A Technology-Gap Model of ‘Premature’ Deindustrialization

By Ippei Fujiwara and Kiminori Matsuyama*

Updated: June 7, 2024

Abstract

We propose a parsimonious mechanism for generating *premature deindustrialization* (PD). In the baseline model, the Baumol effect drives the hump-shaped path of the manufacturing share. Countries follow different paths due to the difference in the sector-specific adoption lags. The condition for PD under which countries differ only in *technology gap* implies that the cross-country productivity dispersion is the largest in agriculture. Moreover, when calibrated to match Rodrik (2016)’s findings, it is the smallest in manufacturing. In three extensions, we add the Engel effect, international trade, and catching-up by late industrializers, to demonstrate the robustness of the mechanism.
(*JEL classifications*: O11, O14, O33)

Keywords: Premature deindustrialization, technology gaps, adoption lags, The Baumol effect, The Engel effect, International Trade, Catching-up, Log-supermodularity

*Fujiwara: Faculty of Economics, Keio University (email: ippei.fujiwara@keio.jp) and Crawford School of Public Policy, Australian National University (email: ippei.fujiwara@anu.edu.au); Matsuyama: Department of Economics, Northwestern University (email: k-matsuyama@northwestern.edu). Arnaud Costinot was the co-editor for this article. We are grateful to Timo Boppart, Francisco Buera, Shinnosuke Kikuchi, Marti Mestieri, Rachel Ngai, Dani Rodrik, Richard Rogerson, Kei-Mu Yi, and seminar participants at (chronologically) Hitotsubashi Conference on International Trade and FDI, Kent, ES China, ES Australia, FRB-Chicago, Monash, Melbourne, ES European Summer, STEG, St. Gallen-Zürich, Oxford, Tokyo, FRB-Dallas, Peking, CIGS, Princeton, Sydney, Iowa State, Chicago, Keio, and Bank of Japan for their feedback. Much of this project was conducted during Matsuyama’s visits to Keio since December 2018, most recently as a visiting professor of its Super Global program in December 2023. And the final version was prepared during his visit at Becker Friedman Institute. Fujiwara acknowledges financial support from JSPS KAKENHI Grant-in-Aid for Scientific Research (A) No. 18H03638. The usual disclaimer applies.

1. Introduction

The share of the manufacturing sector, whether measured in employment or value-added, followed an inverted U-shaped or hump-shaped path over the course of development in most countries, as well-documented by Herrendorf, Rogerson, and Valentinyi (2014; see, e.g., Figure 6.1). Recently, Rodrik (2016) presented the finding that more recent industrializers entered the stage of deindustrialization at lower income levels with lower peak manufacturing shares compared to more advanced economies that had industrialized earlier; see, e.g., his Figures 5, 6, and 7. Rodrik (2016) focused on documenting and establishing these empirical regularities--now widely known as "premature deindustrialization," following the title of his paper—instead of proposing any formal model that captures a particular causal relation nor making any normative statements. Nevertheless, he speculated that globalization may be a cause. Moreover, the word, "premature," may seem to imply some sorts of inefficiencies that call for government interventions.[1] In contrast, we present in this paper a parsimonious mechanism for generating "premature" deindustrialization (hereafter PD) along the efficient equilibrium path of a closed economy, which is robust to opening up for international trade.

In our framework, there are three competitive sectors: agriculture, manufacturing, and services, which produce the consumption goods that are gross complements. The productivity of the frontier technology in each sector grows at an exogenously constant rate, which is the highest in agriculture, the lowest in services, with manufacturing in the middle. In the baseline model, the hump-shaped path of the manufacturing share, along with the declining agricultural share and the increasing service share, is driven solely by such productivity growth rate differences across the sectors, as in Baumol (1967) and Ngai and Pissarides (2007), but countries follow different hump-shaped paths of the manufacturing share due to the differences in their adoption lags in the three sectors. To simplify the exposition, we further assume that the countries differ only in one dimension; their ability to adopt the frontier technology, which we call "technology gap," following Krugman (1985). Unlike in Krugman, however, the technology gap has differential impacts on its adoption lags across sectors. In this setup, we investigate the sufficient and necessary condition under which PD occurs.

[1]Rodrik (2016), while following the terminology used by Dasgupta and Singh (2006) and Palma (2014), explicitly cautioned the reader against drawing any policy implications in his footnote 19. What distinguishes Rodrik's work from these and other works in the literature is its scope, as the latter mostly focus on a particular region, say Latin America or Sub-Saharan Africa, or a particular country, say India or Mexico.

To see the importance of the differential impacts of the technology gap on the adoption lags across sectors, suppose, for the moment, that the technology gap would affect its adoption lags in all sectors uniformly. Then, poorer countries with larger technology gaps reach their peaks later than richer countries, but their delays exactly make up for the larger adoption lags in all sectors. As a result, all countries follow the same path, reaching exactly the same peak manufacturing shares at exactly the same level of the per capita income. Only the timing is different. In other words, PD could not occur.

Instead, suppose that the technology gap leads to a longer adoption lag in services than in agriculture, but the productivity growth rate is sufficiently higher in agriculture than in services such that poorer countries with larger technology gaps are more lagged behind in agricultural productivity than in service productivity.[2] Then, poorer countries reach their peaks later in time than richer countries, but their delays are *not long enough* to make up for their longer adoption lags so that they reach the peaks at lower productivity levels in these two sectors. In other words, they reach their peaks "prematurely." Furthermore, when the impact of the technology gap on the adoption lag in manufacturing is not too large, their peak manufacturing shares stay lower than those in early industrializers. Under these conditions, the baseline model captures the three features of PD; that is, countries with larger technology gaps reach their manufacturing peaks later in time but earlier in per capita income with lower peak manufacturing shares.

In a nutshell, our mechanism suggests that the manufacturing shares among latecomers reach their peaks *later* due to their lower relative productivity in agriculture *but prematurely* due to the long adoption lag in services, while the peak manufacturing shares are *lower* because their productivity is not so behind in manufacturing. Indeed, these conditions for PD jointly imply that the cross-country productivity dispersion is the largest in agriculture, as empirically observed.[3] On the other hand, these conditions impose no restriction on the relative magnitude of the cross-country productivity dispersion between manufacturing and services. However, when our model

[2]This is because the productivity *level* in a sector is *log-submodular* (see e.g., Costinot 2009, Costinot and Vogel 2015) in its productivity *growth rate* and the adoption lag. In words, the *negative* impact of the adoption lag on the productivity *level* is *magnified* by the productivity *growth rate*; That is, even a short adoption lag matters a lot in a rapidly growing sector, while an even long adoption lag matters little in a slowly growing sector.

[3]There seems to be broad consensus on this. See, e.g., Caselli (2005), Restuccia, Yang, and Zhu (2008), and Gollin, Lagakos, and Waugh (2014a).

is calibrated to match the Rodrik's (2016; Table10) findings, the implied productivity dispersion is the smallest in manufacturing.[4]

We also consider three extensions of the baseline model. First, the baseline model assumes that structural change is driven solely by the Baumol effect. Most existing models of structural change, however, rely on the nonhomotheticity of sectoral demand compositions, the Engel effect for short, as the main driver behind the hump-shaped path of manufacturing. Indeed, it has been pointed out that the Baumol effect alone cannot account for many key features of structural change: see, e.g., Boppart (2014) and Comin, Lashkari and Mestieri (2021). Rodrik (2016, p.7) also noted that we need a combination of the Baumol and Engel effects. In view of the importance of the Engel effect as one of the drivers of structural change, we extend the baseline model by adding the Engel effect. As expected, adding the Engel effect on top of the Baumol effect significantly changes the shape of the time path, but it has little effects on the impacts of the peak values, hence on the mechanism of PD presented by the baseline model. Furthermore, if we had relied solely on the Engel effect without the Baumol effect, the technology gap generates PD only under the conditions that would imply, counterfactually, that the cross-country productivity dispersion is the largest in services. Second, we also extend the baseline model to allow for opening up for international trade.[5] Since one implication of PD in our mechanism is that the productivity dispersion across countries is larger in agriculture than in manufacturing, late industrializers have comparative advantage in manufacturing, so that opening up for trade enables them to export manufacturing, which weakens the mechanism, suggesting that PD occurs in spite of, not because of, international trade. This is broadly consistent with another finding by Rodrik; that is, East Asian countries, which grew through manufacturing exports, "suffer" less from PD. Finally, we extend the baseline model to allow for

[4]The conventional view, at least among the trade economists interested in explaining the Balassa-Samuelson hypothesis that services are relatively cheaper in poorer countries, is that the cross-country productivity dispersion is larger in manufacturing than in services, following the seminal study of Kravis, Heston and Summers (1982). Duarte and Restuccia (2010) offers the contrarian view; also related is Rodrik (2013)'s finding of unconditional convergence in manufacturing. Duarte and Restuccia (2010, p.154-156) argued that their finding is not inconsistent with the conventional view, because they look at producer prices, not the expenditure prices. The disagreement may also stem from the fact that the producer prices, capital, and its capacity utilization rate are harder to measure in services, and that home productions are more important in services. Our calibration result is in line with the Duarte-Restuccia-Rodrik view. However, it should be stressed that, although small cross-country productivity dispersion in manufacturing helps, its magnitude relative to the service sector is not crucial for our mechanism.

[5]For structural change and trade, see, e.g., Atkin, Costinot, and Fukui (2022), Cravino and Soleto (2019), Lewis et.al. (2022), Matsuyama (1992, 2009, 2019), Sposi, Yi, and Zhang (2021), and Uy, Yi, and Zhang (2013).

poor countries to catch up by narrowing their technology gaps and show that the main messages carry over, unless the catching up speed is too high.

By focusing on the cross-country heterogeneity in adoption lags across sectors, particularly due to the difference in the technology gap, we do not mean to suggest that the countries differ only in the technology gap, nor we suggest that our mechanism is the sole cause for PD.[6] Nor do we intend to argue that the technology gap alone could explain the patterns of structural change. As the rich literature on structural change has convincingly demonstrated, structural change is a multifaceted phenomenon, which defies any simple explanation. Indeed, there are many important issues that we abstract from, including, but not limited to, sector-specific factor intensities,[7] home production,[8] consumption vs. investment,[9] productivity gaps,[10] endogenous productivity and externalities,[11] and much more. Recent calibration studies in the field incorporate many of these issues to fit the data. While successful in accounting for the data, such complex models with a rich array of moving parts and a large dimension of exogenous heterogeneity across countries obscure the driving forces behind PD. In this paper, we instead opt for a parsimonious approach by tying our hands to restrict ourselves to the cross-country differences in sector-specific adoption lags, particularly, to one dimension of exogenous cross-country heterogeneity, i.e., the technology gap, to explain three dimensions of endogenous cross-country heterogeneity, i.e., the peak time, the peak manufacturing share, and the peak time per capita income.[12]

---

[6]In this respect, Huneeus and Rogerson (2020) deserve special mention. Comparing the two models is not easy because of the difference in objectives, with the different preference specification and different parametrizations for the evolution of sectoral productivity profiles. They propose a model, where the cross-country difference in agricultural productivity growth rate is the sole source of heterogenous development paths in the presence of the subsistence level of agriculture goods consumption. They do not aim to explain the lower peak time per capita income among late industrializers, since explaining PD is not their objective. In contrast, we assume the common productivity growth rate to focus on explaining PD by the relative productivity differences, which depends on the distance to the frontier. In this sense, our approach is closer in spirit to the "distance to the frontier" literature, such as Acemoglu, Aghion, and Zilibotti (2006).
[7]See, e.g., Acemoglu and Guerrieri (2008), Buera et. al. (2021), and Cravino and Soleto (2019).
[8]See, e.g., Ngai and Pissarides (2008).
[9]See, e.g., Garcia-Santana et.al. (2021) and Herrendorf, Rogerson, and Valentinyi (2021).
[10]See, e.g., Caselli (2005), and Gollin, Lagakos, and Waugh (2014b).
[11]See, e.g., Atkin, Costinot, and Fukui (2022) and Matsuyama (1992, 2002, 2019).
[12]In doing so, our parsimonious approach follows the long tradition in international trade, which seeks to explain the patterns of trade across many sectors across many countries with only one dimension of exogenous cross-country heterogeneity at a time, i.e., *other things being equal*, or "an elementary theory," as Costinot (2009) would call it. See, e.g., Krugman (1985), Matsuyama (2005), Costinot (2009) and Costinot and Vogel (2010, 2015).

The rest of the paper is as follows. In Section 2, we set up the baseline model and derive the analytical expressions that show how the peak time, the peak manufacturing share, and the peak time per capita income depends on the adoption lags. In Section 3, we further tie our hands by assuming that adoption lags are proportional to the technology gap, the only source of cross-country heterogeneity, in order to identify the sufficient and necessary condition for PD. We then calibrate the model to match the findings of Rodrik. In Section 4, we add the Engel effect on top of the Baumol effect to demonstrate that the main messages of the baseline model are not affected. We also show that the Engel effect alone could cause PD, but only under the conditions that would generate counterfactual implications. In Section 5, we introduce international trade to show that our mechanism is robust to opening up for trade but weakened by it. In section 6, we allow poorer countries to catch up by narrowing the technology gap over time. We conclude in Section 7.

2. Structural Change, the Baumol Effect, and Adoption Lags

Consider an economy with three competitive sectors, indexed by $j = 1, 2, 3$. Each sector produces a single consumption good, also indexed by $j = 1, 2, 3$. We interpret sector-1 as agriculture, sector-2 as manufacturing and sector-3 as services. In the baseline model, the hump-shaped path of the manufacturing share is driven solely by the Baumol effect, with the sector-specific productivity growth rate being the highest in agriculture and the lowest in services. To this, we add sector-specific adoption lags, which affect the manufacturing peak time, the peak manufacturing share and the peak time per capita income.

