CIRJE-F- 1190

A Technology-Gap Model of Premature Deindustrialization

Ippei Fujiwara
Keio University

Kiminori Matsuyama
Northwestern University and The University of Tokyo

March 2022

A Technology-Gap Model of Premature Deindustrialization*

Ippei Fujiwara
Faculty of Economics, Keio University
ippei.fujiwara@keio.jp;
Crawford School of Public Policy, ANU
ippei.fujiwara@anu.edu.au

Kiminori Matsuyama
Department of Economics
Northwestern University
k-matsuyama@northwestern.edu

First Version, December 4, 2020
This Version, March 23, 2022

Abstract

This paper presents a parsimonious mechanism for generating what Rodrik (2016) called premature deindustrialization (PD); the tendency that, compared to early industrializers, late industrializers reach their peaks of industrialization later in time, but earlier in per capita income, with lower peak manufacturing shares. In the baseline model, the hump-shaped path of the manufacturing sector is solely driven by the Baumol (1967) effect with the productivity growth rates of the frontier technology being the highest in agriculture and the lowest in services. The countries are heterogeneous only in the "technology gap," their capacity to adopt the frontier technology, which might affect adoption lags across sectors differently. In this setup, we show that PD occurs when the following three conditions are met; i) the impact of the technology gap on the adoption lag is larger in services than in agriculture, ii) in spite of its relatively shorter adoption lag, the productivity growth rate is sufficiently higher in agriculture than in services that the cross-country productivity dispersion is larger in agriculture than in services; and iii) the impact of the technology gap on the adoption lag is not too large in manufacturing. It turns out that these conditions for PD jointly imply that the cross-country productivity dispersion is the largest in agriculture.

In the first of the two extensions, we add the Engel effect on top of the Baumol effect so that the hump-shaped path of manufacturing is also shaped by nonhomothetic demand with the income elasticities being the largest in services and the smallest in agriculture. Even though adding the Engel effect to the Baumol effect changes the shape of the path, it does not change the main implications on how the technology gap generates PD. We also show that, if we had relied solely on the Engel effect, PD would occur only under the conditions that would imply that the cross-country productivity dispersion is the largest in services. In the second extension, we allow late industrializers to catch up by narrowing the technology gaps over time and show that the main results carry over, unless the catching-up speed is too high.

Keywords: Premature deindustrialization, technology gaps, adoption lags, The Baumol effect, The Engel effect, Catching-up
JEL classifications: O11, O14, O33

*We thank T. Boppart, F. Buera, S. Kikuchi, M. Mestieri, R. Ngai, D. Rodrik, and K. Yi seminar participants at (chronologically) Hitotsubashi Conference on International Trade and FDI, Kent MaGHiC Lecture, ES China and ES Australia, FRB-Chicago, Monash, Melbourne, ES European Summer, STEG, St.Gallen-Zürich, Oxford, Tokyo, FRB-Dallas, Peking, and CIGS for their feedback. This project started in December 2018 during Matsuyama's visit to Keio University, whose hospitality he gratefully acknowledges. 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). 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, in particular, his Figure 5. Rodrik called this pattern "premature deindustrialization."

In this paper, we present a parsimonious mechanism for generating premature deindustrialization (PD). There are three competitive sectors: agriculture, manufacturing, and services, which produce the consumption goods that are gross complements. As the frontier technology improves in each sector, productivity 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 three sectors, as in Baumol (1967) and Ngai and Pissarides (2007). The only source of heterogeneity across countries is their ability to adopt the frontier technology, which we call "technology gap," following Krugman (1985). Unlike Krugman, however, we allow for the possibility that the extent to which the country's technology gap affects its adoption lags varies across sectors. In this framework, we investigate the conditions under which PD occurs.

To see the importance of the differential impacts on 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, poorer countries follow the same path with richer countries, reaching the same peak manufacturing shares at the same level of the per capita income. Hence, PD could not occur.

Instead, suppose that the technology gap has a larger impact on the adoption lag in services than in agriculture, but 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.[1] Then, poorer countries reach their peaks later in time than earlier industrializers, 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. It turns out that these conditions for PD jointly imply that the cross-country productivity dispersion is the largest in agriculture, as empirically observed.[2] On the other hand, they impose no restriction on the relative magnitude of the cross-country productivity dispersions between manufacturing and services.[3]

In the baseline model, 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). In view of the importance of the Engel effect as the driver of structural change, we extend the baseline model by adding the Engel effect. As expected, 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 on the mechanism of PD presented by the baseline model. Furthermore, if we had relied solely on the Engel effect without the Baumol

---

[1]What is crucial here is 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.

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

[3]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. We remain agnostic, because the relative magnitude of the cross-country productivity dispersions in the two sectors is not crucial for our mechanism.

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. We also extend the baseline model to allow for 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 presenting our technology-gap model of PD, we do not mean to suggest that the countries differ only in the technology gap or that the technology gap is the sole cause for PD.[4] 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 surveyed by Herrendorf, Rogerson, and Valentinyi (2014) and many subsequent studies have 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, the role of trade,[5] sector-specific factor intensities,[6] home production,[7] consumption vs. investment,[8] productivity gaps,[9] endogenous productivity and externalities,[10] and much more. Most recent calibration studies in this 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 heterogeneity across countries obscure what could be driving forces behind PD. In this paper, we instead opt for a parsimonious approach by tying our hands to restrict ourselves to one dimension of exogenous cross-country heterogeneity to explain many dimensions of endogenous cross-country heterogeneity.[11]

The rest of the paper is organized as follows. In Section 2, we set up the three-sector model on structural change, driven solely by the Baumol effect, and show how adoption lags, and derive the analytical expressions for the peak time, the peak manufacturing share, and the

[4]In this respect, Huneeus and Rogerson (2020) deserve special mention. They propose an alternative mechanism, where the cross-country difference in agricultural productivity growth rate is the main driver of PD in the presence of the subsistence level of agriculture goods consumption.
[5]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).
[6]See, e.g., Acemoglu and Guerrieri (2008), Buera et. al. (2021), and Cravino and Soleto (2019).
[7]See, e.g., Ngai and Pissarides (2008).
[8]See, e.g., Garcia-Santana et.al. (2021) and Herrendorf, Rogerson, and Valentinyi (2021).
[9]See, e.g., Caselli (2005), and Gollin, Lagakos, and Waugh (2014b).
[10]See, e.g., Atkin, Costinot, and Fukui (2022) and Matsuyama (1992, 2002, 2019).
[11]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, 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).

peak time per capita income. In Section 3, we introduce the technology gap as the only source of heterogeneity across countries and identify the conditions under which the technology gap causes PD through its differential effects on adoption lags. In Section 4, we extend the baseline model by adding the Engel effect on top of the Baumol effect to demonstrate that the main messages of the baseline model is not affected. We also show that the Engel effect only could cause PD, but only under the conditions that would generate counterfactual implications. In Section 5, we extend the model to allow for poorer countries to catch up. We conclude in Section 6.

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

We 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 our baseline model, the hump-shaped path of the manufacturing share is driven 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 to explore how they affect the timing of the manufacturing peak, 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, which is perfectly mobile across sectors, at the wage rate w and κ_j units of the factor specific to sector-j at the rental price, ρ_j, to earn income, $w + \sum_{j=1}^{3} \rho_j \kappa_j$. It spends its income to consume c_j units of good-j, purchased at the price, p_j, subject to the budget constraint,

$$\sum\nolimits_{j=1}^{3} p_j c_j \leq E = w + \sum\nolimits_{j=1}^{3} \rho_j \kappa_j, \tag{1}$$

to maximize its CES utility

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

where E denotes the household expenditure and $\beta_j > 0$ and $0 < \sigma < 1$, so that the three goods are gross complements. Maximizing eq.(2) subject to eq.(1) yields the household's expenditure shares,

$$m_j \equiv \frac{p_j c_j}{E} = \frac{\beta_j \left(p_j\right)^{1-\sigma}}{\sum_{k=1}^{3} \beta_k (p_k)^{1-\sigma}} = \beta_j \left(\frac{p_j}{P}\right)^{1-\sigma} = \beta_j \left(\frac{E/p_j}{U}\right)^{\sigma-1}, \tag{3}$$

where

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

is the cost-of-living index and

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

is the (real) per capita income.

2.2. Production

Sector-j produces good-j, with its specific factor and labor as the inputs, by using the following Cobb-Douglas Technology:

$$Y_j = A_j \left(\kappa_j L\right)^{\alpha} \left(L_j\right)^{1-\alpha}, \tag{4}$$

where $A_j > 0$ is the TFP of sector-j; $\kappa_j L$ is the total supply of specific factor-j; and L_j is the labor employed in sector-j, with $\alpha \in [0,1)$ being the share of specific factor.[12] Labor employed in the three sectors are subject to the labor supply constraint:

$$\sum_{j=1}^{3} L_j = L. \tag{5}$$

From eq.(4), output per worker and output per capita in sector-j can be expressed as:

$$\frac{Y_j}{L_j} = \tilde{A}_j \left(s_j\right)^{-\alpha}; \qquad \frac{Y_j}{L} = \tilde{A}_j \left(s_j\right)^{1-\alpha} \tag{6}$$

where $\tilde{A}_j \equiv A_j \left(\kappa_j\right)^{\alpha}$ and $s_j \equiv L_j / L$ is the employment share of sector-j with

$$\sum_{j=1}^{3} s_j = 1. \tag{7}$$

[12] We introduce the specific factor, $\alpha > 0$, for two reasons. First, it introduces diminishing returns to labor, which makes the prediction of the model robust to opening up for trade. Second, a higher α amplifies the power of the Baumol effect on structural change.

