovering the posterior distribution of structural parameters (i.e., the static/fixed parameters) given prior distributional assumptions and data. Note that recovery of the posterior does not involve directly estimating $\ln z_t$: these objects are residuals which themselves must be recovered from the posterior predictive distribution *after* it has been parameterized by the estimated static parameters. Thus, $\ln z_t$'s are algebraically determined given the estimated model parameters and data.

Table 1: Production Function Estimates; England + Wales, 1200-1860 ($\rho = 0$)

	(1)	(2)	(3)	(4)
	$\nu, \kappa < 0$	$\nu < 0, \kappa \in (0, 1)$	$\nu \in (0, 1), \kappa < 0$	$\nu, \kappa \in (0, 1)$
ν	-49.090 (2.8129)	-85.765 (3.6532)	0.958 (0.0006)	0.951 (0.0008)
κ	-1.448 (0.0111)	0.965 (0.0005)	-5.376 (0.1791)	0.966 (0.0004)
ω_L	0.508 (0.006)	0.511 (0.005)	0.486 (0.006)	0.490 (0.006)
ω_K	0.518 (0.006)	0.478 (0.006)	0.504 (0.006)	0.480 (0.006)
α	0.439 (0.002)	0.318 (0.001)	0.624 (0.001)	0.443 (0.001)
WAIC ^a	121.319	214.231	146.265	210.699

^a Widely applicable info. criterion, smaller = better (Watanabe 2010; Vehtari, Gelman, and Gabry 2017).

We estimate all models using a Hamiltonian Monte Carlo (HMC) integration technique to re-

³⁶Given data and primitive elasticities, the function $\mathbf{f}_t$ thus outputs a two-dimensional vector, $(\ln z_t^L, \ln z_t^K)^\top$.

³⁷When $\rho \neq 0$ so that the outer nest is not Cobb Douglas, $\mathbf{f}_t$ will also depend on output per capita, y_t , which can be seen by inspecting the Γ^j functions in (30) and (31), and the parameter ρ .

cover the posterior.³⁸ Each of the six separate estimations are run for 2,000 sampling iterations with 1,000 warm-up draws. The convergence properties are robust, and we compare the models' fitness using the widely applicable information criterion (WAIC) which is robust to autocorrelations within the progression of the Hamiltonian Monte Carlo chain (Watanabe 2010; Vehtari, Gelman, and Gabry 2017). For this value, smaller numbers in signed magnitude (including negative numbers) indicate a better fit. Posterior means for the structural elasticities are presented in Table 1. Columns are the separate models (with different prior assumptions) considered. The values in parentheses are standard errors of the posterior means which can be used to assess the statistical significance of the means as point estimates. Estimates from models where $\rho \neq 0$ are shown in Technical Appendix C.1.

From Table 1 the best fitting model is that where $\rho = 0$ and $\nu, \kappa < 0$ from the first column. In this model the upper-tier production function is Cobb-Douglas and the lower-tier nests each feature complementarities, though land and labor-efficiency complementarities are significantly stronger, which can be seen by comparing $\hat{\nu}$ and $\hat{\kappa}$: $-49.090 < -1.448$. This is the best fitting model, both with respect to the information criterion we deploy, as well as for non-targeted output data, which we discuss in the next section.

3.1 Robustness & Diagnostics

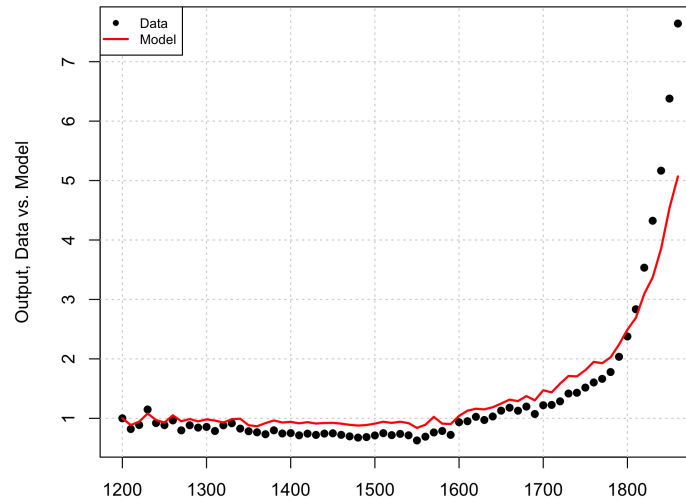


Figure 3: In this figure we present GDP (output) as implied by the time series of the preferred model's ($\rho = 0$ and $\nu, \kappa < 0$) posterior mean and data, each normalized to unity in the year 1200. Other models' predictive fits are featured in Technical Appendix C.2.

As a robustness exercise we also plot model-predicted GDP, Y_t , normalized to unity in 1200

³⁸For detailed explanations of HMC posterior estimation, see Neal (2011), Betancourt and Stein (2011), Gelman et al. (2013a), and Gelman et al. (2013b).

against data from Clark (2010).³⁹ To compute estimates for $\hat{Y}_t$ from the model, we use the period-by-period (i.e., decade-by-decade) posterior means for $\ln \hat{z}_t$ and the posterior means of our structural elasticity estimates. Figure 3 presents $\hat{Y}_t$ against data for the preferred model. The preferred model does a good job picking up the slow acceleration in GDP growth in the early to later 1600s, as well as the transition to exponential growth in the early 1800s.

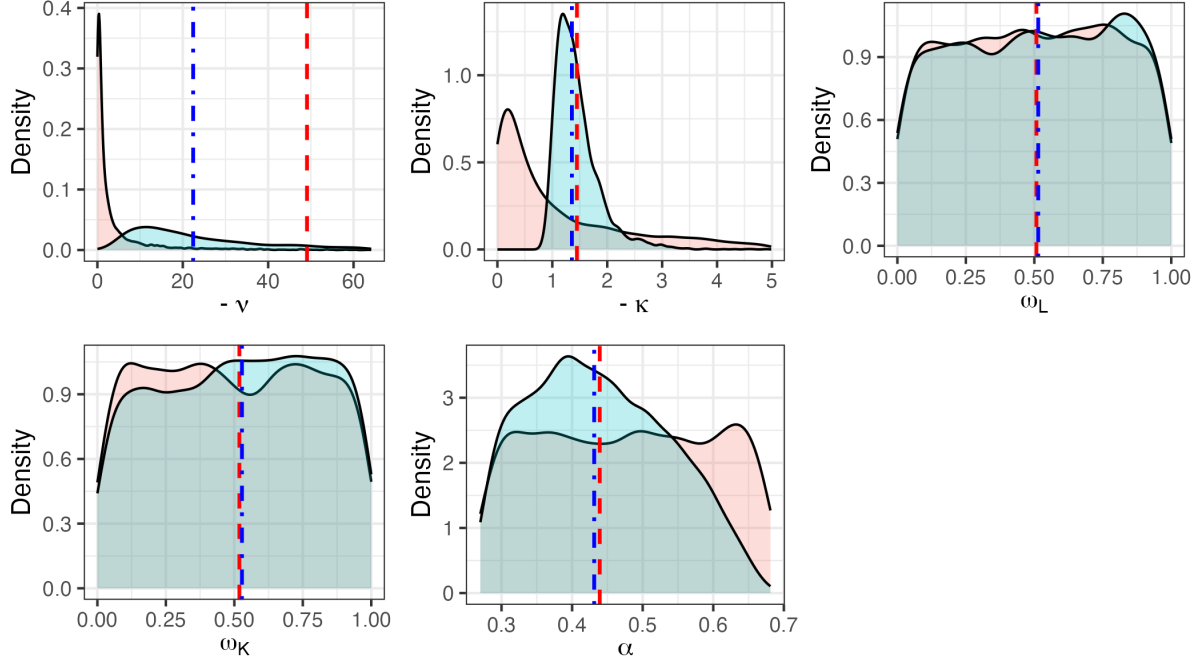


Figure 4: In this figure we plot posterior distributions (teal) against their priors (pink) for the five primary structural parameters of the preferred model ($\rho = 0$ and $v, \kappa < 0$). Posterior means are denoted with red dashed lines, and posterior medians are denoted with blue dashed lines. Similar plots for the alternative models are presented in Technical Appendix C.3.

Given previously documented identification issues in CES production functions when simultaneously trying to estimate factor-augmenting productivities and elasticities (Herrendorf et al. 2015), we also engage in several identification checks. A high-level assessment examines the degree to which posterior distributions differ from their prior assumptions. We also engage in a more granular robustness check, assessing the sensitivity of model outcomes to one-by-one variation in the underlying fixed (not time-varying) parameters.

To assess prior/posterior similarity, we first plot posterior and prior distributions together to visually assess their differences. Note that, while Herrendorf et al. (2015) require a normalization on productivity growth processes to ensure identification of elasticities, our Bayesian method instead imposes what amounts to a distributional restriction on the values of the output elasticities, ω_L and ω_K . Figure 4 presents the prior distributions (pink) against the posterior distributions (teal) for the five structural elasticities. The intercepts of (28) and (29), governed by the

³⁹Note that population is exogenous, so it is sufficient as a robustness exercise to simply compare GDP rather than GDP per capita.

input/output weights ω_L and ω_K , are poorly identified, as can be seen by noting that their posteriors are observationally equivalent to their uniform priors. Restricting ω_L and ω_K to uniform priors on the unit interval amounts to effectively assuming $\hat{\omega}_L = \hat{\omega}_K = \frac{1}{2}$, which is approximately the value of their estimated posterior means. While ω_L and ω_K are poorly identified, the parameters governing the lower-tier nests' elasticities of substitution, ν and κ , appear to be well identified. Note that $-\nu$ and $-\kappa$ have posteriors with small variances that differ substantially from their fat-tailed prior distributions.⁴⁰ Further, α which governs the weight of the land nest in production is also well-identified in a Bayesian sense, in that it has a posterior mean and standard error of 0.439 and 0.002 respectively against a prior assumption that it takes on a uniform distribution on the $[\underline{\alpha}, \bar{\alpha}]$ interval.

While a prior/posterior comparison tells us how much information in data improve prior assumptions, the plots in Figure 4 cannot tell us how sensitive the model is to variation in particular parameters. In Technical Appendix C.6 we fix the different non-time-varying parameters $\{\alpha, \nu, \kappa, \omega_L, \omega_K\}$, one by one, along various grids and assess how this perturbation procedure affects posterior predictive performance. To summarize, the sensitivity checks reveal several things. First, despite its extreme magnitude, ν does *not* exert significant leverage over the predictive power of the model; that is, varying ν does not significantly affect inferences on non-targeted output, $\hat{Y}_t$, as much as varying κ or α . Second, in our preferred model productivity first-order and second-order properties are practically invariant to variation in the fixed parameters; instead, perturbing the fixed parameters just changes the average *levels* of the productivities but not whether they decline and/or rise before, during, and after different epochs of development. This robustness exercise leads us to conclude that *data*, conditional upon the preferred parameterization, as opposed to anomaly of the estimation procedure is primarily responsible for the patterns we estimate in the unobserved series.

A final caveat concerns the two $\rho \neq 0$ specifications reported in Technical Appendix C.1. Read naively, their WAIC values would appear to rival our preferred model.⁴¹ This ranking is spurious. When each model's fit is recomputed by Pareto-smoothed importance-sampling leave-one-out cross-validation (PSIS-LOO), both $\rho \neq 0$ models return a maximum Pareto shape estimate of $\hat{k} = 1.21$, with two decadal observations exceeding the $\hat{k} = 0.7$ reliability threshold in each case, whereas every $\rho = 0$ model in Table 1 stays well below it ($\max \hat{k} \leq 0.61$).⁴² Because $\hat{k} > 0.7$ implies the importance-sampling weights have infinite variance, both the PSIS-LOO and WAIC estimates for these two models are dominated by a handful of high-leverage decades and cannot be trusted (Vehtari, Gelman, and Gabry 2017). We therefore do not read the low nominal WAIC

⁴⁰Recall that we sample the negatives of these numbers, since we impose a strictly positive log-normal distribution on their priors.

⁴¹Indeed, the $\rho \in (0, 1)$, $\nu, \kappa < 0$ model posts a WAIC of 48.76, below the preferred model's 121.32.

⁴²Leave-one-out cross-validation (LOO) scores a model by its expected accuracy in predicting each observation from a posterior refit to the remaining data; PSIS-LOO approximates this cheaply by importance-reweighting the full-sample posterior, and $\hat{k}$ is the estimated tail-shape parameter of a generalized Pareto distribution fit to those importance weights. Values $\hat{k} > 0.7$ signal that the weights are too heavy-tailed for the approximation and to be reliable (Vehtari, Gelman, and Gabry 2017).

of the $\rho \neq 0$ models as evidence against our preferred specification. Finally, fitness problems with the $\rho \neq 0$ models are perhaps most obvious when inspecting the predictive plots in Technical Appendix C.2: it is clear from Figure C.1e that the model vastly over fits Y_t when $\rho < 0$, and for $\rho \in (0, 1)$ fitness featured in Figure C.1d, while such over fitting is not as stark, it does show up earlier in the sample.

3.2 Estimated Factor-augmenting Productivities

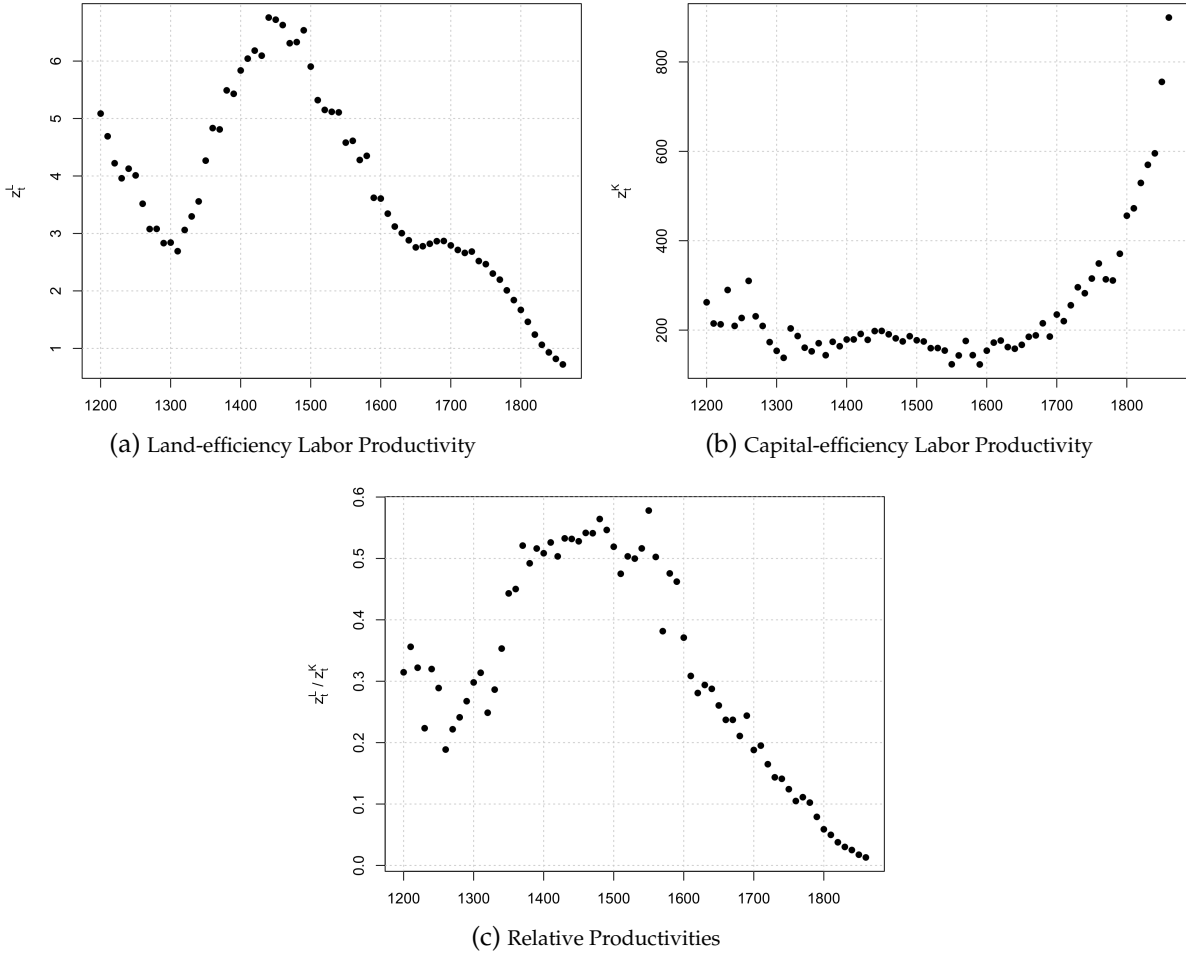


Figure 5: The figure contains estimated posterior-mean productivity levels in panels (a) and (b) and relative land/non-landed capital productivities in panel (c) for the preferred model with $\rho = 0$ and $\nu, \kappa < 0$.

We can back out posterior-mean productivity levels by passing the posterior means of the structural elasticities — $\hat{\nu}$, $\hat{\kappa}$, $\hat{\omega}_L$, $\hat{\omega}_K$, and $\hat{\alpha}$ — along with data in each period through (28) and (29) to get estimates for the time series $\{\hat{z}_t^L, \hat{z}_t^K\}_{t=1200}^{1860}$. Figure 5 presents the productivity levels in panels (a) and (b) and relative productivities, $\hat{z}_t^L / \hat{z}_t^K$, in panel (c) for the preferred model.⁴³

⁴³Productivity plots for the other models are included in Technical Appendix C.4.

Posterior means for each decade are black dots.

From our estimates it is clear that at the onset of modern growth capital-efficiency labor productivity, z_t^K , began to rise in the late 1600s, consistent with the timing of early industrialization. Further, land-efficiency labor productivity, z_t^L , was declining at the onset of industrialization. What is a bit surprising, though, is that the decline began about 200 years before the capital-efficiency labor productivity began to rise. The decline in z_t^L is responsible for driving down z_t^L/z_t^K through approximately 1650, when the decline stabilized at the same time that z_t^K began to grow. Both declining z_t^L and rising z_t^K together force down z_t^L/z_t^K as industrialization deepens after 1730.

The quantitative results thus conform to our theoretical findings. We find that rising z_t^K happens as growth in y_t begins to take off, while z_t^L continues to fall as the input-share reallocation occurs in the mid to late 1700s (see Figure 2).⁴⁴ Further, we can replicate modern growth preceding factor reallocation by approximately 50 to 75 years: since z_t^L and z_t^K estimates are structural errors they ensure that the model *exactly predicts* the income shares from Figure 2. Notice that z_t^K begins to rise in the mid-1600s but the income-share reallocation does not take place until the mid-1700s. Thus, rising z_t^K precedes factor-share reallocation and helps ensure (along with the structural elasticity estimates) that y_t starts to grow *before* such reallocation takes place.⁴⁵

Finally, in our multi-sector re-framing in Technical Appendix D, we decompose variation in z_t^j , $\forall j \in \{L, K\}$, as resulting from variation in the agriculture share of labor and an underlying, unobserved productivity term, A_t^j , that is independent of forces driving labor reallocation. Our decomposition relies on an admittedly coarse assumption that the land-intensive sector is primarily “agriculture.” Our findings, though, are null as to the degree to which reallocative forces may be biasing our interpretations regarding changes in z_t ’s: we find that even when accounting for sectoral labor reallocation, variation in the measured productivities is mostly driven by the underlying unobserved residual, A_t^j . The contribution is strongest for the capital series, so that we can strongly affirm that rising z_t^K at the onset of industrialization is *not* a simple by-product of labor reallocation but reflects actual improvements to capital-utilization efficiency.

3.3 Estimated Elasticities of Substitution

In this section we analyze the posterior average elasticities of substitution using the estimated parameters from our preferred model, where $\rho = 0$ and $\nu, \kappa < 0$.⁴⁶ Specifically, we pass period-by-period posterior mean estimates for $\hat{z}_t$ plus posterior average structural parameters and data through the parameterized elasticity expressions associated with Proposition 1. We derive the fully parameterized elasticities of substitution in Appendix A.4, where we replace $\tilde{r}(r^x, r^y)$ with $\Upsilon^{x,y}(\vartheta_t)$. This gives us three variable elasticities of substitution (land/labor, capital/labor, and land/capital) for each decade in our sample (1200-1860). We plot the time series of these estimates

⁴⁴Recall, in Section 2.3.4 we discussed how both rising z_t^K and declining z_t^L would be necessary in order to reconcile rising θ_t^K and declining θ_t^L in the 1700s when $\rho = 0$.

⁴⁵As mentioned in Section 3.1, all of the first- and second-order productivity patterns hold regardless of how we vary the underlying primitive elasticities — see Technical Appendix C.6.1.

⁴⁶Elasticities of substitution under alternative specifications are plotted in Technical Appendix C.5.

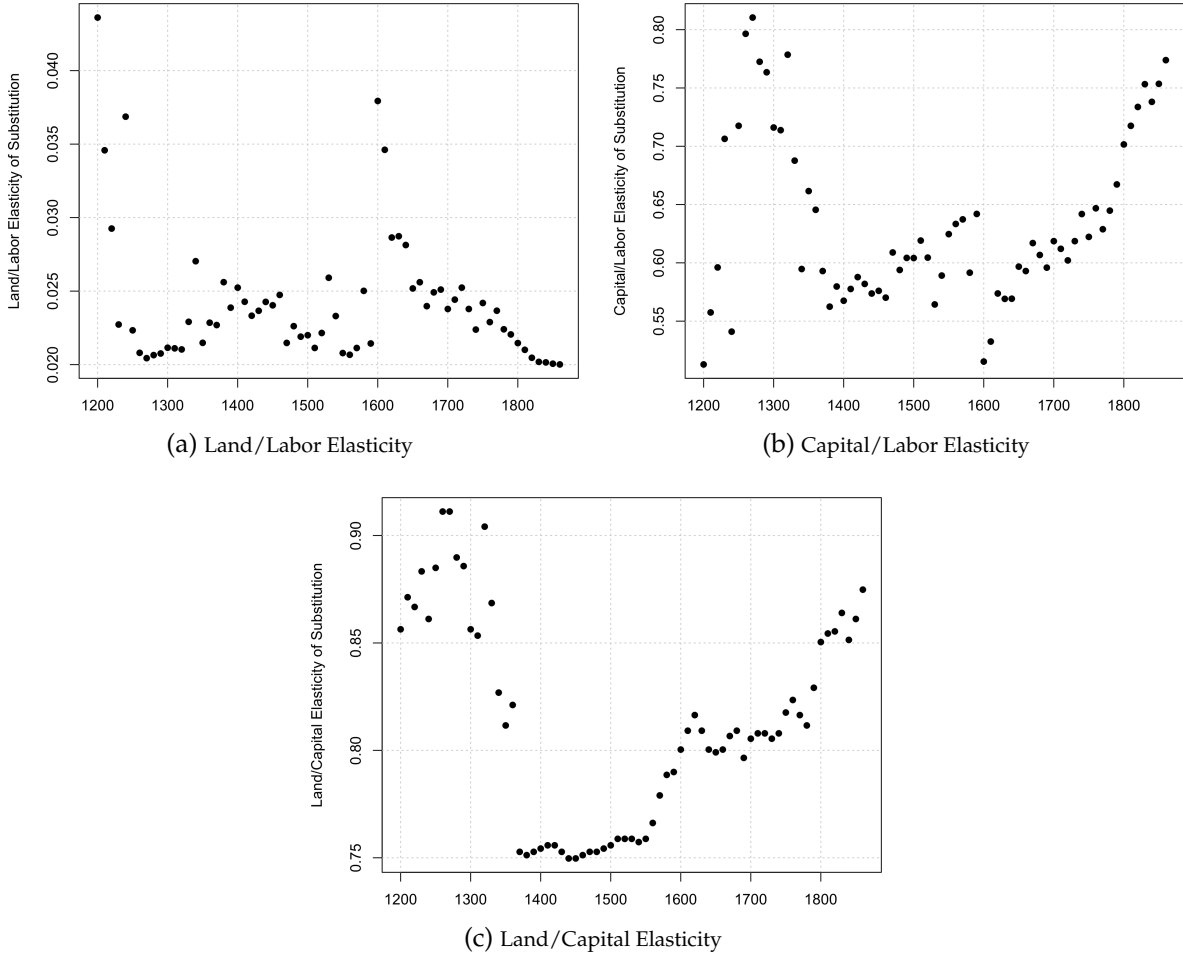


Figure 6: This figure contains estimated posterior-mean substitution elasticities for factor inputs from the preferred model with $\rho = 0$ and $\nu, \kappa < 0$. We compute these elasticities by passing both posterior-mean parameter and productivity estimates, as well as data, through the expressions derived in Appendix A.4.

in Figure 6.

