Service-Led or Service-Biased Growth?

Equilibrium Development Accounting Across Indian Districts.

Tianyu Fan, Michael Peters, and Fabrizio Zilibotti

Yale University*

December 2020

Abstract

In many developing countries, employment declines in the agricultural sector, increases in the service sector, and there is limited industrialization. Is the service sector an engine of growth or is its expansion a mere consequence of productivity growth in goods-producing industries? What are the welfare and distributional effects of different paths of economic growth? To address these questions, we construct and estimate a spatial equilibrium model with nonhomothetic preferences. The estimated model lends itself to a quantitative assessment of the heterogeneous welfare effects of different sources of economic growth. We apply our methodology to India. We find that productivity growth in consumer services is an important driver of structural transformation and for improving living standards, although the growth is highly skewed toward high-income households living in urban districts.

1 Introduction

The major industrialized countries have undergone a similar pattern of structural transformation. At an early stage, a growing industrial sector drew labor from a declining agricultural sector. At a later stage, the employment shares of both agriculture and manufacturing fell while the service sector became the main source of employment growth. This patterns fits the experience of the Western economies. Even in China, industrial employment as a share of GDP has started declining since 2014 at the expense of a growing service sector. However, in other parts of the globe, economic development appears to have taken a different turn. Over the last four decades, the share of manufacturing jobs has barely grown in most developing countries, including fast-growing economies like India and sub-Saharan Africa. There, structural transformation has taken the form of a shift from agriculture to services (see, e.g., Celasun and Gruss (2018)).

To many scholars (e.g., McMillan and Rodrik (2011)), the absence of industrialization is a concern. The traditional view (e.g., Baumol (1967)) is that technical progress in manufacturing is the engine of productivity growth. In contrast, the expansion of the service sector is a by-product of economic development: as countries

*This study is part of a broader research project also involving Philippe Aghion and Robin Burgess. We thank Fabian Eckert, Reto Foellmi, Doug Gollin, Peter Klenow, Samuel Kortum, Rachel Ngai, Richard Rogerson, María Sáez Martí and seminar participants at the Cowles Macro Conference, the STEG Workshop, Penn State University, Peking University, RIDGE, the University of Sankt Gallen, and Yale University. We also thank Sarah Moon, Shengqi Ni, Pariroo Rattan and Huihuang Zhu for excellent research assistance. Financial support from SNSF project "Inequality, Cultural Transmission, and Human Capital Accumulation" (grant 100018 165616) is gratefully acknowledged. Michael Peters thanks the "Minnesota Opportunity and Inclusive Growth Institute" for the hospitality.

grow richer, consumers spend an increasing share of their income on *luxuries*, and their demand triggers a growing provision of services. However, there is little or no productivity growth in service industries.

This pessimistic view of the service sector is not undisputed. Buera and Kaboski (2012) emphasize the importance of the (demand-driven) growth of a skill-intensive service industry in the post-1950s US economy. Hsieh and Rossi-Hansberg (2019) argue that in more recent years ICT has triggered an industrial revolution in the service sector that has been a major source of productivity growth in mature economies during the last few decades.[1] Their view is echoed by Eckert et al. (2020) who argue that productivity growth in ICT service industries is the main cause of the growing urban-rural gap in the United States.

Equilibrium Development Accounting. A hurdle to resolving this controversy is measurement: it is difficult to quantify productivity in many service industries. In this paper, we provide a novel structural methodology to estimate productivity in the service sector and to quantitatively assess its importance in the development process. Our approach is in the vein of the development accounting literature.[2] We do not attempt to explain the determinants of productivity growth but propose a method to measure sectoral productivity across subnational geographical units (such as districts, provinces, MSAs, etc.) and over time. The estimation is disciplined by a theory that stands on two building blocks: (i) nonhomothetic preferences, and (ii) a spatial multisector equilibrium model with inter-regional trade where firms have heterogeneous productivities in different locations.

Spatial Equilibrium. We assume that labor is the only productive factor and that people can work in four sectors of activity: agriculture, manufacturing, producer services (PS), and consumer services (CS). People have heterogeneous productivities (human capital). Labor is perfectly mobile across industries although immobile across locations (*districts*). Thus, there is a single wage per effective unit of labor in each district although wages differ across districts. These assumptions are extreme but can be relaxed by introducing non-prohibitive labor mobility frictions across both sectors and districts. Consumers' preferences are defined over three final items: *food*, *industrial goods*, and *CS*.

Because the service sector is broad and heterogeneous, its growth may have different implications for consumers and producers. Part of the services produced improve households' access to consumption goods (e.g., restaurants or retail) or enter directly their consumption basket (e.g., leisure services). This is what we call CS. Other services (PS) are predominantly inputs to the production of goods, mostly in the industrial sector. These include, among others, business services, corporate law services, and part of transport services. We model CS as directly entering the households' consumption basket and PS as inputs to the production of industrial goods.[3]

We assume that food and industrial goods are traded across districts. In contrast, services must be purchased locally. There is an important difference, though: CS must be consumed locally, so their local productivity impacts directly on the price and ultimately on the availability of CS in each market. In contrast, PS are embedded in industrial goods and their value-added is tradable.

To avoid overly stark predictions about regional specialization, we assume that each district produces a differ-

[1]The authors argue that ICT has made it possible for the most efficient service firms to replicate their superior production process in multiple locations close to consumers. This has allowed large productivity gains in the provision of services.

[2]See, e.g. Caselli (2005) and Hall and Jones (1999). This literature postulates aggregate production functions and uses information on the accumulation of productive factors to fit the data. Our methodology is closer to the structural development accounting of Gancia et al. (2013), who exploit the restrictions imposed by an equilibrium model to identify sectoral productivities. However, in their model, spatial interactions are limited to technological spillovers. Moreover, they only use national data while this paper uses granular information about employment and income in a novel fashion.

[3]If PS were the only source of productivity growth, one might argue there is nothing new under the sun. Industrial technology would simply be turning more intensive in services over time. Ultimately, a properly defined goods-producing sector would continue to be the main growth engine.

entiated variety of agricultural and industrial goods. Preferences over varieties are represented by a standard CES aggregator. This assumption ensures that all consumers demand all regional varieties.

Price Independent Generalized Linear (PIGL) Preferences. The demand side of the economy is populated by consumers endowed with nonhomothetic preferences belonging to the PIGL class. This class of preferences was first introduced by Muellbauer (1976) and has been recently popularized in the literature on growth and structural change by Boppart (2014) and Alder et al. (2019). In general, PIGL preferences do not admit an analytical representation of the utility function but can be described by an indirect utility function. Moreover, they have transparent aggregation properties: the choice of a set of agents endowed with PIGL preferences facing a common price vector can be rationalized as the choice of a representative agent. In our model, this guarantees that we can derive district-level demand functions since, within each district, all agents face the same price vector. Moreover, it allows for a tractable characterization of the spatial equilibrium even though factor prices vary across space.

Positive and Normative Focus. The goal of our analysis is twofold. First, we quantify the extent to which the growth of the service sector is either a source or a consequence of the economic development process. Second, we draw quantitative inferences about the welfare effects associated with different sources of growth and structural change. When agents have nonhomothetic preferences and live in different locations characterized by different provision of local services, the growth of different sectors benefits people differentially–rich versus poor as well as urban versus rural residents. We view this as important, as growth theory is often criticized by development economists for glossing over the distributional implications of economic growth.

An Example. To clarify our conceptual framework, consider the following example. Imagine two small districts R (rich) and P (poor), which are part of a large multi-district economy, and abstract for simplicity from PS. Let A_{rs} denote labor productivity in district r and sector $s \in \{F, M, CS\}$. Consider two polar opposite scenarios.

In the first scenario, $\{A_{RF}, A_{RM}\} = \{\lambda A_{PF}, \lambda A_{PM}\}$, for $\lambda > 1$, i.e., the productivity of the tradable sectors are larger in R than in P by the factor λ. Instead, the productivity of CS is the same in both districts, $A_{RCS} = A_{PCS}$. In equilibrium, workers in district R earn a higher wage. If food is a necessity and CS is a luxury, consumers in R will then spend a higher share of their income on CS while consumers in P spend a higher share of their income on food. Since CS is nontradable, district R will also have a larger employment share in CS, while P will specialize in the production of goods.[4] In this scenario, spatial differences in expenditure and employment shares are entirely driven by income effects in demand. We refer to this case as *service-biased growth.*

The second scenario is one in which productivity in the tradable sectors are identical in R and P, whereas $A_{RCS} > A_{PCS}$. As before, district R is richer, has a larger service sector, and consumers spend a smaller share of their budget on food. However, in this case, any difference stems from a technological gap in the consumer service sector, such as R having a more efficient retail sector. We refer to this scenario as *service-led* growth.

Although the example is a spatial difference, the same argument applies to the analysis of a given district at two points in time. Under *service-biased growth*, the growth of the service sector would be entirely a consequence of the productivity growth in the goods-producing sectors. Under *service-led growth*, productivity growth in the service sector would be the sole cause of productivity growth and structural change.

Identification. Our model allows us to disentangle the relative importance of service-biased versus service-led growth by estimating the variation in sectoral productivities across space and time. Conditional on a set of preference

[4] Because food and industrial goods are tradable, whether P specializes in agriculture or industry depends on its comparative advantages.

and technology parameters, labor productivities $\{A_{rF}, A_{rM}, A_{rPS}, A_{rCS}\}$, and regional labor endowments $\{L_r\}$, the model yields a unique equilibrium vector of wages $\{w_r\}$ and labor allocations $\{L_{rF}, L_{rM}, L_{rPS}, L_{rCS}\}$. Conversely, if we have data for the allocation of labor across sectors in each district and for the local real wage, we can retrieve a unique set of labor productivities $\{A_{rF}, A_{rM}, A_{rPS}, A_{rCS}\}$. After estimating the productivities, we run counterfactual experiments.

The analysis hinges on a set of structural parameters that we also estimate. Among them, the one governing the income elasticities of demand is especially important because it determines the strength of income effects. We estimate this parameter from Engel curves using microdata within districts. Its estimates are precise and robust.

Application to India. We apply our methodology to India. India is a fast-growing economy, with an average annual growth rate of 4.2% during 1987–2011. In this period, the employment share of agriculture declined while that of manufacturing increased only marginally. The lion's share of the process of structural change has then been a shift from agriculture to services.

Our estimation exploits individual geolocalized data on consumption and employment. The main data source is the Employment Schedule of the National Sample Survey (NSS) covering 400 Indian districts between 1987 and 2011. The NSS provides us with information on consumption per capita that we use as a proxy for district-level real wages. In order to split the employment in the service sector into CS and PS, we use data from the EC, a complete census of firms in India, and from the Survey of the Service Firms of India which contains information on whether services are bought by firms or consumers.

We separately estimate cross-sectional distribution of productivities for the years 1987 and 2011. Then, we chain them so that the average growth rate in the model matches the national account statistics for GDP per capita growth in India.[5] The results are interesting in several respects. At the spatial level, there are large productivity differences across districts in both industry and CS. In the CS sector, the productivity gap is especially pronounced between the most urbanized districts and the rest of the economy. Urban locations have higher employment shares in services not only because their inhabitants are richer but also because services are more productively provided.

Distributional Implications of Service-Led Growth. Next, we use our model to quantify the importance of sectoral productivity growth for India's economic development between 1987 and 2011. We find that productivity growth in agriculture and CS are the main drivers of welfare changes in between 1987 and 2011. These effects are largely heterogeneous, in spite of perfect labor mobility across sectors. For poorer households living in rural districts, most benefits of growth accrue from productivity growth in agriculture. In contrast, productivity growth in CS is the main source of welfare gains for richer households, especially those living in urbanized districts. The residents in the top quintile of urbanization (by district) would have been better off taking a 40% income cut in 2011 than moving back to the productivity distribution that the CS sector had in 1987.

Because of nonhomothetic preferences, welfare gains also vary widely across the income distribution ladder. The richest households would have been better off giving up the the entire productivity gains in agriculture and industrial production than renouncing those those accruing from technical progress in the CS sector. Taking stock, much of the Indian growth was service-led and, as such, skewed toward the rich.

Productivity growth in CS is also the key driver of structural change. Had productivity in the service sector stagnated, the employment share in agriculture would have barely declined. In other words, productivity in

[5] There are well-known discrepancies between consumption growth in survey data and national account statistics. We could alternatively use the evolution of consumption from survey data as an anchor for growth. This would rescale results without altering the main results.

agriculture has played no role in structural change: if anything, it has marginally slowed down it.

Open Economy. Our main analysis is based on a closed-economy environment. Since India underwent a significant process of international integration, we extend the analysis to account for international trade. We set the parameters of the model to match the empirical increase in the trade flow of goods. We also consider the export of ICT services, which are especially important in India. The results are broadly similar to those in the closed economy.

We also consider a variety of other robustness analyses. Most notably, we allow for the possibility that the development of the service sector is driven by the increasing educational attainment of the labor force. Although we find some evidence for this mechanism, its effects are less important than those of productivity growth in CS.

In conclusion, our estimates indicate that the variations across time and space of productivity in the service sector are an important aspect of the recent economic development in India.

Related Literature. Our paper contributes to the macroeconomic literature on structural transformation including, among others, Nga, Herrendorf, Rogerson, and Valentinyi (2013, 2014, and 2020), Gollin et al. (2014), Hobijn et al. (2019), and Garcia-Santana et al. (2020).[6] Duarte and Restuccia (2010 and 2019), Buera and Kaboski (2012), and Eckert et al. (2020) focus more narrowly–as we do–on the service sector. In particular, Duarte and Restuccia (2010) document large cross-country productivity differences in service industries. Moreover, these differences have fallen over time much less than in agriculture and industry. Their findings suggest an important role for productivity catch-up in developing countries, a conclusion broadly in line with our results.

Earlier papers emphasizing the importance of nonhomothetic preferences in the growth process include Foellmi and Zweimueller (2006, 2008), and Matsuyama (2000). The more recent literature on structural change with nonhomothetic preferences includes, among others, Boppart (2014) and Alder et al. (2019) who, like us, propose generalizations of PIGL preferences class, and Matsuyama (2019) and Comin et al. (2020), who instead use a class of generalized CES preferences related to Sato (2014). The authors show that these preferences can account accurately for the patterns of structural transformation across several countries. In our paper, we use PIGL preferences because their tractable and transparent aggregation properties are especially suitable to the welfare analysis. Exploring different class of preferences would certainly be interesting but we leave it to future research.

The interaction between spatial equilibrium and structural change is studied by Eckert and Peters (2016), although their model has a different focus. To model inter-regional trade linkages, we build on a large body of literature in economic geography, see e.g. Redding and Rossi-Hansberg (2017) or Allen and Arkolakis (2014).

We also contribute to the vast literature on the economic development of India including, among others, , Aghion et al. (2005, 2008), Akcigit et al. (2020), Basu (2008), Basu and Maertens (2007), Foster and Rosenzweig (1996), Foster and Rosenzweig (2004), Goldberg et al. (2010), Kochhar et al. (2006), and Martin et al. (2017). A number of recent papers in this literature more specifically study the role of the of service sector in India. Among them, see Amirapu and Subramanian (2015), Eichengreen and Gupta (2011), Erumban et al. (2019), Ghose (2014), Gordon and Gupta (2005), Majid (2019), Mitra and Ural (2008), Mukherjee (2013), and Singh and Dasgupta (2016). Many of these papers establish useful stylized facts on which our analysis builds.

Road Map. The structure of the paper is as follows. Section 2 lays out the theoretical framework. Section 3 describes the data and the main empirical patterns in India. Section 4 discusses the estimation method. Section 5 discusses the main results. Section 6 performs robustness analysis. Section 7 concludes. The Appendix contains

[6] Herrendorf et al. (2013) draw an important distinction between value-added and final-expenditure approaches. Our estimation exploits data on income and employment rather than on expenditure. Therefore, our analysis ultimately rests on the value-added approach. For instance, if a restaurant serves a meal, the consumer would be purchasing some tradable food and some local CS.

technical details, a description of the data sources, and additional tables and figures.

2 Theory

We consider a general equilibrium environment with R regions. Consumers have preferences over three goods: agricultural goods (F for *food*), industrial goods (G for *goods*), and CS. There is a single factor of production, which is inelastically supplied: labor. While goods and food are tradable, CS are not traded and must be provided locally. All markets are frictionless and competitive.

2.1 Technology and Preferences

Technology. All goods and services are produced with constant returns to scale technologies such that

$$Y_{rst} = A_{rst} H_{rst},$$

where H_{rst} denotes the amount of human capital employed in the production of sector s goods in region r. While we take total productivity in agriculture ,A_{rFt}, and CS, A_{rCSt}, as exogenous, productivity in the industrial sector, A_{rGt}, is endogenously determined. However, in Section 2.4 we show that the equilibrium allocation in the industrial sector implies that $Y_{rGt} = A_{rGt} H_{rGt}$, where A_{rGt} is a function of structural parameters that does not depend on equilibrium prices. Therefore, we can characterize the general equilibrium taking A_{rGt} as given, and then solve for the equilibrium in the industrial sector.

For the traded commodities, we assume that consumers buy a CES aggregate of differentiated regional varieties with an elasticity of substitution σ. Hence, the price indexes of food and industrial goods, which are common across localities due to free trade, are given by

$$p_{Ft} = \left(\sum_{r=1}^{R} p_{rFt}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \quad \text{and} \quad p_{Gt} = \left(\sum_{r=1}^{R} p_{rGt}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}.$$

Preferences. We model consumers' preferences as stemming from the PIGL class. These preferences have two important advantages for us. First, they allow us to parameterize the extent of nonhomotheticity in a flexible way, which we can easily estimate from individual data. Second, they allow us to derive a simple aggregate demand system, which we can also take to the data.

The class of PIGL preferences is characterized by an indirect utility function of the form:

$$V(e, \mathbf{p}) = \frac{1}{\varepsilon} \left(\frac{e}{B(\mathbf{p})} \right)^{\varepsilon} - D(\mathbf{p}), \tag{1}$$

where e denotes total spending and $\mathbf{p}$ the vector of prices of goods. The function $D(\mathbf{p})$ is homogeneous of degree zero while $B(\mathbf{p})$ is homogeneous of degree one. We parameterize the functions B and D as follows:

$$B(\mathbf{p}) = \prod_{s \in \{F,G,CS\}} p_s^{\omega_s} \quad \text{and} \quad D(\mathbf{p}) = (1-\varepsilon) \left(\sum_{s \in \{F,G,CS\}} \tilde{\nu}_s \ln p_s \right),$$

where $\sum_s \omega_s = 1$ and $\sum_s \tilde{\nu}_s = 0$. This yields the indirect utility function

$$V(e, \mathbf{p}) = \frac{1}{\varepsilon}\left(\frac{e}{\prod_s p_s^{\omega_s}}\right)^{\varepsilon} - \sum_s \tilde{\nu}_s \ln p_s. \tag{2}$$

The sectoral expenditure shares associated with $V(e,p)$ (see Section A-1 in the Appendix) are given by

$$\vartheta_s^h(e, \mathbf{p}) = \omega_s + \tilde{\nu}_s \left(\frac{e}{\prod_s p_s^{\omega_s}}\right)^{-\varepsilon}. \tag{3}$$

Equation (3) highlights that consumers' expenditure shares feature both income and price effects. In particular,

$$\frac{\partial \vartheta_s^h(e, \mathbf{p})}{\partial \ln e} = -\varepsilon \tilde{\nu}_s \times \left(\frac{e}{\prod_s p_s^{\omega_s}}\right)^{-\varepsilon},$$

that is the expenditure share on goods from sector s is increasing in income (i.e. the good is a luxury) if $\tilde{\nu}_s < 0$. If food is a necessity and CS are a luxury, then, $\tilde{\nu}_F > 0$ and $\tilde{\nu}_{CS} < 0$. In this case, the expenditure shares in food CS are, respectively, decreasing and increasing in e and $\lim_{e\to\infty} \vartheta_A^h(e, \mathbf{p}) = \omega_F$ and $\lim_{e\to\infty} \vartheta_{CS}^h(e, \mathbf{p}) = \omega_{CS}$. The strength of this nonhomotheticity is governed by the parameter ε, which, with a slight abuse of notation, we also refer to as the "income elasticity" as a shorthand.

2.2 Heterogeneity and Aggregate Demand

PIGL preferences admit a tractable aggregation. Suppose individuals have heterogeneous abilities, and let w_{rt} denote the wage per efficiency unit of labor. Then, the income and expenditure for individual h is given by $e_{rt}^h = q^h w_{rt}$, where q^h is the number of efficiency units of labor. Let $F_{rt}(q)$ denote the distribution function of q in region r at the t. Empirically, we will relate the spatial variation in the distribution of q to observable differences in human capital. Using (3), the *aggregate* spending share on goods in sector s in region r is then given by:

$$\vartheta_{rs}(w_{rt}, \mathbf{p_{rt}}) \equiv \frac{L_{rt}\int \vartheta_s^h(q w_{rt}, p_{rt})\, q w_{rt} dF_{rt}(q)}{L_{rt}\int q w_{rt} dF_{rt}(q)} = \omega_s + \nu_{rs}\left(\frac{E_{rt}[q]\, w_{rt}}{p_F^{\omega_F} p_G^{\omega_G} p_{CSr}^{\omega_{CS}}}\right)^{-\varepsilon}, \tag{4}$$

where

$$\nu_{rs} \equiv \frac{E_{rt}\left[q^{1-\varepsilon}\right]}{E_{rt}[q]^{1-\varepsilon}} \tilde{\nu}_s. \tag{5}$$

Comparing (4) with (3) shows that the aggregate demand is isomorphic to that of a representative consumer in region r, who earns the average income $E_{rt}[q]\, w_{rt}$ and has the inequality-adjusted preference parameter ν_{rs} instead of the primitive parameter $\tilde{\nu}_s$.

In general, ν_{rs} depends on the local income distribution. The analysis simplifies further if we assume that q follows a Pareto distribution with c.d.f. $F_{rt}(q) = 1 - \left(\underline{q}_{rt}/q\right)^{\zeta}$, with a region-invariant tail parameter ζ. Then, $E_r[q] = \frac{\zeta}{\zeta-1}\underline{q}_r$ and $E_r\left[q^{1-\varepsilon}\right] = \frac{\frac{\zeta}{1-\varepsilon}}{\frac{\zeta}{1-\varepsilon}-1}\underline{q}_r^{1-\varepsilon}$. Equation (5) therefore simplifies to $\nu_{rs} = \nu_s \equiv \frac{\zeta^{\varepsilon}(\zeta-1)^{1-\varepsilon}}{\zeta+\varepsilon-1}\tilde{\nu}_s$.

Thus, if income follows a Pareto distribution with a common tail parameter, all regions have the same "aggregate" parameter ν_s, which is proportional to the primitive individual preference parameter $\tilde{\nu}_s$. In this case, regional demand differences are solely driven by local prices, wages, and human capital $E_{rt}[q] \propto \underline{q}_{rt}$. Hence, we can write

$\vartheta_{rs}\left(w_{rt}, \mathbf{p_{rt}}\right) = \vartheta_s\left(\underline{q}_{rt} w_{rt}, \mathbf{p_{rt}}\right)$. Note that, if CS are a luxury, their local demand is increasing in both local wages w_{rt} and human capital $\underline{q}_{rt}$. Our goal is to disentangle this income effect from that of changes in the productivity of the local CS sector.

