

From Agriculture to Services: Development without Industrialization

Tancredi Rapone*

UAB, BSE

June 13, 2025

Abstract

Is an industrial revolution necessary for sustained long-run growth? This paper builds a unified growth model of the development process featuring structural change, endogenous fertility and misallocation through firm level distortions in order to explore the implications of bypassing manufacturing in today's developing countries. The model is able to replicate several stylized facts including: (i) current GDP per capita is significantly related to a country's peak manufacturing employment share, (ii) fertility is higher (lower) in countries which have a larger (smaller) services share of non-farm labor, (iii) countries with lower peak manufacturing employment shares have lower levels of education. Through the lens of the model, 14% of current GDP per capita gaps are explained by countries having lower peak manufacturing employment shares. Cross-country income inequality is projected to increase over time as countries which bypass industrialization land on balanced growth paths with lower productivity growth.

Keywords: structural change, trade, unified growth.

JEL Classification: O14, O15, O41, F16, J23, O33, E24

*First and foremost, I am deeply grateful to my advisors, Raül Santaeulàlia-Llopis and André Gröger, for their continued support, patience, and guidance throughout this project. I also thank Gabriela Araujo, Sophie Brochet, Francisco Buera, Marcela Eslava, Alejandro Gobieno, Jacob Hess, Luca Looser, Martí Mestieri, Enoch Ntsiful, Luis Felipe Sáenz, Etienne Wasmer, and seminar participants at the BSE PhD Jamboree and the Bellaterra Macro Club for their valuable comments and suggestions. All errors are my own. For comments, suggestions, or critiques, you can reach me at tancredi.rapone@icloud.com.

1 Introduction

Developing countries today have larger shares of employment in the service sector and lower shares of employment in the manufacturing sector when compared to advanced economies at similar levels of GDP per capita, a phenomenon which has been termed "premature de-industrialization" in the literature (Rodrik, 2016). Although this fact has been known for almost a decade, we still have scarce evidence on (i) what causes countries to move out of industrial activities toward services and (ii) what are the implications of bypassing a manufacturing boom on long-term development. The latter point is especially salient given that very few advanced economies have achieved sustained growth without experiencing an industrialization boom at some point in their history (Kuznets, 1973; Chang, 2003; Murphy et al., 1989). In this paper, I build a unified growth model which captures the transition from stagnation to growth with a key role played by the manufacturing sector. Calibrating the model to the US economy over the past 200 years and simulating a transition from stagnation to growth with different peak employment shares in manufacturing reveals that both current *levels* of GDP per capita and *growth rates* are proportional to the peak employment share in manufacturing. The results are not only consistent with the cross country evidence, but also quantitatively significant: reducing the peak employment share of manufacturing by 50% reduces current GDP per capita by 10% and the growth rate of GDP per capita by 6.5%. The mechanism generating this result is straightforward: productivity growth in the services sector is bounded due to physical and technological constraints (Baumol, 1967; Vollrath, 2020). Hence, a transition out of agriculture which bypasses the manufacturing sector not only mechanically lowers the aggregate growth rate, by allocating more resources to the least productive sector, but also lowers the return to investments in education, which further lowers the productivity growth rate of all sectors. Integrating structural transformation to the benchmark unified growth model of Galor and Weil (2000) adds several testable predictions which can be used to discipline the model's parameters, including the relationship between structural change within the non-agricultural sector (movement of labor from manufacturing to services and vice-versa) and total fertility rates. Using both microdata and cross-country aggregates, I show that fertility rates and

structural change are related in a way which, to the best of my knowledge, has not been previously documented in the literature. More specifically, countries with higher shares of manufacturing employment within the non-farm sector have lower fertility, a relationship which holds at all levels of development and within almost all countries in the IPUMS microdata when comparing women in different sectors and controlling for a wide array of demographic characteristics. The model is able to rationalize this given that a lower aggregate growth rate relaxes the quantity-quality tradeoff by lowering the value of educational investment, a feature present in all Unified Growth Models since Galor and Weil (2000).¹ Further counterfactual exercises will explore the dynamic effects of international trade, which is a known factor influencing the premature exit from manufacturing activities in emerging economies (Sposi et al., 2021), but also an increase in the productivity of establishments within manufacturing (Melitz, 2003).

To investigate the drivers of premature de-industrialization, I construct regional trade shocks using IPUMS microdata from 31 high-, middle-, and low-income countries, following the approach of Autor et al. (2013). This allows me to examine how rising international trade exposure influences the sectoral composition of employment across a broad range of national income levels. In a future version of the paper, I will integrate these reduced form results in a model with trade and firm heterogeneity, in the form of distortions and productivity investment, which can be used to explore the impact of industrial and trade policy to understand whether countries should (i) reduce trade exposure to foster the manufacturing sector, (ii) increase trade exposure to make the manufacturing sector more productive and less distorted or (iii) pursue a combination of both objectives through targeted industrial policy and increasing trade openness. These counterfactual results are not yet available in this version of the paper. Appendix B however documents a pervasive negative effect of increasing trade exposure on the manufacturing employment share, across all low and

1. The assumption that returns to human capital are positively affected by the economy's growth rate is widespread in the quantity-quality tradeoff literature since the earliest contributions of Becker et al. (1990); Galor and Weil (2000); Parente and Prescott (2002) and others. This assumption reflects the well-known fact that aggregate productivity growth has an erosion effect on human capital, raising the return to educational investment (Nunn, 2014).

middle income countries, and a positive effect on the employment shares of *low-skill* services in reduced form regressions of the type run by Autor et al. (2013) for the United States.

This paper contributes to several strands of literature on structural change, unified growth, and trade. Economists have long been interested in the movement of labor out of the rural-agricultural sector, dating back to the seminal contributions of Kuznets (1966, 1973), Lewis (1954), and Rostow (1959). This literature saw a revival with formal models linking structural transformation and growth (e.g. Matsuyama, 1992b; Kongsamut et al., 2001; Ngai and Pissarides, 2007; Buera and Kaboski, 2012), which showed how sectoral reallocation accompanies sustained income growth. Cross-country variation in these transitions has renewed interest in the topic (Rodrik, 2016; Huneeus and Rogerson, 2024); see also Gollin and Kaboski (2023) for a recent survey. A related literature emphasizes the role of international trade in shaping structural transformation. Notably, Matsuyama (1992a) and Galor and Mountford (2006, 2008) argued that trade could delay industrialization in countries with high agricultural productivity, by reinforcing comparative advantage in low-income-elasticity sectors—an idea rooted in Prebisch (1959) and Engel’s law. Unified Growth Theory (UGT) provides a framework for understanding the transition from stagnation to growth. However, most UGT models are single-sector and deterministic: development is inevitable, and cross-country income gaps reflect only timing differences (e.g. Galor, 2011; Parente and Prescott, 2002; Hansen and Prescott, 2002; Cervellati and Sunde, 2015). Some extensions introduce sectoral structure—typically a modern, skill-intensive sector and a traditional, land-intensive one—to study fertility dynamics (Doepke, 2004; Galor and Mountford, 2008). To the best of my knowledge, this paper presents the first unified growth model featuring three-sector structural change—agriculture, manufacturing, and services—embedded in an endogenous growth framework. This allows me to revisit a key open question: Is manufacturing essential for sustained development, or can services lead growth? While recent work suggests the latter is possible (Fan et al., 2023; Herrendorf et al., 2022; Juhász et al., 2023), formal models that study this question remain limited. Other papers which explore population and industrialization include reduced form exercises by Ager et al. (2020)

and Wanamaker (2012), which study the role of industrial expansion in triggering the fertility reduction in quasi-experimental settings in the post-bellum US South. More theoretical contributions include the work of Doepke (2004), Bar and Leukhina (2010) and, more recently, Cheung (2023) who build models of endogenous fertility (but exogenous growth) to investigate the determinants of the demographic transition.

The rest of this paper is structured as follows: Section 2 presents cross-country evidence relating manufacturing booms and income gaps, the differences in structural change patterns between late and early starters and the relationship between structural change and fertility. Section 3 presents the baseline model environment. Section 4 presents the calibration strategy and counterfactual exercises. Section 5 empirically documents the relationship between trade exposure and the manufacturing sector, laying the groundwork for future counterfactual exercises. Section 6 concludes.

2 Stylized Facts

This section presents three stylized facts concerning structural change and the demographic transition. Data on modern sectoral employment shares in developing countries is from two sources: the Groningen Growth and Development Center Hamilton and de Vries (2023) and Gapminder, whereas the corresponding historical values for the US are from the 19th century censuses compiled and cleaned by Haines (2001) and FRED for the post 1940 period. I compare the recent evolution of the employment shares in developing countries to that of the US at the turn of the 20th century for two reasons: first, there is high quality and easily accessible data for the US which can be used to arrive at employment shares going as far back as 1880,² secondly, the US provides an example of an early industrializer which developed at considerable distance and high levels of tariff protection from the frontier economy at the time, Great Britain (Chang, 2003). If the import competition channel has played a role in

2. One could go back as far as 1850 although the earlier censuses are generally considered less reliable.

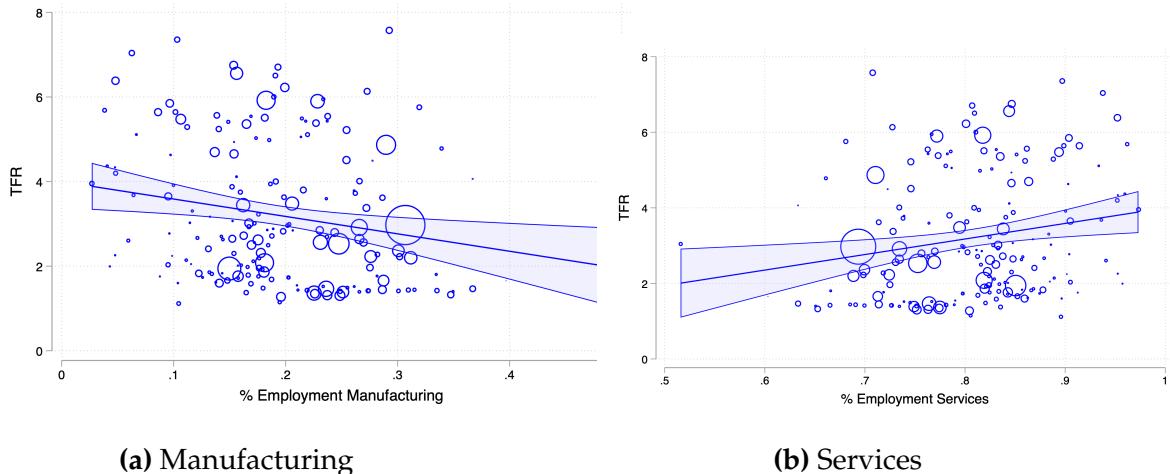


Figure 1: Scatterplot of total fertility rates (TFR) against the share of manufacturing (left) and services (right) in non-agricultural labor. Each dot represents a country average over 1990-2018, size refers to population in the last observed year. Source: Gapminder.

shaping the structural change process historically and now in developing countries, differences should arise in the employment shares of the US as it was developing and those observed in today's developing countries.³ Using this data, I document the following facts:

FACT 1: Countries with higher shares of manufacturing (services) employment in the non-farm sector have lower (higher) total fertility rates. Using IPUMS microdata this relationship is shown to be driven by differential fertility rates within countries: women working in the manufacturing sector have lower fertility rates than those in the - low productivity - services sector.