The elasticity values are consistent with what we would expect given the factor-share patterns observed in Figure 2. The model generates a capital/labor elasticity of substitution that is < 1 which is consistent with estimates for modern economic growth from Juselius (2008), Raurich, Sala, and Sorolla (2012), and Knoblach, Roessler, and Zwerschke (2019). The land/capital and land/labor elasticities are also < 1 ; thus, all inputs are gross complements in all decades of observation.

Elasticities move around even when factor shares do not, because their variation over time reflects variation in factor-input *levels* around variation in output *levels*. This can be seen by comparing Figure 2 (factor shares) with Figure 6 (elasticities of substitution). In fact, elasticity trend changes seem to *precede* factor-share reallocation. For example, in Figure 2 the capital share falls during the Black Death, and capital/labor become more complementary in the 1300s. The Black

Death itself then coincides with a decrease in the capital/labor substitution elasticity, though the decrease begins some 50 years or so before the first major plague in 1348. The capital/labor substitution elasticity then does not significantly rise again until well into the Industrial Revolution (after 1750). Thus, capital and labor were relatively more complementary in the period of Malthusian stagnation from 1400-1650 than before and maintained the same level of complementarities for the first 100 years or so of the Industrial Revolution (1650-1750). Industrialization is then associated with capital and labor becoming more substitutable.

4 Conclusion

We have built an aggregate model that can reconcile key features of economies in *any* stage of development. First, we have built a model that embeds *both* Malthusian and Solow dynamics, where that which dominates depends on the structure of economic activity in the economy (i.e., the production function elasticities). Second, we showed how our model can fit factor income shares given data for land, labor, and capital. Further, we demonstrated that the model is robust and can accurately predict the evolution of GDP per capita even when the estimator is built simply around targeting changing income shares. Third, we showed that elasticities of substitution are endogenously variable despite the nested CES form the model takes.

Taken together, our estimates provide a quantitative anatomy of the transition to modern growth in England + Wales. The labor efficiency of using land began to decline around the mid-1400s, roughly two centuries before the labor efficiency of using capital began to rise in the mid-to-late 1600s. Rising capital-utilization efficiency in turn preceded the reallocation of factor income shares by 50 to 75 years, and the emergence of fully exponential growth by more than a century. These movements are not artifacts of labor reallocation: variation in capital-utilization efficiency is almost entirely attributable to sector-level productivity improvement rather than to the shift of labor out of agriculture (the land-intensive sector), and the pre-1650 decline in land-utilization efficiency cannot be a reallocation phenomenon at all, since the agricultural employment share was essentially constant before 1650. Meanwhile, all three factors remain gross complements throughout the sample, with capital and labor becoming more substitutable only once industrialization was well underway.

These measurements do not identify *which* institutional or technological shock initiated the capital-sector take-off, but they do discipline the set of admissible explanations. Any successful theory of English industrialization must generate a gradual rise in the efficiency of *using* capital beginning in the late 1600s, against the backdrop of a land-utilization efficiency that had already been falling for two centuries. The timing and sectoral incidence we measure are consistent with accounts emphasizing the seventeenth century's institutional transformations: strengthened property rights following the Glorious Revolution (North and Weingast 1989; Acemoglu, Johnson, and Robinson 2005a; Acemoglu, Johnson, and Robinson 2005b), post-1688 financial and administrative reform (Hodgson 2017), knowledge dissemination through England's guild and appren-

ticeship systems (Mokyr 2002), and the emergence of capital-market institutions (e.g., Lloyd’s of London, the London Stock Exchange) out of the commercial sociability of the early coffeehouse era (Jamieson 2001; Withington 2020). Each of these stories operates on precisely the margin our estimates single out, but our measurements cannot adjudicate among them. Instead, what we supply is the quantitative signature (i.e., magnitude and timing of productivity changes) that any candidate explanation must match. Symmetrically, explanations centered on the land sector (e.g., enclosure read either as agricultural modernization or as a gradual withdrawal of customary access to the commons),⁴⁷ bear on the long prior decline in z_t^L but cannot by themselves produce an industrial take-off. Indeed, under our preferred specification a pure capital-side productivity shock leaves the land share of income unchanged (Lemma 4), so reproducing the observed income-share reallocation of the 1700s requires movement in *both* utilization efficiencies. Such mono-causal accounts of the Industrial Revolution are thus inconsistent with our estimates, which instead lend partial credence to multiple theories.

More broadly, nothing about the framework is specific to historical England. Because three factor income shares admit exactly two time-varying residuals, the technology can be estimated for any economy for which factor share data are available. Our measurement exercise also carries a caution for that broader agenda: folding land into the capital stock is not innocuous, since omitting the fixed factor biases Solow-style TFP measurement downward (Section 2.5). In this manner, we have contributed a generalized theory of economic development with capital to the economic growth literature, and a corresponding measurement apparatus that could, with the right data, be applied to analyze *any* economy, including those still in earlier, land-intensive stages of development today.

A Mathematical Appendix

A.1 Proofs

The proofs in this Appendix are for those statements where such proofs are not included in the main text. Also included are additional corollaries required to complete the proofs.

Corollary A.1. *If $g_{z^L} = 1$ and $g_N > 1$, then $g_{\mathcal{L}} < 1$ for $\nu < 0$.*

Proof. Using the per-capita expression for $\mathcal{L}$ note that

$$\tilde{g}_{\mathcal{L}} = \frac{1}{\nu} \left[\ln \left(\omega_L \bar{L}^{\nu} (g_N N)^{-\nu} + (1 - \omega_L) (z^L)^{\nu} \right) - \ln \left(\omega_L \bar{L}^{\nu} N^{-\nu} + (1 - \omega_L) (z^L)^{\nu} \right) \right] < 0. \quad (\text{A1})$$

QED

⁴⁷See, e.g., Wordie (1983), Allen (1992), Ostrom (1990), Neeson (1993), and Overton (1996).

Proof of Lemma 1. If $\tilde{g}_N > 0$ then $\tilde{g}_{\theta^L} > 0$ as long as $-\nu\tilde{g}_N > \nu\tilde{g}_L$ by (22). Since $\nu < 0$ by assumption, this implies $\tilde{g}_N > -\tilde{g}_L = |\tilde{g}_L|$ by Corollary A.1. Thus,

$$|\nu|\tilde{g}_N > \left[\ln \left(\omega_L \bar{L}^\nu (g_N N)^{-\nu} + (1 - \omega_L)(z^L)^\nu \right) - \ln \left(\omega_L \bar{L}^\nu N^{-\nu} + (1 - \omega_L)(z^L)^\nu \right) \right], \quad (\text{A2})$$

which, after some rearrangement implies $g_N^{|\nu|} > 1$. This is true as long as $g_N > 1$, thus proving statement (i). The assumption that GDP per capita is not growing and the assumption that $k \propto y$ in steady state ensures that $\tilde{g}_K = 0$ so $\tilde{g}_{\theta^K} = \kappa\alpha\tilde{g}_L$ from (23). By Corollary A.1 $\tilde{g}_L < 0$, thus proving statement (ii). For statement (iii) note that if both θ^L and θ^K are rising, θ^N must fall. θ^N rises when $\tilde{g}_{\theta^L} < \tilde{g}_{\theta^K}$ and $\kappa > 0$ implying

$$|\nu|(\tilde{g}_N + \tilde{g}_L) < \kappa\alpha|\tilde{g}_L| \quad \Rightarrow \quad \tilde{g}_N < \left(\frac{\kappa\alpha}{|\nu|} + 1 \right) |\tilde{g}_L|,$$

which holds when $\kappa\alpha \gg |\nu|$ and/or $|\tilde{g}_L| \gg 0$. **QED**

Corollary A.2. If $g_{z^K} = 1$ and $g_N > 1$ but $g_Y = 1$, then $g_K < 1$ always.

Proof. Using the per-capita expression for K note that

$$\tilde{g}_K = \frac{1}{\kappa} \left[\ln \left(\omega_K \left(\frac{sY}{\delta} \right)^\kappa (g_N N)^{-\kappa} + (1 - \omega_K)(z^K)^\kappa \right) - \ln \left(\omega_K \left(\frac{sY}{\delta} \right)^\kappa N^{-\kappa} + (1 - \omega_K)(z^K)^\kappa \right) \right]. \quad (\text{A3})$$

In (A3) the term in brackets is > 0 when $\kappa < 0$ and < 0 when $\kappa \in (0, 1)$ implying $\tilde{g}_K < 0$ always. **QED**

Proof of Lemma 2. Even fixing total GDP, Corollary A.1 applies, and $\tilde{g}_{\theta^L} > 0$. When Y is fixed but $g_N > 1$, $\tilde{g}_K < 0$ always by Corollary A.2. When $\kappa < 0$, $\tilde{g}_{\theta^K} < 0$ if $|\tilde{g}_L| < |\tilde{g}_K|$ and vice-versa when $\kappa \in (0, 1)$. Suppose (WOLOG) the land and capital nests are approximately the same scale so that $\mathcal{L}(\ell; z^L)^\nu \approx \mathcal{K}(k; z^K)^\kappa$, then $\tilde{g}_L - \tilde{g}_K > 0$ if $|\kappa| < |\nu|$ providing case (ii). For case (iii) note that θ^N only rises if θ^K falls faster than θ^L rises. This happens when $|\kappa\alpha(\tilde{g}_L - \tilde{g}_K)| > |\nu|(\tilde{g}_N + \tilde{g}_L)$ or (after rearrangement) when $(\kappa\alpha - |\nu|)\tilde{g}_L > |\nu|\tilde{g}_N + \kappa\alpha\tilde{g}_K$. Since $|\nu|\tilde{g}_N > 0$ by assumption, this number must be sufficiently small enough (relative to values of ν , κ , and α) for θ^N to rise. **QED**

Corollary A.3. If $g_N = 1$ and $g_{z^L} > 1$, then $g_L > 1$ for $\nu < 0$.

Proof. Normalize $z^L = 1$, WOLOG. Using the per-capita expression for L note that

$$\tilde{g}_L = \frac{1}{\nu} \left[\ln \left(\omega_L \bar{L}^\nu + (1 - \omega_L)(g_{z^L})^\nu \right) - \ln \left(\omega_L \bar{L}^\nu + (1 - \omega_L) \right) \right]. \quad (\text{A4})$$

Since $\nu < 0$ and $g_{z^L} > 1$, $(g_{z^L})^\nu < 1$ and the big term in brackets in (A4) is negative, implying that $\tilde{g}_L > 0$. QED

Proof of Lemma 3. By assumption $\tilde{g}_N = 0$ so that (22) becomes $\tilde{g}_{\theta^L} = |\nu|\tilde{g}_L > 0$ by Corollary A.3 and $\nu < 0$, proving (i). For (ii) note that $\tilde{g}_K = 0$ since K is independent of z^L , thus making (23), $\tilde{g}_{\theta^K} = \kappa\alpha\tilde{g}_L$, which is < 0 when $\kappa < 0$ and > 0 when $\kappa \in (0, 1)$. For statement (iii) note that the total change in θ^N is $d\theta^N = -d\theta^L - d\theta^K$ which is $d\theta^N = -\theta^L\tilde{g}_{\theta^L} - \theta^K\tilde{g}_{\theta^K}$. Thus, $\tilde{g}_{\theta^N} > 0$ if and only if $\theta^K|\kappa|\alpha > \theta^L|\nu|$. QED

Corollary A.4. *If $g_N = 1$, $g_{z^K} > 1$ and $g_K > 1$ then $g_y > 1$ and $g_K > 1$ for all $\kappa \neq 0$.*

Proof. Normalize $z^K = y = N = 1$, WOLOG. Using the per-capita expression for K note that

$$\tilde{g}_K = \frac{1}{\kappa} \ln \left(\omega_K g_y^\kappa + (1 - \omega_K) g_{z^K}^\kappa \right). \quad (\text{A5})$$

When $\kappa < 0$, $g_x^\kappa < 1$, $\forall x \in \{y, z^K\}$, so that the log term in (A5) is < 0 and $\tilde{g}_K > 0$, but when $\kappa \in (0, 1)$, both $1/\kappa$ and the log term in (A5) are > 0 so $\tilde{g}_K > 0$. QED

Proof of Lemma 4. Clearly, L is independent of z^K and $g_N = 1$, so $g_{\theta^L} = 1$. (23) becomes $\tilde{g}_{\theta^K} = -\kappa\alpha\tilde{g}_K$. By Corollary A.4 $\tilde{g}_K > 0$ always, giving us (ii), and (iii) immediately follows. QED

Proof of Proposition 3. Along a balanced growth path s and $\hat{\delta}$ are constant, so taking logged first differences of (27) gives

$$\tilde{g}_y = \frac{\tilde{g}_A - \alpha\omega_L\tilde{g}_N}{1 - (1 - \alpha)\omega_K}.$$

The denominator is positive because $(1 - \alpha)\omega_K < 1$ for interior α and ω_K . Setting $\tilde{g}_A = \tilde{g}_N = \tilde{g} > 0$ yields $\tilde{g}_y = \tilde{g}(1 - \alpha\omega_L)/(1 - (1 - \alpha)\omega_K) > 0$, since $\alpha\omega_L < 1$, strictly, when α and ω_L are interior. More generally $\tilde{g}_y > 0$ whenever $\tilde{g}_A > \alpha\omega_L\tilde{g}_N$, so A need not grow as fast as N . QED

A.2 Comparing Factor Shares in Hansen and Prescott (2002)

Using the notation of Hansen and Prescott (2002), let aggregate output be $Y_t = Y_{Mt} + Y_{St}$, where M is the Malthusian technology and S is the Solow technology. Let $\alpha_{Kt} = K_{Mt}/K_t$ be the share of capital used in Malthusian production, and the same for α_{Nt} as the share of labor. Malthusian output is $Y_{Mt} = A_{Mt}(\alpha_{Kt}K_t)^\phi(\alpha_{Nt}N_t)^\mu \bar{L}^{1-\phi-\mu}$ and Solow output is $Y_{St} = A_{St}((1 - \alpha_{Kt})K_t)^\theta((1 - \alpha_{Nt})N_t)^{1-\theta}$. In their model factor shares are constant as long as only one of the technologies operates and thus only vary along the transition path. This can be seen by noting that the marginal product of capital is $MP_K = \phi \frac{Y_{Mt}}{K_t} + \theta \frac{Y_{St}}{K_t} = r_t^K$. Letting α_{Yt} be the share of output generated by the Malthusian technology, we get $\theta_t^K = \phi\alpha_{Yt} + \theta(1 - \alpha_{Yt})$ which varies in t only in so much as $\alpha_{Yt} \in (0, 1)$ and also varies in t . In their model this only happens during the transitional period, but not before or after.

A.3 Derivation of Production Function Estimating Equations

The derivation operates on the full CES factor-share conditions in (18) and (19). Note that these objects are functions of z_t^L and z_t^K , each embedded in $\mathcal{L}_t$ and $\mathcal{K}_t$, respectively. We solve for z_t^L as follows:

$$(\theta_t^L)^{\frac{\nu}{\rho-\nu}} (\alpha \omega_L)^{\frac{\nu}{\nu-\rho}} \ell_t^{\frac{\nu^2}{\nu-\rho}} y_t^{\frac{\rho\nu}{\rho-\nu}} - \omega_L \ell_t^\nu = (1 - \omega_L)(z_t^L)^\nu \quad (\text{A6})$$

$$\Leftrightarrow \ell_t^\nu \left[(\theta_t^L)^{\frac{\nu}{\rho-\nu}} (\alpha \omega_L)^{\frac{\nu}{\nu-\rho}} \ell_t^{\frac{\nu^2}{\nu-\rho}} y_t^{\frac{\rho\nu}{\rho-\nu}} - \omega_L \right] = (1 - \omega_L)(z_t^L)^\nu \quad (\text{A7})$$

$$\Leftrightarrow \ln z_t^L = \ln \ell_t + \frac{1}{\nu} \ln \Gamma^L(y_t, \ell_t, \theta_t^L) - \frac{1}{\nu} \ln(1 - \omega_L) \quad (\text{A8})$$

$$\text{where } \Gamma^L(y_t, \ell_t, \theta_t^L) = \left[(\theta_t^L)^{\frac{\nu}{\rho-\nu}} (\alpha \omega_L)^{\frac{\nu}{\nu-\rho}} \ell_t^{\frac{\nu^2}{\nu-\rho}} y_t^{\frac{\rho\nu}{\rho-\nu}} - \omega_L \right]. \quad (\text{A9})$$

When $\rho = 0$ (i.e., Cobb-Douglas upper nest) the function Γ^L can be written strictly as a function of income shares and land per capita: $\Gamma^L(\ell_t, \theta_t^L) = \frac{\alpha \omega_L}{\theta_t^L} - \omega_L$. By symmetry the same sequence of algebraic operations can be applied to (19) to solve for z_t^K :

$$\ln z_t^K = \ln k_t + \frac{1}{\kappa} \ln \Gamma^K(y_t, k_t, \theta_t^K) - \frac{1}{\kappa} \ln(1 - \omega_K) \quad (\text{A10})$$

$$\text{where } \Gamma^K(y_t, k_t, \theta_t^K) = \left[(\theta_t^K)^{\frac{\kappa}{\rho-\kappa}} ((1 - \alpha) \omega_K)^{\frac{\kappa}{\kappa-\rho}} k_t^{\frac{\rho\kappa}{\kappa-\rho}} y_t^{\frac{\rho\kappa}{\rho-\kappa}} - \omega_K \right]. \quad (\text{A11})$$

When $\rho = 0$, $\Gamma^K(k_t, \theta_t^K) = \frac{(1-\alpha)\omega_K}{\theta_t^K} - \omega_K$.

A.4 Parameterized Elasticities of Substitution

To get the parameterized elasticities of substitution corresponding to the generalized expression in Proposition 1, we differentiate (15), (16), and (17) each in ϑ_t and multiply by the ratio $\Upsilon^{x,y}(\vartheta_t)/\vartheta_t$, where we replace the input-price ratio with the MRTS. The derivatives are

$$\begin{aligned} \frac{\partial \Upsilon^{L,N}}{\partial \vartheta_t} = (1 - \nu) \vartheta_t^{-\nu} & \left[\frac{1 - \omega_L}{\omega_L} (z_t^L)^\nu + \frac{(1 - \alpha)(1 - \omega_K)}{\alpha \omega_L} (z_t^K)^\kappa \frac{\mathcal{K}_t(K_t, N_t)^{\rho-\kappa}}{\mathcal{L}_t(\vartheta_t)^{\rho-\nu}} N_t^{\kappa-\rho} \right] \\ & + (\nu - \rho) \left(\frac{(1 - \alpha)(1 - \omega_K)}{\alpha} \right) (z_t^K)^\kappa \mathcal{K}_t(K_t, N_t)^{\rho-\kappa} \mathcal{L}_t(\vartheta_t)^{-\rho} N_t^{\kappa-\rho} \end{aligned} \quad (\text{A12})$$

$$\begin{aligned} \frac{\partial \Upsilon^{K,N}}{\partial \vartheta_t} = (1 - \kappa) \vartheta_t^{-\kappa} & \left[\frac{1 - \omega_K}{\omega_K} (z_t^K)^\kappa + \frac{\alpha(1 - \omega_L)}{(1 - \alpha)\omega_K} (z_t^L)^\nu \frac{\mathcal{L}_t(L_t, N_t)^{\rho-\nu}}{\mathcal{K}_t(\vartheta_t)^{\rho-\kappa}} N_t^{\nu-\rho} \right] \\ & + (\kappa - \rho) \left(\frac{\alpha(1 - \omega_L)}{(1 - \alpha)} \right) (z_t^L)^\nu \mathcal{L}_t(L_t, N_t)^{\rho-\nu} \mathcal{K}_t(\vartheta_t)^{-\rho} N_t^{\nu-\rho} \end{aligned} \quad (\text{A13})$$

$$\begin{aligned}
\frac{\partial \Upsilon^{K,L}}{\partial \vartheta_t} = & (1 - \kappa) \vartheta_t^{-\kappa} \left(\frac{\alpha \omega_L}{(1 - \alpha) \omega_K} \right) \frac{\mathcal{L}_t(L_t, N_t)^{\rho - \nu}}{\mathcal{K}_t(\vartheta_t L_t, N_t)^{\rho - \kappa}} L_t^{\nu - \kappa} \\
& + (\kappa - \rho) \left(\frac{\alpha \omega_L}{(1 - \alpha)} \right) \mathcal{L}_t(L_t, N_t)^{\rho - \nu} \mathcal{K}_t(\vartheta_t L_t, N_t)^{-\rho} L_t^\nu.
\end{aligned} \tag{A14}$$

To get the elasticities we take the inverse of the objects in (A12), (A13), and (A14) and multiply them by $\Upsilon^{x,y}/\vartheta_t$.

Technical Appendices

(For Online Publication Only)

A Theory of Long-run Development with Capital

Nick Pretnar

*Laboratory for Aggregate Economics and Finance, UCSB
University of Austin Texas (UATX)*

Ákos Valentinyi

University of Manchester

B Growth Properties when $\rho \neq 0$

The growth decompositions in Section 2.3 of the main text are derived under the Cobb-Douglas restriction, $\rho = 0$. That restriction is not merely algebraic convenience: it *exactly* decouples the two nests in the income shares and makes aggregate output growth an input-share-weighted sum with constant weights. In this appendix we relax it and characterize how the four growth regimes discussed in Section 2.3 (i.e., stagnation, pure Malthus, Malthusian productivity growth, and Solow growth) change when $\rho \neq 0$. We treat separately the two specific sign configurations that are used on quantitative estimations featured in Appendix C.1:

Case A ($\rho \in (0, 1)$, $\nu, \kappa < 0$) — the land and capital nests are *substitutes* in the upper tier, while land/labor and capital/labor are complements within each lower nest.

Case B ($\rho < 0$, $\nu, \kappa \in (0, 1)$) — the two nests are *complements* in the upper tier, while land/labor and capital/labor are substitutes within each nest.

Throughout, we retain the constant-investment-rate assumption of the main text and, as there, drop time subscripts.

Recall the per-capita technology and its two nests,

$$y = \left(\alpha \mathcal{L}^\rho + (1 - \alpha) \mathcal{K}^\rho \right)^{\frac{1}{\rho}} \quad (\text{B.1})$$

$$\mathcal{L} = \left(\omega_L \ell^\nu + (1 - \omega_L) (z^L)^\nu \right)^{\frac{1}{\nu}}, \quad \mathcal{K} = \left(\omega_K k^\kappa + (1 - \omega_K) (z^K)^\kappa \right)^{\frac{1}{\kappa}}, \quad (\text{B.2})$$

with $\ell = \bar{L}/N$ and $k = K/N$, and the land, capital, and labor income shares,

$$\theta^L = \alpha \omega_L y^{-\rho} \mathcal{L}^{\rho-\nu} \ell^\nu, \quad \theta^K = (1 - \alpha) \omega_K y^{-\rho} \mathcal{K}^{\rho-\kappa} k^\kappa, \quad \theta^N = 1 - \theta^L - \theta^K. \quad (\text{B.3})$$

Define the upper-tier value shares

$$s_{\mathcal{L}} \equiv \frac{\alpha \mathcal{L}^\rho}{\alpha \mathcal{L}^\rho + (1 - \alpha) \mathcal{K}^\rho} = \frac{\alpha \mathcal{L}^\rho}{y^\rho}, \quad s_{\mathcal{K}} \equiv 1 - s_{\mathcal{L}} = \frac{(1 - \alpha) \mathcal{K}^\rho}{y^\rho}, \quad (\text{B.4})$$

and the within-nest input weights

$$\phi_L \equiv \frac{\omega_L \ell^\nu}{\omega_L \ell^\nu + (1 - \omega_L) (z^L)^\nu}, \quad \phi_K \equiv \frac{\omega_K k^\kappa}{\omega_K k^\kappa + (1 - \omega_K) (z^K)^\kappa}. \quad (\text{B.5})$$

For interior weights $0 < \alpha, \omega_L, \omega_K < 1$ and strictly positive inputs, all four objects $s_{\mathcal{L}}, s_{\mathcal{K}}, \phi_L, \phi_K$ lie in $(0, 1)$, for any finite ρ, ν, κ . Several accounting identities that we use repeatedly are

$$\theta^L = s_{\mathcal{L}} \phi_L, \quad \theta^K = s_{\mathcal{K}} \phi_K, \quad \theta^N = s_{\mathcal{L}} (1 - \phi_L) + s_{\mathcal{K}} (1 - \phi_K), \quad (\text{B.6})$$

which follow by substituting (B.4) and (B.5) into (B.3); they express the observed factor shares as

the product of an upper-tier value share and a within-nest weight, and confirm $\theta^L + \theta^K + \theta^N = 1$. As $\rho \rightarrow 0$, so do $s_{\mathcal{L}} \rightarrow \alpha$ and $s_{\mathcal{K}} \rightarrow 1 - \alpha$, and every expression below collapses to its Cobb-Douglas counterpart.