2.3 Spatial Trade Equilibrium

In this section, we characterize the trade equilibrium. Recall that there each variety of food and industrial goods is traded in the national market, whereas the markets for CS are segmented. Therefore, there are $3 \times R$ market clearing conditions for goods and services and R conditions for labor markets. We drop the t subscript if it does not cause any confusion.

The competitive prices are given by

$$p_{rs} = \frac{1}{A_{rs}} w_r.$$

Also, the CES demand structure implies the following regional expenditure shares for traded varieties:

$$\frac{\text{spending on food/ind. good from } r \text{ in } j}{\text{spending on food/ind. good in } j} = \left(\frac{p_{rs}}{p_s}\right)^{1-\sigma} = \frac{w_r^{1-\sigma} A_{rs}^{\sigma-1}}{\sum_{m=1}^{R} w_m^{1-\sigma} A_{ms}^{\sigma-1}}, \qquad \text{for } s = F, G.$$

In the absence of trade costs, the expenditure shares on traded regional varieties do not depend on the location of the buyer. Therefore, we have a set of nationwide market clearing conditions of the form

$$w_r H_{rs} \quad = \quad \left(\frac{w_r^{1-\sigma} A_{rs}^{\sigma-1}}{\sum_{j=1}^{R} w_j^{1-\sigma} A_{js}^{\sigma-1}}\right) \times \sum_{j=1}^{R} \vartheta_s\left(\underline{q}_{rt} w_{rt}, \mathbf{p_{rt}}\right) w_j H_j \qquad \text{for } s = F, G, \tag{6}$$

and a set of district-specific market clearing conditions for nontradable CS

$$w_r H_{rCS} = \vartheta_{CS}\left(\underline{q}_{rt} w_{rt}, \mathbf{p_{rt}}\right) w_r H_r, \tag{7}$$

where ϑ_{rCS} are the aggregate spending shares given in (4). Together with the labor market clearing conditions

$$H_{rF} + H_{rG} + H_{rCS} = H_r, \tag{8}$$

and after setting the industrial good as the numeraire, $p_G = \left(\sum_r \left(\frac{w_r}{A_{rG}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = 1$, we have a total $4R+1$ equations of which, by Walras's Law, only $4R$ are independent. Given a set of productivities, these equations yield a unique solution for the $4R$ unknowns $\{w_r, H_{rF}, H_{rG}, H_{rCS}\}$.

2.4 Equilibrium in the Industrial Sector

So far we have taken A_{rG} as exogenous. In this section, we characterize the equilibrium of the industrial sector that determines A_{rG} as an equilibrium outcome. We assume that each region's industrial good is produced by a continuum of firms using both production workers and PS as inputs. PS are provided by a separate service sector comprising corporate law services, accounting, transport, financial advising, etc. The distinction between CS and PS is important because in many countries, the employment share of PS has increased while that of the manufacturing sector has stagnated or declined. However, PS are inputs to the production of industrial goods. Our theory explicitly allows for structural change to occur within the industrial sector in the form of an adoption of production techniques that are more intensive in PS over time.

Environment. To study and quantify this process of structural change, we introduce a model with heterogeneous firms where PS intensity is related to firm size. Intuitively, in small firms, activities such as accounting are carried out by the manager, while large firms outsource them to professional providers. We formalize this idea by positing a nonhomothetic production function of the following form:

$$y_i = z_i^{1-\alpha-\beta} H_{PMi}^{\alpha} \left(A_{rPS} H_{PSi} + \kappa\right)^{\beta}.$$

Here z_i is firm i's productivity, and H_{PMi} and H_{PSi} denote the inputs of manufacturing production workers and PS, respectively. A_{rPS} denotes the productivity of the PS sector in region r. The parameter $\kappa \geq 0$ governs the nonhomotheticity of firms' technology—if $\kappa > 0$, the input share of PS increases with firms' size. We assume that $\alpha + \beta < 1$, which captures a limited span of managerial control. Decreasing returns to scale guarantee that firms earn profits in equilibrium.

The mass of active firms is endogenously determined via free entry. There are two types of fixed costs. First, to enjoy the opportunity of drawing a realization from the productivity distribution, an entrant firm must pay a sunk labor cost of f_E workers. The productivity z_i is drawn from a Pareto distribution $F_{rt}(z) = 1 - (A_{rMt}/z)^{\lambda}$, where A_{rM} is a lower-bound productivity that parametrizes the state of technology in region r, and $\lambda > 1$ is the tail parameter of the Pareto distribution.

Second, active firms must hire f_O workers in order to start production. Because the overhead cost accrues after observing z_i, some low-productivity firms might opt to remain inactive. Active firms then decide how many production workers and PS to hire. We assume that

$$\frac{f_O}{f_E} > \frac{\beta + (1-\alpha)(\lambda - 1)}{1 - \alpha - \beta}. \tag{9}$$

This assumption ensures that some low-productivity firms will choose not to be active. While inessential, the assumption simplifies the analysis and avoids a taxonomic presentation.

When mapping the theory to the data, we assume employment in the *manufacturing* sector (H_M) includes workers employed in production (H_{PM}), overhead (H_{OM}), and entry (H_{EM}) activities.[7] In our terminology, the *industrial* sector includes both manufacturing and PS workers, $H_G = H_M + H_{PS}$. We label the PS workers *lawyers*, and refer generically to *workers* for manufacturing workers.[8]

Equilibrium. In this section, we summarize—with the aid of Figure 1—the properties of the equilibrium, whose complete formal characterization is deferred to Appendix Section A-2.

Under condition (9), there exist two productivity thresholds, z^* and $z_L(A_{PS})$ such that $A_M < z^* \leq z_L(A_{PS})$, defining three productivity ranges. Low-productivity firms with $z \in [A_M, z^*]$ are inactive; medium-productivity

[7] In the Appendix, we provide a closed-form equilibrium breakdown of H_M into the three activities.

[8] Denote by M the number of active firms. Then, $H_{OM} = M \times f_O$ and $H_{EM} = M \times f_E$, where M is an endogenous variable. In the Appendix, we prove that in equilibrium $M = \frac{1-\alpha-\beta}{\lambda}\frac{H_M}{f_E}$.

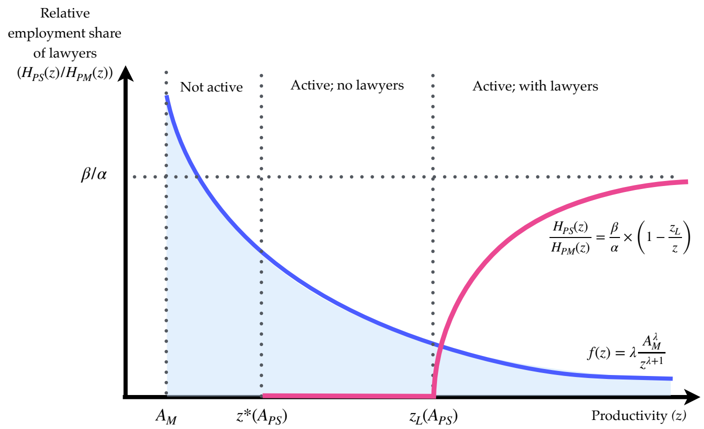

Figure 1: Equilibrium in the Industrial Sector. The figure displays qualitative features of the industrial equilibrium. All firms with $z \geq z^*(A_{PS})$ are active. All firms with $z \geq z_L(A_{PS})$ demand PS. The shaded blue area represents the productivity distribution.

firms with $z \in [z^*, z_L(A_{PS})]$ produce using only production workers;[9] high-productivity firms with $z > z_L(A_{PS})$ demand both workers and lawyers.

For high-productivity firms, there is also an intensive margin: the more productive a firm is, the more intensive in lawyers its technology is. Interestingly, the cutoff $z_L(A_{PS})$ fully determines this intensive margin:

$$\frac{H_{PS}(z)}{H_{PM}(z)} = \frac{\beta}{\alpha} \times \left(1 - \frac{z}{z_L}\right).$$

As z increases, the share of lawyers approaches β/α. The following proposition summarizes the main properties of the equilibrium.

Proposition 1. *The equilibrium production level of the industrial goods sector in region r is given by*

$$Y_{rG} = A_{rG} H_{rG},$$

where $H_{rG} = H_{rM} + H_{rPS}$ and $A_{rG} = A_G(A_{rM}, A_{rPS})$ is given by

$$A_G(A_M, A_{PS}) := \begin{cases} Q_1 A_M^{(1-\alpha-\beta)} \left(\left[1 + \frac{1}{\lambda-1} \frac{\beta}{1-\alpha} \left(\frac{1-\alpha}{\beta} \varsigma(A_{PS}) \right)^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}} \right] \right)^{\frac{1-\alpha-\beta}{\lambda}} & \text{if } \varsigma(A_{PS}) > \frac{\beta}{1-\alpha} \\ Q_2 A_M^{(1-\alpha-\beta)} A_{PS}^{\beta} \left(1 - \varsigma(A_{PS})\right)^{\frac{(1-\lambda)(1-\alpha-\beta)}{\lambda}} & \text{if } \varsigma(A_{PS}) \leq \frac{\beta}{1-\alpha} \end{cases},$$

where $\varsigma(A_{PS}) \equiv \frac{\kappa}{f_O A_{PS}}$, Q_1 and Q_2 are constant terms, and $A_G(A_M, A_{PS})$ is continuous in both arguments.

[9] The medium-productivity range may be empty. In particular, if $A_{PS} > \frac{1-\alpha}{\beta} \frac{\kappa}{f_0}$, then, all active firms hire lawyers, $A_M < z^* = z_L(A_{PS})$.

The employment share of PS is given by:

$$\frac{H_{PS}}{H_G} = \begin{cases} \frac{\beta}{\lambda} \frac{\beta+(\lambda-1)(1-\alpha)}{\beta+(1-\alpha)(\lambda-1)\left(\frac{1-\alpha}{\beta}\varsigma(A_{PS})\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}}} & \text{if } \varsigma(A_{PS}) > \frac{\beta}{1-\alpha} \\ \left(\beta - (1-\alpha-\beta)\,\frac{\lambda-1}{\lambda}\,\frac{\varsigma(A_{PS})}{1-\varsigma(A_{PS})}\right) & \text{if } \varsigma(A_{PS}) \leq \frac{\beta}{1-\alpha} \end{cases}. \tag{10}$$

$H_{PS}/H_G \in [0, \beta]$ *is continuous and strictly increasing in* A_{PS}.

Proof. See Section A-2 in the Appendix. □

Proposition 1 contains two main results. First, it provides a closed-form expression for the total factor productivity A_G, that is increasing in both A_M and A_{PS}. The comparative static of A_{PS} reflects both the direct effect of lawyers being more productive and the indirect effect of selection: an increase in A_{PS} *in*creases z^* and *de*creases z_L and hence induces a reallocation of resources toward more productive firms.

Second, the proposition yields a closed-form expression for the aggregate employment share of lawyers H_{PS}/H_G. The sole determinant of structural change within the industrial sector (i.e., of PS deepening) is the local productivity of the PS sector A_{rPS}. Free entry is key for this stark result. If either the demand for the regional industrial variety or the productivity A_{rM} were to increase, there would be more entry, while the technology choice of the active firms would remain unaffected.

In summary, the model has a tractable recursive structure. The trade equilibrium pins down the employment share of the industrial sector as a function of A_{rM} and A_{rPS}, given preferences and technology in the other sectors. The employment breakdown into manufacturing and PS is then determined by A_{rPS} only.

2.5 Measuring Productivity in the Service Sector

Having presented the full general equilibrium, we can now zoom out on the identification of productivity in the service sectors—which is the focal point of our analysis.

Consumer Services. Using the set of equilibrium conditions (see in particular Appendix Equation (B-3)) and the definition of the expenditure share ϑ_{CS} in (4), our model implies that the local employment share in the CS sector is given by

$$\frac{H_{rCSt}}{H_{rt}} = \omega_{CS} + \nu_{CS} p_F^{\varepsilon\omega_F} p_G^{\varepsilon\omega_G} \times \left(\underbrace{E_{rt}[q]}_{\text{Skills}} \times \underbrace{w_{rt}^{1-\omega_{CS}}}_{\text{Wages}} \times \underbrace{A_{rCSt}^{\omega_{CS}}}_{\text{Productivity}} \right)^{-\varepsilon}, \tag{11}$$

where we have eliminated p_{rCS} using the competitive equilibrium condition $p_{rCS} = w_{rt}/A_{rCSt}$.

Equation (11) highlights the challenge in identifying A_{rCSt}. Because preferences are nonhomothetic and CS are not tradable, employment depends on the local supply of skills ($E_{rt}[q]$), local wages (w_{rt}), and local productivity (A_{rCSt}). Hence, locations like Delhi or Bangalore might have a large employment in the retail service industry either because consumers living there are on average rich or because the local retail sector is highly productive. The former channel describes a *service-biased* growth scenario: the rise of service employment is due to income effects. The latter channel captures a *service-led* growth scenario where rising service employment is driven by rising productivity and the CS sector is a source rather than a corollary of growth.

To solve this identification problem, we leverage both the structure imposed by our theory and additional data. First, the data on earnings, schooling, and an estimate of the returns to schooling allow us to measure local skills and their price. Given an income elasticity ε and the (endogenous) prices of tradable goods p_F and p_G, we can use (11) to identify A_{rCS}. Similar to the traditional approach in development accounting, we use a set of structural parameters to identify productivity in a model-consistent way. However, our inference hinges on solving for a set of equilibrium prices, p_F, p_G and w_r. For this reason, we label our methodology as *equilibrium development accounting.*

Producer services. Here, the identification strategy leverages the recursive structure of Proposition 1. Recall that regional differences in the employment share PS are driven only by differences in A_{rPS}. Specifically, Equation (10) shows that—given α, β and λ—we can infer $\varsigma(A_{rPS})$ from the observed relative employment share of PS relative to manufacturing. Note that, because we can freely chose the units of A_{PS}, it is legitimate to normalize the overhead costs f_O and the non-homotheticity parameter κ to unity. This entails no loss of generality. Our identification relies on the nonhomothetic factor demand functions. If $\kappa = 0$, Equation (10) implies that $\varsigma(A_{PS}) = 0$ irrespective of A_{PS} and the PS employment share would be constant and equal to β.[10]

2.6 The Welfare Impact of Service-Led Growth

To quantify the importance of service-led growth and its distributional implications, we ask the following question: by how much would welfare have changed if productivity had (counterfactually) not not grown in the service sector? When preferences are nonhomothetic, this welfare effect is heterogeneous across households and regions, and depends on their position in the income distribution ladder. The formulation of PIGL preferences allows us address these questions.

Consider first the utilitarian welfare function in location r:

$$\mathcal{U}_{rt}\left(w_{rt}, \mathbf{p_{rt}}\right) = \int V\left(q w_{rt}, p_{rt}\right) dF_{rt}\left(q\right).$$

Using the indirect utility function in (2), we obtain:

$$\mathcal{U}\left(w_{rt}, \mathbf{p_{rt}}\right) = \frac{E_{rt}\left[q^{\varepsilon}\right]}{E_{rt}\left[q\right]^{\varepsilon}} \times \left(\frac{1}{\varepsilon}\left(\frac{E_{rt}\left[q\right] w_{rt}}{p_{Ft}^{\omega_F} p_{Gt}^{\omega_G} p_{rCSt}^{\omega_{CS}}}\right)^{\varepsilon} - \left(\nu_F^{\mathcal{U}} \ln p_{Ft} + \nu_G^{\mathcal{U}} \ln p_{Gt} + \nu_{CS}^{\mathcal{U}} \ln p_{rCSt}\right)\right), \tag{12}$$

where $\nu_s^{\mathcal{U}} \equiv \nu_s \times \left(\left(\zeta - \varepsilon\right)\left(\zeta - \left(1 - \varepsilon\right)\right)\right) / \left(\zeta(\zeta - 1)\right)$. Hence, utilitarian welfare is akin to the indirect utility of a representative agent with average income $E\left[q\right] w_{rt}$ and a scaled taste parameter $\nu_s^{\mathcal{U}}$, that accounts for the income distribution—see Section 2.2. Note that our assumption of Pareto skill distribution with a region-specific parameter $\underline{q}_{rt}$ but a common scale parameter ζ implies that $\nu_s^{\mathcal{U}}$ is constant across time. We show below how this can be identifies from the data.[11]

In addition, we quantify the welfare effects of service-led growth for different parts of the income distribution. Because CS are a luxury and productivity growth in CS lowers the price of such goods, service-led growth is biased toward rich people. In fact, letting $e(\mathbf{p_{rt}}, V)$ denote the expenditure function of achieving a utility level of V given prices $\mathbf{p_{rt}}$, we obtain $\partial \ln e(\mathbf{p_{rt}}, V) / \partial \ln p_{rCSt} = \vartheta_{CS}\left(e, \mathbf{p_{rt}}\right)$. Namely, the percentage increase in expenditure

[10]Intuitively, if the production function was Cobb-Douglas, the relative employment and expenditure shares were independent of productivity. Moreover, in that case, A_M and A_{PS} would both be factor neutral and aggregate TFP A_G, would only depend on $A_M^{1-\alpha-\beta} A_{PS}^{\beta}$. Hence, neither would A_M and A_{PS} be independently identified, nor would it be necessary to do so.

[11]The assumption that ζ is space-time invariant is not essential. We considered more flexible specifications relaxing this restriction. The results were broadly similar.

needed to keep utility constant is equal to the expenditure share on CS. This is larger for richer households, namely, households with high human capital q and households living in high-productivity locations. As we show below, the heterogeneity in the exposure to service-led growth as implied by differences in income and hence the expenditure share is quantitatively large and makes service-led growth significantly pro-rich.

3 Empirical Analysis: Structural Change in India: 1987–2011

In this section, we describe the main empirical patterns that determine the estimates in our model. First, we describe our main data sources. Then, we discuss the features of the process of structural change in India across both space and time.

3.1 Data

Our analysis relies on three datasets.

1. The NSS Employment-Unemployment Schedule, henceforth, the "NSS data."
2. The Economic Census for the years 1990, 1998, 2005, and 2013, henceforth, the "EC."
3. A Special Survey of the Indian Service Sector for the year 2006, henceforth, the "Service Survey."

A more detailed description of these datasets is deferred to Appendix Section B-2. Here, we highlight the main features. The NSS data forms the backbone of our analysis. It is a household survey with detailed information on employment characteristics and households' residence location. We use data for 1987 and 2011. The NSS data allows us to measure sectoral employment shares at the district-year level. Consistent with our theory, we measure employment shares in four sectors: agriculture, manufacturing, PS, and CS. For agriculture and manufacturing, we follow the sectoral classification in the NSS data. The situation is more complicated in the service industry. While, for example, retail workers are clearly part of the CS sector, the distinction is less clear for lawyers as this category includes both corporate lawyers and divorce lawyers, providing, respectively, PS and CS. We rely on combined information from the Economic Census and the Service Survey to allocate service workers to CS and PS (see Section 3.2 below).

The EC is a complete count of all establishments engaged in production or distribution of goods and services in India. The census covers all sectors except crop production and plantation. The EC collects information on each firm's location, industry, employment and the nature of ownership. It covers ca. 24 and 60 million establishments in 1990 and 2013, respectively. As show by earlier studies, most Indian firms are very small, with an average size ranging between two and three employees, over half having a single employee, and only one in a thousand firms employing more than a hundred workers.

The Service Survey was conducted in 2006 and is designed to be representative of India's service sector. It covers 190,282 private enterprises in the following industries: hotels and restaurants, transport, storage and communication, financial intermediation, real estate, and health. In Section B-2 in the Appendix we compare it to the EC and document that it is representative of the distribution of firm size in India.

3.2 Measurement

Measuring Producer and Consumer Services. We aim to distinguish between PS and CS in a way that is consistent with our theory. Ideally, we would want to measure employment in PS and CS with the help of detailed input-output matrices so as to associate the value-added of each firm to the identity of the buyers (either private individuals or firms). To the best of our knowledge, this information is not available.

We exploit then the fact that the Service Survey reports whether a firm is mostly selling to consumers or to other firms. We could in principle calculate the share of employment in every service industry-district cell distinguishing firms selling to other firms from those serving consumers. In practice, even this procedure is not feasible because the Service Survey contains too few firms to precisely estimate these employment shares for each service industry-district cell. Instead, we exploit the fact that the probability of a firm selling to other firms rather than to consumers is highly correlated with firm size—larger firms are more likely to sell to firms. This is shown in Table 1 which displays the share of firms that mainly sell to other firms by employment size. There is a clear pattern that small firms with one or two employees sell almost exclusively to final consumers, while a significant share of large firms sell to other firms. We use this information to classify service firms into providers of PS versus CS.

	Firm size: Number of employees								
	1	2	3	4	5	6-10	11-20	21-50	51+
Producer service share	5.0%	3.8%	6.2%	8.5%	11.5%	12.6%	11.8%	27.6%	42.5%
Number of firms	97337	46571	13227	5156	2777	4841	2830	601	403

Table 1: Share of producer services by firm size. The table reports the share of firms selling to firms (rather than private individuals) in different size categories.

We exploit this pattern reported in Table 1 in the following way. First, we estimate the PS employment share by firm size within service industries. We then use the *district*-specific size distribution from the EC to infer the aggregate PS employment share in district r. Hence, this procedure assumes that the structure of production for firms of equal size does not vary across Indian districts within service industries. The regional variation in PS and CS employment stems from differences in (i) total service employment, (ii) the relative share of different service industries and (iii) the distribution of firm size. In Section B-5 in the Appendix we describe this procedure in more detail.

We exclude from the analysis a subset of service industries for which the categorization into PS and CS is ambiguous. These include public administration and defense, compulsory social security, education, and extraterritorial organizations and bodies.[12]

Finally, we merge construction and utilities with the service sector. Although the construction sector is sometimes included in the industrial sector, the key distinction in our theory is that goods are tradable while services are nontradable. Since construction and utilities are local goods, we find it natural to merge them with services.

We must then break down these activities into PS and CS. The construction sector serves both consumers (e.g., residential housing) and firms (e.g., business construction). We follow a procedure similar to that used for services, which exploits information from the "Informal Non-Agricultural Enterprises Survey 1999–2000" (INAES) dataset. INAES reports the major destination for the sale of final products and services. However, it is only possible to split the destination of construction activities at the national level. We obtain the following breakdown. First, we

[12]The public administration provides services to both individuals and firms. Education affects both households and labor productivity in goods-producing sectors.

remove 12.2% of the construction activity from the sample, which corresponds to the share of government activity (infrastructure and public goods). Then, based on the INAES data, we attribute 86.9% of what is left to CS and 13.1% to PS in every district-year.

Geography. To compare spatial units over time, we create a time-invariant definition of geography. We define regions as Indian districts. Because the boundaries of several districts changed over time, we harmonized them using GIS software, relying on maps for the years 1987, 1991, 2001, and 2011.[13] Appendix Figure B-1 shows the map of India with the actual district borders in 1987 and 2011 (left panel) and the regions we constructed to reconcile discrepancies (right panel).

Human Capital. To be consistent with our theory, we measure each district's endowment of human capital unit $F_{rt}(q)$ and its distribution across sectors in terms of effective units of labor, recognizing that individual workers possess heterogeneous skills.