Figures 1 and 12 show that there is a strong relationship between the direction of the structural change process and the pace of a country along the demographic transition. Previous literature has documented how exit from the agricultural sector is associated with lower fertility due to a decrease in the value of child labor, an increase in the returns to education, and the opportunity cost of parents' time (Ager et al., 2020). However, there is little work on the heterogeneity

3. This pattern generally holds if we replace the US with another advanced economy at the time of its transition out of agriculture, however, the comparison is often less clear as in the case of France, Germany, Italy and others significant political developments over the 19th century, as well as their significant proximity to Great Britain, may contaminate these comparisons.

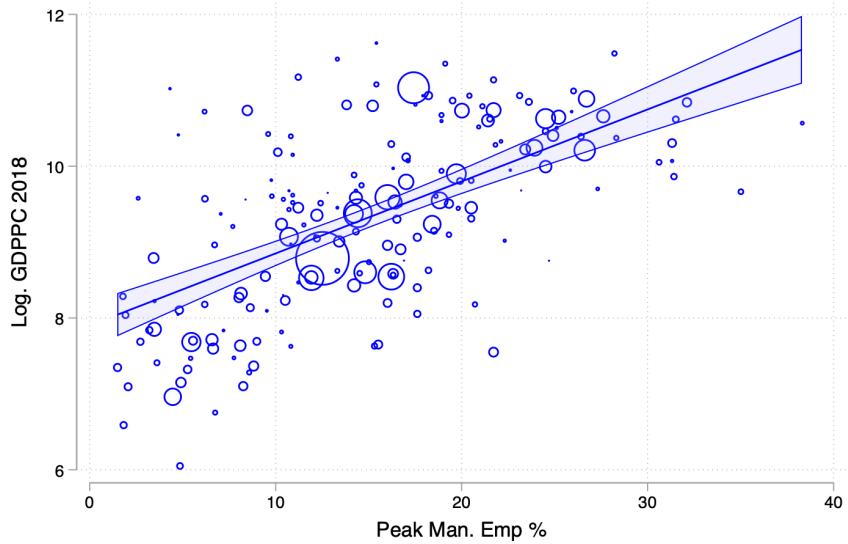


Figure 2: Weighted scatter plot of GDP per capita against the peak employment share of manufacturing in 195 countries. The weights correspond to population in 2018. Data source: Gapminder.

in fertility patterns within the non-farm sector. Heterogeneity within the non-agricultural sector is a strong predictor of a country's fertility rates as countries where the non-farm sector is primarily service based tend to have higher fertility rates. These relationships are statistically significant at the 5% level in the cross-sectional averages over 1990-2018, as well as in panel data regressions with country and year fixed effects as shown in Table 2 in Appendix B.

FACT 2: A country's peak manufacturing employment share is positively related to its current level of GDP per capita.

Figure 2 shows the relationship between peak manufacturing employment shares over the past two centuries and GDP per capita in 2018.⁴ I take 2018 as the "current" gdp per capita to avoid any transitory effect of the Covid-19 pandemic and its fallout which could temporarily alter the distribution of GDP per capita

4. Gapminder provides data for the past 30 years for developing countries, where employment in manufacturing, in any case, did not peak prior to the 1990s.

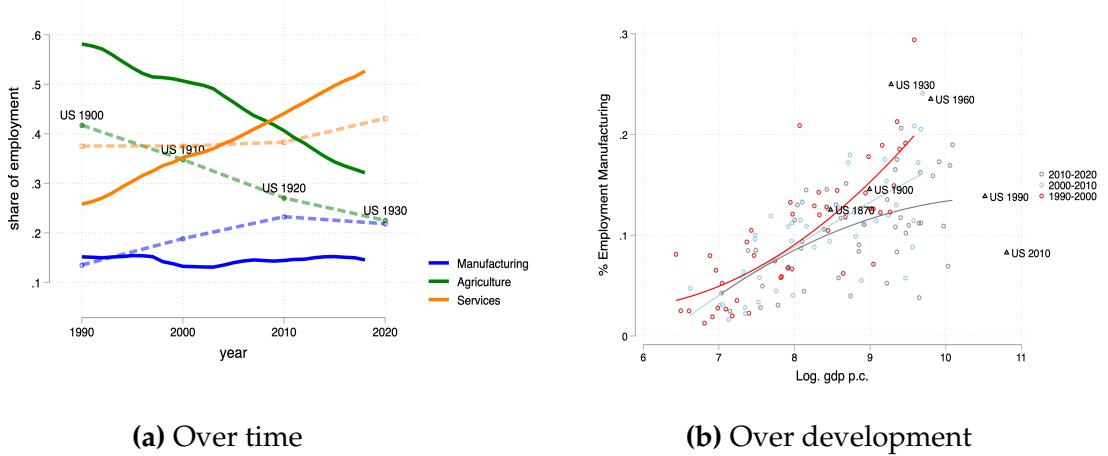


Figure 3: Scatter plot of manufacturing employment shares over development by decade (right) and time series plot of the average of the three employment shares weighted by population compared to the corresponding values for the US one century before (dashed lines) in 45 low and middle income countries (left). Source: GGDC Economic Transformation Database and US Censuses of Population 1880-1930 and FRED.

in recent years.⁵ This confirms the well-known fact that virtually all rich countries have experienced large manufacturing shares of employment, or industrial booms, at some point in their past.

FACT 3: Developing countries over 1990-2018 have lower shares of manufacturing employment compared to the US at similar levels of GDP per capita.

Panel (b) of Figure 3 compares the evolution of sectoral employment shares in the developing countries covered by the GGDC to those of the US at the turn of the century, starting in 1900 when it had a similar level of the manufacturing employment share as that in developing countries in 1990. This exercise shows two striking differences. Firstly, the transition out of agricultural labor is much faster in today's developing countries than it was for the US: over a thirty year period when the US agricultural share declined by 18pp, the agricultural share in developing countries declined by almost 30pp. The same holds for the entry into services, Figure 3 shows that virtually all labor exiting agriculture over 1990-2018 in developing countries has been absorbed by the service sector.

5. This relationship is highly significant and not affected by choosing 2018 or another reference year.

Panel (b) of Figure 3 plots this relationship over levels of GDP per capita instead of time. Virtually all countries in the period between 2010 and 2020 had lower (higher) levels of manufacturing (services) employment shares when compared to the US at a similar level of development.⁶ We can also see that this relationship is relatively recent, in the period spanning 1990-2000 developing countries had, on average, similar levels of manufacturing employment shares, albeit higher services employment shares. Rodrik (2016) first discovered these patterns, although without making an explicit comparison with a (now) developed country at similar levels of GDP per capita. Sposi et al. (2021) showed that international trade through an import competition mechanism is an important driver of the downward trend in the manufacturing employment share's relationship with GDP per capita. In Section B.2 I estimate the effect of increasing trade exposure on the sectoral employment shares finding that trade shocks have not only lowered manufacturing shares of employment, but also raised those of non-tradeable low-skill services.

3 Model

This section presents a general equilibrium model of structural change and the demographic transition. The baseline structure borrows from Galor and Weil (2000) and Lagerlöf (2006) to which I add sectoral and firm heterogeneity, correlated distortions and endogenous productivity investment. Necessary features of the model which we will require to study the role of manufacturing in the development process in a way which is consistent with Section 2 are: (i) multiple sectors, (ii) endogenous fertility choices, (iii) endogenous technological change, (iv) the ability to generate a "takeoff" from stagnation to growth as in other models which study long-run development outcomes (e.g. Galor and Weil, 2000; Cervellati and Sunde, 2015; Parente and Prescott, 2002) and (v) a source of frictions in the manufacturing sector.⁷ While this may seem like a very exigent

6. This statement is clear for levels of GDP per capita above 8000\$, however there is no data from when the US was at a level of GDP below 8000\$ it is difficult to make a comparative statement for the poorest developing countries.

7. Condition (v) is necessary to account for the well known fact that the manufacturing sector is more or less distorted in different countries/time-periods Alfaro et al. (2023), therefore complicating the study of industrial policy in promoting cross-country convergence.

set of modeling features, the model presented below remains highly tractable, with many of its mechanisms being summarized in closed form relationships between key variables.

3.1 Households

Time is discrete and indexed by t . The world is populated by overlapping generations of identical households who live for two periods. Following Galor and Weil (2000), individuals receive education when young and, in adulthood, make decisions about fertility, consumption, labor supply, and their children's education. Preferences are defined over a composite consumption good C_t and the effective number of children is $n_t h_{t+1}$, where h_{t+1} denotes the human capital of each child. The household maximizes:

$$\max_{c_{at}, c_{mt}, c_{st}, n_t, e_t} (1 - \gamma) \ln(C_t) + \gamma \ln(n_t h_{t+1}) \quad (1)$$

subject to the following constraints:

$$\begin{aligned} C_t &\geq \bar{c} \\ p_{at}c_{at} + c_{mt} + p_{st}c_{st} &\leq (1 - n_t(e_t + \tau))w_t, \\ h_{t+1} &= \frac{e_t + \rho\tau}{e_t + \rho\tau + g_t}, \\ 1 &= \sum_{i \in \{a, m, s\}} \omega_i^{1/\sigma} \left(\frac{C_{it}}{C_t^{\varepsilon_i}} \right)^{\frac{\sigma-1}{\sigma}} \end{aligned}$$

where p_{mt} is normalized to 1 and \bar{c} represents a minimum consumption floor. Parameters satisfy $\gamma, \tau, \rho \in (0, 1)$, and $\sigma \in (0, 1)$, sectors are gross complements. By assuming households give the same level of education to all their children, a tradeoff between quantity and quality is generated in the spirit of Becker et al. (1990). Households who have many children face steep education

costs as education takes time away from labor market supply and lowers effective income. The functional form of human capital accumulation is taken from Lagerlöf (2006) and satisfies the assumptions in Galor and Weil (2000), namely $\partial h/\partial e > 0$, $\partial h/\partial g < 0$ and $\partial h/\partial e \partial g > 0$. These assumptions ensure that technological growth has an erosion effect on current human capital by rendering existing knowledge obsolete, which raises the returns to educational investment. On the household side, the model differs from previous literature by adding non-homothetic preferences in the spirit of Comin et al. (2021). This preference structure is implicitly additive with constant price elasticity of substitution across goods, governed by σ , but different income elasticity of demand across goods. The parameter ε_i effectively generates an aggregate income varying weight, following the literature I assume that $\varepsilon_a < \varepsilon_m < \varepsilon_s$, that is agriculture has a lower income elasticity of demand than manufacturing and services.

3.2 Production

The economy consists of three sectors:⁸ agriculture a , manufacturing m , and services s . To bring the model closer to policy, the manufacturing sector features heterogeneous firms with correlated distortions generating resource misallocation. Agriculture and services, on the other hand, operate in a frictionless world where their output can be generated using the following representative firm setup with CRS technology:

$$Y_{it} = K_{it}^{1-\nu} \left((h_t L_{it})^\alpha (X A_{it})^{1-\alpha} \right)^\nu, \quad \alpha, \nu \in (0, 1) \quad (2)$$

where X is a fixed supply of land and A_{it} is the sector-specific level of technology, K is capital and L is labor. To abstract from capital *accumulation*, I assume a fixed world interest rate which determines the level of capital as a function of technology and other inputs to the production function, which implies that the effective production function is DRS in labor:

8. Potentially this could be extended to distinguish between high and low productivity services.

$$K^* = \left(\frac{1-\nu}{\bar{r}} \right)^{\frac{1}{\nu}} (h_t L_{it})^\alpha (X A_{it})^{1-\alpha} \implies Y_{i,t} = \underbrace{\left(\frac{1-\nu}{\bar{r}} \right)^{\frac{1-\nu}{\nu}}}_{\Omega} (h_t L_{it})^\alpha (X A_{it})^{1-\alpha}$$

As there is a single labor market with homogeneous workers, in equilibrium we will have that the FOC for labor will equalize across the sectors yielding:

$$\frac{p_a}{p_s} = \left(\frac{A_s / L_s}{A_a / L_a} \right)^{1-\alpha} \quad (3)$$

Relative prices are pinned down by technology, with the only difference being that here they will also depend on the sectoral employment shares due to the fact that X is in fixed supply. Aside from this variation, necessary to generate the Malthusian relation between wages and population, the agriculture and services sectors of the economy are relatively standard and align very closely with benchmark models of structural transformation (e.g. Herrendorf et al., 2014). The manufacturing sector on the other hand, features heterogeneous firms with firm level distortions in the spirit of Hsieh and Klenow (2009), generating mis-allocation of productive resources in the economy. Each firm i in the manufacturing sector produces output according to:

$$y_i(z_i, l_i) = z_i (h_t l_i)^{\phi_1} (A_{m,t} X)^{\phi_2}$$

where $z_i \in \{z_{low}, \dots, z_{high}\}$ is a discrete firm-specific productivity level, h_t is the labor-augmenting human capital common across sectors, $A_{m,t}$ is the manufacturing-specific technology, and X is the fixed supply of land and the manufacturing goods price p_m is normalized to one. Firms face idiosyncratic distortions τ_i that act as wedges on labor costs, such that they pay $(1 + \tau_i)w_t l_i$ to hire l_i units of labor at the equilibrium wage w_t . Manufacturing firms make two decisions: each period firms chooses labor to maximize profits:

$$\pi(z_i) = \max_{l_i} z_i (h_t l_i)^{\phi_1} (A_{m,t} X)^{\phi_2} - (1 + \tau_i) w_t l_i$$

and between periods firms can pay a fixed investment cost in order to increase the possibility of raising their TFP. The timing is such that these decisions are taken sequentially. In the absence of productivity enhancing investment, each period TFP can go up - increase one step on the ladder - with probability p and down with probability $1 - p$, whereas if the firm pays a fixed investment cost κ this probability goes up to $p_i > p$. The value function of the manufacturing firm is therefore:

$$V(z) = \max_{\mathcal{I} \in \{0,1\}} \left\{ \pi(z) - \kappa \mathcal{I} + \beta \mathbb{E}_{z' | z, \mathcal{I}} [V(z')] \right\} \quad (4)$$

where $\mathcal{I} \in \{0, 1\}$ is the binary innovation decision, κ is the fixed cost of innovation, and $\pi(z)$ denotes current-period profits as a function of firm productivity z . The law of motion for productivity follows a discrete Markov process:

$$\Pr(z' = z_j | z = z_i, \mathcal{I} = 0) = P_{ij}, \quad \Pr(z' = z_j | z = z_i, \mathcal{I} = 1) = P_{ij}^{\text{innov}}$$

where P and P^{innov} are the transition matrices for firms that, respectively, do not and do innovate, and P^{innov} governs the transitions when firms pay the innovation cost. These matrices feature zero entries whenever $\text{abs}(i - j) > 1$, i.e. a firm can only move up one step or down one step each period. Innovation increases the probability of moving to a higher productivity level. The innovation decision is added to the model to endogenize the growth of aggregate productivity in manufacturing A_m as will become clear in Section 3.3. Within a period, the first-order condition for labor yields:

$$l_i(z, w, \tau) = \left(\frac{\phi_1 z_i h_t^{\phi_1} (A_{m,t} X)^{\phi_2}}{(1 + \tau_i) w_t} \right)^{\frac{1}{1-\phi_1}}$$

Letting Φ_t and Ψ_t denote distortion-weighted productivity aggregates:

$$\Phi_t \equiv \int (1 + \tau_i)^{-\frac{1}{1-\phi_1}} z_i^{\frac{1}{1-\phi_1}} di, \quad \Psi_t \equiv \int (1 + \tau_i)^{-\frac{\phi_1}{1-\phi_1}} z_i^{\frac{1}{1-\phi_1}} di$$

Aggregate labor demand in manufacturing is:

$$L_{m,t} = \left(\frac{\phi_1 h_t^{\phi_1} (A_{m,t} X)^{\phi_2}}{w_t} \right)^{\frac{1}{1-\phi_1}} \Phi_t$$

Aggregate manufacturing output is:

$$Y_{m,t} = A_{m,t}^{\frac{\phi_2}{1-\phi_1}} X^{\frac{\phi_2}{1-\phi_1}} h_t^{\frac{\phi_1}{1-\phi_1}} L_{m,t}^{\phi_1} \cdot \left(\phi_1^{\frac{\phi_1}{1-\phi_1}} \Psi_t \Phi_t^{-\phi_1} \right)$$

Output in manufacturing is shaped not only by fundamentals but also by the extent of misallocation, summarized by the term $\Psi_t \Phi_t^{-\phi_1}$. In a frictionless benchmark where distortions τ_i vanish, this term collapses to a one, and all firms operate at the same marginal return to labor. With as long as some firms face positive wedges however, the manufacturing output will be less than its full potential by a factor of $\Psi_t \Phi_t^{-\phi_1} \in (0, 1)$. For analytical tractability we parameterize the output elasticities ϕ_1 and ϕ_2 such that manufacturing output scales with $(A_{m,t} X)^{1-\alpha}$. This requires setting:

$$\phi_2 = (1 - \alpha)^2, \quad \phi_1 = \alpha$$

Under this condition, relative prices across sectors can be derived by equalizing marginal products of labor across sectors, yielding expressions that depend on sectoral employment shares, technology, and the misallocation term. Following these steps, relative prices of agriculture and services are:

$$\begin{aligned} p_a &= C \left(\frac{A_m}{A_a} \right)^{1-\alpha} \left(\frac{s_m}{s_a} \right)^{\alpha-1} \\ p_s &= C \left(\frac{A_m}{A_s} \right)^{1-\alpha} \left(\frac{s_m}{s_s} \right)^{\alpha-1} \end{aligned} \quad (5)$$

where:

$$C = \alpha^{\frac{\alpha}{1-\alpha}} \cdot \Psi \cdot \Phi^{-\alpha} \cdot h_t^{\frac{\alpha^2}{1-\alpha}}$$

is a market level constant. Dividing p_a by p_s yields an expression consistent with (3).

3.3 Population and Technology Dynamics

Each period, population evolves according to fertility decisions:

$$L_{t+1} = n_t L_t. \quad (6)$$

Technological progress in each sector is endogenous and depends on the size of the population up to a sector specific threshold \bar{a}_i :

$$A_{i,t+1} = (1 + (e_t + \rho\tau)a_{it}) A_{it}, \quad \text{where } a_{it} = \min\{\theta L_t, \bar{a}_i\}. \quad (7)$$

Technological progress is exogenously lower in the service sector, with $\bar{a}_s < \bar{a}_a$ and manufacturing productivity growth reaches its full potential only when all firms are investing in innovation, that is, there is an externality whereby firm innovation not only raises individual TFP but also the growth rate of aggregate productivity A_m :

$$\bar{a}_m = \bar{a}_m^{\max} \sum_z \mu(z) \mathcal{I}(z)$$

Where $\mathcal{I}(z)$ is the innovation policy, μ is the stationary distribution over z and $\bar{a}_s < \bar{a}_m^{\max} \leq \bar{a}_a$.

3.4 Equilibrium

A competitive equilibrium in period t consists of household decisions $\{c_{it}, n_t, e_t\}$, firm decisions $\{L_{it}, l_{it}, k_{it}, \mathcal{I}(z)\}$, prices $\{w_t, p_{nt}, p_{st}\}$, technologies $\{A_{it}\}$ and a distribution of firms over productivity levels $\mu(z)$; such that, given prices, the innovation cost κ and a distortion schedule $\tau(z)$:

- **Labor market clears:**

$$L_t = \sum_{i \in \{a, m, s\}} L_{it}. \quad L_m = \sum_z l(z, w, \tau(z)) \mu(z)$$

- **Goods markets clear:**

$$Y_{it} = c_{it}, \quad \text{for all } i \in \{a, m, s\}. Y_{mt} = \sum_z y(z) \mu(z)$$

- **Households optimize:** The household solves 1.

- **Aggregate Distribution:**

$$\mu = \mu \cdot \left[(I - \mathcal{I})P + \mathcal{I}P^{\text{innov}} \right]$$

As in Buera et al. (2020) the economy approaches a BGP when $s_s \rightarrow 1$.

3.5 Characterization

As in standard structural change models (Herrendorf et al., 2014), wages equalize across the sectors since a competitive labor market implies firms face an

infinite labor supply elasticity. The capital labor ratios also will equalize and take the following form:

$$\frac{k_i}{l_i} = \frac{w_t}{\bar{r}} \cdot \frac{1-\nu}{\nu\alpha}, \quad (8)$$

Following Lagerlöf (2006) and Galor and Weil (2000) we can derive closed form expressions for the optimal fertility and education choices. Define $\tilde{z} = \bar{c}/(1-\gamma)$ as the income level which allows the household to cover its subsistence requirement by allocating its preferred time share γ to child rearing. The optimal fertility decision is then:

$$n_t(\tau + e_t) = \begin{cases} \gamma, & \text{if } w_t \geq \tilde{z}, \\ 1 - \frac{\bar{c}}{w_t}, & \text{if } w_t \in (\bar{c}, \tilde{z}), \\ 0, & \text{if } w_t \leq \bar{c}. \end{cases} \quad (9)$$

If income is sufficiently low, the subsistence constraint is binding and the household can only afford to dedicate $1 - \frac{\bar{c}}{w_t}$ of its time to child rearing, the time it has left over after earning enough to consume \bar{c} . The optimality condition for education is slightly more involved, however, we can show that:

Proposition 1. *The household education choice satisfies:*

$$e^* = \max\{0, \sqrt{g_t(1-\rho)\tau} - \rho\tau\} \quad (10)$$

Proof. See Appendix A □

The optimal condition for e shows that education investment is a positive function of the aggregate growth rate in this model. There are several reasons why this is a reasonable prediction, one of which is that technological change is often found to erode the stock of human capital, thereby raising the return to education (Nunn, 2014).

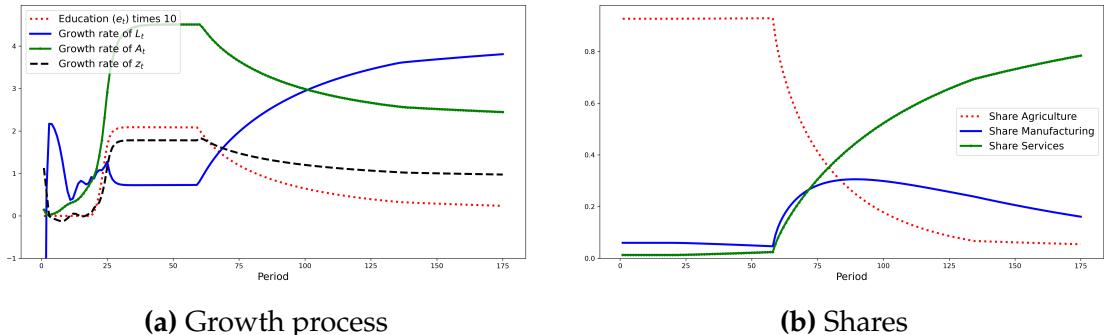


Figure 4: Model simulation starting from an arbitrary set of initial parameters.

Using the Marshallian demand system from 1, the dynamics of population and technology 6, 7, the decisionf for education and fertility 10, 9 and the pricing equations 5 we have a system of 11 equations for 11 endogenous variables, $\{c_{it}, p_{it}, e_t, n_t, A_{i,t+1}, L_{t+1}, h_{t+1}\}_{i \in \{a,m,s\}}$ which we can solve each period and simulate forward starting from arbitrary initial values. The model generates the following qualitative patterns: (1) an initial Malthusian pseudo-steady state where income per capita and population are fluctuating, (2) a post-Malthusian phase where income per capita and population are growing, (3) a sustained growth regime where fertility declines and income per catita grows at a quasi constant rate and (4) a growth slowdown as the economy transitions to the lowest productivity sector (services) with a corresponding rise in fertility. Unlike other unified growth models where development is "inevitable", i.e. given sufficient amount of time the economy reaches a balanced growth path with sustained income per capita growth (Cervellati and Sunde, 2015; Galor and Weil, 2000), in this model the heterogeneity in sector growth rates introduces an additional layer of complexity which makes development a non-trivial outcome of this process. In particular, whether the economy successfully exits the Malthusian trap will be substantially affected by its structural change process. More specifically:

Proposition 2 (Escape from Stagnation and the Onset of Development). *The necessary and sufficient condition to achieve sustained growth in income per capita is that:*

$$e^* > 0 \iff g_t > \frac{\rho^2 \tau}{1 - \rho} \quad \text{where} \quad g_t = \sum_i s_{it} g_{it} \quad (11)$$

Proof. See Appendix A □

In order to escape the Malthusian trap positive investment in education is necessary, which only occurs when the growth rate of the economy crosses a well defined threshold. As the technological constraints on growth are different across sectors, reflected in the assumption that $\bar{a}_a > \bar{a}_m > \bar{a}_s$, the potential growth rate of service sector productivity is lower than that of manufacturing and agriculture (e.g. Vollrath, 2020; Moro et al., 2017), whether the economy is able to takeoff at all is dependent on its structural change process. A service-bias of the structural change process implies a lower probability of takeoff, as it requires stronger parameter restrictions, in particular a small time cost per child and a small baseline human capital level. Moreover, larger, and more productivity correlated, distortions will also lower \bar{a}_m making the manufacturing sector endogenously less productive. Hence, as whether the economy ultimately achieves a takeoff into sustained growth depends on the growth of economy wide aggregate productivity, calculated as in 2, correlated distortions can delay the timing of the economy's takeoff into sustained growth as it reaches its peak growth frontier \bar{a}_m^{\max} only when all firms are innovating. This will eventually occur as long as the investment cost remains fixed as aggregate productivity A_m is growing over time, albeit at a lower rate if at least some firms are not innovating. For the purpose of this exercise, consider a distortion schedule which is a simple linear mapping between distortions τ and z with upper and lower bounds $\bar{\tau}$, meaning that the lowest productivity firm z_{low} receives a subsidy of $\bar{\tau}$ whereas the highest productivity firm z_{high} pays a tax of $\bar{\tau}$. We can see this in Figure 5.

This distortion schedule achieves the intended goal of making innovation less appealing, as increasing z entails a higher distortion, while preserving tractability. We can now plot along the transition path the evolution of misallocation and innovation rates for different distortion schedules $\tau(z)$, which will help

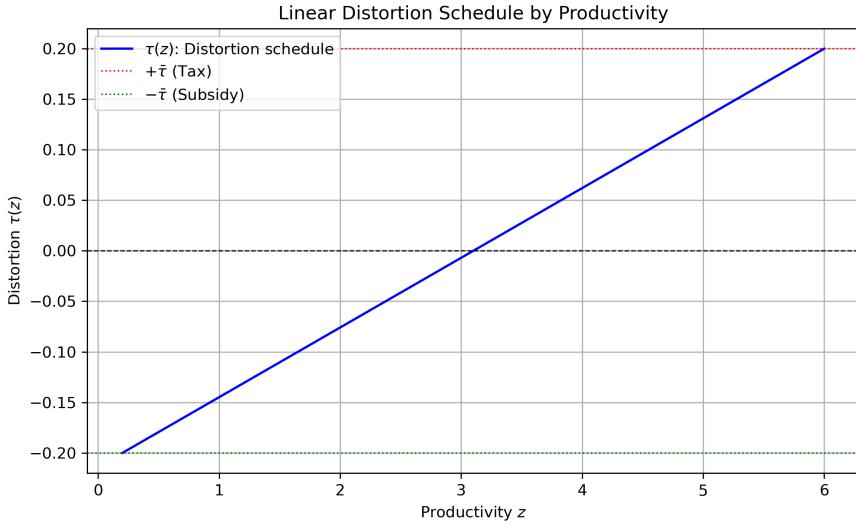


Figure 5: Distortion schedule example, the highest productivity firm pays a tax of 20% of its wage bill, whereas the lowest productivity firm gets a subsidy of 20%.

gain intuition over how correlated distortions can affect the takeoff into sustained growth.

Figure 6 shows that as the economy transitions to sustained growth, the misallocation gap $\Psi_t \Phi_t^{-\phi_1}$ goes down, owing to the fact that a growing aggregate productivity makes investment an attractive prospect even with correlated distortions. This implies that, as the firm innovation rate goes up, the productivity distribution becomes more concentrated at higher levels of z lowering aggregate labor misallocation.

4 Model Calibration

This Section uses data on employment shares and GDP per capita growth from the US over 1800-2020 to calibrate the model. In all counterfactual exercises I will measure misallocation from correlated distortions in relative terms compared to the US. I hence take the US as the benchmark, non distorted economy,

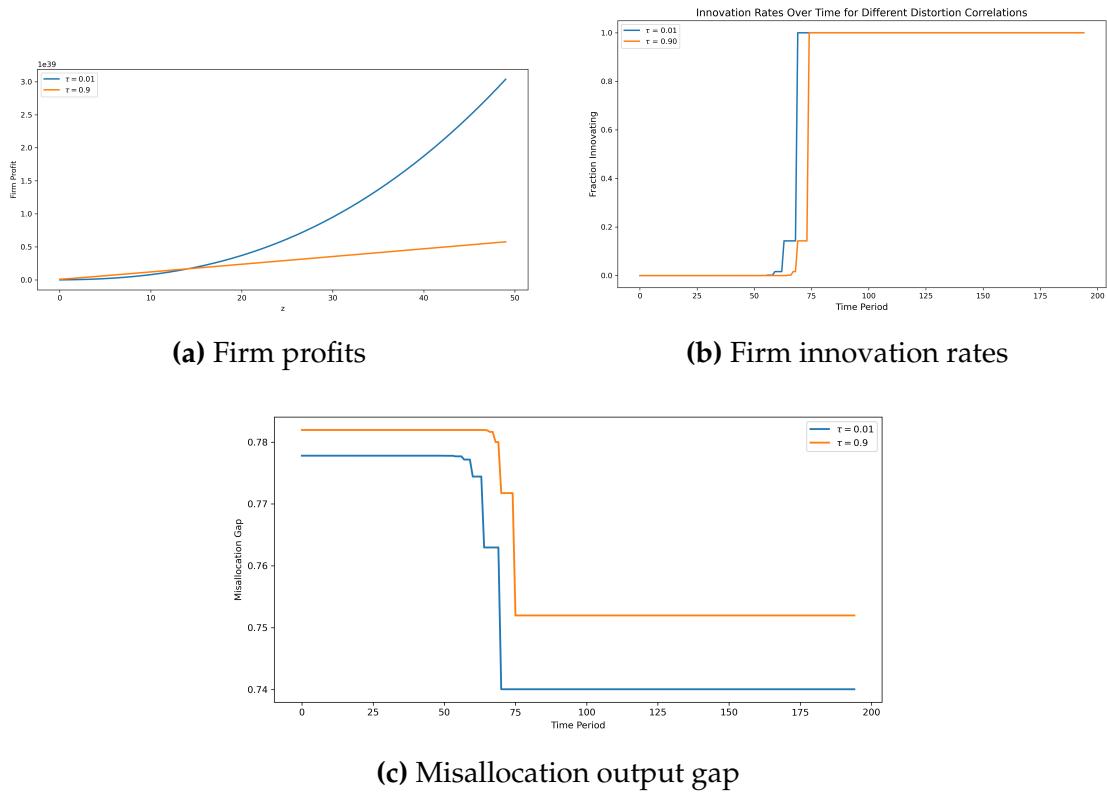


Figure 6: Firm-level outcomes: (a) profits, (b) innovation rates, (c) misallocation

with $\bar{\tau} = 0$, and κ sufficiently low that it is always optimal to invest in productivity, implying that the model effectively collapses to a production structure with three representative firms having DRS technology.

I calibrate the model parameters using the Generalized Method of Moments (GMM), targeting the evolution of employment shares in agriculture, manufacturing, and services over the period 1850–2010, as well as GDP per capita and population growth factors over 1800–2010.⁹ The GMM objective function is as follows:

9. the choice of 1800 as an initial year is due to data constraints, whereas the choice of 2010 as a final year is arbitrary.

$$\begin{aligned}
\mathcal{L}(\theta) = & \sum_{t \in T_{\text{moments}}} \left[\left(s_{a,t}^{\text{sim}} - s_{a,t}^{\text{data}} \right)^2 + \left(s_{m,t}^{\text{sim}} - s_{m,t}^{\text{data}} \right)^2 + \left(s_{s,t}^{\text{sim}} - s_{s,t}^{\text{data}} \right)^2 \right] \\
& + \left(\frac{g_{\text{sim}}^{\text{final}} - g_{\text{data}}^{\text{final}}}{g_{\text{data}}^{\text{final}}} \right)^2 + \left(\frac{g_{\text{sim}}^{\text{mid}} - g_{\text{data}}^{\text{mid}}}{g_{\text{data}}^{\text{mid}}} \right)^2 \\
& + \left(\frac{P_{\text{sim}}^{\text{final}} - P_{\text{data}}^{\text{final}}}{P_{\text{data}}^{\text{final}}} \right)^2 + \left(\frac{P_{\text{sim}}^{\text{mid}} - P_{\text{data}}^{\text{mid}}}{P_{\text{data}}^{\text{mid}}} \right)^2
\end{aligned} \tag{12}$$

Where the "mid" moments for gdp and population correspond to the 1800-1900 population and gdp per capita growth factors, being 4.8 and 13 respectively. These moments are especially important as they help calibrate the timing of the transition.

The model includes a total of 20 parameters, of which 7 are fixed based on values commonly used in the literature or normalized for identification. The remaining 13 parameters are internally calibrated to make the models moments as close as possible to the data. As there are 13 parameters and 52 moments, the system is overidentified. The estimation uses simulated employment shares from the model evaluated at a set of 16 time points, corresponding to years 1850, 1860, ..., 2010 which are then compared to their data counterparts.¹⁰ The GMM optimization then searches over the parameter space to minimize 4 assigning equal weights to all moments. We can see in Figure 4 that the model is able to match the data with a remarkable degree of accuracy. The estimated parameters capture key features of structural transformation and economic growth. The low elasticity of agriculture ($\epsilon_a = 0.0552$) and high elasticity of services ($\epsilon_s = 1.0243$) are consistent Buera et al. (2020) and crucial to generate the desired evolution of employment shares over income.