Firms in each sector chooses the inputs to minimize their cost of production, while taking the wage rate w and the rental price ρ_j given. Under the Cobb-Douglas technology, eq.(4), this leads to:

$$wL_j = (1-\alpha)p_j Y_j; \qquad \rho_j \kappa_j L = \alpha p_j Y_j$$

Thus, $1-\alpha$ is the (common) labor share across all sectors. Using the household budget constraint, eq.(1), this implies the aggregate budget constraint, $EL = \sum_{j=1}^{3} p_j Y_j$. Furthermore, using the labor supply constraint, eq.(5), the sectoral shares measured in labor employment are equal to those measured in value-added.

$$s_j \equiv \frac{L_j}{L} = \frac{p_j Y_j}{\sum_{k=1}^{3} p_k Y_k} = \frac{p_j Y_j}{EL}. \tag{8}$$

2.3 Equilibrium

From eq.(6) and eq.(8), we obtain

$$\frac{E}{p_j} = \frac{Y_j}{L_j} = \tilde{A}_j\left(s_j\right)^{-\alpha}. \tag{9}$$

By inserting this expression in eq.(3), the expenditure shares can now be expressed as:

$$\frac{p_j c_j}{E} = \beta_j \left(\frac{\tilde{A}_j\left(s_j\right)^{-\alpha}}{U}\right)^{\sigma-1}.$$

Since the expenditure shares, eq.(3), and the value-added (and employment) shares, eq.(8), are equal to each other in equilibrium,

$$s_j = \frac{p_j Y_j}{EL} = \frac{p_j c_j}{E} = \beta_j \left(\frac{\tilde{A}_j\left(s_j\right)^{-\alpha}}{U}\right)^{\sigma-1}.$$

By solving this equation for s_j and using the adding up constraint, eq.(7),

$$s_j = \frac{\left[\beta_j^{\frac{1}{\sigma-1}}\tilde{A}_j\right]^{-a}}{U^{-a}} = \frac{\left[\beta_j^{\frac{1}{\sigma-1}}\tilde{A}_j\right]^{-a}}{\sum_{k=1}^{3}\left[\beta_k^{\frac{1}{\sigma-1}}\tilde{A}_k\right]^{-a}} \tag{10}$$

$$U = \left\{\sum_{k=1}^{3}\left[\beta_k^{\frac{1}{\sigma-1}}\tilde{A}_k\right]^{-a}\right\}^{-\frac{1}{a}} \tag{11}$$

where

$$a \equiv \frac{1-\sigma}{1-\alpha(1-\sigma)} = -\frac{\partial \ln(s_j/s_k)}{\partial \ln(\tilde{A}_j/\tilde{A}_k)} > 0.$$

Eq.(10) and eq.(11) 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, $\{\tilde{A}_j\}_{j=1}^{3}$, and $a \equiv \frac{1-\sigma}{1-\alpha(1-\sigma)}$ captures how much high productivity in a sector contributes to its relatively low equilibrium share. A higher α, which makes relative demand less responsive to the relative productivity, amplifies this effect by increasing a.[13]

2.4. Productivity Growth Rates, Adoption Lags and Structural Change

Let us now see how the sectoral shares respond to the sectoral productivities change over time.[14] Suppose that $\{\tilde{A}_j(t)\}_{j=1}^{3}$ change according to:

$$\tilde{A}_j(t) = \bar{A}_j(t-\lambda_j) = \bar{A}_j(0)e^{g_j(t-\lambda_j)} = \bar{A}_j(0)e^{-\lambda_j g_j}e^{g_j t}, \tag{12}$$

with $g_j > 0$. and $\lambda_j \geq 0$, while the other parameters stay constant.[15] 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 $\tilde{A}_j(t) = \bar{A}_j(t-\lambda_j)$, 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. The crucial feature of this specification is that the productivity in sector-j is *log-submodular* in λ_j and g_j:

$$\frac{\partial}{\partial g_j}\left(\frac{\partial}{\partial \lambda_j} \ln e^{-\lambda_j g_j}\right) < 0.$$

Thus, a higher g_j magnifies the negative effect of a larger adoption lag λ_j on productivity. In 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.

---

[13] Gross complementarity, $\sigma < 1$, is crucial here. If the three goods were gross substitutes, $\sigma > 1$, $a \equiv -\frac{\sigma-1}{1+\alpha(\sigma-1)} < 0$, so that high 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. Under Cobb-Douglas, $a = 0$ and relative productivity has no effect on the sectoral shares.

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

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

For the moment, "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.(12) into eq.(11) yields the time path of the real per capita income:

$$U(t) = \left\{\sum_{k=1}^{3}\left[\beta_k^{\frac{1}{\sigma-1}}\tilde{A}_k\right]^{-a}\right\}^{-\frac{1}{a}} = \left\{\sum_{k=1}^{3}\tilde{\beta}_k e^{-ag_k(t-\lambda_k)}\right\}^{-\frac{1}{a}} \tag{13}$$

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.(13) 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 first take the ratio of the shares of two sectors, j and $k \neq j$, given in eq.(10) and using eq.(12), to obtain:

$$\frac{s_j(t)}{s_k(t)} = \left(\frac{\tilde{\beta}_j}{\tilde{\beta}_k}\right)e^{a(\lambda_j g_j-\lambda_k g_k)}e^{a(g_k-g_j)t} \Longrightarrow \frac{d}{dt}\ln\left(\frac{s_j(t)}{s_k(t)}\right) = a\left(g_k - g_j\right) \tag{14}$$

Eq.(14) shows that, with the two sectors producing gross complements ($a > 0$ because $\sigma < 1$), $s_j(t)/s_k(t)$ is decreasing over time if $g_j > g_k$, and increasing 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 of the sector 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, using eq.(9) and eq.(14), the relative price can be expressed as:

$$\left(\frac{p_j(t)}{p_k(t)}\right)^{1-\sigma} = \left[\left(\frac{\beta_j}{\beta_k}\right)^{\alpha}\frac{\bar{A}_k(0)}{\bar{A}_j(0)}\right]^{a} e^{a(\lambda_j g_j-\lambda_k g_k)}e^{a(g_k-g_j)t} \Longrightarrow \frac{d}{dt}\ln\left(\frac{p_j(t)}{p_k(t)}\right) = \frac{a\left(g_k - g_j\right)}{1-\sigma} \tag{15}$$

Eq.(15) shows that $p_j(t)/p_k(t)$ is decreasing over time if $g_j > g_k$, and increasing 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 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, for example, by Herrendorf, Rogerson, Valentinyi (2014), based on the mechanism put forward by Ngai and Pissarides (2007).

That is, the share of agriculture, $s_1(t)$, is decreasing over time. This is because

$$\frac{1}{s_1(t)} - 1 = \frac{s_2(t)}{s_1(t)} + \frac{s_3(t)}{s_1(t)} = \left[\frac{\tilde{\beta}_2}{\tilde{\beta}_1} e^{a(\lambda_2 g_2 - \lambda_1 g_1)}\right] e^{a(g_1 - g_2)t} + \left[\frac{\tilde{\beta}_3}{\tilde{\beta}_1} e^{a(\lambda_3 g_3 - \lambda_1 g_1)}\right] e^{a(g_1 - g_3)t},$$

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

$$\frac{1}{s_3(t)} - 1 = \frac{s_1(t)}{s_3(t)} + \frac{s_2(t)}{s_3(t)} = \left[\frac{\tilde{\beta}_1}{\tilde{\beta}_3} e^{a(\lambda_1 g_1 - \lambda_3 g_3)}\right] e^{-a(g_1 - g_3)t} + \left[\frac{\tilde{\beta}_2}{\tilde{\beta}_3} e^{a(\lambda_2 g_2 - \lambda_3 g_3)}\right] e^{-a(g_2 - g_3)t},$$

with both terms on the RHS exponentially decreasing. In contrast,

$$\frac{1}{s_2(t)} - 1 = \left[\frac{\tilde{\beta}_1}{\tilde{\beta}_2} e^{a(\lambda_1 g_1 - \lambda_2 g_2)}\right] e^{-a(g_1 - g_2)t} + \left[\frac{\tilde{\beta}_3}{\tilde{\beta}_2} e^{a(\lambda_3 g_3 - \lambda_2 g_2)}\right] e^{a(g_2 - g_3)t}, \tag{16}$$

which is U-shaped, since the 1st term of RHS is exponentially decreasing and the 2nd term exponentially increasing. Hence $s_2(t)$ is hump-shaped. By differentiating (16) with respect to t,

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

This shows that $s_2(t)$ is hump-shaped due to the two opposing forces. On one hand, $g_1 > g_2$ pushes labor out of agriculture to manufacturing. On the other hand, $g_2 > g_3$ pulls labor out of manufacturing to services. At earlier stages of development when the share of agriculture is high, the first effect dominates the second, but at later stages when the share of agriculture is low, the second effect dominates the first.

2.5. The Manufacturing Peak

We are now ready to characterize the manufacturing peak. By solving $s_2'(\hat{t}) = 0$, we obtain the peak time:

$$\hat{t} = \frac{\lambda_1 g_1 - \lambda_3 g_3}{g_1 - g_3} + \hat{t}_0, \qquad \text{where} \qquad \hat{t}_0 \equiv \frac{1}{a(g_1 - g_3)} \ln\left(\frac{g_1 - g_2}{g_2 - g_3}\right)\left(\frac{\tilde{\beta}_1}{\tilde{\beta}_3}\right).$$

In what follows, we adopt two normalizations to simplify the notation and facilitate the interpretation. The first normalization is to choose the base year such that

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

and hence

$$\hat{t} = \frac{\lambda_1 g_1 - \lambda_3 g_3}{g_1 - g_3}. \tag{17}$$

This normalization means that the calendar time is set such that the manufacturing sector would reach its peak at $\hat{t} = 0$ *in the absence of adoption lags.* Thus, eq.(17) shows how adoption lags affect the peak time. Notice that a larger λ_1, which makes the relative productivity of the agriculture sector lower and hence agriculture relatively more expensive, causes a further delay in the manufacturing peak, while a larger λ_3, which makes the relative productivity of the service sector lower and hence services relatively more expensive, reduces a delay in the peak. Inserting eq.(17) into eq.(16) and eq.(13) yields the peak manufacturing share, $\widehat{s_2} \equiv s_2(\hat{t})$, and the real per capita income at the peak time, $\widehat{U} \equiv U(\hat{t})$ as follows:

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

$$\widehat{U} = \left\{(\tilde{\beta}_1 + \tilde{\beta}_3) e^{-a\left(\frac{\lambda_1 - \lambda_3}{g_1 - g_3}\right) g_1 g_3} + \tilde{\beta}_2 e^{-a\frac{(\lambda_1 - \lambda_2)g_1 g_2 + (\lambda_2 - \lambda_3)g_2 g_3}{g_1 - g_3}}\right\}^{-\frac{1}{a}}.$$

To facilitate interpreting these expressions, we set $U(0) = 1$ for $\lambda_1 = \lambda_2 = \lambda_3 = 0$, or

$$\tilde{\beta}_1 + \tilde{\beta}_2 + \tilde{\beta}_3 = 1.$$

This normalization means that we choose the real per capita income at the peak time *in the absence of the adoption lags* as the *numeraire.* Then,

$$\frac{1}{\widehat{s_2}} - 1 = \left(\frac{1}{\tilde{\beta}_2} - 1\right) e^{\frac{a[(\lambda_1 - \lambda_2)g_1 g_2 + (\lambda_2 - \lambda_3)g_2 g_3 + (\lambda_3 - \lambda_1)g_3 g_1]}{g_1 - g_3}} \tag{18}$$

$$\widehat{U} = \left\{(1 - \tilde{\beta}_2) e^{-a\left(\frac{\lambda_1 - \lambda_3}{g_1 - g_3}\right) g_1 g_3} + \tilde{\beta}_2 e^{-a\frac{(\lambda_1 - \lambda_2)g_1 g_2 + (\lambda_2 - \lambda_3)g_2 g_3}{g_1 - g_3}}\right\}^{-\frac{1}{a}}. \tag{19}$$

Under these normalizations, the peak time share of sector-j in the absence of the adoption lags would be $\tilde{\beta}_j$. Thus, eq.(18) and eq.(19) show how adoption lags affect the peak manufacturing share and the peak time per capita income.

So far we have looked at the impacts of adoption lags in a single country in isolation without specifying the sources of the adoption lags. In the next section, we introduce cross-country heterogeneity, the technology gap, which generates cross-country variations in adoption lags, and study the cross-country variations.

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. All the countries share the same values of $\tilde{\beta}_j$, g_j, and θ_j for $j = 1,2,3$. Thus, the countries differ in one dimension, λ. 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 for the possibility that the extent to which technology gap affects the adoption lag varies across sectors. That is, θ_j are *sector-specific* parameters, common across countries, capturing the inherent difficulty of adoption in the three sectors. They control how much the technology gap affects the adoption lag and hence productivity in each sector. From eq.(12) and using $(\lambda_1, \lambda_2, \lambda_3) = (\theta_1, \theta_2, \theta_3)\lambda$,

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

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. The crucial feature of this specification is that the adoption lag λ_j is *supermodular* in θ_j and g_j:

$$\frac{\partial}{\partial \theta_j}\left(\frac{\partial \lambda_j}{\partial \lambda}\right) > 0,$$

so that θ_j magnifies the impact of the technology gap on the adoption lag.

By inserting $(\lambda_1, \lambda_2, \lambda_3) = (\theta_1, \theta_2, \theta_3)\lambda$ in eqs. (17)-(19), we obtain the expressions for the peak time, $\hat{t}$, the peak manufacturing share, $\widehat{s_2}$, and the peak time per capita income, $\widehat{U}$, as functions of λ as follows:

$$\hat{t}(\lambda) = \frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3}\lambda. \tag{20}$$

$$\frac{1}{\widehat{s_2}(\lambda)} - 1 = \left(\frac{1}{\tilde{\beta}_2} - 1\right) e^{(g_2 - g_3)\left(\frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3} - \frac{\theta_2 g_2 - \theta_3 g_3}{g_2 - g_3}\right) a\lambda} \tag{21}$$

$$\widehat{U}(\lambda) = \left\{(1 - \tilde{\beta}_2) e^{-g_1 g_3\left(\frac{\theta_1 - \theta_3}{g_1 - g_3}\right) a\lambda} + \tilde{\beta}_2 e^{-g_2\left(\frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3} - \theta_2\right) a\lambda}\right\}^{-\frac{1}{a}} \tag{22}$$

Recall also that, from eq.(13),

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

which is always decreasing in λ. Hence a higher-λ country has a lower per capita income at any point in time.

3.2. The Three Conditions for Premature Deindustrialization

We are now ready to obtain the three conditions for premature deindustrialization. That is, a poorer, higher-λ (hence 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, $\widehat{U}$.

First, from eq.(20), $\hat{t}'(\lambda) > 0$ for all λ, if and only if

$$\theta_1 g_1 > \theta_3 g_3.$$

This condition 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.

Second, from eq.(21), $\widehat{s_2}'(\lambda) < 0$ for all λ, if and only if

$$(g_2 - g_3)\left(\frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3} - \frac{\theta_2 g_2 - \theta_3 g_3}{g_2 - g_3}\right) = \left(\theta_3 - \frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3}\right) g_3 + \left(\frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3} - \theta_2\right) g_2 > 0$$

which holds below the positive-sloped 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. (22) is not generally monotone in λ. However, when the two conditions above 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

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

which holds on the left side of the vertical line $\theta_1/\theta_3 = 1$ in Figure 1. Intuitively, even if $\theta_1 g_1 > \theta_3 g_3$ causes a delay in the peak time, the longer adoption lag in services, $\theta_1 < \theta_3$ means that the delay is not long enough to make up for a high λ, so that $\hat{t}(\lambda) < \theta_1\lambda < \theta_3\lambda$, which makes the peak time agriculture and services productivity of the late industrializers 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.[16] Furthermore, $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$ if $(1-\Theta)(1-\theta_2/\theta_3) < 1 - \theta_1/\theta_3$, where

$$\Theta \equiv 1 - \frac{\tilde{\beta}_2(1-g_3/g_1)}{g_3/g_2 + \tilde{\beta}_2(1-g_3/g_2)} = \frac{\left(1-\tilde{\beta}_2\right) g_3/g_2 + \tilde{\beta}_2\, g_3/g_1}{\left(1-\tilde{\beta}_2\right) g_3/g_2 + \tilde{\beta}_2} < 1,$$

which is decreasing in $\tilde{\beta}_2$ and satisfies $g_3/g_1 < \Theta < 1$ with $\Theta \to 1$ as $\tilde{\beta}_2 \to 0$ and $\Theta \to g_3/g_1$ as $\tilde{\beta}_2 \to 1$. In Figure 1, this condition holds on the left side of the dashed line connecting $(\theta_1/\theta_3\,, \theta_2/\theta_3) = (\Theta, 1)$ and $(\theta_1/\theta_3\,, \theta_2/\theta_3) = (1,1)$. Thus, one *sufficient* condition for $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$ is $(1-g_3/g_1)(1-\theta_2/\theta_3) < 1-\theta_1/\theta_3$, which is equivalent to $\hat{t}(\lambda) < \theta_2\lambda$, which 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, which is why their peak-time per capita income is lower regardless of the sectoral composition of demand.

Notice that these three conditions for premature deindustrialization, depicted in Figure 1, jointly imply $\max\{\theta_1, \theta_2\} < \theta_3$. Thus, the adoption lag is the longest in the service sector in every country with $\lambda > 0$, since the technology gap affects the adoption most negatively, which seems plausible given the intangible nature of the service technology. These three conditions also 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 relative magnitude of $\theta_2 g_2$ and $\theta_3 g_3$, hence the cross-country productivity dispersion in manufacturing may or may not be larger than that in services.

We now illustrate these conditions for premature deindustrialization with some examples.

3.3. Premature Deindustrialization: Some Numerical Illustrations

[16]To see this, note that, even though $\hat{t}(\lambda) > \theta_2\lambda$ implies the two exponential terms in eq.(22) go in the opposite directions as λ goes up, the first term eventually dominates the second term as λ becomes sufficiently large.

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.(20)-(22),

$$\hat{t}(\lambda) = \theta\lambda;\ \ \widehat{s_2}(\lambda) = \tilde{\beta}_2;\ \widehat{U}(\lambda) = 1.$$

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 the same path with early industrializers. Only the timing would be different. Thus, premature deindustrialization does not occur unless the technology gap to have differential impacts across sectors.

Examples 2a-2c. 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 premature deindustrialization region, and hence $\hat{t}'(\lambda) > 0$, $\widehat{s_2}'(\lambda) < 0$, and $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$.

For all the numerical illustrations in this paper, we use the following parameter values: $\alpha = 1/3$, $\sigma = 0.6$ (hence $a = 6/13$), $g_1 = 3.6\% > g_2 = 2.4\% > g_3 = 1.2\%$ for the productivity growth rates.[17] We also choose $\widehat{s_2}(0) = \tilde{\beta}_2 = 1/3$. For these parameter values, the two normalizations, $\hat{t}(0) = 0$ and $\widehat{U}(0) = 1$, are achieved by $\tilde{\beta}_1 = \tilde{\beta}_3 = (1 - \tilde{\beta}_2)/2 = 1/3$.

In Example 2a, $g_3/g_1 = 1/3 < \theta_1/\theta_3 = \theta_2/\theta_3 = 0.5 = g_3/g_2$, 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. Figure 2a illustrates the path of the manufacturing 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.(13) and eq.(16). 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,

[17] 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 Rustuccia (2010), $g_1 = 3.8\% > g_2 = 2.4\% > g_3 = 1.3\%$, but have the ratios of 3/2/1. The results do not change much even if we use the estimates from Comin, Lashkari, and Mestieri (2021).

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 $\widehat{U}(\lambda)$.

Examples 2b and 2c look at the impact of moving along the diagonal. For Example 2b, plotted in Figure 2b, $g_3/g_1 = 1/3 < \theta_1/\theta_3 = \theta_2/\theta_3 = 0.35 < 0.5 = g_3/g_2$, so that $\theta_1 g_1 > \theta_3 g_3 > \theta_2 g_2$, implying that cross-country productivity dispersion is the smallest in manufacturing. For Example 2c, plotted in Figure 2c, $g_3/g_1 = 1/3 < 0.5 = g_3/g_2 < \theta_1/\theta_3 = \theta_2/\theta_3 = 0.75$, so that $\theta_1 g_1 > \theta_2 g_2 > \theta_3 g_3$, implying that cross-country productivity dispersion is the smallest in services. These figures show that, as a decline in $\theta_1/\theta_3 = \theta_2/\theta_3$ 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.

4. 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. More recently, Comin, Lashkari and Mestieri (2021) derived a decomposition of the Baumol effect 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 premature deindustrialization obtained by the baseline model. Furthermore, if we had relied solely on the

Engel effect without the Baumol effect, premature deindustrialization 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 the reason 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.[18] In this paper, we use *isoelastic nonhomothetic CES*, following Comin-Lashkari-Mestieri (2021) and Matsuyama (2019), because it offers a natural extension to the Baumol-Ngai-Pissarides CES framework. More specifically, the utility function of each household, $U = U(c_1, c_2, c_3)$, is given implicitly by

$$\left[\sum_{j=1}^{3} (\beta_j)^{\frac{1}{\sigma}} \left(\frac{c_j}{U^{\varepsilon_j}}\right)^{1-\frac{1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \equiv 1 \tag{23}$$

where $\varepsilon_j > 0$. If $\varepsilon_1 = \varepsilon_2 = \varepsilon_3$, eq.(23) is reduced to the standard homothetic CES, eq.(2). If $\{\varepsilon_j\}_{j=1}^{3}$ vary with j, the relative weights attached to the three goods in eq.(23) 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.(23) subject to the budget constraint, eq.(1) yields the expenditure shares:

$$m_j \equiv \frac{p_j c_j}{E} = \frac{\beta_j \left(U^{\varepsilon_j} p_j\right)^{1-\sigma}}{\sum_{k=1}^{3} \beta_k (U^{\varepsilon_k} p_k)^{1-\sigma}} = \beta_j \left(\frac{U^{\varepsilon_j} p_j}{E}\right)^{1-\sigma}. \tag{24}$$

Here, U is the maximized value of the utility satisfying the indirect utility function, also defined implicitly,

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

[18] 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.

where P is the cost-of-living index and hence $U = E/P$ can be interpreted as the real income per capita.[19]

From eq.(24), it is straightforward to verify that the income elasticity of good-j is given by,

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

Thus, 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[20]. Furthermore, with the constant relative prices, the expenditure share of agriculture $m_1(t)$ is decreasing in $U(t)$, since

$$\frac{1}{m_1(t)} - 1 = \frac{m_2(t)}{m_1(t)} + \frac{m_3(t)}{m_1(t)} = \frac{\beta_2}{\beta_1}\left(\frac{p_2}{p_1}U(t)^{\varepsilon_2-\varepsilon_1}\right)^{1-\sigma} + \frac{\beta_3}{\beta_1}\left(\frac{p_3}{p_1}U(t)^{\varepsilon_3-\varepsilon_1}\right)^{1-\sigma},$$

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

$$\frac{1}{m_3(t)} - 1 = \frac{m_1(t)}{m_3(t)} + \frac{m_2(t)}{m_3(t)} = \frac{\beta_1}{\beta_3}\left(\frac{p_1}{p_3}U(t)^{\varepsilon_1-\varepsilon_3}\right)^{1-\sigma} + \frac{\beta_2}{\beta_3}\left(\frac{p_2}{p_3}U(t)^{\varepsilon_2-\varepsilon_3}\right)^{1-\sigma},$$

with both terms on the RHS exponentially decreasing in $U(t)$. In contrast,

$$\frac{1}{m_2(t)} - 1 = \frac{m_1(t)}{m_2(t)} + \frac{m_3(t)}{m_2(t)} = \frac{\beta_1}{\beta_2}\left(\frac{p_1}{p_2}U(t)^{\varepsilon_1-\varepsilon_2}\right)^{1-\sigma} + \frac{\beta_3}{\beta_2}\left(\frac{p_3}{p_2}U(t)^{\varepsilon_3-\varepsilon_2}\right)^{1-\sigma} \tag{25}$$

with the 1st term of RHS exponentially decreasing in $U(t)$ and the 2nd term exponentially increasing. Hence $m_2(t)$ is hump-shaped in $U(t)$. By differentiating (25) with respect to $U(t)$,

$$\frac{ds_2(t)}{dU(t)} \gtreqless 0 \Leftrightarrow (\varepsilon_2 - \varepsilon_1)\frac{m_1(t)}{m_2(t)} \gtreqless (\varepsilon_3 - \varepsilon_2)\frac{m_3(t)}{m_2(t)} \Leftrightarrow \varepsilon_2 \gtreqless \sum\nolimits_{k=1}^{3} \varepsilon_k m_k(t) \Leftrightarrow \eta_2 \gtreqless 1.$$

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

---

[19]Since the relative weights on the three goods vary continuously with U in eq.(23), 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, we also consider 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. See Appendix A.