2.1 Households

The economy is populated by L identical households. Each household supplies one unit of labor, perfectly mobile across sectors, at the wage rate w, and κ_j units of the managerial skills specific to sector-j, which generate the managerial rents $\rho_j\kappa_j$. Each household consumes c_j units of good-j, purchased at the price, p_j, subject to the budget constraint,

(1)
$$\sum_{j=1}^{3} p_j c_j \leq E = w + \sum_{j=1}^{3} \rho_j \kappa_j,$$

to maximize its CES utility

(2)
$$U(c_1, c_2, c_3) = \left[\sum_{j=1}^{3} (\beta_j)^{\frac{1}{\sigma}} (c_j)^{1-\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}}$$

where E denotes the household expenditure and $\beta_j > 0$ and $0 < \sigma < 1$, so that the three goods are gross complements. This maximization yields the expenditure shares,

(3)
$$m_j \equiv \frac{p_j c_j}{E} = \beta_j \left(\frac{p_j}{P}\right)^{1-\sigma}; \quad P = \left[\sum_{k=1}^{3} \beta_k (p_k)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$$

where P is the cost-of-living index and hence the (real) per capita income is given by:

$$U = \frac{E}{P} = \left[\sum_{k=1}^{3} \beta_k \left(\frac{E}{p_k}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}}.$$

2.2. Production

Good-j is produced in sector-j with the following Cobb-Douglas Technology:

(4)
$$Y_j = \tilde{A}_j(\kappa_j L)^{\alpha}(L_j)^{1-\alpha} = A_j(L)^{\alpha}(L_j)^{1-\alpha} = LA_j(s_j)^{1-\alpha}$$

with $A_j \equiv \tilde{A}_j(\kappa_j)^{\alpha} > 0$. Here, $\kappa_j L$ is the total supply of the managerial skills in sector-j; and L_j the labor employed in sector-j, with $s_j \equiv L_j/L$ the employment share of sector-j, satisfying

(5)
$$\sum_{j=1}^{3} s_j = 1.$$

The parameter, $0 < \alpha < 1$, is the "span of control," which introduces diminishing returns in labor, as in Beckman (1976) and Lucas (1977).[13] As is well-known, this specification leads to $\rho_j \kappa_j L = \alpha p_j Y_j$ and $wL_j = (1-\alpha)p_j Y_j$. That is, the share of the profit (the managerial rents) is equal to α, and the share of labor is equal to $1 - \alpha$. This implies the aggregate budget constraint, $EL = \sum_{j=1}^{3} p_j Y_j$, and that the sectoral shares measured in employment are equal to those measured in value-added,

[13] That is, one unit of managerial skill in sector-j operates a firm, which produces $y_j = \tilde{A}_j(\ell_j)^{1-\alpha}$ units of output with ℓ_j units of labor, and the span of control, $\alpha > 0$, generates managerial rents and diminishing returns in labor. We introduce diminishing returns by assuming $\alpha > 0$ for three reasons. First, if the technologies were linear in labor with $\alpha = 0$, sectoral shares in employment would respond discontinuously to the prices, as should be clear from eq.(6) and eq.(31). This would cause the economy to specialize in a trade equilibrium even with a finite trade cost, no matter how large the trade cost is. Diminishing returns in labor makes the prediction of the model robust to opening up for trade, as seen in section 5. Second, a higher α amplifies the effect so that setting the labor share $1 - \alpha$ in the empirically plausible range helps to generate the impact observed in the data. Moreover, $\alpha > 0$ significantly simplifies the analysis.

$$s_j \equiv \frac{L_j}{L} = \frac{p_j Y_j}{\sum_{k=1}^{3} p_k Y_k} = \frac{p_j Y_j}{EL},$$

which can be combined with eq.(4) to yield

(6)
$$s_j = \left(\frac{p_j A_j}{E}\right)^{1/\alpha} \; ; \; E = \left[\sum_{k=1}^{3} (p_k A_k)^{\frac{1}{\alpha}}\right]^{\alpha}.$$

2.3 Equilibrium

In the closed economy, the expenditure shares must be equal to the value-added (and employment) shares. From eq. (3) and eq.(6), this equilibrium condition can be written as:

$$\left(\frac{p_j A_j}{E}\right)^{1/\alpha} = s_j = m_j = \beta_j \left(\frac{p_j}{P}\right)^{1-\sigma}.$$

Solving this condition for p_j and using eq.(5) yield the equilibrium values of s_j and U as:

(7)
$$s_j = \frac{\left[\beta_j^{\frac{1}{\sigma-1}} A_j\right]^{-a}}{\sum_{k=1}^{3} \left[\beta_k^{\frac{1}{\sigma-1}} A_k\right]^{-a}}$$

(8)
$$U = \frac{E}{P} = \left\{\sum_{k=1}^{3} \left[\beta_k^{\frac{1}{\sigma-1}} A_k\right]^{-a}\right\}^{-\frac{1}{a}}$$

where

$$a \equiv \frac{1-\sigma}{1-\alpha(1-\sigma)} = -\frac{\partial \ln(s_j/s_k)}{\partial \ln(A_j/A_k)} > 0.$$

Eq.(7) and eq.(8) show the equilibrium values of the sectoral shares (measured in employment, value-added, and expenditure) and of the (real) per capita income, as functions of the sectoral productivities, $\{A_j\}_{j=1}^{3}$, and a captures how much relatively *high* productivity in a sector contributes to its relatively *low* equilibrium share. Note that a is increasing in α, because stronger diminishing returns in labor makes relative labor demand less responsive to the relative productivity, thereby amplifying this effect.[14]

[14] Gross complementarity, $\sigma < 1$, is crucial here. If the three goods were gross substitutes, $\sigma > 1$, $a < 0$, which means higher relative productivity in a sector would lead to its relatively high sectoral share. And a higher α, which makes relative labor demand less responsive to relative productivity, would moderate this effect by decreasing a in the absolute value. If $\sigma = 1$, $a = 0$ and relative productivity has no effect on the sectoral shares.

2.4. Productivity Growth Rates, Adoption Lags and Structural Change

We now see how the sectoral shares respond to changes in the sectoral productivities over time.[15] Let $\{A_j(t)\}_{j=1}^{3}$ change according to:

$$A_j(t) = \bar{A}_j(t - \lambda_j) = \bar{A}_j(0)e^{g_j(t-\lambda_j)} = \bar{A}_j(0)e^{-\lambda_j g_j}e^{g_j t}, \tag{9}$$

with $g_j > 0$ and $\lambda_j \geq 0$, while the other parameters stay constant.[16] Here, $\bar{A}_j(t) = \bar{A}_j(0)e^{g_j t}$ is the **frontier technology** in sector-j at time t, which grows at a constant rate $g_j > 0$. With $A_j(t) = \bar{A}_j(t - \lambda_j)$, λ_j represents the **adoption lag** in sector-j. Note that both the growth rates and the adoption lags are *sector-specific.* Note also that the adoption lag in each sector does not affect the productivity growth rate of that sector, but the "level" effect of the adoption lags, $e^{-\lambda_j g_j}$, depends on the growth rate. Note that productivity in sector-j is *log-submodular* in λ_j and g_j, $\frac{\partial}{\partial g_j}\left(\frac{\partial}{\partial \lambda_j}\ln e^{-\lambda_j g_j}\right) < 0$. That is, a higher g_j magnifies the negative effect of a larger λ_j on productivity. In other words, a large adoption lag would not matter much in a sector with slow productivity growth. In contrast, even a small adoption lag would matter a lot in a sector with fast productivity growth. For now, "the base year," $t = 0$, is chosen arbitrarily, but we will later set the calendar time to ease the notation and to facilitate the interpretation.

Inserting eq.(9) into eq.(8) yields the time path of the real per capita income:

$$U(t) = \left\{\sum_{k=1}^{3} \tilde{\beta}_k e^{-a g_k (t-\lambda_k)}\right\}^{-\frac{1}{a}} \tag{10}$$

where $\tilde{\beta}_k \equiv \left((\beta_k)^{\frac{1}{1-\sigma}}/\bar{A}_k(0)\right)^{a} > 0$. Clearly, larger adoption lags would shift down the time path of $U(t)$. Log-differentiating eq.(10) with respect to time shows:

$$g_U(t) \equiv \frac{U'(t)}{U(t)} = \sum_{k=1}^{3} g_k s_k(t),$$

which states the aggregate growth rate is the weighted average of the sectoral growth rates.

To understand the Baumol effect, the productivity growth rate differences across sectors as the driving force behind structural change, let us take the ratio of the shares of two sectors, j and $k \neq j$, given in eq.(7) and using eq.(9), to obtain:

---

[15] With no means to save, the equilibrium path of the economy can be viewed as a sequence of the static equilibria.

[16] Since $A_j \equiv \tilde{A}_j(\kappa_j)^{\alpha}$, g_j could potentially include both the growth rate of $\tilde{A}_j$ and the growth rate of κ_j.

(11) $$\frac{s_j(t)}{s_k(t)} = \left(\frac{\tilde{\beta}_j}{\tilde{\beta}_k}\right) e^{a(\lambda_j g_j - \lambda_k g_k)} e^{a(g_k - g_j)t} \Rightarrow \frac{d}{dt} ln\left(\frac{s_j(t)}{s_k(t)}\right) = a(g_k - g_j)$$

Eq. (11) shows that, with the two sectors producing gross complements ($\sigma < 1 \Rightarrow a > 0$), $s_j(t)/s_k(t)$ declines over time if $g_j > g_k$, and increases over time if $g_j < g_k$. That is, the sectoral shares shift from sectors with faster productivity growth to those with slower productivity growth over time. In contrast, the adoption lags have no effect on the direction nor the speed of the sector share changes *over time*. But they shift the time path, with a higher $\lambda_j g_j - \lambda_k g_k$ raising $s_j(t)/s_k(t)$ at any point in time. Likewise, the relative price changes over time as:

(12) $$\left(\frac{p_j(t)}{p_k(t)}\right)^{1-\sigma} = \left[\left(\frac{\beta_j}{\beta_k}\right)^{\alpha} \frac{\bar{A}_k(0)}{\bar{A}_j(0)}\right]^{a} e^{a(\lambda_j g_j - \lambda_k g_k)} e^{a(g_k - g_j)t} \Rightarrow \frac{d}{dt} ln\left(\frac{p_j(t)}{p_k(t)}\right) = \frac{a(g_k - g_j)}{1-\sigma}$$

Eq.(12) shows that $p_j(t)/p_k(t)$ declines over time if $g_j > g_k$, and increases over time if $g_j < g_k$, so that slower productivity growth causes its relative price to go up over time. In contrast, the adoption lags have no effect on the direction nor the speed of the relative price changes *over time*. But they shift the time path, with a higher $\lambda_j g_j - \lambda_k g_k$ raising $p_j(t)/p_k(t)$ at any point in time.

In what follows, we restrict ourselves to the case of $g_1 > g_2 > g_3 > 0$, to generate the patterns of structural change, well-documented by, e.g. Herrendorf, Rogerson, Valentinyi (2014), based on the mechanism put forward by Ngai and Pissarides (2007). Then, using eq.(11), the agriculture share, $s_1(t)$, is decreasing over time, since,

$$\frac{1}{s_1(t)} - 1 = \left[\frac{\tilde{\beta}_2}{\tilde{\beta}_1} e^{a(\lambda_2 g_2 - \lambda_1 g_1)}\right] e^{a(g_1 - g_2)t} + \left[\frac{\tilde{\beta}_3}{\tilde{\beta}_1} e^{a(\lambda_3 g_3 - \lambda_1 g_1)}\right] e^{a(g_1 - g_3)t},$$

with both terms on the RHS increasing. The service share, $s_3(t)$, is increasing over time since

$$\frac{1}{s_3(t)} - 1 = \left[\frac{\tilde{\beta}_1}{\tilde{\beta}_3} e^{a(\lambda_1 g_1 - \lambda_3 g_3)}\right] e^{-a(g_1 - g_3)t} + \left[\frac{\tilde{\beta}_2}{\tilde{\beta}_3} e^{a(\lambda_2 g_2 - \lambda_3 g_3)}\right] e^{-a(g_2 - g_3)t},$$

with both terms on the RHS decreasing. In contrast, $s_2(t)$ is hump-shaped since

(13) $$\frac{1}{s_2(t)} - 1 = \left[\frac{\tilde{\beta}_1}{\tilde{\beta}_2} e^{a(\lambda_1 g_1 - \lambda_2 g_2)}\right] e^{-a(g_1 - g_2)t} + \left[\frac{\tilde{\beta}_3}{\tilde{\beta}_2} e^{a(\lambda_3 g_3 - \lambda_2 g_2)}\right] e^{a(g_2 - g_3)t}$$

with the 1st term of RHS exponentially decreasing and the 2nd term exponentially increasing. By differentiating eq.(13) with respect to t,

$$s_2'(t) \gtreqless 0 \Leftrightarrow (g_1 - g_2)s_1(t) \gtreqless (g_2 - g_3)s_3(t) \Leftrightarrow g_U(t) = \sum_{k=1}^{3} g_k s_k(t) \gtreqless g_2.$$

This shows that $s_2(t)$ is hump-shaped due to the two opposing forces; $g_1 > g_2$ pushes labor out of agriculture to manufacturing, while $g_2 > g_3$ pulls labor out of manufacturing to services. At earlier stages of development, $s_1(t)/s_3(t)$ is high and the first force dominates the second, but at later stages, $s_1(t)/s_3(t)$ is low and the second force dominates the first.

2.5. The Manufacturing Peak

We now characterize the manufacturing peak time, $\hat{t}$, the peak manufacturing share, $\widehat{s_2} \equiv s_2(\hat{t})$, and the real per capita income at the peak time, $\widehat{U} \equiv U(\hat{t})$. In what follows, we adopt two normalizations, without loss of generality, which helps to simplify the notation and facilitate the interpretation. Our first normalization is to choose the base year $t = 0$, such that

$$\frac{g_2 - g_3}{g_1 - g_2} = \frac{\tilde{\beta}_1}{\tilde{\beta}_3} \equiv \left[\left(\frac{\beta_1}{\beta_3}\right)^{\frac{1}{1-\sigma}} \frac{\bar{A}_3(0)}{\bar{A}_0(0)}\right]^a.$$

Then, solving $s_2'(\hat{t}) = 0$ yields:

(14)
$$\hat{t} = \frac{\lambda_1 g_1 - \lambda_3 g_3}{g_1 - g_3}.$$

This normalization thus effectively sets the calendar time such that the manufacturing sector would reach its peak at $\hat{t} = 0$ *in the absence of adoption lags.* Thus, eq.(14) isolates the impact of the adoption lags on the peak time. Then, inserting eq.(14) into eq.(13) and eq.(10) yields:

(15)[17]
$$\frac{1}{\widehat{s_2}} = 1 + \left(\frac{\tilde{\beta}_1}{\tilde{\beta}_2}\right) e^{a(g_1-g_2)\left(\frac{\lambda_1 g_1 - \lambda_2 g_2}{g_1 - g_2} - \hat{t}\right)} + \left(\frac{\tilde{\beta}_3}{\tilde{\beta}_2}\right) e^{a(g_2-g_3)\left(\hat{t} - \frac{\lambda_2 g_2 - \lambda_3 g_3}{g_2 - g_3}\right)}$$

(16)
$$\widehat{U} = \left\{\sum_{k=1}^{3} \tilde{\beta}_k e^{-a g_k(\hat{t} - \lambda_k)}\right\}^{-\frac{1}{a}}$$

Our second normalization is to set $\sum_{k=1}^{3} \tilde{\beta}_k = \sum_{k=1}^{3} \left((\beta_k)^{\frac{1}{1-\sigma}} / \bar{A}_k(0)\right)^a = 1$. From eq.(16), $\widehat{U} = U(\hat{t}) = U(0) = 1$ for $\lambda_1 = \lambda_2 = \lambda_3 = 0$, under this normalization. We thus choose the peak time real per capita income *in the absence of the adoption lags* as the *numeraire.* Moreover, from

[17] It is easy to verify from eq.(14) that the two exponents in the second and the third terms of RHS of eq.(15) are identical, which can be used to simplify eq.(15) further. However, we express eq.(15) as is, just to keep the symmetry of sector-1 and sector-3. The analogous statements can be made on eq.(18), eq.(25) and eq.(28).

eq.(15), the peak time share of sector-2 in the absence of the adoption lags is $s_2(\hat{t}) = \tilde{\beta}_2$ under these normalizations.

Eqs.(14)-(16) show how the sector-specific adoption lags in the three sectors, $(\lambda_1, \lambda_2, \lambda_3)$, affect the manufacturing peak time, $\hat{t}$, the peak manufacturing share, $\widehat{s_2}$, and the peak time per capita income, $\widehat{U}$. Obviously, we could perfectly account for the cross-county variations in three endogenous variables, $(\hat{t}, \widehat{s_2}, \widehat{U})$, if we allowed ourselves to the 3-dimension of cross-country heterogeneity in $(\lambda_1, \lambda_2, \lambda_3)$. However, it would also make the mechanisms harder to interpret. In what follows, we instead restrict ourselves to a 1-dimension of cross-country heterogeneity to propose a parsimonious mechanism of PD. In a world with at least three countries, this is clearly restrictive and ties our hands, but it allows us to express all three peak values, $(\hat{t}, \widehat{s_2}, \widehat{U})$, as functions of a single variable, which helps to facilitate the interpretation, to offer a much simpler narrative, thereby making the intuition more transparent.

3. Technology Gaps and "Premature" Deindustrialization

3.1 Adoption Lags and Technology Gaps

Now imagine that there are many countries, whose adoption lags are given by

$$(\lambda_1, \lambda_2, \lambda_3) = (\theta_1, \theta_2, \theta_3)\lambda,$$

with λ varying across countries. To keep it simple, we assume that all other parameters are identical across countries, with λ being the only source of country heterogeneity.[18] The idea is that each country tries to adopt the frontier technologies, which keep improving at exogenously constant growth rates, but the countries differ in their ability to adopt, indexed by the *country-specific* parameter, λ, which we shall call **the technology gap**, following Krugman (1985). Unlike Krugman (1985), who made no distinction between the adoption lags and the technology gap by assuming $\lambda_j = \lambda$ in all sectors, we allow the technology gap to have differential impacts on the adoption lag across sectors. That is, θ_j are *sector-specific* parameters, common across countries, capturing the inherent difficulty of adoption in the three sectors, which magnifies the impact of the technology gap on the adoption lag. From eq.(9) and using $\lambda_j = \theta_j\lambda$,

[18] Obviously, countries differ in many dimensions. By this assumption, we choose to zoom into this type of heterogeneity. That is, we are conducting a controlled thought experiment; *Other things being equal,* how this type of heterogeneity alone can generate PD.

$$\frac{A_j(t)}{A_k(t)} = \frac{\bar{A}_j(0)}{\bar{A}_k(0)} e^{-(\theta_j g_j - \theta_k g_k)\lambda} e^{(g_j - g_k)t} \Longrightarrow \frac{\partial}{\partial \lambda} \ln\left(\frac{A_j(t)}{A_k(t)}\right) = -(\theta_j g_j - \theta_k g_k).$$

Thus, the cross-country productivity dispersion is larger in sector-j than in sector-k if $\theta_j g_j > \theta_k g_k$. This is because the negative level effects of λ is *proportional* to $\theta_j g_j$ in sector-j.

By inserting $(\lambda_1, \lambda_2, \lambda_3) = (\theta_1, \theta_2, \theta_3)\lambda$ in eqs. (14)-(16), we obtain:

Proposition 1. Under the Baumol effect, the manufacturing peak time, $\hat{t}$, the peak manufacturing share, $\widehat{s_2}$, and the peak time per capita income, $\widehat{U}$, are functions of the technology gap λ as:

(17)
$$\hat{t}(\lambda) = \frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3} \lambda.$$

(18)
$$\frac{1}{\widehat{s_2}(\lambda)} = 1 + \left(\frac{\tilde{\beta}_1}{\tilde{\beta}_2}\right) e^{a(g_1 - g_2)\left(\frac{\theta_1 g_1 - \theta_2 g_2}{g_1 - g_2}\lambda - \hat{t}(\lambda)\right)} + \left(\frac{\tilde{\beta}_3}{\tilde{\beta}_2}\right) e^{a(g_2 - g_3)\left(\hat{t}(\lambda) - \frac{\theta_2 g_2 - \theta_3 g_3}{g_2 - g_3}\lambda\right)}$$

(19)
$$\widehat{U}(\lambda) = \left\{\sum\nolimits_{k=1}^{3} \tilde{\beta}_k e^{-a g_k[\hat{t}(\lambda) - \theta_k \lambda]}\right\}^{-\frac{1}{a}}$$

Note also that, from eq.(10),

$$U(t;\lambda) = \left\{\sum\nolimits_{k=1}^{3} \tilde{\beta}_k e^{-a g_k t} e^{a g_k \theta_k \lambda}\right\}^{-\frac{1}{a}},$$

which shows that a higher-λ country has a lower per capita income at any point in time.

3.2. The Three Conditions for "Premature" Deindustrialization

From eqs.(17)-(19), we can obtain the three conditions for PD; That is, the conditions under which a poorer, higher-λ (technologically more lagging) country has i) a higher peak time, $\hat{t}$; ii) a lower peak manufacturing share, $\widehat{s_2}$, and iii) a lower peak time per capita income, $\widehat{U}$.

Proposition 2. Premature Deindustrialization (PD) occurs under the three conditions:

a) $\hat{t}'(\lambda) > 0$ for all λ if and only if $\theta_1/\theta_3 > g_3/g_1$.

b) $\widehat{s_2}'(\lambda) < 0$ for all λ if and only if $\left(1 - \frac{g_3}{g_1}\right)\left(\frac{\theta_2}{\theta_3} - 1\right) + \left(1 - \frac{g_3}{g_2}\right)\left(1 - \frac{\theta_1}{\theta_3}\right) < 0$

c) There exists $\lambda_c \geq 0$, such that $\widehat{U}(\lambda) < \widehat{U}(0)$ and $\widehat{U}'(\lambda) < 0$ for all $\lambda > \lambda_c \geq 0$, if and only if $\theta_1/\theta_3 < 1$.

Figure 1 illustrates Proposition 2 by showing the parameter region of PD. Condition a), which follows from eq.(17), holds on the right side of the vertical line, $\theta_1/\theta_3 = g_3/g_1 < 1$ in Figure 1. Intuitively, $\theta_1 g_1 > \theta_3 g_3$ implies that the relative price of the agriculture good is higher and the relative price of services is lower in a higher-λ country, both of which cause a delay in their structural change. Condition b), which follows from eq.(18), holds below the line connecting $(\theta_1/\theta_3, \theta_2/\theta_3) = (g_3/g_1, g_3/g_2)$ and $(\theta_1/\theta_3, \theta_2/\theta_3) = (1,1)$ in Figure 1. Intuitively, with θ_2 sufficiently low, which has no effect on the peak time, $\hat{t}$, the relative price of manufacturing is sufficiently low in a higher-λ country, which keeps its peak manufacturing share, $\widehat{s_2}$, low.

Third, $\widehat{U}(\lambda)$ in eq.(19) is not generally monotone in λ. But, when Conditions a) and b) are met, a sufficiently high-λ country has a lower peak time per capita income, i.e., $\widehat{U}(\lambda) < \widehat{U}(0)$ and $\widehat{U}'(\lambda) < 0$ for all $\lambda > \lambda_c \geq 0$, if and only if $\theta_1 < \theta_3$, which holds on the left side of the vertical line $\theta_1/\theta_3 = 1$ in Figure 1. To see the intuition, note that $\theta_1 < \theta_3$ if and only if

$$0 < \hat{t}(\lambda) = \frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3}\lambda < \theta_1\lambda < \theta_3\lambda.$$

Thus, even if $\theta_1 g_1 > \theta_3 g_3$ causes a delay in the peak time, the delay is *not long enough* to make up for a high λ, so that the peak time agriculture and services productivity of the late industrializers are lower than those of the early industrializers. Even though their peak time manufacturing productivity could be higher if $\hat{t}(\lambda) > \theta_2\lambda$, this is not enough to offset low peak time productivity in agriculture and in services for a sufficiently large λ.[19] Furthermore, $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$ on the left side of the dashed line connecting $(\theta_1/\theta_3, \theta_2/\theta_3) = (\Theta, 1)$ and $(\theta_1/\theta_3, \theta_2/\theta_3) = (1,1)$ in Figure 1, where $g_3/g_1 < \Theta \equiv \frac{(1-\tilde{\beta}_2)g_3/g_2 + \tilde{\beta}_2 g_3/g_1}{(1-\tilde{\beta}_2)g_3/g_2 + \tilde{\beta}_2} < 1$.[20] In particular, $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$ for any $\tilde{\beta}_2 > 0$ if $(1 - g_3/g_1)(1 - \theta_2/\theta_3) < 1 - \theta_1/\theta_3$, which is equivalent to $\hat{t}(\lambda) < \theta_2\lambda$. This means that $\hat{t}(\lambda) - \theta_j\lambda$ is negative and decreasing in λ in all $j = 1,2,3$. In other words, the delay is *not long enough* to compensate the adoption lag in *any* sector, so that the peak time productivity in every sector is lower for late industrializers, and hence their peak-time per capita income is lower regardless of $\tilde{\beta}_2$.

Notice that these three conditions for PD, summarized in Proposition 2 and depicted in Figure 1, jointly imply that $\theta_1 g_1 > \max\{\theta_2 g_2, \theta_3 g_3\}$, thus, cross-country productivity dispersion is the largest in agriculture. However, these conditions do not impose any restriction on the

---

[19] To see this, note that, if $\hat{t}(\lambda) > \theta_2\lambda$, $\tilde{\beta}_2 e^{-\alpha g_2[\hat{t}(\lambda)-\theta_2\lambda]}$ is exponentially decreasing in λ and thus negligible for a sufficiently large λ relative to the exponentially growing terms, $\tilde{\beta}_1 e^{-\alpha g_1[\hat{t}(\lambda)-\theta_1\lambda]}$ and $\tilde{\beta}_3 e^{-\alpha g_3[\hat{t}(\lambda)-\theta_3\lambda]}$.

[20] Moreover, the region in which $\lambda_c > 0$ shrinks as $\tilde{\beta}_2$ declines and disappears as $\tilde{\beta}_2 \to 0$, since Θ is decreasing in $\tilde{\beta}_2$ with $\Theta \to 1$ as $\tilde{\beta}_2 \to 0$. In contrast, the region in which $\lambda_c = 0$ always exists, since $\Theta \to g_3/g_1$ as $\tilde{\beta}_2 \to 1$. In particular, $g_3/g_1 < \theta_1/\theta_3 = \theta_2/\theta_3 < 1$, i.e., ensures $\hat{t}'(\lambda) > 0$, $\widehat{s_2}'(\lambda) < 0$, and $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$.

relative magnitude of $\theta_2 g_2$ and $\theta_3 g_3$, as can be easily verified in Figure 1, in which the horizontal line, $\theta_2/\theta_3 = g_3/g_2$ passes through the parameter region of PD. Thus, the cross-country productivity dispersion may or may not be larger in manufacturing than in services.[21]

We now illustrate these conditions with some examples.

3.3. Premature Deindustrialization: Some Numerical Illustrations

Example 1: First, consider the case where $\theta_1 = \theta_2 = \theta_3 = \theta$, so that $\lambda_1 = \lambda_2 = \lambda_3 = \theta\lambda > 0$, as in Krugman (1985). In Figure 1, this case is depicted by a black dot at the north-east corner of the unit square box. Then, from eqs.(17)-(19), $\hat{t}(\lambda) = \theta\lambda$; $\widehat{s_2}(\lambda) = \tilde{\beta}_2$; $\widehat{U}(\lambda) = 1$. Thus, if technology gaps affect the adoption lags in all the sectors uniformly, they cause a delay in the peak time by $\theta\lambda > 0$, which is exactly the same with the adoption lags in all the sectors. The peak manufacturing share and the peak time per capita income are thus unaffected. The reason is simple. Because the delay is exactly the same with their adoption lags, higher-λ countries, late industrializers, follow exactly the same path with early industrializers. Only the timing differs. Thus, for PD to occur, the technology gap must have differential impacts across sectors.

Example 2. Next, we consider the cases where $g_3/g_1 < \theta_1/\theta_3 = \theta_2/\theta_3 < 1$, i.e., on the diagonal inside the PD region, and hence $\hat{t}'(\lambda) > 0$, $\widehat{s_2}'(\lambda) < 0$, and $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$.[22] In Example 2, $g_3/g_1 = 1/3 < \theta_1/\theta_3 = \theta_2/\theta_3 = 0.5 = g_3/g_2$, so that $\theta_1 g_1 > \theta_2 g_2 = \theta_3 g_3$, depicted in Figure 1 by the black dot at the intersection of the diagonal line $\theta_1/\theta_3 = \theta_2/\theta_3$ and the horizontal line $\theta_2/\theta_3 = g_3/g_2$. Thus, in this example, cross-country productivity dispersion is the same in manufacturing and in services.[23] Figure 2 illustrates the path of the manufacturing

---

[21]Figure 1 shows that $\max\{\theta_1, \theta_2\} < \theta_3$; the adoption lag is the longest in services. The empirical evidence--see Comin and Mestieri (2014, 2018) and Comin et.al. (2008)--, is that, for the extensive margin of adoption (i.e., how long it takes for new vintages of technologies to arrive in a country), they find that it is shorter in services, but, for the intensive margin (i.e., how widely new vintages of technologies are used in a country), they find that it is less intensive in services, so that, when both margins are combined, the empirical support is at best inconclusive.

[22]For all the numerical illustrations in this paper, we use the following parameter values: $\alpha = 1/3$, $\sigma = 0.6$ (hence $a = 6/13$), $g_1 = 3.6\% > g_2 = 2.4\% > g_3 = 1.2\%$ and $\widehat{s_2}(0) = \tilde{\beta}_2 = 1/3$. For these parameter values, the two normalizations, $\hat{t}(0) = 0$ and $\widehat{U}(0) = 1$, are achieved by $\tilde{\beta}_1 = \tilde{\beta}_3 = (1 - \tilde{\beta}_2)/2 = 1/3$. Note that the values of α and σ matter only to the extent they affect a. The sectoral productivity growth rates are chosen to be close enough to the estimates by Duarte and Restuccia (2010; Section IIB), $g_1 = 3.8\% > g_2 = 2.4\% > g_3 = 1.3\%$, but have the ratios of 3/2/1.

[23] Fujiwara and Matsuyama (2023; Examples 2b-2c) look at the effects of moving along the diagonal by changing $\theta_1/\theta_3 = \theta_2/\theta_3$, which may be viewed as the effects of changing θ_3, holding $\theta_1 = \theta_2$. For $g_3/g_1 = 1/3 < \theta_1/\theta_3 = \theta_2/\theta_3 < 0.5 = g_3/g_2$, $\theta_1 g_1 > \theta_3 g_3 > \theta_2 g_2$, so that cross-country productivity dispersion is the smallest in manufacturing. For $g_3/g_1 = 1/3 < 0.5 = g_3/g_2 < \theta_1/\theta_3 = \theta_2/\theta_3$, $\theta_1 g_1 > \theta_2 g_2 > \theta_3 g_3$, so that cross-country productivity dispersion is the smallest in services. These examples show that, as a decline in $\theta_1/\theta_3 = \theta_2/\theta_3$

share for this example. The hump-shaped curves, each capturing the rise and fall of manufacturing, are plotted for $\theta_3\lambda = \lambda = 0$, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, using eq.(10) and eq.(13). The left panel shows the time paths of $s_2(t)$, which shows that a higher λ shifts the curve down and to the right, with the downward-sloping line connecting the peaks, starting from $\hat{t}(0) = 0$ due to the first normalization; this captures an increase in $\hat{t}(\lambda)$ and a decline in $\widehat{s_2}(\lambda)$. The right panel traces the trajectory of $(\ln U(t), s_2(t))$, which shows that a higher λ shifts the curve down and to the left, with the upward-sloping line connecting the peaks, starting from $\ln \widehat{U}(0) = 0$ due to the second normalization; this captures a decline both in $\widehat{s_2}(\lambda)$ and in $\widehat{U}(\lambda)$.