10. Moments using model generated sequences of population and GDP per capita are first converted to yearly series assuming a model period of 20 years (Boldrin and Jones, 2002) and converting using the CAGR formula: $gX_t = 100 \left(\left(1 + \frac{X_{t+1} - X_t}{X_t} \right)^{\frac{1}{20}} - 1 \right)$

Parameter	Value	Description and Source
<i>Fixed Parameters</i>		
T	216	Time horizon in model periods (assumed)
α	0.6	DRS to labor in production (Restuccia and Rogerson, 2008)
σ	0.5	Substitution elasticity in utility (Buera and Kaboski, 2009)
ϵ_m	1.0	Income elasticity of manufacturing (normalization)
r^*	0.02	World interest rate (normalization)
ν	0.5	Capital share in production (standard value)
ω_m	8.0	Utility weight on manufacturing (normalization)
<i>Estimated Parameters (GMM)</i>		
\bar{c}	3.6245	Subsistence consumption threshold
τ	0.2263	Time cost of children
θ	0.00016	Effect of population growth on technology
ω_a	0.8553	Utility weight on agriculture
ω_s	19.5410	Utility weight on services
X	59.9367	Stock of land
γ	0.7713	Utility weight on children (implied)
ρ	0.2364	Baseline human capital formation (implied)
\bar{a}_s	1.4381	Maximum productivity growth in services
\bar{a}_m	2.5548	Maximum productivity growth in manufacturing
\bar{a}_a	2.9853	Maximum productivity growth in agriculture
ϵ_s	1.0243	Income elasticity of services
ϵ_a	0.0552	Income elasticity of agriculture

Table 1: Model Parameters: Calibrated and Fixed

The calibration fits the GDP per capita and population sequences exceptionally well while slightly falling short in the employment shares targeting. This can be improved with better optimization algorithms.

5 Counterfactuals

With the calibrated model in hand, the main counterfactual exercise I perform is to simulate the model for different utility weights of manufacturing consumption ω_m , which alters the peak employment share of the manufacturing sector

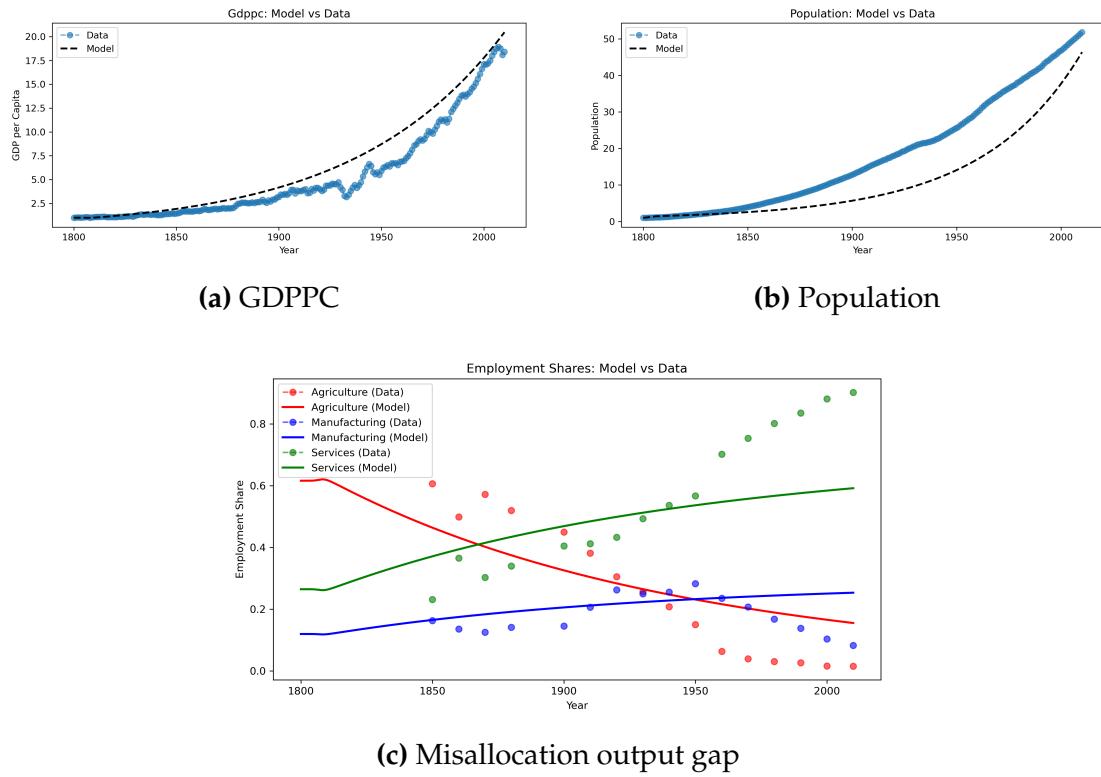


Figure 7: Model vs data: (a) gdp per capita, (b) population, (c) employment shares.

and see the effects this has on GDP per capita at the end of the simulation window. This should recreate the pattern in Figure 2, where we can see that peak manufacturing shares are significantly correlated with current GDP per capita. The counterfactual results are presented in Figure 8 and qualitatively replicate this pattern: a lower ω_m decrease the US' peak manufacturing share over the development trajectory and lowers its current GDP per capita. The mechanism which achieves this in the model is straightforward and intuitive: lowering the employment share of manufacturing lowers the average growth rate of the economy as less labor flows into the highest productivity sector, which in turn discourages education investments.

This is true even for mild productivity growth differentials which arise in the calibration to the US where the service sector peak productivity growth is 1.5%,

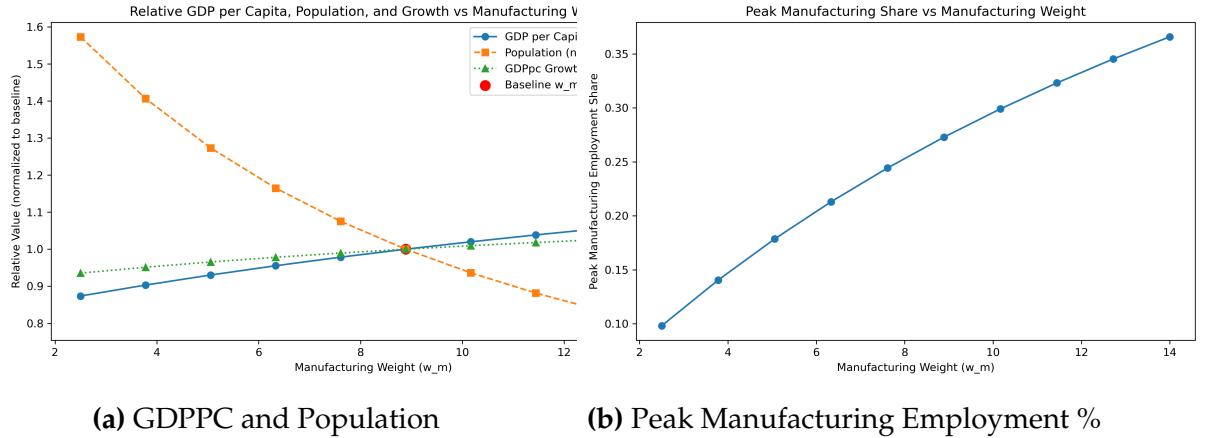


Figure 8: Calibrated model simulations for different values of ω_m .

only one percentage point lower than the growth in manufacturing. The results are also quantitatively significant: halving the peak employment share of manufacturing reduces GDP per capita in 2010 by 10%. Counterfactual paths of population are even more striking as, due to lower aggregate productivity growth, parents chose quantity over quality as the incentive to invest in education is lowered. This results in a 60% larger population in 2010, underscoring how small changes in growth rates, when compounded over a number of years along a balanced growth path can have large implications. With the calibrated model in hand, I turn to the main question of this paper which is to what extent different peak employment shares in manufacturing can explain variation in GDP per capita gaps today. In the benchmark counterfactual exercise, I simulate model economies with many different values of ω_m and associate each country in the world with the model economy which achieves its peak employment share of manufacturing. This is a common type of counterfactual exercise in the macro-development literature (e.g. Ruggieri et al., 2024, 2023), where by varying a parameter in a model calibrated to a developed economy to replicate a specific feature of a developing country can yield insights into how much that feature can explain differential outcomes in cross-country settings. The results are presented in Figure 9.

Both population and GDP per capita growth factors are significantly related to their model counterparts, with model implied growth factors explaining 27%

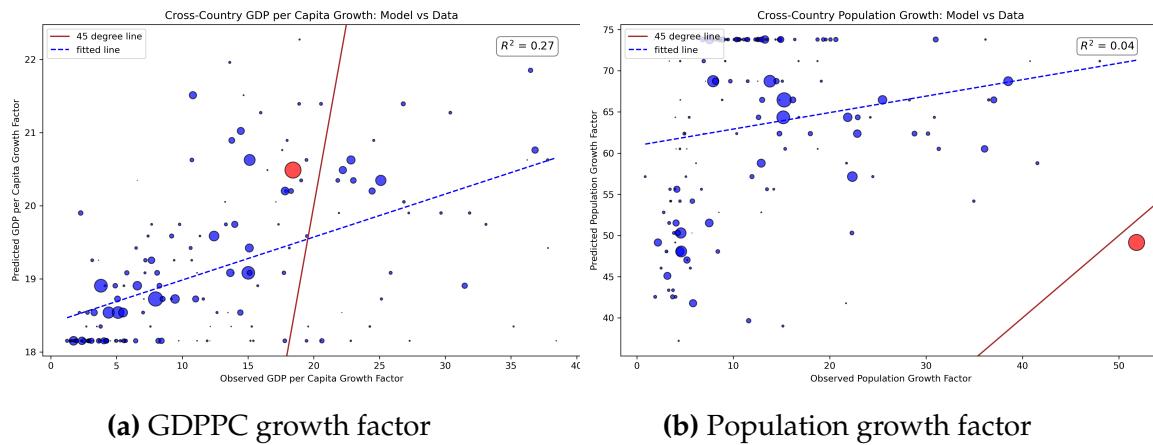


Figure 9: Calibrated model simulations for different values of ω_m . Each dot is a country with the size of the dot corresponding to its population in 2018. The y-axis shows the model economy growth factor whereas the x-axis shows the growth factor in the data.

and 4% of the variation in growth factors in the data respectively. The model is also able to match education gaps across countries even though this is non-targeted. We can see in Figure 10, where I plot a country/year observation

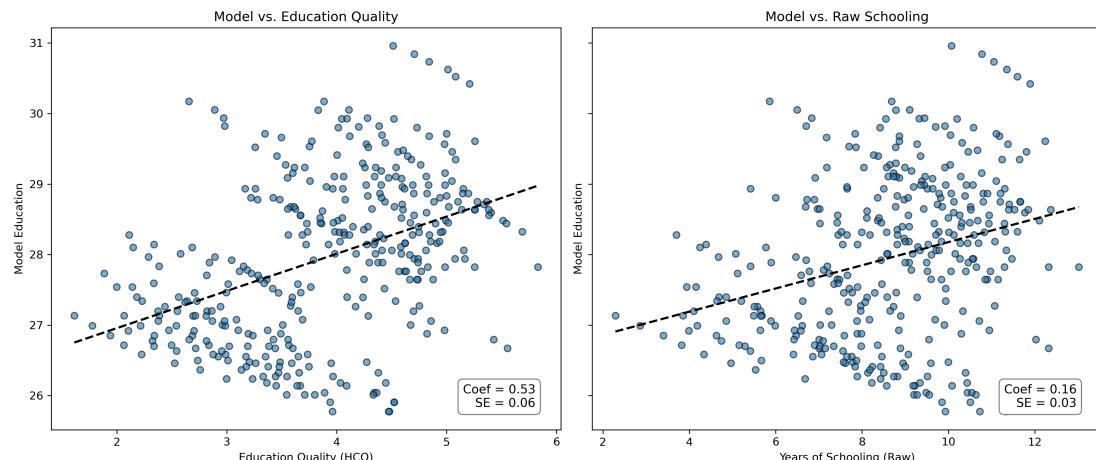


Figure 10: Each dot is a model at the country/year level, y axis shows the Barro Lee (2013) human capital quality index (left) and years of schooling (right) against the model's endogenous education decision.

of educational attainment, against the education choices implied by the model for that country/year observation. The data on educational attainment is from Barro and Lee (2013), who provide basic information on years of schooling and

human capital quality indices based on test scores. The relationship between the data and the model implied education metrics is positive and statistically significant. The final exercise I do is to simulate the distribution of income and population in a world in which countries differ from the US only on their initial conditions and the peak employment share of manufacturing they reach. This essentially involves iterating forward along the model simulation for each country starting from its initial level in 1800 for these two variables. The results

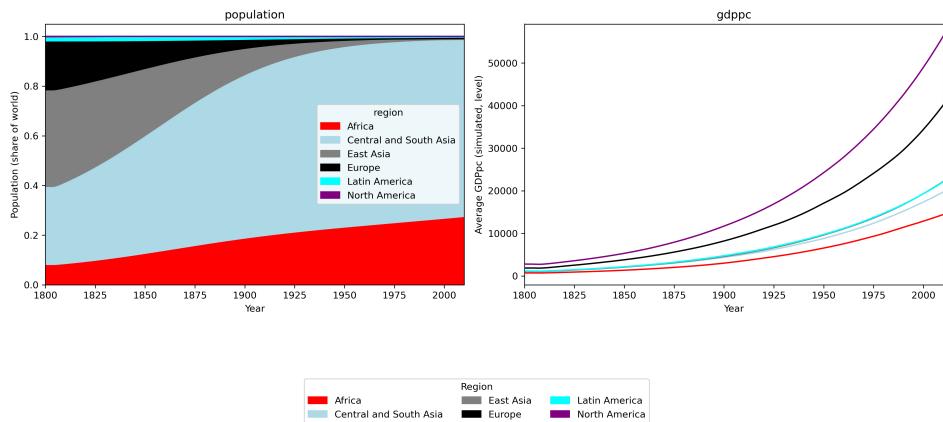


Figure 11: Path of income and population in 175 countries. A country's model implied path is determined by assigning it the model economy which achieves its peak-manufacturing employment share by changing w_m with all other parameters anchored to the US calibration.

are shown in Figure 11, where we can see that changes in the growth factor due to differential levels of industrial attainment contribute to a large part of the rise in world income inequality over the past two centuries. Around 14% of cross-country income gaps today can be explained by whether countries transitioning out of agriculture saw employment go to manufacturing before the rise of the service sector. Due to the general equilibrium feedback mechanism linking aggregate productivity growth with education choices and education choices with sectoral productivity growth, the effect of lower manufacturing shares is present even *after* countries have deindustrialized. This is because by raising education levels, the service sector in countries which had large manufacturing booms is more productive and grows faster, giving rise to transition paths with persistently different growth rates.

6 Conclusion

This paper develops a unified growth model featuring structural change, endogenous fertility, and firm-level distortions to study the long-run consequences of bypassing industrialization. Calibrated to the historical US experience, the model replicates key stylized facts about the relationship between manufacturing employment and development outcomes. Counterfactual simulations reveal that a country's peak manufacturing employment share is a strong predictor of its current GDP per capita and population, with lower industrialization leading to slower productivity growth, lower education investment, and higher fertility. These mechanisms explain up to 14% of today's cross-country income differences. Future work will extend the counterfactual analysis to assess how trade openness and correlated distortions influence countries' ability to industrialize and converge, shedding light on the role of international trade and policy in shaping the global distribution of income.

References

- Ager, Philipp, Benedikt Herz, and Markus Brueckner**, "Structural change and the fertility transition," *Review of Economics and Statistics*, 2020, 102 (4), 806–822.
- Alfaro, Laura, Marcela Eslava, Martí Mestieri, and Luis Felipe Sáenz**, "Jobless Industrialization," Technical Report, Mimeo Universidad de Los Andes 2023.
- Autor, David H, David Dorn, and Gordon H Hanson**, "The China syndrome: Local labor market effects of import competition in the United States," *American economic review*, 2013, 103 (6), 2121–2168.
- , —, —, and **Jae Song**, "Trade adjustment: Worker-level evidence," *The Quarterly Journal of Economics*, 2014, 129 (4), 1799–1860.
- Bar, Michael and Oksana Leukhina**, "Demographic transition and industrial revolution: A macroeconomic investigation," *Review of Economic Dynamics*, 2010, 13 (2), 424–451.
- Barro, Robert J and Jong Wha Lee**, "A new data set of educational attainment in the world, 1950–2010," *Journal of development economics*, 2013, 104, 184–198.
- Baumol, William J**, "Macroeconomics of unbalanced growth: the anatomy of urban crisis," *The American economic review*, 1967, 57 (3), 415–426.
- Becker, Gary S, Kevin M Murphy, and Robert Tamura**, "Human capital, fertility, and economic growth," *Journal of political economy*, 1990, 98 (5, Part 2), S12–S37.
- Boldrin, Michele and Larry E Jones**, "Mortality, fertility, and saving in a Malthusian economy," *Review of Economic Dynamics*, 2002, 5 (4), 775–814.
- Buera, Francisco J and Joseph P Kaboski**, "Can traditional theories of structural change fit the data?," *Journal of the European Economic Association*, 2009, 7 (2-3), 469–477.
- and —, "The rise of the service economy," *American Economic Review*, 2012, 102 (6), 2540–2569.

- , —, Martí Mestieri, and Daniel G O'Connor, "The stable transformation path," Technical Report, National Bureau of Economic Research 2020.
- Cervellati, Matteo and Uwe Sunde**, "The economic and demographic transition, mortality, and comparative development," *American Economic Journal: Macroeconomics*, 2015, 7 (3), 189–225.
- Chang, Ha-Joon**, "Kicking away the ladder: Infant industry promotion in historical perspective," *Oxford Development Studies*, 2003, 31 (1), 21–32.
- Cheung, T Terry**, "Schooling, Skill Demand, and Differential Fertility in the Process of Structural Transformation," *American Economic Journal: Macroeconomics*, 2023, 15 (4), 305–330.
- Comin, Diego, Danial Lashkari, and Martí Mestieri**, "Structural change with long-run income and price effects," *Econometrica*, 2021, 89 (1), 311–374.
- Dix-Carneiro, Rafael and Brian K Kovak**, "Trade liberalization and regional dynamics," *American Economic Review*, 2017, 107 (10), 2908–2946.
- Doepke, Matthias**, "Accounting for fertility decline during the transition to growth," *Journal of Economic growth*, 2004, 9, 347–383.
- Fan, Tianyu, Michael Peters, and Fabrizio Zilibotti**, "Growing like India—the unequal effects of service-led growth," *Econometrica*, 2023, 91 (4), 1457–1494.
- Galor, Oded**, *Unified growth theory*, Princeton University Press, 2011.
- and Andrew Mountford, "Trade and the great divergence: the family connection," *American economic review*, 2006, 96 (2), 299–303.
 - and —, "Trading population for productivity: theory and evidence," *The Review of economic studies*, 2008, 75 (4), 1143–1179.
 - and David N Weil, "Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond," *American economic review*, 2000, 90 (4), 806–828.
- Gollin, Douglas and Joseph P Kaboski**, "New views of structural transformation: insights from recent literature," 2023.

Haines, Michael R, "Historical demographic, economic and social data: The united states, 1790-1970," *Department of Economics. Colgate University, Computer File*, 2001.

Hamilton, C. and G. J. de Vries, "The Structural Transformation of Transition Economies," Research Memorandum 196, GGDC 2023.

Hansen, Gary D and Edward C Prescott, "Malthus to solow," *American economic review*, 2002, 92 (4), 1205–1217.

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi, "Growth and structural transformation," *Handbook of economic growth*, 2014, 2, 855–941.

— ,— , and **Ákos Valentinyi**, "New evidence on sectoral labor productivity: Implications for industrialization and development," Technical Report, National Bureau of Economic Research 2022.

Hsieh, Chang-Tai and Peter J Klenow, "Misallocation and manufacturing TFP in China and India," *The Quarterly journal of economics*, 2009, 124 (4), 1403–1448.

Huneeus, Federico and Richard Rogerson, "Heterogeneous paths of industrialization," *Review of Economic Studies*, 2024, 91 (3), 1746–1774.

Juhász, Réka, Nathan Lane, and Dani Rodrik, "The new economics of industrial policy," *Annual Review of Economics*, 2023, 16.

Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie, "Beyond balanced growth," *The Review of Economic Studies*, 2001, 68 (4), 869–882.

Kovak, Brian K, "Regional effects of trade reform: What is the correct measure of liberalization?," *American Economic Review*, 2013, 103 (5), 1960–1976.

Kuznets, Simon, "Modern economic growth: findings and reflections," *The American economic review*, 1973, 63 (3), 247–258.

Kuznets, Simon Smith, *Economic growth and structure: selected essays*, Yale University Press, 1966.

- Lagerlöf, Nils-Petter**, "The Galor–Weil model revisited: A quantitative exercise," *Review of Economic dynamics*, 2006, 9 (1), 116–142.
- Lewis, William Arthur**, "Economic development with unlimited supplies of labour," *The Manchester School*, 1954, 22 (2), 139–191.
- Matsuyama, Kiminori**, "Agricultural productivity, comparative advantage, and economic growth," *Journal of economic theory*, 1992, 58 (2), 317–334.
- , "A simple model of sectoral adjustment," *The Review of Economic Studies*, 1992, 59 (2), 375–387.
- Melitz, Marc J**, "The impact of trade on intra-industry reallocations and aggregate industry productivity," *econometrica*, 2003, 71 (6), 1695–1725.
- Moro, Alessio, Solmaz Moslehi, and Satoshi Tanaka**, "Does home production drive structural transformation?," *American Economic Journal Macroeconomics*, 2017, pp. 116–146.
- Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny**, "Industrialization and the big push," *Journal of political economy*, 1989, 97 (5), 1003–1026.
- Ngai, L Rachel and Christopher A Pissarides**, "Structural change in a multi-sector model of growth," *American economic review*, 2007, 97 (1), 429–443.
- Nunn, Nathan**, "Historical development," *Handbook of economic growth*, 2014, 2, 347–402.
- Parente, Stephen L and Edward C Prescott**, *Barriers to riches*, MIT press, 2002.
- Prebisch, Raul**, "Commercial policy in the underdeveloped countries," *The American economic review*, 1959, 49 (2), 251–273.
- Restuccia, Diego and Richard Rogerson**, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic dynamics*, 2008, 11 (4), 707–720.
- Rodrik, Dani**, "Premature Deindustrialization," *Journal of Economic Growth*, 2016, 21 (1), 1–33.

Rostow, Walt Whitman, “The stages of economic growth,” *The economic history review*, 1959, 12 (1), 1–16.

Ruggieri, Alessandro, Nezih Guner, and Tristany Armangué-Jubert, “Labor Market Power and Development,” Available at SSRN 4656579, 2023.

—, **Tristany Armangué-Jubert, and Tancredi Rapone**, “Firm Dynamics, Monopsony, and Aggregate Productivity Differences,” *Monopsony, and Aggregate Productivity Differences**(June 30, 2024), 2024.

Ruggles, Steven, Lara Cleveland, Rodrigo Lovaton, Sula Sarkar, Matthew Sobek, Derek Burk, Dan Ehrlich, Quinn Heimann, Jane Lee, and Nate Merrill, “Integrated Public Use Microdata Series, International: Version 7.6,” 2025. [dataset].

Sposi, Michael, Kei-Mu Yi, and Jing Zhang, “Trade integration, global value chains, and capital accumulation,” *IMF Economic Review*, 2021, 69 (3), 505–539.

Vollrath, Dietrich, *Fully Grown: why a stagnant economy is a sign of success*, University of Chicago Press, 2020.

Wanamaker, Marianne H, “Industrialization and fertility in the nineteenth century: Evidence from South Carolina,” *The Journal of Economic History*, 2012, 72 (1), 168–196.

A Proofs

Proof of Proposition 1. We show that the optimal education decision e_t solves the same condition in both regimes of the model.

Regime I: When $w_t \geq \tilde{z}$, the subsistence constraint is slack, and the household chooses fertility and education jointly. Substituting the optimal fertility rule $n_t = \gamma/(\tau + e_t)$ into utility, the problem reduces to:

$$\max_{e_t} \ln \left(\frac{\gamma}{\tau + e_t} \cdot \frac{e_t + \rho\tau}{e_t + \rho\tau + g_t} \right).$$

Taking the first-order condition yields:

$$-\frac{1}{\tau + e_t} + \frac{1}{e_t + \rho\tau} - \frac{1}{e_t + \rho\tau + g_t} = 0,$$

which after algebraic manipulation gives:

$$e_t = \sqrt{g_t \tau (1 - \rho)} - \rho\tau.$$

Regime II: When $w_t \in (\bar{c}, \tilde{z})$, the subsistence constraint binds. The household no longer chooses n_t freely, but education is still a choice variable. In this case, the constraint determines fertility:

$$n_t = \frac{1 - \bar{c}/w_t}{\tau + e_t},$$

and utility becomes:

$$u(e_t) = \text{const} + \gamma \ln \left(\frac{1 - \bar{c}/w_t}{\tau + e_t} \cdot \frac{e_t + \rho\tau}{e_t + \rho\tau + g_t} \right).$$

This again reduces to:

$$\max_{e_t} \ln \left(\frac{1}{\tau + e_t} \cdot \frac{e_t + \rho\tau}{e_t + \rho\tau + g_t} \right),$$

as the multiplicative constants do not affect the first-order condition. Taking the derivative gives the same FOC as in Regime I:

$$-\frac{1}{\tau + e_t} + \frac{1}{e_t + \rho\tau} - \frac{1}{e_t + \rho\tau + g_t} = 0.$$

Solving as before, we find:

$$e_t = \sqrt{g_t\tau(1 - \rho)} - \rho\tau.$$

Since education must be non-negative, we conclude that in both regimes, the household chooses:

$$e_t^* = \max \left\{ 0, \sqrt{g_t\tau(1 - \rho)} - \rho\tau \right\}.$$

□

Proof of Proposition 2. Let $R_t = w_t(1 - n_t(e_t + \tau))$ denote effective income per capita, where the wage is determined by

$$w_t = \alpha \left(\frac{1 - \beta}{\bar{r}} \right)^{\frac{1 - \beta}{\beta}} h_t^\alpha \left(\frac{A_{mt}X}{L_{mt}} \right)^{1 - \alpha},$$

and human capital evolves as

$$h_{t+1} = \frac{e_t + \rho\tau}{e_t + \rho\tau + g_t}, \quad \text{with} \quad g_t = \sum_{i \in \{a, m, s\}} s_{it} g_{i,t}, \quad g_{i,t} = (e_t + \rho\tau) a_{i,t},$$

and $a_{i,t} = \min\{\theta L_t, \bar{a}_i\}$. Households choose fertility according to

$$n_t(e_t + \tau) = \begin{cases} \gamma, & \text{if } w_t \geq \tilde{z} = \bar{c}/(1 - \gamma), \\ 1 - \bar{c}/w_t, & \text{if } w_t \in (\bar{c}, \tilde{z}), \\ 0, & \text{if } w_t \leq \bar{c}. \end{cases}$$

Then the economy evolves in one of the following regimes:

- **Regime I (Modern Growth Regime):** If $w_t \geq \bar{z}$, then $n_t(e_t + \tau) = \gamma$ and $R_t = w_t(1 - \gamma)$. Effective income per capita grows as long as the growth rate of technology in the manufacturing sector exceeds the growth of manufacturing labor.
- **Regime II (Subsistence-Constrained Regime):** If $w_t \in (\bar{c}, \bar{z})$, then $1 - n_t(e_t + \tau) = \bar{c}/w_t$ and $R_t = \bar{c}$. Effective income per capita is constant.
- **Regime III (Collapse Regime):** If $w_t \leq \bar{c}$, then $n_t = 0$ and $L_{t+1} = 0$. Population collapses unless this regime is avoided. We assume parameters (e.g., large X) are such that $w_t > \bar{c}$ always.

Now we need to prove that a transition from Regime II to Regime I can happen if and only if the household chooses a positive education level e . In Regime II, the wage evolves according to:

$$w_t = \alpha \left(\frac{1 - \beta}{\bar{r}} \right)^{\frac{1-\beta}{\beta}} h_t^\alpha \left(\frac{A_{mt} X}{L_{mt}} \right)^{1-\alpha},$$

and the corresponding growth rate is:

$$\frac{\Delta w_t}{w_t} = \alpha \cdot \frac{\Delta h_t}{h_t} + (1 - \alpha) \left(\frac{\Delta A_{mt}}{A_{mt}} - \frac{\Delta L_{mt}}{L_{mt}} \right).$$

We examine each component:

1. Human capital growth. The law of motion implies:

$$h_t = \frac{e_{t-1} + \rho\tau}{e_{t-1} + \rho\tau + g_{t-1}}, \quad h_{t+1} = \frac{e_t + \rho\tau}{e_t + \rho\tau + g_t}.$$

Hence,

$$\frac{\Delta h_t}{h_t} = \left(\frac{e_t + \rho\tau}{e_t + \rho\tau + g_t} \cdot \frac{e_{t-1} + \rho\tau + g_{t-1}}{e_{t-1} + \rho\tau} \right) - 1.$$

This is strictly positive for $e_t > e_{t-1}$, and weakly increasing in e_t .

2. Technology growth. The sectoral technology grows at:

$$\frac{\Delta A_{mt}}{A_{mt}} = g_{m,t} = (e_t + \rho\tau) \cdot a_{m,t}, \quad \text{with } a_{m,t} = \min\{\theta L_t, \bar{a}_m\}.$$

This is strictly increasing in e_t .

3. Labor growth. The fertility rule gives:

$$n_t = \frac{1 - \bar{c}/w_t}{e_t + \tau} \Rightarrow \frac{\Delta L_t}{L_t} = n_t - 1 = \frac{1 - \bar{c}/w_t}{e_t + \tau} - 1.$$

Thus, labor growth is decreasing in e_t , which increases wage growth.

Define the function:

$$\Gamma(e_t, w_t) := (e_t + \rho\tau)a_{m,t} - \left(\frac{1 - \bar{c}/w_t}{e_t + \tau} - 1 \right),$$

so that:

$$\frac{\Delta w_t}{w_t} = \alpha \cdot \frac{\Delta h_t}{h_t} + (1 - \alpha) \cdot \Gamma(e_t, w_t).$$

Now evaluate $\Gamma(0, w_t)$:

$$\Gamma(0, w_t) = \rho\tau a_{m,t} - \left(\frac{1 - \bar{c}/w_t}{\tau} - 1 \right),$$

and multiply both sides by τ and rearrange:

$$\Gamma(0, w_t) > 0 \Leftrightarrow \rho\tau^2 a_{m,t} + \tau > 1 - \frac{\bar{c}}{w_t}.$$

Thus, if the parameter condition $\rho\tau^2 a_{m,t} + \tau < 1$ holds, and $w_t > \bar{c}$, then:

$$\Gamma(0, w_t) > 0 \Rightarrow \Gamma(e_t, w_t) > \Gamma(0, w_t) > 0, \quad \text{for all } e_t > 0.$$

Moreover, since $\frac{\Delta h_t}{h_t} \geq 0$ for $e_t > 0$, it follows that:

$$\frac{\Delta w_t}{w_t} > 0.$$

Hence, wages grow over time as long as $e_t > 0$, and by continuity, will eventually satisfy $w_t \geq \tilde{z}$, completing the transition to Regime I. \square

B Empirical Analysis

B.1 Fertility and Structural Change.

In this section I provide further evidence on the cross country relationship between fertility and structural change. In Figure 1 I show that fertility rates are, on average, strongly related to a country's composition within its manufacturing sector. This relationship also holds if we disaggregate the data at the country/year level. We can see this by running the following regression:

$$tfr_{c,t} = \alpha_c + \delta_t + \beta \left(\frac{\%Man_{c,t}}{1 - \%Agr_{c,t}} \right) + \Gamma' X_{c,t} + \varepsilon_{c,t}$$

Where X is a vector of controls including log of gdp per capita and the agricultural share of employment. Table 2 shows that the manufacturing share of the non-farm labor force is a strong negative predictor of a country's total fertility rates even when controlling for gdp per capita, the agricultural employment share and country and year fixed effects.

The estimated effect of the manufacturing employment share on total fertility is not only statistically significant, but also economically meaningful. When controlling for log of gdp per capita, the agricultural employment share, country and year fixed effects, a 1pp increase in the manufacturing employment share of non-farm labor is associated with a reduction of total fertility of 0.45 or 14% of the sample average. To understand whether the relationship between manufacturing and total fertility rates is driven by unobserved variation within

Table 2: OLS Regressions

	(1)	(2)	(3)	(4)	(5)
% Man./ (1-% Agr.)	-2.522* (-2.03)	-3.352*** (-4.22)	-3.352*** (-4.22)	-4.204*** (-4.80)	-4.025*** (-4.59)
Log GDP p.c.		-0.790*** (-5.60)	-0.790*** (-5.60)	-0.763*** (-5.35)	0.399*** (3.91)
% Agr.		1.708* (2.45)	1.708* (2.45)	1.695* (2.41)	1.759** (2.72)
N	5249	5249	5249	5249	5249
R2	0.0161	0.644	0.644	0.666	0.964
Country FE	no	no	no	no	yes
Year FE	no	no	no	yes	yes

t statistics in parentheses

Dependent variable is total fertility rate. Standard errors are clustered at the country level.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

countries and years, we can use microdata from IPUMS to quantify the fertility differentials across the sectors of the economy within a country for multiple samples/years. The total sample includes 46 million women aged between 15 and 45 from 89 low and middle income countries, which comprises all samples available through IPUMS between the years of 1970 and 2020 from low and middle income countries. When a woman is not working, I assign her the sector of the husband. When a woman is unmarried and does not work or when both the husband and the woman have no reported sector, I am forced to drop them, which leaves me with a sample of over 40 million women in 89 different low and middle income countries. Using this dataset, I regress the woman's total fertility rate on a wide array of controls to remove selection driven by age, rural-urban status, country, time period and their educational attainment. I then estimate the residuals from model 13 and average them by sector.¹¹

$$nchild_{i,c,t} = \alpha_{a,c,t} + \delta_{e,c,t} + \gamma_{n,c,t} + \xi_{u,c,t} + \varepsilon_{i,c,t} \quad (13)$$

11. This is equivalent to showing the sector fixed effect in a model with no constant so that the reference category is the sample average.