We first record how output and the nests grow. We adopt the instantaneous log-growth convention $\tilde{g}_x \equiv \dot{x}/x$ where $\dot{x} = dx/dt$ is the derivative of x with respect to time, which holds up to a first-order Taylor approximation.^{B1} The following lemmas hold for any ρ, ν, κ and underpin all subsequent results.

Lemma B.1. (*Output Growth*) *Along the path, aggregate output per capita grows as the value-share-weighted average of the two nests,*

$$\tilde{g}_y = s_{\mathcal{L}} \tilde{g}_{\mathcal{L}} + s_{\mathcal{K}} \tilde{g}_{\mathcal{K}}. \quad (\text{B.8})$$

Proof. Raise (B.1) to the power ρ to get $y^\rho = \alpha \mathcal{L}^\rho + (1 - \alpha) \mathcal{K}^\rho$. Differentiate with respect to time:

$$\rho y^{\rho-1} \dot{y} = \alpha \rho \mathcal{L}^{\rho-1} \dot{\mathcal{L}} + (1 - \alpha) \rho \mathcal{K}^{\rho-1} \dot{\mathcal{K}}.$$

Divide both sides by ρy^ρ (nonzero since $\rho \neq 0$ and $y > 0$):

$$\frac{\dot{y}}{y} = \frac{\alpha \mathcal{L}^\rho}{y^\rho} \frac{\dot{\mathcal{L}}}{\mathcal{L}} + \frac{(1 - \alpha) \mathcal{K}^\rho}{y^\rho} \frac{\dot{\mathcal{K}}}{\mathcal{K}}.$$

By (B.4) the coefficients are $s_{\mathcal{L}}$ and $s_{\mathcal{K}}$, giving (B.8). As $\rho \rightarrow 0$ the weights tend to α and $1 - \alpha$, recovering $\tilde{g}_y = \alpha \tilde{g}_{\mathcal{L}} + (1 - \alpha) \tilde{g}_{\mathcal{K}}$ from Section 2.3. **QED**

Lemma B.2. (*Inner-nest Growth*) *The nests grow according to*

$$\tilde{g}_{\mathcal{L}} = -\phi_L \tilde{g}_N + (1 - \phi_L) \tilde{g}_{z^L}, \quad \tilde{g}_{\mathcal{K}} = \phi_K \tilde{g}_k + (1 - \phi_K) \tilde{g}_{z^K}, \quad (\text{B.9})$$

where $\tilde{g}_\ell = -\tilde{g}_N$ because $\ell = \bar{L}/N$ with $\bar{L}$ fixed.

Proof. Raise the land nest in (B.2) to the power ν , $\mathcal{L}^\nu = \omega_L \ell^\nu + (1 - \omega_L)(z^L)^\nu$. Differentiate and divide by $\nu \mathcal{L}^\nu$:

$$\frac{\dot{\mathcal{L}}}{\mathcal{L}} = \frac{\omega_L \ell^\nu}{\mathcal{L}^\nu} \frac{\dot{\ell}}{\ell} + \frac{(1 - \omega_L)(z^L)^\nu}{\mathcal{L}^\nu} \frac{\dot{z}^L}{z^L} = \phi_L \tilde{g}_\ell + (1 - \phi_L) \tilde{g}_{z^L},$$

using (B.5) and $\mathcal{L}^\nu = \omega_L \ell^\nu + (1 - \omega_L)(z^L)^\nu$. Substituting $\tilde{g}_\ell = -\tilde{g}_N$ gives the first identity. The capital nest is identical with $(\kappa, \omega_K, k, z^K)$ in place of $(\nu, \omega_L, \ell, z^L)$, and k growing rather than shrinking with N . **QED**

^{B1}This instantaneous representation of the log gross growth rate holds for small enough values of the net growth rate, g_x :

$$\tilde{g}_x = g_x + \frac{1}{2} g_x^2 + \mathcal{O}(g_x^3). \quad (\text{B.7})$$

Using both Lemmas B.1 and B.2 yields the central result of this appendix: the income-share growth rates equal their $\rho = 0$ counterparts plus an explicit cross-nest correction proportional to ρ .

Proposition B.1. *(Income-share Growth, General ρ) Under Assumption 1 the land and capital income shares grow according to*

$$\tilde{g}_{\theta^L} = \underbrace{-\nu(\tilde{g}_N + \tilde{g}_L)}_{\text{Cobb-Douglas term}} + \underbrace{\rho s_K(\tilde{g}_L - \tilde{g}_K)}_{\text{cross-nest correction}}, \quad (\text{B.10})$$

$$\tilde{g}_{\theta^K} = (\kappa - \rho) s_L(\tilde{g}_L - \tilde{g}_K), \quad (\text{B.11})$$

$$\tilde{g}_{\theta^N} = -\frac{1}{\theta^N} (\theta^L \tilde{g}_{\theta^L} + \theta^K \tilde{g}_{\theta^K}). \quad (\text{B.12})$$

Setting $\rho = 0$ (so $s_L = \alpha$, $s_K = 1 - \alpha$) recovers exactly the expressions $\tilde{g}_{\theta^L} = -\nu(\tilde{g}_N + \tilde{g}_L)$ and $\tilde{g}_{\theta^K} = \kappa\alpha(\tilde{g}_L - \tilde{g}_K)$ of Section 2.3.

Proof. Take logs of the land share in (B.3):

$$\ln \theta^L = \ln(\alpha \omega_L) - \rho \ln y + (\rho - \nu) \ln \mathcal{L} + \nu \ln \ell.$$

Differentiating in time and using $\tilde{g}_\ell = -\tilde{g}_N$,

$$\tilde{g}_{\theta^L} = -\rho \tilde{g}_y + (\rho - \nu) \tilde{g}_L - \nu \tilde{g}_N. \quad (\text{B.13})$$

Substitute $\tilde{g}_y = s_L \tilde{g}_L + s_K \tilde{g}_K$ from Lemma B.1 and collect the $\tilde{g}_L$ terms, using $-\rho s_L + \rho = \rho(1 - s_L) = \rho s_K$:

$$\begin{aligned} \tilde{g}_{\theta^L} &= -\rho(s_L \tilde{g}_L + s_K \tilde{g}_K) + (\rho - \nu) \tilde{g}_L - \nu \tilde{g}_N \\ &= (\rho s_K - \nu) \tilde{g}_L - \rho s_K \tilde{g}_K - \nu \tilde{g}_N \\ &= -\nu(\tilde{g}_N + \tilde{g}_L) + \rho s_K(\tilde{g}_L - \tilde{g}_K), \end{aligned}$$

which is (B.10). For the capital share, taking logs of (B.3),

$$\ln \theta^K = \ln((1 - \alpha) \omega_K) - \rho \ln y + (\rho - \kappa) \ln \mathcal{K} + \kappa \ln k,$$

so $\tilde{g}_{\theta^K} = -\rho \tilde{g}_y + (\rho - \kappa) \tilde{g}_K + \kappa \tilde{g}_k$. Impose Assumption 1 ($\tilde{g}_k = \tilde{g}_y$):

$$\tilde{g}_{\theta^K} = -\rho \tilde{g}_y + (\rho - \kappa) \tilde{g}_K + \kappa \tilde{g}_y = (\kappa - \rho) \tilde{g}_y + (\rho - \kappa) \tilde{g}_K = (\kappa - \rho)(\tilde{g}_y - \tilde{g}_K).$$

Finally, $\tilde{g}_y - \tilde{g}_K = s_L \tilde{g}_L + s_K \tilde{g}_K - \tilde{g}_K = s_L \tilde{g}_L - (1 - s_K) \tilde{g}_K = s_L(\tilde{g}_L - \tilde{g}_K)$, giving (B.11). Equation (B.12) is immediate from $\theta^N = 1 - \theta^L - \theta^K$, which gives $\dot{\theta}^N = -\dot{\theta}^L - \dot{\theta}^K$ and hence $\tilde{g}_{\theta^N} = \dot{\theta}^N / \theta^N = -(\theta^L \tilde{g}_{\theta^L} + \theta^K \tilde{g}_{\theta^K}) / \theta^N$, using $\dot{\theta}^x = \theta^x \tilde{g}_{\theta^x}$. QED

Proposition B.1 isolates exactly what $\rho = 0$ suppresses. First, the constant land weight α in (23) is replaced by the endogenous value share s_L , so the magnitude of factor redistribution depends on where the economy sits on its development path. Second, the parameter governing the capital share flips from κ to the *composite* elasticity $(\kappa - \rho)$ since $\rho \neq 0$: the sign thresholds of Section 2.3 move from $\kappa = 0$ to $\kappa = \rho$. Third, and most consequentially, a cross-nest term $\rho s_K(\tilde{g}_L - \tilde{g}_K)$ appears in the land share that disappears when $\rho = 0$. We now trace these three channels through the four regimes under the two different parametric cases A and B above. In each case we compute the pair of growth rates, $(\tilde{g}_L, \tilde{g}_K)$, from Lemma B.2 under the regime's defining restrictions and then substitute them into Proposition B.1. The regime definitions are those of Section 2.3: 1) stagnation ($\tilde{g}_N > 0, \tilde{g}_y = 0, \tilde{g}_{z^L} = \tilde{g}_{z^K} = 0$); 2) pure Malthus ($\tilde{g}_N > 0, \tilde{g}_y < 0, \tilde{g}_{z^L} = \tilde{g}_{z^K} = 0$); 3) Malthusian productivity growth ($\tilde{g}_{z^L} > 0, \tilde{g}_y = 0, \tilde{g}_N = \tilde{g}_{z^K} = 0$); and 4) Solow growth ($\tilde{g}_{z^K} > 0 \Rightarrow \tilde{g}_y > 0, \tilde{g}_N = \tilde{g}_{z^L} = 0$).

B.1 Case A: $\rho \in (0, 1), \nu, \kappa < 0$

Proposition B.2. (Factor Reallocation When Nests Are Substitutes) Let $\rho \in (0, 1)$ and $\nu, \kappa < 0$, and maintain Assumption 1. Then:

- i. **Stagnation** ($g_N > 1$ but $g_y = 1$ with $g_{z^L} = g_{z^K} = 1$): $\tilde{g}_{\theta^L} = [|\nu|(1 - \phi_L) - \rho s_K \phi_L] \tilde{g}_N$, so θ^L rises if and only if $|\nu|(1 - \phi_L) > \rho s_K \phi_L$ and falls otherwise; $\tilde{g}_{\theta^K} = (\rho - \kappa) s_L \phi_L \tilde{g}_N > 0$, so θ^K rises always.
- ii. **Pure Malthus** ($g_N > 1$ but $g_y < 1$ with $g_{z^L} = g_{z^K} = 1$): $\tilde{g}_{\theta^K} = (\rho - \kappa) s_L (\phi_K |\tilde{g}_y| - \phi_L \tilde{g}_N)$, so θ^K rises if and only if the output contraction is fast relative to population growth, $\phi_K |\tilde{g}_y| > \phi_L \tilde{g}_N$, and falls otherwise; θ^L rises whenever $|\nu|(1 - \phi_L) \tilde{g}_N + \rho s_K (\phi_K |\tilde{g}_y| - \phi_L \tilde{g}_N) > 0$, which holds in particular when θ^K rises.
- iii. **Malthusian Productivity Growth** ($g_{z^L} > 1$ but $g_y = 1$ with $g_N = g_{z^K} = 1$): $\tilde{g}_{\theta^L} = (\rho s_K + |\nu|(1 - \phi_L) \tilde{g}_{z^L}) > 0$ and $\tilde{g}_{\theta^K} = -(\rho + |\kappa|) s_L (1 - \phi_L) \tilde{g}_{z^L} < 0$, so θ^L rises and θ^K falls always.
- iv. **Solow Growth** ($g_{z^K} > 1$ leading to $g_y > 1$ with $g_N = g_{z^L} = 1$): $\tilde{g}_{\theta^L} = -\rho s_K \tilde{g}_K < 0$ and $\tilde{g}_{\theta^K} = (\rho - \kappa) s_L \tilde{g}_K > 0$: the land share falls and the capital share rises, with $\tilde{g}_K = (1 - \phi_K) \tilde{g}_{z^K} / (1 - s_K \phi_K) > 0$.

Proof. We use Lemma B.2 and $\tilde{g}_k = \tilde{g}_y$ (Assumption 1) throughout, and substitute into (B.10) and (B.11). Recall $s_L, s_K, \phi_L, \phi_K \in (0, 1)$; with $\nu, \kappa < 0$ we write $\nu = -|\nu|$, $\kappa = -|\kappa|$, and with $\rho \in (0, 1)$ we have $\rho - \kappa = \rho + |\kappa| > 0$ and $\rho s_K > 0$.

(i) *Stagnation.* Here $\tilde{g}_{z^L} = \tilde{g}_{z^K} = 0$ and $\tilde{g}_y = 0$, so $\tilde{g}_k = 0$. Lemma B.2 gives $\tilde{g}_L = -\phi_L \tilde{g}_N$ and $\tilde{g}_K = 0$. Then $\tilde{g}_N + \tilde{g}_L = (1 - \phi_L) \tilde{g}_N$ and $\tilde{g}_L - \tilde{g}_K = -\phi_L \tilde{g}_N$. Substituting into (B.10),

$$\tilde{g}_{\theta^L} = -\nu(1 - \phi_L) \tilde{g}_N + \rho s_K (-\phi_L \tilde{g}_N) = [|\nu|(1 - \phi_L) - \rho s_K \phi_L] \tilde{g}_N,$$

which is positive if and only if $|\nu|(1 - \phi_L) > \rho s_K \phi_L$. From (B.11), $\tilde{g}_{\theta^K} = (\kappa - \rho) s_L (-\phi_L \tilde{g}_N) = (\rho - \kappa) s_L \phi_L \tilde{g}_N > 0$ since $\rho - \kappa > 0$ and $\tilde{g}_N > 0$.

(ii) *Pure Malthus*. Again $\tilde{g}_{z^L} = \tilde{g}_{z^K} = 0$, but now $\tilde{g}_y < 0$ (imposed), so $\tilde{g}_k = \tilde{g}_y < 0$. Lemma B.2 gives $\tilde{g}_L = -\phi_L \tilde{g}_N < 0$ and $\tilde{g}_K = \phi_K \tilde{g}_y < 0$, so

$$\tilde{g}_L - \tilde{g}_K = -\phi_L \tilde{g}_N - \phi_K \tilde{g}_y = \phi_K |\tilde{g}_y| - \phi_L \tilde{g}_N,$$

using $\tilde{g}_y = -|\tilde{g}_y|$. From (B.11), $\tilde{g}_{\theta^K} = (\kappa - \rho)s_L(\tilde{g}_L - \tilde{g}_K) = (\rho - \kappa)s_L(\phi_K |\tilde{g}_y| - \phi_L \tilde{g}_N)$; since $\rho - \kappa > 0$ this is positive if and only if $\phi_K |\tilde{g}_y| > \phi_L \tilde{g}_N$. For the land share, (B.10) gives $\tilde{g}_{\theta^L} = |\nu|(1 - \phi_L)\tilde{g}_N + \rho s_K(\phi_K |\tilde{g}_y| - \phi_L \tilde{g}_N)$; the first term is positive, so the sum is positive whenever the bracket is nonnegative, i.e. whenever θ^K rises (a sufficient, not necessary, condition).

(iii) *Malthusian productivity growth*. Here $\tilde{g}_N = \tilde{g}_{z^K} = 0$ and $\tilde{g}_y = 0 \Rightarrow \tilde{g}_k = 0$, while $\tilde{g}_{z^L} > 0$. Lemma B.2 gives $\tilde{g}_L = (1 - \phi_L)\tilde{g}_{z^L} > 0$ and $\tilde{g}_K = 0$. Then (B.10) yields $\tilde{g}_{\theta^L} = -\nu \tilde{g}_L + \rho s_K \tilde{g}_L = (|\nu| + \rho s_K)(1 - \phi_L)\tilde{g}_{z^L} > 0$, and (B.11) yields $\tilde{g}_{\theta^K} = (\kappa - \rho)s_L \tilde{g}_L = -(|\kappa| + \rho)s_L(1 - \phi_L)\tilde{g}_{z^L} < 0$.

(iv) *Solow growth*. Here $\tilde{g}_N = \tilde{g}_{z^L} = 0$, so $\tilde{g}_L = 0$, while $\tilde{g}_{z^K} > 0$. Capital-nest and output growth are simultaneously determined: Lemma B.2 gives $\tilde{g}_K = \phi_K \tilde{g}_y + (1 - \phi_K)\tilde{g}_{z^K}$, while Lemma B.1 with $\tilde{g}_L = 0$ gives $\tilde{g}_y = s_K \tilde{g}_K$. Substituting the latter into the former and solving,

$$\tilde{g}_K = \frac{(1 - \phi_K)\tilde{g}_{z^K}}{1 - s_K \phi_K} > 0, \quad \tilde{g}_y = s_K \tilde{g}_K > 0,$$

where $1 - s_K \phi_K > 0$ since $s_K, \phi_K \in (0, 1)$. Now $\tilde{g}_L - \tilde{g}_K = -\tilde{g}_K < 0$, so (B.10) gives $\tilde{g}_{\theta^L} = -\nu(0) + \rho s_K(-\tilde{g}_K) = -\rho s_K \tilde{g}_K < 0$, and (B.11) gives $\tilde{g}_{\theta^K} = (\kappa - \rho)s_L(-\tilde{g}_K) = (\rho - \kappa)s_L \tilde{g}_K > 0$. **QED**

The economically important result is part (iv). Under $\rho = 0$, Lemma 4 of the main text states that land's share is *invariant* to Solow (capital-utilization-driven) growth, so reproducing the observed decline in land's share during industrialization requires an auxiliary force — either a decline in z^L or N (or both). When $\rho \in (0, 1)$ this is no longer needed: because the nests are substitutes, capital deepening bids resources away from the land nest and the land share falls *mechanically*, $\tilde{g}_{\theta^L} = -\rho s_K \tilde{g}_K < 0$, at the same time as θ^K rises. Case A thus reproduces both stylized facts of the industrial transition from the single Solow channel. The cost is that the “ θ^L increases always” conclusions of the stagnation and Malthus regimes (parts i–ii) become conditional: the same substitutability that helps in the Solow regime can, when population grows, pull the land share down through the correction term.

Gross substitution effects in the upper-tier yield more subtle reallocative properties outside of the Solow case (iv). For example, under stagnation (i), population growth independent of growth in output per capita, actually leads to a rising capital share *and* reallocation away from land if the underlying input weights (both upper-tier and within-nest) are sufficiently small. One could interpret this as follows: in relatively small countries population growth actually forces capital to become more important given the productivity of land achieves significantly diminishing returns relatively faster than in an otherwise larger country. Under the pure Malthusian case (ii) the capital share will rise if the output contraction is considerably severe and/or the capital weights in the $\mathcal{K}$ inner nest are relatively large compared to the land weights in the $\mathcal{K}$ inner nest; further, both

capital owners and land owners win out at the relative expense of workers when output drops severely in conjunction with a population decline. Finally, in the case of Malthusian productivity growth (iii), we actually get a cleaner result when $\rho \in (0, 1)$ than under the Cobb-Douglas outer nest: rising z^L ensures land owners win at the expense of capital owners. Whether or not laborers benefit from rising land-utilization efficiency depends on the relative magnitudes of the underlying, primitive elasticities.

B.2 Case B: $\rho < 0$, $\nu, \kappa \in (0, 1)$

Proposition B.3. (*Factor Reallocation When Nests Are Complements*) Let $\rho < 0$ and $\nu, \kappa \in (0, 1)$, and maintain Assumption 1. Then:

- i. **Stagnation** ($g_N > 1$ but $g_y = 1$ with $g_{z^L} = g_{z^K} = 1$): $\tilde{g}_{\theta^L} = [|\rho|s_K\phi_L - \nu(1 - \phi_L)]\tilde{g}_N$, so θ^L rises if and only if $|\rho|s_K\phi_L > \nu(1 - \phi_L)$; $\tilde{g}_{\theta^K} = -(|\rho| + \kappa)s_L\phi_L\tilde{g}_N < 0$, so θ^K falls always.
- ii. **Pure Malthus** ($g_N > 1$ but $g_y < 1$ with $g_{z^L} = g_{z^K} = 1$): $\tilde{g}_{\theta^K} = (\kappa + |\rho|)s_L(\phi_K|\tilde{g}_y| - \phi_L\tilde{g}_N)$, so θ^K rises if and only if $\phi_K|\tilde{g}_y| > \phi_L\tilde{g}_N$; the land share obeys $\tilde{g}_{\theta^L} = -\nu(1 - \phi_L)\tilde{g}_N + |\rho|s_K(\phi_L\tilde{g}_N - \phi_K|\tilde{g}_y|)$, of ambiguous sign.
- iii. **Malthusian Productivity Growth** ($g_{z^L} > 1$ but $g_y = 1$ with $g_N = g_{z^K} = 1$): $\tilde{g}_{\theta^L} = -(\nu + |\rho|s_K)(1 - \phi_L)\tilde{g}_{z^L} < 0$ and $\tilde{g}_{\theta^K} = (\kappa + |\rho|)s_L(1 - \phi_L)\tilde{g}_{z^L} > 0$: the land share falls and the capital share rises.
- iv. **Solow Growth** ($g_{z^K} > 1$ leading to $g_y > 1$ with $g_N = g_{z^L} = 1$): $\tilde{g}_{\theta^L} = |\rho|s_K\tilde{g}_K > 0$ and $\tilde{g}_{\theta^K} = -(\kappa + |\rho|)s_L\tilde{g}_K < 0$: the land share rises and the capital share falls, with $\tilde{g}_K = (1 - \phi_K)\tilde{g}_{z^K}/(1 - s_K\phi_K) > 0$.

Proof. The nest growth rates under each regime are exactly as computed in the proof of Proposition B.2, since Lemmas B.2 and B.1 and Assumption 1 do not depend on the signs of ν, κ, ρ . Only the signing changes: now $\nu = +\nu, \kappa = +\kappa$ with $\nu, \kappa \in (0, 1)$, and $\rho = -|\rho| < 0$, so $\kappa - \rho = \kappa + |\rho| > 0$, $\rho s_K = -|\rho|s_K < 0$, and $\rho - \kappa = -(\kappa + |\rho|) < 0$.

(i) *Stagnation.* With $\tilde{g}_L = -\phi_L\tilde{g}_N, \tilde{g}_K = 0$: (B.10) gives $\tilde{g}_{\theta^L} = -\nu(1 - \phi_L)\tilde{g}_N + \rho s_K(-\phi_L\tilde{g}_N) = [|\rho|s_K\phi_L - \nu(1 - \phi_L)]\tilde{g}_N$, which is positive if and only if $|\rho|s_K\phi_L > \nu(1 - \phi_L)$. Equation (B.11) gives $\tilde{g}_{\theta^K} = (\kappa - \rho)s_L(-\phi_L\tilde{g}_N) = -(\kappa + |\rho|)s_L\phi_L\tilde{g}_N < 0$.

(ii) *Pure Malthus.* With $\tilde{g}_L - \tilde{g}_K = \phi_K|\tilde{g}_y| - \phi_L\tilde{g}_N$ as before: (B.11) gives $\tilde{g}_{\theta^K} = (\kappa + |\rho|)s_L(\phi_K|\tilde{g}_y| - \phi_L\tilde{g}_N)$, which again is positive if and only if $\phi_K|\tilde{g}_y| > \phi_L\tilde{g}_N$. Equation (B.10) gives $\tilde{g}_{\theta^L} = -\nu(1 - \phi_L)\tilde{g}_N - |\rho|s_K(\phi_K|\tilde{g}_y| - \phi_L\tilde{g}_N)$; the first term is negative and the second has the sign of $\phi_L\tilde{g}_N - \phi_K|\tilde{g}_y|$, so the total is ambiguous, rising only if the correction term (positive when population growth dominates the output contraction) outweighs the direct substitution term.

(iii) *Malthusian productivity growth.* With $\tilde{g}_L = (1 - \phi_L)\tilde{g}_{z^L} > 0, \tilde{g}_K = 0$: (B.10) gives $\tilde{g}_{\theta^L} = (-\nu + \rho s_K)\tilde{g}_L = -(\nu + |\rho|s_K)(1 - \phi_L)\tilde{g}_{z^L} < 0$, since both $\nu > 0$ and $|\rho|s_K > 0$. Equation (B.11) gives $\tilde{g}_{\theta^K} = (\kappa - \rho)s_L\tilde{g}_L = (\kappa + |\rho|)s_L(1 - \phi_L)\tilde{g}_{z^L} > 0$.

(iv) *Solow growth.* With $\tilde{g}_L = 0$ and $\tilde{g}_K = (1 - \phi_K)\tilde{g}_{z^K}/(1 - s_K\phi_K) > 0$ as derived before: (B.10) gives $\tilde{g}_{\theta^L} = \rho s_K(-\tilde{g}_K) = |\rho|s_K\tilde{g}_K > 0$, and (B.11) gives $\tilde{g}_{\theta^K} = (\kappa - \rho)s_L(-\tilde{g}_K) = -(\kappa + |\rho|)s_L\tilde{g}_K < 0$. QED

Case B actually *precludes* the notion that the outer nest actually features gross complementarities and/or capital/efficiency-labor (or land/efficiency-labor) substitutabilities in empirical reality, given the factor-share reallocation patterns we observe in data. In the isolated Solow channel (iv), a rise in capital-utilization efficiency now *lowers* the capital share and *raises* the land share. Two forces combine: within a substitutable capital nest ($\kappa \in (0, 1)$) the gains from capital deepening accrue to labor rather than to capital, and the complementary upper tier ($\rho < 0$) channels resources *toward* the relatively scarce land nest, raising θ^L . Under case B reproducing the observed joint movement in the data where the capital share rises as the land share falls in a growing economy would therefore require a *large* offsetting decline in z^L working against the mechanism, rather than being delivered by it. This is an economic counterpart to the statistical verdict in Technical Appendix C.1: the $\rho < 0, \nu, \kappa \in (0, 1)$ specification estimated in that appendix not only carries unreliable fit statistics but also, as this appendix shows, theoretically implies share dynamics that run against the data, whereas the preferred $\rho = 0, \nu, \kappa < 0$ model and its near neighbor case A do not.

Gross complementarities and within-nest substitution effects always lead population growth absent growth in GDP per capita to drive down the capital share. In fact, in the stagnation setting population growth can actually be *anti-Malthusian*, reallocating income *toward* labor if labor's within-nest weights are sufficiently large and land/efficiency-labor substitution effects sufficiently weak ($\nu \approx 0$). Further, within-nest substitution effects mean that productivity increases in one nest (all else equal) lead to the income share from the primary factor in the *opposite* nest to rise. This is sort of “cost disease” for factor inputs in that the input whose utilization labor is less efficient gets a greater share of payments.

B.3 Scope and Caveats

Three qualifications delimit these results and mark them as robustness companions to the main analysis rather than replacements for it. First, because s_L, s_K, ϕ_L , and ϕ_K are endogenous and move along the path, the sign conditions above are *local*: unlike the unconditional statements available under $\rho = 0$, several conclusions (the stagnation and Malthus land shares in both cases) hold only conditional on the prevailing value share and on the relative speeds of population, output, and productivity growth. Second, we have used the constant-investment-rate relation $\tilde{g}_k = \tilde{g}_y$ as a maintained hypothesis, exactly as in Section 2.3; we do not assert existence of a balanced growth path for the CES-outer-nest technology with a fixed factor, which is well known to be delicate (the Uzawa steady-state-growth theorem requires labor-augmenting technical change in the aggregator), but confine attention to local decompositions along the path. Third, all statements isolate a single channel at a time; the observed data reflect the superposition of these channels,

so the propositions are diagnostic of *directional forces*, not of net movements. With these qualifications, the two cases sharpen the reading of the $\rho = 0$ analysis: they show precisely which of its conclusions are knife-edge to the Cobb-Douglas upper tier (the invariance of θ^L to Solow growth, and the unconditional Malthusian rise in θ^L) and which survive, and they supply a structural reason beyond the fit diagnostics to prefer the $\rho = 0$ specification of the main text.

C Estimation Appendix

C.1 Estimated Models with ρ as a Free Parameter (Full Sample)

In this appendix we estimate models where $\rho \neq 0$ (i.e., allowed to be a free parameter) using the full 1200-1860 sample. Such models now feature six static free parameters, rather than five. We consider two different models with a free value of ρ : in one model we assume the land and capital aggregates are substitutes (i.e., $\rho \in (0, 1)$) while land/efficiency-labor and capital/efficiency-labor are complements (i.e., $\nu, \kappa < 0$); in the other model we flip the signs — $\rho < 0$ and $\nu, \kappa \in (0, 1)$. When $\rho \neq 0$ we set its prior to be either $\rho \sim \mathcal{U}[0, 1]$ or $-\rho \sim \mathcal{LN}(0, 2^2)$. Parameter values for these alternative estimators are presented in Table C.1. Note that estimates for ρ , since it is an explicitly targeted parameter, are included at the bottom of the table in addition to all of the parameters also estimated in the models restricted to have an upper-tier Cobb-Douglas aggregator. We discuss fitness properties of the models in Table C.1 relative to those in Table 1 in the main text in Section 3.1. Recall, we prefer the $\rho = 0$ with $\nu, \kappa < 0$ model both for its fitness properties and the robustness of its estimator relative to other $\rho = 0$ models and those presented in this appendix.

Table C.1: Production Function Estimates; England + Wales, 1200-1860 ($\rho \neq 0$)

	(1)	(2)
	$\rho \in (0, 1); \nu, \kappa < 0$	$\rho < 0; \nu, \kappa \in (0, 1)$
ν	−12.227 (10.728)	0.947 (0.024)
κ	−0.736 (0.109)	0.383 (0.249)
ω_L	0.167 (0.097)	0.078 (0.064)
ω_K	0.302 (0.023)	0.102 (0.119)
α	0.532 (0.024)	0.473 (0.005)
ρ	0.890 (0.077)	−78.855 (31.695)
WAIC ^a	48.764	301.280

^a Widely applicable info. criterion, smaller = better (Watanabe 2010; Vehtari, Gelman, and Gabry 2017).

C.2 Predicted Output Diagnostics (Alternative Models, Full Sample)

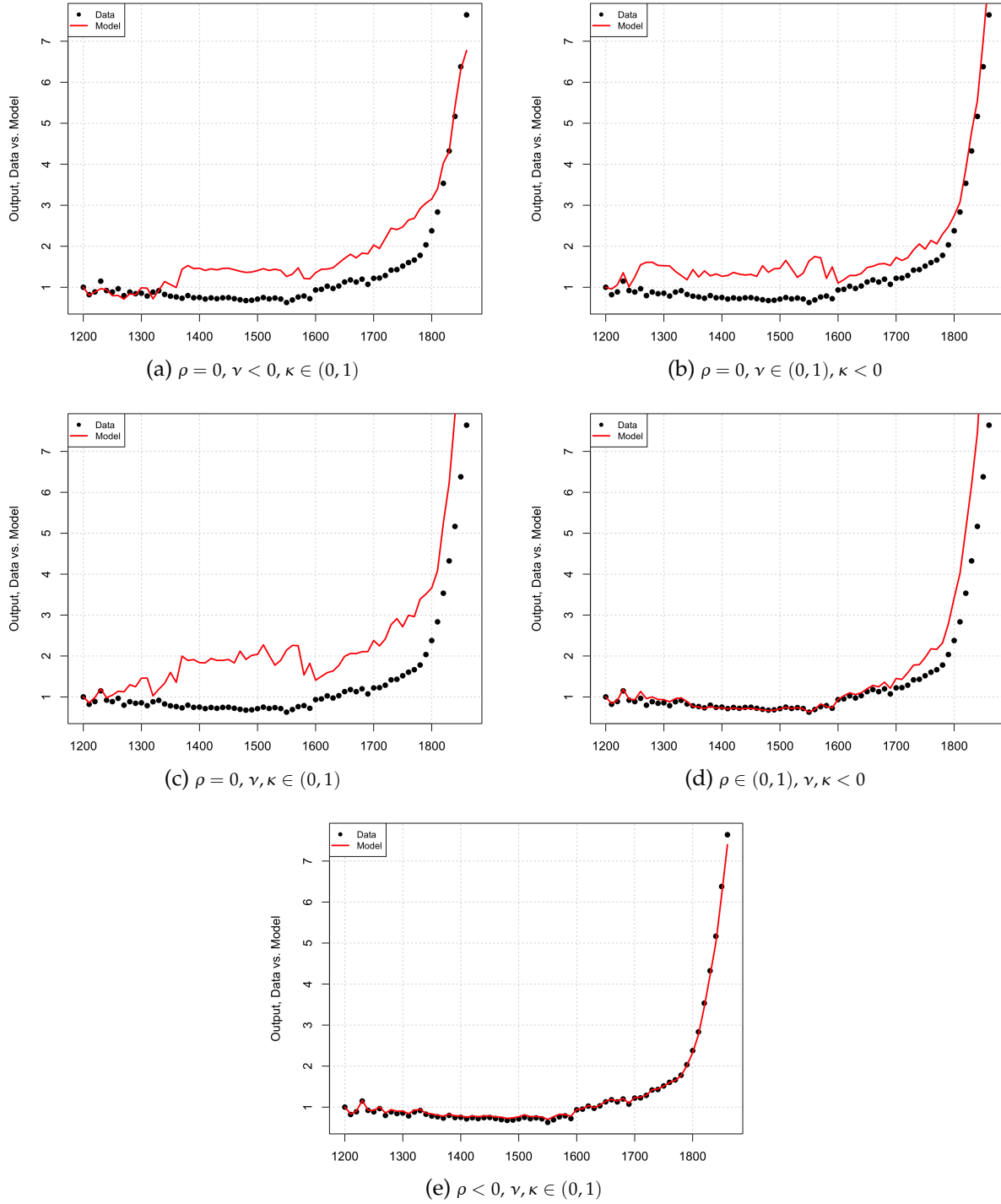


Figure C.1: This figure presents the fitness for non-targeted $\{\hat{Y}_t\}_{t=1200}^{1860}$ for all full-sample, single-sector models except the $\rho = 0, \nu, \kappa < 0$ model featured in Figure 3.

C.3 Prior & Posterior Diagnostics (Alternative Models, Full Sample)

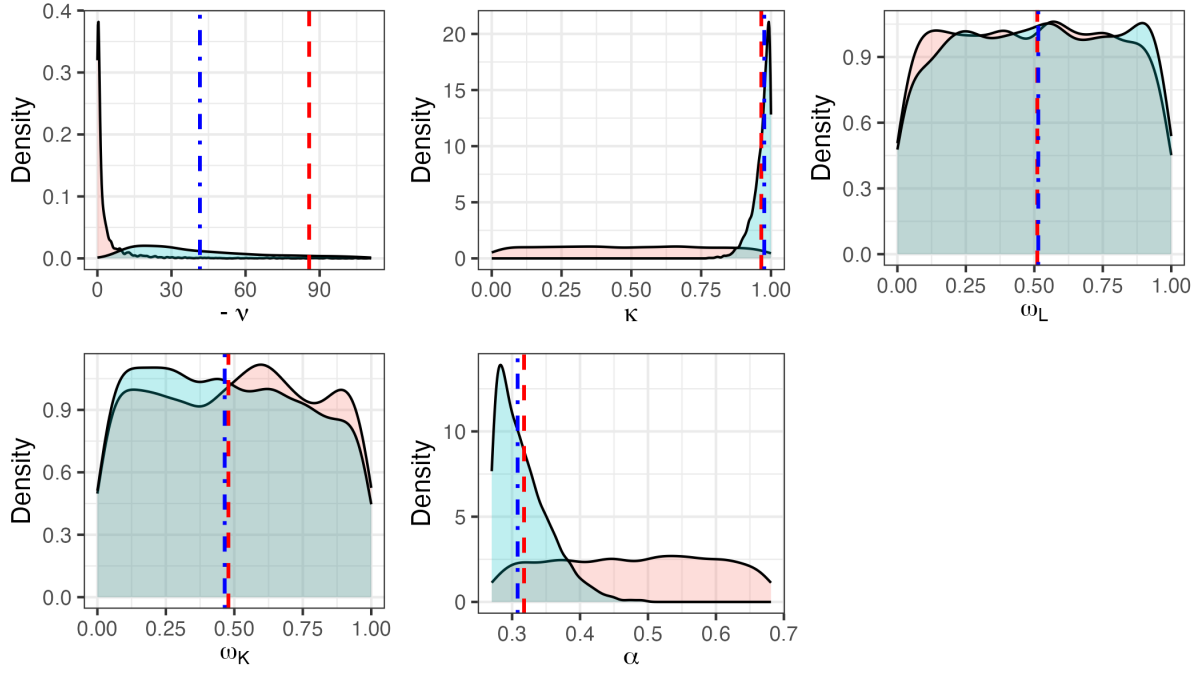


Figure C.2: This figure is analogous to Figure 4 in the main text for the $\rho = 0$, $v < 0$, and $\kappa \in (0, 1)$ model.

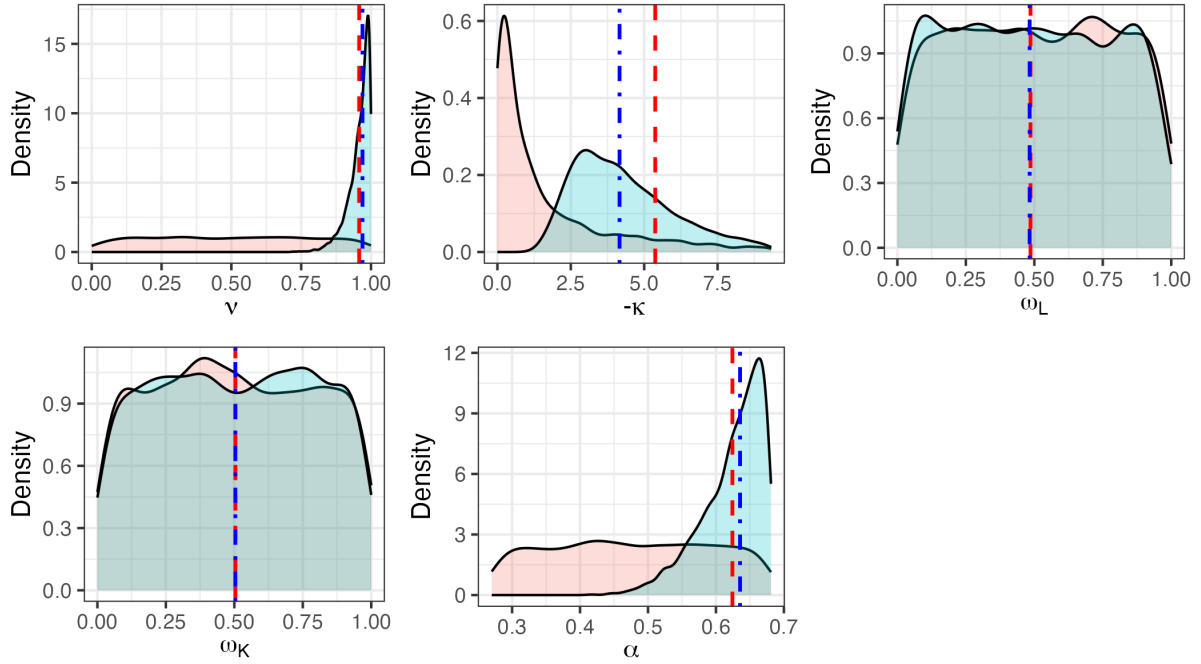


Figure C.3: This figure is analogous to Figure 4 in the main text for the $\rho = 0$, $v \in (0, 1)$, and $\kappa < 0$ model.

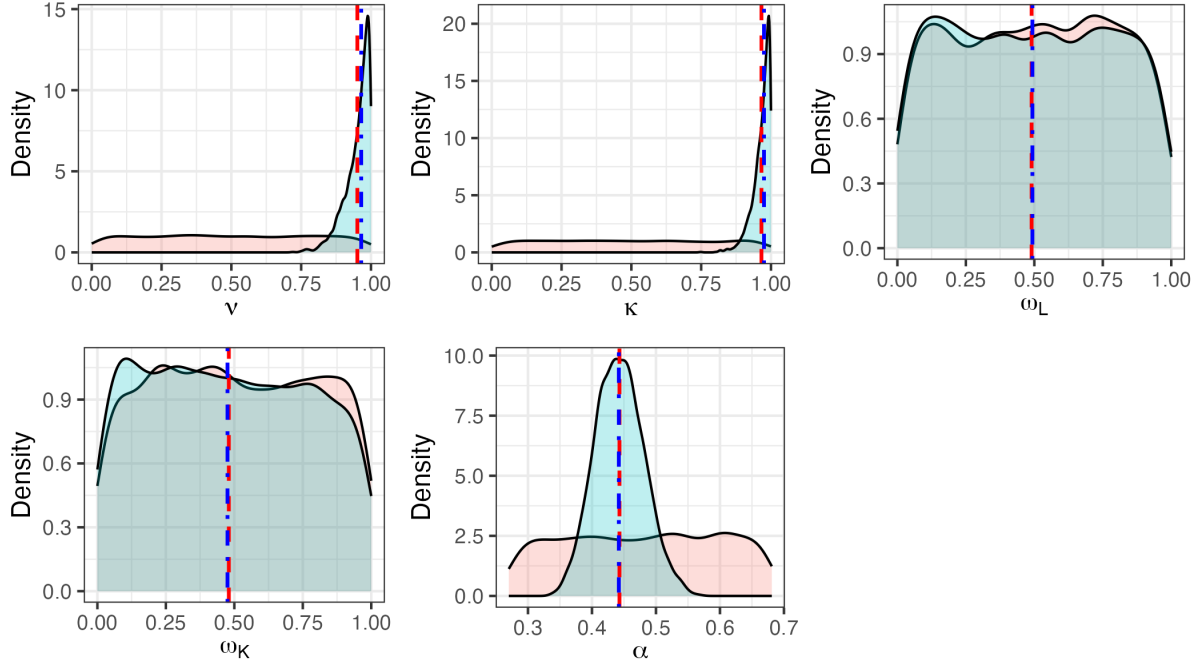


Figure C.4: This figure is analogous to Figure 4 in the main text for the $\rho = 0$ and $\nu, \kappa \in (0, 1)$ model.

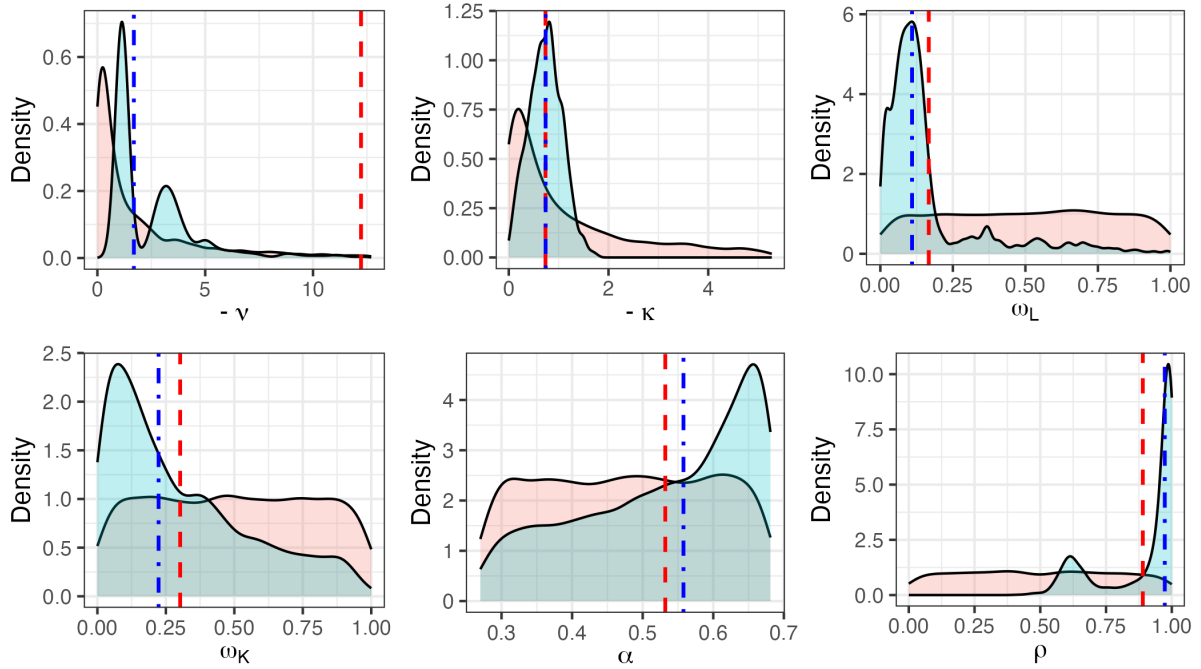


Figure C.5: This figure is analogous to Figure 4 in the main text for the $\rho \in (0, 1)$ and $\nu, \kappa < 0$ model.

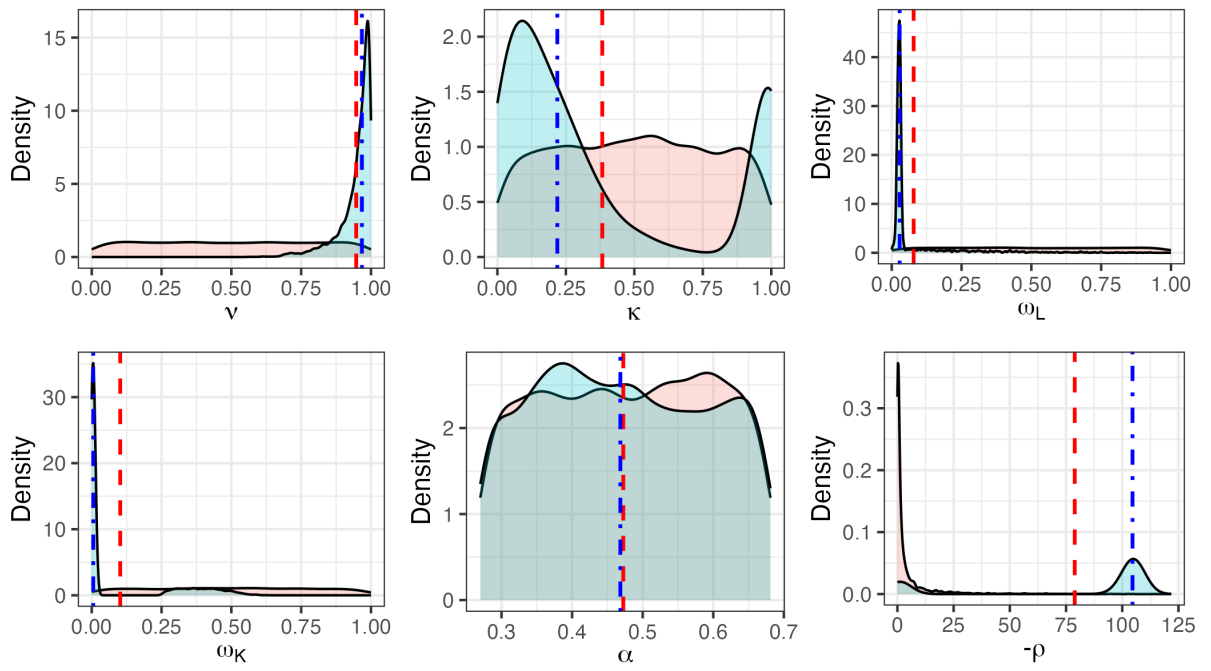


Figure C.6: This figure is analogous to Figure 4 in the main text for the $\rho < 0$ and $\nu, \kappa \in (0, 1)$ model.

C.4 Estimated Factor-augmenting Productivities (Alternative Models, Full Sample)

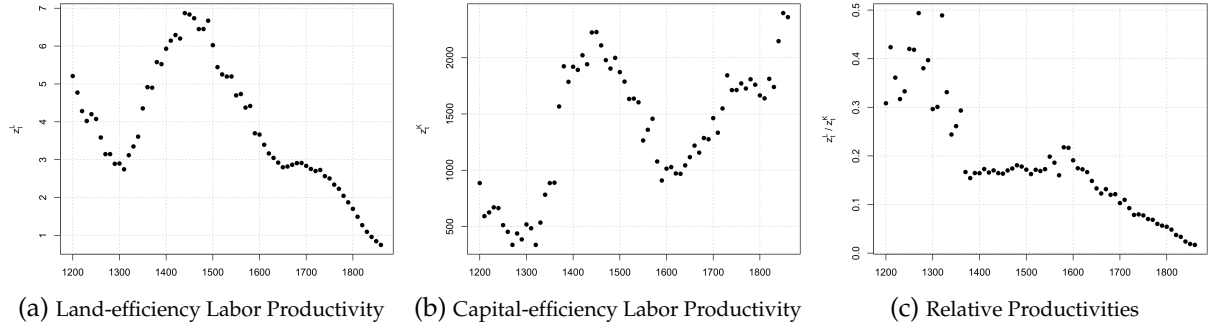


Figure C.7: The figure contains estimated posterior-mean productivity levels in panels (a) and (b) and relative land/non-landed capital productivities in panel (c) for the model with $\rho = 0$, $\nu < 0$, and $\kappa \in (0, 1)$.

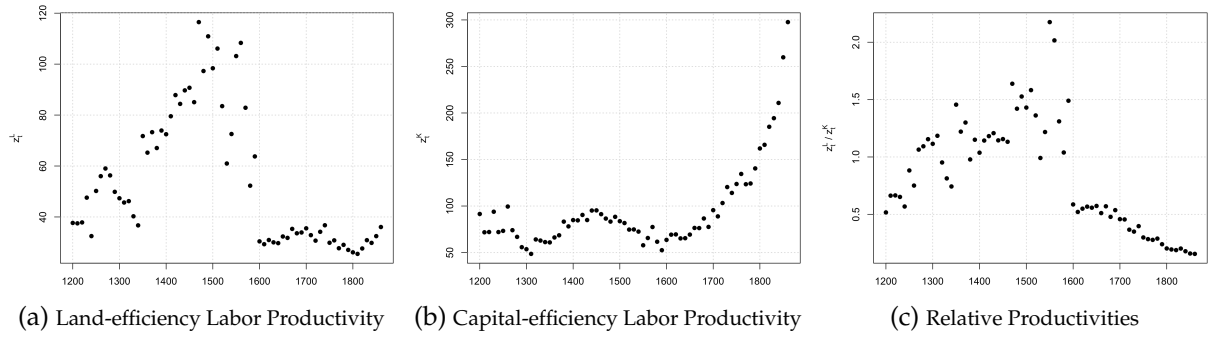


Figure C.8: The figure contains estimated posterior-mean productivity levels in panels (a) and (b) and relative land/non-landed capital productivities in panel (c) for the model with $\rho = 0$, $\nu \in (0, 1)$, and $\kappa < 0$.

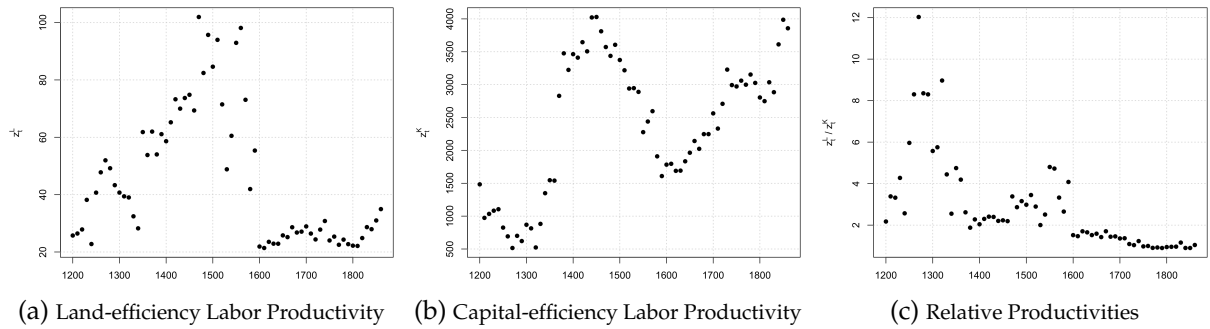


Figure C.9: The figure contains estimated posterior-mean productivity levels in panels (a) and (b) and relative land/non-landed capital productivities in panel (c) for the model with $\rho = 0$ and $\nu, \kappa \in (0, 1)$.

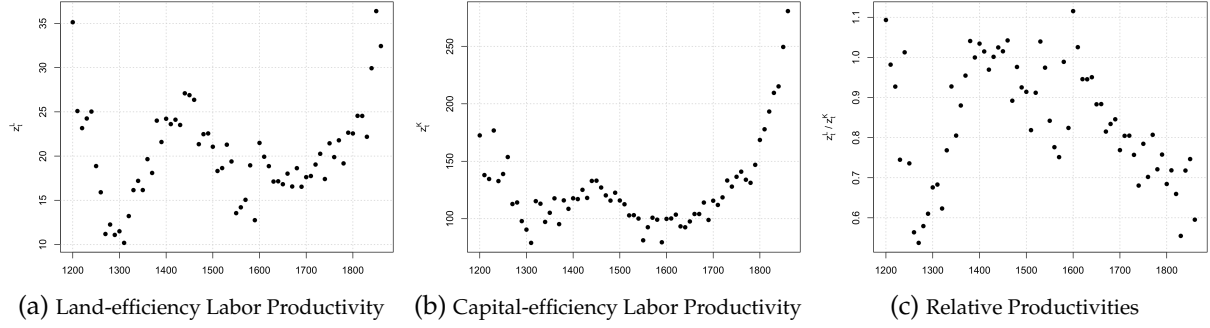


Figure C.10: The figure contains estimated posterior-mean productivity levels in panels (a) and (b) and relative land/non-landed capital productivities in panel (c) for the model with $\rho \in (0, 1)$ and $\nu, \kappa < 0$.

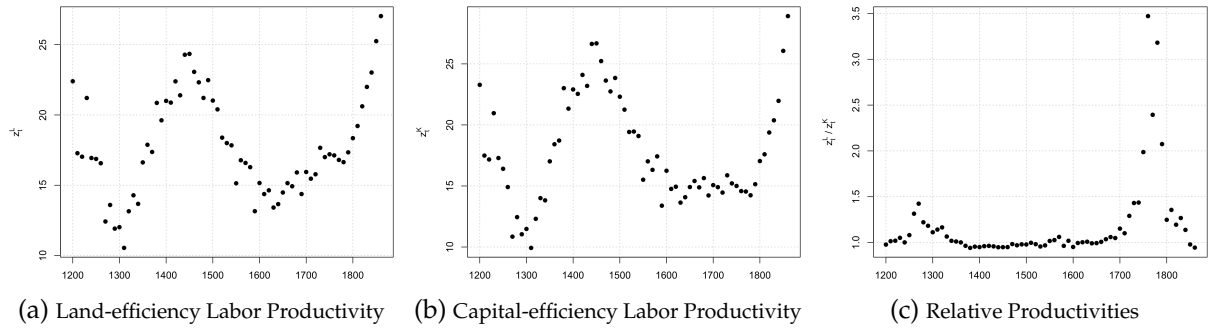


Figure C.11: The figure contains estimated posterior-mean productivity levels in panels (a) and (b) and relative land/non-landed capital productivities in panel (c) for the model with $\rho < 0$ and $\nu, \kappa \in (0, 1)$.

C.5 Estimated Elasticities of Substitution (Alternative Models, Full Sample)

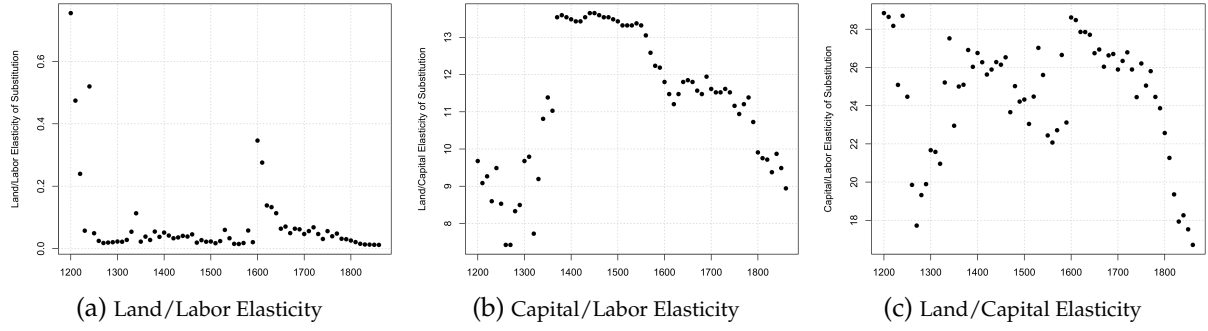


Figure C.12: This figure plots the estimated elasticities of substitution for the model with $\rho = 0$, $\nu < 0$, and $\kappa \in (0, 1)$.

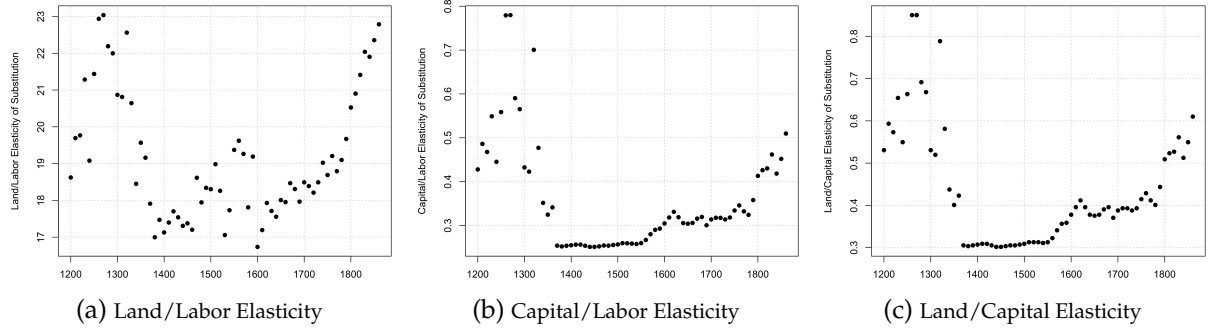


Figure C.13: This figure plots the estimated elasticities of substitution for the model with $\rho = 0$, $\nu \in (0, 1)$, and $\kappa < 0$.

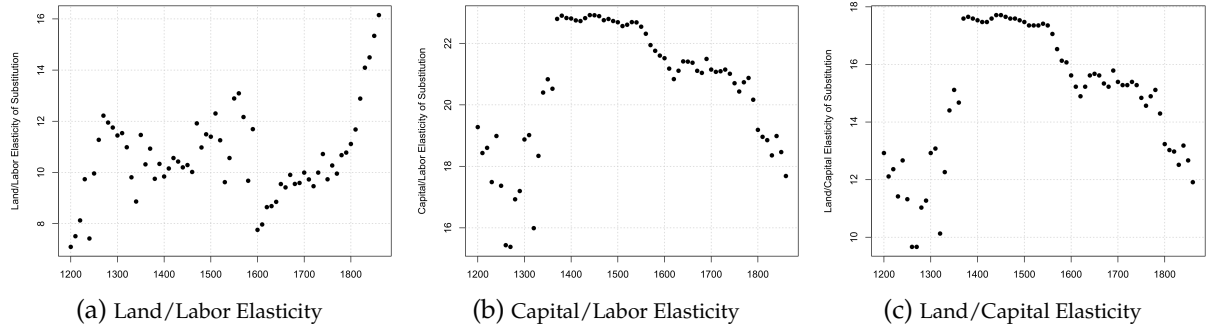


Figure C.14: This figure plots the estimated elasticities of substitution for the model with $\rho = 0$ and $\nu, \kappa \in (0, 1)$.

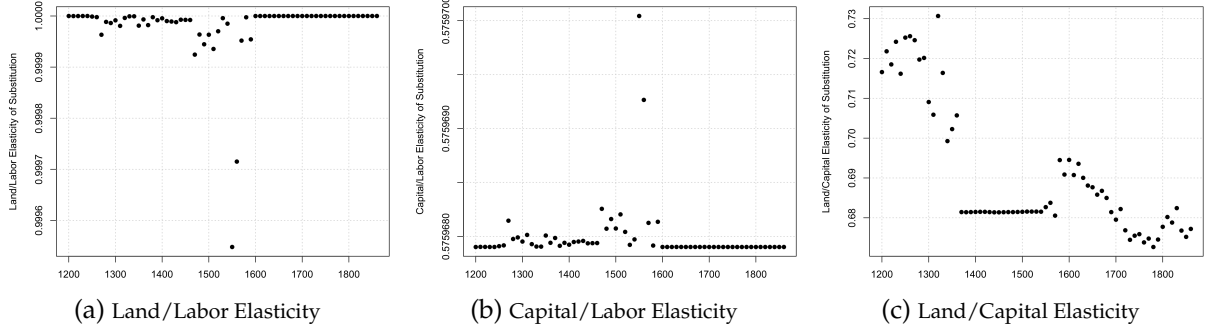


Figure C.15: This figure plots the estimated elasticities of substitution for the model with $\rho \in (0, 1)$ and $\nu, \kappa < 0$.

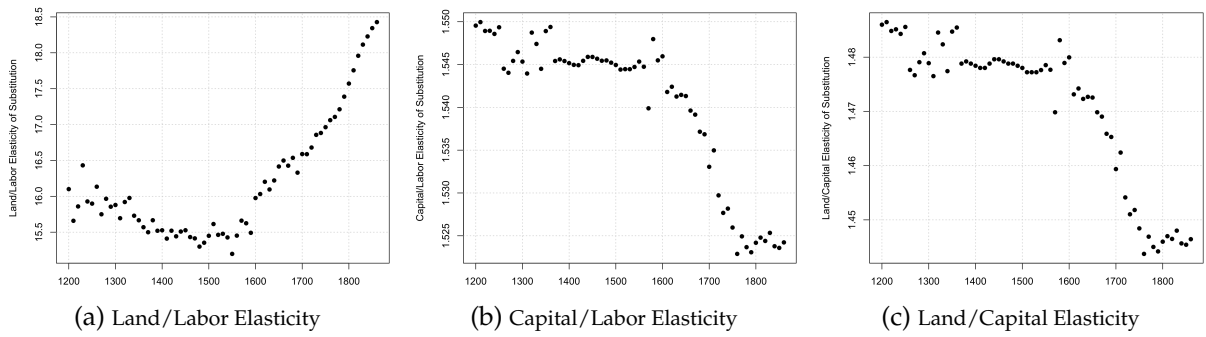


Figure C.16: This figure plots the estimated elasticities of substitution for the model with $\rho < 0$ and $\nu, \kappa \in (0, 1)$.

C.6 Sensitivity Checks via Perturbation (Preferred Model, Full Sample)

A natural concern about the estimated model is that it is unclear how much each of the fixed structural parameters $\{\alpha, \nu, \kappa, \omega_L, \omega_K\}$ drive outcomes. This appendix answers that question directly for the preferred specification ($\rho = 0, \nu, \kappa < 0$). We exploit a feature of the estimator: given the data and *any* values of the five structural parameters, the log productivities, $\ln z_t^L$ and $\ln z_t^K$, are exact algebraic residuals of the factor-share conditions (28) and (29), and model output follows from the production function. The mapping from parameters to the vector $(z_t^L, z_t^K, \hat{Y}_t)$ is therefore a known, deterministic function that we can re-evaluate at will.

Because z_t and $\hat{Y}_t$ are computed at every posterior draw and we report their posterior averages (averages over all computed draws), any sensitivity statement must also be a posterior average over all draws, *not* the map evaluated at the posterior mean point-estimate, which would subject the calculation to errors due to Jensen’s inequality. We therefore proceed as follows. To assess a candidate value of one parameter, we fix that parameter (and only that parameter, one parameter at a time) to a constant value in *every* one of the posterior draws, leaving all other parameters at their drawn values. We then recompute $\{z_t^L, z_t^K\}_t$ from (28) and (29) and the implied $\hat{Y}_t$ *draw by draw*. We then average across posterior draws. Every path we report is thus the mean of a re-simulated distribution obtained by grid-fixing a single parameter. We do this along a grid for each parameter, one at a time, and we complement it with a variance decomposition of the outcomes over the untouched posterior draws. Because the recovered productivities reconcile the factor income shares *exactly* in every draw, the shares carry no residual; the informative dimension of fit is the non-targeted output series, $\{\hat{Y}_t\}_t$.

Table C.2: Output-fit leverage of each structural parameter (draw-averaged).

	grid range	RMSE _{lo}	RMSE _{hi}	RMSE _{min}	RMSE _{max}	range
κ	$[-6, -0.25]$	0.1920	0.5133	0.1029	0.5133	0.4104
α	$[0.30, 0.66]$	0.1397	0.2816	0.1257	0.2816	0.1559
ν	$[-90, -3]$	0.1227	0.1058	0.1058	0.1227	0.0170
ω_K	$[0.30, 0.70]$	0.1251	0.1288	0.1251	0.1288	0.0036
ω_L	$[0.30, 0.70]$	0.1190	0.1190	0.1190	0.1190	0.0001

RMSE of $\ln \hat{Y}_t$ vs. $\ln Y_t$ (normalized to 1200), log points. Baseline (nothing fixed) = 0.1190.

RMSE_{lo}/RMSE_{hi} are at the low/high ends of the grid; RMSE_{min}/RMSE_{max} over the whole grid.

The baseline RMSE from the preferred model is 0.119. Table C.2 presents the grid ranges we consider for each fixed parameter, as well as the different RMSE’s over this range, where “lo” corresponds to the RMSE associated with the smallest grid parameter and “hi” the RMSE associated with the largest. Meanwhile, “min” represents the minimum RMSE over the grid (in the case where the RMSE varies non-monotonically over the grid, this value may be different from either “lo” or “hi”) and analogously for “max.” The number of grid points differ depending on the parameter we are fixing. Figure C.17 plots the posterior RMSE’s when fixing each parameter

at a specific value. Note that, while at first glance these sensitivity analyses reveal that only very marginal RMSE improvements could be achieved by fixing one of the parameters (e.g., ν and ω_L), a caveat is in order. Because we are *fixing* one parameter and then resimulating the non-targeted output path over the whole posterior distribution, the reduction in RMSE is not just mechanical but also a statistical by-product of the fact that the re-simulated posterior is determined by one fewer random variable.

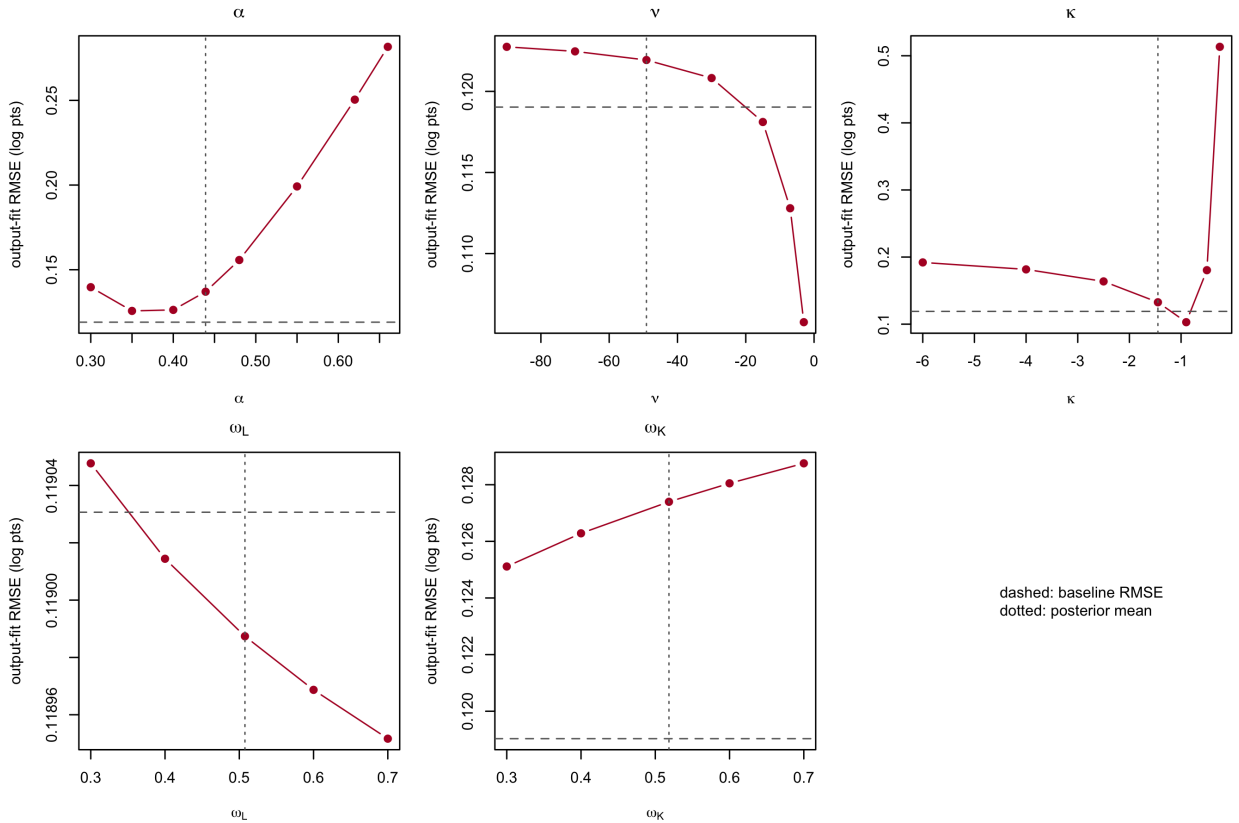


Figure C.17: We plot the output-fit RMSE of the normalized (year 1200) predictive series $\{\hat{Y}_t\}_t$ relative to the normalized data. Dashed lines are the baseline RMSE, while the dotted lines (vertical) are the posterior means at each particular parameter being simulated.

While it may be clear from Figure C.17 that variation in κ (the primitive capital/capital-efficiency labor elasticity) has the most leverage over posterior fitness, Figure C.18 makes this more explicit. Here, we show the induced range in RMSE over the chosen grid we simulate. Clearly, moving κ around changes output fitness the most, followed by the outer CES weight on the land nest, α . Despite its rather large magnitude in Table 1, variation in ν has little affect on model fitness.

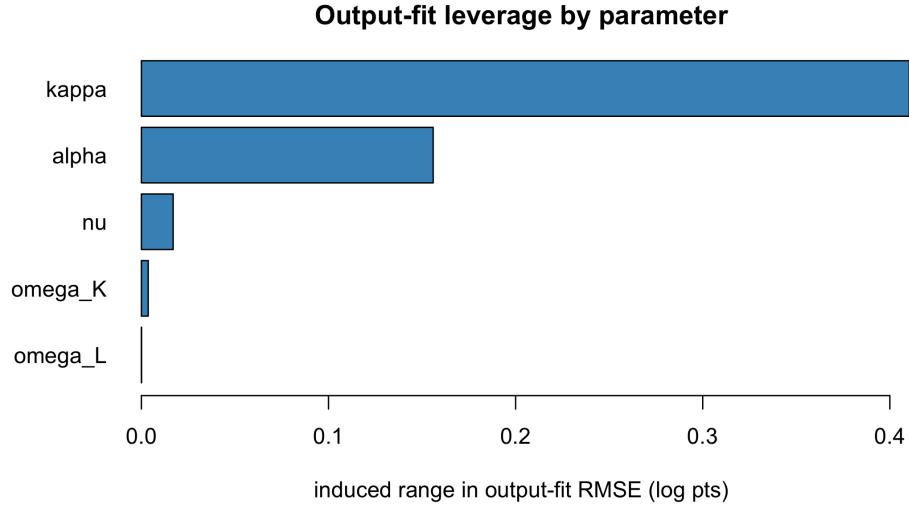


Figure C.18: This figure shows the simulated range of RMSE outcomes (as measured for predictive $\hat{Y}_t$ versus data).