3.4. Premature Deindustrialization: Some Calibration Results

How big is the parameter region of PD shown in Figure 1? And what are the values of the relative adoption lags across sectors, $(\lambda_1/\lambda_3, \lambda_2/\lambda_3) = (\theta_1/\theta_3, \theta_2/\theta_3)$ that replicate the findings of Rodrik (2016)? And what are the implied productivity differences across sectors? We now turn to these questions.

The size of the parameter region of PD depends solely on $(g_3/g_1, g_3/g_2)$, for which we use $(g_1, g_2, g_3) = (3.8\%, 2.4\%, 1.3\%)$ by Duarte and Restuccia (2010; Section IIB).[24] For the peak values, we follow Rodrik (2016), who divided countries into two groups: the pre-1990 peaked and the post-1990 peaked. The former is considered early industrializers, while the latter is late industrializers. Here is the summary of his findings, translated in our notation, with $\lambda > 0$ the technology gap of the post-1990, while setting $\lambda = 0$ for that of the pre-1990 without any loss of generality.

- From Rodrik (2016; Figure 5), $\hat{t}(\lambda) - \hat{t}(0) = \hat{t}(\lambda) = 25$.
- From Rodrik (2016; Table 10), $\widehat{s_2}(0) = 21.5\% > \widehat{s_2}(\lambda) = 18.9\%$, and $\widehat{U}(\lambda)/\widehat{U}(0) = \widehat{U}(\lambda) = 4273/11048$, in terms of the employment shares, while $\widehat{s_2}(0) = 27.9\% > \widehat{s_2}(\lambda) = 24.1\%$, and $\widehat{U}(\lambda)/\widehat{U}(0) = \widehat{U}(\lambda) = 20537/47099$, in terms of the value-added shares.

---

magnifies the impact of a higher λ on $\widehat{s_2}(\lambda)$ and $\widehat{U}(\lambda)$ but reduces the impact on $\hat{t}(\lambda)$, while an increase in $\theta_1/\theta_3 = \theta_2/\theta_3$ has the opposite effects, and, as $\theta_1/\theta_3 = \theta_2/\theta_3$ goes to one, the impacts of $\widehat{s_2}(\lambda)$ and $\widehat{U}(\lambda)$ disappear, as seen in Example 1.

[24] We also used $(g_1, g_2, g_3) = (2.9\%, 1.3\%, 1.1\%)$ by Comin, Lashkari and Mestieri (2021; TABLE VIII) for the robustness in Fujiwara and Matsuyama (2024), but the results are very similar. Here, we choose the numbers by Duarte and Restuccia (2010) because they use the value-added perspective just like Rodrik (2016), while Comin, Lashkari and Mestieri (2021) use the final expenditure perspective.

From eqs. (17)-(19), some algebra yields

$$\theta_1\lambda = \hat{t}(\lambda) - \frac{1}{g_1}\ln\left(\widehat{U}(\lambda)\right) + \frac{1}{ag_1}\ln\left(\frac{1-\widehat{s_2}(\lambda)}{1-\widehat{s_2}(0)}\right),$$

$$\theta_2\lambda = \hat{t}(\lambda) - \frac{1}{g_2}\ln\left(\widehat{U}(\lambda)\right) + \frac{1}{ag_2}\ln\left(\frac{\widehat{s_2}(\lambda)}{\widehat{s_2}(0)}\right).$$

$$\theta_3\lambda = \hat{t}(\lambda) - \frac{1}{g_3}\ln\left(\widehat{U}(\lambda)\right) + \frac{1}{ag_3}\ln\left(\frac{1-\widehat{s_2}(\lambda)}{1-\widehat{s_2}(0)}\right).$$

Thus, for $(g_1, g_2, g_3) = (3.8\%, 2.4\%, 1.3\%)$, the sectoral productivity levels of the late industrializers relative to the early industrializers are $(e^{-g_1\theta_1\lambda}, e^{-g_2\theta_2\lambda}, e^{-g_3\theta_3\lambda}) \approx (13.9\%, 28.1\%, 26.0\%)$ with $(\theta_1/\theta_3, \theta_2/\theta_3) \approx (0.501, 0.512)$ and $\Theta \approx 0.779$ for employment shares, and $(e^{-g_1\theta_1\lambda}, e^{-g_2\theta_2\lambda}, e^{-g_3\theta_3\lambda}) \approx (15.1\%, 32.9\%, 28.2\%)$ with $(\theta_1/\theta_3, \theta_2/\theta_3) \approx (0.511, 0.476)$ and $\Theta \approx 0.726$ for value-added shares, as shown in Figure 3. Note that, in both cases, $\theta_1 g_1 > \theta_3 g_3 > \theta_2 g_2 \Leftrightarrow e^{-\theta_1 g_1\lambda} < e^{-\theta_3 g_3\lambda} < e^{-\theta_2 g_2\lambda}$. The productivity differences across countries are thus not only the largest in agriculture as the model predicts, but also the smallest in manufacturing, which is reminiscent of Rodrik (2013).

4. The First Extension: Adding the Engel Effect

In our baseline model, structural change is driven solely by the Baumol effect, i.e., differential productivity growth rates across sectors. In contrast, most existing models of structural change rely on the nonhomotheticity of sectoral demand compositions, the Engel effect for short, as the main driver behind the hump-shaped path of manufacturing. Indeed, the Baumol effect alone cannot explain another key feature of structural change, as pointed out by Boppart (2014); The path of the manufacturing share exhibits a hump-shape even when it is measured in the real expenditure. Rodrik (2016, p.7) also noted that we need a combination of the Baumol and Engel effects. More recently, Comin, Lashkari and Mestieri (2021) derived a decomposition of the Baumol versus Engel effects in their model that feature both and showed that 75% of structural change can be attributed to the Engel effect, with the remaining 25% to the Baumol effect.

In view of the importance of the Engel effect as the driver of structural change, we now extend our baseline model by adding the Engel effect. Not surprisingly, combining the Engel effect with the Baumol effect significantly changes the shape of the time path, but it has little effects on the impacts of the peak values, hence the main implications on PD obtained by the baseline model. Furthermore, if we had relied solely on the Engel effect without the Baumol effect, PD would occur only under the conditions that would imply, counterfactually, that the cross-country productivity dispersion has to be the largest in services. This is why we build our baseline model relying on the Baumol effect.

4.1. Isoelastic Nonhomothetic CES.

Many different ways of introducing nonhomothetic sectoral demand compositions have been used in the structural change literature.[25] In this paper, we use *isoelastic nonhomothetic CES*, following Comin-Lashkari-Mestieri (2021) and Matsuyama (2019), because it offers a natural extension to the Ngai-Pissarides CES framework. More specifically, the utility function of each household, $U = U(c_1, c_2, c_3)$, is given implicitly by

(20)
$$\left[\sum_{j=1}^{3} (\beta_j)^{\frac{1}{\sigma}} \left(\frac{c_j}{U^{\varepsilon_j}}\right)^{1-\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \equiv 1$$

where $\varepsilon_j > 0$. If $\varepsilon_1 = \varepsilon_2 = \varepsilon_3$, eq.(20) is reduced to the homothetic CES, eq.(2). If $\{\varepsilon_j\}_{j=1}^{3}$ vary with j, the relative weights attached to the three goods in eq.(20) change systematically, making it nonhomothetic. In particular, $\varepsilon_1 < \varepsilon_2 < \varepsilon_3$ implies that the income elasticity of agriculture is less than one and that of services greater than one.

In what follows, we normalize $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 = 3$, without loss of generality. Maximizing eq.(20) subject to the budget constraint, eq.(1) yields the expenditure shares:

(21)
$$m_j \equiv \frac{p_j c_j}{E} = \beta_j \left(\frac{U^{\varepsilon_j} p_j}{E}\right)^{1-\sigma},$$

where U here is the maximized value of the utility given by the indirect utility function,

$$\left[\sum_{j=1}^{3} \beta_j \left(\frac{U^{\varepsilon_j} p_j}{E}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \equiv \left[\sum_{j=1}^{3} \beta_j \left(\frac{U^{\varepsilon_j - 1} p_j}{P}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \equiv 1,$$

---

[25] For example, Matsuyama (1992, 2009) and Kongsamut et.al. (2001) use Stone-Geary preferences; Hierarchical preferences are used by Murphy, Shleifer, and Vishny (1989), Matsuyama (2000, 2002), Foellmi and Zweimüller (2008, 2014), Buera and Kaboski (2012a,b); PIGL by Boppart (2014). Huneeus and Rogerson (2020) use the subsistence level of the agricultural good, combined with PIGL over manufacturing and services. For the relative strengths and weaknesses of different classes of nonhomothetic preferences, see Matsuyama (2023).

also defined implicitly, and P is the cost-of-living index and hence $U = E/P$ can be interpreted as the real income per capita.[26]

From eq.(21), it is simple to verify that the income elasticity of good-j is given by,

$$\eta_j \equiv \frac{d \ln c_j}{d \ln(U)} = 1 + (1-\sigma)\left\{\varepsilon_j - \sum\nolimits_{k=1}^{3} m_k \varepsilon_k\right\}.$$

With $\sigma < 1, 0 < \varepsilon_1 < \varepsilon_2 < \varepsilon_3 = 3 - \varepsilon_1 - \varepsilon_2$ implies $\eta_1 < \eta_2 < \eta_3$ and $\eta_1 < 1 < \eta_3$. Thus, the income elasticity of demand for agriculture is always the lowest and that for services the highest.[27] Furthermore, with the constant relative prices, the expenditure share of agriculture $m_1(t)$ is decreasing in $U(t)$, since

$$\frac{1}{m_1(t)} - 1 = \frac{\beta_2}{\beta_1}\left(\frac{p_2}{p_1} U(t)^{\varepsilon_2 - \varepsilon_1}\right)^{1-\sigma} + \frac{\beta_3}{\beta_1}\left(\frac{p_3}{p_1} U(t)^{\varepsilon_3 - \varepsilon_1}\right)^{1-\sigma},$$

with both terms on the RHS increasing in $U(t)$. Likewise, the expenditure share of services, $m_3(t)$ is increasing in $U(t)$, since

$$\frac{1}{m_3(t)} - 1 = \frac{\beta_1}{\beta_3}\left(\frac{p_1}{p_3} U(t)^{\varepsilon_1 - \varepsilon_3}\right)^{1-\sigma} + \frac{\beta_2}{\beta_3}\left(\frac{p_2}{p_3} U(t)^{\varepsilon_2 - \varepsilon_3}\right)^{1-\sigma},$$

with both terms on the RHS decreasing in $U(t)$. In contrast, $m_2(t)$ is hump-shaped in $U(t)$, since

(22) $$\frac{1}{m_2(t)} - 1 = \frac{\beta_1}{\beta_2}\left(\frac{p_1}{p_2} U(t)^{\varepsilon_1 - \varepsilon_2}\right)^{1-\sigma} + \frac{\beta_3}{\beta_2}\left(\frac{p_3}{p_2} U(t)^{\varepsilon_3 - \varepsilon_2}\right)^{1-\sigma}$$

with the 1st term of RHS exponentially decreasing in $U(t)$ and the 2nd term exponentially increasing. By differentiating eq.(22) with respect to $U(t)$,

$$\frac{ds_2(t)}{dU(t)} \gtreqless 0 \Leftrightarrow (\varepsilon_2 - \varepsilon_1) m_1(t) \gtreqless (\varepsilon_3 - \varepsilon_2) m_3(t) \Leftrightarrow \varepsilon_2 \gtreqless \sum\nolimits_{k=1}^{3} \varepsilon_k m_k(t) \Leftrightarrow \eta_2 \gtreqless 1.$$

This shows that $m_2(t)$ is hump-shaped due to the two opposing forces; $\varepsilon_2 > \varepsilon_1$ pushes labor out of agriculture to manufacturing, while $\varepsilon_3 > \varepsilon_2$ pulls labor out of manufacturing to services. At

---

[26] Since the relative weights on the three goods vary continuously with U in eq.(20), the relative weights in the ideal (i.e., model-implied) cost-of-living index, P, as defined here, also vary with U. Of course, this is not the cost-of-living index used to calculate the real per capita GDP in practice. However, Comin, Lashkari, and Mestieri (2021; section 6.3) showed the model-implied cost-of-living index and the cost-of-living index used in practice are highly correlated; see, e.g., their Figure 5. Nevertheless, Fujiwara and Matsuyama (2023, Appendix A) considered an alternative measure of development, the share of non-agriculture, $s_n(t) = s_2(t) + s_3(t) = 1 - s_1(t)$, as in Huneeus and Rogerson (2020), and show that the results are qualitatively unchanged.

[27] Note that $\{\varepsilon_j\}_{j=1}^{3}$ themselves are not the income elasticities. But they are the parameters that jointly control the income elasticities $\{\eta_j\}_{j=1}^{3}$, which are variables, as η_j is decreasing in $\sum_{k=1}^{3} m_k \varepsilon_k$.

earlier stages of development when $m_1(t)/m_3(t)$ is high, the first force dominates the second, but at later stages when $m_1(t)/m_3(t)$ is low, the second force dominates the first.

4.2. Adding the Engel Effect on top of the Baumol Effect

Fujiwara and Matsuyama (2023; Section 4) showed the steps for solving the peak time values, $\hat{t}$, $\widehat{s_2}$, and $\widehat{U}$, in the presence of both the Baumol and Engel effects. Generally, these values can be obtained only numerically. However, they can be solved analytically as functions of λ under the following condition:

(23)
$$0 \leq \mu \equiv \frac{\varepsilon_2 - \varepsilon_1}{g_1 - g_2} = \frac{\varepsilon_3 - \varepsilon_2}{g_2 - g_3} < \frac{1}{g_1 - \bar{g}},$$

where $\bar{g} \equiv (g_1 + g_2 + g_3)/3$ is the (simple) average growth rate, as follows.[28]

Proposition 3: Suppose that eq.(23) holds. Then, for each μ, the manufacturing peak time, $\hat{t}$, the peak manufacturing share, $\widehat{s_2}$, and the peak time per capita income, $\widehat{U}$, are written explicitly as functions of the technology gap, λ, as follows:

(24)
$$\hat{t}(\lambda;\mu) = \frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3}\lambda - \mu \ln \widehat{U}(\lambda;\mu) = \hat{t}(\lambda;0) - \frac{\mu}{1+\mu\bar{g}} \ln \widehat{U}(\lambda;0)$$

(25)
$$\frac{1}{\widehat{s_2}(\lambda;\mu)} = 1 + \left(\frac{\tilde{\beta}_1}{\tilde{\beta}_2}\right) e^{a(g_1-g_2)\left(\frac{\theta_1 g_1 - \theta_2 g_2}{g_1 - g_2}\lambda - \hat{t}(\lambda;0)\right)} + \left(\frac{\tilde{\beta}_3}{\tilde{\beta}_2}\right) e^{a(g_2-g_3)\left(\hat{t}(\lambda;0) - \frac{\theta_2 g_2 - \theta_3 g_3}{g_2 - g_3}\lambda\right)} = \frac{1}{\widehat{s_2}(\lambda;0)}$$

(26)
$$\widehat{U}(\lambda;\mu) = \left\{\sum_{k=1}^{3} \tilde{\beta}_k e^{-a g_k[\hat{t}(\lambda;0) - \theta_k \lambda]}\right\}^{-\frac{1}{a}\left(\frac{1}{1+\mu\bar{g}}\right)} = \widehat{U}(\lambda;0)^{\left(\frac{1}{1+\mu\bar{g}}\right)}$$

This proposition allows us to study analytically the effect of adding the Engel effect on top of the Baumol effect by increasing μ, for given $g_1 > g_2 > g_3 > 0$. By comparing eqs.(24)-(26) with eqs.(17)-(19), it is easy to see $\widehat{s_2}'(\lambda;\mu) < 0$ and $\widehat{U}'(\lambda;\mu) < 0$ for all $\lambda > 0$ under the same condition; and $\hat{t}'(\lambda;\mu) > 0$ for all $\lambda > 0$ under a weaker condition**.** Thus, adding more nonhomotheticity while satisfying eq.(23) does not change the main messages of the baseline model. Moreover,

- $\hat{t}(0;\mu) = 0, \widehat{s_2}(0;\mu) = \tilde{\beta}_2, \widehat{U}(0;\mu) = 1$; a higher μ has no effect on the country with $\lambda = 0$;

[28] To see why eq.(23) helps to keep the peak values analytically solvable, recall that, under the Baumol effect only, the hump-shaped time path of the manufacturing share is due to the two opposing forces whose relative strength is $(g_1 - g_2)/(g_2 - g_3)$. Likewise, under the Engel effect with the constant relative prices, the manufacturing share is hump-shaped in U due to the two opposing forces whose relative strength is $(\varepsilon_2 - \varepsilon_1)/(\varepsilon_3 - \varepsilon_2)$. Thus, by adding the Engel effect on top of the Baumol effect in a way $(\varepsilon_2 - \varepsilon_1)/(\varepsilon_3 - \varepsilon_2) = (g_1 - g_2)/(g_2 - g_3)$ holds, we do not change the relative strength of the two opposing forces that create the hump-shape. (The upper bound of μ needs to be imposed to ensure $\varepsilon_1 > 0$.)

- a higher μ causes a further delay in $\hat{t}(\lambda;\mu)$ for every country with $\lambda > 0$, from eq.(24);
- a higher μ has no effect on $\widehat{s_2}(\lambda;\mu)$, from eq.(25);
- A higher μ makes a decline in $\widehat{U}(\lambda;\mu)$ smaller for every country with $\lambda > 0$, from eq.(26).

Figure 4a illustrates these results. We use the same parameter values as in Figure 2. Because $g_1 = 3.6\% > g_2 = 2.4\% > g_3 = 1.2\%$, $g_1 - g_2 = g_2 - g_3 = 1.2\%$. Hence, $\varepsilon_1 = 1 - \epsilon < \varepsilon_2 = 1 < \varepsilon_3 = 1 + \epsilon$ for $0 < \epsilon = (1.2\%)\mu < 1$. It is easy to see why increasing the importance of the Engel effect causes a longer delay in the peak time, a higher $\hat{t}(\lambda;\mu)$, and hence the peak time occurs less prematurely, a higher $\widehat{U}(\lambda;\mu)$, since the driver of structural change is a change in time under the Baumol effect, while it is a change in U under the Engel effect.[29]

Thus, adding the Engel effect on top of the Baumol effect does not change fundamentally how technology gaps affect the peak time values and hence the conditions for PD. However, it should be noted that nonhomotheticity has significant effects on the time path of structural change. This can be seen in Figure 4b, which plots the paths of the manufacturing share against time and against log-per capita income $\ln U(t)$ for two cases: the homothetic case in Figure 2 and the analytical solvable case in Figure 4a with $\epsilon = 0.6$. Adding the Engel effect makes the hump-shaped noticeably much sharper, which indicates nonhomotheticity can speed up the pace of structural change. However, we do not see any noticeable changes in the peak time values.

4.3. Premature Deindustrialization (PD) through the Engel Effect Only

One may wonder what happens if we rely *solely* on the Engel effect, by removing the Baumol effect with $g_1 = g_2 = g_3 = \bar{g} > 0$, while keeping $0 < \varepsilon_1 < \varepsilon_2 < \varepsilon_3 = 3 - \varepsilon_1 - \varepsilon_2$. Then, under the two normalizations, $\tilde{\beta}_1(\varepsilon_2 - \varepsilon_1) = \tilde{\beta}_3(\varepsilon_3 - \varepsilon_2)$; $\tilde{\beta}_1 + \tilde{\beta}_2 + \tilde{\beta}_3 = 1$, which ensures $\hat{t}(0) = 0$ and $\widehat{U}(0) = 1$, we obtain

---

[29] Obviously, eq.(23) is a knife-edge condition. In particular, the result that adding the Engel effect has no effect on the peak values of the frontier country and no effect on $\widehat{s_2}(\lambda)$ for every country with $\lambda > 0$ is not robust. For $(\varepsilon_2 - \varepsilon_1)/(\varepsilon_3 - \varepsilon_2) \neq (g_1 - g_2)/(g_2 - g_3)$, the peak values need to be solved numerically. For example, using the values close to the estimates by Comin, Lashkari, and Mestieri (2021, TABLE VIII), Fujiwara and Matsuyama (2023; Figure 3b) reports the peak values for $(\varepsilon_2 - \varepsilon_1)/(\varepsilon_3 - \varepsilon_2) = 4 > 1 = (g_1 - g_2)/(g_2 - g_3)$, where $\varepsilon_1 = 1 - \epsilon < \varepsilon_2 = 1 + \epsilon/3 < \varepsilon_3 = 1 + 2\epsilon/3$ for $0 < \epsilon < 1$. Since changing the relative magnitude of the Engel effect to the Baumol effect affects the peak time values even for the frontier country in these cases, the peak time values relative to the country with $\lambda = 0$ are plotted. These plots suggest that, relative to the frontier country, a higher ϵ, more nonhomotheticity, causes a higher-λ country to have a further delay in $\hat{t}$ and a smaller decline in $\widehat{U}$, similar to the analytical case. As for $\widehat{s_2}$, a higher ϵ makes a decline in $\widehat{s_2}$ larger for every country with $\lambda > 0$. We also looked at some cases where $(\varepsilon_2 - \varepsilon_1)/(\varepsilon_3 - \varepsilon_2) < (g_1 - g_2)/(g_2 - g_3)$. The effects of more nonhomotheticity on $\hat{t}$ and $\widehat{U}$ are qualitatively the same. The effect on $\widehat{s_2}$ is a smaller decline for every country with $\lambda > 0$.

Proposition 4. The manufacturing peak time, $\hat{t}$, the peak manufacturing share, $\widehat{s_2}$, and the peak time per capita income, $\widehat{U}$, are written as functions of the technology gap, λ, as follows:

(27)
$$\hat{t}(\lambda) = \frac{1}{a\bar{g}} ln\left\{\sum_{k=1}^{3} \tilde{\beta}_k e^{a(\theta_k \bar{g}\lambda + \varepsilon_k\, ln\, \widehat{U}(\lambda))}\right\}$$

(28)
$$\frac{1}{\widehat{s_2}(\lambda)} = 1 + \left(\frac{\tilde{\beta}_1}{\tilde{\beta}_2}\right) e^{a(\varepsilon_2-\varepsilon_1)\left(-\frac{\theta_2-\theta_1}{\varepsilon_2-\varepsilon_1}\bar{g}\lambda - ln\, \widehat{U}(\lambda)\right)} + \left(\frac{\tilde{\beta}_3}{\tilde{\beta}_2}\right) e^{a(\varepsilon_3-\varepsilon_2)\left(ln\, \widehat{U}(\lambda) - \frac{\theta_2-\theta_3}{\varepsilon_3-\varepsilon_2}\bar{g}\lambda\right)}$$

(29)
$$ln\, \widehat{U}(\lambda) = \frac{\theta_1 - \theta_3}{\varepsilon_3 - \varepsilon_1}\bar{g}\lambda$$

from which we obtain the three conditions for PD as follows.

Proposition 5. Premature Deindustrialization occurs under the following three conditions:

a) $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$ if and only if $\theta_1/\theta_3 < 1$.

b) $\widehat{s_2}'(\lambda) < 0$ for all λ if and only if $\left(1 - \frac{\varepsilon_2}{\varepsilon_3}\right)\left(\frac{\theta_1}{\theta_3} - 1\right) + \left(1 - \frac{\varepsilon_1}{\varepsilon_3}\right)\left(1 - \frac{\theta_2}{\theta_3}\right) > 0$.

c) There exists $\lambda_c \geq 0$, such that $\hat{t}(\lambda) > \hat{t}(0)$ and $\hat{t}'(\lambda) > 0$ for all $\lambda > \lambda_c \geq 0$, if and only if $\theta_1/\theta_3 > \varepsilon_1/\varepsilon_3$.

Figure 5 illustrates Proposition 5 by showing the parameter region of PD under the Engel effect only. Condition a), which follows from eq.(29), holds on the left side of the vertical line, $\theta_1/\theta_3 = 1$ in Figure 5. Intuitively, when $\theta_1 < \theta_3$, the price of the income elastic services is high relative to the income inelastic agriculture in a higher-λ country, which makes it necessary to reallocate labor to services at earlier stage of development. Condition b), which follows from eq.(28), holds below the line connecting $(\theta_1/\theta_3, \theta_2/\theta_3) = (\varepsilon_1/\varepsilon_3, \varepsilon_2/\varepsilon_3)$ and $(\theta_1/\theta_3, \theta_2/\theta_3) = (1,1)$ in Figure 5. Intuitively, with θ_2 sufficiently low, which has no effect on $\widehat{U}(\lambda)$, the relative price of manufacturing is sufficiently low in a higher-λ country, which keeps the peak manufacturing share low. Third, $\hat{t}(\lambda)$ in eq.(27) is not generally monotone. But, when Conditions a) and b) are met, a sufficiently large λ makes the country peak later, i.e., $\hat{t}(\lambda) > \hat{t}(0)$ and $\hat{t}'(\lambda) > 0$ if and only if $\theta_1/\theta_3 > \varepsilon_1/\varepsilon_3$, which holds on the right side of the vertical line, $\theta_1/\theta_3 = \varepsilon_1/\varepsilon_3$ in Figure 5. Furthermore, $\hat{t}'(\lambda) > 0$ for all $\lambda > 0$ above the dashed line connecting $(\theta_1/\theta_3, \theta_2/\theta_3) = (\varepsilon_1/\varepsilon_3, \varepsilon_2/\varepsilon_3)$ and $(\theta_1/\theta_3, \theta_2/\theta_3) = (\Theta_E, 0)$, where $\Theta_E \equiv \frac{\tilde{\beta}_2\varepsilon_2/\varepsilon_3 + (1-\tilde{\beta}_2)\varepsilon_1/\varepsilon_3}{\tilde{\beta}_2\varepsilon_2/\varepsilon_3 + (1-\tilde{\beta}_2)}$ satisfies $\varepsilon_1/\varepsilon_3 < \Theta_E < 1$; it is increasing in $\tilde{\beta}_2$ with $\Theta_E \to \varepsilon_1/\varepsilon_3$ as $\tilde{\beta}_2 \to 0$ and $\Theta_E \to 1$ as $\tilde{\beta}_2 \to 1$.

Thus, even with the Engel effect only, the heterogeneity of the technology gap could cause PD. However, with $g_1 = g_2 = g_3 = \bar{g}$, these conditions imply $\theta_1\bar{g}, \theta_2\bar{g} < \theta_3\bar{g}$, that is, when cross-country productivity dispersion is *the largest in the service sector*, which is counterfactual. Precisely for this reason we used the Baumol effect only in our baseline model and added the Engel effect in an extension, not the other way around.

5. **The Second Extension: Opening Up for International Trade**

In the baseline model, we assume no international trade in order to stress that our mechanism does not hinge on the presence of international trade. Of course, countries do trade with each other in reality. How does the analysis need to be modified if when the countries trade in agriculture and manufacturing? Recall that one implication for PD in the baseline model is

$$\frac{\partial}{\partial\lambda}\ln\left(\frac{A_1(t)}{A_2(t)}\right) = -(\theta_1 g_1 - \theta_2 g_2) < 0.$$

Thus, a high-λ country has comparative advantage in manufacturing relative to a low-λ country, and a low-λ country has comparative advantage in agriculture relative to a high-λ country. This means that opening up for trade in these two sectors would weaken PD by allowing the high-λ country to export manufacturing to the low-λ country. This prediction is in line with the finding by Rodrik (2016) that East Asian countries, which experienced export-led growth in manufacturing "suffer" less from PD compared to Latin American countries.

To see this more formally, let us go back to the baseline model and introduce trade between two countries, whose technology gaps are given by $\lambda^1 < \lambda^2$, where superscript 1 indicates the leader and superscript 2 the laggard. Their population sizes are L^1 and L^2. Trade in agriculture and manufacturing is subject to the iceberg costs, such that only $e^{-\tau_1} < 1$ (i.e., $\tau_1 > 0$) fraction of the agriculture goods shipped and only $e^{-\tau_2} < 1$ (i.e., $\tau_2 > 0$) fraction of the manufacturing goods shipped arrive to the destination. Services are nontradeable. Let us denote the price of good-j in country-c by p_j^c. Then, from eq.(3), the expenditure share of sector-j in country-c is

$$m_j^c = \beta_j\left(\frac{p_j^c}{P^c}\right)^{1-\sigma}; \qquad P^c = \left[\sum_{k=1}^{3}\beta_k(p_k^c)^{1-\sigma}\right]^{1/(1-\sigma)}. \tag{30}$$

Likewise, from eq.(6), the employment (and value-added) share of j in c is:

(31) $$s_j^c = \left(A_j^c\right)^{\frac{1}{\alpha}}\left(\frac{p_j^c}{E^c}\right)^{\frac{1}{\alpha}}; \qquad E^c = \left[\sum_{k=1}^{3}(A_k^c)^{\frac{1}{\alpha}}(p_k^c)^{\frac{1}{\alpha}}\right]^{\alpha}.$$

Because services are nontradeable, its expenditure share must be equal to the value-added (and employment) share in each country, and hence $s_3^1 = m_3^1$ and $s_3^2 = m_3^2$.