To measure the distribution of human capital across sectors within a district, we rely on the sectoral distribution of earnings. Since local labor markets are frictionless, there is a single wage per efficiency units in each district. Hence, differences in earnings must reflect heterogeneity in the endowment of effective units of labor. To measure the distribution of human capital across districts, we follow the approach in the development accounting literature and leverage the data on the regional distribution of schooling. More formally, we assume that individual human capital q_i is partly determined by the level of schooling s_i, that is

$$q_i = \tilde{q}\left(s_i, \upsilon_i\right) = \exp\left(\rho s_i\right) \times \upsilon_i$$

where s_i denotes the number of years education, ρ is the annual return to schooling and υ_i is an idiosyncratic shock, which we assume to be iid across districts and years and which satisfies $E[\upsilon_i] = 1$. For skills to follow a Pareto distribution, the regional lower bound $\underline{q}_{rt}$ must satisfy

$$E_{rt}[q_i] = E_{rt}[\exp\left(\rho s_i\right)] = \frac{\zeta}{\zeta - 1}\underline{q}_{rt},$$

and hence can be identified from the observed regional distribution of schooling for a given return to skill ρ. We estimate ρ using Mincerian regression—see Section 4.2.

To measure schooling attainment s_i, we classify people into four educational groups: (i) less than primary school; (ii) primary and upper primary/middle school; (iii) secondary school; and (iv) more than secondary school. We associate each step in the education ladder with three extra years of education, consistent with the organization of schools in India.

3.3 Spatial Structural Change in India: 1987–2011

We now use this data to measure the structural transformation in India, both across time and across space. Consider first Panels (a) and (b) in Figure 2, which display the time evolution of sectoral employment shares. Panel (a) uses a standard classification of sectors in the national account statistics. We separate public services and education from other services, because these two categories are excluded from our analysis. The figure suggests that excluding public services and education is inconsequential insofar as their employment share stays approximately constant

[13] Our regions are defined so that they have the same boundaries over time. To keep the number of regions as large as possible, a region is always the smallest area that covers a single district or a set of districts with consistent boundaries over time. In the end, we obtain 370 regions. Section B-1 in the Appendix describes in detail how we constructed this crosswalk.

over time at a 5% level. Panel (b) uses the sectoral classification we adopt in our analysis. Recall that services here include here construction and utilities.

Two facts are apparent: First, agriculture is the largest employment source, accounting for more than 40% of total employment in 2011 (more than 50% in Panel (b)). Second, the structural transformation in India is mostly an outflow of agriculture and an inflow into CS. Employment in the industrial sector is essentially stagnant. During this time period, income per capita grew by a factor of 3 (see Section C-1 in the Appendix).

We now turn to the spatial heterogeneity across Indian districts. We focus on urbanization as our measure of spatial heterogeneity. This as a mere descriptive device. In Section C-1 in the Appendix we show that there is a strong positive correlation between urbanization and the expenditure per capita in the NSS data for 2011. Thus, we take the urbanization rate as a proxy for economic development across Indian districts.

In Panels (c) and (d) of Figure 2 we display sectoral employment shares by urbanization quantiles. As expected, richer urban locations have lower employment shares in agriculture and specialize in the production of services and industrial goods. Over time, the share of agriculture declines. Between 1987 and 2011 the structural transformation was especially fast in the most urbanized districts. In 1987, agriculture was the main sector of activity even in the top quintile of urbanization. In contrast, in 2011, more than half of the working population was employed in the service sector.[14]

4 Estimation Method and Identification

In this section we describe our estimation methodology. Our approach is in the tradition of development accounting, which has a long history in macro and development economics (see e.g. Caselli (2005), Hall and Jones (1999), and Gancia et al. (2013)). While these studies infer productivity at the country-level from an aggregate production function, our methodology allows us to estimate the entire distribution of sectoral productivity $\{A_{rst}\}$. We do so by relying on the entire equilibrium structure of our model and hence refer to our method as *equilibrium development accounting.*

The centerpiece of our methodology is the distinction between structural parameters and local fundamentals. Our model is characterized by 12 structural parameters describing preferences, technologies, and the distribution of skills

$$\left\{ \underbrace{\varepsilon, \nu_{CS}, \nu_F, \omega_{CS}, \omega_F, \sigma}_{\text{Preference parameters}}, \underbrace{\lambda, \beta, \alpha, f_O, f_E, \kappa}_{\text{Manufacturing technology}}, \underbrace{\rho, \zeta}_{\text{Human capital}} \right\}.$$

In terms of local fundamentals, each region is characterized by a 4-tuple of regional productivity levels in each of the four sectors of production. Define the set of such 4-tuples by $\mathbf{A_t} = \{A_{rFt}, A_{rMt}, A_{rCSt}, A_{rPSt}\}$. Given the set of structural parameters in $\mathbf{A_t}$, there is a unique mapping from the equilibrium skill prices $\{w_{rt}\}$ and sectoral employment allocations $\{H_{rst}\}$ to the underlying productivity fundamentals in $\mathbf{A_t}$. In Section 4.1 we describe this procedure. In Section 4.2 we describe how we estimate the structural parameters in (13).

[14] The difference is even larger when one looks at earnings instead of employment, see Section C-1 in the Appendix.

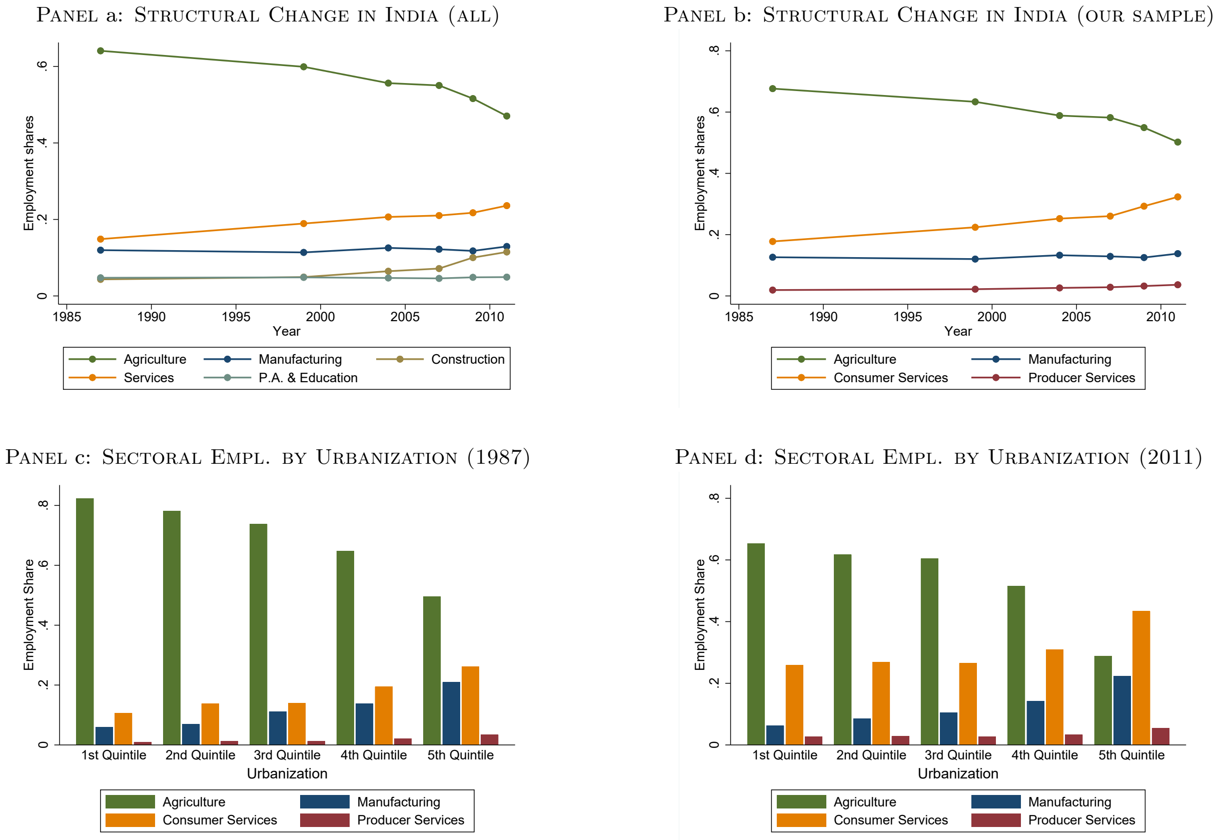

Figure 2: Structural Change in India: 1987–2011. In Panels (a) and (b) we show the evolution of sectoral employment shares over time. Panel (a) shows as separate categories Public Administration&Education and Construction&Utilities. Panel (b) excludes Public Administration and Education and merges Construction and Utilities with the service sector as described in the text. Panel (b) is based on the classification we use in our analysis. In Panels (c) and (d) we plot the sectoral employment shares by urbanization quintile in 1987 and 2011.

4.1 Estimation of Productivity Fundamentals

In this section, we assume that the structural parameter vector $\mathbf{A}$ is known and that we have data on local wages and sectoral employment allocations as well as time-series data on relative prices and aggregate income. These are the data displayed in Figure 2. As we show in detail in Appendix Section A-3, the equilibrium conditions for tradable goods in (6), for CS in (B-3) and for the labor market in (8), uniquely identify a set of local productivity fundamentals.

We discussed the identification of CS in Section 2.5. In particular Equation (11) implies that we can uniquely solve for A_{rCSt} as

$$A_{rCSt} = \left(\frac{(-\nu_{CS})}{\omega_{CS} - \frac{H_{rCSt}}{H_{rt}}} \right)^{\frac{1}{\omega_{CS}} \frac{1}{\varepsilon}} p_F^{\frac{\omega_F}{\omega_{CS}}} p_G^{\frac{\omega_G}{\omega_{CS}}} \left(E_{rt}\left[q\right] \times w_{rt}^{1-\omega_{CS}} \right)^{-\frac{1}{\omega_{CS}}} . \tag{13}$$

Controlling for the level of human capital $E_{rt}\left[q\right]$ and the equilibrium factor price w_{rt}, CS productivity is increasing in the observed employment share $\frac{H_{rCSt}}{H_{rt}}$.[15] Similarly, holding the employment share $\frac{H_{rCSt}}{H_{rt}}$ constant, CS productivity A_{rCSt} is decreasing in both human capital and factor prices. Structurally decomposing the observed variation in employment shares into the part that is service-led (i.e. A_{rCSt}) versus the part that is service-biased (i.e. $E_{rt}[q]\, w_{rt}^{1-\omega_{CS}}$) is a key aspect of our equilibrium accounting methodology.

The procedure to estimate productivity in agriculture and in the industrial sector is different, given the tradable nature of these sectors. Equation (6) implies that relative productivity in agriculture across two locations is given by

$$\frac{A_{rF}}{A_{jF}} = \left(\frac{H_{rF}}{H_{jF}}\right)^{\frac{1}{\sigma-1}} \left(\frac{w_r}{w_j}\right)^{\frac{\sigma}{\sigma-1}}. \tag{14}$$

Hence, relative sectoral productivity differences can be inferred from relative skill prices and relative factor inputs (in units of human capital) given the elasticity of substitution σ. Note that no other preference parameters are involved in this estimation because food and industrial goods are freely tradable so the local demand is dissociated from the local income.

While we can use (14) to estimate the relative sectoral productivity, we still need additional restrictions to estimate the productivity in agriculture and industry. As we show in the Appendix, we can do so using the time-series data on the relative price of food (relative to goods), and aggregate GDP. Intuitively, the relative price is informative about the level of agricultural productivity (relative to industrial goods) and the change in aggregate GDP determines the extent to which productivity in both tradable commodities grew.

Equations (13) and (14) underscore the sense in which our methodology is an accounting procedure: for given parameters we estimate sectoral productivity fundamentals that exactly rationalize the observed data on wages, human capital and sectoral factor inputs as equilibrium outcomes.

4.2 Estimation of Structural Parameters

In the previous section, we took the structural parameters as given. In this section we describe how we estimate these parameters. While we estimate all parameters simultaneously, we describe our identification strategy by referring to the main empirical moments, which identify a given parameter.

The Income Elasticity ε. The single most important parameter in our analysis is the income elasticity ε, which determines how quickly demand shifts away from agricultural goods as incomes rise. The strength of this income effect is essential in identifying the extent to which the variation in employment structure across space and time depends on productivity differences. To estimate ε, we use the cross-sectional relationship between household income and household expenditure shares and estimate ε via indirect inference. In particular, we estimate the following Engel curve using the Indian household data:

$$\ln \vartheta_{F,i} = \delta_r + \beta \times \ln e_i + u_i \tag{15}$$

and estimate ε to target the coefficient β. Here, $\vartheta_{F,i}$ is the expenditure share of food of individual i, δ_r is a region fixed effect, and $\ln e_i$ denotes total spending. While β is not an explicit structural parameter in our theory, there is

[15] Recall that if CS are a luxury, then, $\nu_{CS} < 0$ and $\frac{H_{rCSt}}{H_{rt}} < \omega_{CS}$.

	$\ln \vartheta_F$				
	All	2004	2007	2009	2011
$\ln e$	-0.275***	-0.269***	-0.273***	-0.258***	-0.297***
	(0.00073)	(0.00184)	(0.00128)	(0.00136)	(0.00150)
Year×District FE	Yes				
District FE		Yes	Yes	Yes	Yes
N	570,511	98,763	188,920	143,401	139,427
R^2	0.462	0.434	0.471	0.421	0.431

Table 2: ENGEL CURVES IN INDIA. The table shows the estimated coefficient β of the regression equation (15). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

a tight connection between the structural parameter ε and the regression coefficient β. Note that our theory implies that

$$\ln \vartheta_F^h (e, \mathbf{p_r}) = \ln \left(\omega_F + \nu_F^h \left(\frac{e}{p_F^{\omega_F} \times p_G^{\omega_G} \times p_{CSr}^{\omega_{CS}}} \right)^{-\varepsilon} \right).$$

cf. (3). Hence, if $\omega_F \approx 0$ (which is the case in our structural estimation), our theory implies that

$$\ln \vartheta_F^h (e, \mathbf{p_r}) = \ln \left(\nu_F^h \left(p_F^{\omega_F} \times p_G^{\omega_G} \times p_{CSr}^{\omega_{CS}} \right)^{\varepsilon} \right) - \varepsilon \ln e.$$

The first term, which includes the region-specific price of consumer services p_{rCSt}, is absorbed in the district fixed effect δ_r in (15). The estimated income elasticity β directly coincides with the structural parameter ε.

Table 2 reports the regression results are reported in. The income elasticity is estimated to be in the range between -0.26 and -0.3, very similar across years. In Appendix Section C-1 that the constant elasticity between expenditure and agricultural expenditure shares is a good approximation of the empirical relationship between income and food share for a large part of the expenditure distribution.

Other Preference Parameters $\nu_{CS}, \nu_F, \omega_{CS}, \omega_F, \sigma$. The market-level demands depend on the regional preference parameters ν_{rCS} and ν_{rF}, which are in turn related to the primitive microlevel preference parameters $\tilde{\nu}_{CS}$ and $\tilde{\nu}_F$–cf. Equation (5). In principle one can retrieve the primitive parameters without further restrictions by allowing for regional variation in ν_{rCS} and ν_{rF} and estimating for each region the term $E\left[q^{1-\varepsilon}\right]/E\left[q\right]^{1-\varepsilon}$ from the data. Alternatively, one can impose restrictions on the income distribution ensuring that ν_{rF} and ν_{rCS} are constant over time and across space. As discussed in Section 2.2, a sufficient condition for this to hold true is that q is Pareto-distributed. We pursued both strategies finding very small differences; intuitively, the cross sectional variation in the term $E\left[q^{1-\varepsilon}\right]/E\left[q\right]^{1-\varepsilon}$ is small and has only a negligible influence on the estimation results (details are available upon request). In the remainder of the paper, we focus on the latter strategy by assuming that q follows a Pareto distribution with a region-invariant tail parameter ζ. This distributional form restriction becomes essential to quantify the welfare consequences of structural change in Section 5.2.1 below.[16]

Equation (4) shows that the taste shifters ν_{CS} and ν_F determine sectoral spending and employment when holding income and prices constant. In Appendix Section A-3, we prove that the taste shifter for CS ν_{CS} is not separately identified from the productivity in CS A_{rCSt}. Hence, without loss of generality, we can normalize it to

[16]The value of the Pareto tail ζ is immaterial for the estimation as it turns out to simply scale up or down the estimated $\tilde{\nu}_s$ terms. However, the value of ζ matters for welfare comparisons. For this reason, we return to it in Section 5.2.1.

Parameter	Target	Value
	Preference parameters	
ϵ	Engel Curve	0.297
ω_F	Agricultural spending share US	0.01
ω_{CS}	Agricultural Employment share 2011	0.7
ν_F	Agricultural Employment share 1987	1.249
ν_{CS}	Normalization	-1
σ	Set exogenously	3
	Production function parameters	
λ	Tail of the employment distribution	1.42
β	Employment share of lawyers in the US	0.7
α	Profit share	0.158
f_O	Normalization	1
f_E	Normalization	1
κ	Normalization	1
	Skill parameters	
ρ	Mincerian schooling returns	0.056
ζ	Earnings distribution within regions	2

Table 3: STRUCTURAL PARAMETERS. The table summarizes the estimated structural parameters. The details of the estimation are discussed in the text.

-1. The taste shifter for agricultural products, ν_F, can then be directly identified from the aggregate agricultural employment share in a given year. We opt to match it in the year 1987. This implies that $\nu_F = 1.249$. Given the normalization of $\nu_{CS} = -1$, this implies that $\nu_M = -(\nu_F + \nu_{CS}) = -0.249$. Hence, manufacturing products are also luxury goods as their expenditure share is increasing in income. However, their income elasticity is below the one for consumer services.

To identify the share parameters ω_{CS} and ω_F, recall that $\vartheta_F^h(e,p) > \lim_{e\to\infty}\vartheta_F^h(e,p) = \omega_F$ and that $\vartheta_{CS}^h(e,p) < \lim_{e\to\infty}\vartheta_{CS}^h(e,p) = \omega_{CS}$. Hence, the expenditure share on food (consumer services) approaches ω_F (ω_{CS}) from above (below) as income becomes large. In the United States, which we take an example of a rich economy, where nonhomothetic demand is less important, the agricultural employment share is about 1%. Hence, we take $\omega_F = 0.01$. For ω_{CS}, we follow a similar strategy as for ν_F and match the aggregate sectoral employment shares in a given year. Given our interest in the long-run growth experience of India, we opt to match sectoral employment in 2011. This implies that $\omega_{CS} = 0.7$.[17]

Finally, we set the inter-regional trade elasticity σ to a consensus estimate in the literature. As our baseline estimate we assume that $\sigma = 3$, but we entertain different values in Section 6 when we discuss the robustness of our results.

Technology parameters: $\lambda, \beta, \alpha, f_O, f_E$ and κ. Proposition 1 establishes that all allocations only depend on $A_{rPSt} \times f_O/\kappa$. Hence, the parameters κ and f_O are not separately identified from A_{rPSt}. Since we are not interested in the scale of the average productivity, we normalize $f_O = \kappa = 1$. Similarly, the entry costs f_E is not separately identified from the level of productivity A_{rMt} as long as some firms are "discarded" after their efficiency draw z is observed, that is, condition (9) is satisfied.

[17] Note that our model implies that regional *employment* shares in CS are bounded by ω_{CS} from above. As we discuss in more detail in Section B-4 in the Appendix, there are only seven districts in our Indian data that feature employment shares in CS that exceed ω_{CS}. Because these districts are very small and account for less than 1% of employment, we drop them from our analysis.

We identify the tail of the productivity distribution λ from the employment distribution of the EC. As we show in detail in Section B-10 in the Appendix, our model implies that, like the distribution of productivity, the distribution of employment for large firms is also Pareto with shape λ. We find an estimate of $\lambda = 1.42$, which is very precisely estimated.

We then pick α and β to jointly match a profit share of 10% and the long-run share of lawyers within the industrial sector. More specifically, consider a situation in which the productivity of lawyers A_{PS} becomes large. Our model implies that $\lim_{A_{PS}\to\infty} H_S = \beta H_G$. In the United States, lawyers account for about 28% of employment in the goods producing sector and production workers for 12%. This suggests that $\beta = \frac{0.28}{0.28+0.12} = 0.7$. Given λ and β the parameter α is tied to the profit share because $\alpha + \beta$ determines the returns to scale and hence the share accruing to the fixed factor. In particular, our model implies that the profit share equals $(1 - \alpha - \beta)/\lambda$. For $\beta = 0.7$ and $\lambda = 1.42$, a profit rate of 10% requires that $\alpha = 0.158$. In Section 6, we show that our results are not overly sensitive to these choices.

Skill Parameters ζ, ρ. To estimate the returns to schooling ρ, note that our specification of skills $q_i = exp(\rho s_i)\upsilon_i$ implies that individual log earnings of individual i in region r at time t, y_{irt} are given by the usual Mincerian regression

$$\ln y_{irt} = \ln w_{rt} + \rho s_i + \ln \upsilon_i.$$

Hence, we can estimate ρ from the within-region variation between earnings and education, which we can measure from the NSS data.

We proxy earnings by individual consumption, and regress its logarithm on the education level controlling for year-district fixed effects. We estimate an average annual rate of return of 5.6% . While this is in the lower end of standard Mincerian regressions, recall that we are using data on consumption rather than income. In Section 6 we discuss the robustness of our results with respect to the Mincerian estimate.

The Mincerian regressions allow us to decompose the earning differences across district-time into observed human capital heterogeneity and residual productivity differences that we structurally estimate. In particular, we can calculate the average amount of human capital per region as $E_{rt}[q] = \sum_e \exp(\rho \times e)\ell_r(e)$, where $\ell_r(e)$ denotes the share of people in region r with e years of education. Hence, the distribution of educational attainment across space determines the spatial distribution of human capital.

Finally, we estimate the tail parameter of the skill distribution ζ. This parameter does not affect the equilibrium conditions given that we estimate the aggregate preference parameter ν_s directly. Hence, our estimate of regional productivity does not depend on the value of zeta. An estimate of ζ is only required once we want to calculate welfare—see Section 5.2.1 below. To estimate ζ, recall that the distribution of income in region r is given by $G_r(y) = 1 - \left(\frac{\underline{q}_r w_r}{y}\right)^{\zeta}$, implying that $\ln(1 - G_r(y)) = \zeta \ln\left(\underline{q}_r w_r\right) - \zeta \ln y$. We therefore estimate ζ from a cross-sectional regression

$$\ln(1 - G_r(y_i)) \quad = \quad \delta_r + \beta \ln y_i + u_{ir},$$

where δ_r is a district fixed effect and $\{y_i\}$ is grid of the income distribution. In practice, we pick a grid of 200 points and consider a support of regional incomes above the median as the Pareto distribution is a better fit to the left tail of the income distribution. This procedure yields an estimate of $\zeta \approx 2$ (see Appendix Section B-9).