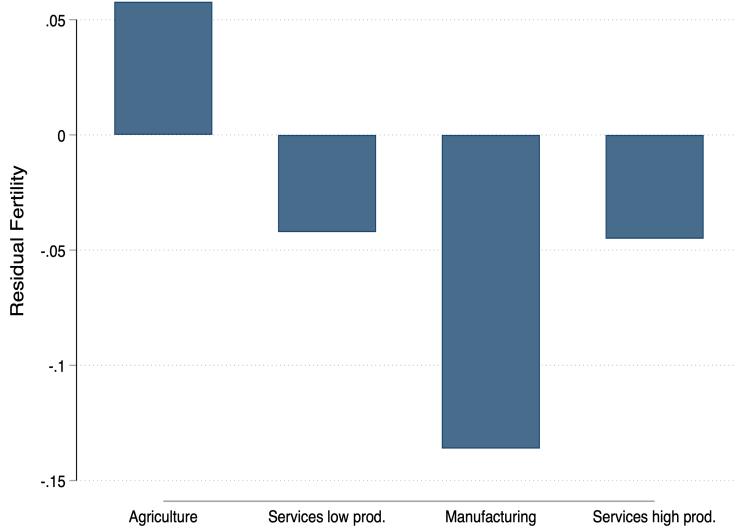


Figure 12: Average residuals by income group and sector from a regression of total fertility on country-year fixed effects, urban dummy and age fixed effects. The data comprises samples of women between 15 and 45 from 89 low and middle income countries. Where women do not work or have missing sector information, I assign them to the husband's sector. Source: IPUMS International Survey Data.

Where i, c, t, a, e, n, u refer to individual, country, year, age, education level (detailed classification), employment status and urban status respectively. The regression shown in equation 13 includes 4027 fixed effects, controlling for time and country varying variables which could affect fertility. This exercise shows that there are strong differences in average fertility rates across the sectors of the economy. Not surprisingly, women working in agriculture have the highest residual fertility rates. However, consistently with the cross country data in Figure 1 there is heterogeneity within the non-farm sector with manufacturing women having the lowest residual fertility rates. Considering the sample average fertility rate of 2.1, this is approximately a 5% difference.

B.2 Estimating Trade shocks

This section builds on the descriptive evidence presented in Section 2 by using IPUMS microdata to run shift-share regressions of employment shares in

different sectors against a regional trade shock. The data encompasses 31 countries, each generally available for at least three different samples between the years from 1970 to 2020, when the largest trade liberalizations took place. These samples are nationally representative and contain from 1 to 10 % of the population (see Ruggles et al. (2025) for more information on this data).

The empirical methodology used is very similar to that in Autor et al. (2013), where the shares of employment in trade affected sectors will provide variation in the incidence of trade at the sub-national level. The basic idea is that aggregate changes in imports, which are assumed to be exogenous to regional outcomes, affect regions differently depending on the shares of people working in industries affected by the imported goods. Aggregating these changes in traded imports, and weighting by the regional shares, provides regional variation in the incidence of trade, which simultaneously solves two problems: it increases the observational units in a setting where data on trade is only available at the country level, and provides an argument for identification as long as regional demand and supply shocks are exogenous to national level trade patterns. Contrary to Autor et al. (2013) however, the identification I employ relies on the exogeneity of global - not national - trade patterns to the evolution of regional outcomes. This is due to trade data limitations as I cannot obtain imports for most of the countries in my sample covering the entire time period, hence I use global changes in trade flows as the shifter for all countries.¹² This type of exercise has previously been possible only with detailed administrative data, as mapping workers to a product code classification system such as SITC4 or HS4 requires precise information on the sector in which the worker was active which goes beyond the standard 2-digit classification available in harmonized survey data (e.g. Autor et al., 2013, 2014; Dix-Carneiro and Kovak, 2017). To circumvent this issue, I build a crosswalk linking over 10 thousand textual descriptions of industries present in the IPUMS International data,¹³ to SITC4 codes. I am hence able to accomplish to construct the classical shock measure employed in the literature without requiring administrative data containing 4-digit industry classifiers. This means that I can explore the relationship between

12. In principle this should strengthen the identification claim as global trade patterns are more credibly exogenous to the changes in regional outcomes of particular countries.

13. This is IPUMS variable IND, which has varying codes by country/year sample.

trade shocks and regional outcomes at the *global* level, which would not be possible unless one had administrative data from *all* of the 31 countries which provide survey data through IPUMS International.¹⁴ The drawback is that IPUMS international generally provides samples comprising between 1% and 10% of the population. If one had access to the full universe, it would be possible to disaggregate regions at the level of commuting zones as done in Autor et al. (2013) and Dix-Carneiro and Kovak (2017), thereby gaining additional degrees of freedom. In order to not run into small sample issues when disaggregating the labor force at the SITC4 level, I conduct empirical exercises at the Level-1 administrative region, which is slightly larger than what is typically considered a region in most of the literature.¹⁵

Table 3: Industry Descriptions and Corresponding SITC4 Codes

Sample	Industry Description	SITC4 Code
Ghana 2000	Agriculture and animal farming combined	0111, 0112, 0113,...
Ghana 2010	Agriculture and hunting	0111, 0112, 0113,...
Ethiopia 1984	Agriculture and livestock production	0111, 0112, 0113,...

Table 3 illustrates one of the main contributions of the paper: the construction of a crosswalk mapping textual industry descriptions to SITC4 codes. In the raw (non-harmonized) IPUMS data industries are reported in much greater detail than the harmonized variables, with considerable breakdown in the tradeable goods sector. The challenge is that these industry descriptions vary not only by country but also by sample. While the core classification may remain fixed, e.g. one category for "plants" and "animals" as in Table 3, the textual description of this category can still vary across samples within the same country. Using the ChatGPT API I map over 10 thousand text based industry classifications to SITC4 codes. This crosswalk is consistent across API calls.¹⁶

Using this dataset I construct regional trade shocks as follows:

$$\text{shock}_{i,c} = \sum_j s_{i,c,<2000,j} (M_{>2000,j} - M_{<2000,j}) \quad (14)$$

14. I am currently working on expanding this to over 70 countries.

15. For instance Autor et al. (2013) use commuting zones as regions whereas in the US Level-1 regions correspond to the states.

16. Creating the crosswalk several times shows similar results.

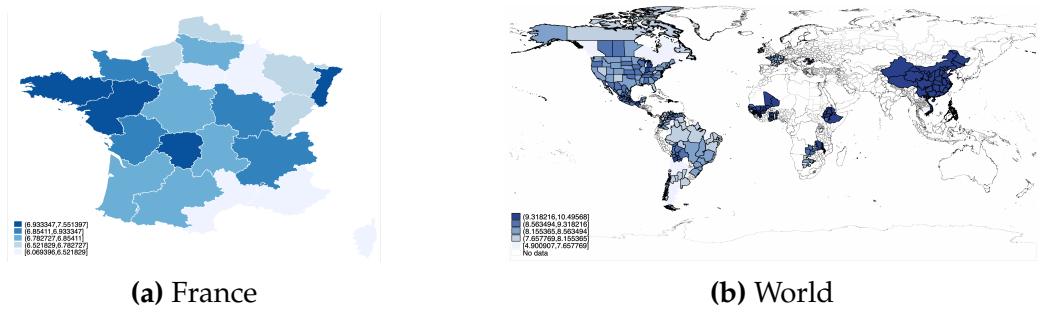


Figure 13: Trade shocks across the world. The figure shows the estimated trade shocks following 14 for all the countries in the sample (right) and zooms in on France (left). Source: Author's calculations using IPUMS microdata.

Where $s_{i,c,<2000,j}$ is the share of employment in industry j and region i of country c before the year 2000 and $M_{>2000,j} - M_{<2000,j}$ is the change in total imports of product j at the global level between the years from 1970 to 2000 and 2000 to 2020. Figure 13 shows the spatial distribution of these shocks in the countries available in this sample. Zooming in on a particular country such as France (a country the author is familiar with), shows that the the distribution of these shocks matches what one would expect: highly diversified economies, such as the Île-de-France, and service focused regions, such as the southern tourist economy, experience lower trade shocks whereas the region focused on the production of goods such as the industrial North-West, experience larger shocks. This methodology for constructing trade shocks is similar to that in Autor et al. (2013) but varies from that used by Autor et al. (2014) and Dix-Carneiro and Kovak (2017) in that they use variation in tariff changes at the industry level as a shifter instead of the amount of traded goods. This is preferable as tariff changes are a more relevant measure of trade protection than imports which also reflect local demand and supply shocks in the country of origin (Kovak, 2013). However, the global scope of this analysis means that constructing variation in tariff changes at the industry level would in practice be extremely cumbersome due to the widespread use of non-tariff barriers in the pre-2000 period.¹⁷ I then run the following regression model:

17. I am exploring data sources to find a way to automate this.

$$y_{i,c,>2000} - y_{i,c,<2000} = \alpha_c + \beta \ln(\text{shock}_{i,c}) + \epsilon_{i,c}$$

Where y is a regional outcome variable, α is a country fixed effect and ϵ is a classical error term. Guided by the stylized facts I consider five outcome variables: employment shares in all three sectors plus fertility. I further divide employment in the services sector by the high-productivity FIRE services and group the rest in the low-productivity category.¹⁸

Table 4: OLS Regressions

	(1)	(2)	(3)	(4)	(5)
	Manufacturing	Serv. High	Serv. Low	Agriculture	Fertility
(mean) ln_shock	-0.0167*	-0.000955	0.0225**	-0.00671	0.0117
	(-2.30)	(-0.66)	(2.73)	(-1.13)	(0.29)
N	460	460	460	460	448
R2	0.692	0.828	0.539	0.577	0.663
Country FE	yes	yes	yes	yes	yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results are consistent with the stylized facts and quantitatively meaningful. A doubling of trade exposure (i.e., a one-unit increase in $\log(\text{shock})$) is associated with a 1.88 percentage point decline in manufacturing employment, a statistically significant effect at the 5% level. This represents a 13.4% reduction relative to the average manufacturing share of 14% in the sample. The shock has no significant effect on high-productivity services, while it leads to a 2.67 percentage point increase in low-productivity service employment, also statistically significant at the 5% level. Agricultural employment falls by about 0.69 percentage points, although this effect is not statistically significant, and the average agricultural share remains high (above 30%). The estimated effect on fertility is positive but small and statistically insignificant, with a point estimate of 0.0306. Overall, the results point to a reallocation of labor away from manufacturing and toward lower productivity services in response to trade exposure, with limited direct effects on agriculture or fertility.

18. FIRE comprises finance, insurance real estate, health care and education.