For $g_1\theta_1 > g_2\theta_2$, the leader, Country-1, has comparative advantage in agriculture and hence may export agriculture. Since the agriculture goods produced in Country-1 cost $e^{\tau_1}p_1^1$ per unit in Country-2, they are not traded if $e^{\tau_1}p_1^1 > p_1^2 > p_1^1$. If they are traded, $e^{\tau_1}p_1^1 = p_1^2 > p_1^1$ and $[s_1^1 - m_1^1]E^1L^1 = [m_1^2 - s_1^2]E^2L^2 > 0$ is the total value of the agricultural trade.[30] Thus, we have the following complementarity condition;

$$[s_1^1 - m_1^1]E^1L^1 = [m_1^2 - s_1^2]E^2L^2 \geq 0; \qquad e^{\tau_1}p_1^1 \geq p_1^2 > p_1^1.$$

Likewise, the laggard, Country-2, has comparative advantage in manufacturing and hence may export manufacturing, so that we have the following complementarity condition:

$$[m_2^1 - s_2^1]E^1L^1 = [s_2^2 - m_2^2]E^2L^2 \geq 0; \qquad e^{\tau_2}p_2^2 \geq p_2^1 > p_2^2.$$

Indeed, these complementarity conditions can be consolidated into:

$$[s_1^1 - m_1^1]E^1L^1 = [s_2^2 - m_2^2]E^2L^2 \geq 0; \qquad e^{\tau_1+\tau_2} \geq \frac{p_1^2}{p_1^1}\frac{p_2^1}{p_2^2} > 1,$$

because the total value of the agriculture trade must be equal to the total value of the manufacturing trade in equilibrium.[31] Finally, from eq.(12), the ratio of the relative price difference in the absence of trade would be:

$$\frac{p_1^2}{p_1^1}\frac{p_2^1}{p_2^2} = \left[\frac{A_1^2(t)}{A_1^1(t)}\frac{A_2^1(t)}{A_2^2(t)}\right]^{-\frac{a}{(1-\sigma)}} = e^{\frac{a(g_1\theta_1 - g_2\theta_2)(\lambda^2-\lambda^1)}{(1-\sigma)}} \equiv e^{T_+} > 1,$$

where $T_+ \equiv \frac{a(g_1\theta_1 - g_2\theta_2)(\lambda^2-\lambda^1)}{1-\sigma} > 0$.

Thus, T_+ is a prohibitive level of the trade cost; for $\tau_1 + \tau_2 \geq T_+ > 0$, trade does not occur. Note that T_+ is proportional to $(g_1\theta_1 - g_2\theta_2)(\lambda^2 - \lambda^1)$, the strength of comparative advantage, which is in turn proportional to $\lambda^2 - \lambda^1 > 0$, the distance between the two countries in their technology gaps. For $0 < \tau_1 + \tau_2 < T_+$, Country-2, the laggard, exports manufacturing

[30] To see this, the volume of the agriculture import of Country-2 is equal to $e^{-\tau_1}$ fraction of the volume of the agriculture goods shipped from Country-1, so that $e^{-\tau_1}[Y_1^1 - c_1^1L^1] = c_1^2L^2 - Y_1^2$. By multiplying both sides by $e^{\tau_1}p_1^1 = p_1^2$; $p_1^1[Y_1^1 - c_1^1L^1] = p_1^2[c_1^2L^2 - Y_1^2]$, which is equivalent to $[s_1^1 - m_1^1]E^1L^1 = [m_1^2 - s_1^2]E^2L^2$.

[31] To see this, note that $s_3^c = m_3^c$ implies $s_1^c + s_2^c = 1 - s_3^c = 1 - m_3^c = m_1^c + m_2^c$ and hence $s_1^c - m_1^c = m_2^c - s_2^c$. Thus, the first condition of each of the two complementarity conditions is equivalent.

to Country-1, the leader, which exports agriculture to Country-2, with $e^{\tau_1}p_1^1 = p_1^2$; $e^{\tau_2}p_2^2 = p_2^1$, $[s_1^1 - m_1^1]E^1L^1 = [m_2^1 - s_2^1]E^1L^1 = [s_2^2 - m_2^2]E^2L^2 = [m_1^2 - s_1^2]E^2L^2 > 0$; $s_3^1 = m_3^1$ and $s_3^2 = m_3^2$, and eqs.(30)-(31). And the employment (and the value-added) share of manufacturing goes up, while its expenditure share goes down, in Country-2.

But what are the effects on $\hat{t}$, the peak time, $\widehat{s_2}$, the peak manufacturing employment and value-added share, and $\widehat{U}$, and the peak time per capita income? In spite of our best effort, we are unable to answer this question analytically. So, we have solved them numerically. Using the calibrated values obtained in Section 3.4 and $L_1/L_2 = 1$, we plot how $\hat{t}$, $\widehat{s_2}$, and $\widehat{U}$ respond to

$$\tau \equiv \frac{\tau_1 + \tau_2}{T_+} \equiv \frac{(1-\sigma)(\tau_1 + \tau_2)}{a(g_1\theta_1 - g_2\theta_2)(\lambda^2 - \lambda^1)} < 1,$$

the trade cost normalized such that no trade would take place if $\tau \geq 1$. In Figure 6, we use the values calibrated to match Rodrik (2016; Table 10)'s finding of PD between pre-1990 and post-1990 countries in terms of the employment shares with the productivity growth rates taken from Duarte and Restuccia (2010; Section IIB). This figure shows that our mechanism for PD is robust to opening up for trade, but, as τ declines from 1, it becomes weaker in the sense that the differences between the two countries in $\hat{t}$ and in $\hat{s}_2$ becomes smaller (but those in $\widehat{U}$ and in $\widehat{m}_2$ become larger). Figure 6 also shows that, as τ declines further, the reversal of $\hat{s}_2$ occurs at $\tau \approx 0.91$, or $e^{\tau_1+\tau_2} = e^{\tau T_+} \approx 2.242$ and the reversal of $\hat{t}$ occurs at $\tau \approx 0.85$, or $e^{\tau_1+\tau_2} = e^{\tau T_+} \approx 1.986$. Since $e^{\tau_1+\tau_2} = (p_1^2/p_1^1)(p_2^1/p_2^2)$ when trade occurs, these critical values correspond to the goods being, on average, $\sqrt{2.242} \approx 1.497$ and $\sqrt{1.986} \approx 1.41$ times higher in the importing country.[32] Thus, PD holds as long as the trade cost accounts for more than 1/3 of the imported good prices. This seems well within the empirically plausible range of the trade cost in the "tradeable" goods. For example, Burstein, Neves, and Rebelo (2003) found the share of the distribution cost in the imported goods prices, which should be included in our "trade costs," to be around 42% for the US and higher for other countries.

[32]The results are almost identical if we use the values calibrated to match Rodrik (2016; Table 10)'s finding in terms of the value-added shares, with the reversal of $\hat{s}_2$ at $\tau \approx 0.90$ or $e^{\tau_1+\tau_2} \approx 2.244$, and the reversal of $\hat{t}$ at $\tau \approx 0.87$ or $e^{\tau_1+\tau_2} \approx 2.185$. These correspond to the goods being $\sqrt{2.244} \approx 1.498$ and $\sqrt{2.185} \approx 1.478$ times higher in the importing country. If the productivity growth rates are taken from Comin-Lashkari and Mestieri (2021; TABLE VIII) instead, the results are similar, except that the reversal of $\hat{s}_2$ occurs at a lower trade cost than the reversal of $\hat{t}$. [The numbers are $\tau \approx 0.77$ or $e^{\tau_1+\tau_2} \approx 1.947$ and $\tau \approx 0.96$, or $e^{\tau_1+\tau_2} \approx 2.295$, so that $\sqrt{1.947} \approx 1.395 < \sqrt{2.295} \approx 1.515$ for the employment shares and $\tau \approx 0.76$ or $e^{\tau_1+\tau_2} \approx 2.068$ and $\tau \approx 0.96$, or $e^{\tau_1+\tau_2} \approx 2.504$, so that $\sqrt{2.068} \approx 1.438$ and $\sqrt{2.504} \approx 1.582$ for the value-added shares.]

6. The Third Extension: Introducing Catching Up

Until now, we have assumed that the technology gap λ is a time-invariant parameter. This implies that the sectoral productivity growth rate is constant over time and identical across countries.[33] We have made this assumption to focus on the "level effect" of the technology gap, by shutting down the potential "growth effect."

What happens if we allow, in the baseline model, latecomers to achieve a higher productivity growth in each sector by *narrowing a technology gap* over time. More specifically, suppose that countries differ only in the *initial* value of the technology gap, λ_0, but technology gap shrinks exponentially over time at the common rate, $g_\lambda > 0$. Thus,

$$\lambda_t = \lambda_0 e^{-g_\lambda t}, \qquad g_\lambda > 0,$$

which preserves the country ranking and $\tilde{A}_j(t) = \bar{A}_j(0)e^{g_j(t-\theta_j\lambda_t)}$. Then, the time path of the manufacturing share is given by:

$$\frac{1}{s_2(t)} = \left(\frac{\tilde{\beta}_1}{\tilde{\beta}_2}\right) e^{a[(\theta_1 g_1 - \theta_2 g_2)\lambda_t - (g_1 - g_2)t]} + 1 + \left(\frac{\tilde{\beta}_3}{\tilde{\beta}_2}\right) e^{a[(\theta_3 g_3 - \theta_2 g_2)\lambda_t + (g_2 - g_3)t]}.$$

Fujiwara and Matsuyama (2023; Section 5) derived the peak time, $\hat{t}(\lambda_0)$, the peak manufacturing share, $\hat{s}_2(\lambda_0)$, and the peak time per capita income, $\widehat{U}(\lambda_0)$, as functions of the initial value, λ_0, but these peak time values need to be solved numerically. Figure 7 shows the result for the same parameter values with Example 2. These numerical solutions suggest that higher-λ countries peak later in time and have lower peak manufacturing shares for a given g_λ. For the peak time per capita income, they have lower peak time per capita income, *unless g_λ is too high*.[34] This result makes sense because, in the baseline model, higher-λ countries have lower peak time per capita income because a delay caused by higher-λ is not long enough to make up for their longer adoption lags. With a fast catching up, these countries can narrow their gaps and experience faster productivity growth during that delay. Another notable feature is that the effect of a higher

[33] In contrast, even with a time-invariant technology gap, the aggregate growth rate, $g_U(t) \equiv U'(t)/U(t) = \sum_{k=1}^{3} g_k s_k(t)$, declines over time due to the reallocation from high productivity growth sectors to low productivity growth sectors. This can be verified as $g_U'(t) = g_1 s_1'(t) + g_2 s_2'(t) + g_3 s_3'(t) = (g_1 - g_2)s_1'(t) + (g_3 - g_2)s_3'(t) < 0$. This is what Nordhaus (2008) called the sixth symptom of the Baumol's diseases.

[34] One route through which catching up could take place is global technology diffusion. See Acemoglu (2009, Ch.18) and Comin and Mestieri (2018), just to name a few. According to Comin and Mestieri (2018, TABLE III), the extensive margin of the adoption lags shrinks at the rate around 1.0%, but the intensive margin of the adoption lags diverges. This suggests $g_\lambda < 1.0\%$, when the two margins are combined.

g_λ is nonmonotonic, and the graphs of $\hat{t}(\lambda_0)$ for different values of g_λ, though all monotonically increasing λ_0, cross with each other.

7. Concluding Remarks

In this paper, we presented a parsimonious mechanism for generating what Rodrik (2016) called premature deindustrialization (PD). In the baseline model, the hump-shaped path of the manufacturing share, along with the declining agricultural share and the increasing service share, is caused by the productivity growth rate differences across the three sectors, as in Baumol (1967) and Ngai and Pissarides (2007). Countries follow different hump-shaped paths of the manufacturing share due to the differences in their adoption lags in the three sectors. To keep it simple, we further assume that the countries differ only in one dimension, the "technology gap," which has differential effects on the adoption lags in the three sectors. In this setup, we identified the sufficient and necessary condition for PD, i.e., when countries with larger technology gaps reach their manufacturing peaks later in time, but earlier in per capita income with lower peak manufacturing shares. Intuitively, our mechanism suggests that the manufacturing shares among latecomers reach their peaks *later* due to their lower productivity in agriculture *but prematurely* due to the long adoption lag in services, while the peak manufacturing shares are *lower* because their productivity is not so behind in manufacturing.

It turns out that this condition for PD implies that the cross-country productivity dispersion is the largest in agriculture. In contrast, the relative magnitude of the cross-country productivity dispersions in manufacturing and services does not play a crucial role. However, when the model is calibrated to match Rodrik's findings, we found that the cross-country productivity dispersion is the smallest in manufacturing.

In the baseline model, the sectoral demand composition is generated by homothetic CES (to focus on the Baumol effect). Furthermore, it is a closed economy (to isolate the role of trade), and there is no catching up in technology adoption by late industrializers (to isolate the level effect of the technology gap from its growth effect). To demonstrate the robustness of our mechanism, we considered the following three extensions. First, we added the Engel effect on top of the Baumol effect so that the hump-shaped path of the manufacturing share is also shaped by nonhomothetic demand with the income elasticities being the largest in services and the smallest in agriculture, using *isoelastic nonhomothetic CES* preferences introduced by Comin, Lashkari, and Mestieri (2021). Even though combining the Engel effect with the Baumol effect changes the shape of the time path, it does not change the main implications on how the technology gap generates PD. We also showed that, if we had relied solely on the Engel effect without the Baumol effect, PD would have occurred only under the conditions that would imply, counterfactually, that the cross-country productivity dispersion is the largest in the service sector. Second, we also introduced international trade to demonstrate that the mechanism is robust, even though it weakens the mechanism, which is consistent with another finding by Rodrik that East Asian countries that export manufacturing "suffer" less for PD. Finally, in the third extension, we allowed late industrializers to catch up by narrowing the technology gaps over time and showed that the main results carry over, unless the catching-up speed is too high.

Throughout the analysis, we assumed that the productivity growth rates of the frontier technology in each sector, as well as the technology gap of each country, are exogenous, and that the resources are allocated in competitive equilibrium. The equilibrium allocation is hence efficient in the absence of any distortions.[35] Furthermore, because of the diminishing returns to labor due to the presence of the managerial skill specific to the sector, the prediction of our analysis is robust to introducing trade. A natural next step is to open up the black boxes and offer micro foundations for the productivity growth rates and the technology gaps through innovation and imitation by profit-seeking firms and/or human capital accumulations. Such extensions naturally introduce the market size effects and externalities with nontrivial welfare implications, as known in the directed technological change literature, surveyed by Acemoglu (2009, Ch.15). Furthermore, if the productivity growth rate differences across sectors respond endogenously to the market size differences, the Baumol effect and the Engel effect become intrinsically intertwined, as in Matsuyama (2019) and Bohr, Mestieri and Yavuz (2021). The market size effects on endogenous productivity growth could also create another potential cause for PD, international trade, where the manufacturing firms based on late industrializers have disadvantages in competing against those based on earlier industrializers in the world market.[36]

[35]In our mechanism, deindustrialization is "premature" in the sense that late industrializers reach their manufacturing peaks at lower productivity levels than early industrializers, because the delays in reaching the peaks caused by larger technology gaps are *shorter* than the adoption lags they create.

[36]This is indeed what Rodrik (2016) might have in mind when he speculated globalization as a possible cause for PD. In Fujiwara and Matsuyama (Work in Progress), we seek to rationalize his hypothesis, using a variant of the two-country monopolistic competition model of trade in Matsuyama (2019), in which the rich country enjoys comparative advantage in manufacturing relative to the poor country through the home market effect as in Krugman (1980).

Needless to say, we do not believe that that our mechanism is the sole cause for PD. In principle, there may be many complementary mechanisms that could be jointly responsible for generating PD in the real world. We are merely suggesting that our mechanism should be one of them, as it can go a long way toward understanding PD. We hope that our analysis will stimulate further research on structural change, particularly on PD, such that there will be many complementary mechanisms, which could be combined to generate a better understanding of PD.

References

Acemoglu, Daron. 2009. *Introduction to Modern Economic Growth*, Princeton, Princeton University Press.

Acemoglu, Daron, Philip Aghion, and Fabrizio Zilibotti. 2006, "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association,* 4 (1), 37-74.

Acemoglu, Daron, and Veronica Guerrieri. 2008. "Capital Deepening and Nonbalanced Economic Growth," *Journal of Political Economy,* 116(3),

Atkin, David, Arnaud Costinot, and Masao Fukui. 2022. "Globalization and the Ladder of Development: Pushed to the Top or Held at the Bottom," Unpublished.

Baumol, William J., 1967. "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crises," *American Economic Review,* 57, 415-426.

Beckman, Martin J., 1977. "Management Production Functions and the Theory of the Firm," *Journal of Economic Theory*, 14, 1-18.

Bohr, Clement E., Martí Mestieri and Emre Enes Yavuz. 2021. "Engel's treadmill: the perpetual pursuit of cornucopia," Unpublished.

Boppart, Timo. 2014. "Structural Change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences," *Econometrica,* 82, 6, 2167-2196.

Buera, Francisco J., and Joseph P. Kaboski. 2012a, "The Rise of the Service Economy," *American Economic Review.* 102(6), 2540-69.

Buera, Francisco J., and Joseph P. Kaboski, 2012b, "Scale and the Origins of Structural Change," *Journal of Economic Theory.* 147(2), 684-712.

Buera, Francisco J., Joseph P. Kaboski, Richard Rogerson, and Juan I. Vizcaino. 2022, "Skill-Biased Structural Change," *Review of Economic Studies,* 89(2), 592-625.

Burstein, Ariel T., Joao C. Neves, and Sergio Rebelo. 2003. "Distribution costs and real exchange rate dynamics during exchange-rate based stabilizations," *Journal of Monetary Economics*, 50(6), 1189-1214.

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In *Handbook of Economic Growth*, vol. 1A, edited by Philippe Aghion and Steven N. Durlauf, 679–742. New York: Elsevier.

Comin, Diego, Bart Hobijn, and Emilie Rovito, 2008, "Technology Usage Lags," *Journal of Economic Growth*, 13, 237-256.

Comin, Diego, and Martí Mestieri, 2014, "Technology Diffusion: Measurement, Causes and Consequences," *Handbook of Economic Growth.*

Comin, Diego, and Martí Mestieri. 2018, "If Technology Has Arrived Everywhere, Why Has Income Diverged?" *American Economic Journal: Macroeconomics,* 10(3), 137-178.

Comin, Diego, Danial Lashkari, and Martí Mestieri, 2021, "Structural Transformation with Long-run Income and Price Effects," *Econometrica*, 89, 1, 311-374.

Costinot, Arnaud. 2009. "An Elementary Theory of Comparative Advantage," *Econometrica,* 77, 1165-1192.

Costinot, Arnaud, and Jonathan Vogel, 2010, "Matching and Inequality in the World Economy," *Journal of Political Economy,* 118, 747-786.

Costinot, Arnaud, and Jonathan Vogel, 2015. "Beyond Ricardo: Assignment Models in International Trade," *Annual Review of Economics,* 7, 31-62.

Cravino, Javier, and Sebastien Sotelo. 2019. "Trade-Induced Structural Change and the Skill Premium," *American Economic Journal: Macroeconomics,* 11(3), 289-326.

Dasgupta, Sukti, and Ajit Singh, “Manufacturing, services, and premature deindustrialization in developing countries: A Kaldorian analysis. UNU-WIDER, United Nations University Research Paper, No.2006/49.
Duarte, Margarida, and Diego Restuccia, 2010, “The Role of the Structural Transformation in Aggregate Productivity,” *Quarterly Journal of Economics,* 125(1), 129-173.
Foellmi, Reto, and Josef Zweimüller. 2006. “Income Distribution and Demand-Induced Innovations,” *Review of Economic Studies*. 73(4), 941-960.
Foellmi, Reto, and Josef Zweimüller. 2014. “The Macroeconomics of Model T,” *Journal of Economic Theory*. 153, 617-647.
Fujiwara, Ippei, and Kiminori Matsuyama. “A Global View of Premature Deindustrialization,” Work in Progress, Keio and Northwestern Universities.
Fujiwara, Ippei, and Kiminori Matsuyama. 2023. “A Technology-Gap Model of Premature Deindustrialization,” CEPR Discussion Paper #15530-Version 3.
Fujiwara, Ippei, and Kiminori Matsuyama. 2024. “A Technology-Gap Model of Premature Deindustrialization,” CEPR Discussion Paper #15530-Version 4.
Garcia-Santana, Manuel., Josep Pijoan-Mas, and Lucciano Villacorta. 2021. “Investment Demand and Structural Change,” *Econometrica*, 89(6), 2751-2785.
Gollin, Douglas, David Lagakos, and Michael E. Waugh, 2014a, “Agricultural Productivity Differences across Countries,” *American Economic Review, Papers & Proceedings,* 104(5): 165-170.
Gollin, Douglas, David Lagakos, and Michael E. Waugh, 2014b, “The Agricultural Productivity Gap,” *Quarterly Journal of Economics,* 9129(2), 939-993.
Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi. 2014. “Growth and Structural Transformation,” in *Handbook of Economic Growth*, ed. by Philippe. Aghion and Steven N. Durlauf, 855–941.
Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi. 2021. “Structural Change in Investment and Consumption: A Unified Analysis, *The Review of Economic Studies,* 88(3), 1311-1346.
Huneeus, Federico, and Richard Rogerson. 2020. “Heterogeneous Paths of Industrialization,” Princeton University.
Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie. 2001. “Beyond Balanced Growth,” *Review of Economic Studies,* 68: 869-882.
Kravis, Irving, Alan Heston, and Robert Summers. 1982. “The share of services in economic growth” *In Global Econometrics: Essays in Honor of Lawrence R. Klein,* ed. F. G. Adams and Bert Hickman, Cambridge: MIT Press.
Krugman, Paul., 1980. “Scale Economies, Product Differentiation, and the Pattern of Trade,” *American Economic Review*, 70(5), 950-959.
Krugman, Paul. 1985. A ‘Technology-Gap’ Model of International Trade. In *Structural Adjustment in Developed Open Economies*, ed. K. Jungenfelt and D. Hague. London, Macmillan.
Lewis, Logan, Ryan Monarch, Michael Sposi, and Jing Zhang, 2022. “Structural Change and Global Trade,” *Journal of the European Economic Association,* 20(1): 476-512.
Lucas, Robert E. Jr., 1978, “On the Size Distribution of Business Firms,” *The Bell Journal of Economics,* Autumn, Vol.9, no.2: 508-523.
Matsuyama, Kiminori. 1992. "Agricultural Productivity, Comparative Advantage, and Economic Growth," *Journal of Economic Theory* 58: 317-334.

Matsuyama, Kiminori. 2000. "A Ricardian Model with a Continuum of Goods under Nonhomothetic Preferences: Demand Complementarities, Income Distribution and North-South Trade," *Journal of Political Economy* 108: 1093-1120.
Matsuyama, Kiminori. 2002. "The Rise of Mass Consumption Societies," *Journal of Political Economy,* 110: 1035-1070.
Matsuyama, Kiminori. 2005. “Credit Market Imperfections and Patterns of International Trade and Capital Flows,” *Journal of European Economic Association,* 3(2-3), 714-723.
Matsuyama, Kiminori. 2009. "Structural Change in an Interdependent World: A Global View of Manufacturing Decline," *Journal of the European Economic Association.* 7: 478-486.
Matsuyama, Kiminori. 2019. “Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade,” *Econometrica,* 87(2), 497-528.
Matsuyama, Kiminori. 2023. “Non-CES Aggregators: A Guided Tour,” *Annual Review of Economics,* 15, 235-265.
Murphy, Kevin, Andrei Shleifer, and Robert Vishny. 1989. “Income Distribution, Market Size, and Industrialization,” *Quarterly Journal of Economics,* 104, 537-564.
Ngai, L. Rachel, and Christopher A. Pissarides, 2007, “Structural Change in a Multisector Model of Growth,” *American Economic Review,* 97, 429–443.
Ngai, L. Rachel, and Christopher A. Pissarides, 2008, “Trends in Hours and Economic Growth,” *Review of Economic Dynamics,* 11(2): 239-256.
Nordhaus, William D., 2008. “Baumol’s diseases: A macroeconomic analysis,” *The BE Journal of Macroeconomics.*8(1).
Palma, José Gabriel, 2014. “Industrialization, ‘premature” deindustrialization and the Dutch Disease. Revista NECAT 3(5), 7-23.
Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu. 2008. “Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis,” *Journal of Monetary Economics,* 55(2), 234-250.
Rodrik, Dani. 2013. “Unconditional Convergence in Manufacturing,” *Quarterly Journal of Economics,* 128(1), 165-204.
Rodrik, Dani. 2016. “Premature Deindustrialization,” *Journal of Economic Growth,* 21, 1-33.
Sposi, Michael, Kei-Mu Yi, and Jing Zhang. 2021. “Deindustrialization and Industry Polarization,” NBER Working Paper No.29483.
Uy, Timothy, Kei-Mu Yi, and Jing Zhang. 2013. “Structural Change in an Open Economy,” *Journal of Monetary Economics*, 60, 667-682.

Figure 1: Conditions for PD only with the Baumol Effect

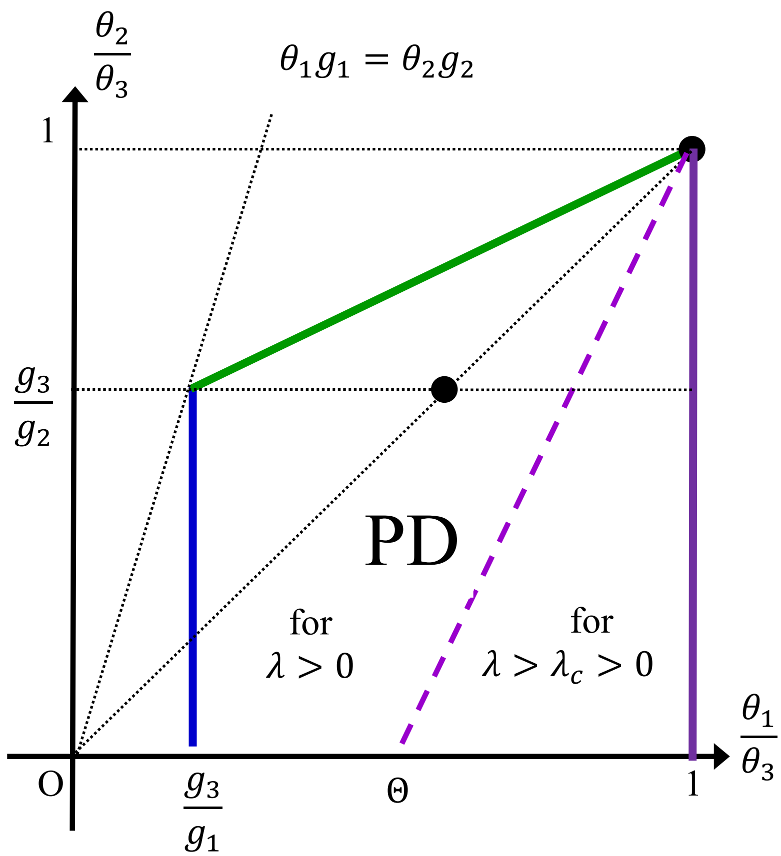


Figure 2: Premature Deindustrialization under the Baumol Effect: Numerical Illustration

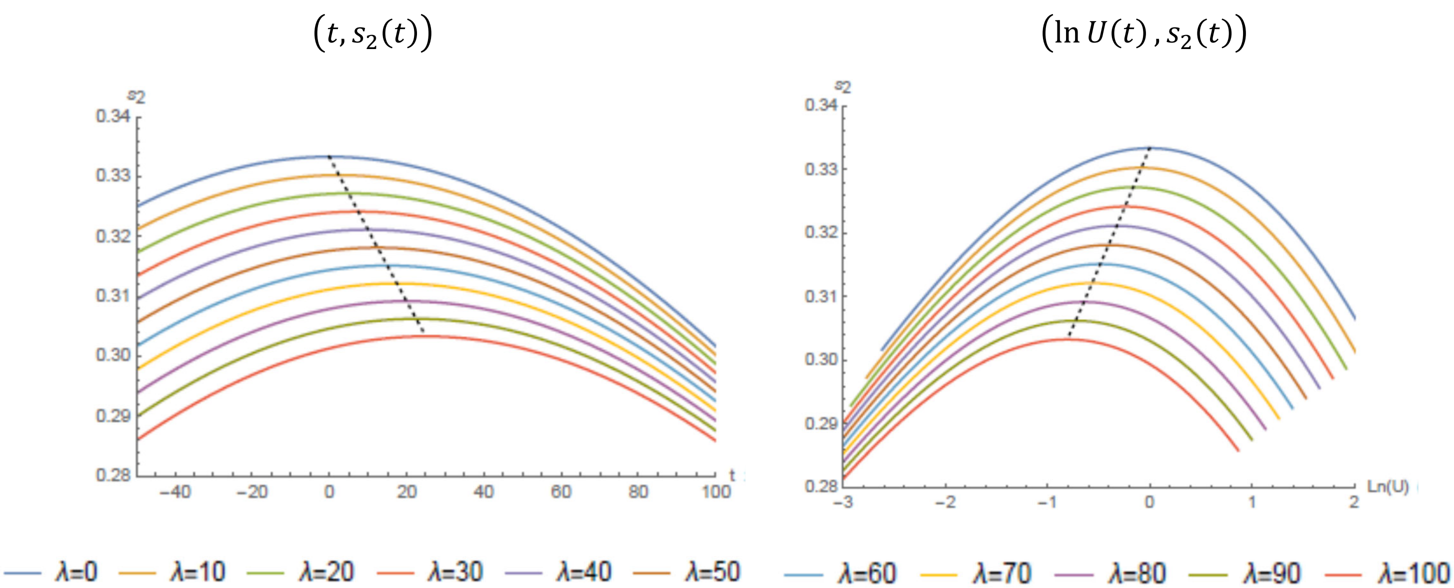


Figure 3: Premature Deindustrialization: Calibration

g_1 = 3.8% > g_2 > 2.4% > g_3 = 1.3%, as in Duarte-Restuccia (2010)