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

manufacturing to services. At earlier stages of development when the share of agriculture is high, the first effect dominates the second, but at later stages when the share of agriculture is low, the second effect dominates the first.

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

We now add the above Engel's mechanism of structural change on top of the production structure assumed in the baseline model. Then, by following the same step as before, one could show that the equilibrium shares, $s_j(t) = m_j(t)$ are given by:

$$s_j(t) = \frac{\left[\beta_j^{\frac{1}{\sigma-1}} \tilde{A}_j(t)\right]^{-a}}{[U(t)^{\varepsilon_j}]^{-a}}, \qquad \text{where } \sum_{k=1}^{3} \frac{\left[\beta_k^{\frac{1}{\sigma-1}} \tilde{A}_k(t)\right]^{-a}}{[U(t)^{\varepsilon_k}]^{-a}} \equiv 1,$$

where $\tilde{A}_j(t) = \bar{A}_j(t-\lambda_j) = \bar{A}_j(0)e^{g_j(t-\theta_j\lambda)}$. From this, $s_2(t)$ can be written as:

$$\frac{1}{s_2(t)} - 1 = U(t)^{a(\varepsilon_1-\varepsilon_2)} \left[\frac{\tilde{\beta}_1}{\tilde{\beta}_2} e^{a(\theta_1 g_1 - \theta_2 g_2)\lambda}\right] e^{-a(g_1-g_2)t} + U(t)^{a(\varepsilon_3-\varepsilon_2)} \left[\frac{\tilde{\beta}_3}{\tilde{\beta}_2} e^{a(\theta_3 g_3 - \theta_2 g_2)\lambda}\right] e^{a(g_2-g_3)t} \tag{26}$$

where $U(t)$ satisfies

$$U(t)^{a\varepsilon_1}\tilde{\beta}_1 e^{-ag_1(t-\theta_1\lambda)} + U(t)^{a\varepsilon_2}\tilde{\beta}_2 e^{-ag_2(t-\theta_2\lambda)} + U(t)^{a\varepsilon_3}\tilde{\beta}_3 e^{-ag_3(t-\theta_3\lambda)} \equiv 1 \tag{27}$$

By differentiating eq.(26) subject to eq.(27), the equation for $s_2'(t) = 0$ can be rewritten as:

$$(g_1-g_2) - (g_2-g_3)U^{a(\varepsilon_3-\varepsilon_2)}\left[\frac{\tilde{\beta}_3}{\tilde{\beta}_1}\right] e^{a(\theta_3 g_3-\theta_1 g_1)\lambda} e^{a(g_1-g_3)t}$$
$$= \frac{\left\{(\varepsilon_1-\varepsilon_2) + (\varepsilon_3-\varepsilon_2)U^{a(\varepsilon_3-\varepsilon_1)}\left[\frac{\tilde{\beta}_3}{\tilde{\beta}_1}\right] e^{a(\theta_3 g_3-\theta_1 g_1)\lambda} e^{a(g_1-g_3)t}\right\}\left\{g_1 U^{a(\varepsilon_1-\varepsilon_2)}\tilde{\beta}_1 e^{-ag_1(t-\theta_1\lambda)} + g_2\tilde{\beta}_2 e^{-ag_2(t-\theta_2\lambda)} + g_3 U^{a(\varepsilon_3-\varepsilon_2)}\tilde{\beta}_3 e^{-ag_3(t-\theta_3\lambda)}\right\}}{\varepsilon_1 U^{a(\varepsilon_1-\varepsilon_2)}\tilde{\beta}_1 e^{-ag_1(t-\theta_1\lambda)} + \varepsilon_2\tilde{\beta}_2 e^{-ag_2(t-\theta_2\lambda)} + \varepsilon_3 U^{a(\varepsilon_3-\varepsilon_2)}\tilde{\beta}_3 e^{-ag_3(t-\theta_3\lambda)}} \tag{28}$$

In principle, we need to solve eqs.(27)-(28) simultaneously for the peak time $\hat{t}$ and the peak time per capita income $\widehat{U} = U(\hat{t})$. Then, by inserting the solutions $t = \hat{t}$ and $U = \widehat{U}$ into eq. (26), we obtain the peak manufacturing share $\widehat{s_2}$.

4.3. Analytically Solvable "Unbiased" Cases

The peak time values, $\hat{t}$, $\widehat{s_2}$, and $\widehat{U}$, cannot be expressed analytically in the presence of both the Baumol and Engel effects, except when the following equality holds.

$$\frac{\varepsilon_2-\varepsilon_1}{\varepsilon_3-\varepsilon_2} = \frac{g_1-g_2}{g_2-g_3}.$$

Recall that, under the Baumol effect only, the time path of the manufacturing share is hump-shaped path due to the two opposing forces whose relative strength is given by the RHS of the above equality. Likewise, under the Engel effect with the constant relative prices, the manufacturing share becomes hump-shaped in U is due to the two opposing forces whose relative strength is given by the LHS of the above equality. Thus, by adding the Engel effect on top of the Baumol effect and increasing its relative magnitude by increasing $\varepsilon_2 - \varepsilon_1$ and $\varepsilon_3 - \varepsilon_2$ in such a way that the LHS does not change, we do not change the relative strength of the two opposing forces that create the hump-shape. We call this analytically solvable case as "unbiased."

More specifically, we set

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

where $\bar{g} \equiv (g_1 + g_2 + g_3)/3$ is the (simple) average growth rate, and study the effect of increasing μ, for given $g_1 > g_2 > g_3 > 0$. (The upper bound of μ is needed to ensure $\varepsilon_1 > 0$.) Then, some algebra yields the peak time, the peak manufacturing share and the peak time per capita income can be analytically expressed as:

$$\hat{t}(\lambda) = \frac{\theta_1 g_1 - \theta_3 g_3}{g_1 - g_3}\lambda - \ln\left\{(1-\tilde{\beta}_2)e^{-g_1 g_3\left(\frac{\theta_1-\theta_3}{g_1-g_3}\right)a\lambda} + \tilde{\beta}_2 e^{-g_2\left(\frac{\theta_1 g_1-\theta_3 g_3}{g_1-g_3}-\theta_2\right)a\lambda}\right\}^{-\frac{1}{a}\left(\frac{\mu}{1+\mu\bar{g}}\right)} \tag{29}$$

$$\frac{1}{\widehat{s_2}(\lambda)} - 1 = \left(\frac{1}{\tilde{\beta}_2} - 1\right)e^{(g_2-g_3)\left(\frac{\theta_1 g_1-\theta_3 g_3}{g_1-g_3}-\frac{\theta_2 g_2-\theta_3 g_3}{g_2-g_3}\right)a\lambda} \tag{30}$$

$$\widehat{U}(\lambda) = \left\{(1-\tilde{\beta}_2)e^{-g_1 g_3\left(\frac{\theta_1-\theta_3}{g_1-g_3}\right)a\lambda} + \tilde{\beta}_2 e^{-g_2\left(\frac{\theta_1 g_1-\theta_3 g_3}{g_1-g_3}-\theta_2\right)a\lambda}\right\}^{-\frac{1}{a}\left(\frac{1}{1+\mu\bar{g}}\right)} \tag{31}$$

By comparing eq.(29)-(31) with eq.(20)-(22), one could easily verify $\widehat{s_2}'(\lambda) < 0$ and $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$ under the same condition; and $\hat{t}'(\lambda) > 0$ for all $\lambda > 0$ under a weaker condition**.** Thus, at least for the (analytically solvable) unbiased case, introducing the Engel effect and adding more nonhomotheticity do not change the main messages of the baseline model. Furthermore, it is easy to see that

- $\hat{t}(0) = 0, \widehat{s_2}(0) = \tilde{\beta}_2, \widehat{U}(0) = 1$; a higher μ has no effect on the country with $\lambda = 0$;
- a higher μ causes a further delay in $\hat{t}(\lambda)$ for every country with $\lambda > 0$, from eq.(29);
- a higher μ has no effect on $\widehat{s_2}(\lambda)$ for every country with $\lambda > 0$, from eq.(30);
- A higher μ makes a decline in $\widehat{U}(\lambda)$ smaller for every country with $\lambda > 0$, from eq.(31).

Figure 3a illustrates these results. We use the same parameter values as in Figure 2a. In particular, $g_1 = 3.6\% > g_2 = 2.4\% > g_3 = 1.2\%$, so that $g_1 - g_2 = g_2 - g_3 = 1.2\%$. 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 relative to the Baumol effect cause a longer delay in the peak time, a higher $\hat{t}(\lambda)$ and hence the peak time occurs less prematurely a higher $\widehat{U}(\lambda)$, because 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.

4.4. Empirically More Plausible "Biased" Case

Obviously, the unbiased case is a knife-edge case. 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 "biased" cases of $(\varepsilon_2 - \varepsilon_1)/(\varepsilon_3 - \varepsilon_2) \neq (g_1 - g_2)/(g_2 - g_3)$, we can solve for the peak values only numerically. Instead of reporting all possible cases, however, we show the case of $(\varepsilon_2 - \varepsilon_1)/(\varepsilon_3 - \varepsilon_2) = 4 > 1 = (g_1 - g_2)/(g_2 - g_3)$, using the values close to the estimates by Comin, Lashkari, and Mestieri (2021). Figure 3b reports the numerical solution in this case for the peak values, 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 biased cases, we plot the peak time values relative to the country with $\lambda = 0$. 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 unbiased case. As for $\widehat{s_2}$, adding more nonhomotheticity makes a decline in $\widehat{s_2}$ larger for every country with $\lambda > 0$. [21]

These results suggest that 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 premature deindustrialization. However, it should be noted that nonhomotheticity has significant effects on the path of structural change. To see this, we plot in Figure 3c the paths of the manufacturing share against time and against log-per capita income $\ln U(t)$ for three different cases: the homothetic case in Figure 2a, the unbiased case in Figure 3a with $\epsilon = 0.6$ and the biased case of Figure 3b with $\epsilon = 0.6$. Adding the Engel effect makes the hump-shaped

---

[21] We tried some biased cases in the opposite direction, $(\varepsilon_2 - \varepsilon_1)/(\varepsilon_3 - \varepsilon_2) < (g_1 - g_2)/(g_2 - g_3)$. 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 $\lambda > 0$.

noticeably much sharper, which indicates nonhomotheticity can spend up the pace of structural change. In contrast, we do not see any noticeable changes in the peak time values.[22]

4.5. Premature Deindustrialization (PD) through the Engel (Income) 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

$$\left(\frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_3 - \varepsilon_2}\right)\frac{\tilde{\beta}_1}{\tilde{\beta}_3} = 1;\;\; \tilde{\beta}_1 + \tilde{\beta}_2 + \tilde{\beta}_3 = 1$$

which ensures $\hat{t}(0) = 0$ and $\widehat{U}(0) = 1$, we obtain

$$\hat{t}(\lambda) = \frac{1}{a\bar{g}}\ln\left\{(1-\tilde{\beta}_2)e^{\frac{(\varepsilon_3\theta_1 - \varepsilon_1\theta_3)}{(\varepsilon_3 - \varepsilon_1)}a\bar{g}\lambda} + \tilde{\beta}_2 e^{\left(\theta_2 + \frac{(\theta_1 - \theta_3)}{(\varepsilon_3 - \varepsilon_1)}\varepsilon_2\right)a\bar{g}\lambda}\right\} \tag{32}$$

$$\frac{1}{\widehat{s_2}(\lambda)} - 1 = \left(\frac{1}{\tilde{\beta}_2} - 1\right)e^{(\varepsilon_3 - \varepsilon_2)\left(\frac{\theta_1 - \theta_3}{\varepsilon_3 - \varepsilon_1} - \frac{\theta_2 - \theta_3}{\varepsilon_3 - \varepsilon_2}\right)a\bar{g}\lambda} \tag{33}$$

$$\ln \widehat{U}(\lambda) = \frac{\theta_1 - \theta_3}{\varepsilon_3 - \varepsilon_1}\bar{g}\lambda \tag{34}$$

From these equations, we can obtain the three conditions for premature deindustrialization. First, from eq.(34), $\widehat{U}'(\lambda) < 0$ for all $\lambda > 0$ if and only if

$$\theta_1 < \theta_3.$$

This condition holds on the left side of the vertical line, $\theta_1/\theta_3 = 1$ in Figure 4. Intuitively, when θ_1 is relatively smaller than θ_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. Second, from eq.(33), $\widehat{s_2}'(\lambda) < 0$ for all $\lambda > 0$ if and only if

$$(\varepsilon_3 - \varepsilon_2)\left(\frac{\theta_1 - \theta_3}{\varepsilon_3 - \varepsilon_1} - \frac{\theta_2 - \theta_3}{\varepsilon_3 - \varepsilon_2}\right) = \frac{(\varepsilon_3 - \varepsilon_2)\theta_1 + (\varepsilon_2 - \varepsilon_1)\theta_3}{(\varepsilon_3 - \varepsilon_1)} - \theta_2 > 0.$$

In Figure 4, this condition holds below the line connecting the two points, $(\theta_1/\theta_3, \theta_2/\theta_3) = (\varepsilon_1/\varepsilon_3, \varepsilon_2/\varepsilon_3)$ and $(\theta_1/\theta_3, \theta_2/\theta_3) = (1,1)$. Intuitively, with a sufficiently low θ_2, which has no effect on $\widehat{U}(\lambda)$, the price of manufacturing is low relative to both agriculture and services in a higher-λ country, which keeps the manufacturing share low. Finally, $\hat{t}(\lambda)$ in eq.(32) is not

[22]See also Figure A, in which we use, instead of U, an alternative measure of economic development, the share of non-agriculture, $s_n(t) = s_2(t) + s_3(t) = 1 - s_1(t)$, which do not affect the results qualitatively.

generally monotone. However, when the above two conditions are met, one could show that 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 $\theta_1/\theta_3 > \varepsilon_1/\varepsilon_3$. This condition holds on the right side of the vertical line, $\theta_1/\theta_3 = \varepsilon_1/\varepsilon_3$ in Figure 4. Furthermore, $\hat{t}'(\lambda) > 0$ for all $\lambda > 0$ if and only if $(\Theta_E - \varepsilon_1/\varepsilon_3)[1 - (\varepsilon_3/\varepsilon_2)(\theta_2/\theta_3)] < \theta_1/\theta_3 - \varepsilon_1/\varepsilon_3 < 1 - \varepsilon_1/\varepsilon_3$, where

$$\Theta_E \equiv 1 - \frac{(1-\tilde{\beta}_2)(1-\varepsilon_1/\varepsilon_3)}{1-\tilde{\beta}_2(1-\varepsilon_2/\varepsilon_3)} = \frac{\tilde{\beta}_2\,\varepsilon_2/\varepsilon_3 + (1-\tilde{\beta}_2)\,\varepsilon_1/\varepsilon_3}{\tilde{\beta}_2\,\varepsilon_2/\varepsilon_3 + (1-\tilde{\beta}_2)},$$

which is increasing in $\tilde{\beta}_2$ and satisfies $\varepsilon_1/\varepsilon_3 < \Theta_E < 1$ with $\Theta_E \to \varepsilon_1/\varepsilon_3$ as $\tilde{\beta}_2 \to 0$ and $\Theta_E \to 1$ as $\tilde{\beta}_2 \to 1$. In Figure 4, this condition holds above the dashed line connecting $(\theta_1/\theta_3, \theta_2/\theta_3) = (\varepsilon_1/\varepsilon_3, \varepsilon_2/\varepsilon_3)$ and $(\theta_1/\theta_3, \theta_2/\theta_3) = (\Theta_E, 0)$.

Thus, even with the Engel effect only, the heterogeneity of the technology gap could cause premature deindustrialization. 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. Introducing Catching Up

Until now, we have assumed that the country-specific parameter, the technology gap λ is time-invariant. This implies that the sectoral productivity growth rate is constant over time and identical across countries.[23] We have made this assumption to focus on the "level effect" of the technology gap, by shutting down the potential "growth effect."

In this section, we allow for the possibility that latecomers may 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,

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

[23] 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.

which preserves the ranking of countries 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:

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

Again, we can obtain the peak time, $\hat{t}(\lambda_0)$, now a function of the initial value of λ_t, by solving $s_2{}'(t) = 0$. By setting the calendar time such that $\hat{t}(0) = 0$, we obtain the peak time, $\hat{t}(\lambda_0)$,

$$\hat{t}(\lambda_0) = \frac{\theta_1 g_1-\theta_3 g_3}{g_1 - g_3}\lambda_{\hat{t}} + D(g_\lambda\lambda_{\hat{t}}),$$

where

$$D(g_\lambda\lambda_{\hat{t}}) \equiv \frac{1}{a(g_1 - g_3)}\ln\left[\left(\frac{g_1 - g_2 + (\theta_1 g_1-\theta_2 g_2)g_\lambda\lambda_{\hat{t}}}{g_2 - g_3 - (\theta_3 g_3-\theta_2 g_2)g_\lambda\lambda_{\hat{t}}}\right)\left(\frac{g_2 - g_3}{g_1 - g_2}\right)\right].$$

Hence the peak manufacturing share, $\hat{s}_2(\lambda_0)$ and the peak time per capita income $\widehat{U}(\lambda_0)$ are:

$$\frac{1}{\hat{s}_2(\lambda_0)} - 1 = \left(\frac{1}{\tilde{\beta}_2} - 1\right)\left[\frac{(g_2 - g_3)e^{a(g_2-g_1)D(g_\lambda\lambda_{\hat{t}})} + (g_1 - g_2)e^{a(g_2-g_3)D(g_\lambda\lambda_{\hat{t}})}}{g_1 - g_3}\right]\left[e^{\frac{a(g_1-g_2)(g_2-g_3)}{(g_1-g_3)}}\right]^{\left(\frac{\theta_1 g_1-\theta_2 g_2}{g_1-g_2}+\frac{\theta_3 g_3-\theta_2 g_2}{g_2-g_3}\right)\lambda_{\hat{t}}}$$

$$\widehat{U}(\lambda_0) = \left\{\left(\tilde{\beta}_1 e^{-ag_1 D(g_\lambda\lambda_{\hat{t}})} + \tilde{\beta}_3 e^{-ag_3 D(g_\lambda\lambda_{\hat{t}})}\right)e^{-a\frac{(\theta_1-\theta_3)g_1 g_3}{g_1-g_3}\lambda_{\hat{t}}} + \left(\tilde{\beta}_2 e^{-ag_2 D(g_\lambda\lambda_{\hat{t}})}\right)e^{-a\frac{(\theta_1-\theta_2)g_1 g_2+(\theta_2-\theta_3)g_2 g_3}{g_1-g_3}\lambda_{\hat{t}}}\right\}^{-\frac{1}{a}}$$

For $g_\lambda = 0, D(g_\lambda\lambda_{\hat{t}}) = D(0) = 0$, and $\lambda_{\hat{t}} = \lambda_0$ so that these expressions become identical with the baseline model. For $g_\lambda > 0$, these expressions need to be solved numerically. Figure 5 shows the result for the same parameter values with Example 2a. 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*.[24] 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 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.

6. Concluding Remarks

[24]One route through which catching up could take place is global technology diffusion. See Acemoglu (2008, Ch.18) and Comin and Mestieri (2018), just to name a few. The estimates by Comin and Mestieri suggest $g_\lambda < 1.0\%$.

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). The countries are heterogenous only in one dimension, in their "technology gap," the country's capacity to adopt the frontier technologies, as Krugman (1985), but the technology gap affects the adoption lags in the three sectors differently, unlike Krugman (1985). In this setup, we identified the conditions 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. We found that the heterogeneity in the technology gap generates PD when the following three conditions are met; i) the impact of the technology gap on the adoption lag is larger in service than in agriculture, ii) although the adoption lag is shorter in agriculture than in services, the productivity growth rate is sufficiently higher in agriculture than in services that the cross-country productivity dispersion is larger in agriculture than in services; and iii) the impact of the technology gap on the adoption lag is not too large in manufacturing. It turns out that these three conditions for PD jointly imply 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.

In the baseline model, the sectoral demand composition is generated by homothetic CES (to focus on the Baumol effect) 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 two extensions. In the first extension, 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. In the second 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 have 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. Thus, in the absence of any distortions, the equilibrium allocation is efficient.[25] This also makes the prediction of our analysis 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. 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). 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.[26] We hope that the analysis presented in this paper will provide a useful building block for future research.

References

Acemoglu, D., Introduction to Modern Economic, Princeton, Princeton University Press, 2009.

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

Atkin, D., A. Costinot, and Fukui, 2022, "Globalization and the ladder of development: Pushed to the top or held at the bottom," MIT.

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

Boppart, T., 2014, "Structural Change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences," Econometrica

[25]In this model, 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.

[26]This is indeed what Rodrik (2016) hypothesized as a possible cause for PD. In Fujiwara and Matsuyama (2021), we 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).

Buera, F.J., and J.P. Kaboski, 2012a, "The Rise of the Service Economy," American Economic Review.
Buera, F.J., and J.P. Kaboski, 2012b, "Scale and the Origins of Structural Change," Journal of Economic Theory.
Buera, F.J., J.P. Kaboski, R. Rogerson, and J.I. Vizcaino, 2021, "Skill-Biased Structural Change," Review of Economic Studies,
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, D., and M. Mestieri, 2018, "If Technology has arrived everywhere, why has income diverged?" American Economic Journal: Macroeconomics
Comin, D., D. Lashkari, and M. Mestieri, 2021, "Structural Transformation with Long-run Income and Price Effects," Econometrica, 89, 1, 311-374.
Costinot, A., "An Elementary Theory of Comparative Advantage," Econometrica, 77 (2009), 1165-1192.
Costinot, A., and J. Vogel, 2010, "Matching and Inequality in the World Economy," Journal of Political Economy, 118, 747-786.
Costinot, A., and J. Vogel, "Beyond Ricardo: Assignment Models in International Trade," Annual Review of Economics, 7 (2015), 31-62.
Cravino, J., and S. Sotelo, 2019, "Trade-Induced Structural Change and the Skill Premium," American Economic Journal: Macroeconomics,
Duarte, M., and D. Restuccia, 2010, "The Role of the Structural Transformation in Aggregate Productivity," Quarterly Journal of Economics,
Foellmi, R., and J. Zweimüller, 2006, "Income Distribution and Demand-Induced Innovations," Review of Economic Studies.
Foellmi, R., and J. Zweimüller, 2014, "The Macroeconomics of Model T," Journal of Economic Theory.
Fujiwara, I., and K. Matsuyama, 2021, "A Global View of Premature Deindustrialization," Work in Progress, Keio and Northwestern Universities.
Garcia-Santana, M., J. Pijoan-Mas, and L. Villacorta, 2021, "Investment Demand and Structural Change," Econometrica, 89(6), 2751-2785.
Gollin, D., D. Lagakos, and M.E. Waugh, 2014a, "Agricultural Productivity Differences across Countries," American Economic Review, Papers & Proceedings, 104(5): 165-170.
Gollin, D., D. Lagakos, and M.E. Waugh, 2014b, "Agricultural Productivity Gap," Quarterly Journal of Economics, 939-993.
Herrendorf, B., R. Rogerson, and A. Valentinyi, 2014, "Growth and Structural Transformation," in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf, 855–941.
Herrendorf, B., R. Rogerson, and A. Valentinyi, 2021, "Structural Change in Investment and Consumption: A Unified Analysis, The Review of Economic Studies, 88(3), 1311-1346.
Huneeus, F., and R. Rogerson, 2020, "Heterogeneous Paths of Industrialization," Princeton University.
Kongsamut, P., R. Rebelo, and D. Xie, "Beyond Balanced Growth," Review of Economic Studies, 68 (2001): 869-882.
Kravis, I., Heston, A. and Summers, R., 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, P., “Scale Economies, Product Differentiation, and the Pattern of Trade,” American Economic Review, (December 1980), 950-959.
Krugman, P. 1985, A ‘Technology-Gap’ Model of International Trade. In *Structural Adjustment in Developed Open Economies*, ed. K. Jungenfelt and D. Hague. London, Macmillan.
Lewis, L., R. Monarch, M. Sposi, and J. Zhang, 2022, “Structural Change and Global Trade,” Journal of the European Economic Association, 20(1): 476-512.
Matsuyama, K., "Agricultural Productivity, Comparative Advantage, and Economic Growth," Journal of Economic Theory 58 (December 1992): 317-334.
Matsuyama, K., "A Ricardian Model with a Continuum of Goods under Nonhomothetic Preferences: Demand Complementarities, Income Distribution and North-South Trade," Journal of Political Economy 108 (December 2000): 1093-1120.
Matsuyama, K., "The Rise of Mass Consumption Societies," Journal of Political Economy, 110 (October 2002): 1035-1070.
Matsuyama, K., 2005, “Credit Market Imperfections and Patterns of International Trade and Capital Flows,” Journal of European Economic Association, 3(2-3), 714-723.
Matsuyama, K., "Structural Change in an Interdependent World: A Global View of Manufacturing Decline," Journal of the European Economic Association. 7 (April-May 2009): 478-486.
Matsuyama, K., 2019, “Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade” Econometrica, 497-528.
Murphy, K., A. Shleifer, and R. Vishny, “Income Distribution, Market Size, and Industrialization,” Quarterly Journal of Economics, 104, 1989, 537-564.
Ngai, L. R., and C. A. Pissarides, 2007, “Structural Change in a Multisector Model of Growth,” American Economic Review, 97, 429–443.
Ngai, L. R., and C. A. Pissarides, 2008, “Trends in Hours and Economic Growth,” Review of Economic Dynamics, 11(2): 239-256.
Nordhaus, W.D., 2008, “Baumol’s diseases: A macroeconomic analysis,” The BE Journal of Macroeconomics.
Restuccia, D., D.T. Yang, and X. Zhu, “Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis,” Journal of Monetary Economics, 55(2), 234-250.
Rodrik, D., 2013, “Unconditional Convergence in Manufacturing,” The Quarterly Journal of Economics, 128(1), 165-204.
Rodrik, D., 2016, “Premature Deindustrialization,” Journal of Economic Growth, 21, 1-33.
Sposi, M., K.M. Yi, and J. Zhang, 2021, “Deindustrialization and Industry Polarization,” Chicago Fed.
Uy, T., K-M. Yi, and J. Zhang, “Structural Change in an Open Economy,” Journal of Monetary Economics, 60 (2013), 667-682.

Figure 1: Conditions for PD only with the Baumol (Relative Price) Effect

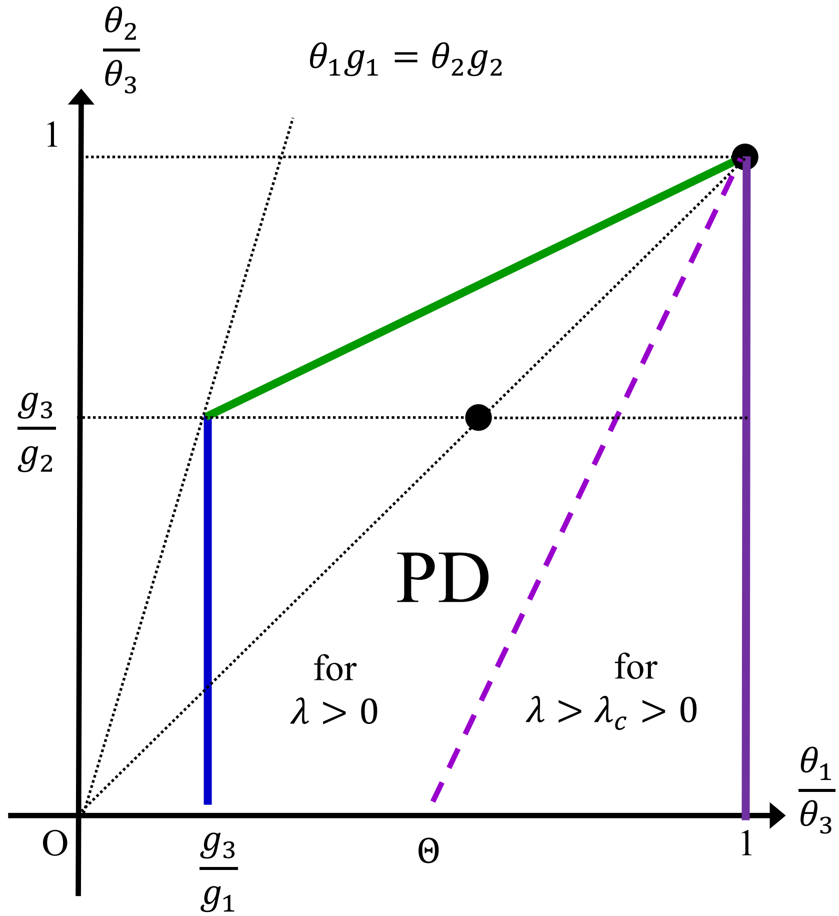


Figure 2: Premature Deindustrialization under the Baumol Effect: Numerical Illustrations

Example 2a: $\frac{\theta_1}{\theta_3} = \frac{\theta_2}{\theta_3} = 0.5 = \frac{g_3}{g_2}$ Example 2b: $\frac{\theta_1}{\theta_3} = \frac{\theta_2}{\theta_3} = 0.35 < \frac{g_3}{g_2}$ Example 2c: $\frac{\theta_1}{\theta_3} = \frac{\theta_2}{\theta_3} = 0.75 > \frac{g_3}{g_2}$

$\left(t, s_2(t)\right)$ $\left(\ln U(t), s_2(t)\right)$

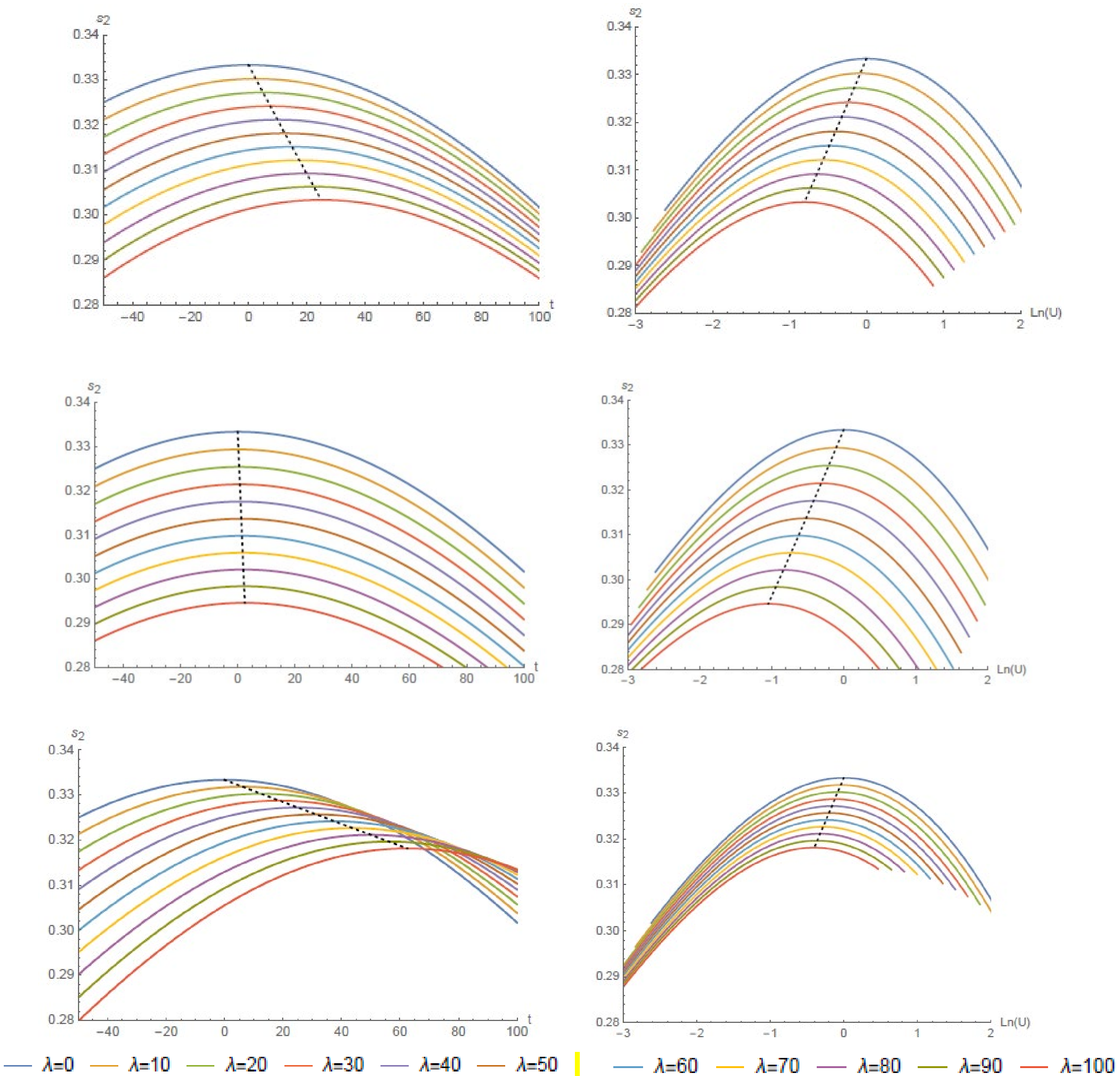
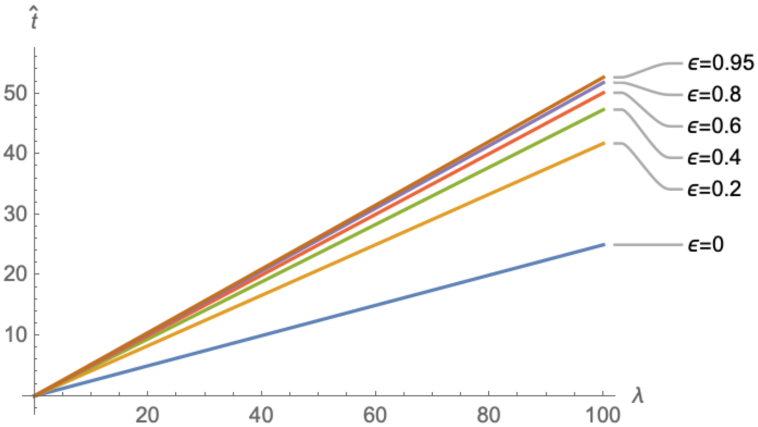
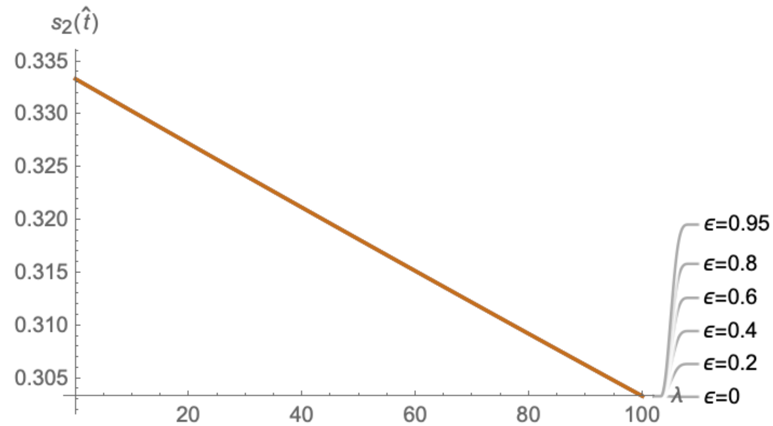


Figure 3a: (Analytically Solvable) "Unbiased" Case: $g_1 - g_2 = g_2 - g_3 = 1.2\% > 0 \Rightarrow \varepsilon_1 = 1 - \epsilon < \varepsilon_2 = 1 < \varepsilon_3 = 1 + \epsilon$ for $0 < \epsilon = (1.2\%)\mu < 1$

Peak Time



Peak Time
Manufacturing Share



Peak Time
Per Capita Income

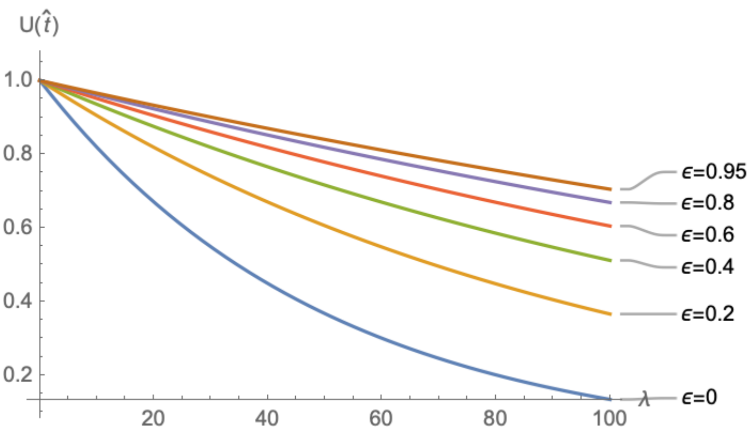
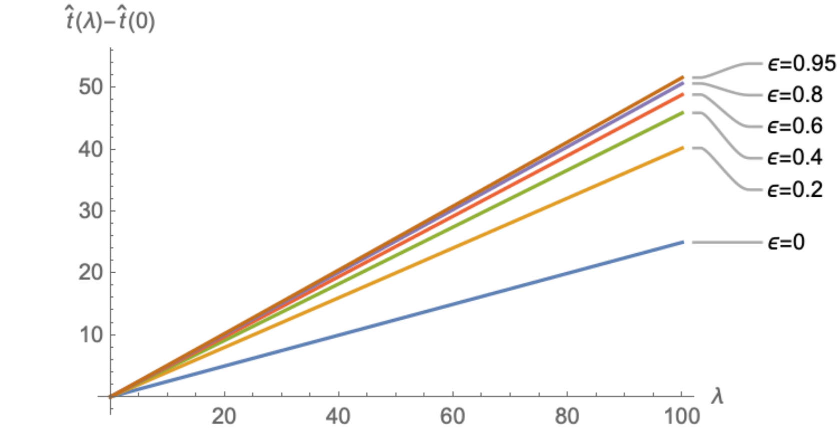


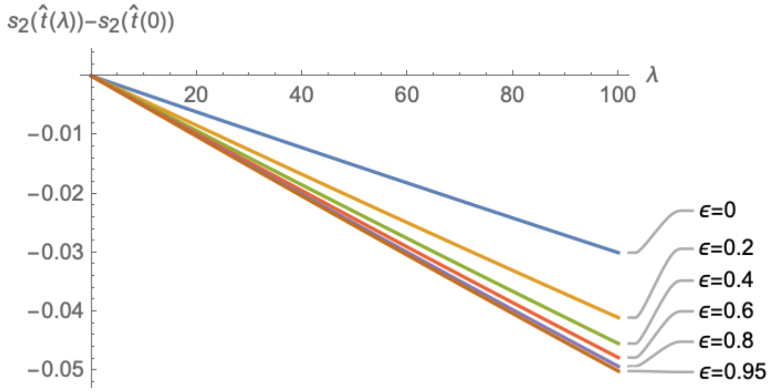
Figure 3b: (Empirically More Plausible) Biased Case:

$\varepsilon_1 = 1 - \epsilon < \ \varepsilon_2 = 1 + \frac{\epsilon}{3} < \varepsilon_3 = 1 + \frac{2\epsilon}{3}$ for $0 < \epsilon < 1 \Longrightarrow \frac{g_1 - g_2}{g_2 - g_3} = 1 < \frac{\varepsilon_2 - \varepsilon_1}{\varepsilon_3 - \varepsilon_2} = 4.$

Peak Time



Peak Time
Manufacturing
Share



Peak Time
Per Capita Income

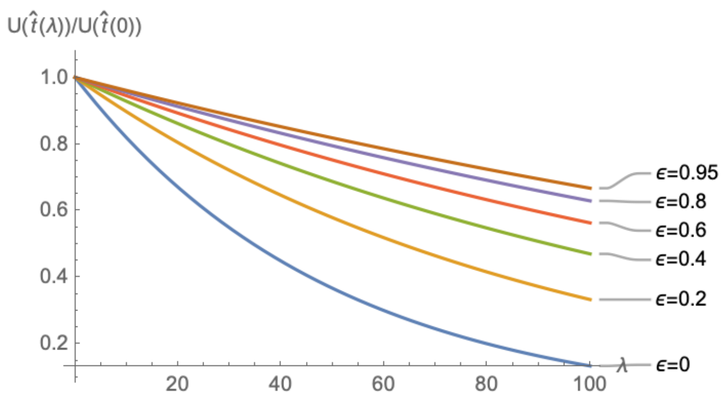


Figure 3c: Stronger nonhomotheticity significantly changes the shape of the time paths, but has little effects on λ affects $\hat{t}, \widehat{s_2}$, and $\widehat{U}$.

$\left(t, s_2(t)\right)$ $\quad$ $\left(\ln U(t), s_2(t)\right)$

$\epsilon = 0$
Homothetic case
$(\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1)$

$\epsilon = 0.6$
Unbiased case
$(\varepsilon_1 = .4 < \varepsilon_2 = 1 < \varepsilon_3 = 1.6)$

$\epsilon = 0.6$
Biased case
$(\varepsilon_1 = .4 < \varepsilon_2 = 1.2 < \varepsilon_3 = 1.4)$

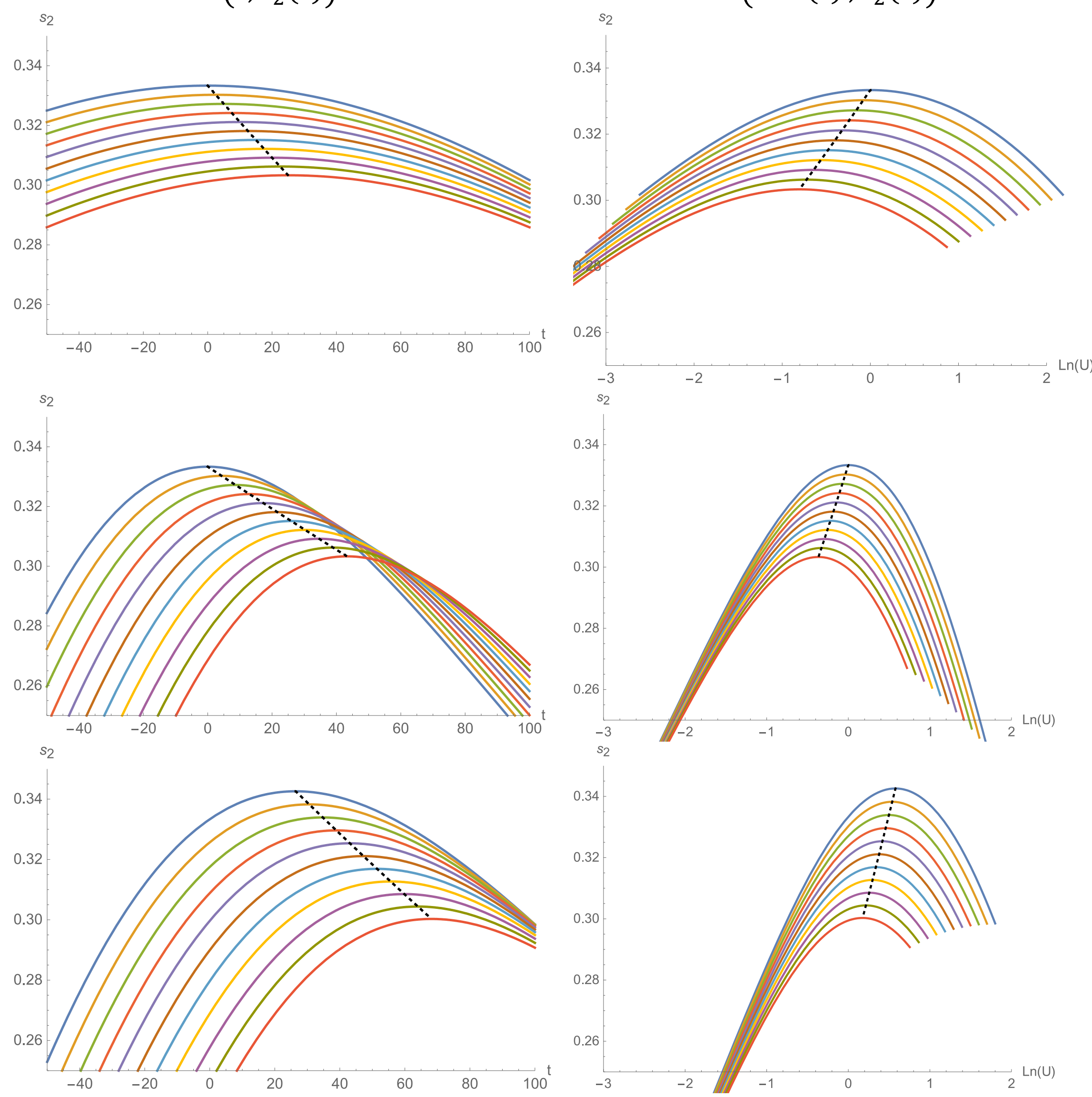


Figure 4: Conditions for Premature Deindustrialization (PD) only with the Engel Effect

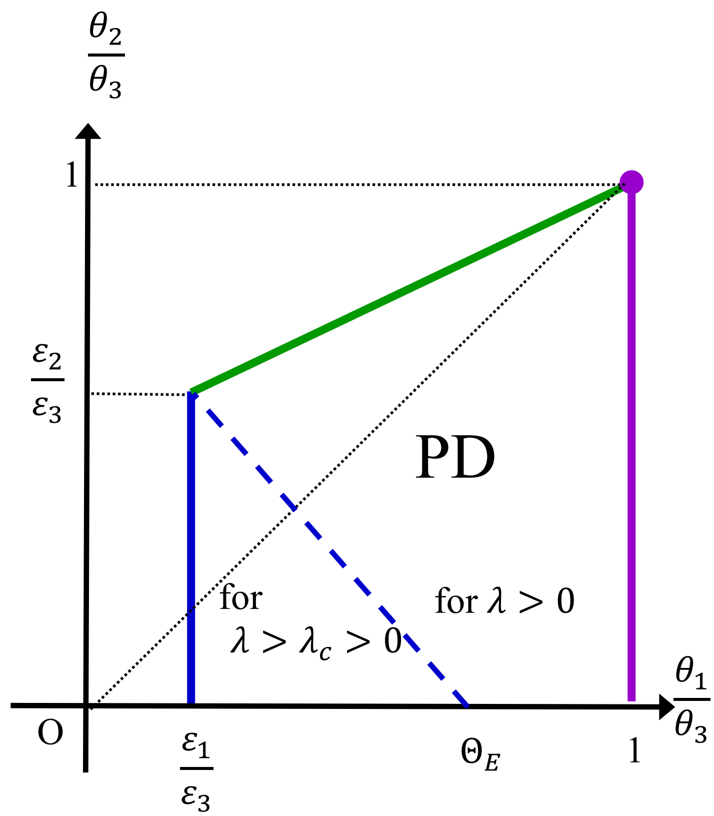
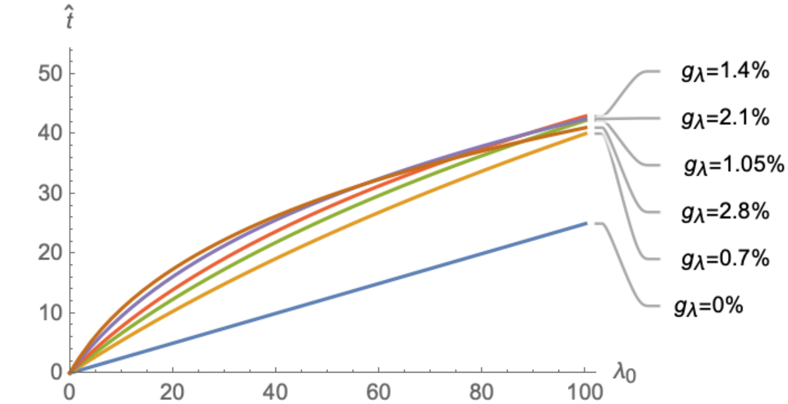
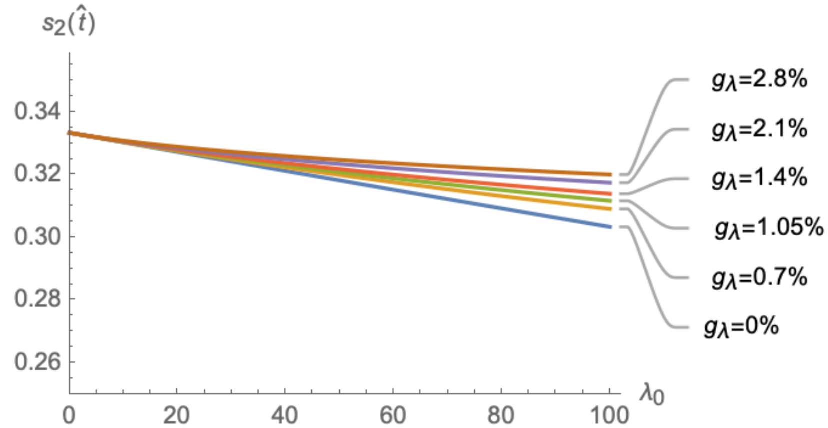


Figure 5: Catching Up

Peak Time

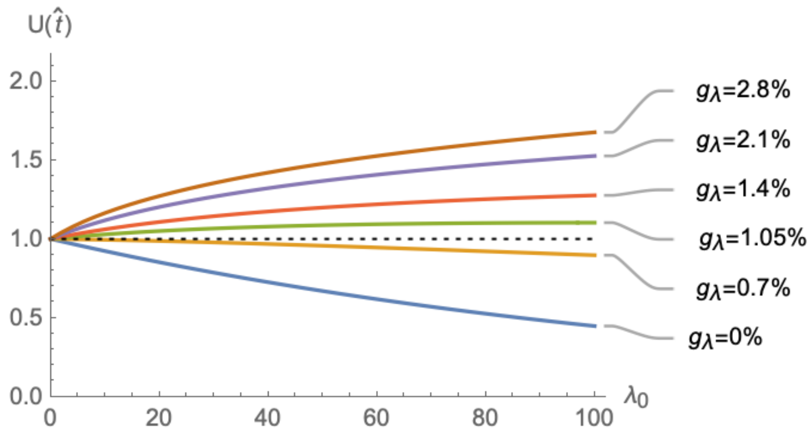


Peak Time
Manufacturing Share



Peak Time

Per Capita Income

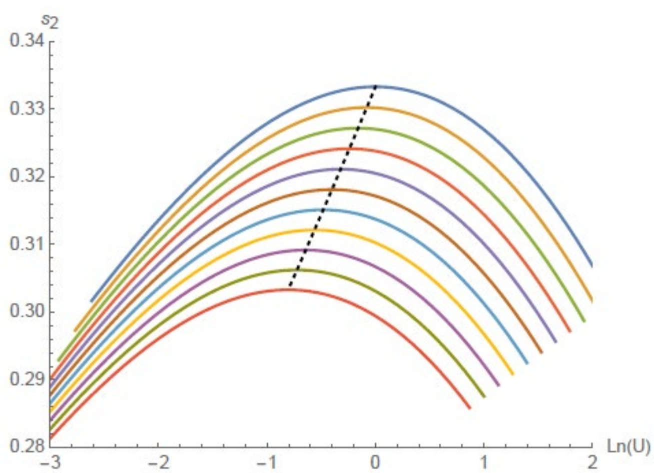


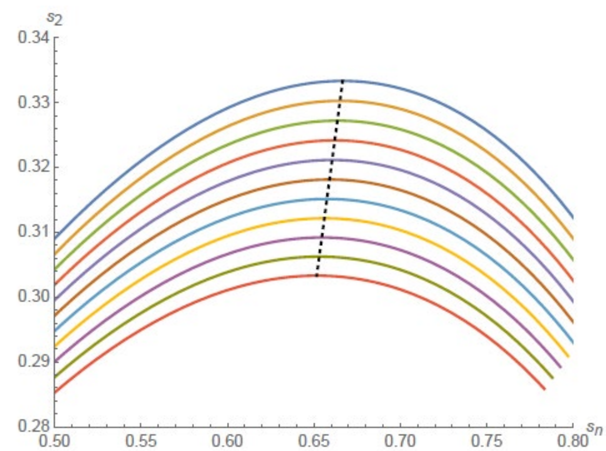
Appendix A

Figure A: Non-agricultural share $1 - s_1(t) = s_2(t) + s_3(t) \equiv s_n(t)$ as a measure of development

Baseline Homothetic Case:

$\left(\ln U(t), s_2(t)\right)$ $\left(s_n(t), s_2(t)\right)$



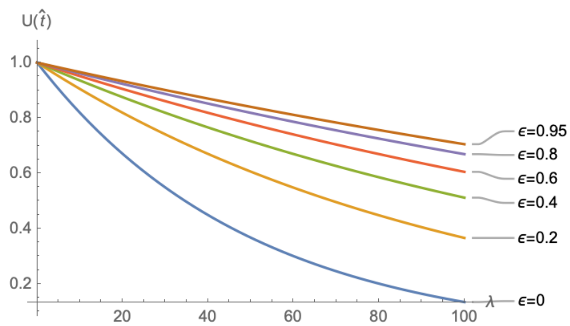


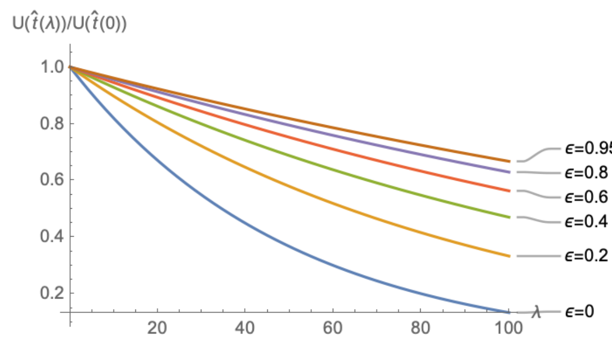
Nonhomothetic Cases:

Unbiased: $\varepsilon_1 = 1 - \epsilon < \varepsilon_2 = 1 < \varepsilon_3 = 1 + \epsilon$

Biased: $\varepsilon_1 = 1 - \epsilon < \varepsilon_2 = 1 + \epsilon/3 < \varepsilon_3 = 1 + 2\epsilon/3$

$\ln U(\hat{t})$





$s_n(\hat{t})$