C.6.1 Productivities Under Sensitivity Checks

How are the productivity residuals affected by variation in each of the specific, underlying primitive parameters? Generally speaking, we find that the capital-utilization productivities are more sensitive to varying the fixed parameters than the land-utilization ones. Not surprisingly, fixing κ at values far away from its posterior mean has a huge impact on the level of estimates z_t^K , but variation in ν has a significantly smaller impact on its associated productivity, z_t^L . Both ω_K and α dictate greater leverage over the capital series than the land series, too. Notable, though, is that why the levels of the series may vary in the underlying value of the fixed parameter we target, the first- and second-order properties do not seem to change at all: the land-utilization efficiency always falls and the capital-utilization efficiency always rises, regardless of the levels of the fixed parameters. All of this can be gleaned by inspecting Figure C.19 below.

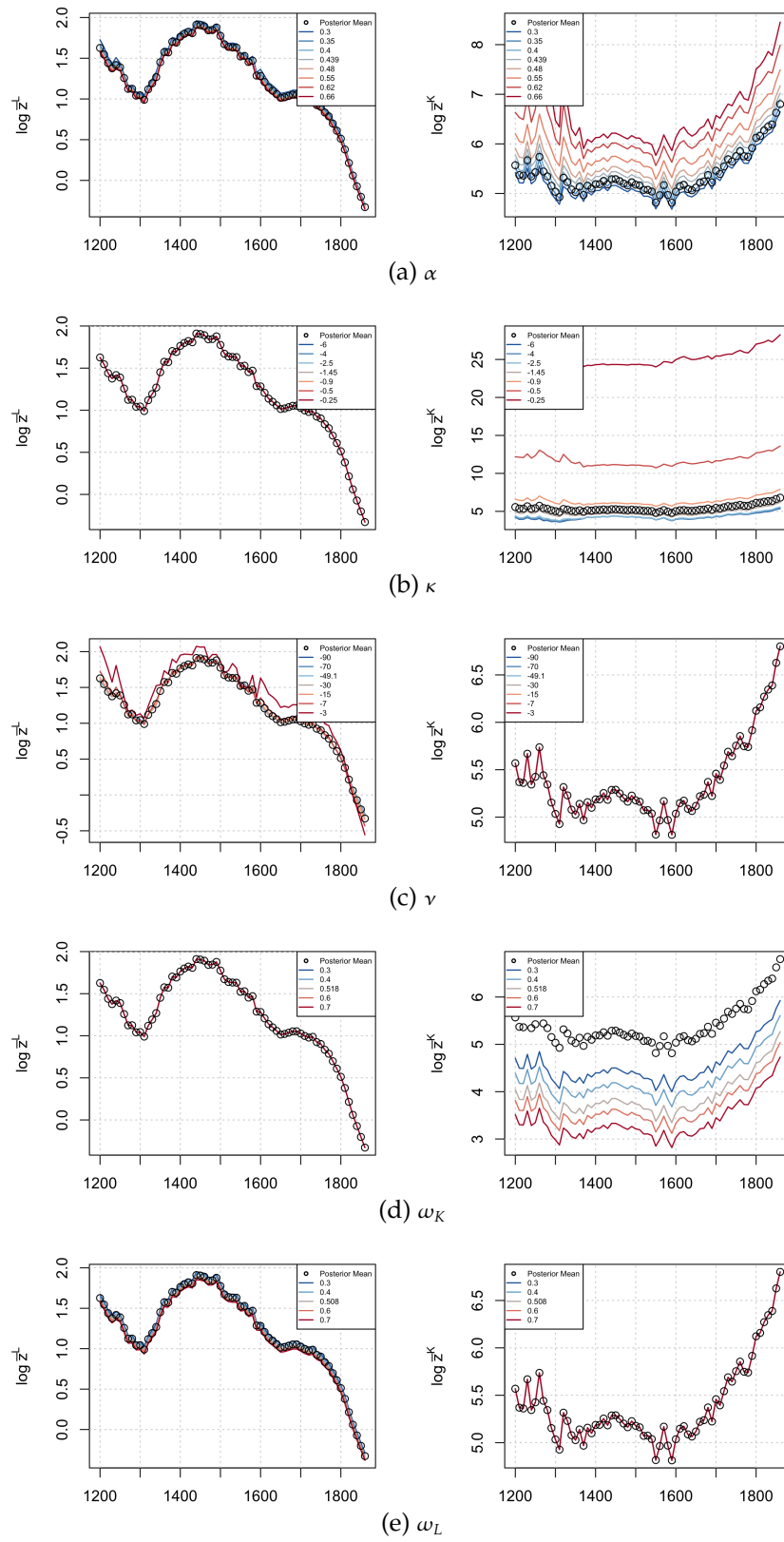


Figure C.19: Here, we plot the productivity series under all of the perturbations outlined. Levels move around, but first- and second-order properties are unchanged.

C.7 Estimated Models on Truncated Data (1500-1860 Sample)

In this appendix we re-estimate the model on a truncated dataset starting in 1500. Table C.3 contains the posterior estimates for all six models under the six different parametric restrictions. Comparing the results in Table C.3 to those in Tables 1 and C.1, we notice several things. First, focusing on the preferred model in column (1) (i.e., $\rho = 0$, $\nu, \kappa < 0$, which is still our preferred model under the truncated sample as discussed in Technical Appendix C.7.1 below), notice that land/efficiency-labor complementarities (ν) are weaker under the truncated sample, though still strong. Second, capital/efficiency-labor complementarities are actually significantly stronger ($\kappa < 0$). This is to be expected when training a model on a dataset restricted to eras of capital-intensive industrialization.

Generally, though, few things change relative to our predictions in the main text, including the qualitative properties of the factor-augmenting productivities, featured in Technical Appendix C.7.2. In Technical Appendix C.7.1 we do find that fitness comparisons are but murkier than in the main text, but still conclude that the $\rho = 0$, $\nu, \kappa < 0$ model is the best one for largely the same reasons.

Table C.3: Production Function Estimates; England + Wales, 1500-1860

	$\rho = 0$				$\rho \neq 0$	
	(1) $\nu, \kappa < 0$	(2) $\nu < 0, \kappa \in (0, 1)$	(3) $\nu \in (0, 1), \kappa < 0$	(4) $\nu, \kappa \in (0, 1)$	(5) $\rho \in (0, 1); \nu, \kappa < 0$	(6) $\rho < 0; \nu, \kappa \in (0, 1)$
ν	-24.406 (1.9176)	-49.364 (2.5386)	0.948 (0.0008)	0.907 (0.0013)	-9.403 (7.3381)	0.960 (0.0136)
κ	-16.111 (1.2342)	0.863 (0.0021)	-78.605 (8.6286)	0.771 (0.0032)	-0.199 (0.0278)	0.677 (0.1255)
ω_L	0.512 (0.007)	0.508 (0.005)	0.491 (0.006)	0.491 (0.006)	0.188 (0.131)	0.087 (0.042)
ω_K	0.489 (0.007)	0.471 (0.007)	0.505 (0.006)	0.474 (0.006)	0.535 (0.067)	0.086 (0.104)
α	0.451 (0.003)	0.440 (0.002)	0.631 (0.001)	0.596 (0.001)	0.567 (0.053)	0.492 (0.004)
ρ	0	0	0	0	0.798 (0.1571)	-77.516 (31.0239)
WAIC ^a	52.414	61.722	83.598	46.226	-6.361	245.171

^a Widely applicable info. criterion, smaller = better (Watanabe 2010; Vehtari, Gelman, and Gabry 2017).

C.7.1 Predicted Output Diagnostics (1500-1860 Sample)

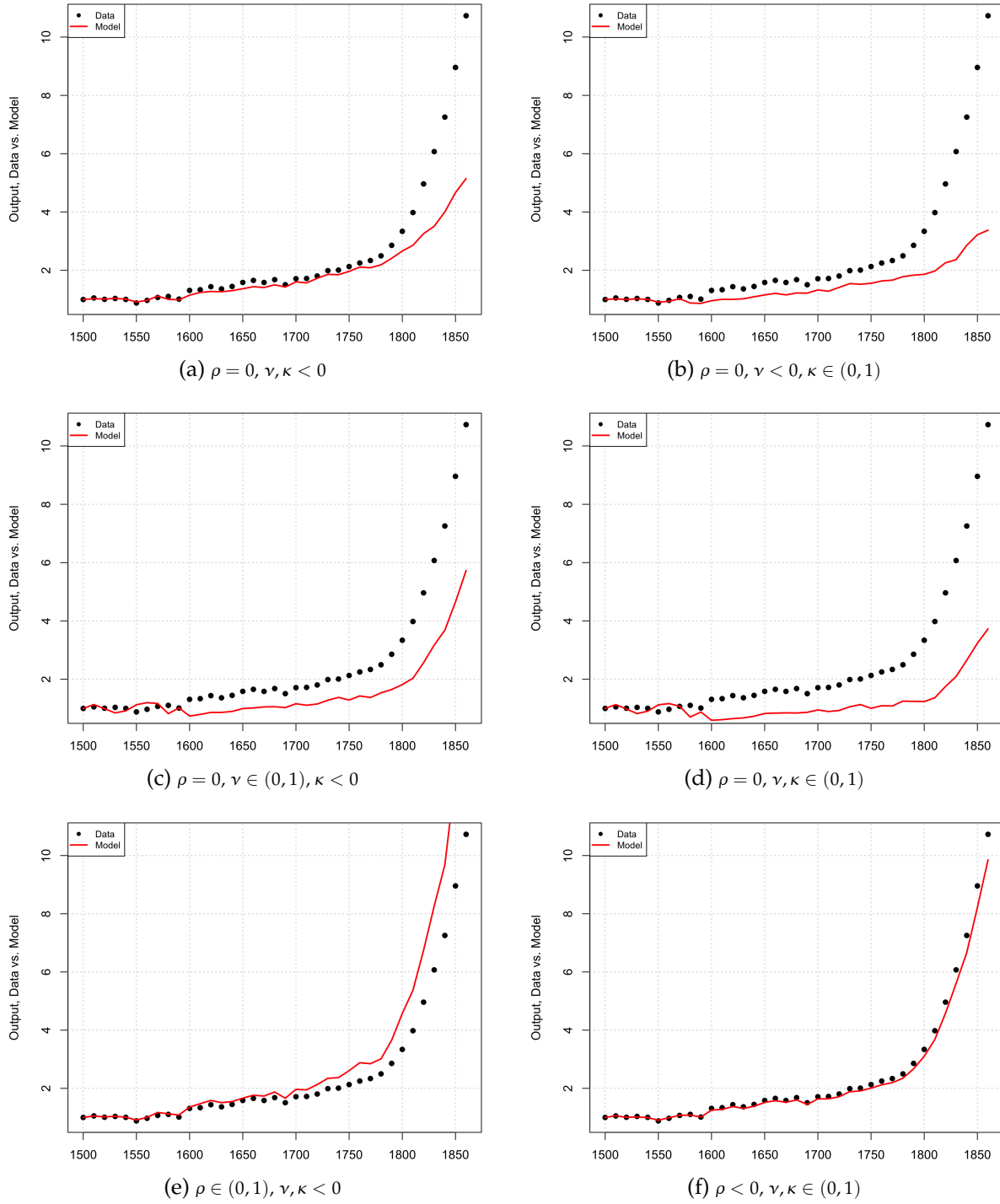


Figure C.20: This figure presents the fitness for non-targeted $\{\hat{Y}_t\}_{t=1500}^{1860}$ for all single-sector, truncated data models.

According to both WAIC and LOO cross validation, the $\rho = 0$ model with $\nu, \kappa \in (0, 1)$ is the best fitting model, though the $\rho = 0, \nu, \kappa < 0$ (i.e., the preferred model from the main text) is very similar. The non-targeted fitness eye test on $\hat{Y}_t$ in Figure C.20 suggests that the Cobb Douglas outer nest with inner-nest complementarities model in Figure C.20a is the best fitting model that is not biased by poor Hamiltonian Monte Carlo sampling properties, just like in the main text. Estimates for models featured in Figures C.20e and C.20f are generally unreliable due to such models' poor predictive sampling properties: this is the same problem that plagues the models with $\rho \neq 0$ on the full sample discussed in Technical Appendix C.1. Given the fitness properties and parametric magnitudes are similar when estimating on 1500-1860 data as 1200-1860 data, we conclude that using the full Clark (2010) dataset does not induce any significant bias in our results.

C.7.2 Estimated Factor-augmenting Productivities (Preferred Model, 1500-1860 Sample)

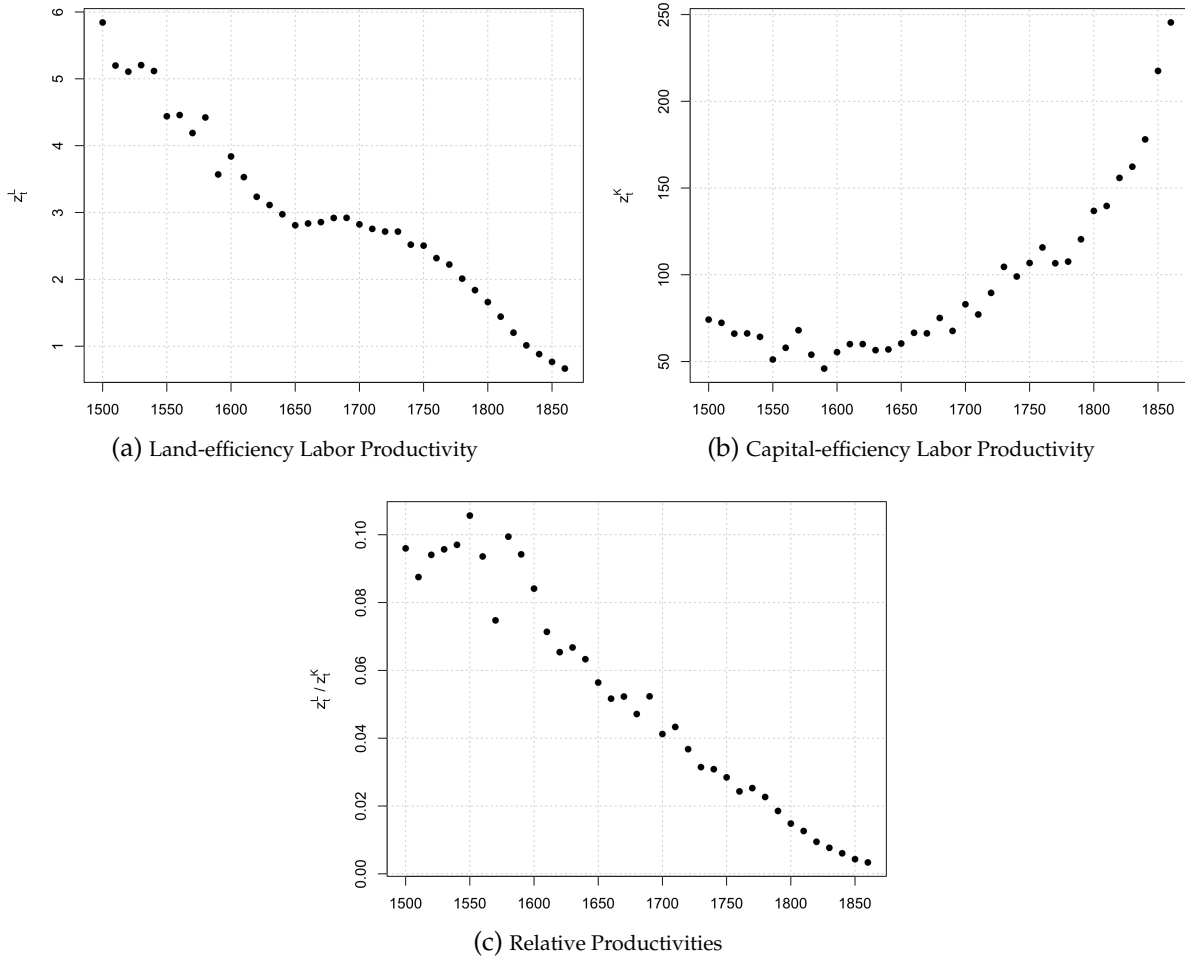


Figure C.21: The figure is analogous to Figure 5 in the main text for the preferred model but as estimated over the truncated sample, 1500-1860.

D Multi-sector Re-interpretation

This appendix considers a re-interpretation of the single-sector model through a multi-sector lens. Think of $\mathcal{L}_t$ as a “land-intensive” sector (e.g., agriculture in a Malthusian setting) and $\mathcal{K}_t$ as a “capital-intensive” sector (e.g., non-agriculture otherwise). Think of the outer nests in (1) and (2) as final good aggregators over the two sectoral outputs; assume this final goods sector is perfectly competitive. In this re-interpretation all land is used in the “land-intensive” sector and all capital is used in the “capital-intensive” sector, with labor split across the two sectors; we can think of such land- and capital-intensive sectors as the “intermediate” sectors. We write the outputs of these intermediate sectors as $Y_t^L \equiv \mathcal{L}_t$ and $Y_t^K \equiv \mathcal{K}_t$.

Let n_t^L be the fraction of aggregate labor in sector L , so that total labor in sector L is $N_t^L = n_t^L N_t$, and total labor in sector K is $N_t^K = (1 - n_t^L) N_t$. Let A_t^L and A_t^K be the primitive factor-biased technical change terms. The intermediate goods are produced using a sector-specific factor and labor following the same functional forms as the inner nests, $\mathcal{L}_t$ and $\mathcal{K}_t$:

$$Y_t^L = \left(\omega_L L_t^\gamma + (1 - \omega_L) (A_t^L N_t^L)^\gamma \right)^{1/\gamma}, \quad (\text{D.1})$$

$$Y_t^K = \left(\omega_K K_t^\kappa + (1 - \omega_K) (A_t^K N_t^K)^\kappa \right)^{1/\kappa}. \quad (\text{D.2})$$

The sector-specific labor terms are such that $N_t^L + N_t^K = N_t$.

D.1 Equilibrium Aggregation of Two-sector Economy

We now show that the single-sector model of the main text absorbs the equilibrium of the two-sector re-interpretation. Take the final good as numeraire. The final-good firm’s first-order conditions give the intermediate-good prices:^{D1}

$$P_t^L = \alpha \left(\frac{Y_t}{Y_t^L} \right)^{1-\rho}, \quad (\text{D.3})$$

$$P_t^K = (1 - \alpha) \left(\frac{Y_t}{Y_t^K} \right)^{1-\rho}. \quad (\text{D.4})$$

Labor mobility requires the equilibrium value marginal product of labor to be the same in both intermediate sectors:

$$w_t = P_t^L (1 - \omega_L) A_t^L \left(\frac{Y_t^L}{A_t^L N_t^L} \right)^{1-\gamma} = P_t^K (1 - \omega_K) A_t^K \left(\frac{Y_t^K}{A_t^K N_t^K} \right)^{1-\kappa}. \quad (\text{D.5})$$

^{D1}Note that we *do not possess data* for sectoral prices, hence limiting our analyses to a single-sector economy; this appendix is merely an illustrative exercise designed to assuage concerns that our model *cannot* be used for a two-sector analysis: indeed, it certainly can.

Now write $N_t^L = n_t^L N_t$ and $N_t^K = (1 - n_t^L) N_t$, and note that with a slight change of arguments

$$\mathcal{L}_t(n_t^L; A_t^L)^\nu \equiv \omega_L L_t^\nu + (1 - \omega_L) \underbrace{(A_t^L n_t^L N_t)^\nu}_{z_t^L}, \quad (\text{D.6})$$

$$\mathcal{K}_t(n_t^L; A_t^K)^\kappa \equiv \omega_K K_t^\kappa + (1 - \omega_K) \underbrace{[A_t^K (1 - n_t^L) N_t]^\kappa}_{z_t^K}. \quad (\text{D.7})$$

Substituting (D.3), (D.4), (D.6), and (D.7) into (D.5), canceling the common factor $Y_t^{1-\rho}$, and multiplying by N_t gives one equation for the equilibrium labor share n_t^L :

$$\begin{aligned} & \alpha(1 - \omega_L)(A_t^L N_t)^\nu (n_t^L)^{\nu-1} \mathcal{L}_t(n_t^L; A_t^L)^{\rho-\nu} \\ &= (1 - \alpha)(1 - \omega_K)(A_t^K N_t)^\kappa (1 - n_t^L)^{\kappa-1} \mathcal{K}_t(n_t^L; A_t^K)^{\rho-\kappa}. \end{aligned} \quad (\text{D.8})$$

Under the usual concavity and interiority conditions, this equation determines the equilibrium sectoral labor allocation. Equivalently, n_t^L maximizes final output over feasible labor allocations. Note that, in general it must be solved implicitly.

It is useful to restate (D.8) in the per-capita units in which the model is actually estimated. Both intermediate technologies in (D.1) and (D.2) are linearly homogeneous in their two arguments, so, writing land per capita as $\ell_t = L_t/N_t$ and non-landed capital per capita as $k_t = K_t/N_t$, the population level factors straight out of each composite:

$$\mathcal{L}_t(n_t^L; A_t^L) = N_t \mathcal{L}_t(\ell_t; z_t^L), \quad \mathcal{K}_t(n_t^L; A_t^K) = N_t \mathcal{K}_t(k_t; z_t^K), \quad (\text{D.9})$$

where $z_t^L = A_t^L n_t^L$ and $z_t^K = A_t^K (1 - n_t^L)$ are the labor-efficiency composites underbraced in (D.6) and (D.7), and where

$$\mathcal{L}_t(\ell_t; z_t^L) \equiv \left(\omega_L \ell_t^\nu + (1 - \omega_L) (z_t^L)^\nu \right)^{1/\nu}, \quad \mathcal{K}_t(k_t; z_t^K) \equiv \left(\omega_K k_t^\kappa + (1 - \omega_K) (z_t^K)^\kappa \right)^{1/\kappa} \quad (\text{D.10})$$

are exactly the per-capita nests of the main text. Substituting (D.9) into (D.8), the left-hand side carries $N_t^\nu N_t^{\rho-\nu} = N_t^\rho$ and the right-hand side carries $N_t^\kappa N_t^{\rho-\kappa} = N_t^\rho$; both sides are thus homogeneous of degree ρ in the population level. Dividing through by N_t^ρ leaves the equilibrium condition entirely in per-capita terms:

$$\begin{aligned} & \alpha(1 - \omega_L)(A_t^L)^\nu (n_t^L)^{\nu-1} \mathcal{L}_t(\ell_t; z_t^L)^{\rho-\nu} \\ &= (1 - \alpha)(1 - \omega_K)(A_t^K)^\kappa (1 - n_t^L)^{\kappa-1} \mathcal{K}_t(k_t; z_t^K)^{\rho-\kappa}, \end{aligned} \quad (\text{D.11})$$

or, collecting the productivity and allocation terms back into the composites themselves,

$$\alpha(1 - \omega_L) A_t^L (z_t^L)^{\nu-1} \mathcal{L}_t(\ell_t; z_t^L)^{\rho-\nu} = (1 - \alpha)(1 - \omega_K) A_t^K (z_t^K)^{\kappa-1} \mathcal{K}_t(k_t; z_t^K)^{\rho-\kappa}. \quad (\text{D.12})$$

The scale of the economy has dropped out: the equilibrium sectoral labor allocation depends on the aggregate state only through land per capita, capital per capita, and the two sectoral productivities. This is precisely what makes the two-sector re-interpretation compatible with the estimation of the main text, where the technology (6) and the factor-share conditions beginning with (18) are all written in (ℓ_t, k_t) . It does *not*, however, mean that the population level is irrelevant for the labor allocation. Because land is fixed at $L_t = \bar{L}$, N_t continues to enter (D.11) through $\ell_t = \bar{L}/N_t$ — the same Malthusian channel that operates in the one-sector model.

Now, absorb the equilibrium sectoral labor allocation into the two labor-efficiency composites:

$$z_t^L \equiv A_t^L n_t^L, \quad z_t^K \equiv A_t^K (1 - n_t^L). \quad (\text{D.13})$$

Substitution into the intermediates and then the final goods producer's technology yields the equilibrium technology

$$Y_t(L_t, K_t, N_t) = \left\{ \alpha \left[\omega_L L_t^\gamma + (1 - \omega_L) (z_t^L N_t)^\gamma \right]^{\rho/\gamma} + (1 - \alpha) \left[\omega_K K_t^\kappa + (1 - \omega_K) (z_t^K N_t)^\kappa \right]^{\rho/\kappa} \right\}^{1/\rho}, \quad (\text{D.14})$$

which converges to the Cobb-Douglas outer-nest aggregator as $\rho \rightarrow 0$.

D.2 Implied Productivity Growth Rates

Given this split it is clear that aggregate output from the two-sector economy takes the form of the main text's equations (1) or (2), with the productivities now embedding the sectoral labor shares — $z_t^L \equiv A_t^L n_t^L$ and $z_t^K \equiv A_t^K (1 - n_t^L)$. Under the multi-sector interpretation the A 's can be thought of as the true underlying sectoral labor-augmenting productivities. The growth rates of z_t^L and z_t^K thus mask both a reallocation effect and sector-specific, Harrod-neutral technical change (Harrod 1939):

$$\tilde{g}_{z^L} = \tilde{g}_{A^L} + \tilde{g}_{n^L}, \quad \tilde{g}_{z^K} = \tilde{g}_{A^K} + \tilde{g}_{1-n^L}. \quad (\text{D.15})$$

As industrialization occurs labor transitions out of agriculture, so that we expect $\tilde{g}_{n^L} < 0$ during this process. If we were to observe z_t^L declining, then we could conclude that either both A_t^L and n_t^L are falling *or* the labor share is falling faster than the underlying agricultural productivity is improving — $|\tilde{g}_{n^L}| > \tilde{g}_{A^L}$. Continuing to invoke the multi-sector re-interpretation here, this is actually testable, as we can recover A_t^L given estimates for z_t^L and data on the agricultural labor share, n_t^L .

Of course for quantitative recovery to be feasible, we must be able to map the multi-sector model back to the single-sector one. Since land *only* features in $\mathcal{L}_t$ and capital *only* features in $\mathcal{K}_t$, the factor share equations in (18) through (21) of the main text are still valid, with n_t^L entering only via $\mathcal{L}_t(\ell_t; A_t^L n_t^L)$ and $\mathcal{K}_t(k_t; A_t^K (1 - n_t^L))$. Thus, regardless of the single- or multi-sector interpretation, if one estimates the model treating z_t^L and z_t^K as residuals, A_t^L and A_t^K can still be recovered

pending data for n_t^L .

Further, note that in the limit as $\rho \rightarrow -1$ (perfect substitutes in the outer tier) the outer aggregator becomes $Y_t \rightarrow \alpha \mathcal{L}_t + (1 - \alpha) \mathcal{K}_t$, and the economy exhibits a corner solution: all labor is allocated to the sector with higher shadow value of the marginal product. This nests a version of Hansen and Prescott (2002) in which the Malthusian and Solow technologies are CES rather than Cobb-Douglas (and the Malthusian technology does not include capital). The novelty of our approach is that, if you choose to take a multi-sector interpretation, our model provides a theory of *smooth*, continuous transition from land-intensive to capital-intensive production when $\rho < 1$.

In the remainder of this appendix, we outline how to identify n_t^L from the Clark (2010) data. Then, we discuss identification of A_t 's given z_t 's (from the Bayesian estimators in the main text) and n_t^L , reporting various results under the multi-sector re-interpretation of our main-text estimations. After this, we re-estimate the model, leveraging information we have for n_t^L *directly* in the Hamiltonian Monte Carlo posterior estimation routine and assess whether such information passed directly to our estimator (as opposed to being used to indirectly infer the values of A_t 's) changes anything regarding inferences: it does not.

D.3 Indirectly Inferring Sector-level Productivities Given z_t 's

We take n_t^L directly from the data. Clark (2010) reports the share of the labor force in the primary sector by decade from 1200 to 1860 (Table 1), which we use as the labor share of the land-intensive sector. That series is assumed constant at 0.60 prior to 1680 and, from 1750 onward, rests on benchmark employment estimates interpolated between benchmark years. With n_t^L in hand and posterior estimates for z_t 's from the procedures outlined in the main text, the implied sector-level productivities are as follows:^{D2}

$$A_t^L = \frac{z_t^L}{n_t^L}, \quad A_t^K = \frac{z_t^K}{1 - n_t^L}. \quad (\text{D.16})$$

We back out A_t 's from our preferred model with $\rho = 0$ and $\nu, \kappa < 0$, using Clark (2010) data for n_t^L . Note that estimates for n_t^L from Clark (2010) are essentially constant at $n_t^L = 0.6$ from 1200 until ≈ 1650 , with a decline from 0.6 to < 0.3 by 1860. The decompositions in (D.15) thus are only non-trivial after 1650.

Figure D.1 reports the growth decompositions from (D.15), as implied by the procedure outlined above on estimates of z_t 's from the preferred model. It is apparent that labor reallocation plays only a minimal role in driving variation in z_t^L and z_t^K over time. In fact the correlation between $\tilde{g}_{z_t^L}$ and $\tilde{g}_{A_t^L}$ is ≈ 0.82 , with the correlation between $\tilde{g}_{z_t^L}$ and $\tilde{g}_{n_t^L}$ being ≈ 0.53 . On the capital productivity side of things, growth in A_t^K dominates with the correlation between $\tilde{g}_{z_t^K}$ and $\tilde{g}_{A_t^K}$ almost perfect at ≈ 0.99 , while the correlation between $\tilde{g}_{z_t^K}$ and $\tilde{g}_{1-n_t^L}$ at ≈ 0.09 . We conclude that, as implied by this particular exercise wherein we take the posterior estimates of z_t 's as given,

^{D2}Note, under the procedure outlined here, we take estimates for z_t 's as given and back out the A_t 's given n_t^L 's: we do not completely re-estimate the model using the information that n_t^L provides; rather, we do that in the next section.

backing out A_t 's given data on n_t^L , missing the effects of labor reallocation does not seriously bias our single-sector model's inferences.

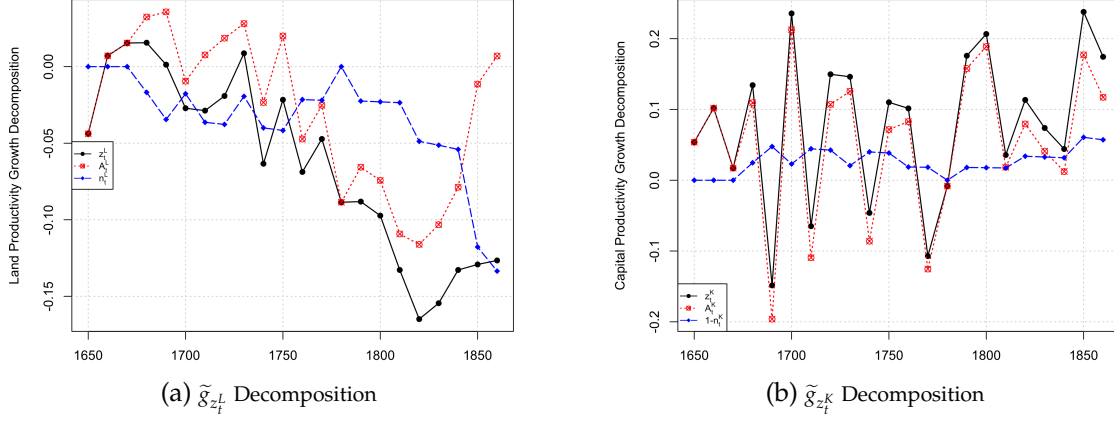


Figure D.1: This figure shows the land and capital productivity growth decompositions from expression (D.15) under the multi-sector re-interpretation. The reallocation of labor from land-intensive to capital-intensive activities has only a minimal effect on changes in z_t 's from 1650 to 1860.

D.4 The Labor Allocation Implied by Wage Equalization

The decomposition in Technical Appendix D.3 takes the allocation n_t^L as given and reads the sectoral productivities off of it. The equilibrium condition (D.11) runs in the opposite direction: given productivities and per-capita factor endowments, it *determines* n_t^L . Because labor mobility is nowhere imposed in the single-sector estimation — the posterior in the main text is identified entirely from the factor-share conditions — comparing the allocation the condition implies against the Clark (2010) primary-sector share is an over-identifying check on the multi-sector re-interpretation.

We run the check draw by draw. For each of the 4,000 posterior draws of $(\alpha, \omega_L, \omega_K, \nu, \kappa)$ from the preferred $\rho = 0, \nu, \kappa < 0$ model, and for each of the 67 decades, we hold (A_t^L, A_t^K) at the values backed out in (D.16) and solve (D.11) for n_t^L . The solution is well behaved. With $\nu, \kappa < 0$ the left-hand side of (D.11) is strictly decreasing in n_t^L and diverges as $n_t^L \rightarrow_+ 0$, while the right-hand side is strictly increasing and diverges as $n_t^L \rightarrow_- 1$, so an interior root exists and is unique at every draw. A bracketing root-finder locates it in all $4,000 \times 67$ draw-decade cells, with a maximum absolute residual of order 10^{-13} .

The implied allocation tracks the Clark (2010) share closely. Posterior means correlate with it at 0.95 in levels and 0.82 in decadal log changes. The mean absolute gap is 0.059 and the RMSE 0.065, against a mean posterior standard deviation of 0.061; the Clark (2010) share lies inside the 95% posterior credible interval of the model-implied share in 66 of the 67 decades, the single exception being 1860, where the reported value of 0.280 falls just below the lower bound of 0.282. The shape of the path is also right: the model-implied share fluctuates in a narrow band between

0.51 and 0.58 until the 1680s and then declines steadily to 0.32 by 1860, so the condition locates the onset of the reallocation out of agriculture in the same decades that the data do.

Two systematic discrepancies remain. First, there is a level gap before 1650, where the Clark (2010) share is 0.600 against a model-implied average of 0.529. Part of this gap is an artifact of the source series, which is assumed constant at 0.600 for every decade from 1200 to 1670; the model instead reads modest decade-to-decade movement over the same window, and no comparison of levels there is very informative. Second, the model generates a smaller total reallocation: from its peak to 1860 the Clark (2010) share falls by 0.320 while the model-implied share falls by 0.256, about four-fifths as much. The two-sector reading thus reproduces the direction and timing of the movement of labor out of agriculture, but understates its magnitude.

We stress that this is a coherence check on the two-sector representation, not an independent validation of n_t^L . The productivities entering the exercise are themselves constructed from the Clark (2010) share through (D.16), so that series is not withheld from the calculation; and the posterior intervals are wide — a posterior standard deviation of 0.061 against a series whose whole range is 0.28 to 0.60 — so a failure to reject is weak evidence. What the exercise does establish is that a condition playing no role in estimation has a unique interior solution at every posterior draw, and that the solution has the right level, the right turning point, and the right direction of travel. Had (D.11) implied a corner allocation, or a share rising across the industrial period, the multi-sector re-interpretation would be in trouble. It does not.

D.5 Re-estimating the Posterior Treating n_t^L as Data

In this appendix we re-estimate all six of the parameterizations, including all four $\rho = 0$ parameterizations from the main text and the alternative parameterizations in Technical Appendix C.1, letting information present in variation in n_t^L contribute to the posterior parameter estimates. That is, the residuals of the models we estimate here are $\{A_t^L, A_t^K\}_t$ not $\{z_t^L, z_t^K\}_t$, so that the likelihood internalizes variation in $\{n_t^L\}_t$ over time. Table D.1 shows the estimates of the fixed parameters under this estimation procedure. Few (if any) qualitative differences exist relative to the primary specifications featured in both Table 1 of the main text and Technical Appendix C.1. Further, fitness and robustness results are similar, as are productivity estimates from the preferred model before and after the decomposition (see Figure D.3).

Table D.1: Production Function Estimates, Multi-Sector Prior; England + Wales, 1200-1860

	$\rho = 0$				$\rho \neq 0$	
	(1) $\nu, \kappa < 0$	(2) $\nu < 0, \kappa \in (0, 1)$	(3) $\nu \in (0, 1), \kappa < 0$	(4) $\nu, \kappa \in (0, 1)$	(5) $\rho \in (0, 1); \nu, \kappa < 0$	(6) $\rho < 0; \nu, \kappa \in (0, 1)$
ν	-59.578 (3.9328)	-102.897 (3.6854)	0.962 (0.0005)	0.974 (0.0004)	-3.344 (1.4920)	0.974 (0.0013)
κ	-2.069 (0.0186)	0.969 (0.0005)	-6.589 (0.0905)	0.974 (0.0003)	-1.598 (0.1519)	0.939 (0.0196)
ω_L	0.500 (0.006)	0.511 (0.006)	0.486 (0.006)	0.488 (0.006)	0.167 (0.155)	0.032 (0.002)
ω_K	0.517 (0.006)	0.471 (0.006)	0.506 (0.005)	0.465 (0.006)	0.384 (0.010)	0.134 (0.161)
α	0.456 (0.002)	0.314 (0.001)	0.639 (0.001)	0.463 (0.001)	0.534 (0.007)	0.478 (0.002)
ρ	0	0	0	0	0.804 (0.1346)	-1.387 (0.2189)
WAIC ^a	60.868	149.590	76.553	174.500	-60.114	87.383

^a Widely applicable info. criterion, smaller = better (Watanabe 2010; Vehtari, Gelman, and Gabry 2017).

D.5.1 Model Fitness of $\hat{Y}_t$ After Re-estimation (All Models)

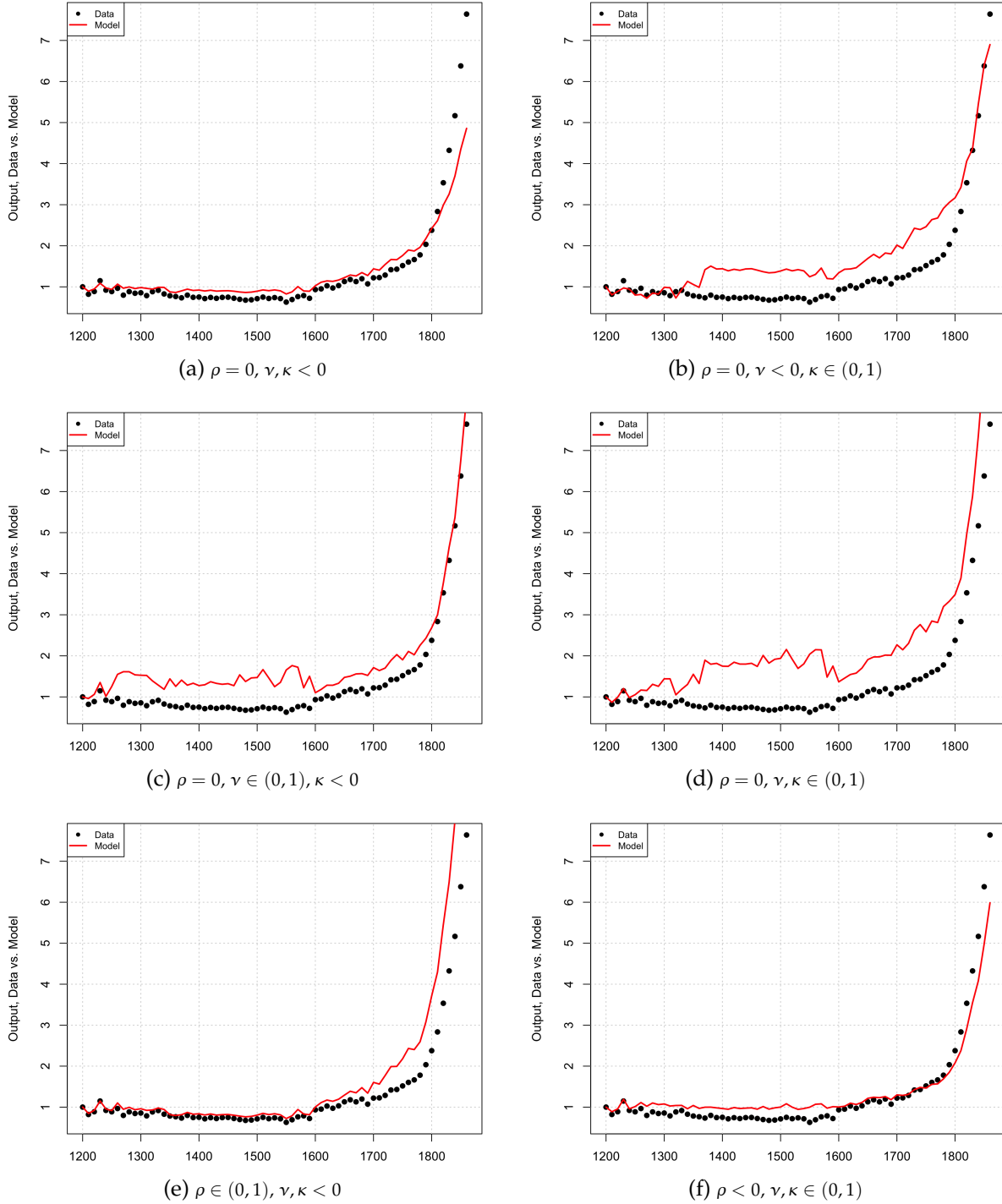


Figure D.2: This figure presents the fitness for non-targeted $\{\hat{Y}_t\}_{t=1200}^{1860}$ for all multi-sector models, re-estimated featuring n_t^L as data.

D.5.2 Productivity Growth Decomposition After Re-estimation (Preferred Model)

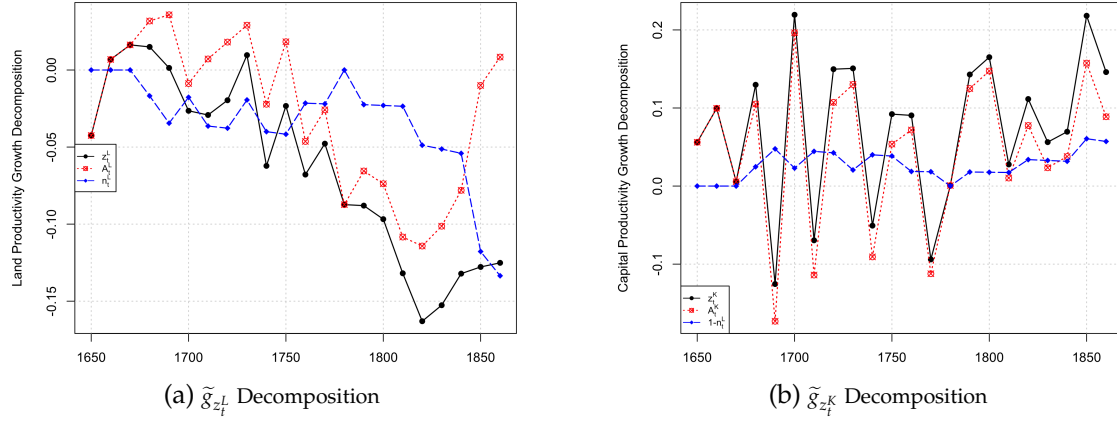


Figure D.3: This is the analog of Figure D.1 but for the $\rho = 0, \nu, \kappa < 0$ model with re-estimated parameters. Notice the compositions are almost identical, so re-estimation does not meaningfully affect inferences.

E Comparing Clark (2010) Factor Share Data with Allen (2007)

To address concerns regarding the provenance of Clark (2010) labor share estimates, which are constructed using aggregate national accounting procedures, in this appendix we turn to the micro dataset underlying the work in Allen (1992) and Allen (2007). Allen (2007) features both wage and price indices using actual employment, time use, and expenditure data from 1450-1850. In this appendix we seek to understand whether the productivity trends we estimate survive when the Allen (2007) data is used to construct the factor-income shares: they do.

How do the Clark (2010) and Allen (2007) datasets compare? Clark (2010) reports nominal factor income shares for land, labor, and non-landed capital for England at decadal frequency, 1200–1860 (his Table 13), together with population, farmland rents, and the rate of return on capital (his Table 7). The Clark (2010) shares are the output of a single, internally consistent national-accounting exercise: the land, labor, and capital shares are constructed jointly so as to exhaust national income. While Clark’s data are constructed using aggregate methods, Allen (1992) use manorial accounts, probate inventories, and market records to assemble annual price and wage series for England and the south midlands from 1450 to 1850. These include separate input-price indices for land (rents per acre, by land type and tenure), labor (day-wages), and capital (implied rental rates), a consumer price index, and a real-rent (land-return) series. Allen’s data differ from Clark’s in several ways that make them an informative external check. First, they are prices, not shares. Allen measures the rental price of an acre and the wage of a day’s labor directly, with no national-accounting identity imposed; income shares must be reconstructed from them by combining prices with quantities. Second, Allen’s rents and wages rest on a different documentary base than Clark’s aggregates and are not constructed to exhaust ex-ante known national income, but represent true aggregation over micro data.

Because Allen reports factor prices rather than shares, we map his series into a counterfactual income share comparable with Clark’s. Under cost-share accounting, the land/labor income-share ratio satisfies

$$\underbrace{\frac{\theta_t^L}{\theta_t^N}}_{\text{Allen-implied share ratio}} = \frac{r_t^L L_t}{w_t N_t} = \underbrace{\frac{r_t^L}{w_t}}_{\text{rent/wage from Allen}} \times \underbrace{\frac{L_t}{N_t}}_{\text{land per worker from Clark}}. \quad (\text{E.1})$$

Letting the non-landed capital share be that from Clark (2010), equation (E.1) along with the simplex normalization on shares gives us two equations in two unknowns, so we get Allen-implied estimates for both θ_t^L and θ_t^N .

The imputed share differences are small, with averages (land and labor) at ≈ 0.026 over the 1450–1850 sample and maximum absolute differences of ≈ 0.077 . Figure E.1 shows the Clark (2010) land and labor shares plotted against the imputation of land/labor shares using Allen (1992) and Allen (2007). Notice the time series track each other closely over the overlapping sample period. At first glance Figure E.1 does *not* suggest any suspicion is warranted regarding utilization

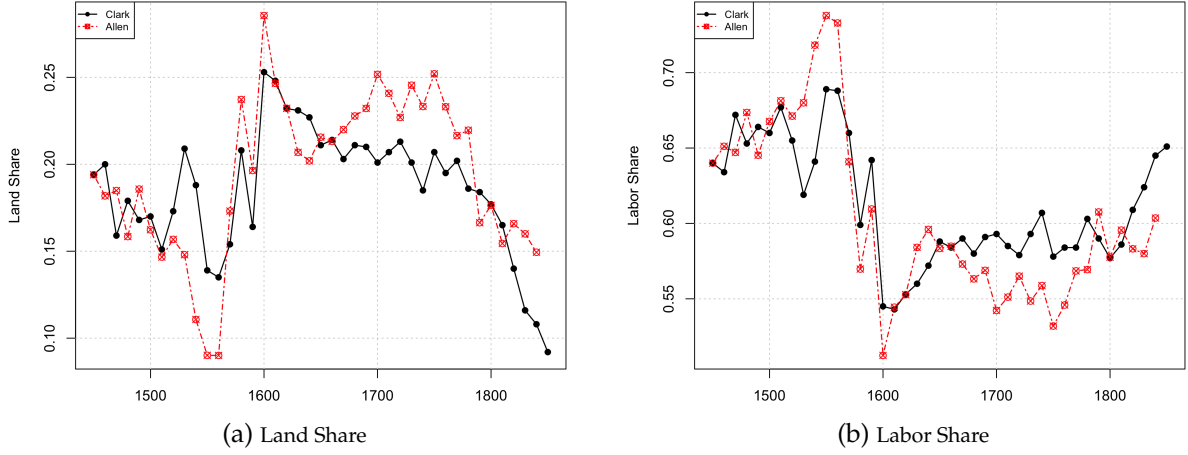


Figure E.1: We compare Clark (2010) and Allen (2007) land and labor share time series (1450–1850) following the imputation procedure outlined in this appendix.

of share data from Clark (2010) for more granular model inferences.

To understand if productivity trends significantly change when using the Allen (2007) data, we then take the posterior estimated parameters from the preferred $\rho = 0, \nu, \kappa < 0$ model and simulate the implied posterior mean paths of $\{z_t^L, z_t^K\}_{t=1450}^{1850}$ using the Allen-derived factor shares. Figure E.2 shows the Clark-implied posterior means in black against the Allen-implied means in red; the time series are practically identical.

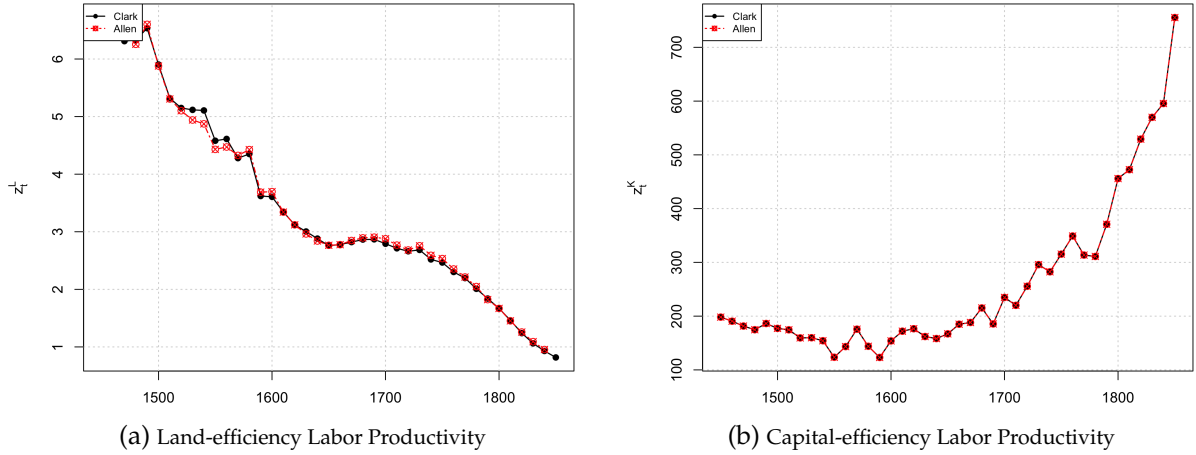


Figure E.2: We compare productivities as backed out of the original estimator ($\rho = 0; \nu, \kappa < 0$) using Clark (2010) data against those implied by using Allen (2007) data over the time series 1450 to 1850.

F Growth-Accounting Bias: TFP vs. Population Growth Contours

This appendix supplements the main text's discussion of the growth properties of the Cobb-Douglas special case ($\rho = \nu = \kappa = 0$) from Section 2.5, illustrating graphically how the relative contributions of TFP growth and population growth to per-capita output growth depend on the structural elasticities α , ω_L , and ω_K . Steady-state growth in output per capita, g_y , decomposes into a linear function of TFP growth, g_A , and population growth, g_N :^{F1}

$$g_y = \underbrace{\frac{1}{1 - (1 - \alpha)\omega_K} g_A}_{\text{TFP Growth Contribution}} - \underbrace{\frac{\alpha\omega_L}{1 - (1 - \alpha)\omega_K} g_N}_{\text{Pop. Growth Contribution}}. \quad (\text{F.1})$$

All growth rates are expressed in logged differences (e.g., $g_y = \ln y_t - \ln y_{t-1}$). Population growth g_N enters expression (F.1) negatively, while TFP growth g_A enters positively.

Figure F.1 plots contours of g_y over the space of g_A (horizontal axis) and g_N (vertical axis) for several combinations of α and ω_L , holding $\omega_K = 0.33$. Green regions denote $g_y > 0$; red regions denote $g_y < 0$. The sign of g_y turns on $\alpha\omega_L$ alone: from (F.1), $g_y > 0$ if and only if $g_A > \alpha\omega_L g_N$, so ω_K scales the magnitude of per-capita growth without moving the boundary between the green and red regions. The main text's result that TFP growth dominates when TFP and population grow proportionally is evident from the green values along the 45-degree line in each panel. However, if production is more land-intensive (α close to one), high population growth coupled with low TFP growth pushes $g_y < 0$; likewise, even for smaller α , a high land intensity within the land/labor nest (ω_L close to one) drives $g_y < 0$ unless TFP growth is sufficiently high. Because a classic Solow accounting exercise omits land as a fixed factor, it understates this competition between TFP and population growth — the source of the downward TFP-measurement bias discussed in the main text.

^{F1}This can be seen by taking logged first differences of equation (27) from the main text and expanding out the corresponding coefficients.

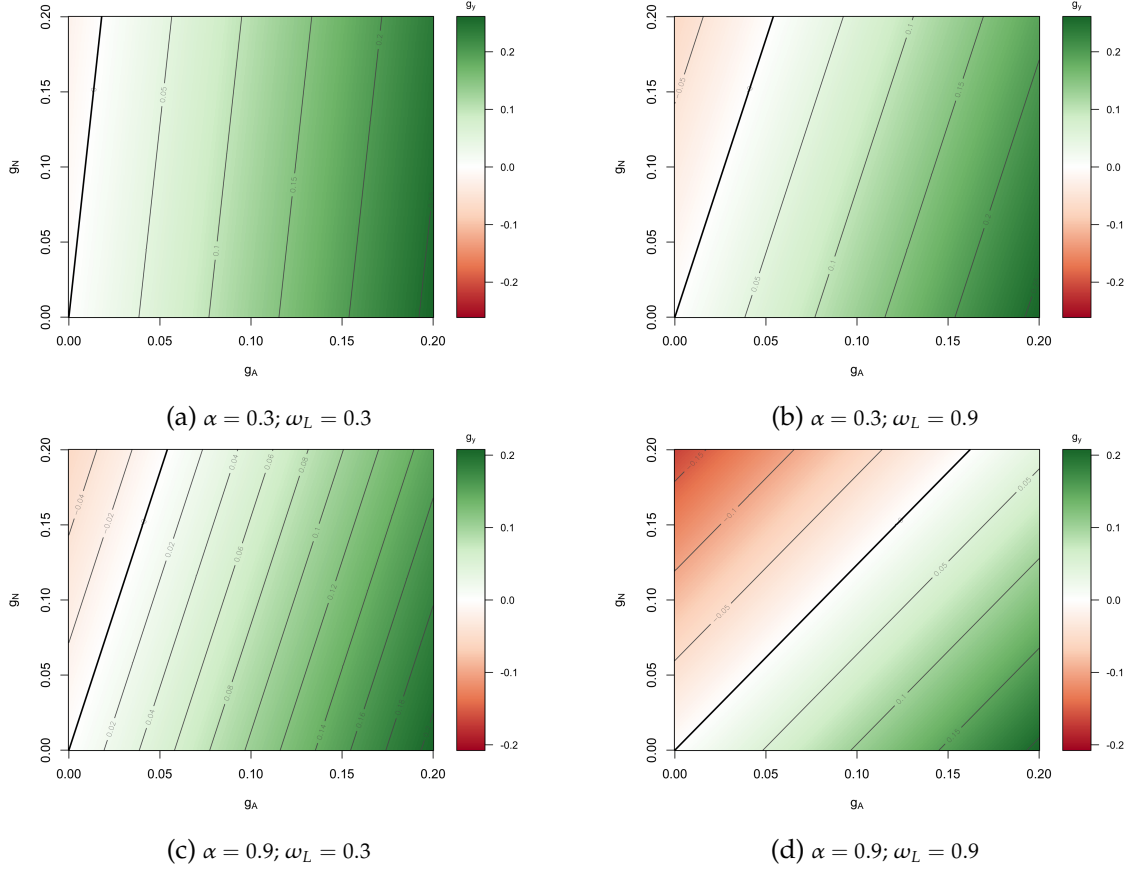


Figure F.1: Contours of g_y (growth in output per capita) for $\omega_K = 0.33$ and different values of α and ω_L . The population growth rate, g_N , is on the vertical axis; the TFP growth rate, g_A , is on the horizontal axis. Green values denote $g_y > 0$; red values denote $g_y < 0$.

References

- Acemoglu, D., S. Johnson, and J. A. Robinson (2005a). "Institutions as a Fundamental Cause of Long-Run Growth". *Handbook of Economic Growth: Volume 1A*. Ed. by Philippe Aghion and Steven N. Durlauf. North Holland: Elsevier (cit. on pp. 3, 34).
- (2005b). "The Rise of Europe: Atlantic Trade, Institutional Change and Economic Growth". *American Economic Review* 95.3, 546–579 (cit. on pp. 3, 34).
- Allen, Robert C. (1992). *Enclosure and the Yeoman: The Agricultural Development of the South Midlands, 1450–1850*. Oxford University Press (cit. on pp. 26, 35, 73).
- (2007). *English Agricultural Prices*. Spreadsheet `enclyeo.xls` with documentation `datasources.doc`; agricultural output prices and land, labour, and capital input-price indices plus a real-rent TFP series for England and the south midlands. Data underlying Allen, *Enclosure and the Yeoman* (1992). URL: <https://www.nuffield.ox.ac.uk/people/sites/allen-research-pages/> (visited on 06/20/2026) (cit. on pp. 26, 73, 74).
- Alvarez-Cuadrado, Francisco and Markus Poschke (2011). "Structural Change Out of Agriculture: Labor Push versus Labor Pull". *American Economic Journal: Macroeconomics* 3.3, pp. 127–158. DOI: [10.1257/mac.3.3.127](https://doi.org/10.1257/mac.3.3.127) (cit. on p. 6).
- Bakker, Bas B., Florian Niesert, and Michael Seiferling (2023). *Unveiling the Hidden Impact of Urban Land Rents on Total Factor Productivity*. IMF Working Paper WP/23/170. Presents cross-country estimates of urban land rents as a share of GDP; Great Britain is estimated at roughly 21% (see Figure 4.1). International Monetary Fund. DOI: [10.5089/9798400252808.001](https://doi.org/10.5089/9798400252808.001). URL: <https://www.imf.org/en/publications/wp/issues/2023/08/25/unveiling-the-hidden-impact-of-urban-land-rents-on-total-factor-productivity-538514rt> (cit. on p. 5).
- Bar, Michael and Oksana Leukhina (2010a). "Demographic transition and industrial revolution: A macroeconomic investigation". *Review of Economic Dynamics* 13.2, pp. 424–451 (cit. on pp. 1, 2, 5, 9, 20, 23).
- (2010b). "The role of mortality in the transmission of knowledge". *Journal of Economic Growth* 15, pp. 291–321 (cit. on p. 20).
- Becker, Gary S, Kevin M Murphy, and Robert Tamura (1990). "Human capital, fertility, and economic growth". *Journal of political economy* 98.5, Part 2, S12–S37 (cit. on pp. 1, 2).
- Betancourt, Michael and Leo Stein (2011). "The Geometry of Hamiltonian Monte Carlo". *arXiv:1112.4118* (cit. on p. 28).
- Caunedo, Julieta and Elisa Keller (2021). "Capital Obsolescence and Agricultural Productivity". *The Quarterly Journal of Economics* 136.1, pp. 505–561. DOI: [10.1093/qje/qjaa028](https://doi.org/10.1093/qje/qjaa028) (cit. on pp. 6, 8, 24).
- Clark, Gregory (2010). *The macroeconomic aggregates for England, 1209–2008*. Ed. by A.J. Field. Vol. 27. Emerald Publishing Group Limited, Leeds UK, pp. 51–140 (cit. on pp. 3–5, 9, 21, 25–27, 29, 63, 67–69, 73, 74).

- Dennis, Benjamin N. and Talan B. Işcan (2009). "Engel versus Baumol: Accounting for Structural Change Using Two Centuries of U.S. Data". *Explorations in Economic History* 46.2, pp. 186–202. DOI: [10.1016/j.eeh.2008.10.002](https://doi.org/10.1016/j.eeh.2008.10.002) (cit. on p. 6).
- Duffy, John and Chris Papageorgiou (2000). "A cross-country empirical investigation of the aggregate production function specification". *Journal of Economic Growth* 5, pp. 87–120 (cit. on p. 7).
- Eaton, Jonathan and Samuel Kortum (1996). "Measuring technology diffusion and the international sources of growth". *Eastern economic journal* 22.4, pp. 401–410 (cit. on p. 10).
- Galor, Oded and Omer Moav (2002). "Natural Selection and the Origin of Economic Growth". *Quarterly Journal of Economics* (cit. on pp. 2, 10).
- Galor, Oded and David N. Weil (2000). "Population, Technology and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond". *American Economic Review* 90, pp. 806–828 (cit. on pp. 2, 10).
- Gelman, Andrew et al. (2013a). "Computation in R and Stan". *Bayesian Data Analysis*. Third. Chapman & Hall. Chap. Appendix C (cit. on p. 28).
- (2013b). "Computationally efficient Markov chain simulation". *Bayesian Data Analysis*. Third. Chapman & Hall. Chap. 12 (cit. on p. 28).
- Gollin, Douglas and Joseph P. Kaboski (2023). "New Views of Structural Transformation: Insights from Recent Literature". *Oxford Development Studies* 51.4, pp. 339–361. DOI: [10.1080/13600818.2023.2280748](https://doi.org/10.1080/13600818.2023.2280748). URL: <https://www.tandfonline.com/doi/full/10.1080/13600818.2023.2280748> (cit. on p. 8).
- Gollin, Douglas, David Lagakos, and Michael E. Waugh (2014). "The Agricultural Productivity Gap". *The Quarterly Journal of Economics* 129.2, pp. 939–993 (cit. on p. 23).
- Gollin, Douglas, Stephen L. Parente, and Richard Rogerson (2002). "The Role of Agriculture in Development". *American Economic Review* 92.2, pp. 160–164. DOI: [10.1257/000282802320189177](https://doi.org/10.1257/000282802320189177) (cit. on p. 23).
- (2004). "Farm Work, Home Work, and International Productivity Differences". *Review of Economic Dynamics* 7.4, pp. 827–850. DOI: [10.1016/j.red.2004.05.003](https://doi.org/10.1016/j.red.2004.05.003) (cit. on p. 23).
- (2007). "The Food Problem and the Evolution of International Income Levels". *Journal of Monetary Economics* 54.4, pp. 1230–1255. DOI: [10.1016/j.jmoneco.2006.04.002](https://doi.org/10.1016/j.jmoneco.2006.04.002) (cit. on p. 23).
- Gollin, Douglas and Richard Rogerson (2025). "Structural Change and Macro Development: Beyond the One-Sector Growth Model". *Oxford Review of Economic Policy* 41.2, pp. 291–312. DOI: [10.1093/oxrep/graf024](https://doi.org/10.1093/oxrep/graf024). URL: <https://academic.oup.com/oxrep/article-abstract/doi/10.1093/oxrep/graf024/8257194> (cit. on pp. 8, 23).
- Goodfriend, Marvin and John McDermott (1995). "Early Development". *American Economic Review* 85.1, pp. 116–133 (cit. on pp. 2, 10).
- Grossman, Gene M and Elhanan Helpman (1991). "Trade, knowledge spillovers, and growth". *European economic review* 35.2-3, pp. 517–526 (cit. on p. 10).

- Growiec, Jakub and Jakub Mućk (2020). "Isoelastic elasticity of substitution production functions". *Macroeconomic Dynamics* 24.7, pp. 1597–1634 (cit. on p. 12).
- Hansen, Gary D. and Edward C. Prescott (2002). "Malthus to Solow". *American Economic Review* 92.4, pp. 1205–1217 (cit. on pp. 1, 2, 5, 7–9, 15, 23, 37, 67).
- Harrod, R. F. (1939). "An Essay in Dynamic Theory". *The Economic Journal* 49.193, pp. 14–33 (cit. on p. 66).
- Herrendorf, Berthold et al. (2015). "Sectoral Technology and Structural Transformation". *American Economic Journal: Macroeconomics* 7.4, pp. 104–133 (cit. on pp. 5, 25, 26, 29).
- Hicks, John R (1932). "Marginal productivity and the principle of variation". *Economica* 35, pp. 79–88 (cit. on pp. 12, 13).
- Hodgson, Geoffrey M. (2017). "1688 and all that: property rights, the Glorious Revolution and the rise of British capitalism". *Journal of Institutional Economics* 13.1, pp. 79–107 (cit. on p. 34).
- Huneus, Federico and Richard Rogerson (2024). "Heterogeneous Paths of Industrialization". *The Review of Economic Studies* 91.3, pp. 1746–1774. DOI: [10.1093/restud/rdad053](https://doi.org/10.1093/restud/rdad053) (cit. on p. 6).
- Jamieson, Ross W. (2001). "The Essence of Commodification: Caffeine Dependencies in the Early Modern World". *Journal of Social History* 35.2, pp. 269–294 (cit. on pp. 3, 35).
- Jones, Charles and Dietrich Vollrath (2013). *INTRODUCTION TO ECONOMIC GROWTH 3rd ed* (cit. on pp. 7, 21).
- Jones, Charles I (1995). "R & D-based models of economic growth". *Journal of political Economy* 103.4, pp. 759–784 (cit. on p. 10).
- Jones, Charles I. (2001). "Was the Industrial Revolution Inevitable? Economic Growth over the Very Long Run". *Advances in Macroeconomics* 1 (cit. on pp. 2, 10).
- Jones, Charles I (2002). "Sources of US economic growth in a world of ideas". *American economic review* 92.1, pp. 220–239 (cit. on p. 10).
- Juselius, Mikael (2008). "Long-run relationships between labor and capital: Indirect evidence on the elasticity of substitution". *Journal of Macroeconomics* 30.2, pp. 739–756 (cit. on pp. 5, 11, 33).
- Kaldor, Nicholas (1961). "Capital Accumulation and Economic Growth". *Capital Accumulation and Economic Growth*. Ed. by Friedrich Lutz and Douglas Hague. MacMillan and Co. Ltd. Chap. 10, pp. 177–222 (cit. on p. 7).
- Karabarbounis, Loukas and Brent Neiman (Oct. 2013). "The Global Decline of the Labor Share*". *The Quarterly Journal of Economics* 129.1, pp. 61–103. ISSN: 0033-5533. DOI: [10.1093/qje/qjt032](https://doi.org/10.1093/qje/qjt032). eprint: <https://academic.oup.com/qje/article-pdf/129/1/61/30626867/qjt032.pdf>. URL: <https://doi.org/10.1093/qje/qjt032> (cit. on pp. 5, 9, 11).
- Katz, Arnold J. (2017). *Imputing Rents to Owner-Occupied Housing by Directly Modelling Their Distribution*. BEA Working Paper WP2017-7. Reports BEA estimates of net "space rent" (the land component of housing rents) as a share of personal income; see Figure 2 for the time series (about 3% in 2013). U.S. Bureau of Economic Analysis. URL: <https://www.bea.gov/sites/default/files/papers/WP2017-7.pdf> (cit. on p. 5).

- Knoblach, Michael, Martin Roessler, and Patrick Zwerschke (2019). "The Elasticity of Substitution Between Capital and Labour in the US Economy: A Meta-Regression Analysis". *Oxford Bulletin of Economics and Statistics* (cit. on pp. 5, 11, 33).
- Lewandowski, Daniel, Dorota Kurowicka, and Harry Joe (2009). "Generating random correlation matrices based on vines and extended onion method". *Journal of Multivariate Analysis* 100.9 (cit. on p. 26).
- Lucas Jr, Robert E (1988). "On the mechanics of economic development". *Journal of monetary economics* 22.1, pp. 3–42 (cit. on p. 10).
- Maddison Project (2020). *Maddison Project Database 2020*. URL: <https://www.rug.nl/ggdc/historicaldevelopment/maddison/releases/maddison-project-database-2020?lang=en> (visited on 03/23/2024) (cit. on pp. 3, 4).
- Malthus, Thomas (1798). *An Essay on the Principle of Population*. J. Johnson, London (cit. on pp. 2, 17, 18).
- Mokyr, Joel (2002). *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton University Press. ISBN: 9780691120133. URL: <http://www.jstor.org/stable/j.ctt7rz25> (visited on 03/27/2025) (cit. on pp. 3, 35).
- Neal, Radford (2011). "MCMC using Hamiltonian dynamics". *Handbook of Markov Chain Monte Carlo*. Ed. by Steve Brooks et al. Chapman & Hall. Chap. 5 (cit. on pp. 25, 28).
- Neeson, J. M. (1993). *Commoners: Common Right, Enclosure and Social Change in England, 1700–1820*. Cambridge University Press (cit. on p. 35).
- Ngai, L. Rachel and Christopher A. Pissarides (2007). "Structural Change in a Multisector Model of Growth". *American Economic Review* 97.1, pp. 429–443. DOI: [10.1257/aer.97.1.429](https://doi.org/10.1257/aer.97.1.429) (cit. on p. 5).
- North, D. C. and B. R. Weingast (1989). "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England". *Journal of Economic History* 49.4, 803–832 (cit. on pp. 3, 34).
- North, Douglass C. (1981). *Structure and Change in Economic History*. W. W. Norton (cit. on p. 20).
- Ostrom, Elinor (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press (cit. on pp. 20, 35).
- Overton, Mark (1996). *Agricultural Revolution in England: The Transformation of the Agrarian Economy 1500–1850*. Cambridge University Press (cit. on p. 35).
- Raurich, Xavier, Hector Sala, and Valeri Sorolla (2012). "Factor shares, the price markup, and the elasticity of substitution between capital and labor". *Journal of Macroeconomics* 34.1, pp. 181–198 (cit. on pp. 5, 11, 33).
- Ravikumar, B. and Guillaume Vandenbroucke (2025). "On the Transition to Sustained Growth: The Importance of Recent Agricultural Employment". *Review of Economic Dynamics* 58 (cit. on pp. 1, 2, 5, 9, 23).

- Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu (2008). "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis". *Journal of Monetary Economics* 55.2, pp. 234–250. DOI: [10.1016/j.jmoneco.2007.11.006](https://doi.org/10.1016/j.jmoneco.2007.11.006) (cit. on p. 23).
- Romer, Paul M (1986). "Increasing returns and long-run growth". *Journal of political economy* 94.5, pp. 1002–1037 (cit. on p. 10).
- (1990). "Endogenous technological change". *Journal of political Economy* 98.5, Part 2, S71–S102 (cit. on p. 10).
- Solow, Robert M (1956). "A contribution to the theory of economic growth". *The quarterly journal of economics* 70.1, pp. 65–94 (cit. on pp. 1, 2, 19, 21).
- (1957). "Technical change and the aggregate production function". *The review of Economics and Statistics* 39.3, pp. 312–320 (cit. on pp. 1–3, 6, 19, 24).
- Stokey, Nancy L (1988). "Learning by doing and the introduction of new goods". *journal of Political Economy* 96.4, pp. 701–717 (cit. on p. 10).
- (1991). "Human capital, product quality, and growth". *The Quarterly Journal of Economics* 106.2, pp. 587–616 (cit. on p. 10).
- Tamura, Robert (1991). "Income convergence in an endogeneous growth model". *Journal of Political Economy* 99.3, pp. 522–540 (cit. on p. 10).
- (1998). "From agriculture to industry: Human capital and specialization". *Unpublished manuscript (Clemson University)* (cit. on p. 10).
- (2002). "Human capital and the switch from agriculture to industry". *Journal of Economic Dynamics and Control* 27.2, pp. 207–242 (cit. on pp. 2, 10).
- U.S. Bureau of Economic Analysis (2025). *Fixed Assets Interactive Data Tables*. <https://www.bea.gov/itable/fixed-assets>. Accessed: 2026-01-15. URL: <https://www.bea.gov/itable/fixed-assets> (cit. on p. 6).
- Vehtari, Aki, Andrew Gelman, and Jonah Gabry (2017). "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC". *Statistics and Computing* 27.5, pp. 1413–1432 (cit. on pp. 27, 28, 30, 48, 61, 70).
- Watanabe, Sumio (2010). "Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory". *Journal of Machine Learning Research* 11, pp. 3571–3594 (cit. on pp. 27, 28, 48, 61, 70).
- Withington, Phil (2020). "Where Was the Coffee in Early Modern England?" *The Journal of Modern History* 92, pp. 40–75 (cit. on pp. 3, 35).
- Wordie, J. R. (1983). "The Chronology of English Enclosure, 1500–1914". *The Economic History Review* 36.4, pp. 483–505. DOI: [10.1111/j.1468-0289.1983.tb00213.x](https://doi.org/10.1111/j.1468-0289.1983.tb00213.x) (cit. on p. 35).
- Yang, Dennis Tao and Xiaodong Zhu (2013). "Modernization of Agriculture and Long-Term Growth". *Journal of Monetary Economics* 60.3, pp. 367–382. DOI: [10.1016/j.jmoneco.2013.02.002](https://doi.org/10.1016/j.jmoneco.2013.02.002) (cit. on p. 6).
- Zhang, Shangfeng et al. (2022). "Measurement of factor price distortion: A new production function method with time-varying elasticity". *Technological Forecasting and Social Change* 175, p. 121363.

ISSN: 0040-1625. DOI: <https://doi.org/10.1016/j.techfore.2021.121363>. URL: <https://www.sciencedirect.com/science/article/pii/S0040162521007940> (cit. on p. 12).