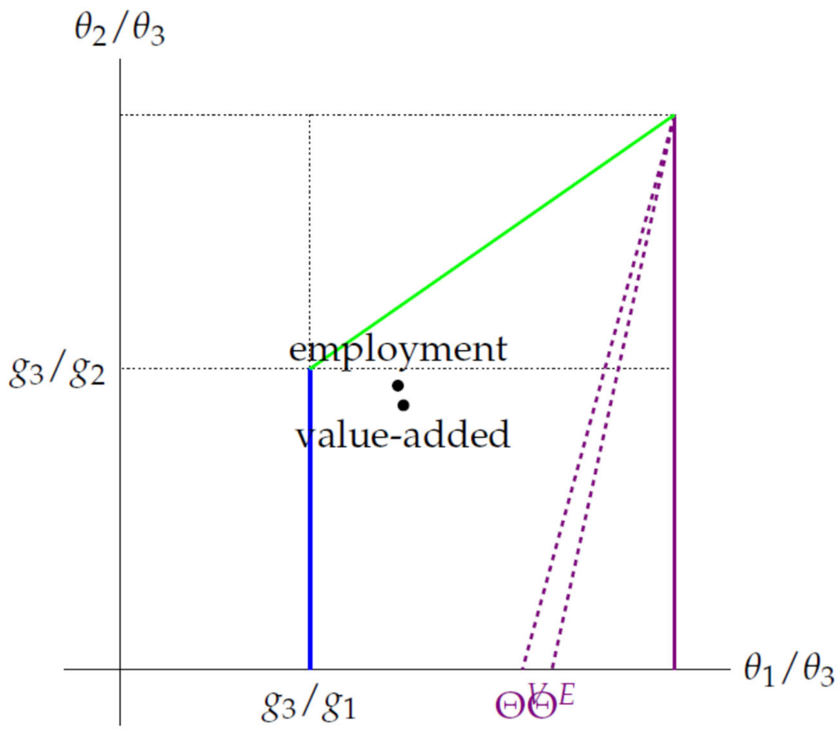
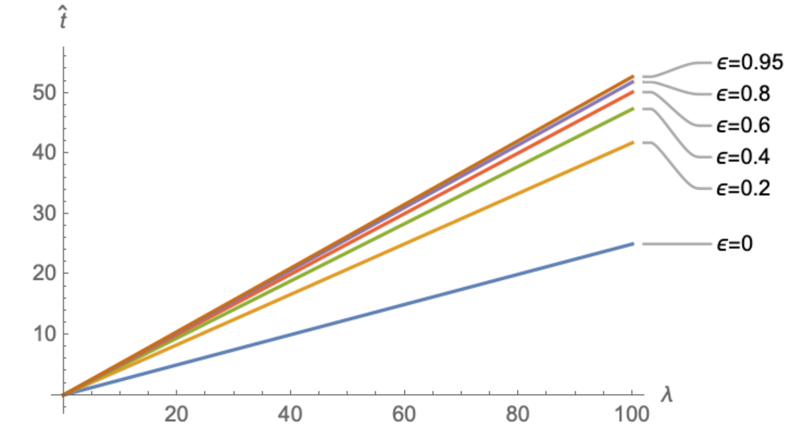


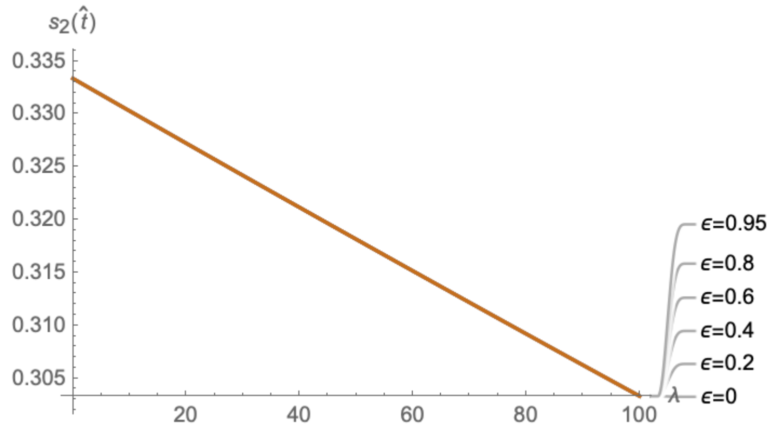
Figure 4: Introducing the Engel Effect

Figure 4a: $g_1 - g_2 = g_2 - g_3 = 1.2\% > 0 \Rightarrow \varepsilon_1 = 1 - \epsilon < \varepsilon_2 = 1 < \varepsilon_3 = 1 + \epsilon$ for $0 < \epsilon = (1.2\%)\mu < 1$

Peak Time



Peak Manufacturing Share



Peak Time Per Capita Income

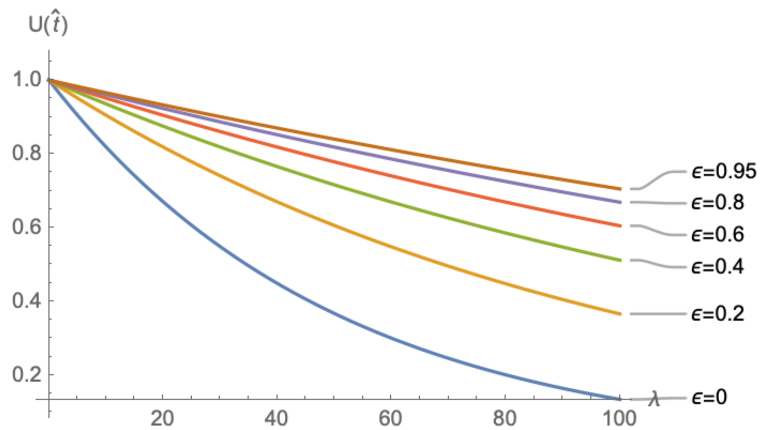


Figure 4b: Stronger nonhomotheticity significantly changes the shape of the time paths, but has little effects on how λ affects $\hat{t}, \widehat{s_2}$, and $\widehat{U}$.

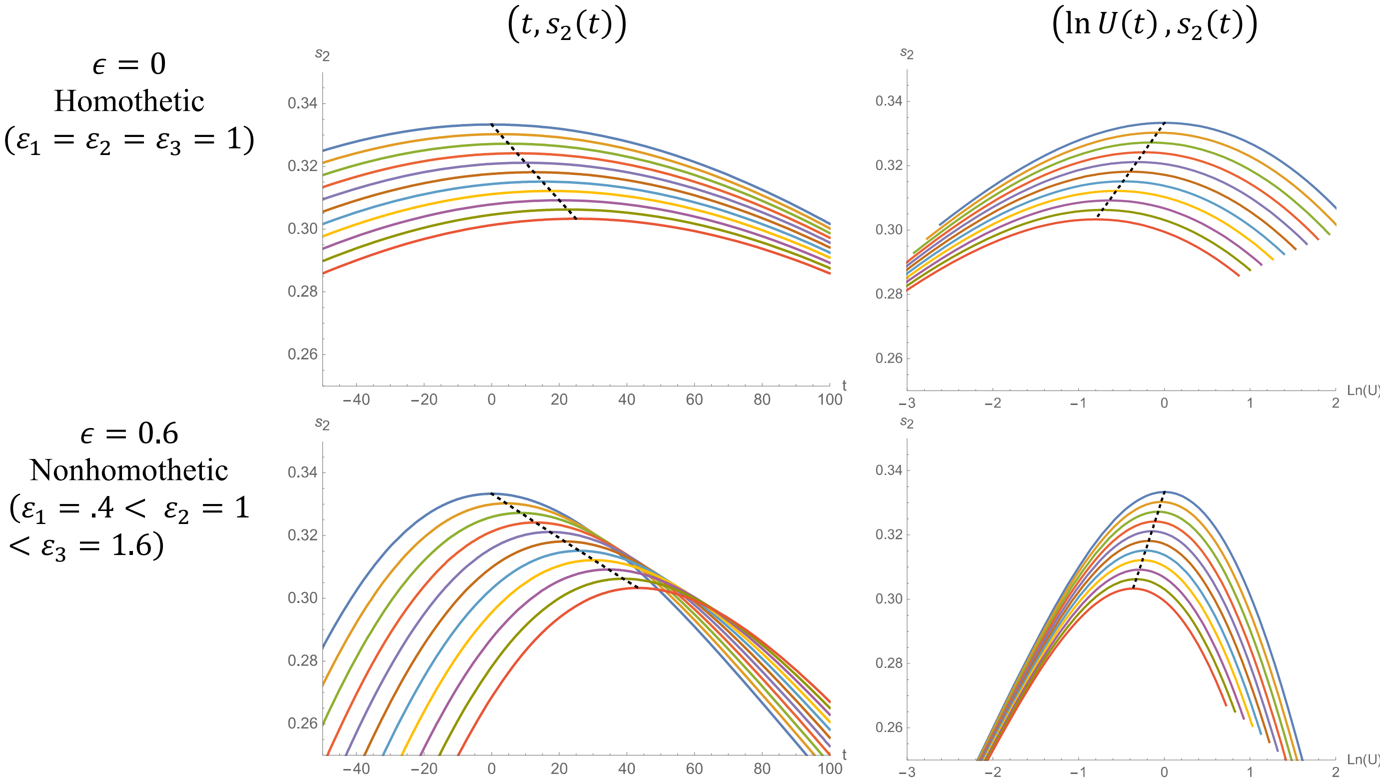


Figure 5: Conditions for Premature Deindustrialization (PD) only with the Engel Effect

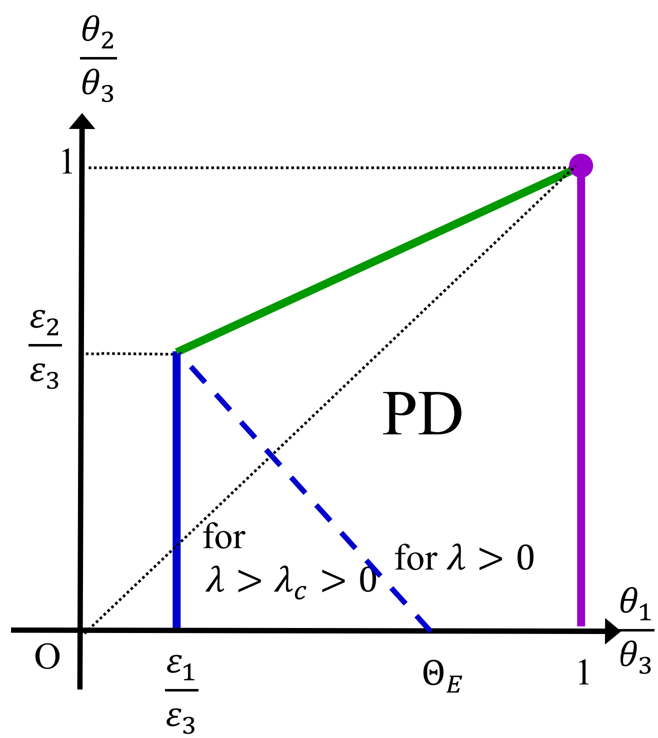
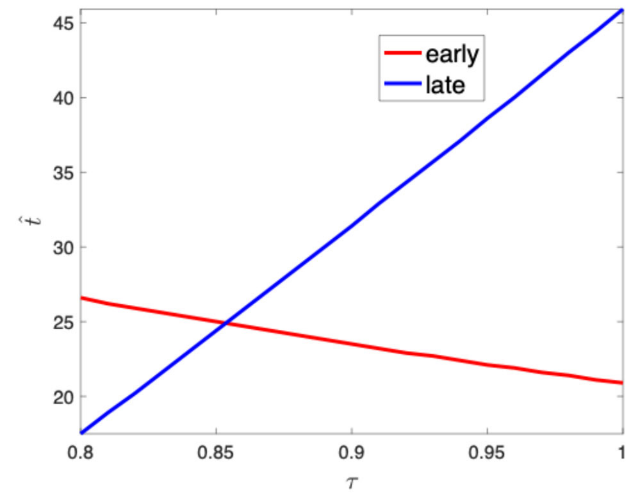
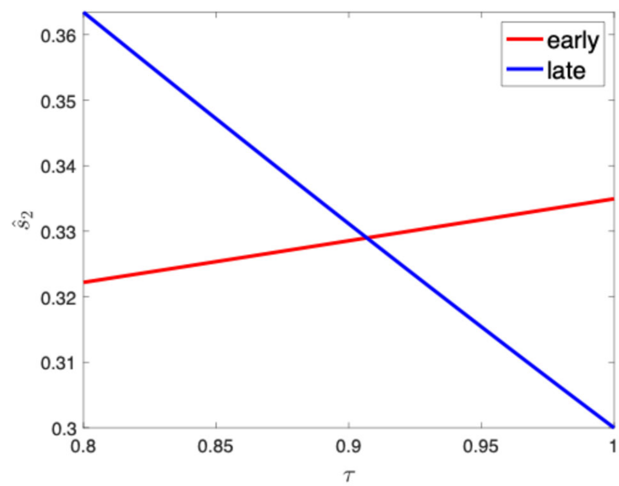


Figure 6: The Effects of Opening Up for Trade

Peak Time



Peak Manufacturing Share



Peak Time Per Capita Income

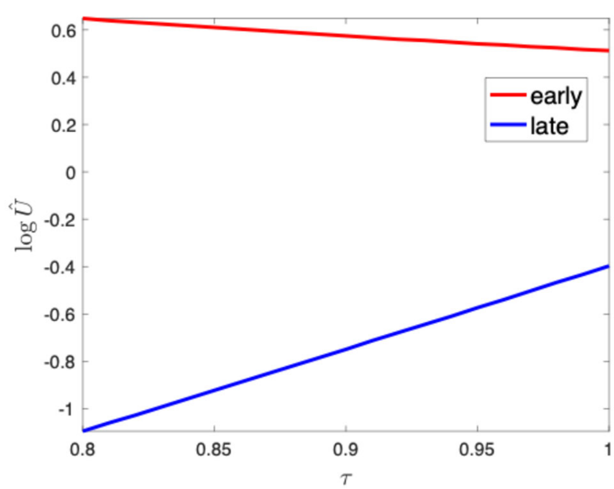
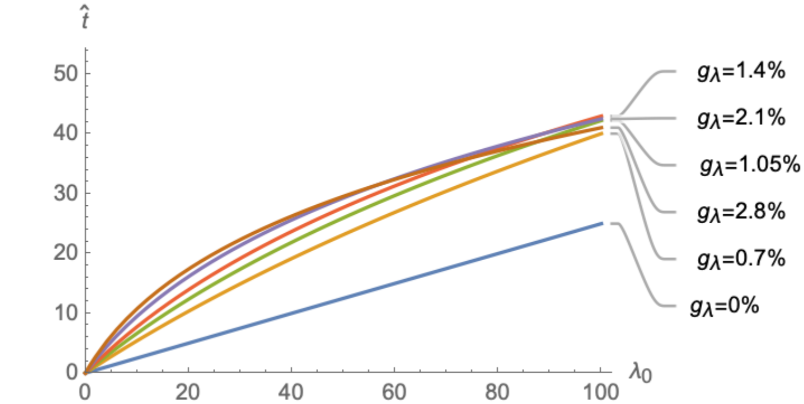
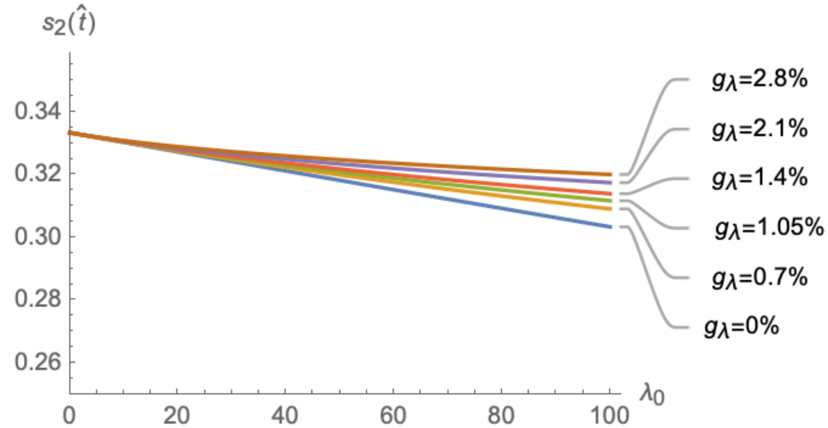


Figure 7: Catching Up

Peak Time



Peak Manufacturing Share



Peak Time Per Capita Income