5 Estimation Results and Counterfactuals

5.1 Estimation Results: Productivity Fundamentals

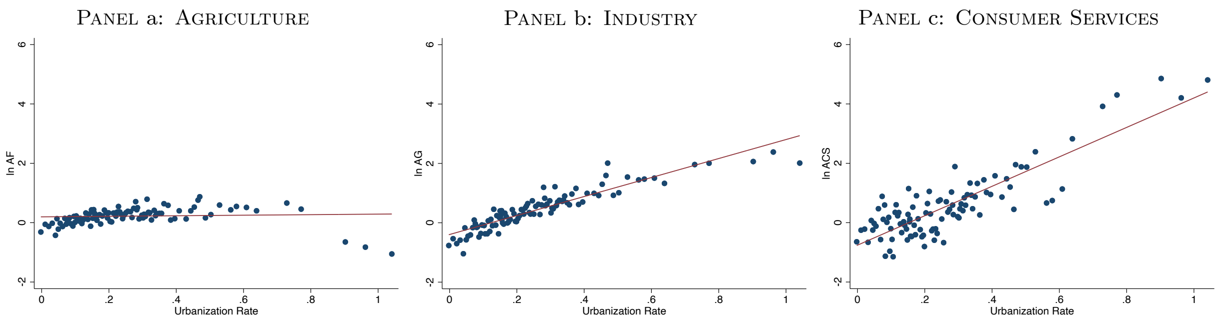

Figure 3: Estimated Sectoral Productivities. The figure shows a bin scatter plot of the estimated sectoral labor productivities in agriculture, industry, and CS across urbanization rate bins. Each plot is constructed by pooling the estimates for 1987 and 2011 after absorbing the year effects and subtracting the (logarithm of) mean sectoral productivity.

We estimate a full set of sector-district productivities separately for 1987 and 2011. Then, we chain them so as to match the aggregate growth rate of GDP per capita in India.[18]

Figure 3 displays a bin scatter plot of the distribution of the (logarithm of the) estimated labor productivities in the agricultural, industrial, and CS sector as functions of the urbanization rate. The relationship between productivity and urbanization is increasing in industry (Panel (b)) and CS (Panel (c)). For agriculture (Panel (b)), the relationship is hump-shaped. The declining portion corresponding to districts with an urbanization rate above 50% likely reflects the scarcity of land (a factor of production from which we abstract) in urban areas.

Remarkably, the productivity dispersion and its correlation with urbanization is strongest in the CS sector. The large employment share of CS in urbanized districts appears to be a consequence not only of high wages (the Baumol effect) or of an abundance of human capital, but also of a high productivity relative to rural areas. Within the traded sector, the dispersion of productivity is significantly larger in the industrial than in the agricultural sector. To understand why, note that that a district's relative productivity in either the food or industrial sector is identified by the district's earning share in that sector relative to the national expenditure in food and goods, respectively–cf. Appendix Equations (A-19)–(A-20). The "compressed" productivity distribution in agriculture reflects the empirical observation that wages are negatively correlated with the employment share of agriculture across districts. Namely, poor districts (where wages are low) have many people employed in agriculture, while richer districts (where wages are high) have fewer people working there. This results in a compressed distribution of earning shares. In contrast, wages are positively correlated with the employment share of industry, implying a wider dispersion in productivity.

Panel (b) of Figure 3 displays the productivity distribution in the industrial sector altogether. Recall that A_G is determined by two separate primitive productivities—A_M and A_{PS}—that we can separately estimate. Both

[18] We measure GDP in terms of numeraire industrial good. Because of nonhomothetic preferences, it is not feasible to define a standard consumption price index. For comparison, we calculated wage growth for a fictitious agent endowed with the median wage and living in a district in which the supply of CS is at the median level. Based on the consumption basket of such an individual in 1987 and 2011, we calculated real wage growth using a Laspeyres and a Paasche index. The resulting real wage growth in the two cases is 2.57 and 1.36, respectively. Our calibration yields a wage growth factor of 2.36, which is in-between.

A_M and A_{PS} are increasing in urbanization, which indicates, on the one hand, that manufacturing productivity increases with urbanization. On the other hand, industrial firms in urban districts use technologies that are more intensive in PS. Quantitatively, A_M accounts for the lion's share of the variation in A_G, with the PS sector playing a far less important role.[19]

Figure 3 focuses on the level of sectoral productivity. In Appendix Figure C-5 we report the distribution of *growth* rates of CS productivity between 1987 and 2011. This is estimated to be positive in the vast majority of districts, indicating a significant extent of service-led growth between 1987 and 2011. Note that our methodology could have delivered constant (or even falling) productivity in CS in spite of the positive employment growth. This would have indicated that employment growth is exclusively driven by income effects. However, we find that, if prices and productivity had been constant across districts, income effects would have been insufficient to explain the different employment shares in the CS sector.

5.2 Counterfactuals: The Importance of Service-led Growth

In this section, we study the implications of sectoral productivity growth for welfare and the structural transformation of the economy during the period 1987–2011. Toward this aim, we run counterfactual experiments in which we assume that each of the sectoral productivities is set back to its 1987 level. This allows us to assess the relative importance of technological progress in different sectors for improving living standards and for structural change. We also study how the welfare effects vary across space and across ladders of the income distribution. As we shall see, our quantitative analysis uncovers a great deal of heterogeneity in both dimensions.

5.2.1 Counterfactual Welfare Effects: Methodology

To measure welfare changes, we calculate equivalent variations at the district level relative to the *status quo* in 2011. Namely, we calculate the percentage income reduction that an agent (either a "representative agent" or an individual with a specific income) would be willing to accept in 2011 to avoid resetting a particular sector's productivity to its 1987 level.

Measurement. We compare two allocations $\{e_{rt}, p_{rt}\}_r$ and $\left\{e_{rt}^{CF}, p_{rt}^{CF}\right\}_r$, where "CF" stands for *counterfactual*, and focus on the concept of equivalent variation. More formally, we define $\overline{\omega}\left(\left(e_{rt}^{CF}, p_{rt}^{CF}\right)|p_{rt}\right)$ to be the income the representative agent in region r facing the equilibrium prices p_{rt} would require to achieve the utility given in (12) with income and prices $\left\{e_{rt}^{CF}, p_{rt}^{CF}\right\}$:$\mathcal{U}\left(\overline{\omega}\left(\left(e_{rt}^{CF}, p_{rt}^{CF}\right)|p_{rt}\right), p_{rt}\right) = \mathcal{U}\left(e_{rt}^{CF}, p_{rt}^{CF}\right)$. [20]

Given the welfare-equivalent income $\overline{\omega}\left(\left(e_{rt}^{CF}, P_{rt}^{CF}\right)|P_{rt}\right)$, we can calculate aggregate welfare as follows:

[19] The relationship between A_PS and urbanization (not shown in the picture) is noisy at low levels of urbanization, where the employment share of PS is very small and susceptible to measurement error. The correlation becomes robust at higher levels of urbanization.

[20] Equation (12) implies that

$$\frac{1}{\varepsilon}\left(\frac{E[q]\,\overline{\omega}\left(\left(e_{rt}^{CF}, p_{rt}^{CF}\right)|p_{rt}\right)}{\prod_s p_{rst}^{\omega_s}}\right)^{\varepsilon} - \sum_s \nu_s^{Welfare}\ln p_{rst} = \frac{1}{\varepsilon}\left(\frac{E[q]\,e_{rt}^{CF}}{\prod_s \left(p_{rst}^{CF}\right)^{\omega_s}}\right)^{\varepsilon} - \sum_s \nu_s^{Welfare}\ln p_{rst}^{CF},$$

so that

$$E[q]\,\overline{\omega}\left(\left(e_{rt}^{CF}, p_{rt}^{CF}\right)|p_{rt}\right) = \left(\left(E[q]\,e_{rt}^{CF}\right)^{\varepsilon}\left(\frac{\prod_s p_{rst}^{\omega_s}}{\prod_s \left(p_{rst}^{CF}\right)^{\omega_s}}\right)^{\varepsilon} - \left(\prod_s p_{rst}^{\omega_s}\right)^{\varepsilon}\varepsilon\left(\sum_s \nu_s^{Welfare}\ln\frac{p_{rst}^{CF}}{p_{rst}}\right)\right)^{1/\varepsilon}.$$

Hence, given vectors of prices p_{rst}^{CF} and p_{rst} and incomes e_{rt}^{CF}, we can calculate $\overline{\omega}\left(\left(e_{rt}^{CF}, P_{rt}^{CF}\right)|P_{rt}\right)$ for a given distribution of q, $F(q)$.

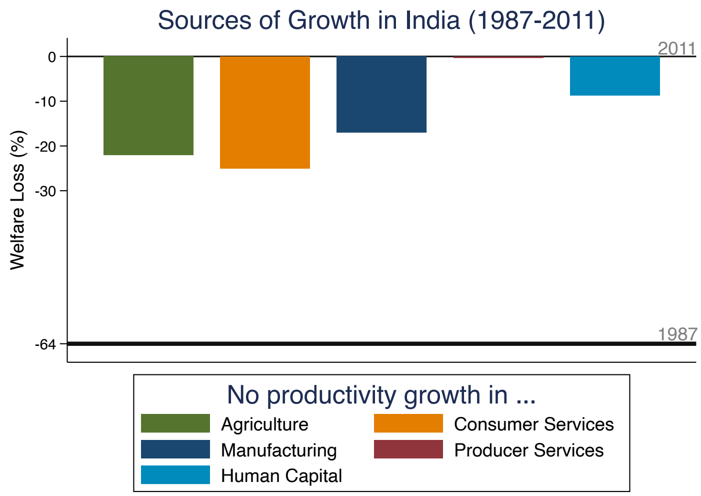

Figure 4: Average Welfare Effects. The figure displays the average percentage welfare losses (equivalent variations) associated with counterfactually setting productivity in agriculture, CS, manufacturing, and PS, as well as the level and distribution of human capital, at the respective 1987 level in all Indian districts. Each district is weighted by the size of its total employment. For comparison, the figure also shows the welfare loss of resetting all productivities and human capital to the 1987 level.

1. Welfare in the calibrated model in 2011 is

$$Welfare^{2011} = \sum_r L_{r2011} e_{r2011}.$$

   Note that this happens to be the same as GDP in 2011.

2. Then take a counterfactual, for example, reducing A_{rF2011} to A_{rF1987}. Aggregate Welfare is the then given by

$$Welfare\left(A_{rF2011} \to A_{rF1987}\right) = \sum_r L_{r2011} \overline{\omega}\left(\left(e_{rt}^{CF}, P_{rt}^{CF}\right) | P_{r2011}\right)$$

   where e_{rt}^{CF} and P_{rt}^{CF} are the income and prices in the counterfactual.

3. Finally, we calculate the percentage difference between $Welfare^{2011}$ and $Welfare\left(A_{rF2011} \to A_{rF1987}\right)$ as our measure of welfare loss.

Trimming Outliers. When we counterfactually set CS productivities to their 1987 level, a few (mostly very small) districts display very large productivity swings implying large equivalent variations. Because the average welfare effects are somewhat sensitive to these outliers, we trim the top and bottom 2% of districts in terms of counterfactual productivity changes. The details are discussed in the Appendix, where we also report robustness results to trimming (see Table B-11).

5.2.2 Counterfactual Welfare Effects: Results

We first consider the average welfare effects obtained by first calculating the counterfactual equilibrium and welfare changes in each district and then aggregating at the national level. Next, we break down the welfare effects across Indian districts and the income distribution ladder.

Average Effects. Figure 4 shows that productivity growth in agriculture and CS are the most important sources of welfare improvement between 1987 and 2011. The equivalent variations are, respectively, 22% and 25% of the 2011 income. The salient role of agriculture is hardly surprising given the large employment share of this sector in India. The large welfare effects of productivity growth in the CS sector (approximately 3/8th of the total effect) is more surprising. Remarkably, productivity welfare in CS has a significantly larger welfare effect than productivity growth in the industrial sector. The role of the PS sector is altogether negligible. The welfare effect stemming from human capital accumulation is modest in comparison, a mere 9% of 2011 income.

Heterogeneous Effects by Urbanization. A centerpiece of our contribution is the quantification of heterogeneous welfare effects across geography and income distribution. To this aim, we first group districts by quintiles of urbanization in 2011. The average urbanization rates of the five quintiles are, respectively: 6.4%, 12.1%, 19.6%, 29.3%, and 56.6%. Then, we calculate equivalent variations when productivity growth in each sector is shut down in the entire economy. The effects are broken down by quintiles of urbanization. Panel (a) of Figure 5 shows the percentage income per capita that, on average, people in each quintile are prepared to sacrifice.

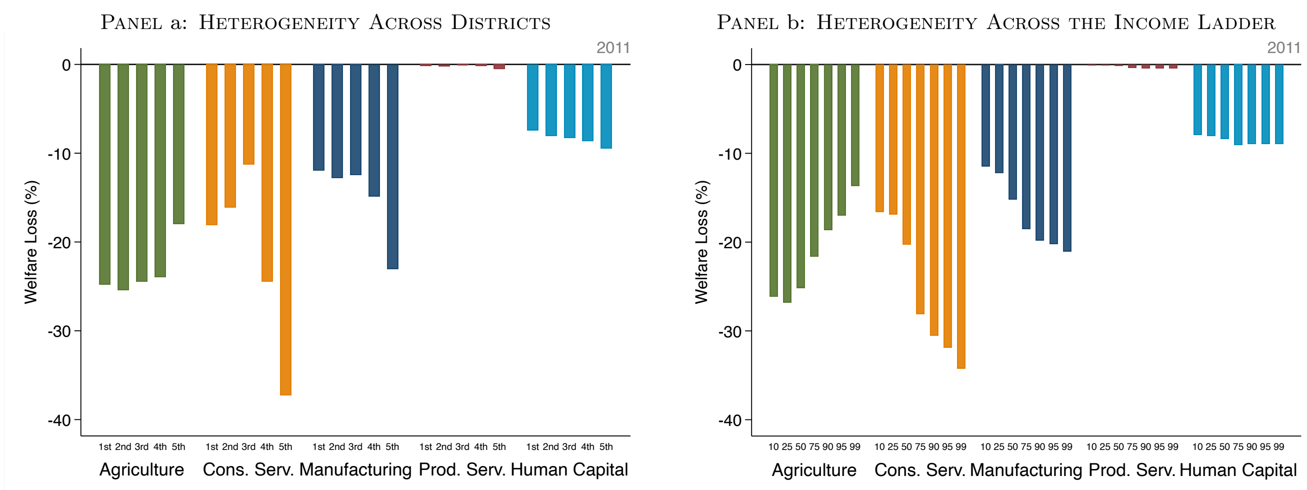

Figure 5: The Heterogeneous Welfare Impact of Service-Led Growth. The figure displays the percentage welfare losses (equivalent variations) associated with counterfactually setting productivity in agriculture, CS, manufacturing, and PS, as well as human capital, at the respective 1987 level, broken down by urbanization quintile in 2011 (Panel (a)) and by the 20th, 50th and 80th percentile of the income distribution in 2011 (Panel (b)).

The welfare effect of productivity growth in agriculture is decreasing with urbanization. On average, households in the lowest quintile of urbanization are prepared to sacrifice 25% of their 2011 income to avoid going back to the 1987 productivity level in agriculture. The equivalent variation declines sharply in the top quintile, where productivity growth in agriculture is only worth 18% of the 2011 income. The reason is twofold: on the one hand, the agriculture share of GDP declines with urbanization; on the other hand, more-urbanized districts are richer, and food represents a smaller share of household expenditure. The pattern is reversed for manufacturing: the equivalent variation increases from 12% to 23% over the urbanization quintiles. This is consistent with our estimated Engel curves showing that industrial goods are a luxury.

The heterogeneity of welfare effects is even stronger for CS. The importance of productivity growth in CS increases sharply at the fourth and, especially, at the fifth urbanization quintile. For the most urbanized quintile, the observed productivity growth in CS is worth 38% of the 2011 income. Resetting productivity at the 1987 level would have two effects. On the one hand, it would induce a reallocation toward the production of the local varieties of food and industrial goods worsening the terms of trade. On the other hand, it would reduce the local availability of CS. Owing again to nonhomothetic preferences, CS are especially valuable to rich consumers in urbanized areas.

In summary, the welfare effects of sectoral productivity growth are heavily skewed. In urban areas, the standards of living grew mostly because of productivity growth in CS and manufacturing. In contrast, technical progress in agriculture is the main source of welfare gains in the three least urbanized quintiles. We conclude that service-led growth is especially important for urban districts.

Heterogeneous Effects by Income. While less urbanized districts are poorer, there are poor and rich households living in each district. In Panel (b) of Figure 5 we decompose the welfare effects across the income distribution ladder. We focus on the 10th, 20th, 50th, 75th, 90th, 95th, and 99th percentiles of the income distribution. As expected, the benefits of productivity growth in CS and (to a lesser extent) manufacturing are sharply increasing in income, while the opposite is true for agriculture. Note that the equivalent variation for the top 99% relative to CS is smaller than for the average of the top quintile of the urbanization distribution. The reason is that not all the rich people live in cities, and the consumption of CS is local.[21]

5.2.3 Counterfactual Structural Change

We now turn to the analysis of the process of structural change. Figure 6 displays the effects of sectoral technical progress on structural change nationwide. Each of the four panels focuses on one sector. The four bars show the actual employment share in 2011 (grey bar) and the counterfactual employment when each of the sectoral productivities is set to its 1987 level.[22] The dashed horizontal line shows the sectoral employment in 1987 for reference.

Productivity growth in CS (orange bars) is the single most important source of structural change. Had productivity stagnated in CS, the employment share of agriculture in 2011 would have been above 60% (Panel (a)), while that of manufacturing and CS would be 2 and 9 percentage points lower (Panels (b) and (c), respectively). In other words, India would have undergone hardly any structural transformation between 1987 and 2011.

In contrast, productivity growth in agriculture (green bars) appears to have marginally increased employment in agriculture—the green bar is lower than the grey bar in Panel (a)—and slowed down employment growth in manufacturing and CS–see Panels (b) and (c). Our findings run against the view that productivity growth in agriculture is a precondition for industrialization. They are instead in line with the findings of Foster and Rosenzweig (2004) on the effects of the Green Revolution and those of Kelly et al. (2020), who document a negative effect of agricultural productivity on the industrial revolution across British regions. Appendix Figure C-6 shows the breakdown of the effects by urbanization quintiles. In conclusion, service-led growth explains the lion's share of structural change, especially in urban areas, but also in less urbanized districts.

[21] In Appendix Figure A-5, we show the welfare effect for people living at selected percentiles of the local income distribution in each district. The results are qualitatively similar.

[22] The figure shows results for employment in effective units of labor, which we label *employment* with a slight abuse of terminology.

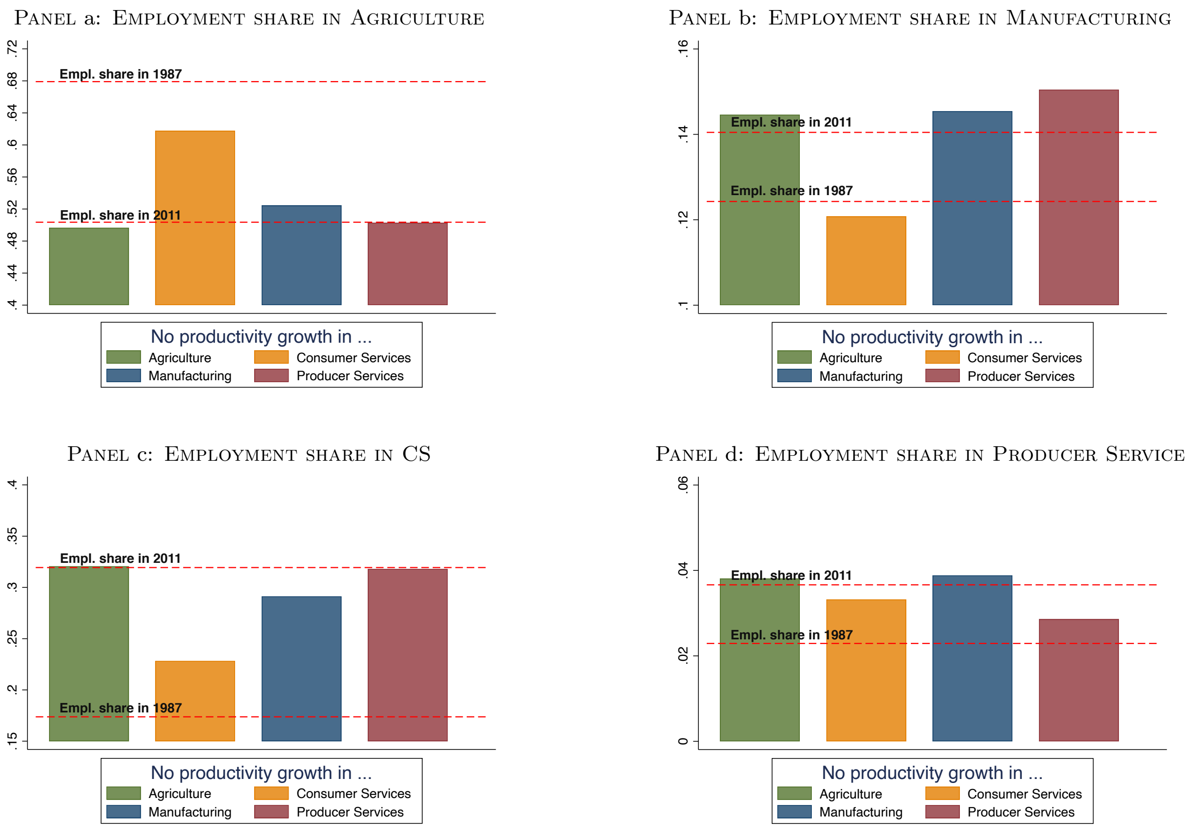

Figure 6: Sectoral Productivity Growth and Structural Change. Each panel in the figure shows the 2011 employment share in one sector and the counterfactual employment shares in the same sector corresponding to setting the productivity in agriculture, CS, manufacturing, and PS at their respective 1987 levels. The dashed horizontal line shows employment in 1987, for reference.

6 Robustness

In this section we discuss the robustness of our results.

6.1 Open Economy

Thus far, we have treated India as a closed economy. However, international trade has become increasingly important for India, which is today among the fifteen largest exporting nations worldwide. In this section we extend our model to an open economy environment. For brevity, we only summarize the main features of the extended model. The technical analysis can be found in Appendix Section A-5.

We assume that consumers, both in India and in the rest of the world, consume industrial goods sourced from many countries. Different national varieties, which are in turn CES aggregates of regional varieties, enter as imperfect substitutes into a CES utility function. In addition, we recognize a specific comparative advantage of India in ICT services. Because in our model services are local goods, this extension requires us to introduce a new separate category of services that is traded internationally.

More specifically, we assume that India exports both domestic goods and ICT services. The foreign demand system for both these exported products is CES; the foreign sector purchases a bundle of regional varieties of ICT services. For simplicity, we assume that export ICT services are not sold in the domestic market. In our estimation, we assume balanced trade. In the line with the empirical observation, India runs a trade deficit in goods and a surplus in ICT services.

As in the baseline case, estimating the model requires us to invert these relationships, so as to identify each region's sectoral productivities, expenditure shares, and parameters from the observable distribution of employment and expenditure, and from the trade flows. Relative to the closed-economy environment, in the equilibrium with trade, we estimate the following additional parameters: $\left\{[A_{rICT}]_{r=1}^{R}, \Upsilon_{ICT}, \Upsilon_G, \eta\right\}$. Here, A_{rICT} is the regional productivity in ICT productivity, Υ_{ICT} and Υ_G parametrize the foreign demand for ICT, and Indian goods, and η is the demand elasticity of international consumers. We externally calibrate η based on the evidence in the trade literature. Then, Υ_G and Υ_{ICT} are estimated to match the total trade flow and its composition between goods and services. The model nests the no-trade equilibrium as a special case in which $\Upsilon_G = \Upsilon_{ICT} = 0$.

Calibration. We set the trade elasticity to $\eta = 5$ following Simonovska and Waugh (2011). According to World Bank data, the export of goods and merchandise increased from 11.3 billions (4.1% of GDP) in 1987 to 302.9 billions (16.6% of GDP) in current USD. The manufacturing sector accounted for 66% of such merchandise exports in 1987 and for 62% in 2011. According to the OECD, the domestic value-added in gross exports amounts to 83.9% of exports for India (there is no time series, so we assume this to be constant). In accordance with these data, we assume that the value-added export of trade has increased from 13.9% in 1987 to 53.6% in 2011 as a share of the GDP in the manufacturing sector.[23]

We classify as ICT service workers all those employed in the following service industries: (i) telecommunications, (ii) computer programming, (iii) consultancy and related activities software publishing, and (iv) information service activities. In our NSS data, this comprises 2.15% of employment in the service sector and 0.77% of total employment in 2011 (in 1987, it was a less than 0.1% of total employment). ICT workers earn on average higher wages than other workers. When one consider the earning share, they account for 4.21% of earnings in the service sector and 1.8% of total earnings in 2011 (in 1987, it was 0.12% of total earnings).[24] Since ICT export was negligible in 1987, we assume it was zero. The target moment is then the revenue share of ICT in 2011. Again, we refer to Section A-5 in the Appendix for more details in the measurement.

Results. Table 4 summarizes the results by reporting the aggregate welfare loss.[25] In the first row we replicate our baseline results reported in Figure 4 for comparison: welfare would decline by 25% and 20% in the absence of productivity growth in CS and agriculture, respectively.

The second row contains the same results for the open economy version of our model: shutting down productivity growth in the CS sector reduces welfare by 20%. Although the effects are slightly smaller than in our baseline calibration, they are fairly similar. In contrast to our baseline model, we now explicitly attribute the increase in employment in the ICT sector to rising foreign and lower trade costs. Earlier, some of these workers were counted as CS workers and hence contributed to our inference of rising productivity growth in CS. However, because the ICT sector is relatively small, the quantitative change is small. For comparison, we also report the results of shutting

[23] This corresponds to an increase in the value-added of exports from 6.26 billions to 157.6 billions (in current USD).

[24] If we multiply 0.67% by the total size of the labor force in 2011, our estimate corresponds to 3.1 million workers being employed in the ICT sector. This is in the ballpark of existing statistics.

[25] For brevity we focus here on CS and agriculture. The full results are available upon request.

down productivity in the agricultural sector, where the differences between our baseline calibration and the open economy model are minuscule.

The model also allows us to assess the welfare effects of international trade. The welfare consequences of shutting down trade in goods and ICT services—reported in the last two columns—are modest: they are equivalent to a fall in consumption of less than 5% in the open economy equilibrium for 2011.

	No productivity growth in ...		No trade in ...	
	Consumer Services	Agriculture	Goods	ICT Exports
Baseline	25%	22%	-	-
Open Economy	22%	20%	3.7%	0.5%
Open Economy (large ICT)	21%	17%	3.8%	1.1%

Table 4: COUNTERFACTUAL WELFARE LOSSES WITH INTERNATIONAL TRADE. The table reports the welfare loss in the absence of productivity growth in CS and agriculture and the gains from trading goods and ICT services. In the second row, we measure ICT employment directly in the data. In the third row we assume that ICT employment is twice as large as in the data.

An alternative calibration in which the ICT sector is twice as large as in our data yields similar results. We also checked that the results are not sensitive to the calibration of the trade elasticity. In Appendix Section A-5, we show that any $\eta \in [4, 6]$ yields quantitatively similar results.

6.2 Sensitivity to Structural Parameters

In this section, we assess the sensitivity of our main results to changes in the parameters governing preferences and skills. On the preference side, we focus on the asymptotic expenditure share on food ω_F and the income elasticity ε. For the distribution of skills we focus on the Mincerian returns ρ and the tail parameter of the skill distribution ζ. All results are based on re-estimating the entire set of district-sector productivities for different parameter values.

Preferences. The parameter ω_F was calibrated so that the asymptotic GDP share of agriculture is 1%, corresponding to the output share of the US farming sector in 2017. However, the GDP share of agriculture is larger than 1% in many European countries, at about 2% in Italy and France, and about 3% in Spain. It is therefore useful to consider a range of larger values of ω_F. Panel (a) of Figure 7 shows the welfare counterfactual results for $\omega_F \in [0.1, 0.6]$. The equivalent variation is essentially unchanged when we shut down productivity growth in either A_F or A_M. In the case of A_{CS}, the figure shows a downward sloping relationship, implying that increasing ω_F magnifies the importance of productivity growth in CS. We conclude that our benchmark calibration is conservative. The quantitative effects of changing ω_F are not negligible but not very large either: the equivalent variation goes up from 25% to 28% as we move from $\omega_F = 0.1$ to $\omega_F = 0.6$.

Panel (b) of Figure 7 focuses on the income elasticity ε. We expect the results to be sensitive to this parameter, which governs the importance of income effects. In particular, a high income elasticity would tend to attribute a large share of the growth of the CS sector to income effects, scaling back productivity growth. Conversely, a low income elasticity would require large productivity growth to explain the observed expansion of the local CS sector. Consequently, we expect the welfare effect of counterfactually shutting down productivity growth in the CS sector to be decreasing in ε. The results shown in Panel (b) of Figure 7 confirm our expectation, and show that changing ε yields large quantitative differences. When $\varepsilon = 0.2$, the equivalent variation goes up to around 33%.

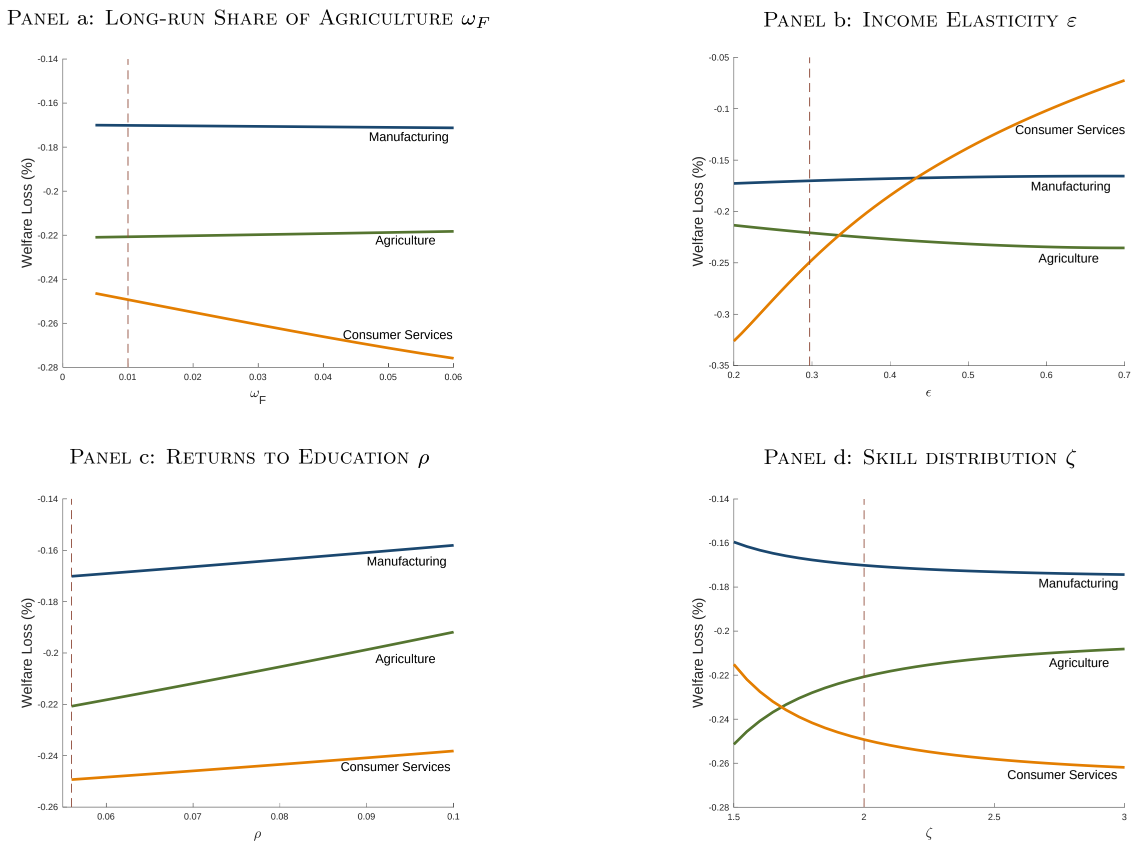

Figure 7: Robustness Analysis. Panels (a), (b), (c), and (d) show the welfare effects as a function of the preference parameters ω_F, ε, the Mincerian rates of return to education ρ and the tail parameter of the skill distribution ζ. The vertical dashed line corresponds to the parameter value in our benchmark analysis.

Conversely, when $\varepsilon = 0.7$, the equivalent variation falls to a mere 7%. Interestingly, ε has no significant effects on the counterfactuals involving A_F and A_M.

We recall that ε is inferred from within-district individual expenditure data (see Table 2 in Section 4.2). Therefore, we view this robustness analysis as reassuring. Our results are identified by an income elasticity that is precisely estimated and are less sensitive to externally calibrated parameters that are subject to more uncertainty.

Skills. In the lower panels of Figure 7, we focus on the skill distribution. When estimating the return to education ρ using our microdata we find an annual 5.6% return, which is in the lower end of typical Mincerian regressions. A potential concern is that we use data on consumption that might reflect consumption sharing within households with different skills and education level. This might result in an attenuation bias. For this reason, we consider alternative scenarios in which the return to education is higher, up to an annual 10% that is an upper bound of the typical estimates.

Panel (c) of Figure 7 shows the results. A higher return to education increases the importance of human capital accumulation up from 9% to 15% when the return to schooling is 10% (not shown in the picture). This reduces the effect of technical progress in agriculture and industry, whereas its impact on CS is far smaller. Intuitively, a higher return to education implies that human capital accumulation is responsible for a larger share of wage growth, thereby reducing the estimated technical change in tradable sectors. However, technical progress in the CS sector is mainly identified by the extent of structural change. Thus, human capital accumulation has negligible consequences on the estimated productivity growth in the CS sector. In conclusion, while our estimated return to education is low, a higher return would increase the relative importance of technical progress in CS relative to the other sectors.

Panel (d) of Figure 7 shows the effect of changing the tail parameter of the skill distribution ζ. This parameter mostly affects our decomposition of productivity growth into agriculture and CS: the higher the ζ, the higher the importance of CS growth relative to agricultural productivity. This result reflects the importance of nonhomothetic demand. The smaller the ζ, the higher income inequality. And because higher inequality increases aggregate demand for CS for a given average wage, due to nonhomothetic demand, less productivity growth is "required" to explain the rise in CS employment if ζ was small. Figure 7 shows that this intuition is borne out, but the effects are quantitatively moderate.

Production function parameters. We identify the production function parameters α and β from two moments: the profit share and the long-run employment share of PS workers relative to production workers. The parameters α and β affect the mapping between the observed employment share L_{rPS}/L_{rG} and the level of PS productivity A_{rPS}. For our baseline results we assumed a profit share of 10% and a long-run employment share of 0.7.

Quantitatively, our results are essentially insensitive to such choices. First of all, recall, that our estimates of the productivity in CS A_{CS}, agriculture A_F, and the industrial sector A_G are *independent* of α and β. Such parameters only affect the relative decomposition of A_G into the manufacturing piece (i.e. A_M) and the part accounted for by PS, i.e., A_{PS}. Second, the Indian PS sector is still sufficiently unproductive that the employment share is well below its asymptotic level of $\beta = 0.7$. And in that range, the mapping between the observed employment share and PS productivity A_{PS} is almost insensitive to α and β.

6.3 A More General Parametrization of PIGL Preferences.

Our parametrized indirect utility function (2) entails restrictions on the price elasticity that are stronger than in the general class of PIGL preferences—see Equation (1). In particular, we set $D(\mathbf{p}) = (1-\varepsilon)\left(\sum_s \tilde{\nu}_s \ln p_s\right)$. In this section, we partially relax this restriction by assuming a more general CES specification where

$$D(\mathbf{p}) = \frac{1}{\gamma}\left[\left(\prod_j p_j^{\tilde{\nu}_j}\right)^{\gamma} - 1\right], \tag{16}$$

and where we continue to assume $\sum_j \nu_j \leq 1$. This class of preferences nests our benchmark logarithmic specification as a particular case as $\gamma \to 0$. The CES formulation yields the following (individual) expenditure shares:

$$\vartheta_s^h(e, \mathbf{p}) = \omega_s + \tilde{\nu}_j \prod_j p_j^{\tilde{\nu}_j \gamma}\left(\frac{e}{B(\mathbf{p})}\right)^{-\varepsilon},$$

where $B(\mathbf{p})$ is as in the benchmark model. It is easy to show that aggregation properties carry over to this generalization.

The CES formulation contains an additional parameter (γ) that is not easy to identify with our data. For this reason, we limit ourselves to a sensitivity analysis of the parameter γ in a neighborhood of the benchmark specification (2). We show results obtained from reestimating the model for alternative values of γ.

We run experiments in which we reduce each of the productivities A_F, A_G, and A_{CS} by 20% in all districts in 2011.[26] We find that the welfare effect of productivity growth in CS is increasing in γ, with moderate quantitative effects. The welfare effect of productivity growth in agriculture is also increasing in γ (and more sensitive to γ relative to CS). Finally, the welfare effect of productivity growth in manufacturing is not affected by γ. We conclude that the welfare results are robust to this generalization within the class of PIGL preferences. See Section B-11 in the Appendix for details.

6.4 Imperfect Substitution and Skill Bias in Technology

Thus far, we have allowed individual heterogeneity in human capital. However, we have maintained that workers endowed with different efficiency units are perfect substitutes for one another. In this section, we generalize the model by assuming that workers with different educational attainments are imperfect substitutes in production (see Section B-11 in the Appendix for details).

6.4.1 Model Environment.

We continue to allow for heterogeneous productivities across workers of the same educational group. A worker's wage is a draw from a skill-specific Pareto distribution with the same tail parameter as in our baseline analysis.[27] The technology admits differences in both TFPs and skill intensity across sector-districts and over time. The main goal of the extension is to capture the effect of variations in skill endowment at the district level on the process of structural change. For instance, agricultural workers have on average lower educational attainment that those employed in service industries. Thus, an increase in the skill endowment could be responsible for part of the growth of employment in the CS sector at the expense of agriculture (cf. Porzio et al. (2020) and Schoellman and Hendricks (2020)). By ignoring specialization based on skills, our Ricardian model could exaggerate the importance of technology (and, conversely, understate the importance of human capital accumulation) for the development of the service sector.

Table 5 reports the share of skilled workers by sector and across time. We define workers to be skilled if they have completed secondary school. It is apparent that there are significant differences in the skill content by sector and over time.

	Agriculture	Goods	Cons. Services	Aggregate
1987	2.6%	12%	16%	5.6%
2011	14%	35%	39%	24%

Table 5: Skill-Employment Shares. The table reports the share of skilled workers in each sector and in the aggregate economy. We define workers as skilled if their schooling equals or exceeds secondary school.

More specifically, we assume the production functions to be of the following form:

[26] The experiment is close in spirit to resetting TFP to its 1987 level in each district. However, doing the latter raises some computational problems because in numerous locations the counterfactual equilibrium features corner solutions with zero employment in CS.

[27] It is impossible to separately identify the lower bound of the Pareto distribution from which each individual draws the realization of human capital from the level of the technology parameters. Therefore, we normalize the lower bound to unity for both skill groups. Since we are only interested in changes over time in TFP, this is immaterial.

$$Y_{rs} = A_{rst} \left(\left(H^{-}_{rst}\right)^{\frac{\rho-1}{\rho}} + \left(Z_{rst} H^{+}_{rst}\right)^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}},$$

where H^{+} and H^{-} denote high- and low-skilled workers, respectively. In the data, we classify the two groups as people with at least secondary school and with less than secondary school. The parameter Z_{rst} captures the skill bias of technology.[28] Instead, we assume the elasticity of substitution to be constant across sector-districts. We externally calibrate $\rho = 1.8$, which is in the consensus region (see, e.g., Ciccone and Peri (2005) and Gancia et al. (2013)). Below we show that our conclusions do not hinge on the particular calibration of ρ.

For simplicity, we assume that within the industrial sector, there are no differences in factor intensity between the manufacturing workers and those producing PS and ignore the fact that in reality PS workers tend to have higher educational attainments. More formally, $Z_{rMt} = Z_{rPSt}$, for all r and t. This assumption, which we impose for symmetry with our benchmark model, ensures that the aggregation result highlighted in Proposition 1 continues to hold. Because our earlier results show that the PS sector plays a negligible role in our results, we do not expect this assumption to significantly affect any of the quantitative results in this section.[29]

The model with heterogeneous skill-intensities is exactly identified (see Section B-11 in the Appendix). Given the set of technological parameters and the factor endowments of each district, and given the preference parameters, H^{+}_{rt} and H^{-}_{rt}, the equilibrium pins down the sectoral employment of the two types of workers in all districts and the relative skill prices. Conversely, if we know the distribution of employment at the sectoral level and the skill premium, we can retrieve the full set of technology parameters from the equilibrium condition of the model. Hence, our calibrated model perfectly rationalizes the data of sectoral earnings shares by skill group and average earnings by skill group for each region in India.

6.4.2 Results

The model with heterogeneous skills allows us to uncover additional facts about the skill bias in technology. First, across districts, $Z_{r\sigma}$ increases in the level of urbanization for all sectors $\sigma \in F, G, CS$. This reflects the empirical observation that the skill premium is higher in urban than in rural districts. Second, there is skill-biased technical change: over time, $Z_{r\sigma}$ increases in all sectors—see Figure 8.

Although our accounting approach does not uncover causal links, the evidence is consistent with models of directed technical change and directed technology adoption such as Acemoglu and Zilibotti (2001) and Gancia et al. (2013), where firms adopt more skill-intensive technologies in response to the wider availability of skilled workers.

Next, we turn to counterfactuals. Since technology is described by two parameters, we run counterfactual experiments by simultaneously resetting A_{rs} and Z_{rs} from the 2011 to the 1987 level. Figure 9 shows the results about welfare. The results are remarkably similar to the benchmark model. The left panel shows the average effects–cf. Figure 4. As expected, changes in human capital are now more salient, although they continue to be less important than changes in technology. Across sectors, technical progress in agriculture is now slightly more

[28] It is important to allow the skill bias of technology to vary across space. If Z were constant across districts, the model would predict skill premia to be lower in skill-rich regions. However, this contradicts the observation that both the relative supply of skills and the skill premium are positively correlated with urbanization.

[29] If PS and manufacturing services were provided with different skill intensities, aggregate productivity in the industrial sector (A_{rGt}) would depend on the relative skill prices w^{+}_{rt} and w^{-}_{rt}. Hence, all equilibrium allocations would depend on the production structure in the industrial sector. While, empirically, the PS sector is more skill-intensive than the manufacturing sector, this difference is smaller than the difference between the industrial sector and the other sectors.

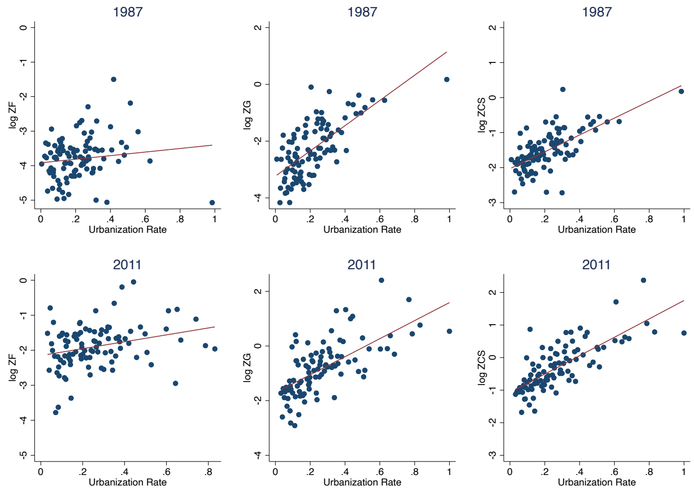

Figure 8: Skill Bias of Technology. The figure shows the distribution of Z_{rF}, Z_{rG}, Z_{rCS} across districts of different urbanization (binned scatter plots) in 1987 (upper panels) and 2011 (lower panels).

important than in CS, although the welfare effects of both continue to be very large. The right panel shows the distribution across quintile of urbanization. The pattern is similar to that in the benchmark model.

Finally, Appendix Figure C-7 shows the effect of the counterfactual shutdowns of productivity growth on structural change. The results are similar to those in benchmark economy–cf. Figure 6. Overall, this section shows that the main results of the paper are robust to an environment in which workers of different educational attainment are imperfect substitutes and in which we allow for skill-biased technical change.

7 Conclusion

In this paper, we propose a new methodology of structural development accounting based on a spatial equilibrium model to estimate sectoral productivities across regions and over time. The methodology allows us to separate income effects in demand from the effect of changes in relative prices arising from heterogeneous sectoral productivity growth. The split is disciplined by income elasticities that we estimate from microdata controlling for price heterogeneity in nontradable sectors.

The estimated model allows us to determine the importance of different sectors as an engine of growth and structural transformation. Moreover, it lends to a quantitative analysis of welfare and distributional effects associated with different sources of productivity growth.

We apply the methodology to India, a country experiencing a pronounced transition from agriculture to services with only a modest growth of the industrial sector.

A counterfactual analysis based on the estimated model indicates that productivity growth in agriculture has a major impact on the welfare of poor Indian districts and poor households. In contrast, the effect of productivity growth in consumer services is skewed toward urban districts and rich households. For the top quintile of the urbanization distribution, the benefits from productivity growth in consumer services exceed those from all other sectors of the economy jointly. The reason is twofold. First, productivity in services grows faster in urban districts.

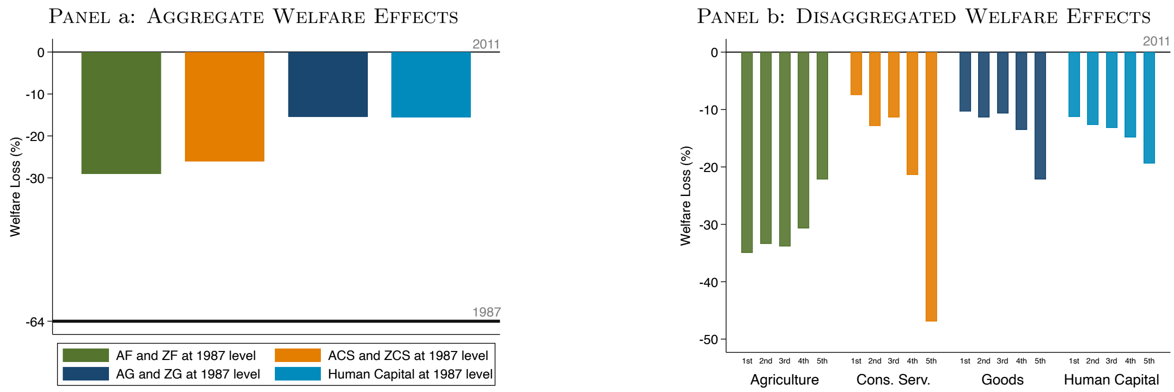

Figure 9: Welfare Effects with Imperfect Substitution in Skills. The figure shows the equivalent variation associated with setting productivity (both A_{rs} and Z_{rs}) in agriculture, CS, manufacturing, PS, and average human capital at the respective 1987 levels in all Indian districts. Panel (a) shows the average effect nationwide, while Panel (b) shows the breakdown by quintiles of urbanization.

Second, richer consumers care more about the consumption of services owing to nonhomothetic preferences.

We also find that productivity growth in consumer services is the main driver of the structural transformation of the Indian economy. Had productivity stagnated in the service sector, the structural composition of employment in 2011 would have remained similar to that of 1987. In particular, cities would be more industrial and offer less amenities. In contrast, technical progress in agriculture did not promote structural change. This finding is in line with a growing body of literature–including Kelly et al. (2020), Moscona (2019), Foster and Rosenzweig (2004)–who document similar findings for India, within and across countries, and for the British Industrial Revolution. We test the robustness of our results along several dimensions, including international trade, changes in preference parameters, inequality, the return to education, and different skill intensities in technology across sectors, districts, and time. There is a single critical parameter to which the results are sensitive: the income elasticity in the PIGL preferences. Reassuringly, this parameter is very precisely estimated (and stable over time) from microdata within Indian districts.

Our approach has several limitations that we hope to overcome in future research. First, we conjecture that producer services have played a more salient role in the recent experience of mature economies–as emphasized by Eckert et al. (2020). Second, we do not attempt to endogenize technical change, human capital accumulation, and migration. Third, one could test the robustness of the results to other classes of preferences and technologies. In spite of these and other limitations, our framework can be a useful tool for analysis for researchers planning to use survey data as the backbone of macroeconomic analyses of the development process.

References

Acemoglu, D. and F. Zilibotti: 2001, 'Productivity Differences'. *Quarterly Journal of Economics* **116**(2), 563–606.

Aghion, P., R. Burgess, S. Redding, and F. Zilibotti: 2005, 'Entry Liberalization and Inequality in Industrial Performance'. *Journal of the European Economic Association* **3**(2-3), 291–302.

Akcigit, U., H. Alp, and M. Peters: 2020, 'Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries'. Forthcoming in the *American Economic Review*.

Alder, S., T. Boppart, and A. Muller: 2019, 'A Theory of Structural Change that Can Fit the Data'. CEPR Discussion Paper 13469.

Allen, T. and C. Arkolakis: 2014, 'Trade and the Topography of the Spatial Economy'. *Quarterly Journal of Economics* **129**(3), 1085–1140.

Amirapu, A. and A. Subramanian: 2015, 'Manufacturing or Services? An Indian Illustration of a Development Dilemma'. Centre for Global Development Working Paper 409.

Basu, K.: 2008, 'The Enigma of India'. *Journal of Economic Literature* **46**(2), 396–406.

Basu, K. and A. Maertens: 2007, 'The Pattern and Causes of Economic Growth in India'. *Oxford Review of Economic Policy* **23**(2), 143–167.

Baumol, W.: 1967, 'Macroeconomics of Unbalanced Growth: The Anatomy of the Urban Crisis'. *American Economic Review* **57**(3), 415–426.

Boppart, T.: 2014, 'Structural Change and the Kaldor Facts in a Growth Model With Relative Price Effects and Non-Gorman Preferences'. *Econometrica* **82**(6), 2167–2196.

Buera, F. J. and J. P. Kaboski: 2012, 'The Rise of the Service Economy'. *American Economic Review* **102**(6), 2540–69.

Caselli, F.: 2005, 'Accounting for Cross-Country Income Differences'. In: P. Aghion and S. Durlauf (eds.): *Handbook of Economic Growth*, Vol. 1A. Elsevier B.V., Chapt. 9, pp. 679–741.

Celasun, O. and B. Gruss: 2018, 'The declining share of manufacturing jobs'. Column published in vox.eu, CEPR Economic Portal, https://voxeu.org/article/declining-share-manufacturing-jobs.

Ciccone, A. and G. Peri: 2005, 'Long-Run Substitutability Between More and Less Educated Workers: Evidence from U.S. States, 1950-1990'. *Review of Economics and Statistics* **87**(4), 652–663.

Comin, D., D. Lashkari, and M. Mestieri: 2020, 'Structural Change with Long-Run Income and Price Effects'. Forthcoming in *Econometrica*.

Duarte, M. and D. Restuccia: 2010, 'The Role of the Structural Transformation in Aggregate Productivity*'. *Quarterly Journal of Economics* **125**(1), 129–173.

Duarte, M. and D. Restuccia: 2019, 'Relative Prices and Sectoral Productivity'. *Journal of the European Economic Association* **18**(3), 1400–1443.

Eckert, F., S. Ganapati, and C. Walsh: 2020, 'Skilled Scalable Services: The New Urban Bias in Economic Growth'. Minneapolis Reserve Bank of Minneapolis, Institute Working Papaer 25.

Eckert, F. and M. Peters: 2016, 'Spatial Structural Change'. Working Paper.

Eichengreen, B. and P. Gupta: 2011, 'The Service Sector as India's Road to Economic Growth'. *India Policy Forum, National Council of Applied Economic Research* **7**(1), 1–42.

Erumban, A. A., D. K. Das, S. Aggarwal, and P. C. Das: 2019, 'Structural change and economic growth in India'. *Structural Change and Economic Dynamics* **51**, 186–202.

Foellmi, R. and J. Zweimüller: 2006, 'Income Distribution and Demand-Induced Innovations'. *Review of Economic Studies* **73**(4), 941–960.

Foellmi, R. and J. Zweimüller: 2008, 'Structural change, Engel's consumption cycles and Kaldor's facts of economic growth'. *Journal of Monetary Economics* **55**(7), 1317–1328.

Foster, A. and M. Rosenzweig: 1996, 'Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution'. *American Economic Review* **86**(4), 931–953.

Foster, A. and M. Rosenzweig: 2004, 'Agricultural Productivity Growth, Rural Economic Diversity, and Economic Reforms: India, 1970–2000'. *Economic Development and Cultural Change* **52**(3), 509–542.

Gancia, G., , A. Müller, and F. Zilibotti: 2013, 'Structural Development Accounting'. In: D. Acemoglu, M. Arellano, and E. Dekel (eds.): *Advances in Economics and Econometrics: Theory and Applications (Tenth World Congress of the Econometric Society)*. Cambridge University Press, pp. 373–418.

Garcia-Santana, M., J. Pijoan-Mas, and L. Villacorta: 2020, 'Investment Demand and Structural Change'. Mimeo, Universitat Pompeu Fabra.

Ghose, A.: 2014, 'India's Services Led Growth'. Institute for Human Development New Delhi, Working Paper 01/2014.

Goldberg, P., A. K. Khandelwal, N. Pavcnik, and P. Topalova: 2010, 'Imported Intermediate Inputs and Domestic Product Growth: Evidence from India'. *Quarterly Journal of Economics* **125**(4), 1727–1767.

Gollin, D., D. Lagakos, and M. Waugh: 2014, 'The Agricultural Productivity Gap'. *Quarterly Journal of Economics* **129**(2), 939–993.

Gordon, J. and P. Gupta: 2005, 'Understanding India's Services Revolution'. In: W. Tseng and C. David (eds.): *India's and China's Recent Experience with Reform and Growth. Procyclicality of Financial Systems in Asia*. Palgrave Macmillan, London, pp. 49–84.

Hall, R. and C. Jones: 1999, 'Why Do Some Countries Produce So Much More Ouput Per Worker Than Others?'. *Quarterly Journal of Economics* **114**(1), 83–116.

Herrendorf, B., R. Rogerson, and A. Valentinyi: 2013, 'Two Perspectives on Preferences and Structural Transformation'. *American Economic Review* **103**(7), 2752–89.

Herrendorf, B., R. Rogerson, and A. Valentinyi: 2014, 'Growth and Structural Transformation'. In: *Handbook of economic growth*, Vol. 2. Elsevier, pp. 855–941.

Herrendorf, B., R. Rogerson, and Valentinyi: 2020, 'Structural Change in Investment and Consumption—A Unified Analysis'. *Review of Economic Studies*.

Hobijn, B., T. Schoellman, and A. Vindas: 2019, 'Structural Transformation by Cohort'. Technical report, Technical Report, Arizona State University.

Hsieh, C.-T. and E. Rossi-Hansberg: 2019, 'The Industrial Revolution in Services'. NBER Working Paper No. 25968.

Kelly, M., J. Mokyr, and C. O'Grada: 2020, 'The Mechanics of the Industrial Revolution'. CEPR Discussion Paper No. DP14884.

Kochhar, K., U. Kumar, R. Rajan, A. Subramanian, and I. Tokatlidis: 2006, 'India's Pattern of Development: What Happened, What Follows?'. *Journal of Monetary Economics* **53**(5), 981–1019.

Majid, N.: 2019, 'Structural Change and Employment in India'. ILO/SIDA Partnership on Employment Working Paper No.1.

Martin, L., S. Nataraj, and A. Harrison: 2017, 'In with the Big, Out with the Small: Removing Small-Scale Reservations in India'. *American Economic Review* **107**(2), 354–386.

Matsuyama, K.: 2000, 'A Ricardian Model with a Continuum of Goods under Nonhomothetic Preferences: Demand Complementarities, Income Distribution, and North-South Trade'. *Journal of Political Economy* **108**(6), 1093–1120.

Matsuyama, K.: 2019, 'Engel's Law in the Global Economy: Demand-Induced Patterns of Structural Change, Innovation, and Trade'. *Econometrica* **87**(2), 497–528.

McMillan, M. and D. Rodrik: 2011, 'Globalization, Structural Change, and Productivity Growth'. In: M. Bacchetta and M. Jansen (eds.): *Making Globalization Socially Sustainable*. International Labor Organization, pp. 49–84.

Mitra, D. and B. P. Ural: 2008, 'Indian Manufacturing: A Slow Sector in a Rapidly Growing Economy'. *Journal of International Trade & Economic Development* **17**(4), 525–559.

Moscona, J.: 2019, 'Agricultural Development and Structural Change, Within and Across Countries'. Working Paper, MIT.

Muellbauer, J.: 1976, 'Community Preferences and the Representative Consumer'. *Econometrica* **44**(5), 979–999.

Mukherjee, A.: 2013, ‘The Service Sector in India’. Asian Development Bank Economics Working Paper Series No. 352.

Porzio, T., F. Rossi, and G. Santangelo: 2020, ‘The Human Side of Structural Transformation’. CEPR Discussion Paper No. DP15110.

Redding, S. J. and E. Rossi-Hansberg: 2017, ‘Quantitative Spatial Economics’. *Annual Review of Economics* **9**(1), 21–58.

Sato, R.: 2014, ‘The Most General Class of CES Functions’. *Econometrica* **43**, 999–1003.

Schoellman, T. and L. Hendricks: 2020, ‘Skilled Labor Productivity and Cross-Country Income Differences’. Mimeo, Federal Reserve of Minneapolis.

Simonovska, I. and M. E. Waugh: 2014, ‘The elasticity of trade: Estimates and evidence’. *Journal of international Economics* **92**(1), 34–50.

Singh, A. and S. Dasgupta: 2016, ‘Will Services Be the New Engine of Economic Growth in India?’. Working Papers id:11176, eSocialSciences.

APPENDIX A: TECHNICAL DETAILS

In this section, we discuss technical material referred in the text.

A-1 Derivation of Expenditure Shares (Equation (3))

Using the definition of $B(\mathbf{p})$ and $D(\mathbf{p})$, the indirect utility function in (2) is given by

$$V(e,\mathbf{p}) = (1-\varepsilon)\left[\frac{1}{\varepsilon}\left(\frac{e}{\prod_s p_{st}^{\omega_s}}\right)^{\varepsilon} - \left(\sum_s \nu_s \ln p_{st}\right)\right]. \tag{A-1}$$

Roy's Identify implies that the expenditure share on sector s is given by

$$\vartheta_s(e,\mathbf{p}) = -\frac{\frac{\partial V(e(\mathbf{p},u),\mathbf{p})}{\partial p_s} p_s}{\frac{\partial V(e(\mathbf{p},u),\mathbf{p})}{\partial e} e}.$$

Using (A-1), it follows that

$$\vartheta_s(e,\mathbf{p}) = -\frac{(1-\varepsilon)\left(-\omega_s\left(\frac{e}{\prod_s p_{st}^{\omega_s}}\right)^{\varepsilon} - \nu_s\right)}{(1-\varepsilon)\left(\frac{e}{\prod_s p_{st}^{\omega_s}}\right)^{\varepsilon}} = \omega_s + \nu_s\left(\frac{e}{\prod_s p_{st}^{\omega_s}}\right)^{-\varepsilon}$$

A-2 Equilibrium in the Industrial Sector

In this section we characterize the equilibrium in the industrial sector. The technical details are deferred to the online appendix. As highlighted in Proposition 1, we have to distinguish two cases. In particular, recall the definition

$$\varsigma(A_{PS}) \equiv \frac{\kappa}{f_0 A_{PS}}. \tag{A-2}$$

Henceforth, we simply write ς. Below we will show that some active firms do not hire lawyers if and only if

$$\varsigma \geq \frac{\beta}{1-\alpha}. \tag{A-3}$$

Note that ς is deceasing in A_{PS} (see (A-2)). Hence, condition (A-3) requires the productivity of lawyers A_{PS} to be low enough.

Firm-level allocations

We first solve for the firm-level allocations, i.e. firm profits, firm employment and the productivity cutoff. Let p_G denote the price of the industrial good. If active, firm z_i solves the maximization problem

$$\pi(z_i) = \max_{H_{PMi},H_{PSi}\geq 0}\left\{p_G z_i^{1-\alpha-\beta} H_{PMi}^{\alpha}\left(A_{PS}H_{PSi}+\kappa\right)^{\beta} - w\left(H_{PMi}+H_{PSi}\right) - f_O w\right\}. \tag{A-4}$$

where $f_O w$ denotes the overhead costs. Note that we explicitly impose the constraint that $H_{PSi} \geq 0$. Firms operate if and only if $\pi(z_i) \geq 0$. We denote the productivity threshold by z^*, i.e., $\pi(z^*) = 0$. Under the condition (9), $z^* > A_M$, namely, there is a range of low-productivity firms that choose to be inactive.

Proposition 2. *Suppose that* $\varsigma \geq \frac{\beta}{1-\alpha}$*, where* ς *is given in* (A-2)*. Let* z^* *denote the endogenous productivity threshold, such that firm with* $z_i \geq z^*$ *will produce in equilibrium. Define*

$$z_L = z^* \left(\frac{1-\alpha}{\beta}\varsigma\right)^{\frac{1-\alpha}{1-\alpha-\beta}}, \tag{A-5}$$

Then:

1. *The productivity threshold is given by*

$$z^* = \left(\frac{w}{p_G}\frac{1}{\kappa^\beta \alpha}\left(\frac{\alpha}{1-\alpha}f_O\right)^{1-\alpha}\right)^{\frac{1}{1-\alpha-\beta}}. \tag{A-6}$$

2. *Optimal factor demands are given by*

$$H_{PS}(z_i) = \begin{cases} 0 & \text{if } z_i < z_L \\ \varsigma \frac{z_i - z_L}{z_L} f_0 & \text{if } z_i \geq z_L \end{cases}, \tag{A-7}$$

i.e. a firm hires lawyers if and only if $z \geq z_L$*. Moreover,*

$$H_{PM}(z_i) = \begin{cases} \frac{\alpha}{\beta} f_O \varsigma \left(\frac{z_i}{z_L}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} & \text{if } z_i < z_L \\ \frac{\alpha}{\beta}\varsigma \frac{z_i}{z_L} f_0 & \text{if } z_i \geq z_L \end{cases}. \tag{A-8}$$

3. *Firm-level profits are given by*

$$\pi(z_i) = \begin{cases} \left(\frac{1-\alpha}{\beta}\varsigma\left(\frac{z_i}{z_L}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} - 1\right) f_O w & \text{if } z_i < z_L \\ \left(\varsigma\left(1 + \left(\frac{1-\alpha-\beta}{\beta}\right)\frac{z_i}{z_L}\right) - 1\right) f_O w & \text{if } z_i \geq z_L \end{cases}. \tag{A-9}$$

Proof. See Section OA-1.1 in the Appendix. □

Note that (A-5) determines z_L directly as a function of z^* and that under our assumption that $\varsigma > \frac{\beta}{1-\alpha}$ indeed $z^* < z_L$ and all firms with $z_i \in [z^*, z_L]$ do not hire lawyers. As $\varsigma \to \frac{\beta}{1-\alpha}$, we have $z^* \to z_L$. Note also that the profit function in (A-9) is concave in z as long firms do not hire lawyers but linear in z once they hire lawyers.

Proposition 3. *Suppose that* $\varsigma < \frac{\beta}{1-\alpha}$*, where* ς *is given in* (A-2)*. Let* $\tilde{z}$ *denote the endogenous productivity threshold, such that firm with* $z_i \geq \tilde{z}$ *will produce in equilibrium. Then:*

1. *The productivity threshold is given by*

$$\tilde{z} = \frac{1}{1-\alpha-\beta}\left(\frac{w}{p_G}\right)^{\frac{1}{1-\alpha-\beta}}\left(\frac{1}{\alpha}\right)^{\frac{\alpha}{1-\alpha-\beta}}\left(\frac{1}{\beta A_{PS}}\right)^{\frac{\beta}{1-\alpha-\beta}} f_O(1-\varsigma). \tag{A-10}$$

2. *Optimal factor demands are given by*

$$\begin{aligned} H_{PMi} &= \frac{\alpha}{1-\alpha-\beta} f_O (1-\varsigma)\frac{z_i}{\tilde{z}} \\ H_{PSi} &= \frac{\beta}{1-\alpha-\beta} f_O (1-\varsigma)\frac{z_i}{\tilde{z}} - \frac{\kappa}{A_{PS}}. \end{aligned}$$

3. Firm-level profits are given by

$$\pi(z_i) = \pi(z_i) = \left(\frac{z-\tilde{z}}{\tilde{z}}\right) f_O (1-\varsigma) w. \tag{A-11}$$

Proof. See Section OA-1.1 in the Appendix. □

Free Entry and the Equilibrium Wage

Free entry requires that the cost of entry are equal to the expected profits, i.e.

$$f_E w = E[\pi] = \int \pi(x) f(x) dx.$$

This condition allows us to solve for the equilibrium real wage $\frac{w}{p_G}$.

Proposition 4. *Suppose that* $\varsigma \geq \frac{\beta}{1-\alpha}$. *Then*

$$\left(\frac{z_L}{A_M}\right)^{\lambda} = \frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left[\left(\varsigma\frac{1-\alpha}{\beta}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} + \frac{\varsigma}{\lambda-1}\right]\frac{f_O}{f_E} \tag{A-12}$$

$$\left(\frac{z^*}{A_M}\right)^{\lambda} = \frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left[1+\frac{1}{\lambda-1}\left(\frac{\beta}{1-\alpha}\right)^{\frac{(1-\alpha)\lambda}{1-\alpha-\beta}}\varsigma^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}\right]\frac{f_O}{f_E}, \tag{A-13}$$

and

$$\frac{w}{p_G} = \left(\frac{z_L}{\kappa}\right)^{1-\alpha-\beta}\alpha^{\alpha}(\beta A_{PS})^{1-\alpha} \tag{A-14}$$

$$= \left(\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left[1+\frac{1}{\lambda-1}\left(\frac{\beta}{1-\alpha}\right)^{\frac{(1-\alpha)\lambda}{1-\alpha-\beta}}\varsigma^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}\right]\frac{f_O}{f_E}\right)^{\frac{1-\alpha-\beta}{\lambda}} A_M^{1-\alpha-\beta}\kappa^{\beta}\alpha^{\alpha}\left(\frac{1-\alpha}{f_0}\right)^{1-\alpha} \tag{A-15}$$

Suppose that $\varsigma < \frac{\beta}{1-\alpha}$. *Then*

$$\left(\frac{\tilde{z}}{A_M}\right)^{\lambda} = \frac{1}{\lambda-1}\frac{f_O}{f_E}(1-\varsigma), \tag{A-16}$$

and

$$\frac{w}{p_G} = \left(\frac{\tilde{z}(1-\alpha-\beta)}{f_O(1-\varsigma)}\right)^{1-\alpha-\beta}(\beta A_{PS})^{\beta}(\alpha)^{\alpha} \tag{A-17}$$

$$= (1-\alpha-\beta)^{1-\alpha-\beta}\left(\frac{1}{\lambda-1}\frac{1}{f_E}\right)^{\frac{1-\alpha-\beta}{\lambda}}\left(\frac{1}{f_O(1-\varsigma)}\right)^{\frac{\lambda-1}{\lambda}(1-\alpha-\beta)} A_M^{1-\alpha-\beta}\left(\beta\frac{\kappa}{f_0\varsigma}\right)^{\beta}(\alpha)^{\alpha}. \tag{A-18}$$

Proof. See Section OA-1.3 in the Appendix. □

Proposition 4 characterizes the cutoffs and the real wage in terms of parameters. In particular, all cutoffs z_L are independent of the wage w. Note also that ς is decreasing in A_{PS} so that $\frac{\partial z_L}{\partial A_{PS}} < 0$, i.e. if lawyers become more productive, the cutoff to hire lawyers declines. Moreover, $\frac{\partial z^*}{\partial A_{PS}} > 0$.[30] Because ς is decreasing in A_{PS}, the

[30]Note that we assumed that $z^* > A_M$. Hence, we need to impose that

$$\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left[1+\frac{1}{\lambda-1}\left(\frac{\beta}{1-\alpha}\right)^{\frac{(1-\alpha)\lambda}{1-\alpha-\beta}}\varsigma^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}\right]\frac{f_O}{f_E} < 1$$

real wage is increasing in A_{PS} (see Section OA-1.5 in the Online Appendix).

Aggregate Labor Allocations

Now consider aggregate employment. In our economy, workers in the manufacturing sector are used for production work (H_{PM}), to provide PS (H_{PS}), pay for overhead (H_{OM}) and generate new business ideas (H_{EM}). Hence, labor market clearing requires that

$$H_G = H_{PS} + H_{PM} + H_{EM} + H_{OM}.$$

Proposition 5. *The number of entry and production workers is given by*

$$\begin{aligned} H_{EM} &= \frac{1-\alpha-\beta}{\lambda} H_G \\ H_{PM} &= \alpha H_G \end{aligned}$$

independent of ς. The number of firms, M, is given by

$$M = \frac{1-\alpha-\beta}{\lambda}\frac{H_G}{f_E}.$$

independent of ς. The number of lawyers and overhead workers is given by

$$H_{PS} = \begin{cases} \frac{\sigma(\varsigma)}{1+\sigma(\varsigma)}\frac{\beta+(\lambda-1)(1-\alpha)}{\lambda} H_G & \text{if } \varsigma \geq \frac{\beta}{1-\alpha} \\ \left(\beta - (1-\alpha-\beta)\frac{\lambda-1}{\lambda}\frac{\varsigma}{1-\varsigma}\right) H_G & \text{if } \varsigma < \frac{\beta}{1-\alpha} \end{cases}$$

and

$$H_{OM} = \begin{cases} \frac{1}{1+\sigma(\varsigma)} \times \frac{\beta+(\lambda-1)(1-\alpha)}{\lambda} H_G & \text{if } \varsigma \geq \frac{\beta}{1-\alpha} \\ (1-\alpha-\beta)\frac{\lambda-1}{\lambda}\frac{1}{1-\varsigma} H_G & \text{if } \varsigma < \frac{\beta}{1-\alpha} \end{cases},$$

where

$$\sigma(\varsigma) \equiv \frac{1}{\lambda-1}\left(\frac{\beta}{1-\alpha}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} \varsigma^{-\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}}.$$

Proof. See Section OA-1.4 in the Appendix. □

Note that $H_{PS} = \Xi(A_{PS})$, where $\Xi(A_{PS})$ is consistent with the expression in Proposition 1 for the case in which $\frac{\kappa}{f_O A_{PS}} > \frac{\beta}{1-\alpha}$. In particular, H_{PS} is increasing in A_{PS}, whereas H_{OM} is decreasing in A_{PS}. Interestingly, their sum is independent of A_{PS}, i.e.

$$\frac{H_{OM} + H_{PS}}{M} = \left(\frac{\beta + (\lambda-1)(1-\alpha)}{1-\alpha-\beta}\right) f_E.$$

This also implies that the endogenous number of firms M is given by

$$M = \frac{1-\alpha-\beta}{\lambda}\frac{H_G}{f_E}.$$

Hence, the number of ideas, which is generated does not depend on A_{PS}. But the number of ideas, which are actually implemented is decreasing in A_{PS} as the production cutoff z^* is increasing in A_{PS}. Hence, improvements in the productivity of lawyers induce selection by truncating the productivity distribution. Note that all these allocations are independent of A_M. This is in contract to the the micro-level, where employment shares vary systematically with firm productivity. In particular, (A-7) and (A-8) imply that for firms that hire lawyers (i.e.

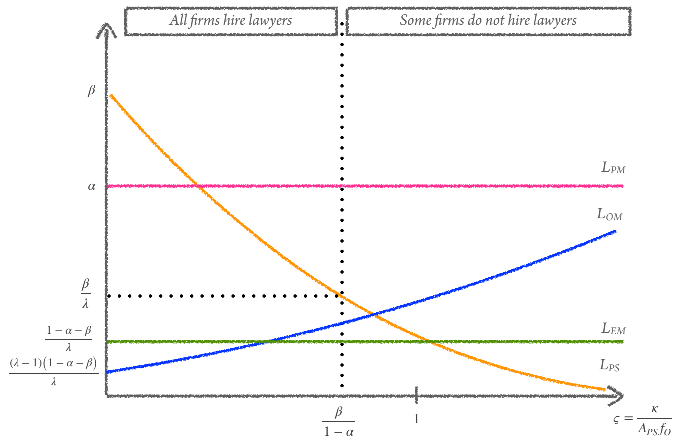

Figure A-1: Aggregate labor allocations

$z_i \geq z_L$) we have

$$\frac{H_{PS}(z)}{H_{PM}(z)} = \frac{\beta}{\alpha}\left(1 - \frac{z_i}{z_L}\right).$$

However, the aggregate employment of production workers hired by (large) firms who hire a positive share of lawyers (i.e. $z_i > z_L$) is given by

$$\frac{\int_{z \geq z_L} H_{PM}(z)\, dG(z)}{\int_{z \geq z_L} H_{PS}(z)\, dG(z)} = \lambda \frac{\alpha}{\beta}.$$

Hence, even though there is micro-heterogeneity in the intensity of hiring lawyers, the aggregate employment share of lawyers (among firms who hire lawyers) is constant and depends explicitly on the tail of the productivity distribution λ. A thicker tail, i.e. λ smaller, *increases* the aggregate employment share of lawyers by shifting resources towards large firms, which are lawyer intensive.

Figure A-1 depicts the allocation of employment as a function of $\varsigma = \frac{\kappa}{A_{PS} f_O}$ for both cases discussed above. Note that all employments are continuous at the threshold $\varsigma = \frac{\beta}{1-\alpha}$.

Aggregate Manufacturing Productivity

The free entry condition ensures that the industrial sector as a whole does not generate any profits. Hence, aggregate revenue is equal to aggregate labor payments

$$p_G Y_G = w H_G.$$

Proposition 6. *Let aggregate productivity A_G be defined by*

$$\frac{Y_G}{H_G} = A_G.$$

Then,

$$A_G = \begin{cases} Q_2 \left(\frac{1}{1-\varsigma}\right)^{\frac{\lambda-1}{\lambda}(1-\alpha-\beta)} \left(\frac{1}{\varsigma}\right)^{\beta} A_M^{1-\alpha-\beta} & \text{if } \varsigma < \frac{\beta}{1-\alpha} \\ Q_1 \left(1 + \frac{1}{\lambda-1}\frac{\beta}{1-\alpha}\left(\frac{1-\alpha}{\beta}\varsigma\right)^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}\right)^{\frac{1-\alpha-\beta}{\lambda}} A_M^{1-\alpha-\beta} & \text{if } \varsigma \geq \frac{\beta}{1-\alpha} \end{cases},$$

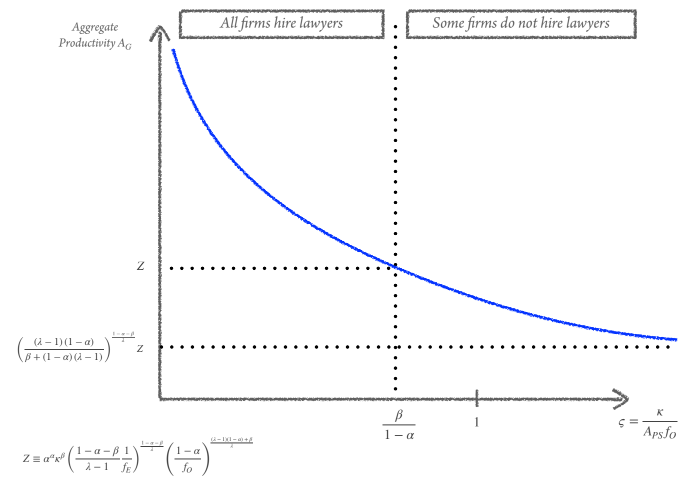

Figure A-2: Aggregate productivity A_G

where

$$Q_1 = \left(\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\frac{f_O}{f_E}\right)^{\frac{1-\alpha-\beta}{\lambda}} \kappa^{\beta}\alpha^{\alpha}\left(\frac{1-\alpha}{f_O}\right)^{1-\alpha}$$

and

$$Q_2 = \alpha^{\alpha}(1-\alpha-\beta)^{1-\alpha-\beta}\left(\frac{1}{\lambda-1}\frac{1}{f_E}\right)^{\frac{1-\alpha-\beta}{\lambda}}(\beta\kappa)^{\beta}\left(\frac{1}{f_O}\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{\lambda}}.$$

Proposition 6 follows directly from the fact that $A_G = w/p_G$ and the solution for w/p_G from Proposition 4. The importance of Proposition 6 is that it shows that the manufacturing sector is characterized by an aggregate production function for the industrial good sector, where total productivity in industrial production A_G is fully determined from parameters: the productivity of lawyers A_{PS} (encapsulated in ς), the level of productivity A_M, the overhead cost f_O and the entry cost f_E . Note that A_G is continuous in ς and satisfies

$$\lim_{\varsigma\to\infty} A_G = \left(\frac{(\lambda-1)(1-\alpha)}{\beta+(1-\alpha)(\lambda-1)}\right)^{\frac{1-\alpha-\beta}{\lambda}} \kappa^{\beta}\alpha^{\alpha}\left(\frac{1-\alpha-\beta}{\lambda-1}\frac{1}{f_E}\right)^{\frac{1-\alpha-\beta}{\lambda}}\left(\frac{1-\alpha}{f_0}\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{\lambda}} A_M^{1-\alpha-\beta}$$

and

$$A_G\left(\varsigma = \frac{\beta}{1-\alpha}\right) = \alpha^{\alpha}\kappa^{\beta}\left(\frac{1-\alpha-\beta}{\lambda-1}\frac{1}{f_E}\right)^{\frac{1-\alpha-\beta}{\lambda}}\left(\frac{1-\alpha}{f_O}\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{\lambda}}.$$

Figure A-2 depicts aggregate productivity A_G as a function of $\varsigma = \frac{\kappa}{A_{PS} f_O}$ for both cases discussed above. As for the employment allocations, aggregate productivity A_G is also continuous in ς.

A-3 Spatial Accounting

Consider a single period for now. We observe $\left\{[w_r^D]_r, H_{rF}, H_{rG}, H_{rCS}\right\}_r$. We indicate the observed wages by w_r^D to distinguish them from the model wages w_r as we did not pick a numeraire yet - see below. We want to infer $[A_{rF}, A_{rG}, A_{rCS}]_r$.

Step 1: Getting *relative* food productivities and *relative* manufacturing productivities Again, it is useful to write productivities as

$$\begin{aligned} A_{rF} &= A_F a_{rF} \text{ with } \sum_{r=1}^{R} a_{rF}^{\sigma-1} = 1 \\ A_{rG} &= A_G a_{rG} \text{ with } \sum_{r=1}^{R} a_{rG}^{\sigma-1} = 1. \end{aligned}$$

Then,

$$a_{rF} = \left(\frac{H_{rF} w_r^{\sigma}}{\sum_{r=1}^{R} (H_{rF}) w_r^{\sigma}} \right)^{\frac{1}{\sigma-1}} \tag{A-19}$$

$$a_{rG} = \left(\frac{H_{rG} w_r^{\sigma}}{\sum_{r=1}^{R} (H_{rG}) w_r^{\sigma}} \right)^{\frac{1}{\sigma-1}} \tag{A-20}$$

This means we have $R+2$ unknowns left

$$A_F, A_G, \{A_{rCS}\}_{r=1}^{R}.$$

Note a_{rF} and a_{rG} are insensitive to the level of w_r.

Step 2: Getting A_F and A_G The two prices of the tradable goods are

$$\begin{aligned} p_{Ft} &= \left(\sum_r \left(\frac{w_r}{A_{rF}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} = \frac{1}{A_F} \left(\sum_r \left(\frac{w_r}{a_{rF}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \\ p_{Gt} &= \left(\sum_r \left(\frac{w_r}{A_{rG}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} = \frac{1}{A_G} \left(\sum_r \left(\frac{w_r}{A_{rG}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}. \end{aligned}$$

Note that $\left(\sum_r \left(\frac{w_r}{a_{rF}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$ and $\left(\sum_r \left(\frac{w_r}{A_{rG}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$ are observable from (A-19) and (A-20) and are homogenous of degree 1 as a_{rF} and a_{rG} is homogeneous of degree zero. Hence, let us write

$$\Lambda_F^w \equiv \left(\sum_r \left(\frac{w_r}{a_{rF}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{A-21}$$

$$\Lambda_G^w \equiv \left(\sum_r \left(\frac{w_r}{A_{rG}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{A-22}$$

where $\Lambda_s^w(w)$ is known. The superscript "w" indicates that this is HD1 in the level of wages. To determine A_F and A_G we need two restrictions:

1. First we choose the manufacturing good as the numeraire. This determines A_G as

   $$p_{Gt} = 1 \qquad \Longrightarrow \qquad A_G = \Lambda_G^w.$$

   Note that an increase in w by a common factor increases A_G by the same amount.

2. Now suppose we observe the relative price of agriculture relative to manufacturing p_t^{AG}. Then

$$p_t^{AG} = \frac{p_t^A}{p_t^G} = \frac{A_F^{-1}\Lambda_F^w}{A_G^{-1}\Lambda_G^w} = A_F^{-1}\Lambda_F^w.$$

Hence, given p_t^{AG}, we can identify A_F. For a given relative price p_t^{AG}, a common increase in wages w increases A_F by the same amount.

Step 3: Getting $\{A_{rCS}\}_{r=1}^R$ and the structural parameters To derive (A-19) and (A-20) we used $R-1$ equations, i.e. all relative trade shares in the tradable sectors. This means that we still have the R equations for the non-tradable CS sector and the two aggregate resource constraints for the traceable goods. These are

$$\frac{H_{rCS}}{H_r} = \omega_{CS} + \nu_{CS}\left(p_{At}^{\omega_A} p_{rCSt}^{\omega_{CS}} p_{Gt}^{\omega_G}\right)^{\varepsilon}\left(E_{rt}[q]w_{rt}\right)^{-\varepsilon} \quad \text{(A-23)}$$

$$\sum_r w_r H_{rF} = \sum_{j=1}^R \left(\omega_A + \nu_A\left(p_{At}^{\omega_A} p_{CSjt}^{\omega_{CS}} p_{Gt}^{\omega_G}\right)^{\varepsilon}\left(E_{rt}[q]w_{rt}\right)^{-\varepsilon}\right) w_j H_j \quad \text{(A-24)}$$

$$\sum_r w_r H_{rG} = \sum_{j=1}^R \left((1-\omega_A-\omega_{CS}) - (\nu_A+\nu_{CS})\left(p_{At}^{\omega_A} p_{CSjt}^{\omega_{CS}} p_{Gt}^{\omega_G}\right)^{\varepsilon}\left(E_{rt}[q]w_{rt}\right)^{-\varepsilon}\right) w_j H_j \quad \text{(A-25)}$$

Note first that equation (A-25) is redundant, it is implied by (A-23) and (A-24) due to Walras' Law. Substituting the numeraire assumption $p_{Gt} = 1$ and the fact that $p_{At} = p_t^{AG}$, where p_t^{AG} is the relative price we are targeting and $p_{rCSt} = \frac{w_{rt}}{A_{rCSt}}$ yields

$$\frac{H_{rCS}}{H_r} = \omega_{CS} + \nu_{CS}\left(p_t^{AG}\right)^{\varepsilon\omega_A}\left(\frac{w_{rt}}{A_{rCS}}\right)^{\varepsilon\omega_{CS}}\left(E_{rt}[q]w_{rt}\right)^{-\varepsilon} \quad \text{(A-26)}$$

$$\sum_r w_r H_{rF} = \omega_A \sum_{j=1}^R w_j H_j + \nu_A\left(p_t^{AG}\right)^{\varepsilon\omega_A}\sum_{j=1}^R\left(\frac{w_{jt}}{A_{CSj}}\right)^{\varepsilon\omega_{CS}}\left(E_{rt}[q]\right)^{-\epsilon} w_{jt}^{1-\varepsilon} H_j. \quad \text{(A-27)}$$

For a given year these are $R+1$ equations in R productivities $\{A_{rCS}\}$ and 4 structural parameters $(\omega_{CS}, \omega_A, \nu_A, \nu_{CS})$ (recall that we take ε as given because we estimate it from the expenditure shares). If we have T years, we have

$$\begin{aligned}\text{Number of unknowns} &= TR+4\\ \text{Number of equations} &= T(R+1) = TR+T.\end{aligned}$$

Hence, by insisting that preferences are constant over time, we add over-identifying restrictions if we add additional years to our analysis. Note that (A-26) implies that

$$\left(p_t^{AG}\right)^{\varepsilon\omega_A}\left(\frac{w_{rt}}{A_{rCS}}\right)^{\varepsilon\omega_{CS}} = -\frac{1}{\nu_{CS}}\left(\omega_{CS} - \frac{H_{rCS}}{H_r}\right)\left(E_{rt}[q]w_{rt}\right)^{\varepsilon}.$$

Hence, substituting this into (A-27) yields

$$\sum_r w_r H_{rF} = \omega_A \sum_{r=1}^R w_r H_r - \frac{\nu_A}{\nu_{CS}}\sum_{r=1}^R\left(\omega_{CS} - \frac{H_{rCS}}{H_r}\right) w_{rt} H_r. \quad \text{(A-28)}$$

Given the data this is a single equation in ω_A, $\frac{\nu_A}{\nu_{CS}}$ and ω_{CS}. Note that (A-28) is HDZ, i.e. a common increase in the level of wages will leave ω_A, $\frac{\nu_A}{\nu_{CS}}$ and ω_{CS} unchanged. This leaves us with R equations for consumer service employment

$$\frac{H_{rCS}}{H_r} = \omega_{CS} + \nu_{CS} \left(p_t^{AG}\right)^{\varepsilon \omega_A} \left(\frac{w_{rt}}{A_{rCS}}\right)^{\varepsilon \omega_{CS}} \left(E_{rt}[q] w_{rt}\right)^{-\varepsilon} = \omega_{CS} + \nu_{CS} A_{rCS}^{-\varepsilon \omega_{CS}} \left(p_t^{AG}\right)^{\varepsilon \omega_A} \left(E_{rt}[q]\right)^{-\epsilon} w_{rt}^{-\varepsilon(1-\omega_{CS})}$$

From (A-26) it is seen that ν_{CS} is not separately identified from the *level* of productivity in the consumer service sector: holding ω_{CS} and ε fixed, the data on wages w_{rt} and employment shares $\frac{L_{rCS}}{L_r}$ identifies $\nu_{CS} A_{rCS}^{-\varepsilon \omega_{CS}}$. In Section OA-2 in the Online Appendix we provide more details on the computational algorithm how we estimate the productivity fundamentals.

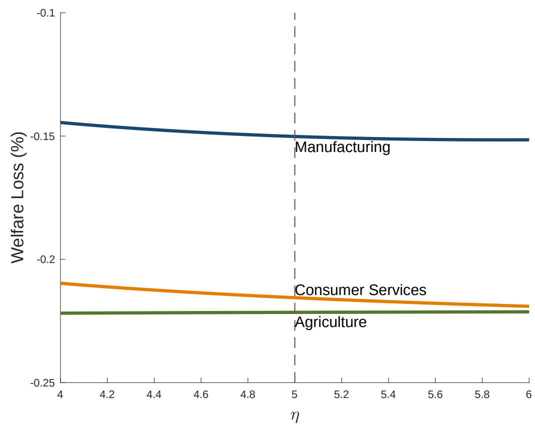

Figure A-3: Robustness Tests on Trade Elasticity.

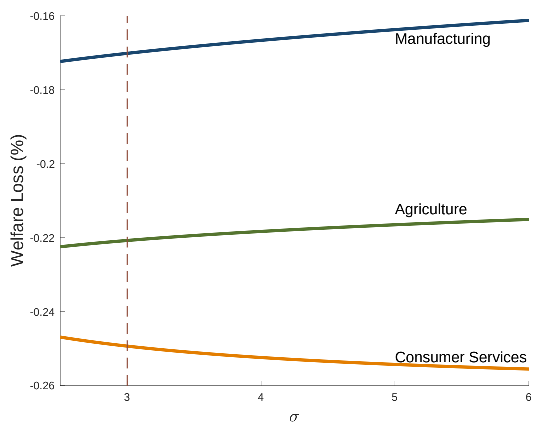

Figure A-4: Robustness Tests on Variety Elasticity.

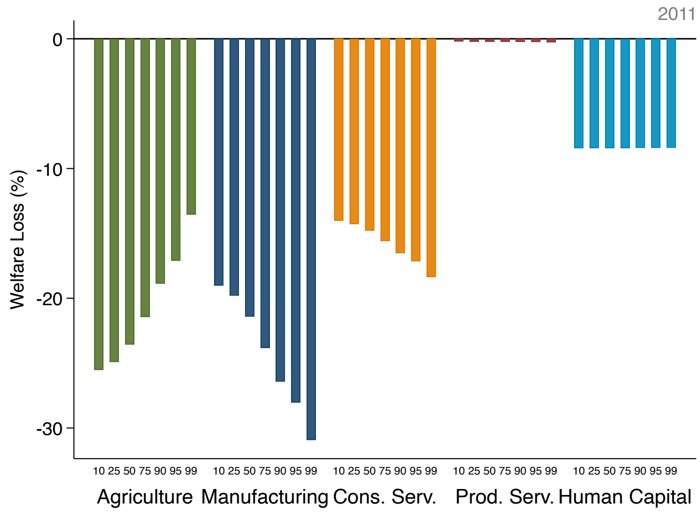

Figure A-5: WELFARE LOSS OF DIFFERENT LOCAL INCOME LEVELS

A-4 Welfare loss of Different Local Income Levels

A-5 Open economy

In this model we present the formal analysis for the open economy extension discussed in Section 6.1.

A-5.1 Environment and Equilibrium

We assume that the consumption of the physical good of consumers in India is a combination of domestic and imported goods with a constant elasticity of substitution η:

$$C_G = \left(C_{G,D}^{\frac{\eta-1}{\eta}} + \varphi C_{G,ROW}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}.$$

Here, $C_{G,D}$ and $C_{G,ROW}$ are the physical quantities of the domestic and imported physical good, φ is a taste parameter capturing the preference for the imported good, and η is the elasticity of substitution that we interpret as a trade elasticity.

Letting $p_{G,D}$ and $p_{G,ROW}$ denote the respective prices, the price index of the bundle C_G is given by

$$P_G = \left(p_{G,D}^{1-\eta} + \varphi^{\eta} p_{G,ROW}^{1-\eta} \right)^{\frac{1}{1-\eta}}. \tag{A-29}$$

The expenditure share on Indian goods is $\frac{p_{G,D} C_G}{P_G C_G} = \left(\frac{P_{G,D}}{P_G}\right)^{1-\eta}$. Combining this expression with Equation (A-29) yields the expenditure shares

$$
\begin{aligned}
\frac{p_{G,D}C_{G,D}}{P_G C_G} &= \frac{\varphi^{-\eta}\left(\frac{P_{G,D}}{p_{G,ROW}}\right)^{1-\eta}}{1+\varphi^{-\eta}\left(\frac{P_{G,D}}{p_{G,ROW}}\right)^{1-\eta}}, \\
\frac{p_{G,ROW}C_{G,ROW}}{P_G C_G} &= \frac{1}{1+\varphi^{-\eta}\left(\frac{P_{G,D}}{p_{G,ROW}}\right)^{1-\eta}}.
\end{aligned}
$$

Although we do not model explicitly trade costs, these are reflected in the relative price of foreign goods.

The Indian economy is assumed to export both domestic goods and a special category of services that is traded internationally: ICT exports. Consider first the export of goods. We model total spending on Indian goods (in term of domestic goods) from the foreigner sector as

$$
X_{G,D} = \frac{\varphi^{-\eta}\left(\frac{P_{G,D}}{p_{G,ROW}}\right)^{1-\eta}}{1+\varphi^{-\eta}\left(\frac{P_{G,D}}{p_{G,ROW}}\right)^{1-\eta}}\Upsilon_G,
$$

where $X_{G,D}$ are total export sales from India, Υ_G is a demand shifter for goods and $p_{G,ROW}$ denotes the price of goods in the foreign sector. For simplicity we assume the price elasticity of exports and imports to be the same and equal to η. .

Consider, next, the exported (ICT) services. For simplicity, we assume that ICT services are not sold in the domestic market but only internationally. We assume that the foreign sector buys a bundle of regional varieties ICT services

$$
Y_{ICTt} = \left(\sum_{r=1}^{R}(y_{rICTt})^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}},
$$

where y_{rICTt} denotes the quantity of services produced in region r and exported to the rest of the world. ICT services are produced in region r according to the following production function:

$$
y_{rICTt} = A_{rICTt}L_{rt}.
$$

Hence, the price of ICT services is given by

$$
p_{ICTt} = \left(\sum_r p_{rICTt}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = \left(\sum_r \left(\frac{w_{rt}}{A_{rICTt}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}}.
$$

As we do for goods, we model the import demand for ICT services as

$$
X_{ICT} = p_{ICTt}^{1-\eta}\Upsilon_{ICT}.
$$

Again, any trade costs are subsumed in the demand shifter Υ_{ICT}.

Equilibrium

The equilibrium with trade is pinned down by the following equilibrium conditions:

1. Market clearing for agricultural goods:

$$
w_r H_{rF} = \left(\frac{w_r^{1-\sigma}A_{rF}^{\sigma-1}}{\sum_{j=1}^{R} w_j^{1-\sigma}A_{jF}^{\sigma-1}}\right)\times\left(\sum_{j=1}^{R}\vartheta_{jF}\left(w_j-\tau_j\right)H_j\right)
$$

2. Market clearing for manufacturing goods:

$$w_r H_{rG} = \left(\frac{w_r^{1-\sigma} A_{rG}^{\sigma-1}}{\sum_{j=1}^{R} w_j^{1-\sigma} A_{jG}^{\sigma-1}}\right) \times \underbrace{\left(\underbrace{\frac{\varphi^{-\eta}\left(\frac{\left(\sum_j w_j^{1-\sigma} A_{jG}^{\sigma-1}\right)^{\frac{1}{1-\sigma}}}{p_{G,ROW}}\right)^{1-\eta}}{\varphi^{-\eta}\left(\frac{\left(\sum_j w_j^{1-\sigma} A_{jG}^{\sigma-1}\right)^{\frac{1}{1-\sigma}}}{p_{G,ROW}}\right)^{1-\eta}+1} \sum_{j=1}^{R} \vartheta_{jG}\left(w_j-\tau_j\right) H_j}_{\text{Domestic spending}} + \underbrace{\left(\sum_j w_j^{1-\sigma} A_{jG}^{\sigma-1}\right)^{\frac{1-\eta}{1-\sigma}} \Upsilon_G}_{\text{Total exports}} \right)}_{\text{Aggregate demand for physical goods}}$$

3. Market clearing for local PS:

$$w_r L_{rCS} = \vartheta_{rCS}\left(w_j - \tau_j\right) L_r.$$

4. Market clearing for local ICT services:

$$w_r H_{rICT} = \left(\frac{w_r^{1-\sigma} A_{rICT}^{\sigma-1}}{\sum_{j=1}^{R} w_j^{1-\sigma} A_{jICT}^{\sigma-1}}\right) \times \underbrace{\left(\sum_j w_j^{1-\sigma} A_{jICT}^{\sigma-1}\right)^{\frac{1-\eta}{1-\sigma}} \Upsilon_{ICT}}_{\text{ICT exports}}$$

5. Labor market clearing:

$$H_{rF} + H_{rG} + H_{rCS} + H_{rICT} = H_r.$$

6. Balanced Trade:

$$\underbrace{\left(\left(\sum_j w_j^{1-\sigma} A_{jG}^{\sigma-1}\right)^{\frac{1-\eta}{1-\sigma}} \Upsilon_G + \left(\sum_j w_j^{1-\sigma} A_{jICT}^{\sigma-1}\right)^{\frac{1-\eta}{1-\sigma}} \Upsilon_{ICT}\right)}_{\text{Exports}} = \underbrace{\frac{\sum_{j=1}^{R} \vartheta_{jG}\left(w_j-\tau_j\right) H_j}{\varphi^{-\eta}\left(\frac{\left(\sum_j w_j^{1-\sigma} A_{jG}^{\sigma-1}\right)^{\frac{1}{1-\sigma}}}{p_{G,ROW}}\right)^{1-\eta}+1}}_{\text{Imports}}$$

The balanced trade implies that the trade balance term that distributed to households $\tau_j = 0$.

Letting $x \equiv \varphi^{\eta} {p_{G,ROW}}^{1-\eta}$ denote the (scaled) terms of trade, these are $5R+1$ equations in $5R+1$ unknowns $\{x, \{w_r, H_{rF}, H_{rG}, H_{rCS}, H_{rICT}\}_r\}$. Again, we can pick a numeraire

$$p_{G,IND} = \left(\sum_r \left(\frac{w_r}{A_{rG}}\right)^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = 1.$$

Given the productivities $\{A_{rF}, A_{rG}, A_{rCS}, A_{rICT}\}_r$, the population distribution $\{H_r\}_r$, the demand shilfters of the foreign sector $(\Upsilon_{ICT}, \Upsilon_G)$ and the other preference parameters of the model, we can calculate $\{x, \{w_r, H_{rF}, H_{rG}, H_{rCS}, H_{rICT}\}_r\}$.

A-5.2 Identification of Trade Equilibrium Parameters

For the economy with trade we need to identify the following additional objects:

$$\left\{ [A_{rICT}]_{r=1}^{R}, \Upsilon_G, \Upsilon_{ICT} \right\}.$$

There are $R+2$ unknowns. For these $R+2$ unknowns we have the following conditions

1. Relative ICT payments across localities for ICT exports:

$$\frac{w_r H_{rICT}}{w_j H_{jICT}} = \frac{w_r^{1-\sigma} A_{rICT}^{\sigma-1}}{w_j^{1-\sigma} A_{jICT}^{\sigma-1}}$$

   These are $R-1$ equations to determine A_{rICT} up to scale, i.e.

$$A_{rICT} = A_{ICT} a_{rICT} \text{ with } \sum_r a_{rICT}^{\sigma-1} = 1$$

   yields

$$a_{rICT} = \left(\frac{H_{rICT} w_r^{\sigma}}{\sum_j H_{jICT} w_j^{\sigma}} \right)^{\frac{1}{\sigma-1}}$$

   Because the level of ICT productivity A_{ICT} is not separately identified from the aggregate demand shifter Υ_{ICT}, without loss of generality we can set $A_{ICT} = 1$.[31]

2. To identify Υ_{ICT} we use that

$$\begin{aligned} \sum_r w_r L_{rICT} &= \sum_r \left(\frac{w_r^{1-\sigma} A_{rICT}^{\sigma-1}}{\sum_{j=1}^{R} w_j^{1-\sigma} A_{jICT}^{\sigma-1}} \right) \left(\sum_j w_j^{1-\sigma} A_{jICT}^{\sigma-1} \right)^{\frac{1-\eta}{1-\sigma}} \Upsilon_{ICT} \\ &= \left(\sum_j w_j^{1-\sigma} a_{jICT}^{\sigma-1} \right)^{\frac{1-\eta}{1-\sigma}} \Upsilon_{ICT}. \end{aligned} \tag{A-30}$$

   The RHS is total value added of the ICT sector, which we can calculate directly in the data. Given that w_j and a_{jICT} is observed, we can calculate Υ_{ICT}.

3. To identify Υ_G we use a moment about the share of manufacturing value added that is exported. Our model implies that:

$$\text{Total value added in manufacturing} = \sum_r w_r H_{rG}$$

   and

$$\text{Total value added of exports} = \left(\sum_j w_j^{1-\sigma} A_{jG}^{\sigma-1} \right)^{\frac{1-\eta}{1-\sigma}} \Upsilon_G$$

[31] To see this, note that the equilibrium condition for ICT exports implies that

$$w_r H_{rICT} = \left(\frac{w_r^{1-\sigma} A_{rICT}^{\sigma-1}}{\sum_j^R w_j^{1-\sigma} A_{jICT}^{\sigma-1}} \right) \left(\sum_j w_j^{1-\sigma} A_{jICT}^{\sigma-1} \right)^{\frac{1-\eta}{1-\sigma}} \Upsilon_{ICT} = \left(\frac{w_r^{1-\sigma} a_{rICT}^{\sigma-1}}{\sum_j^R w_j^{1-\sigma} a_{jICT}^{\sigma-1}} \right) \left(\sum_j w_j^{1-\sigma} a_{jICT}^{\sigma-1} \right)^{\frac{1-\eta}{1-\sigma}} A_{ICT}^{\eta-1} \Upsilon_{ICT}$$

Hence, Υ_{ICT} and A_{ICT} are not separately identified.

Hence, the share of value added in the manufacturing sector is

$$M_1 = \frac{\left(\sum_j w_j^{1-\sigma} A_{jG}^{\sigma-1}\right)^{\frac{1-\eta}{1-\sigma}} \Upsilon_G}{\sum_r w_r H_{rG}} = \frac{P_{G,IND}^{1-\eta} \Upsilon_G}{\sum_r w_r H_{rG}} = \frac{\Upsilon_G}{\sum_r w_r H_{rG}} \tag{A-31}$$

Hence, for a given moment of the export share of manufacturing M_1 and data on $\{w_j, L_{jG}\}_j$ we can solve for Υ_G.

APPENDIX B: DATA AND MEASUREMENT

In this section, we discuss data and empirical issues discussed in the text.

B-1 Geography

The border of numerous Indian districts have changed between 1987 and 2011. The left panel of Figure B-1 plots the districts' boundaries in 2004 and 2011. The purple line represents the boundaries in 2004 and the red line represents the boundaries in 2011, we can see boundary changes across the country.

The most common type of change is a partition–one district has been separated into several districts in the later year. The second type is a boundary moving–the shared boundary of two districts has been moved. The third one is merge–two districts have been merged into a sinle district. In addition, there were several complicated changes over the years. Figure B-2 we plots two examples of boundary changes: partition (left panel) and boundary moving (right panel).

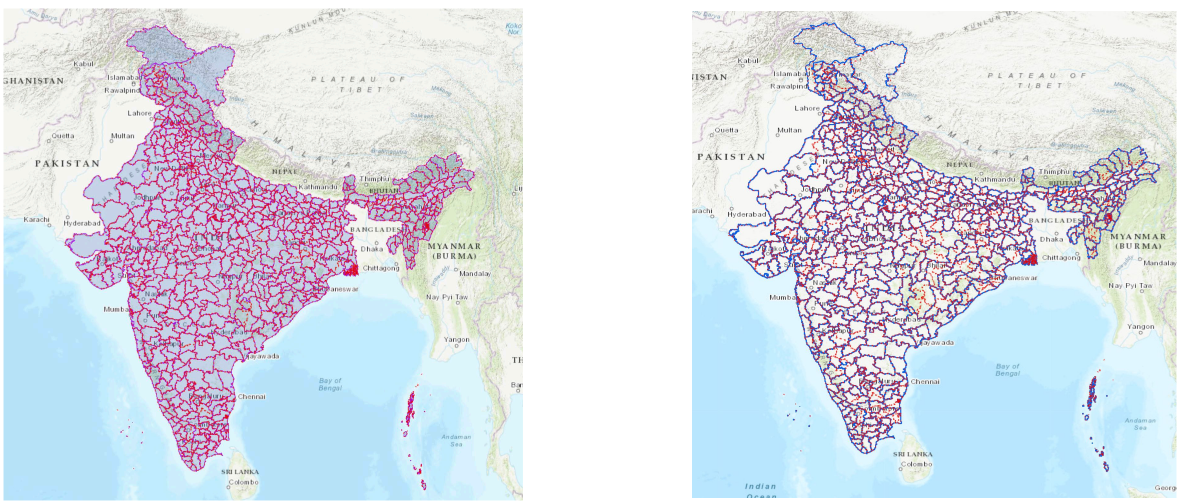

Figure B-1: District Borders in India: 1987 - 2011. The left figure plots the districts' boundaries in 2004 and 2011. The purple line represents the boundaries in 2004 and the red line represents the boundaries in 2011. The right figure shows the official Indian districts in the year 2011 (dashed red lines) and the time-invariant geographical units (*districts* we construct (solid blue lines) upon which our analysis is based.

In order to carry out the analysis on a panel of districts with a consistent geography, we construct regions that have consistent boundaries in 1987 and 2011. To keep the number of regions as large as possible, a region is always the smallest area that covers a single district or a set of districts with consistent boundaries over time. For instance, in the case of partition, the region is constructed as the district in the previous year. In the case of boundary moving, a region is constructed as the sum of two districts. We do it for three years' maps and arrive at a regional map with consistent boundaries from 1991 to 2011. Once we have regions with consistent boundaries, we can map our micro-data including NSS, Economic Census, and USS to the regions to construct panel data. The right panel of Figure shows the official Indian districts in the year 2011 (dashed red lines) and the time-invariant geographical units (*districts* we construct (solid blue lines) upon which our analysis is based.

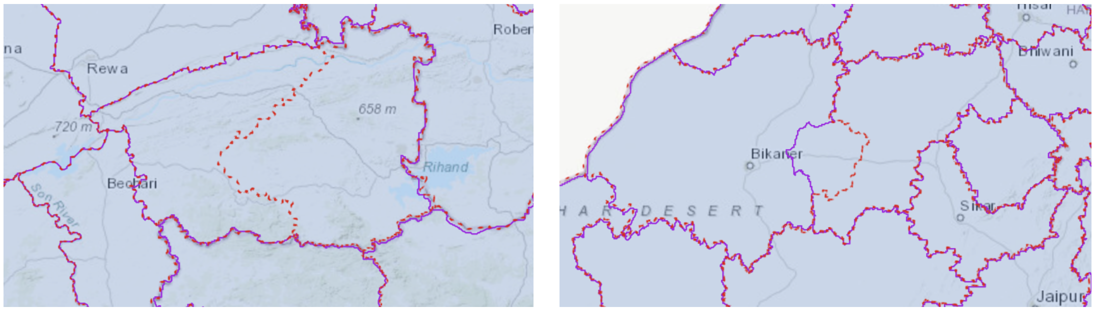

Figure B-2: Two Types of Boundary Change

B-2 Data

We use the following datasets

B-2.1 NSS

(Description of NSS data; which years do we have; number of observations per year; we take the following variables ...., a table with important summary statistics)

The National Sample Survey (NSS) is a representative survey conducted by the Government of India to collect socio-economic data at the household level since 1950. Each round of survey consists of several schedules to cover different subjects like “Consumer Expenditure”, “Employment and Unemployment”, “Participation in Education”, etc. We use data from rounds 43, 55, 60, 64, 66 and 68 of NSS, which span the years 1987 to 2011.

Geographical Coverage: The survey covers the whole of the Indian Union except a few regions due to unfavorable field conditions. For example, Ladakh and Kargil districts of Jammu & Kashmir, some interior villages of Nagaland, and villages in Andaman and Nicobar Islands are not covered in some rounds of the survey.

Information Coverage: We mainly use the “Household Consumer Expenditure” schedule and “Employment and Unemployment” schedule. Both schedules collect Household characteristics (such as Household size, Household type (agriculture \non-agriculture\regular wage earning\casual labour\self-employed), land ownership, and religions) and Demographic particulars of household members (such as Sex, Age, Marital Status, and Education). In addition, the “Household Consumer Expenditure” schedule collects household expenditure information on different kinds of goods and services. The ”Employment and Unemployment” schedule collects information such as employment status, occupation, job sectors, contract type, time disposition.

Concepts and Definitions:

Employment: An individual is defined as being employed if his/her usual principal activity status is one of the following: (i) worked in household enterprises (self-employed); (ii) worked as helper in household enterprises; (iii) worked as regular salaried/wage employee; (iv) worked as casual wage labour in public works; (v) worked as casual wage labour in other types of work

Education level: we classify individual's education into four levels: (i) Less than primary; (ii). Primary, upper primary, and middle; (iii). Secondary; (iv). More than secondary.

Household size The size of a household is the total number of persons in the household

Expenditure/Consumption: The survey collects households' consumption of various kinks of food, entertainment, sundry articles,consumer services, rent/ house rent and so on during last 30 days (monthly-based expenditure), and consumption of clothing, bedding, footwear, education and medical goods and services, and various kinds of durable goods during last 365 days (yearly-based expenditure). Total monthly household consumer expenditure is the sum of monthly-based expenditure and (30/365) * yearly-based expenditure

Expenditure per capita: Total household consumer expenditure divided by household size. In the main analysis, we top-code expenditure data at 98th percentiles to deal with extreme values.

Income: Total household consumer expenditure divided by the number of people between 15 and 65 in the household. In the main analysis, we top-code wage data at 98th percentiles to deal with extreme values.

Table B-1 reports important summary statistics of NSS.

Table B-1: National Sample Survey of India: summary statistics

round	Year	households	individuals	household size	employment rate	urbanization rate	expenditure per capita	Income per labor
43	1987-1988	126,353	654,903	5.18	36.89%	22.60%	2066.23	3873.32
55	1999-2000	107,215	596,688	5.57	36.53%	25.42%	6161.96	10929.28
60	2004	59,042	303,233	5.14	36.92%	25.51%	7318.92	12673.22
64	2007-2008	125,578	572,254	4.56	37.06%	26.32%	9713.66	16216.31
66	2009-2010	100,957	459,784	4.55	36.49%	27.32%	12987.22	21074.68
68	2011-2012	101,717	456,970	4.49	35.44%	28.84%	17507.15	28142.28

B-2.2 Economic Census

(Description of data; which years do we have; number of observations per year; we take the following variables .... a table with important summary statistics)

India Economic Census (EC) is the complete count of all establishments (i.e. units engaged in production or distribution of goods and services not for the purpose of sole consumption) located within the country. The Censuses were conducted in the years 1977, 1980, 1990, 1998, 2005, 2013, 2019. Micro-level data in 1990, 1998, 2005, 2013 are public available.

Industries Coverage: The Census covers all sectors excluding crop production and plantation. The Censuses in 2005 and 2013 exclude some public sectors like public administration, defence and compulsory social security.

Geographical Coverage: The Census covers all States and Union Territories of the country except for the year 1990, which covers all States/UTs except Jammu and Kashmir.

Information Coverage: The Census collects information such as firms' location, industry, ownership, owner's social Group, nature of operation, finance source, employment (male/female, hired/not hired). The following table reports important summary statistics of Censuses.

Census	Year	Number	Total	Employment distribution			
		of firms	employment	Average	1 employee	Less than 5 employees	More than 100 employees
Third EC	1990	24216790	74570280	3.08	53.77%	91.24%	0.13%
Fourth EC	1998	30348881	83308504	2.75	51.18%	91.71%	0.11%
Fifth EC	2005	41826989	100904120	2.41	55.76%	93.17%	0.12%
Sixth EC	2013	58495359	131293872	2.24	55.47%	93.44%	0.06%

Table B-2: The Economic Census: Summary Statistics

B-2.3 Service Sector in India: 2006-2007

The Service Sector in India (2006-2007) dataset was a part of an integrated survey under NSSO(National Sample Survey Organisation) in its 63rd round of survey of enterprises and households. In the 62nd round of NSSO(2001-2002), the dataset was called Unorganized Service Sector. With, the inclusion of financial sector and large firms, the dataset is renamed as Service Sector in India and is designed to be representative for India's service sector. Table 2 shows the consistency of Service Sector in India by comparing it's characteristic of firm sizes with that of the Economic Census dataset.

Industries Coverage (Covered): The Service Sector survey covers a broad ranges of service sectors,including hotels and restaurants (Section H of NIC 04); transport, storage and communication (I); financial intermediation (J); real estate, renting and business activities (K); education (M); health and social work (N) and other community, social and personal service activities (O)

Industries Coverage (Not Covered): However, the industries does not cover the following industries: Railways Transportation, Air transport, Pipeline transport; Monetary intermediation(central banks, commercial banks, etc); Trade Unions. Moreover, the following units are not surveyed: Government and Public sector enterprises, Government aided education institutions, Service sector units appeared in the *Annual Survey of Industries frame(ASI 2004-2005).*
Geographical Coverage: The survey covers the whole of Indian Union except for 4 districts and some remote interior villages.
information Coverage: In the context of our analysis, the survey provides two important information:

- The number of employees in a surveyed unit
- Where is the major destination of output for a surveyed unit: Firm, Households, Non-resident, and so on

Locations and Number: The survey covered the whole of the Indian Union except (i) Leh (Ladakh), Kargil, Punch and Rajauri districts of Jammu & Kashmir, (ii) interior villages situated beyond 5 km of a bus route in Nagaland, (iii) villages of Andaman and Nicobar Islands, which remain inaccessible throughout the year. The survey was conducted in a total number of 5573 villages and 7698 urban blocks. A total of 1,90,282 enterprises (including 438 list frame enterprises) were ultimately surveyed.

To check the representativeness of this survey, we compare firm size from Economic Census 2005 and this survey. The following table reports the average firm size and the share of firms with less than 5 employees for both the Service Survey and the Economic Census.

Sector	Number of firms		Average employment		Less than 5 employees	
	Census	Service Survey	Census	Service Survey	Census	Service Survey
Hotels and restaurants	1,499,101	30,744	2.52	2.49	0.90	0.91
Land transport; transport via pipelines	1,317,904	41,065	1.67	1.24	0.97	0.99
Post and telecommunications	723,119	22,885	2.06	1.41	0.96	0.99
Other business activities	519,696	10,610	2.81	1.92	0.90	0.95
Renting of machinery and household goods	365,246	5,387	2.00	1.77	0.94	0.97
Financial intermediation	221,953	12,984	6.27	4.15	0.63	0.79
Transport activities; travel agencies	188,474	2,101	3.40	3.33	0.86	0.85
Real estate activities	70,128	3,648	2.18	1.64	0.93	0.96
Computer and related activities	66,414	1,060	6.01	13.45	0.83	0.86
Activities auxiliary to financial intermediation	45,449	2,601	2.41	1.77	0.93	0.96
Insurance and pension funding	26,087	746	5.52	2.30	0.83	0.99
Water transport	7,914	174	4.35	1.92	0.90	0.98
Research and development	2,097	5	16.66	4.58	0.66	0.89

Table B-3: ECONOMIC CENSUS AND SERVICE SURVEY. The table reports statistics about firms' number and employment from the Economic Census 2005 and Service Survey 2006.

B-2.4 Informal Non-Agricultural Enterprises Survey 1999-2000 (INAES)

We use this dataset when splitting the construction sector into consumer service and producer service. The enterprise survey is one schedule of NSS 55th round. It covers all informal enterprises in the non-agricultural sector of the economy, excluding those engaged in mining & quarrying and electricity, gas & water supply[32]. The survey collects

[32] In India, organised and unorganised sectors are defined as follows:
organised sector: factories registered under Sections 2(m)(i) and 2(m)(ii) of the Factories Act, 1948, where 2(m)(i) includes manufacturing factories which employ 10 or more workers with power, and 2(m)(ii) includes manufacturing factories which 20 or more without power.
unorganised sector: all enterprises not covered in the organised sector
Informal sector is a subset of the unorganised sector. The unorganised sector includes four types of enterprises: 1) unincorporated proprietary enterprises; 2) partnership enterprises; 3) enterprises run by cooperative societies, trusts, private; 4) public limited companies; The informal sector only includes firms in category 1) and 2).

information on operational characteristics, expenses, value-added, fixed asset, loan, factor income. In particular, the survey asks the major destination agency for sale of final product/service. This information helps us to identify if a construction firm is consumer-oriented or producer-oriented

B-3 Industry Classification

We divide economic activities into 6 industries: Agriculture, Manufacture, Construction and Utility, Service, Information and Communications Technology (ICT), and Public based on 2-digit (sometimes 3-digit) level of India National Industrial Classification (NIC). