Productivity Growth and Structural Transformation

Roberto M. Samaniego* Juliana Y. Sun†

November 12, 2013

Abstract

Economies tend to diversify and then re-specialize as they develop. An economy with many industries with different productivity growth rates may display these "stages of diversification" as a result of productivity-driven structural change if initially resources are concentrated in industries other than those that dominate economic structure in the long run. A calibrated multi-industry growth model with many countries replicates the main features of the "stages of diversification". We also present evidence that countries systematically shift resources towards manufacturing industries with rapid productivity growth, and towards sectors with low productivity growth, consistent with the model.

Keywords: Structural transformation, stages of diversification, productivity differences, economic development.

JEL Codes: O11 O14 O33 O41.

*Roberto M. Samaniego, George Washington University, 2115 G St NW Suite 340, Washington, DC 20052. Tel: (202) 994-6150. Fax: (202) 994-6147. Email: roberto@gwu.edu.

†Juliana Y Sun, School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903. Tel: (65) 6808 5456, Fax: (65) 6828 0833, Email: yusun@smu.edu.sg.

1 Introduction

Economic development is a joint process of economic growth and economic restructuring.[1] It is well known that economic development tends to involve a shift of resources away from agriculture and towards services.[2] In addition, Imbs and Wacziarg (2003, hereafter IW) show that there exist "stages of diversification": along the development path, countries begin with employment concentrated in a few industries, diversifying until they reach a certain threshold in income per capita beyond which they re-specialize. The finding that industrial specialization is U-shaped along the development path is more general than just the rise of services and the decline of agriculture: IW document the "stages" across 9 broad sectors and also across 28 manufacturing industries.[3] Although this specialization pattern is arguably the broadest stylized fact of structural transformation, the literature on structural transformation and growth has not accounted for the "stages of diversification."

This paper aims to fill this gap in the literature. We show that persistent productivity differences across industries can account for the observed patterns of economic restructuring along the development path. We study a multi-sector model economy that highlights productivity growth differences across sectors, and also across manufacturing industries. We show that the pattern of diversification followed by specialization can be accounted for simply by the dynamics of industry structure resulting from these differences, and we present empirical evidence that supports our hypothesis, underlining the importance of productivity growth as a factor of structural transformation.

Consider the following intuition. Suppose that markets are competitive, and that there are two goods that are *substitutes* in consumption. Then, persistent differences in productivity growth rates lead to an increase in the GDP share of the industry with the most rapid

[1] The World Bank (2012) defines economic development as:

> Qualitative change and restructuring in a country's economy in connection with technological and social progress. The main indicator of economic development is increasing GNP per capita (or GDP per capita), reflecting an increase in the economic productivity and average material wellbeing of a country's population. Economic development is closely linked with economic growth.

[2] See Gollin, Parente and Rogerson (2007), Restuccia, Yang and Zhu (2008) and Duarte and Restuccia (2010) for recent work documenting and analyzing the sources and consequences of these empirical regularities.

[3] Koren and Tenreyro (2007) report a similar finding across 19 manufacturing industries and across 19 sectors.

productivity growth, as the good it produces registers a decline in its relative price. However, if the economy starts out being specialized in the *other* industry, then the economy diversifies until half of resources are devoted to each industry, after which it begins to specialize again. The economy exhibits a "U" shaped pattern of specialization along the development path. Conversely, suppose the goods are *complements*. Then, persistent differences in productivity growth rates lead to an increase in the GDP share of the industry with the *slowest* productivity growth, also generating a "U" shape if the economy starts out specialized in the other industry.[4]

First, we provide crucial evidence that, as they grow, countries *indeed systematically shift resources among industries with different total factor productivity (TFP) growth rates*. The literature suggests that goods are *substitutes* within manufacturing, whereas goods are *complements* across broad sectors: we would then expect that resources shift towards *high-TFP growth industries* within manufacturing, whereas resources shift towards *low-TFP growth sectors*. This is exactly what we find in the data, providing strong evidence that structural change along the development path is indeed related to productivity growth differences, as we hypothesize.[5]

Then, we develop a multi-industry growth model in which productivity growth rates differ across industries. We calibrate initial productivity levels so as to reproduce the industry and sectoral composition of each of the 51 countries in the IW dataset in 1963, and allow the structure of the model economies to evolve over time based on persistent productivity growth differences across industries. Along the development path, the labor shares of different industries evolve due to disparities between their TFP growth rates. Applying the same non-parametric method as IW to the model-generated series of industrial specialization, the calibrated model generates the U-shaped specialization pattern found in IW. Our results hold both within manufacturing and across broad sectors, and are robust to a number of variations in the calibration procedure. We conclude that differences in TFP growth across industries indeed contribute to the "stages of diversification," and that an important characteristic of the process of economic development is the reallocation of resources among industries with

[4]Understanding the initial conditions themselves is beyond the scope of the paper. However, we do not impose that initial conditions are skewed one way or another: initial conditions are drawn from the data. We also show that our ability to match the U shape is not driven simply by the fact that we match initial conditions

[5]Koren and Tenreyro (2007) find that along the development path economies shift resources towards low-volatility industries. We find that our industry productivity growth measures are not significantly correlated with their industry volatility measures, so our findings offer evidence for a distinct feature of structural transformation.

different rates of productivity growth.

Industry productivity growth rates are calibrated using US data. The use of common industry TFP growth rates across countries provides a clean experiment, as empirical country-specific rates could be influenced by industrial structure.[6] At the same time, this assumption is consistent with the finding in Rodrik (2012) that there is unconditional convergence in labor productivity across countries among disaggregated manufacturing industries. We perform several robustness checks, using different productivity measures and allowing for productivity convergence dynamics, finding that the results are robust to variation in productivity growth rates across countries, as well as to different approaches to productivity measurement.

Ngai and Pissarides (2007) and Acemoglu and Guerrieri (2008) develop growth models where industry TFP differences generate structural change in the manner of our paper. We build on their theoretical insights to provide quantitative evidence that productivity change is not only an important factor of structural change across sectors but also within manufacturing, and moreover this mechanism can account for the stages of diversification. The literature has suggested several mechanisms underlying structural transformation, and we provide evidence supporting the importance of the productivity growth mechanism using industry and sector data from a broad set of countries, complementing the results of Herrendorf, Herrington and Valentinyi (2013), who show that sectoral productivity growth differences are the main factor behind structural transformation in the US among broad sectors.

We also make methodolgical contribution to the computation of growth models along an unbalanced growth path. We follow Rogerson (2008), Duarte and Restuccia (2010) and others in computing transition dynamics in multi-sector growth models without capital.[7] This is for simplicity but also without loss of generality if capital shares are similar across industries. We also show that, if we allow capital shares to vary across industries, then not only do our results remain robust but also capital share differences on their own are unable

[6] Industry size is viewed by some as a potential determinant of productivity-enhancing R&D and hence possibly of productivity growth, although the evidence is mixed – see the survey of Cohen (2011). We wish to isolate the impact of productivity growth on economic structure from any feedback that might occur.

[7] Duarte and Restuccia (2010) also assume that preferences over the agricultural good are different above and below a subsistence threshold, in line with other papers that focus on agriculture such as Gollin, Parente and Rogerson (2007) and Restuccia, Yang and Zhu (2008). We abstract from such non-homotheticity of preferences to focus on the productivity mechanism: also, our industries and sectors are much more highly disaggregated, so that patterns of structural change among them are unlikely to depend on a subsistence requirement in one particular sector or industry.

to account for the stages of diversification. However, in a technical appendix, we develop a growth model with capital, and show that our results are indeed similar in that context. We do not include it in the main text because it turns out that there are important technical complications involved in computing such a model. In particular, the conditions shown in Ngai and Pissarides (2007) to be required for a balanced growth path do not hold empirically (the elasticity of substitution among capital goods is not one). Our results in the model with capital are important not only for robustness, but also because our simulation procedure for computing a model economy with many sectors in transition is of independent interest, since it can be used in many contexts where the model does not admit a balanced growth path (for example, when production functions are not of Cobb-Douglas form as in Krusell et al (2000) or He and Liu (2008)).

Our work also complements the literature on the link between trade and development by establishing a benchmark characterizing the evolution of economic structure in a closed economy. The literature tends to interpret the "stages of diversification" as being related to trade.[8] However, before concluding that international trade is an important determinant of the evolution of economic structure along the development path, we need to understand how economic structure evolves *under autarky*. If a closed economy can generate the stages of diversification using well-understood and empirically supported mechanisms, Ockham's principle of parsimony indicates that a key role for trade in generating the "stages" should not be assumed *a priori*. Indeed, we show that even in a closed economy persistent total factor productivity (TFP) growth differences across industries are sufficient to generate a U-shaped pattern of specialization that is very close to the pattern in the data.

Recent work by Imbs, Montenegro and Wacziarg (2012) relates the stages of diversification to, first, integration of markets within countries, followed by integration of markets across countries. In fact, our account is complementary to theirs. Comparing their theory with data, they find that their predicted patterns of integration and specialization are observed more clearly among tradeables than among non-tradeables. In contrast, our model accounts well for the "stages" both within manufacturing (often identified with tradeables) and also across broad sectors (including many non-tradeable services). Indeed our sector-level results are the strongest, suggesting that growth-theoretic concerns might be most important for non-tradeables, whereas trade-theoretic mechanisms should have some traction within manufacturing.

The rest of the paper is organized as follows. Section 2 provides evidence concerning

[8]See IW, and also Kalemli-Ozcan, Sorensen and Yosha (2003), Rodrik (2008) and Imbs et al (2012).

the evolution of economic structure along the development path, including evidence linking structural transformation to shifts in resources among industries with different rates of productivity growth. Section 3 describes the model, and Section 4 delivers the main results. Section 5 provides evidence supporting the predictions of the model regarding shifts in resources among industries and sectors with different TFP growth rates. Section 6 concludes with a discussion of possibilities for future work.

2 Diversification and TFP Growth differences

We begin by describing the stylized facts of how industrial structure evolves along the development path. First, we summarize the methodology and results of IW. Then, we provide new evidence that, along the development path, countries do indeed systematically shift resources among industries and sectors based on industry rates of productivity growth. This supports the theory of productivity-driven structural change in general, not just as it relates to the stages of diversification, so these results are important and of general interest.

2.1 Economic structure along the development path

IW use a nonparametric methodology to investigate the relationship between sectorial diversification and income. Manufacturing industry data are drawn from the INDSTAT3 database distributed by UNIDO, whereas sector-level data are provided by the ILO, and data on aggregate income per capita are from the Penn World Tables. The industry (or sector) share is defined as the share of manufacturing employment.

IW use the Gini coefficient of industry shares $GINI_{c,t}$ to measure the degree of industrial concentration in any country c at date t: the more equal the industry shares (i.e. the lower the Gini), the more diversified the economic structure.[9] Then, they apply a procedure related to robust locally weighted scatterplot smoothing (lowess) to uncover the link between income per capita $GDP_{c,t}$ and specialization. Specifically, they regress the Gini coefficients of industrial specialization on income per capita with country fixed effects, using rolling income windows.

$$GINI_{c,t} = \hat{\alpha}_c(x) + \hat{\beta}(x) GDP_{c,t} + \varepsilon_{c,t}, \; GDP_{c,t} \in [x - \Delta/2, x + \Delta/2]. \tag{1}$$

[9] Although IW focus on Gini coefficients, they use several other measures of specialization for robustness. We focus on Gini coefficients too, but results for other measures are available upon request.

The income interval size Δ is fixed at $5,000 (in 1985 dollars) and the midpoint x of the interval gradually moves away from the previous income range (the increment across regressions is $25). Then, they plot the fitted Gini coefficients estimated at the midpoint of the income interval in each regression. They find a U-shaped relationship between Gini coefficients and income levels. Their U-shaped relationship is robust across sectors that account for the entire private economy of the countries concerned[10] (ILO data) and within manufacturing (UNIDO data). Figure 1 reproduces their main results.

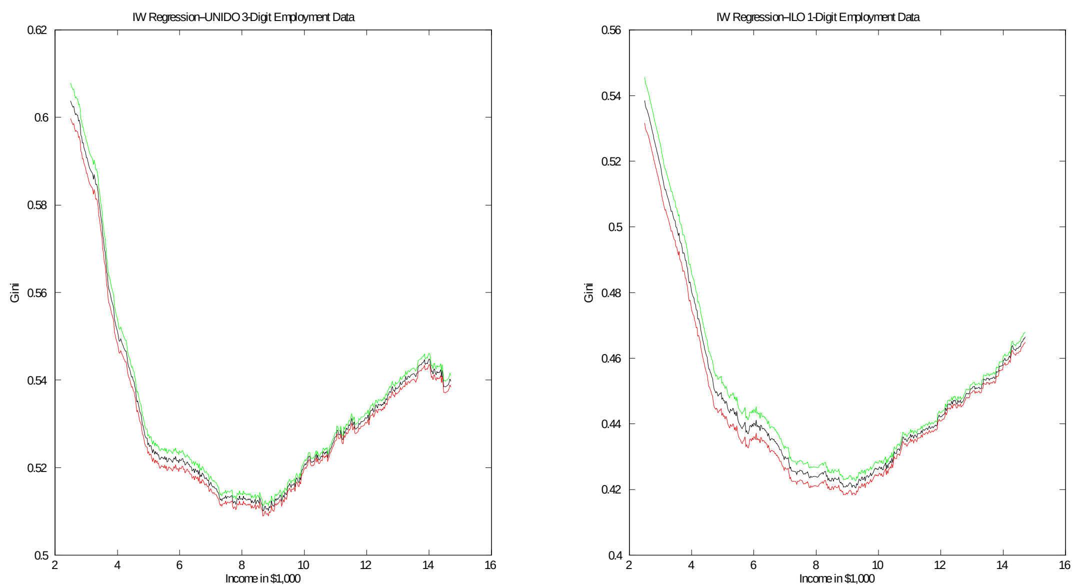

Figure 1. IW results. Each point in the Figure is the fitted value $\hat{\alpha}_c(x) + \hat{\beta}(x)x$ from equation (1), where x is GDP per capita. Confidence bands represent two standard errors of the coefficient $\hat{\beta}(x)$. The left panel is industry concentration within manufacturing estimated using INDSTAT3 data provided by UNIDO. The right panel is sectorial concentration across the entire economy estimated using ILO data. The middle line is the point estimate, whereas the other lines reflect the point estimate plus or minus the standard error of $\hat{\beta}(x)$, as in IW.

[10]The 9 sectors are Agriculture, Hunting, Forestry and Fishing; Mining and Quarrying; Mining and Quarrying; Electricity, Gas and Water; Construction; Wholesale and Retail Trade and Restaurants and Hotels; Transport, Storage and Communication; Financing, Insurance, Real Estate and Business Services; and Community, Social and Personal Services.

2.2 Productivity and structural change

This paper will argue that these patterns can be related to systematic shifts of resources between industries that experience different rates of productivity growth. Ngai and Pissarides (2007) and Acemoglu and Guerrieri (2008) show that, if goods are substitutes, then in equilibrium resources should shift towards industries with rapid productivity growth. If goods are complements, then in equilibrium resources should shift towards industries with low productivity growth. This productivity mechanism of structural change would then be capable of generating the stages of diversification, depending on initial conditions.

To see this, suppose that there are two industries, and that "specialization" is measured using the Gini coefficient. Let g_i be the productivity growth factor of industry $i \in \{1, 2\}$, and let ε be the elasticity of substitution between these goods. If s_{jt} is the share of industry j, then the Gini coefficient equals $0.5 - \min\{s_{1t}, 1 - s_{1t}\}$.[11] Now suppose that $g_1 < g_2$. Then, if $\varepsilon > 1$, for a sufficiently low initial share of industry 1 the economy will start off specialized in industry 1 whereas in all periods thereafter the share of 2 will increase and that of 1 will decrease. Thus, the minimum of the two industry shares (s_{2t}) will grow until its share reaches 0.5 and the Gini coefficient has dropped to 0. After this, the minimum of the two becomes s_{1t} and, as its share continues to decrease, the Gini coefficient rises again. Thus, for a time, specialization decreases, until s_{1t} drops below half – after which specialization will begin increasing again. Alternatively, if $\varepsilon < 1$, for sufficiently low initial productivity in industry 2 the economy will start off specialized in industry 2, whereas in all periods thereafter the share of 1 will increase, and the same dynamics obtain.

Along the development path, do we indeed observe resources moving between industries systematically based on their productivity growth rates? To see this, we require two further pieces of information. First, we require estimates of the elasticities of substitution across goods – to be precise, we require knowledge of whether these elasticities are greater or less than unity. Second, we require measures of productivity growth for the industries and sectors considered in IW.

Across broad sectors, it is generally thought that the elasticity of substitution is less than one – see Ngai and Pissarides (2004) and Herrendorf, Rogerson and Valentinyi (2013). However, *within manufacturing*, the elasticity of substitution is thought to be *more than one* – see Anderson and Van Wincoop (2004) and Ilyina and Samaniego (2012). Thus, our

[11] To see this, note that the Lorenz curve of industry composition when there are 2 industries is a line joining $(0, 0)$ to $(0.5, \min\{s_1, s_2\})$ and another line joining $(0.5, \min\{s_1, s_2\})$ to $(1, 1)$. The Gini coefficient is defined as the integral of the area above this line.

theory has two strong predictions:

1. within manufacturing, as countries develop they shift resources towards high-TFP growth industries;

2. in contrast, across broad sectors, as countries develop they shift resources towards low-TFP growth sectors.

Indeed, if verified, these predictions are of independent interest, as they constitute a hitherto unknown pattern of structural transformation along the development path, and as they substantiate the importance of productivity growth differences as an important mechanism underlying structural transformation.

To test whether the data support these predictions, we first need a measure of productivity growth g_i for different sectors and different manufacturing industries. We obtain these using US data in a manner described in detail in our quantitative experiments in Section 4.

Using these measures, we compute a time series for the weighted average value g_i across sectors and (separately) within manufacturing for each country and at each date.[12] The weights are the employment shares of each sector as reported in the ILO data, or of each manufacturing industry as reported in the UNIDO data. Finally, we examine the empirical relationship between average TFP growth across sectors and GDP per capita (as well as between average TFP growth across manufacturing industries and GDP per capita) by applying the nonparametric method in IW to this measure instead of the Gini coefficient of industry shares. This tells us whether there are systematic shifts in resources among industries with different TFP growth rates along the development path. TFP growth rates g_i in this experiment are assumed constant in each industry across time and across countries, so any observed patterns are solely due to patterns of specialization along the development path.

Figure 2 shows the estimated curves of industry- and sector-weighted average TFP growth rates. Within manufacturing, there is a mostly positive relationship with income, indicating that, behind the "stages of diversification", economic structure shifts towards industries with rapid TFP growth. In the ILO sector data, by contrast, there is a negative relationship. These results strongly support the idea that TFP growth differences are an important factor or structural transformation along the development path, and it is striking that the opposite is happening within manufacturing and across broad sectors, consistent with the

[12] For this exercise we normalize g_i so that the mean measure is zero and the standard deviation is one.

view in the literature that manufacturing goods are substitutes whereas sector-level goods are complements.

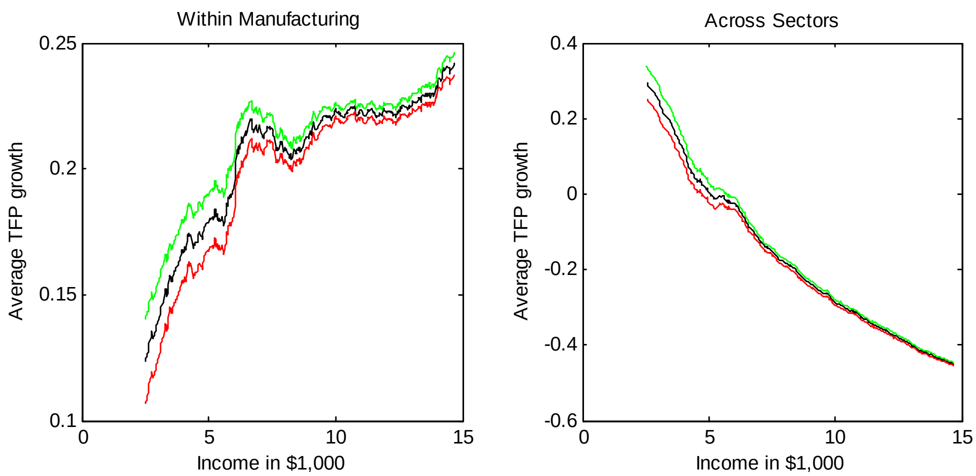

Figure 2 – Trends in average TFP growthwithin manufacturing and across sectors along the development path. The middle line is the point estimate, whereas the other lines reflect the point estimate plus or minus the standard error of $\hat{\beta}(x)$, as in IW.

Another way of testing these predictions is to examine whether changes in the shares of individual industries along the development path are systematically related to g_i. For example, within manufacturing, we should observe that the share of the highest-TFP growth industry (Non-electrical machinery) rises with income per head, whereas the share of the lowest-TFP growth industry (Tobacco) declines with income per head. Figure 3 shows that this is exactly the case, where the curves are estimated by applying the IW method to the share of each industry. More broadly, comparing across all industries in manufacturing, we expect the slope of the relationship between the share of a given industry and GDP to be *positively* related to its level of g_i. Indeed, the correlation between the slopes and g_i is 0.60. On the other hand, across *sectors*, the slope of the relationship between sector shares and GDP should be *negatively* related to g_i. The correlation coefficient is -0.65.

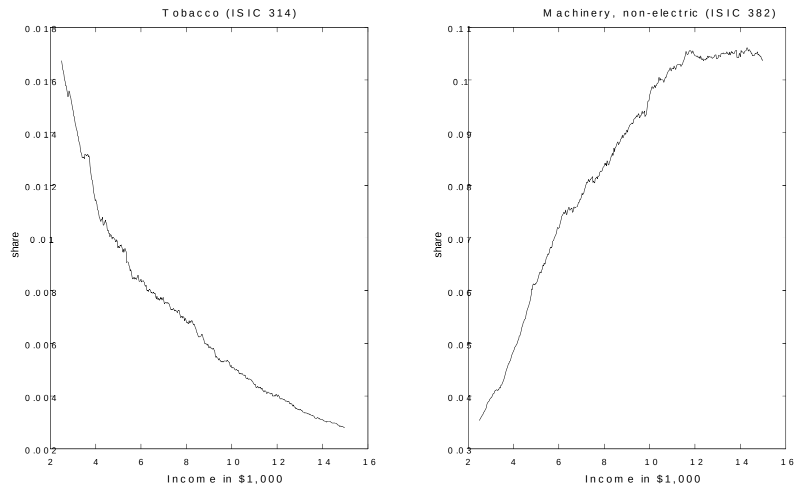

Figure 3 – Trends in industry shares along the development path. The curves represent the link between industry shares and income per head for the highest- and lowest-TFP growth industries, Non-electrical Machinery and Tobacco, computed using the IW methodology.

3 Model economy

We now present a simple model that can account for these shifts in employment among industries and sectors, as well as accounting for the stages of diversification.

Suppose there are N competitive industries, and C countries, with production functions of the form:

$$y_{cit} = A_{ct} A_{it} n_{cit} \tag{2}$$

where y_{ict} and n_{ict} are output and labor in industry i, country c at date t, productivity A_{it} grows according to $A_{it} = A_{i0} g_i^t$ and A_{ct} is arbitrary for now. The growth factor g_i may vary across industries. As is common in this literature we abstract for now from capital to focus on the productivity mechanism, but discuss capital later in Section 5.[13]

Producers solve the problem

$$\max_{n_{it}} \{p_{cit} y_{cit} - w_{ct} n_{cit}\} \tag{3}$$

where p_{ict} is the price of good i, and w_{ct} is the wage.

[13] We abstract from intermediate goods because it would not significantly affect results. See the Appendix for details.

Assume all these goods are consumed and that preferences display constant elasticity of substitution, so that

$$u(\mathbf{y}_{ct}) = \left[\sum_{i=1}^{N} \xi_i \times y_{ict}^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}, \quad \sum_{i=1}^{N} \xi_i = 1 \tag{4}$$

where ε is the elasticity of substitution among goods and $\mathbf{y}_{ct}=\{y_{1ct}, ..., y_{Nct}\}$.

Henceforth we suppress the country subscripts c except where needed. Let v_{it} be value added in industry i, so $v_{it} = p_{it}y_{it}$ where p_{it} is the price of good i. Then define the growth factor of value added G_{it} as:

$$G_{it} = v_{i,t+1}/v_{it}.$$

On the demand side, the consumer's first order conditions imply $\frac{p_{it}}{p_{jt}} = \left(\frac{y_{j,t}}{y_{i,t}}\right)^{\frac{1}{\varepsilon_s}} \frac{\xi_i}{\xi_j}$, so that

$$\frac{G_{it}}{G_{jt}} = \left[\frac{\frac{p_{i,t+1}}{p_{it}}}{\frac{p_{j,t+1}}{p_{jt}}}\right]^{1-\varepsilon}. \tag{5}$$

On the supply side, it is straightforward to show that $\frac{A_{it}}{A_{jt}} = \frac{p_{jt}}{p_{it}}$. Thus, in equilibrium equation (5) becomes:

$$\frac{G_{it}}{G_{jt}} = \left[\frac{g_i}{g_j}\right]^{\varepsilon-1}. \tag{6}$$

Let $s_{i,t}$ be the share of manufacturing value added (or employment[14]) of industry i at date t. Given shares $s_{i,t}$ for one year t, we can compute shares for the next year $t+1$ by multiplying $s_{i,t}$ by $g_i^{\varepsilon-1}$, re-scaling so the shares add to one,[15] and repeating this procedure to get predicted shares for as many years as desired. Thus, given initial conditions, a value of ε, and productivity growth factors g_i, we can compute model-generated industry shares of manufacturing, and subject the resulting industry structure to the same nonparametric methodology as in IW to study whether productivity differences might be able to generate a U-shaped specialization pattern.

---

[14]Notice that, while we defined $G_{it} = v_{i,t+1}/v_{it}$, (6) would also hold if $G_{it} = n_{i,t+1}/n_{it}$. to see this, remember the household's first order conditions imply that $\frac{p_{it}}{p_{jt}} = \left(\frac{y_{j,t}}{y_{i,t}}\right)^{\frac{1}{\varepsilon_s}} \frac{\xi_{s,i}}{\xi_{s,j}}$. Plugging in the production functions and recalling that capital labor ratios are constant across industries yields $\frac{p_{it}}{p_{jt}} = \left(\frac{A_{jt}n_{jt}}{A_{it}n_{it}^{1-\alpha}}\right)^{\frac{1}{\varepsilon_s}} \frac{\xi_{s,i}}{\xi_{s,j}}$. Rearranging, we have that $\frac{n_{it}}{n_{jt}} = \frac{p_{it}y_{it}}{p_{jt}y_{jt}}$.

[15]Given $s_{i,t}$, let $z_{i,t+1} = s_{i,t}g_i^{\varepsilon-1}$. Then $s_{i,t+1} = z_{i,t+1}/\sum_{j=1}^{N} z_{j,t+1}$.

4 Quantitative Experiments

4.1 Calibration

We now examine whether the model presented above generates a U-shaped specialization pattern for the 9 sectors in the ILO data, which cover the entire private economy. To do this requires values of g_i for each industry or sector, as well as the elasticity ε, preference parameters ξ_i, initial productivity levels $A_{i,0}$ and a series for country productivity $A_{c,t}$.

To measure g_i for different sectors, we first take the manufacturing sector value of g_i to equal the average value of total factor productivity derived from the NBER manufacturing productivity database.[16] Then, we calibrate g_i for the other sectors using their inverse price growth rates relative to manufacturing, using price data drawn from the US Bureau of Economic Analysis. See Table 2 in the Appendix for the g_i values.

We obtain an estimate of ε by observing that equation (6) is equivalent to:

$$\log G_i = a + (\varepsilon - 1)\log g_i + \epsilon_i \tag{7}$$

where $a = \log G_j - \log g_j$ for some arbitrary industry j and ϵ_i is any unmodeled noise in the relationship. Thus, regressing value-added growth rates (or employment growth rates) on TFP growth rates yields a coefficient equal to $\varepsilon - 1$.

Using US data, and also pooling across countries, we estimate that ε is not significantly different from zero.[17] This is consistent with Herrendorf, Rogerson and Valentinyi (2013), who estimate that $\varepsilon = 0$ across agriculture, manufacturing and services. Thus, we set $\varepsilon = 0$ across sectors. Notice that the fact that we have an estimate of ε that is significantly different from one is strong corroborative evidence that industry productivity growth rates are indeed linked to structural change.

[16] We use total factor productivity rather than labor productivity because it is without loss of generality if factor shares are the same across industries (so the same values of g_i can be used in a model with capital, such as the one in the technical appendix). To see this, suppose the production function includes capital. In this case, $y_{it} = A_{it}k_{it}^{\alpha}n_{it}^{1-\alpha}$. The first order conditions can be written $p_{it}\alpha y_{it}/K_{it} = r_t$ (where r is the interest rate) and $p_{it}(1-\alpha)y_{it}/n_{it} = w_t$. Dividing one condition by the other we get that $\frac{K_{it}}{n_{it}} = \frac{\alpha w_t}{(1-\alpha)r_t}$. Then, dividing any of the first order conditions for industry i by that for j yields the result that $\frac{p_{it}}{p_{jt}} = \frac{A_{jt}}{A_{it}}$, as before. Later in the paper we explore the implications of allowing for differences in capital shares.

[17] We estimated equation (7) using US time series from the ILO database, allowing for autocorrelated errors and obtained an estimate of $\varepsilon - 1 = -0.84$ with a standard error of 0.31. Thus, the estimate of ε is not significantly different from zero. We repeated this procedure pooling the data for all countries and using country fixed effects, obtaining an estimate for $\varepsilon - 1$ of -1.16, with a standard error 0.19 such that once more ε is not significantly different from zero.

As a benchmark we assume that the values of g_i are constant across countries. The use of common TFP growth rates across countries provides the cleanest experiment, since it is only the initial conditions that vary across countries. For example, empirical country-specific TFP growth rates could be influenced by industrial structure: as discussed in the survey of Cohen (2011), industry size is viewed by some as a potential determinant of productivity-enhancing R&D and hence possibly of productivity growth.[18] Later we will check the robustness of our results by allowing for convergence dynamics. We do not use country-specific measures of industry TFP growth because they do not exist for a wide sample of countries.[19]

We set the utility weights $\xi_i = 1/N$, where N is the number of sectors or industries in the model. This is without loss of generality as it amounts to a normalization of the units used to measure production in each industry.[20] We set A_{i0} to match the sector shares in the initial year as reported in the ILO data.

Finally, we select A_{ct} to match GDP per person for each country in each year. Since A_{ct} affects all industries equally, it does not impact economic structure. We also develop and calibrate a version of the model that has capital and also generates a series for real GDP: that model is relegated to the Appendix because it turns out to entail considerable computational complications that obscure the simplicity of our results.

Our simulation procedure is as follows. For each country, we match the initial industry shares in 1969 drawn from the ILO database, and simulate a time series of future industry shares from 1970 until 1992 using equation (6). We then include the same country-time pairs as IW, so that we have a model-generated unbalanced panel of industry shares that is of the same dimensions as that in the IW database. We apply the IW non-parametric procedure to these pseudodata, and compare the fitted curves that link economic structure to GDP per head.

We also perform this exercise independently for the 28 manufacturing industries examined in IW – these are the industries in the ISIC revision 2 industry classification used by the UNIDO INDSTAT3 database. This requires an independent calibration of the model. The procedure is exactly the same as above, except for some differences in the data used to set

[18] At the same time it is worth noting that the evidence regarding this link is inconclusive, see Cohen (2011) and Ngai and Samaniego (2011).

[19] While in principle we could compute country-industry specific TFP growth rates using the country-specific UNIDO data, we found that in many cases the resulting rates were absurdly high or low. We interpret this as indicating that the investment data in those countries are likely subject to significant measurement error.

[20] For example, suppose I measure apples in numbers of apples and find that $\xi_{apples} = 1$ and $A_{apples,0} = 1$. I could choose to measure apples in units of "half an apple", then $\xi_{apples} = 0.5$ and $A_{apples,0} = 2$.

up the calibration.[21]

For our manufacturing simulation, we estimate g_i using the TFP values in the NBER manufacturing productivity database. Note that the NBER industry classification is 4-digit SIC. We use Domar weights to convert NBER SIC industry TFP growth values into values for the ISIC revision 2 classification used by UNIDO. The value of g_i is the industry average over time.[22] See Table 1 in the Appendix for the values of g_i.

For manufacturing data we require a different value of ε. We use the value $\varepsilon = 3.75$, which is estimated in Ilyina and Samaniego (2012) using equation (7) as above.[23] Notice that by this estimate $\varepsilon > 1$, consistent with independent estimates from the literature on international trade.[24] Again, the fact that we have an estimate of ε that is significantly different from one is strong corroborative evidence that industry productivity growth differences are indeed linked to structural transformation. Finally, as before, we use the initial industry shares in 1963 from the UNIDO employment data, and simulate a time series of future industry shares until 1992 using equation (6). To the subsample of these data that correspond to the entries in the IW data, we apply the IW non-parametric procedure to extract the model link between income and specialization.

4.2 Results

We follow the IW methodology and regress Gini coefficients generated from our simulation on income per capita. Our results display a similar U-shaped relation between sector concentration and income levels: see Figure 4. In addition, the turning point is roughly $9,000, as found by IW. This is the same across broad sectors and within manufacturing data, something that lends weight to the empirical relevance of the productivity mechanism.

[21] In the technical appendix we also develop a model with capital where there are broad sectors, one of which is manufacturing which is further disaggregated into sub-industries. Results turn out to be very similar.

[22] For robustness we also estimated g_i in three other ways, all of which generate estimates that are highly correlated with each other. First, we used the perpetual inventory method to derive productivity growth measures for the US from the UNIDO data itself. Second, we used the growth gap between industry value added and industry production indices in the UNIDO data as indicators of relative price changes in the US. Third, we derived a measure of g_i from the observed changes in employment patterms around the world: this measure of g_i provides an extremely tight fit to the data. Details are in the Appendix.

[23] Ilyina and Samaniego (2012) estimate this coefficient using the industry TFP and value-added data reported in Jorgenson et al (2007).

[24] See the survey of Anderson and Van Wincoop (2004).

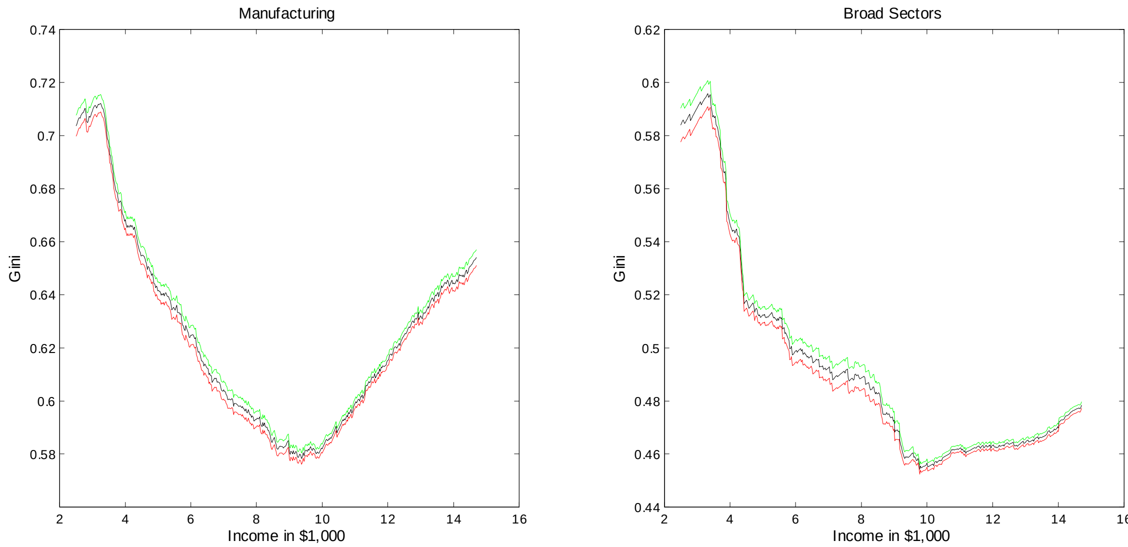

Figure 4. Industry structure along the development path in the simple model. The left panel is the relationship between income and specialization within manufacturing. The right panel is the same relationship across broad sectors.

While having constant values of g_i across countries, it is interesting to examine how sensitive the results are to this assumption. This requires imposing some structure on the manner in which productivity growth might vary across countries. We draw on the finding of Rodrik (2012) that there is unconditional convergence in productivity levels across manufacturing industries. This would suggest that, if a less developed economy c has productivity growth factor g_{ict} at date t, then $g_{ict} \rightarrow g_i$ over time.[25]

Suppose that

$$g_{ict} = g_{ct} g_i f\left(x_{c,t}\right)$$

where $x_{c,t} = Y_{c,t}/Y_{US,t}$ is the relative GDP gap between country c and the United States, and $g_{ct} = A_{c,t+1}/A_{c,t}$ is a country-specific productivity term. Thus, productivity convergence is a function of relative income, as is typically assumed in the empirical growth literature – see for example Barro and Sala-i-Martin (1992). In equilibrium, changes in the industry

[25] Unfortunately industry TFP data exist for only a small set of mostly developed countries. The finding of Rodrik (2012) relates to labor productivity: however it is hard to think of reasonable conditions under which labor productivity would converge while TFP does not. Note also that Rodrik (2012) tests for convergence in productivity *levels*, which is a stronger condition than convergence in productivity growth rates.

shares of manufacturing $G_{ict} = s_{i,c,t+1}/s_{ict}$ follow

$$\frac{G_{ict}}{G_{jct}} = \left(\frac{g_{ict}}{g_{jct}}\right)^{\varepsilon-1} = \left(\frac{g_{ct}g_i f(x_{c,t})}{g_{ct}g_j f(x_{c,t})}\right)^{\varepsilon-1} = \left(\frac{g_i}{g_j}\right)^{\varepsilon-1}.$$

Thus the only way in which convergence could affect the results is if the convergence function f is different *across industries.* For example, suppose that

$$g_{ict} = g_{ct} g_i f_i(x_{c,t}) = g_{ct} g_i^{x_{c,t}^{\eta}}$$

If $\eta < 0$, poorer countries experience disproportionately rapid convergence in high-g_i industries, capturing the idea that in high-g_i industries there is greater room for catchup. If $\eta > 0$, then the reverse is the case – catchup is relatively slow in high-g_i industries, as in the model of Ilyina and Samaniego (2012) where barriers to technology transfer in less-developed economies limit the R&D that would be necessary for them to catch up in high-tech industries.[26] Duarte and Restuccia (2010) find that labor productivity growth rates in agriculture is 1.4 percent higher than in manufacturing in the US, and that this gap varies across the 29 countries in their data by between -0.8 percent and 3.5 percent and is correlated with income, suggesting that empirically $\eta > 0$.

To explore the impact of industry-specific convergence, we solve the manufacturing model with values $\eta \in \{-0.3, 0.3\}$. A value of -0.3 implies that the difference in $\log g_i$ between any two industries or sectors is doubled in a country with 10 percent of US GDP per head. A value of 0.3 implies that the difference in $\log g_i$ between any two industries or sectors is halved in a country with 10 percent of US GDP per head. In general, negative values of η lower than -0.3 or -1 may imply huge industry productivity growth differences across industries that are not reasonable, whereas there is nothing a priori unreasonable about positive values: even $\eta \to \infty$ just means that TFP growth rates are roughly the same in all industries.

The results show that, although the exact shape of the specialization curve is sensitive to the value of η, the U-shape in manufacturing is preserved, see Figure 5. The results with $\eta = 0.3$ are better in the sense that, when $\eta = 0.3$, the initial diversification stage is monotonic, whereas in the benchmark results the lowest income countries seemed to display a little specialization at first. Thus, the results are robust to allowing for convergence in industry productivity growth rates – particularly if convergence is slower in high-TFP growth manufacturing industries, e.g. due to barriers to technology transfer.

[26]Note that in the calibrated model of Ilyina and Samaniego (2012) productivity growth in poor countries is relatively fast and convergence does take place: it is in a *relative* sense that catchup is slower in high-g_i industries.

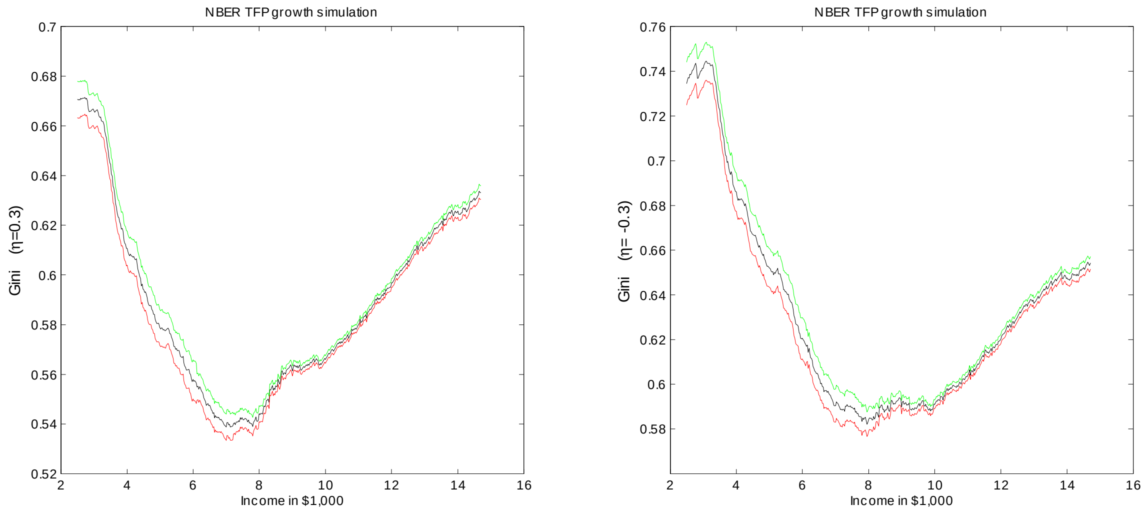

Figure 5 - Industry structure along the development path allowing for convergence in TFP growth rate across countries. Productivity growth g_{ict} equals $g_{ct}g_i^{x_{c,t}^{\eta}}$ where the output gap $x_{c,t}$ is defined as $Y_{c,t}/Y_{US,t}$.

4.3 Goodness of fit

Our aim is to see whether the calibrated model can generate a U-shaped pattern of specialization. At the same time, since we explore different versions of the model, it is useful to have some measures to compare the relative fit of different specialization curves.

First, we adopt a broad criterion for whether or not the model is generating "stages of diversification." Recall that the IW procedure involves regressing specialization measures on income within a window of income $[x - \Delta, x + \Delta]$, and then plotting the fitted value. We now study in addition the link between income and the *coefficients* generated in each of these regressions.

The reason for doing so is as follows. Out of necessity, we set the initial economic structure so as to match the data. As it happens, the initial conditions alone (ignoring data for future years) display a downward slope with respect to income, similar to the first "stage" in IW. If the model is itself generating the "stages" then not only should the fitted values in the IW regression display a "U" shape but the coefficients $\hat{\beta}(x)$ in these regressions should be increasing in income x, just as the slope of a "U" shape is monotonically increasing. The best fit will be obtained if, in addition, there is a threshold level of income below which the coefficients are negative and above which they are positive. IW find that the data do indeed display such a threshold.

We also report two quantitative measures of fit. Let $GINI_{IW}(x)$ equal the predicted Gini coefficient from IW, and let $GINI_{MODEL}(x)$ be the model Gini, each evaluated at income level $x \in [\underline{x}, \bar{x}]$. Define also $\overline{GINI}_{IW}$ and $\overline{GINI}_{MODEL}$ to be the respective means of each of these measures. One measure of goodness of fit could be simply the average distance between the IW curve and any curve generated from pseudodata. However, there are two reasons why a given curve might fail to exactly fit the empirical IW curve. One is that the given curve has a different shape from the IW curve. The other is that the given curve simply has a different mean. We find it useful to distinguish between these two. The former tells us about how well the model matches the *pattern* of specialization, whereas the second tells us about whether the model matches the general *level* of specialization in the data. Thus, we develop two measures of fit:

$$
\begin{aligned}
D_{mean} &= \overline{GINI}_{MODEL} - \overline{GINI}_{IW} \\
D_{shape} &= \left[\frac{\int_{\underline{x}}^{\bar{x}} \left(GINI_{IW}(x) - GINI_{MODEL}(x) + D_{mean} \right)^2 dx}{\bar{x} - \underline{x}} \right]^{1/2}
\end{aligned}
$$

D_{mean} is simply the difference in means, whereas D_{shape} is the average difference between the two lines minus each of their means.

To interpret these ginis (or their differences) consider the following. Suppose all the industries are exactly the same size, except for one industry, which is larger than all the others. Then with sufficiently many industries, the Gini coefficient is the share of the largest industry.

First, we look at whether the model coefficients $\hat{\beta}(x)$ are increasing. This implies that industry dynamics in the model economy are generating the U-shape, as opposed to having the initial "diversification" stage being driven by initial conditions.

Across broad sectors Figure 6 shows that, for the benchmark value of $\eta = 0$ (no convergence, or symmetric industry convergence) the coefficients show that the model itself is indeed generating the "stages." The coefficients are broadly increasing in income, starting negative and ending positive. Thus it is not the initial conditions that create the "U" shape: the model itself generates it. Indeed, the same holds for values of η that are positive, or negative but not too large in magnitude. Thus we conclude that our sector-level results are very robust to different convergence dynamics.

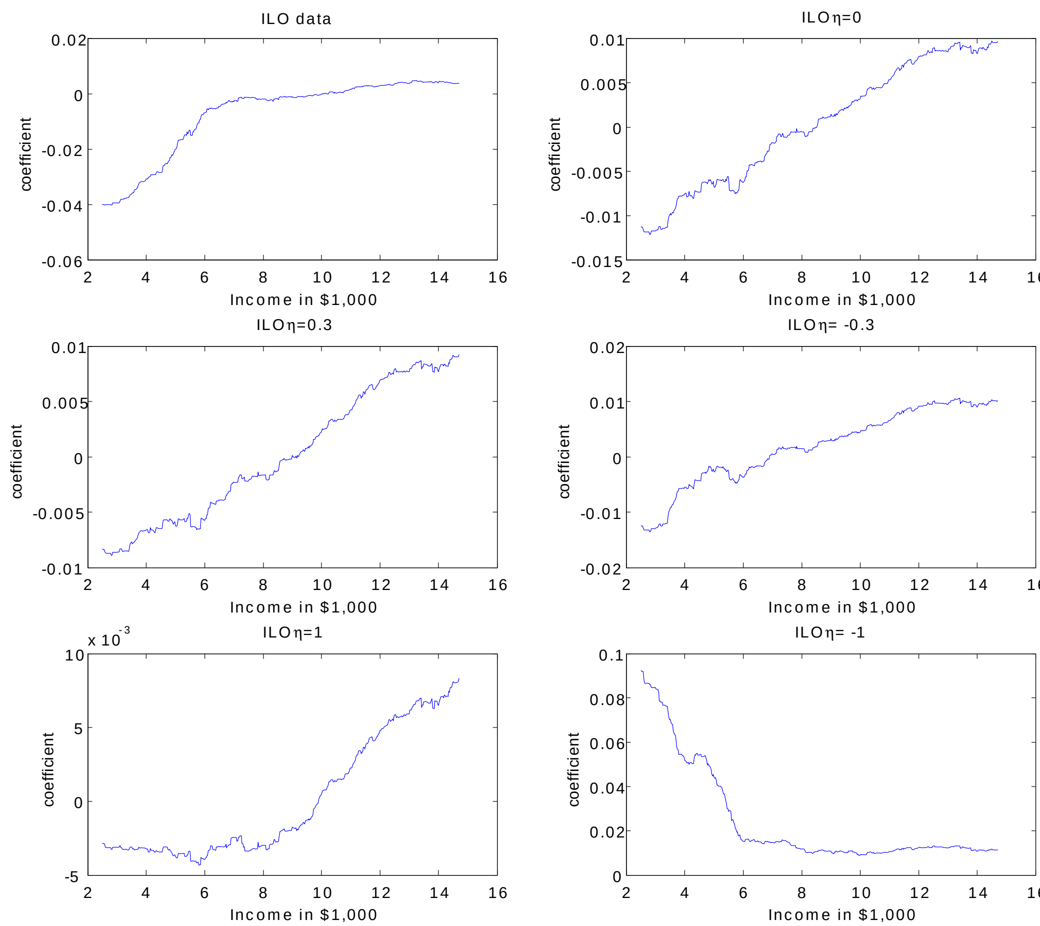

Figure 6 – Coefficients $\hat{\beta}(x)$ in the Gini regressions over income according to the IW procedure, where $GINI_{c,t}$ equals $\hat{\alpha}_c(x) + \hat{\beta}(x)\,GDP_{c,t} + \varepsilon_{c,t}$, for $GDP_{c,t}$ values in the window $[x - \Delta/2, x + \Delta/2]$, across broad sectors in the ILO data. The top left panel is based on data from IW, the other panels are model-generated.

Figure 7 plots the regression coefficients within manufacturing. In this case, when $\eta = 0$ the coefficients are all positive and not upward-sloping below about \$6000. Thus suggests that the initial conditions play a part in matching the downward slope below this level of income. The same is true when $\eta < 0$. However, when $\eta = 0.3$, we find that the coefficients are upward-sloping. When $\eta = 1$, we find that they are upward-sloping and negative below about \$5000. Thus, the best fit of the model is when $\eta > 0$, so that TFP growth differences among industries are compressed in developing countries (as found by Ilyina and Samaniego (2012)).

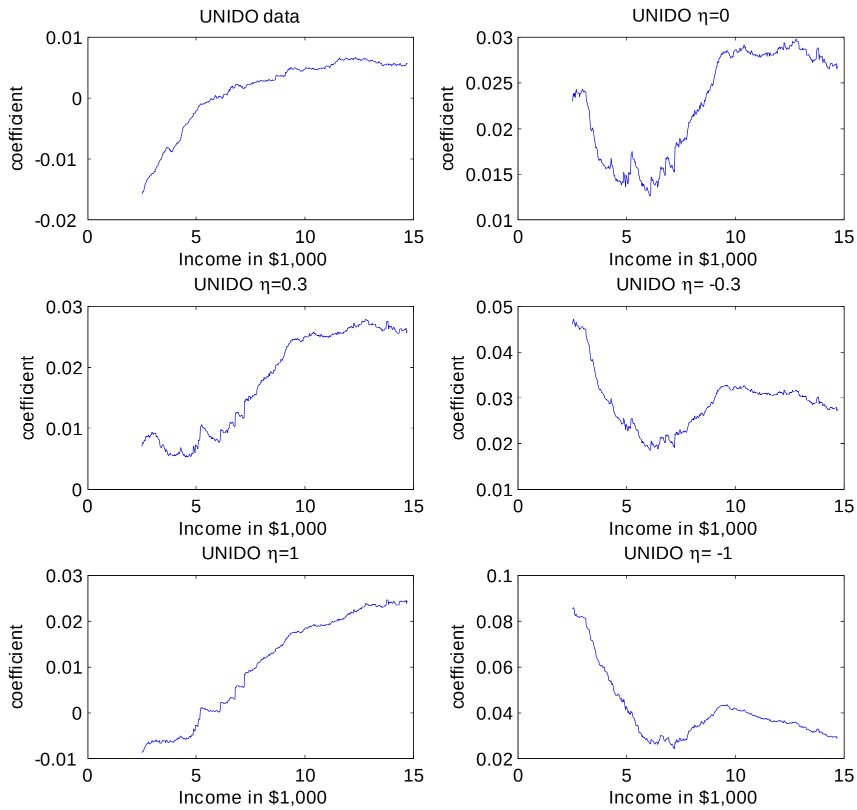

Figure 7 – Coefficients $\hat{\beta}(x)$ in the Gini regressions over income according to the IW procedure, where $GINI_{c,t}$ equals $\hat{\alpha}_c(x) + \hat{\beta}(x) GDP_{c,t} + \varepsilon_{c,t}$, for $GDP_{c,t}$ values in the window $[x - \Delta/2, x + \Delta/2]$, across manufacturing industries in the UNIDO data. The top left panel is based on data from IW, the other panels are model-generated.

Next, we investigate how the models compare in terms of D_{mean} (the difference of the means) and D_{shape} (the average distance between lines minus their means).

For the sector level results, we find that the fit of the model is U-shaped in η, see Figure 8. There is a little extra specialization when $\eta = 0$, a difference in the average Gini coefficient of about 3.5. This is equivalent to comparing two countries where all of the industries except one are the same size, but in one the largest industry has share 0.53, whereas in the other it is 0.56 – not a large difference. In terms of shape, again $\eta = 0$ seems to deliver the minimum, with an average shape difference again of about 3.5. We conclude that the benchmark model fits the sector results very well.

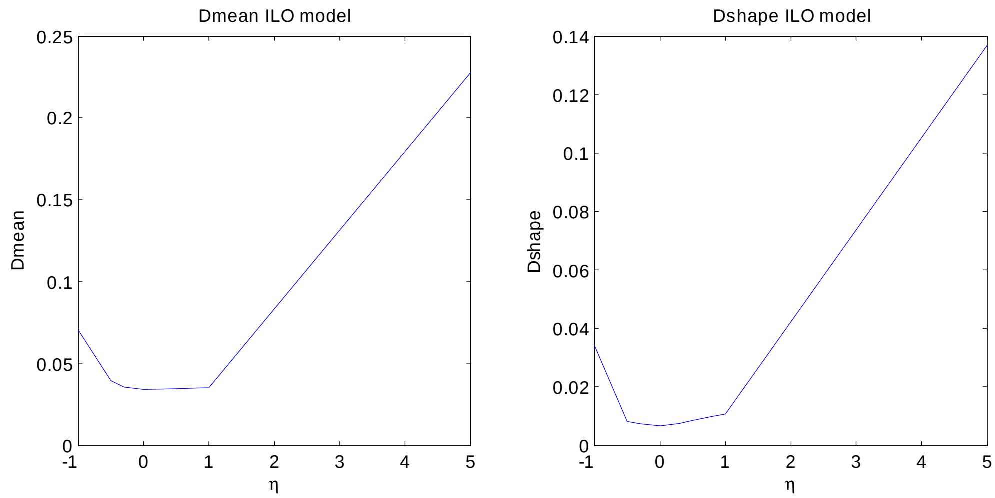

Figure 8– D_{mean} (the difference of the means) and D_{shape} (the average distance between lines minus their means) against η in the calibration for sectors, allowing for convergence in TFP growth rate across countries. g_{ict} equals $g_{ct} g_i^{x_{c,t}^{\eta}}$, where the output gap $x_{c,t}$ is defined as $Y_{c,t}/Y_{US,t}$.

For the manufacturing level results, we find that the best match in terms of means is around $\eta = 5$, see Figure 9. In other words, the model best matches the level of specialization in the data when industry TFP growth rates are compressed. At that point the match is very close, with a difference in mean Ginis of well under one. Interestingly there is no monotonic relationship between D_{shape} and η. The global minimum is around $\eta = 4$, but over the range considered there is little variation in D_{shape} over this range of η values. In fact, ignoring the fact that they have different means, the fitted curves for $\eta \in [0, 5]$ are visually indistinguishable from each other. In addition, the detrended specialization curves are very close to those of the data, especially for the sector level results (Figure 10). Thus, once more, the manufacturing results point once more towards an environment where industry TFP growth rates are compressed in less developed economies.

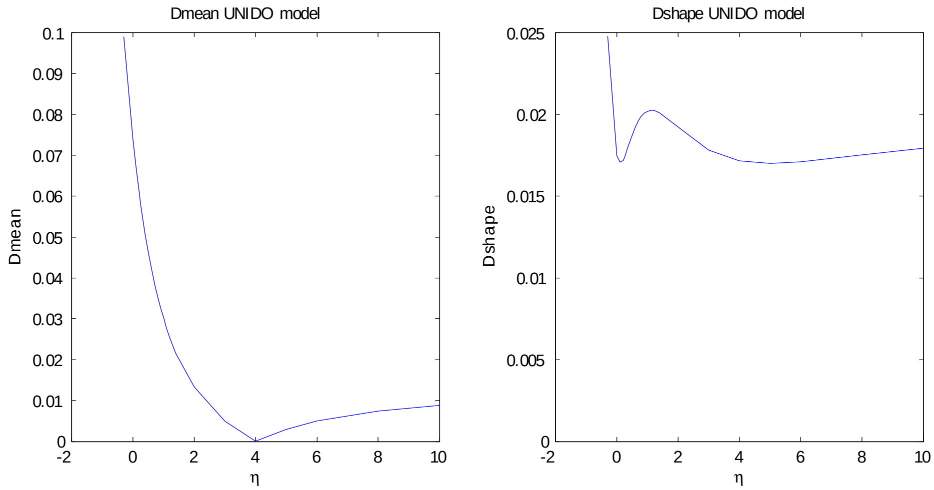

Figure 9–D_{mean} (the difference of the means) and D_{shape} (the average distance between lines minus their means) against η in manufacturing calibration, allowing for convergence in TFP growth rate across countries (g_{ict} equals $g_{ct}g_i^{x_{c,t}^{\eta}}$, where the output gap $x_{c,t}$ is defined as $Y_{c,t}/Y_{US,t}$).

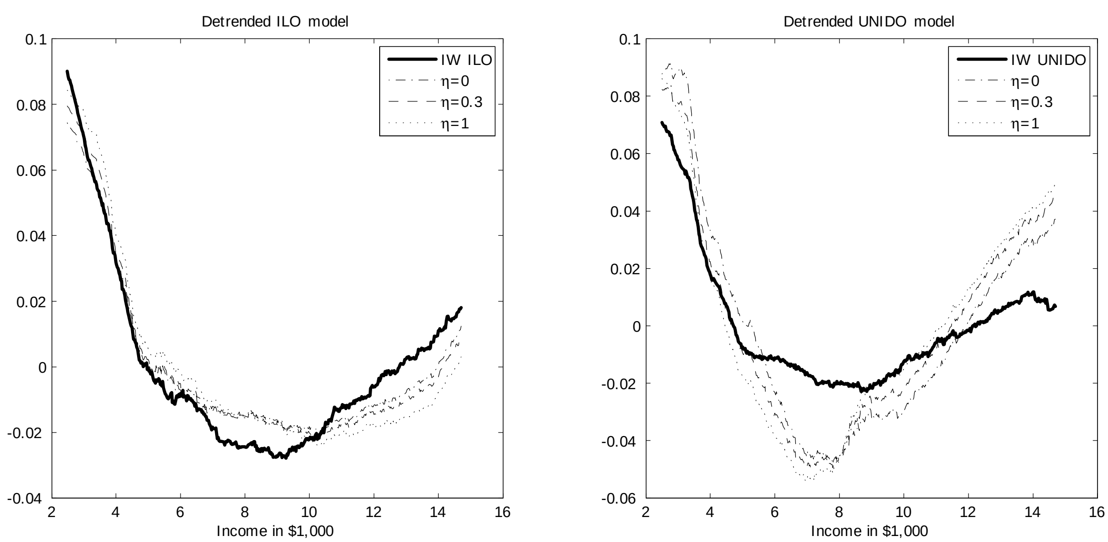

Figure 10 – Fitted measures of industry concentration (Ginis) for the data in IW and for model-generated pseudodata, each minus the mean value of the Gini ($GINI\,(x) - \overline{GINI}$ for both model and the data). The left figure is ILO data and simulations. The right one is for manufacturing data and simulations.

5 Factors of structural transformation

Our model focuses on productivity differences as the motor of structural change. However, there are other theories of long-run structural change that imply a shift in resources towards particular industries in the long run. As long as countries begin specialized in industries other than those that dominate in the long run, those models too might predict stages of diversification.

At least four general equilibrium frameworks have recently been developed to think about long-term structural change:[27]

1. Ngai and Pissarides (2007, NP) emphasize persistent productivity differences across industries, as we do.

2. Ilyina and Samaniego (2012, IS) emphasize productivity differences driven by differences in desired R&D intensity. This theory is not at odds with that of NP, but digs deeper as to the underlying causes of TFP growth differences.

3. Acemoglu and Guerrieri (2008, AG) consider both productivity differences and differences in capital shares. Specifically they predict that TFP growth rates divided by labor shares determine which industries tend to dominate in the long run. Differences in capital shares across industries could be a factor of structural change, since capital deepening along the growth path could have a differential impact on industries based on this factor.

4. Buera and Kaboski (2012) argue that structural change is affected by industry differences in firm size, with poorer countries less able to afford large-scale technologies.

Is there a way to see whether any of these mechanisms is related to the "stages"? To answer this question, we try two approaches. First, we adapt our model to allow for differences in capital shares across industries. Second, we see whether measures of R&D intensity, labor intensity and firm size display any clear relationship to the stages of development.

[27] In addition, Kongsamut et al (2000) relate structural change to differences in requirements for different goods – so that, for example, instead of utility over consumption of good i being defined over c_i, it is defined over $c_i - \bar{c}_i$, where $\bar{c}_i > 0$ is a constant. We do not think of this account of structural change as being relevant for the "stages" because $\bar{c}_i$ is generally interpreted in terms of requirements for agricultural goods, something that is not obviously relevant for disaggregated manufacturing industries. At the sector level, we have 9 sectors, so again requirements would likely be relevant for at most one sector.

5.1 Capital shares

So far we have assumed that there is no capital. First, if capital shares are the same across industries then the pattern of structural change implied by the model would be the same. Suppose the production function includes capital. In this case, $y_{it} = A_{it}k_{it}^{\alpha}n_{it}^{1-\alpha}$. The first order conditions can be written $p_{it}\alpha y_{it}/K_{it} = r_t$ (where r is the interest rate) and $p_{it}(1-\alpha)y_{it}/n_{it} = w_t$. Dividing one condition by the other we get that $\frac{K_{it}}{n_{it}} = \frac{\alpha w_t}{(1-\alpha)r_t}$. Then, dividing any of the first order conditions for industry i by that for j yields the result that $\frac{p_{it}}{p_{jt}} = \frac{A_{jt}}{A_{it}}$ so, as before, structural change is driven by equation (6).

Thus, the question is whether industry *differences* in capital shares might affect our results.

Consider the model presented earlier, except that the production structure allows capital shares α_i to vary across industries.

$$y_{it} = A_{it}K_{it}^{\alpha_i}n_{it}^{1-\alpha_i}. \tag{8}$$

It is no longer the case that $\frac{p_{it}}{p_{jt}} = \frac{A_{jt}}{A_{it}}$ in equilibrium. Instead, $\frac{p_{it}}{p_{jt}}$ becomes a function of the capital shares α_i, the wage w_t and the interest rate r_t. Letting G_i equal value added growth in industry i, in equilibrium it can be shown that:

$$\frac{G_i}{G_j} = \left(\frac{g_i}{g_j}\right)^{\varepsilon-1}\left(\frac{g_w}{g_r}\right)^{(\alpha_i-\alpha_j)(\varepsilon-1)} \tag{9}$$

which reduces to the earlier expression (6) when the capital shares are the same. In expression (9), $\frac{g_w}{g_r}$ is the growth factor of the wage rate divided by the growth factor of the interest rate. Thus, we could perform a simple experiment to gauge the implications of differences in capital shares for our results, as long as we had data on capital shares and a reasonable value for $\frac{g_w}{g_r}$.

Notice that expression (9) is equivalent to

$$\log G_i = c + (\varepsilon-1)\log g_i + (\varepsilon-1)\log\left(\frac{g_w}{g_r}\right)\alpha_i \tag{10}$$

where $c = \log G_j - (\varepsilon-1)\log g_j - (\varepsilon-1)\,\alpha_j\log\left(\frac{g_w}{g_r}\right)$ for some arbitrary industry j. Thus, we can estimate whether differences in capital shares are an important factor of structural change by estimating equation (10) and seeing whether capital shares carry a significant coefficient. Moreover, the coefficient on capital shares equals $(\varepsilon-1)\log\left(\frac{g_w}{g_r}\right)$ so that estimating this equation provides all the inputs necessary for us to simulate structural change in the model using equation (9) and assuming $\frac{g_w}{g_r}$ is relatively constant across time and space.

For manufacturing, we measure α_i using the INDSTAT4 database 1977-1990, as one minus the labor share of value added in each industry, as in Ilyina and Samaniego (2011). We do not have comparable data for broad sectors.

When we estimate (10), we find several things. First, the estimate of ε hardly changes, from 3.75 to 3.77, and remains significant at the 1% level. Second, the coefficient on α_i is significant at the 10% level only, and implies a value of $\frac{g_w}{g_r}$ of about 1.02 (although statistical significance improves to the 5% level if we allow for heteroskedasticity). This provides weak evidence that capital shares matter for structural change within manufacturing, while underlining the fact that productivity growth differences are important.

We simulate the basic model in two ways. First we use (9) to generate model industry shares. Second, we isolate the role of different capital shares by assuming that all industry productivity growth rates are equal, but allowing capital shares to differ. In both cases we impose the estimated value of $\frac{g_w}{g_r}$.

We find that allowing for capital share differences preserves the U shape. We do this with $\eta = 0$ (when the U-shape was partly dependent on initial conditions) and with $\eta = 1$ (a value of η which displays monotonic coefficients), suggesting that the presence of differences in α_i does not significantly affect the results. Moreover, when we simulate the model assuming TFP growth rates are the same (i.e. $g_i = g_j \ \forall i, j$) we find that there is no U, only a downward slope. Investigating the *coefficients* we find that they are almost flat and close to zero. In other words, the model itself with alpha differences generates no noticeable structural change, and the downward slope is purely generated by the initial conditions. We conclude that differences in capital shares are not significantly related to the stages of diversification.

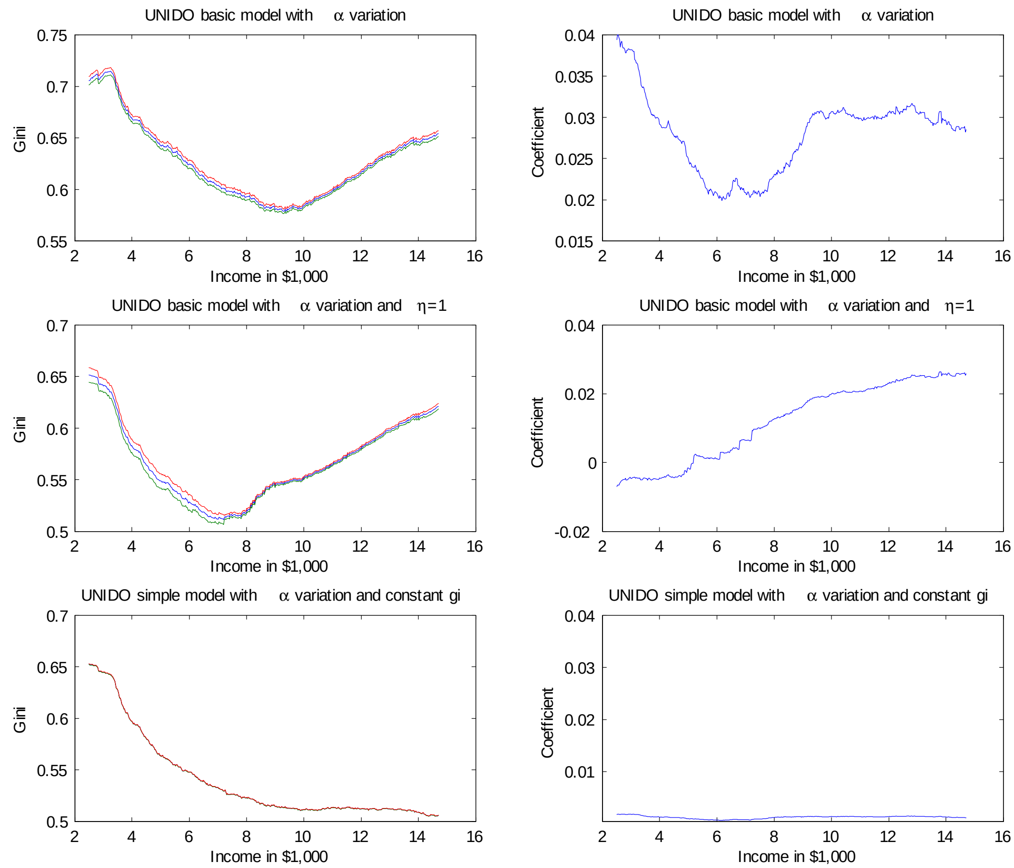

Figure 11 — Industry structure along the development path allowing for capital share variation (i.e., y_{it} equals $A_{it}K_{it}^{\alpha_i}n_{it}^{1-\alpha_i}$). The left panels represent the relationship between income and specialization within manufacturing, based on the IW regression where $GINI_{c,t}$ equals $\hat{\alpha}_c(x) + \hat{\beta}(x)\,GDP_{c,t} + \varepsilon_{c,t}$, for $GDP_{c,t}$ in the interval $[x - \Delta/2, x + \Delta/2]$. The right panels represent the coefficients $\hat{\beta}(x)$ from these regressions. The upper panels represent the model with α_i variation and $\eta = 0$ (no industry-specific productivity convergence).The middle panel is the model with η equal to one. The bottom panel is the model with α_i variation across industries setting $g_i = g_j \forall i, j$.

5.2 Other factors

To see whether structural change appears *empirically* related to any of the factors of structural change other than TFP growth rates (R&D intensity, labor intensity, firm size), we repeat the experiment illustrated in Figure 2 and compute series for the weighted average of each of these measures (R&D intensity, etc.) for each country over time. Industry R&D intensity and labor intensity measures are 3-decade averages of the values reported in Ilyina and Samaniego (2011).[28] The industry firm size is the average number of employees per establishment in the US over the period 1963-1992, as reported by UNIDO in INDSTAT3. Again, each measure is normalized so that the mean measure is zero and the standard deviation is one. Then, as before, weighted averages are computed for each country-year, where the weights are value added shares of each industry in total manufacturing, computed using UNIDO data. Finally, we apply the same nonparametric method to these measures, examining their relationship to real GDP per capita.

Figure 12 shows the link between each of these measures and income using the IW method. Average R&D displays a positive relationship with income within manufacturing, indicating that, behind the "stages of diversification", economic structure shifts towards industries with rapid TFP growth, and industries with high R&D intensity. Figure 12 also shows a link between average R&D intensity and income for the broad sectors in the ILO data:[29] in this case the trend is downwards, consistent with the assumption that $\varepsilon < 1$ across sectors. These results support the assumption that TFP growth differences can be a driving force behind structure change along the development path, and that behind the scenes TFP growth is related to R&D intensity.

Regarding the other measures, AG argue that differences in labor shares could be a driving force behind structural change. We can see that labor intensity shows a hump-shaped relationship with income. In particular, labor intensity declines beyond the income level of $10,000. Thus, while labor shares may be a factor of structural change, they do not appear to play a consistent role throughout the "stages."[30] This justifies our focus on a model with productivity differences, abstracting from differences in labor shares. As for firm size, average firm size declines along the development path before flattening out, which does not obviously support the idea that countries are more able to overcome large optimal scales of production as they develop. Again, this does not mean that the scale of production is not relevant for understanding some aspects of economic structure (e.g. the Buera and

[28] R&D intensity is the median R&D spending as a share of capital expenditures among firms in Compustat. Labor intensity is the wage share of value added in the INDSTAT3 database.

[29] For most measures, unfortunately we do not have data for sectors outside of manufacturing.

[30] More specifically, AG argue that the relevant variable is the productivity growth rate divided by the labor share. The weighted average of this variable across manufacturing industries behaves in a manner simmilar to the labor share.

Kaboski (2012) model is about resource shifts between the home and the market), but it is not obviously related to the "stages."

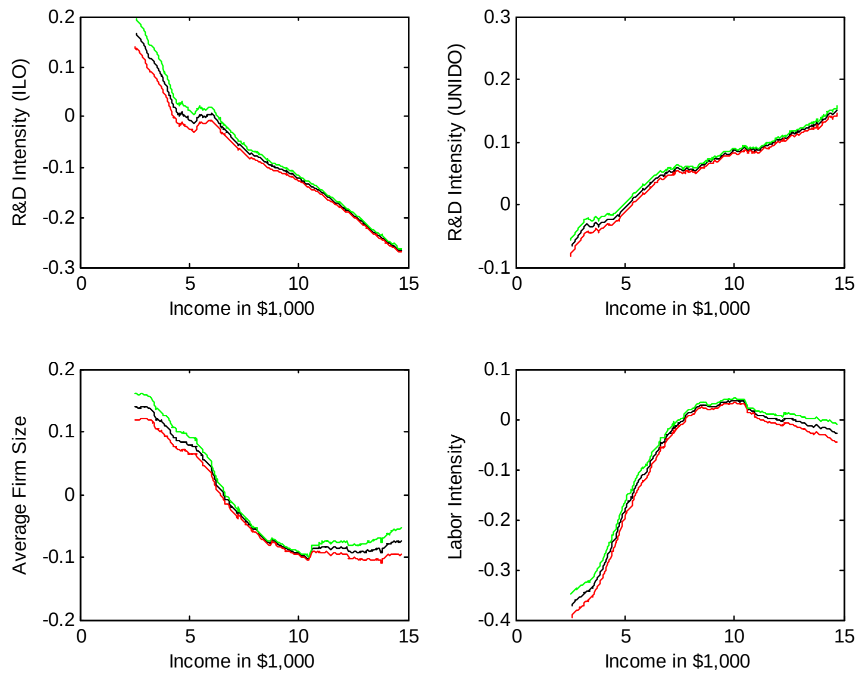

Figure 12 – Trends in average firm size, R&D intensity and labor intensity within manufacturing along the development path. The pattern of R&D intensity is consistent with TFP growth being related to R&D spending. The other measures do not display a monotonic relationship with income.

6 Concluding Remarks

This paper fills an important gap in the literature on structural change and growth by accounting for the broadest pattern of structural transformation along the development path: the "stages of diversification." We develop a multi-sector model in which differential TFP growth rates across industries and sectors lead to structural transformation. We find that the model accounts for the pattern of diversification followed by specialization – stages of diversification – that is well-known in the literature. The results are robust to a variety of extensions and modifications, and hold both across sectors and within manufacturing. This does not rule out a role for other factors, such as differences in factor shares or international trade. However, the paper provides quantitative evidence that productivity differences can account on their own for the dynamics of industrial structure along the development path. Importantly, the theory indicates that flows of resources may be different depending on the level of aggregation, something that the data confirm.

In general the literature tends to interpret the "stages" as being related to trade – most notably in the recent theory of integration and trade presented by Imbs et al (2012). This paper shows that even in a closed economy persistent total factor productivity (TFP) growth differences across industries are sufficient to generate a U-shaped pattern of specialization. It would be interesting in future work to develop a model that nests all the various possibilities, to estimate the contribution of different factors to the evolution of economic structure. For example, IW find that open economies tend to re-specialize at a lower income level, something that can only be analyzed in an open economy model.

In the paper we take initial conditions as given for our quantitative experiments. The results suggest that poorer countries tend to begin specialized in industries where TFP growth is low. Although it is beyond the scope of this paper, it is interesting to think about why initial conditions might be biased in this way. One possibility is that there are non-homothetic preferences (see Kongsamut et al (2000)), so that consumption patterns in poor countries are dominated by subsistence considerations that wear off later. If manufacturing industries that produce goods necessary for subsistence (e.g. food products) happen to have slow TFP growth, whereas sectors that are necessary for subsistence (e.g. agriculture) so happen to have rapid TFP growth, then we would observe these initial conditions. Another possibility that does not necessarily hinge on non-homothetic preferences involves the transition from a "traditional" technology with low productivity growth to a "modern" technology with more rapid productivity growth. Ngai (2004) shows that small differences across countries in barriers to technology adoption can lead to very large differences in income by delaying the transition from the "traditional" to the "modern" technologies. Initial conditions would be determined by the traditional technology and the date of transition between technologies. The idea that the transition between the "traditional" and "modern" technologies could explain economic structure as well as income levels is an interesting topic for future research.

References

[1] Acemoglu, Daron, and Veronica Guerrieri. 2008. "Capital Deepening and Nonbalanced Economic Growth." Journal of Political Economy 116 (3): 467-498.

[2] James E. Anderson & Eric van Wincoop, 2004. "Trade Costs," Journal of Economic Literature 42(3), 691-751.

[3] Barro, Robert J. 1991. "Economic Growth in a Cross Section of Countries." Quarterly Journal of Economics 106 (2): 407-443.

[4] Barro, Robert J and Sala-i-Martin, Xavier, 1992. "Convergence," Journal of Political Economy 100(2), 223-51.

[5] Buera, Francisco J. & Kaboski, Joseph P., 2012. "Scale and the origins of structural change," Journal of Economic Theory 147(2), 684-712.

[6] Cohen, Wesley M. 2011. Fifty Years of Empirical studies of innovative activity and performance. In B. Hall & N. Rosenberg (ed.) Handbook of the Economics of Innovation, Chapter 7, 129-213.

[7] Colucci, Domenico. 2001. "Limited Computational Ability and Approximation of Dynamical Systems." Computational Economics 17: 155-178.

[8] Cummins, Jason G. and Giovanni L. Violante, 2002. "Investment-Specific Technical Change in the US (1947-2000): Measurement and Macroeconomic Consequences," Review of Economic Dynamics 5(2), 243-284.

[9] Duarte, Margarida and Diego Restuccia. 2010. "The Role of the Structural Transformation in Aggregate productivity." Quarterly Journal of Economics 125 (1): 129-173.

[10] Gollin, Douglas & Parente, Stephen L. & Rogerson, Richard, 2007. "The food problem and the evolution of international income levels," Journal of Monetary Economics 54(4) 1230-1255.

[11] Greenwood, Jeremy, Zvi Hercowitz and Per Krusell. 1997. "Long-Run Implications of Investment-Specific Technological Change." The American Economic Review 87 (3): 342-362.

[12] He, Hui and Zheng Liu. 2008. "Investment-Specific Technological Change, Skill Accumulation, and Wage Inequality." Review of Economic Dynamics 11, 314-334.

[13] Herrendorf, Berthold, Christopher Herrington and Akos Valentinyi. 2013. Technology and Structural Transformation. Mimeo: Arizona State University.

[14] Herrendorf, Berthold, Richard Rogerson and Akos Valentinyi. 2013. "Two Perspectives on Preferences and Structural Transformation." American Economic Review, forthcoming.

[15] Hsieh, Chang-Tai and Klenow, Peter. 2007. "Relative Prices and Relative Prosperity," American Economic Review 97, 562-585.

[16] Imbs, Jean, and Romain Wacziarg. 2003. "Stages of Diversification." American Economic Review 93 (1): 63-86.

[17] Imbs, Jean, Claudio Montenegro and Romain Wacziarg. 2012. Economic Integration and Structural Change. Mimeo: UCLA.

[18] Ilyina, Anna, and Roberto M. Samaniego. 2011. "Technology and Financial Development." Journal of Money, Credit and Banking 43 (5), 899-921.

[19] Ilyina, Anna, and Roberto M. Samaniego. 2012. "Structural Change and Financing Constraints." Journal of Monetary Economics 59 (2), 166-179.

[20] Jermann, Urban. 1998. "Asset Pricing in Production Economies." Journal of Monetary Economics 41(2), 257-275.

[21] Jorgensen, Dale, Mun S. Ho, Jon Samuels and Kevin J. Stiroh, 2007, "Industry Origins of the American Productivity Resurgence," Economic Systems Research 19(2), 229-252.

[22] Kalemli-Ozcan, Sebnem & Bent E. Sorensen & Oved Yosha, 2003. "Risk Sharing and Industrial Specialization: Regional and International Evidence," American Economic Review 93(3), 903-918.

[23] Koren, Miklós and Silvana Tenreyro, 2007. Volatility and Development. Quarterly Journal of Economics 122(1), 243-287.

[24] Krusell, Per & Lee E. Ohanian & José-Victor Ríos-Rull & Giovanni L. Violante, 2000. "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," Econometrica 68(5), 1029-1054.

[25] Ngai, L. Rachel, and Christopher A. Pissarides. 2004. "Balanced Growth with Structural Change." CEP Discussion Paper 627.

[26] Ngai, L. Rachel, and Christopher A. Pissarides. 2007. "Structural Change in a Multisector Model of Growth." The American Economic Review 97 (1): 429-443.

[27] Ngai, L. Rachel, and Roberto M. Samaniego. 2011. "Accounting for Research and Productivity Growth Across Industries." Review of Economic Dynamics 14 (3): 475-495.

[28] Rodrik, Dani. 2008. "Industrial Development; Stylized Facts and Policies." In *Industrial Development for the 21st Century.* Edited by David O'Connor and Monica Kjollerstrom. United Nations: New York.

[29] Rodrik, Dani. 2012. "Unconditional Convergence in Manufacturing." Mimeo: Harvard University.

[30] Restuccia, Diego & Yang, Dennis Tao & Zhu, Xiaodong, 2008. "Agriculture and aggregate productivity: A quantitative cross-country analysis." Journal of Monetary Economics 55(2), 234-250.

[31] Rogerson, Richard. 2008. "Structural Transformation and the Deterioration of European Labor Market Outcomes." Journal of Political Economy, 116(2), 235-259.

[32] Sala-i-Martin, Xavier X. 1997. "I Just Ran Two Million Regressions." American Economic Review 87 (2): 178-183.

[33] World Bank. Glossary. http://www.worldbank.org/depweb/english/beyond/global/glossary.html. Last checked 3/30/2012.

Appendix

A Alternative Measurement of productivity in Manufacturing

We measure productivity using the NBER Manufacturing Productivity Database. The data are more disaggregated that the ISIC3 Classification we need for the UNIDO data, so we aggregate them using Domar weights.

In addition, we use an alternative way of measuring TFP growth rates. Using the UNIDO dataset, we compute the TFP growth rates of 28 UNIDO manufacturing industries of the United States using the following equation:

$$\ln(TFP_{it}) = \ln(Y_{it}) - (1-\alpha)\ln(L_{it}) - \alpha\ln(K_{it}) \tag{11}$$

where Y_{it} is the production index. This requires computing the capital stock at the industry level. The UNIDO dataset provides investment data but not capital stock data K_{it}, so we use a perpetual inventory method

$$K_{it+1} = (1-\delta)K_{it} + I_{it}q_{it} \tag{12}$$

to compute growth rate of capital stock, where I_{it} is investment and q_{it} represents investment-specific technical progress[31]. Then the growth rate of K_{it} is the sum of growth rates of I and q. We set $q_{it} = g_{iq}^t$, so that growth rates of q_i vary across industries. We use growth factor g_{iq} from IS. Also, $\delta = 0.06$ and $\alpha = 0.3$. These are standard numbers in the literature.[32] Then, if $\Gamma(x)$ is the log growth rate of x over the time period in the data, note that

$$\ln g_i = \Gamma(Y_i) - (1-\alpha)\Gamma(L_i) - \alpha\Gamma(K_i) \tag{13}$$

We obtain $\Gamma(Y_i)$ and $\Gamma(L_i)$ from UNIDO, and set $\Gamma(K_i) = \Gamma(I_i) + \log g_{iq}$, which is the long run relationship in (12).

In addition, equation (5) suggests using inverse price growth rates to measure industry TFP growth. The price index is computed using value added divided by the production index from the UNIDO dataset.[33] Both TFP and price growth rates are averages over the period $1963-1992$ (data are available upon request). TFP growth rates computed this way are highly correlated with those derived from the NBER data, with a correlation coefficient of 0.6 (significant at the 5 percent level). The TFP growth and price growth series based on UNIDO data are highly negatively correlated with a coefficient of -0.9 (significant at the 5 percent level). All of this is encouraging as to the robustness of the productivity measures.

We simulate industry shares following equation (5) for UNIDO price growth and (6) for TFP growth and apply nonparametric methodology to model simulated Gini coefficients on income. Again, we obtain a U-shape in both cases, see Figure 13.

[31] We allow for investment-specific technical progress because the model is one with many industries where productivity growth rates in capital-producing industries may be different from productivity growth elsewhere.

[32] The value of δ is from Greenwood, Hercowitz and Krusell (1997) and is a value typical in models with investment-specific technical change, in other words where $g_q > 1$.

[33] Recall that value added $v_{it} = p_{it}y_{it}$. The assumption is that growth in the UNIDO industrial production index proxies for growth in y_{it}.

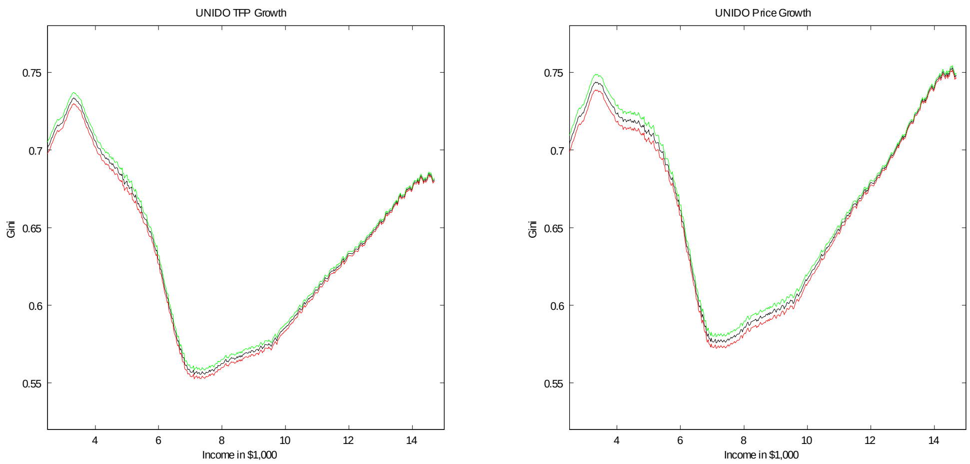

Figure 13. IW nonparametric regression using simulated industry concentration measures based on equation (6) using INDSTAT3 TFP growth rates, and based on equation (5) using INDSTAT3 price growth rates.

Finally, using the model itself, we can map between productivity growth rates and changes over time in employment shares by means of equation (6), setting the errors $\epsilon_t = 0$. We use this procedure to back out a measure of productivity growth we call g_i^{median}, which will be the median productivity growth factor implied by changes in employment patterns around the world (details below). We repeat the above experiments using g_i^{median} for all countries, again providing a clean experiment that abstracts from all country differences other than initial conditions.

All that is required to use (6) to back out productivity growth values from country employment data is knowing g_{jct} for some benchmark industry j in each country. We assume that in all countries and dates the productivity growth rate is the same in industry 342, Printing and Publishing. We choose this industry because in the NBER data $g_j \simeq 1$, for industry 342, i.e. there is essentially no total factor productivity growth in this industry. At the sector level, we assume that productivity growth is the same in all countries for Community, Social and Personal Services, which has the lowest g_i value among the sectors.

Having generated series for g_{ict} as described above, we compute g_{ic}, the average of the time series for each country and industry. The measure g_i^{median} is simply the median value of g_{ic} for each i. Note that, since the country with the median value will vary by industry i, no country's employment patterns correspond to g_i^{median}, even though g_i^{median} is derived from the employment data.

First, it turns out that g_i^{median} is very highly correlated with the values of g_i calibrated using NBER data. Within manufacturing the correlation is 0.40 and across sectors it is 0.70, both significant at the five percent level. Second, Figure 14 displays the link between specialization and development using the basic model, using g_i^{median} as a measure of productivity growth. Again, the findings are robust. When we apply the measures of fit to these curves, we find that the curves in Figure 13 fit a bit worse than when we measure g_i using NBER data, and that the coefficients are not monotonic unless we allow for convergence parameter $\eta > 0$, as before. However, the curves using g_i^{median} provide an extremely tight fit to the data, with D_{mean} and D_{shape} an order of magnitude smaller than before. Also, with g_i^{median} the coefficients are also monotonic and negative below about \$5000 even when $\eta = 0$. Thus, the model essentially reproduces the empirical curves with g_i^{median}.

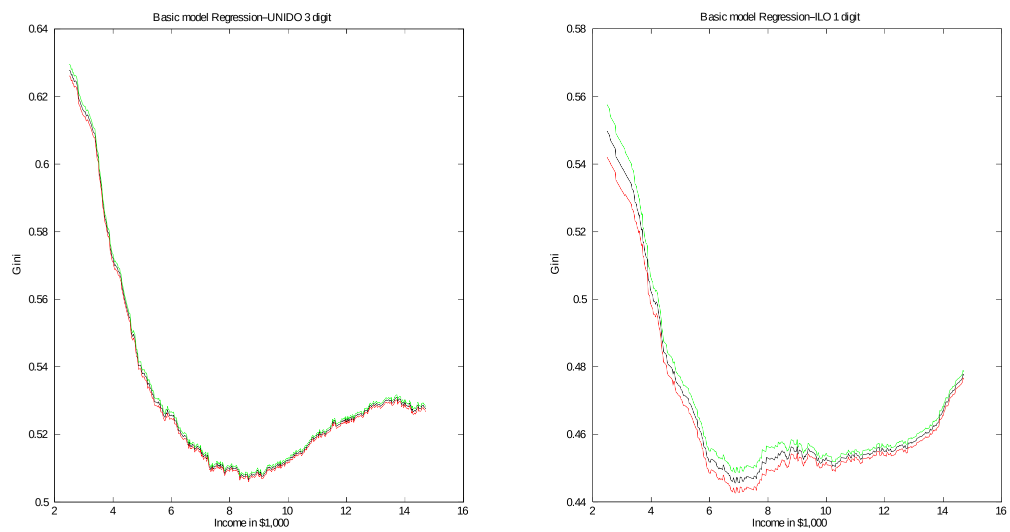

Figure 14 - Patterns of specialization in the basic model using g_i^{median}. This is the median TFP growth rate across countries measured using employment patterns as described in the text.

B Intermediate goods

We have abstracted from the existence of intermediate goods. There is a question as to whether results might change if we allowed for intermediates. For one thing, when there are intermediate goods, Ngai and Samaniego (2009) point out that TFP growth measures computed using gross output data (as is the case for the NBER numbers) understate TFP growth in a value added model.

Consider that the input-output matrix is largely diagonal: industries tend to use intermediates produced within the industry. Let us abstract from the off-diagonal elements. In this case, the production function is

$$y_{it} = A_{it} \left(K_{it}^{\alpha} n_{it}^{1-\alpha} \right)^{1-\psi} x^{\psi} \tag{14}$$

where x_{it} are intermediate goods and ψ is the intermediate goods share. Producers solve the problem

$$\max_{k_{it}, n_{it}, x_{it}} \left\{ p_{it} y_{it} - w_t n_{it} - r_t K_{it} - p_{it} x_{it} \right\}. \tag{15}$$

Solving for optimal use of x_{it}, It is easy to show that this is equivalent to solving

$$\max_{k_{it}, n_{it}} \left\{ p_{it} \tilde{y}_{it} - w_t n_{it} - r_t K_{it} \right\}. \tag{16}$$

where $\tilde{y}_{it} = y_{it} - x_{it}$ is value added in terms of good i, with the value-added production function

$$\tilde{y}_{it} = p_{it} \tilde{A}_{it} K_{it}^{\alpha} n_{it}^{1-\alpha} \tag{17}$$
$$\tilde{A}_{it} = A_{it}^{\frac{1}{1-\psi}} \left[\psi^{\frac{\psi}{1-\psi}} - \psi^{\frac{1}{1-\psi}} \right].$$

While (17) is of the same form as (23), note that the growth factor of $\tilde{A}_{it}$ is equal to $g_i^{\frac{1}{1-\psi}} > g_i$.

At the same time, this does not matter for the results. Recall that what affects rates of structural change in the model is the combination of g_i and ε, not one or the other in isolation. If we regress the log value added growth on the log real value-added productivity factor $g_i^{\frac{1}{1-\psi}}$, we would obtain a different value of epsilon. Recall that (6) is equivalent to $\log G_i = \alpha + (\varepsilon - 1) \log g_i + \epsilon_i$ where $\alpha = \log G_j - \log g_j$ for some arbitrary industry j and ϵ_i is any unmodeled noise. If we use g_i instead of $g_i^{\frac{1}{1-\psi}}$ in this equation, we would have an estimated elasticity $\tilde{\varepsilon}$ where $\tilde{\varepsilon} - 1 = (\varepsilon - 1)(1 - \psi)$ – a lower value, since $\psi < 1$. However, structural change within sectors would be driven by the following relationship

$$\frac{G_{it}}{G_{jt}} = \left[\frac{g_i^{\frac{1}{1-\psi}}}{g_j^{\frac{1}{1-\psi}}} \right]^{\tilde{\varepsilon}-1} = \left[\frac{g_i}{g_j} \right]^{\frac{\tilde{\varepsilon}-1}{1-\psi}} = \left[\frac{g_i}{g_j} \right]^{\varepsilon-1}, \tag{18}$$

which is exactly equivalent quantitatively to patterns of structural change in our model without intermediates. Thus, significant off-diagonal elements in the input-output tables would be required to change our quantitative results.

A similar intuition regards the possibility of adjustment costs in the reallocation of capital cross industries. Given other parameters, capital adjustment costs could slow the reallocation of resources across industries. However in the presence of adjustment costs the value of ε required to match the link between industry growth and TFP growth in the data would be larger. Thus, results would not be affected.

C Industry TFP growth data

Below we report the industry productivity growth rates used in the model calibrations. See the text for the measurement strategies.

The manufacturing measures are computed from the NBER manufacturing productivity database. That database is at the 4-digit SIC level, so we aggregate the measures using Domar weights. The sector level values are computed using relative price growth information from the US BEA. First, we compute the value for TFP growth in manufacturing using the NBER manufacturing productivity database. Then, the productivity growth factor in any given sector is the value for manufacturing times the factor by which the price of the good produced in that sector declines relative to manufacturing.

Table 1: NBER TFP Growth Rates for the ISIC revision 2 industry classification. Source: NBER productivity database and authors' calculations.

Industry	ISIC code	NBER TFP Growth Rate
Food products	311	0.0101
Beverages	313	0.0303
Tobacco	314	-0.0345
Textiles	321	0.0269
Apparel	322	0.0121
Leather	323	-0.0034
Footwear	324	-0.0035
Wood products	331	0.0113
Furniture, except metal	332	0.0066
Paper and products	341	0.0088
Printing and publishing	342	-0.0022
Industrial chemicals	351	0.0214
Other chemicals	352	0.0135
Petroleum refineries	353	0.0196
Misc. pet. and coal products	354	0.0223
Rubber products	355	0.0142
Plastic products	356	0.0339
Pottery, china, earthenware	361	0.0078
Glass and products	362	0.0051
Other non-metallic mineral prod.	369	0.0120
Iron and steel	371	0.0047
Non-ferrous metals	372	0.0016
Fabricated metal products	381	0.0029
Machinery, except electrical	382	0.0285
Machinery, electric	383	0.0347
Transport equipment	384	0.0160
Prof. & sci. equip.	385	0.0126
Other manufactured prod.	390	0.0089

Table 2: Productivity growth across broad ILO Sectors. The manufacturing sector is composed of the 28 ISIC2 industries. The value for Manufacturing in the table is the average over the sample period in the NBER data.

	ILO 1-Digit Classification (9 sectors)	Growth Factor g
1	Agriculture, Hunting, Forestry and Fishing	1.027
2	Mining and Quarrying	0.995
3	Manufacturing*	1.012
4	Electricity, Gas and Water	0.987
5	Construction	0.983
6	Wholesale and Retail Trade and Restaurants and Hotels	1.011
7	Transport, Storage and Communication	1.006
8	Financing, Insurance, Real Estate and Business Services	0.982
9	Community, Social and Personal Services	0.974

Productivity Growth and Structural Transformation: Technical Appendix

D General equilibrium model with capital

We now develop a multi-industry growth model with capital and ask whether the mechanisms described above can generate stages of diversification at the industry or sector levels in this environment in a general equilibrium framework.

D.1 Preferences and Technology

Time is discrete and there is a $[0,1]$ continuum of agents. There are S sectors, each of which produces an aggregate of I industries. Let I_s be the set of industries that supplies sector s. We focus on the case in which each industry supplies only one sector, so that $I_s \cap I_{s'} = \varnothing$, $\forall s \neq s'$. Note that this is without loss of generality, as one could have two industries identical in all ways that are distinguished by the fact that they provide a given good to two different sectors.

We assume that sectors $s \in \{1, ..., S-1\}$ produce consumption goods. Only one sector,

S, produces capital goods. Now for each sector $s \in \{1, ...S\}$, the production function has the CES form:

$$y_{st} = \left[\sum_{i \in I_s} \xi_i \times u_{s,i,t}^{\frac{\varepsilon_s - 1}{\varepsilon_s}} \right]^{\frac{\varepsilon_s}{\varepsilon_s - 1}}, \quad \sum_{i \in I_s} \xi_i = 1, \qquad s = 1, ..., S \tag{19}$$

where u_{sit} is use of good i by sector s, ξ_i is the weight on good i, and ε_s is the elasticity of substitution among goods within sector s.

Agents consume a CES aggregate c_t of the output of the different consumption sectors:

$$c_t = \left[\sum_{s=1}^{S-1} \zeta_s y_{st}^{\frac{\varepsilon - 1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon - 1}}.$$

Finally, agents have isoelastic preferences over c_t and discount the future using a factor $\beta < 1$, so that:

$$\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\theta} - 1}{1 - \theta}. \tag{20}$$

They are endowed with one unit of labor every period which they supply inelastically, and start period zero with capital K_0.

Let q_{st} be the price of the sector aggregate s, with r_t as the interest rate and w_t as the wage. Agents choose expenditure on each good so as to maximize (20) subject to the budget constraint

$$\sum_{s=1}^{S} q_{st} y_{st} \leq \sum_{s=1}^{S} \sum_{i \in I_s} r_t K_{it} + \sum_{s=1}^{S} \sum_{i \in I_s} w_t n_{it} \tag{21}$$

and the capital accumulation equation

$$K_{t+1} = y_{St} + (1 - \delta) K_t. \tag{22}$$

On the supply side, each industry features a Cobb-Douglas production function:

$$y_{it} = A_{it} K_{it}^{\alpha} n_{it}^{1-\alpha}, \;\; A_{it} = A_{i0} g_i^t \tag{23}$$

where $g_i = A_{i,t+1}/A_{it}$ is the TFP growth factor of industry i and A_{i0} is given. Producers maximize profits

$$\max_{n_{it}, K_{it}} \{ p_{it} y_{it} - w_t n_{it} - r_t K_{it} \} \tag{24}$$

subject to (23), where p_{it} is the output price of industry i at time t. Capital and labor are freely mobile across sectors – we discuss this assumption further in Section B.

D.2 Equilibrium

The producers' first order conditions imply that the capital labor ratio is constant across industries, which implies that $A_{it}p_{it} = A_{jt}p_{jt}$. Thus, as in related models, goods that experience rapid productivity growth display a decline in their relative price. This result, combined with the consumer's first order conditions implies that the ratio of value added $p_{it}y_{it}$ in any two industries in the same sector s depends on parameters and the productivity terms.

$$\frac{p_{it}y_{it}}{p_{jt}y_{jt}} = \left(\frac{\xi_{s,i}}{\xi_{s,j}}\right)^{\varepsilon_s} \left(\frac{A_{it}}{A_{jt}}\right)^{\varepsilon_s - 1} = \frac{n_{it}}{n_{jt}} \qquad \forall s \tag{25}$$

Notice that the same relationship holds for the ratio of employment – just as with the basic model – except that it only holds comparing industries that are in the same sector.

Define the growth factor of employment (or value added) in industry i as

$$G_{it} \equiv \frac{n_{i,t+1}}{n_{i,t}} = \frac{p_{i,t+1}y_{i,t+1}}{p_{it}y_{it}}. \tag{26}$$

Then, the expression G_{it}/G_{jt} then denotes the growth of employment (or value added) in industry i relative to industry j. Using (25) we have that

$$\frac{G_{it}}{G_{j,t}} = \left(\frac{g_i}{g_j}\right)^{\varepsilon_s - 1} \qquad \forall s. \tag{27}$$

Consequently, within sectors, structural change depends on relative TFP growth factors $\frac{g_i}{g_j}$ and on the elasticity of substitution ε_s. For comparing industries *across* sectors requires characterizing shifts in expenditure across sectors, as well as investment behavior.

D.3 Sectorial and Aggregate Growth

Notice that in equilibrium we can aggregate the industries in a given sector into a sectorial production function. To see this, define q_{st} as the price index for final goods in sector s, so that

$$q_{st}y_{st} = \sum_{i \in I_S} p_{it}A_{it}k_t^{\alpha}n_{it}$$

where k_t is the equilibrium capital-labor ratio, which is common across industries. Define input use in sector s as $K_{st} = \sum_{i \in I_s} K_{it}$ and $n_{st} = \sum_{i \in I_s} n_{it}$. Then, define a sectorial production function:

$$y_{st} = A_{st}K_{st}^{\alpha}n_{st}^{1-\alpha}, \ A_{st} = A_{s0}\bar{g}_s^t \tag{28}$$

where $\bar{g}_{st} = A_{s,t+1}/A_{st}$.

The problem of the sector s firms and the industry $i \in I_s$ firms can be combined as

$$\max_{n_{it}} \left\{ q_{st} \left[\sum_{i \in I_s} \xi_i \times (A_{it} k_t^{\alpha} n_{it})^{\frac{\varepsilon_s - 1}{\varepsilon_s}} \right]^{\frac{\varepsilon_s}{\varepsilon_s - 1}} - r_t k_t \sum n_{it} - w_t \sum n_{it} \right\} \tag{29}$$

The first order conditions imply that:

$$\frac{n_{jt}}{n_{it}} = \left(\frac{\xi_j}{\xi_i} \right)^{\varepsilon_s} \left(\frac{A_{it}}{A_{jt}} \right)^{1-\varepsilon_s} \tag{30}$$

We also have that $\sum_i n_i = n_s$ by definition, so we can use (30) write n_i in terms of n_s. Substituting this back into problem (29), we have that a sector s firm solves the problem

$$\max_{n_{it}} \{ q_{st} A_{st} k_t^{\alpha} n_s - r k n_{st} - w n_{st} \}$$

where

$$A_{st} = \left[\sum_{i \in I} \xi_{s,i}^{\varepsilon_s} \times A_{it}^{\varepsilon_s - 1} \right]^{\frac{1}{\varepsilon_s - 1}} = \left[\sum_{i \in I} \xi_{s,i}^{\varepsilon_s} \times A_{i0}^{\varepsilon_s - 1} g_i^{t(\varepsilon_s - 1)} \right]^{\frac{1}{\varepsilon_s - 1}}. \tag{31}$$

Recalling that $\bar{g}_{st} = A_{s,t+1}/A_{st}$, we have that

$$\bar{g}_{st} = \prod_{i \in I_s} g_i^{x_{it}/X_{st}} \tag{32}$$

where

$$x_{it} = \xi_{s,i}^{\varepsilon_s} A_{it}^{\varepsilon_s - 1},\ X_{st} = \sum_{i \in I_s} x_{it}.$$

Since the total production of consumption sectors $c_t = \left[\sum_{s=1}^{S-1} \zeta_s y_{st}^{\frac{\varepsilon - 1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon - 1}}$, we can also aggregate all the consumption goods production sectors in the same fashion. Then we have that

$$c_t = A_{c_t} K_{ct}^{\alpha} n_{ct}^{1-\alpha},\ A_{c_t} = \left[\sum_{s=1}^{S-1} \zeta_s^{\varepsilon} \times A_{st}^{\varepsilon - 1} \right]^{\frac{1}{\varepsilon - 1}} \tag{33}$$

As a result, the aggregate behavior of the model economy with many sectors is the same as that of a 2-sector economy that produces c_t using technology (33) and produces capital goods using technology (28). In the consumption goods sector, firms maximize

$$\max_{K_{c_t}, n_{c_t}} \left\{ p_{ct} A_{ct} K_{ct}^{\alpha} n_{ct}^{1-\alpha} - r_t K_{ct} - w_t n_{ct} \right\}$$

$$\text{where} \qquad A_{c_t} = \left[\sum_{s=1}^{S-1} \zeta_s^{\varepsilon} \times A_{st}^{\varepsilon - 1} \right]^{\frac{1}{\varepsilon - 1}}$$

whereas in the capital goods sector:

$$\max_{K_{h_t}, n_{h_t}} \left\{ p_{St} A_{St} K_{St}^{\alpha} n_{St}^{1-\alpha} - r_t K_{St} - w_t n_{St} \right\}$$

$$\text{where} \qquad A_{St} = \left[\sum_{i \in I_S} \xi_i^{\varepsilon_S} \times A_{it}^{\varepsilon_S - 1} \right]^{\frac{1}{\varepsilon_S - 1}}$$

Consumers choose consumption c_t and investment y_{St} to solve:

$$\max_{c_t, h_t} \left\{ \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\theta} - 1}{1-\theta} \right\} \tag{34}$$

$$s.t. \qquad p_{c_t} c_t + p_{S_t} y_{St} \leq r_t K_t + w_t \tag{35}$$

$$K_{t+1} = K_t (1-\delta) + y_{St} \tag{36}$$

$$K_0 \text{ given.} \tag{37}$$

In equilibrium, capital and labor markets must clear at all dates, so

$$c_t = A_{c_t} K_{c_t}^{\alpha} n_{c_t}^{1-\alpha} \tag{38}$$

$$y_{St} = A_{S_t} K_{S_t}^{\alpha} n_{S_t}^{1-\alpha}$$

$$K_t = K_{S_t} + K_{c_t} \tag{39}$$

$$n_{c_t} + n_{S_t} = 1 \tag{40}$$

It will be convenient to set $p_{S,t} = 1 \forall t$, so that consumption goods prices p_{c_t} are expressed relative price to the price of capital goods.

Solving the 2-sector problem and using the equilibrium conditions, we obtain expressions for labor shares in the capital goods sector n_{St} and the consumption goods' sector $n_{ct} = 1 - n_{St}$ along an unbalanced growth path. These turn out to be functions only of the productivity growth rates g_i, parameters, and of the equilibrium growth rate of aggregate consumption $g_{c_t} = \frac{p_{c,t+1} c_{t+1}}{p_{ct} c_t}$ which is endogenous. This will be true at all dates except possibly date zero, where n_{St} is determined by the initial condition K_0.

Define real GDP as $y_t = y_{St} + p_{ct} c_t$. Notice it is measured in units of capital.

Proposition 1 *Equilibrium exists and is unique. In equilibrium, the growth factors of total capital K, capital per capita k, and total output y depend on the growth factors of TFP in the consumption and capital sectors and on the growth factor of consumption sector (as well as parameters):*

$$g_{k_t} = \frac{k_{t+1}}{k_t} = \frac{K_{t+1}}{K_t} = \bar{g}_{A_{St}}^{\frac{1}{1-\alpha}} \left(\frac{r_t}{r_{t+1}} \right)^{\frac{1}{1-\alpha}} \tag{41}$$

and

$$g_{yt} = \frac{y_{t+1}}{y_t} = \bar{g}_{A_{St}}^{\frac{1}{1-\alpha}} \left(\frac{r_t}{r_{t+1}} \right)^{\frac{\alpha}{1-\alpha}} \tag{42}$$

where GDP is defined as $y_t = y_{St} + p_{ct} c_t$ *and the equilibrium interest rate is* $r_t = \frac{\left(\frac{\bar{g}_{A_{St-1}}}{\bar{g}_{A_{ct-1}}}\right)^{1-\theta} g_{c_{t-1}}^{\theta}}{\beta} - 1 + \delta$ *for* $t > 0$*. At date zero,* r_0 *is determined by market clearing given* K_0*.*

Proposition 2 *The model economy converges to a balanced growth path where in each sector*

$$\lim_{t\to\infty} A_{st} = A_{jt} \text{ where } j = \begin{cases} \arg\max_{i\in I_s} \{A_i\} & \text{if } \varepsilon_s > 1 \\ \arg\min_{i\in I_s} \{A_i\} & \text{if } \varepsilon_s < 1 \end{cases},$$

and

$$\lim_{t\to\infty} A_{ct} = A_{st} \text{ where } s = \begin{cases} \arg\max_{s<S} \{A_s\} & \text{if } \varepsilon > 1 \\ \arg\min_{s<S} \{A_s\} & \text{if } \varepsilon < 1 \end{cases}.$$

Recalling that the only endogenous variable that affects r_t for $t > 0$ is g_{ct},[34] Proposition 1 implies that we can compute the equilibrium for the multi-industry model economy in transition, provided we can derive the series for g_{ct}. The economy with many consumption goods sectors will asymptotically converge to an economy with one consumption sector which has either the highest or lowest TFP growth rate depending on the elasticity of substitution. The same occurs within the capital goods sector. As a result, the expression r_t converges to some constant r and, although in general the model does not possess a balanced growth path (see Ngai and Pissarides (2007)), it converges to one. This suggests that the equilibrium may be computed by finding a sufficiently good approximation to the series for g_{ct}. In the limit, since by assumption $\varepsilon_s \neq 1$ for all $s \leq S$, one industry will end up dominating each sector. However, we wish to study the behavior of the model economy in transition, where sectors are relatively diversified.

E UNIDO Calibration: Manufacturing

In the remainder of the paper we will focus on a particular type of equilibrium. Observe that the capital stock will be set to satisfy the Euler equation (41) at all dates except date zero. In other words, the investment share of the model economy will in general be smooth over time, except between dates zero and one. The model will be calibrated to the available data and, since the initial year in which data for a given country become available has no economic

[34] In general, at $t = 0$, the value of r_0 is determined by market clearing and the value of K_0.

content, it is difficult to justify why the first year we have data for (generally 1963) happens to be the only date when the intertemporal optimization condition (41) is not satisfied. For this reason, we focus on an equilibrium where this does not occur.

Definition 3 *An Euler Growth Path (EGP) is an equilibrium and an initial condition* K_0 *such that equation* (41) *holds at date zero.*

The Euler growth path is a generalization of a balanced growth path which exists in models that do not exhibit balanced growth. For our benchmark results, we calibrate the model to match an Euler growth path by matching the composition of manufacturing but not necessarily its *size*. Details are in Section G. Nonetheless, it is important to underline that our results concerning the structure of the economy turn out not to hinge on whether we focus on an Euler growth path: results on the equilibrium calibrated to match the initial conditions in the data are qualitatively indistinguishable.

Calibrating the model economy requires a choice of industries, and values of the following parameters and variables.

1. Technological parameters α, δ.
2. Preference coefficients $\xi_{s,i}, \zeta_i, \beta$.
3. Elasticities of substitution ε_s for $s \leq S$, and ε, the elasticity across consumption sectors
4. The intertemporal elasticity parameter θ.
5. Productivity growth values g_i.
6. Productivity initial conditions $A_{i0.}$

We provide two selections of industries. In this section, we calibrate the model so as to focus on the "stages" in manufacturing in the UNIDO data. In the following section, we disaggregate the non-manufacturing sectors further to focus on the "stages" across broad sectors in the ILO data. We calibrate the model twice because the data required for the ILO sector calibration are available for fewer countries. We refer to them as the UNIDO or manufacturing calibration and the ILO or sector calibration. Details are provided below. The simulation requires computing transition dynamics in a model without a balanced growth path, and the procedure is described in Section G.

Table 3: Sectors and Industries in the model economy

	Industries			
Sector	Agriculture	Services, etc	ISIC Manuf	Construction
Agriculture	X	-	-	-
Services, etc	-	X	-	-
Manufacturing (not capital)	-	-	X	-
Manufacturing (capital)	-	-	X	X

For the UNIDO calibration, we group all industries into four sectors: Agriculture, Services, Capital and Non-capital manufacturing. Agriculture, services and non-capital manufacturing sectors produce consumption goods, and the capital sector only produces capital goods. Industries include agriculture, services, the 28 UNIDO manufacturing industries, and construction (see Table 3). Thus, the agriculture and services industries only contain one industry. The UNIDO industries serve either the capital or the non-capital manufacturing sectors. We assigned an industry to the capital sector if the US NIPA tables count it in their "fixed asset" tables (see Table 4). Construction serves the capital sector too. The initial shares of agriculture, services, manufacturing and construction sectors out of GDP are derived from World Development Indicators data (WDI).[35]

1. We assume that $\delta = 0.06$ as in Greenwood et al (1997): this is a standard values in models in which the productivity of the investment technology exceeds that in the consumption sector. We use a standard value for the capital share, $\alpha = 0.3$.

2. To calibrate the utility weights $\xi_{s,i}$, it should be noted that in a sense these weights are arbitrary, as they depend on the exact unit of measurement for good i.[36] Thus, without loss of generality, we set $\xi_{s,i} = \frac{1}{I_s}$, where I_s is the number of industries in sector s. The same applies to ζ_i, the utility weight at the sector level, so $\zeta_i = \frac{1}{S-1}$. We set $\beta = 0.95$, a standard value.

[35] For countries with missing data, we use predicted values computed by regressing sector shares on income, income squared and UNIDO industry shares in the manufacturing sector for all countries and years in our sample.

[36] For example, if I measure apples and get $\xi_{s,apples} = 2$ (and $A_{apples,0} = 3$), I could choose to measure apples in units of "half an apple" and then $\xi_{s,apples} = 1$ (and $A_{apples,0} = 6$).

Table 4: Capital good-producing manufacturing industries

Industry	ISIC code
Wood products	331
Furniture, except metal	332
Fabricated metal products	381
Machinery, except electrical	382
Machinery, electric	383
Transport equipment	384
Prof. & sci. equip.	385
Other manufactured prod.	390

3. For each sector, equation (27) is equivalent to $\log G_i = \alpha + (\varepsilon - 1) \log g_i + \epsilon_i$ where $\alpha = \log G_j - \log g_j$ for some arbitrary industry j and ϵ_i is any unmodeled noise in the relationship. We regress U.S. value added growth rates on TFP growth rates[37] for capital and non-capital manufacturing goods respectively, finding that they were not statistically significantly different: $\varepsilon_{noncapmanuf} = \varepsilon_{capital}$. Pooling the data, we estimate that $\varepsilon_{noncapmanuf} = \varepsilon_{capital} = 3.73$. Across sectors, we use the value $\varepsilon = 0$, as discussed earlier.

4. The preference parameter θ is calibrated so that in the long run the investment share of GDP converges to 12 percent, which is roughly the share in the US: investment shares in transition turn out not to be very different. This implies that $\theta = 3$: typical values used in calibration fall in the range $\theta \in [1, 5]$,[38] so it is encouraging that our value falls in the middle.

5. Productivity growth values g_i are drawn from the NBER productivity database, as described in Section 2. We use the average value over the period 1963-1992. See

[37] We use data from Jorgenson et al (2007): although they are a little more disaggregated, we want a value estimated at roughly the same level of aggregation as the UNIDO data. The UNIDO data themselves are too few so we were unable to obtain a good estimate from them directly. The estimate of ε is slightly different from Ilyina and Samaniego (2012) because of the inclusion of Construction in the set of capital goods producing industries.

[38] Growth models tend to use $\theta = 1$, whereas asset pricing studies tend to use larger values, see for example Jermann (1998).

Table 1. To calibrate the growth factors of the consumption goods sectors, first we use equation (31) to compute TFP growth in the capital sector (excluding construction) over the period 1963-1992, and get the average value $g_S = 1.0241$. According to NIPA, the relative price of construction has risen at a rate of 0.0109 each year relative to other capital. This means the growth factor of construction sector $g_{construction} = g_S/e^{0.0109} = 1.0130$. For the services sector, the relative price of services has risen at a rate of 0.0103 each year relative to other capital. This means the growth factor of services sector $g_{services}$ is then $g_S/e^{0.0103} = 1.0136$. For agriculture, we have that the relative price of agriculture has dropped at a rate of 0.004 each year relative to other capital. So the growth factor of the agricultural sector is $g_{agriculture} = g_S/e^{-0.004} = 1.0282$.

6. For the initial productivities of the capital and consumption sectors, we initially set $A_{capital,0} = 1$ and $A_{consumption,0} = 1$. The former is a normalization, and the latter is without loss of generality because the size of the non-investment sectors is independent of the level of $A_{consumption,0} = 1$.[39] Then, using (25) and (31), for the capital sector industries $i \in I_S$, we set initial TFP to equal $A_{i0} = \left[\frac{n_{i0}}{\sum \xi_i^{\varepsilon_S} n_{i0}}\right]^{\frac{1}{\varepsilon_S - 1}}$, thus matching the initial share of capital industries in each country. For the consumption sectors, set A_{s0} (where $s \in \{$services, agriculture and non-capital manufacturing$\}$) so as to match the initial share of that sector in each country: $A_{s0} = \left[\frac{n_{s0}}{\sum \zeta^{\varepsilon} n_{s0}}\right]^{\frac{1}{\varepsilon - 1}}$. Finally, for industry productivity in non-capital manufacturing, we have again that $A_{i0} = \left[\frac{n_{i0} A_{s0}}{\sum \xi_i^{\varepsilon_s} n_i}\right]^{\frac{1}{\varepsilon_s - 1}}$. Industry shares are drawn from UNIDO and sector shares are based on the WDI.[40] Finally, we multiply A_{i0} in all industries and sectors by a country-specific constant so that the country GDP per head relative to US GDP per head in the initial year is the same as in the data.

Finally, there are many country factors the literature has related to economic growth which are not featured in the model (see for example Barro (1991) or Sala-i-Martin (1997)). We add a country-specific productivity growth term that affects all industries equally, and calibrate it to match average GDP growth rates in each country over the sample period. This term could be interpreted as capturing policies that affect technological diffusion, trends in

[39] Proof available upon request. We could calibrate the ratio of these producivity terms to match the relative price of capital reported in the initial year in the Penn World Tables 7.1 but this does not affect results – as a result of the Proposition.

[40] As mentioned, an adjustment to industry shares is required due to our focus on an EGP: see Section G for details.

Table 5: Calibrated Parameters: Baseline Model

$g_{construction}$	$g_{services}$	$g_{agriculture}$	$\varepsilon_{capital}$	$\varepsilon_{consumption}$
1.0130	1.0136	1.0282	3.73	0

$\varepsilon_{noncapmanuf}$	θ	δ	α	β
3.73	3	0.06	0.3	0.95

policy, or any of the factors commonly included in growth regressions. In any case, results are not sensitive to the inclusion of this term.

E.1 Simulation

For each country we use initial conditions from 1963[41] as starting points, and simulate the share of GDP n_{it} of any industry or sector along unbalanced growth path.[42]

Figure 15 shows the estimated curve of Gini using industry shares simulated in our baseline model. We can see that our baseline model is able to capture the U-shape of stages of diversification very well. Thus, the results derived using the simple model are robust to allowing the composition and size of the capital sector to evolve independently of the non-capital manufacturing sector, and to allowing the model to generate the GDP series as well as just industry structure. It is notable that, in the full growth model, the re-specialization after the turning point is slower than the initial specialization, just as in IW.[43]

[41] For some countries initial data in 1963 are not available: then we use the earliest available year.

[42] In our derivations, the model measures GDP in terms of capital goods (remember we normalize capital goods price to 1 and consumption goods prices are expressed as relative to capital goods price). In the data, however, GDP is measured in terms of consumption goods, see Greenwood et al (1997). Since in our model, $A_{h_t} = p_{c_t} A_{c_t}$, we can express the GDP growth factor measured in units of *consumption* using the formula $\widetilde{g}_{yt} = g_{yt} \frac{g_{Ah}}{g_{Ac}}$. This is the notion of GDP we use in the graphs below. The model simulated using GDP defined in terms of capital goods yields very similar results. An issue here is that the values of A_{h_0} and A_{c_0} are arbitrary in the calibration, but not when we wish to express cross-country GDP in common units. We handle this by assuming that the real GDP data are measured in units of consumption and are internationally comparable, and then use the model to compute growth rates (which do not depend on GDP levels) extrapolating from initial GDP in the data.

[43] If we extend the simulated curve from \$15,000 to \$20,000, the rising right hand side of the curve continues to increase linearly.

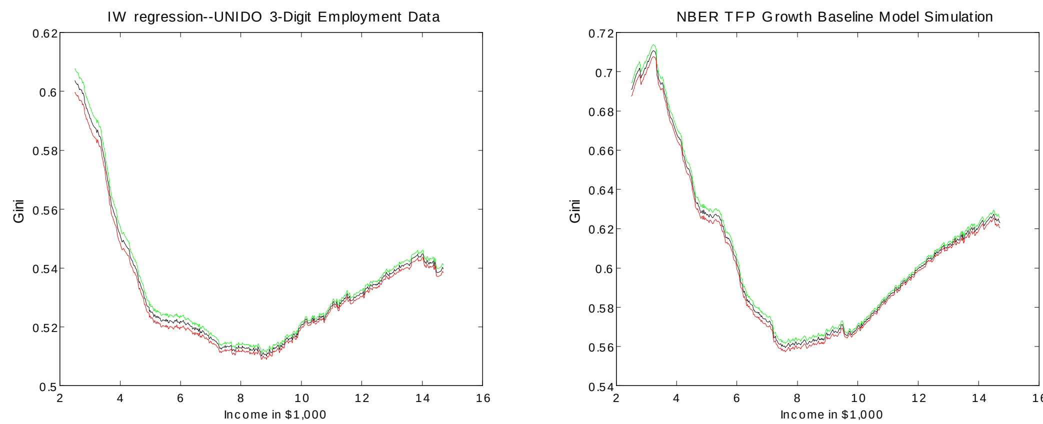

Figure 15. Industry structure along the development path in the full model. The left panel is the relationship between income and specialization within manufacturing reported in IW. The right panel is the same relationship in the pseudo-data generated from the model economy. The range of income is the same as that reported in IW.

F ILO Calibration: Broad sectors

IW report how economic structure evolves along the development path for 9 broad sectors, drawing on data from the ILO. In this section, we will show that TFP growth differences in our model can explain economic structure through the economy at sector level. We recalibrate the model economy so that $S = 10$, where the capital and non-capital manufacturing sectors are the same as in the UNIDO calibration, whereas the other 8 sectors correspond to the non-manufacturing sectors in the ILO 1-digit data. Thus, whereas before we had capital, non-capital manufacturing, agriculture and services, we now disaggregate services into several new sectors. See Table 2 for the list of non-manufacturing sectors. Within manufacturing, we still use 28 UNIDO industries, of which 8 produce capital goods as the UNIDO calibration (Table 4). The definition of the capital goods sector is the same as the UNIDO calibration, i.e. 8 UNIDO industries and construction. All other sectors in Table 2 produce consumption goods. The initial sector shares are taken from the ILO dataset.[44] As earlier, we set $\varepsilon = 0$. All other parameters are calibrated as before.

Figure 16 shows the estimated link between income and sectoral concentration using the

[44] Again, an adjustment is required due to the focus on an EGP. See Appendix.

9 ILO sector shares simulated in our model. It is clear that economic structure at sector level still displays a U-shape. Again, the turning point is around $9,000 as in the data.

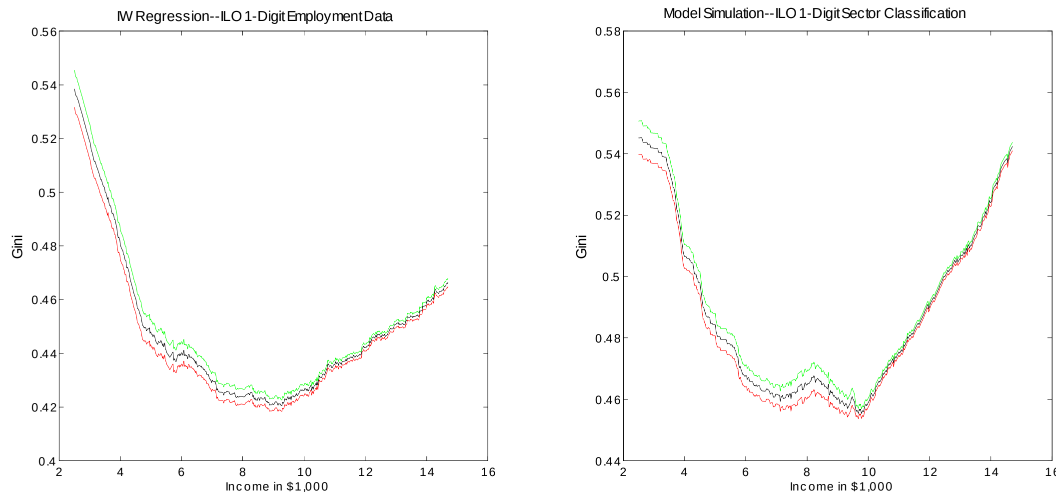

Figure 16. Economic Structure along the Development path: results for the entire economy, using the ILO 1-Digit Sector Classification. The left panel represents the ILO employment data, whereas the right panel is based on model-generated shares.

Within the manufacturing sector (for both capital and non-capital manufacturing industries), resources will shift towards high-TFP growth industries, as calibrated elasticities of substitution are above one. However across *sectors*, because the elasticity of substitution across the consumption sectors is *less* than unity, the model predicts that the economy will shift resources towards the *slowest* TFP growth sector. During this process of structural change, the economy displays a U-shaped pattern of stages of diversification at both levels of disaggregation. This is exactly what we observe in the data in Figure 2 in the main text.

A lot of attention has been devoted to explaining changes in the share of agriculture and services along the development path – see Ngai and Pissarides (2004, 2007), Rogerson (2008) and Duarte and Restuccia (2010), among others. Although not designed to do so, the model matches extremely well the observed changes in the link between agricultural shares and service shares and income. See Figure 17.[45]

[45] The model does not produce a hump shape in the share manufacturing, which some authors have focused on (e.g. Buera and Kaboski (2012)). Pooling lots of countries and using the IW method, the data do seem to display a hump but it is weak compared to the dramatic rise of services and fall in agriculture, which the model does reproduce. Duarte and Restuccia (2010) show that a model with our basic mechanisms plus non-homothetic preferences can reproduce the hump.

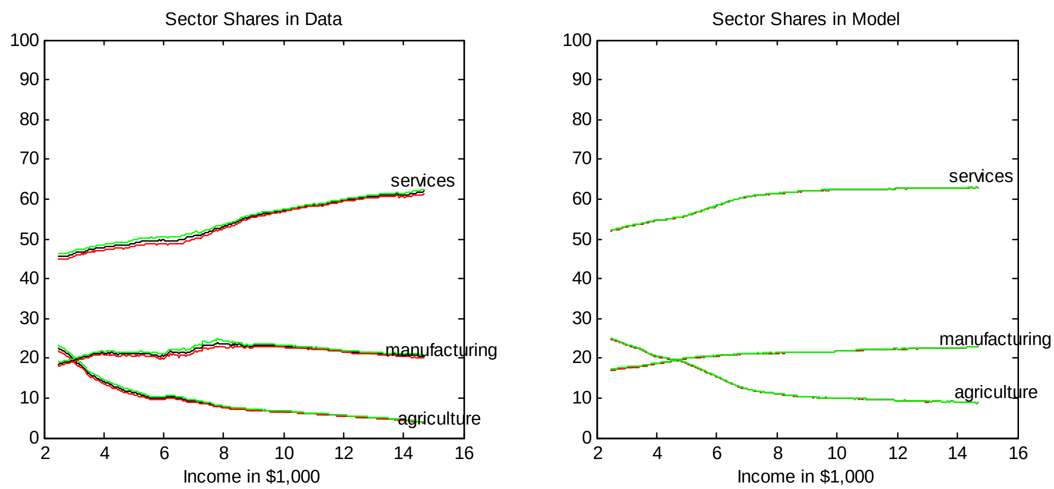

Figure 17. IW regressions for shares of agriculture, manufacturing and services. The left panel is based on the ILO data. The right panel is based on the model-generated shares.

G Simulation procedure

Simulating the model requires overcoming two distinct problems.

The first concerns matching the model with the data. Notice that the model reduces to a 2 sector model where consumption and investment are made by different sectors. As shown in Greenwood, Hercowitz and Krusell (1997), this is the same as a one-sector model with investment specific technical change. In the one-sector growth model, the equilibrium for any initial conditions is a jump to the stable branch of a saddle path that leads to the long run equilibrium (which in this case is the model where the capital sector has converged to contain only one industry). Thus, for general initial conditions K_0, the share of investment will jump after period 1, so that the structure of the manufacturing sector will change abruptly after period zero (and smoothly thereafter).

We handle this problem in two ways. First, we computed everything without worrying about the jump. Second, we calibrated the model so as to focus on an Euler growth path – which are the results reported in the paper (results were very similar either way).

In this second case we did not set the initial value of the capital stock K_0 to match the investment share of GDP in each country. The reason was that, in all other periods after $t = 0$, the investment share will follow the Euler equation. It seems arbitrary to assume that in all countries the Euler equation is satisfied in all years except 1963, or whatever happens to be the year for which data are initially available. As a result, we assume that the Euler equation is also satisfied at date zero. We call this an "Euler growth path" or EGP. To do this requires setting the investment share of GDP at a value that is different from that in the data. At the same time, it is critical that we preserve the composition of manufacturing. Hence, we adopted a recursive strategy. We know from the data the composition of investment in year zero. Given an assumption on the investment to GDP ratio, we can preserve the ratio of capital to manufacturing and find a value for the size of manufacturing that preserves its composition.[46] Then we check whether the assumption on the investment to GDP ratio corresponds to an EGP.[47] If not, then we generate another guess based on the predicted EGP value from the last iteration. We find that 3 loops is sufficient for very tight convergence. When we regress data on initial manufacturing shares on the model initial manufacturing shares, we find a coefficient of 1.16 (positive and close to one) and an intercept of -0.026 (close to zero), both significant at the 1% level. We take this to imply that, in general, our procedure does not significantly distort the sector structure of the model economy.

The second computational issue we confront is the fact that we are simulating a model economy that does not have a balanced growth path (although it converges to one). Recall that the aggregate behavior of the model is the same as a one-sector model with investment specific technical change. In the one-sector growth model, any approximation to the saddle path will "shadow" it for a period of time, eventually diverging infinitely from it: see Colucci (2001). As a result, we adopt a procedure to provide this "shadowing" without suffering an eventual divergence.

The procedure is to assume limited computational ability among the agents, a procedure we call "rolling windows of consciousness." Specifically, the structure of the model economy can be computed exactly given the investment share of employment. This can be computed exactly given a series for g_{ct}, which is determined by Euler equation (65) and the transversality condition. The Euler equation converges uniformly to $g_{ct} = \overline{g}_{A_{St}}^{\frac{1}{1-\alpha}}$, where $\overline{g}_{A_{St}}$ is the weighted-average productivity in the manufacturing industry and which is known for any t given initial conditions and using equation (6). We assume that an agent at date t acts as though difference equation (65) characterizes g_{ct} up to period $t+T$, whereas after $t+T$ the agent believes that $g_{ct} = \overline{g}_{A_{St}}^{\frac{1}{1-\alpha}}$. Notice that this is distinct from simulating the transition

[46] Other sectors are resized so that, relative to each other, shares of GDP are preserved too.

[47] Recall that computing the equilibrium, including the initial share of investment, requires a series for g_c, which in turn depends on sector productivity growth rates. However, sector productivity growth rates depend on the initial composition and size of the economy. This is why an iterative procedure is necessary to find an EGP.

path of the economy assuming that there is some fixed date T in the future after which productivity dynamics in the model economy change and correspond to a BGP:[48] we found that procedure to be sensitive to the value of T. Instead, the horizon over which the agent uses the exact Euler equation (6) does not change over time.

We tried $T = 50$, 90 and 200. For $T = 90$, the error between the realized value of g_{c1} and the value forecast by the agent when making investment and consumption decisions in period 0 is about 1% of the actual value (because the series for g_{ct} converges uniformly to its long run value, the forecast errors are the highest in the first period). For $T = 200$ these values are indistinguishable to eight decimal places. At the same time, for all these values of T, the Gini nonparametric regression results such as Figure 15 were indistinguishable regardless of the value of T.

This indicates two results. First, this procedure could yield an arbitrarily accurate approximation to the correct aggregate equilibrium dynamics, given a sufficiently large (but finite) value of T. This is distinct from the shadowing property, which provides arbitrarily precise approximations only for a finite period, after which there is increasing divergence. Second, industry dynamics are robust even to using values of T such that aggregate dynamics are computed with some degree of imprecision.

[48] See for example He and Liu (2008).

H Proofs

Proof of decentralized economy. For consumers:

$$\max_{y_{st}} \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\theta} - 1}{1-\theta}$$

$$c_t = \left[\sum_{s=1}^{S-1} \zeta_s y_{st}^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}$$

$$s.t. \quad \sum_{s=1}^{S-1} q_{st} y_{st} + q_{St} y_{St} = \sum_{s=1}^{S} \sum_{i \in I_s} r_t K_{it} + \sum_{s=1}^{S} \sum_{i \in I_s} w_t n_{it}$$

Capital and labor market clearing conditions are:

$$K_t = \sum_{s=1}^{S} \sum_{i \in I_s} K_{it}$$

$$1 = \sum_{s=1}^{S} \sum_{i \in I_s} n_{it}$$

F.O.C w.r.t y_{st} :

$$\frac{q_{st}}{q_{s't}} = \left(\frac{y_{s't}}{y_{st}} \right)^{\frac{1}{\varepsilon}} \frac{\zeta_s}{\zeta_{s'}} \qquad s, s' = 1, ..., S-1 \tag{43}$$

or

$$\frac{y_{st}}{y_{s't}} = \left(\frac{\zeta_s}{\zeta_{s'}} \frac{p_{s't}}{p_{st}} \right)^{\varepsilon} \qquad s, s' = 1, ..., S-1 \tag{44}$$

Final Goods Sector s maximizes profit:

$$\max_{u_{s,i,t}} q_{st} y_{st} - \sum_{i \in I} p_{it} u_{s,i,t}$$

$$= q_{st} \left[\sum_{i \in I} \xi_{s,i} \times y_{i,t}^{\frac{\varepsilon_s - 1}{\varepsilon_s}} \right]^{\frac{\varepsilon_s}{\varepsilon_s - 1}} - \sum_{i \in I} p_{it} y_{i,t}$$

F.O.C w.r.t $y_{i,t}$:

$$q_{st} \left[\sum_{i \in I} \xi_{s,i} \times y_{i,t}^{\frac{\varepsilon_s - 1}{\varepsilon_s}} \right] \xi_{s,i} y_{i,t}^{\frac{-1}{\varepsilon_s}} = p_{it}$$

similarly for $y_{j,t}$:

$$q_{st} \left[\sum_{i \in I} \xi_{s,i} \times y_{i,t}^{\frac{\varepsilon_s - 1}{\varepsilon_s}} \right] \xi_{s,j} y_{j,t}^{\frac{-1}{\varepsilon_s}} = p_{jt}$$

So we have:

$$\frac{p_{it}}{p_{jt}} = \left(\frac{y_{j,t}}{y_{i,t}} \right)^{\frac{1}{\varepsilon_s}} \frac{\xi_{s,i}}{\xi_{s,j}} \tag{45}$$

or

$$\frac{y_{i,t}}{y_{j,t}} = \left(\frac{\xi_{s,i}}{\xi_{s,j}}\frac{p_{jt}}{p_{it}}\right)^{\varepsilon_s} \tag{46}$$

For industry i in a given sector:

$$\max p_{it}A_{it}K_{it}^{\alpha}n_{it}^{1-\alpha} - r_tK_{it} - w_tn_{it}$$

F.O.C w.r.t K_{it} :

$$p_{it}\alpha A_{it}K_{it}^{\alpha-1}n_{it}^{1-\alpha} = r_t \tag{47}$$

F.O.C w.r.t n_{it} :

$$p_{it}(1-\alpha)A_{it}K_{it}^{\alpha}n_{it}^{-\alpha} = w_t \tag{48}$$

Dividing one F.O.C. by the other we get that

$$\frac{1-\alpha}{\alpha}\left(\frac{K_{it}}{n_{it}}\right) = \frac{w_t}{r_t} \Rightarrow k_t = \frac{w_t}{r_t} \times \frac{\alpha}{1-\alpha} \tag{49}$$

where the capital labor ratio $k_t \equiv K_{it}/n_{it}$ is a constant across industries. Applying this result to (47) implies that

$$\frac{A_{it}}{A_{jt}} = \frac{p_{jt}}{p_{it}} \tag{50}$$

Using (46), (49) and (50) yields

$$\frac{n_{it}}{n_{jt}} = \left(\frac{\xi_{s,i}}{\xi_{s,j}}\right)^{\varepsilon_s}\left(\frac{A_{it}}{A_{jt}}\right)^{\varepsilon_s-1} \tag{51}$$

which, rearranging (45), implies that $\frac{n_{it}}{n_{jt}} = \frac{p_{it}y_{it}}{p_{jt}y_{jt}}$. Define the industry i growth factor as :

$$G_{it} = \frac{p_{i,t+1}y_{i,t+1}}{p_{it}y_{it}}$$

and the expression G_{it}/G_{jt} then denotes the growth of industry i relative to industry j

$$\begin{aligned} \frac{G_{it}}{G_{j,t}} &= \frac{\frac{p_{i,t+1}y_{i,t+1}}{p_{it}y_{it}}}{\frac{p_{j,t+1}y_{j,t+1}}{p_{jt}y_{jt}}} = \frac{\frac{p_{it+1}}{p_{jt+1}}\left(\frac{\xi_{s,i}}{\xi_{s,j}}\frac{p_{jt+1}}{p_{it+1}}\right)^{\varepsilon_s}}{\frac{p_{it}}{p_{jt}}\left(\frac{\xi_{s,i}}{\xi_{s,j}}\frac{p_{jt}}{p_{it}}\right)^{\varepsilon_s}} \\ &= \frac{\left(\frac{p_{it+1}}{p_{jt+1}}\right)^{1-\varepsilon_s}}{\left(\frac{p_{it}}{p_{jt}}\right)^{1-\varepsilon_s}} = \frac{\left(\frac{A_{it+1}}{A_{it}}\right)^{\varepsilon_s-1}}{\left(\frac{A_{jt+1}}{A_{jt}}\right)^{\varepsilon_s-1}} \\ &= \left(\frac{g_i}{g_j}\right)^{\varepsilon_s-1} \end{aligned}$$

■

Proof of Proposition 1. Solving the 2 sector problem and using the equilibrium conditions, we have:

$$A_{St} = p_{ct} A_{ct} \tag{52}$$

$$\begin{aligned} r_t &= \frac{\frac{p_{c_t} c_t^{\theta}}{p_{c_{t-1}} c_{t-1}^{\theta}}}{\beta} - 1 + \delta = \frac{\left(\frac{\overline{g}_{A_{S,t-1}}}{\overline{g}_{A_{c,t-1}}}\right)^{1-\theta} g_{c_{t-1}}^{\theta}}{\beta} - 1 + \delta \text{ IF } \beta \neq 0 & (53)\\ \text{where } g_{ct-1} &\equiv \frac{p_t c_t}{p_{t-1} c_{t-1}} \text{ is the growth factor of aggregate consumption} & (54)\\ \overline{g}_{A_{ct-1}} &= \frac{A_{ct}}{A_{c,t-1}}, \overline{g}_{A_{S,t-1}} = \frac{A_{St}}{A_{S,t-1}} \text{ are known} & (55)\end{aligned}$$

Let $\phi_t = \alpha^{-1} r_t^{\frac{-\alpha}{1-\alpha}} - \overline{g}_{A_{S,t}}^{\frac{1}{1-\alpha}} r_{t+1}^{\frac{-1}{1-\alpha}} + (1-\delta) r_t^{\frac{-1}{1-\alpha}}$

$$\begin{aligned} k_t &= \frac{K_{S_t}}{n_{S_t}} = \frac{K_{c_t}}{n_{c_t}} = \left(\frac{\alpha A_{S_t}}{r_t}\right)^{\frac{1}{1-\alpha}} \\ K_t &= k_t \end{aligned}$$

The growth factor of capital per capita in each sector is:

$$g_{k_t} = \frac{k_{t+1}}{k_t} = \overline{g}_{A_{St}}^{\frac{1}{1-\alpha}} \left(\frac{r_t}{r_{t+1}}\right)^{\frac{1}{1-\alpha}} \tag{56}$$

Similarly, we get aggregate capital growth factor:

$$g_{K_t} = g_{k_t}$$

Using (52) and (36), we derive capital sector output, i.e., investment:

$$\begin{aligned} y_{St} &= \left(\frac{\alpha A_{S_{t+1}}}{r_{t+1}}\right)^{\frac{1}{1-\alpha}} - (1-\delta)\left(\frac{\alpha A_{S_t}}{r_t}\right)^{\frac{1}{1-\alpha}} \\ &= (\alpha A_{S_t})^{\frac{1}{1-\alpha}} \left[\left(\frac{\overline{g}_{A_{S,t}}}{r_{t+1}}\right)^{\frac{1}{1-\alpha}} - (1-\delta)\left(\frac{1}{r_t}\right)^{\frac{1}{1-\alpha}}\right] & (57)\end{aligned}$$

and the growth factor of investment y_{St} is:

$$g_{y_{St}} = \frac{y_{S,t+1}}{y_{St}} = \overline{g}_{A_{St}}^{\frac{1}{1-\alpha}} \frac{\left(\frac{\overline{g}_{A_{S,t+1}}}{r_{t+2}}\right)^{\frac{1}{1-\alpha}} - (1-\delta)\left(\frac{1}{r_{t+1}}\right)^{\frac{1}{1-\alpha}}}{\left(\frac{\overline{g}_{A_{St}}}{r_{t+1}}\right)^{\frac{1}{1-\alpha}} - (1-\delta)\left(\frac{1}{r_t}\right)^{\frac{1}{1-\alpha}}}$$

so that the labor in capital sector is:

$$n_{S,t} = \alpha \left[\frac{1}{r_t}\left(\frac{\overline{g}_{A_{S,t}} r_t}{r_{t+1}}\right)^{\frac{1}{1-\alpha}} - \frac{(1-\delta)}{r_t}\right] \tag{58}$$

and the growth factor of n_{ht} is:

$$g_{n_{St}} = \frac{n_{S,t+1}}{n_{St}} = \frac{\frac{1}{r_{t+1}}\left(\frac{\overline{g}_{A_{S,t+1}} r_{t+1}}{r_{t+2}}\right)^{\frac{1}{1-\alpha}} - \frac{(1-\delta)}{r_{t+1}}}{\frac{1}{r_t}\left(\frac{\overline{g}_{A_{St}} r_t}{r_{t+1}}\right)^{\frac{1}{1-\alpha}} - \frac{(1-\delta)}{r_t}} \tag{59}$$

Notice that n_{St} (and hence $n_{ct} = 1 - n_{S,t}$) is independent of the level of technology in c and y_S as long as the interest rate is too. We can get capital in capital sector:

$$K_{S,t} = \alpha \left[\frac{1}{r_t}\left(\frac{\overline{g}_{A_{S,t}} r_t}{r_{t+1}}\right)^{\frac{1}{1-\alpha}} - \frac{(1-\delta)}{r_{t-1}}\right]\left(\frac{\alpha A_{S_t}}{r_t}\right)^{\frac{1}{1-\alpha}} \tag{60}$$

Define the aggregate output per capita as $y_t = y_{St} + p_{ct}c_{ct}$. Since $K_{ct} = K_t - K_{St}$ and $n_{ct} = 1 - n_{St}$,

$$\begin{aligned} y_t &= y_{St} + p_{ct}c_{ct} \\ &= A_{St}K_{St}^{\alpha}n_{St}^{1-\alpha} + p_{c_t}A_{c_t}K_{c_t}^{\alpha}n_{c_t}^{1-\alpha} \\ &= A_{S_t}k_t^{\frac{\alpha}{1-\alpha}} = \left(\frac{\alpha}{r_t}\right)^{\frac{\alpha}{1-\alpha}} A_{S_t}^{\frac{1}{1-\alpha}} \end{aligned} \tag{61}$$

and its growth factor is:

$$g_{yt} = \frac{y_{t+1}}{y_t} = \overline{g}_{A_{St}}^{\frac{1}{1-\alpha}}\left(\frac{r_t}{r_{t+1}}\right)^{\frac{\alpha}{1-\alpha}} \tag{62}$$

Aggregate consumption is:

$$C_t = p_{ct}c_t = y_t - y_{S,t} \tag{63}$$

$$\begin{aligned} &= \left(\frac{\alpha}{r_t}\right)^{\frac{\alpha}{1-\alpha}} A_{S,t}^{\frac{1}{1-\alpha}} \\ &\quad - (\alpha A_{S_t})^{\frac{1}{1-\alpha}}\left[\left(\frac{\overline{g}_{A_{St}}}{r_{t+1}}\right)^{\frac{1}{1-\alpha}} - (1-\delta)\left(\frac{1}{r_t}\right)^{\frac{1}{1-\alpha}}\right] \end{aligned} \tag{64}$$

The growth factor of consumption is:

$$g_{ct} = \frac{C_{t+1}}{C_t} = \overline{g}_{A_{St}}^{\frac{1}{1-\alpha}}\frac{\phi_{t+1}}{\phi_t}. \tag{65}$$

Notice that as $t \to \infty$ the expressions for $\overline{g}_{A_{St}}$ and $\overline{g}_{Act}$ converge to constants, so the difference equation for g_{ct} converges uniformly to that which characterizes the model of investment-specific technical change in Greenwood, Hercowitz and Krusell (1997). Thus, the result that the transversality condition picks out a single equilibrium solution in that model extends to our case too. ■

Proof of Proposition 2. Corollary of the proof of Proposition 1 and (27). ■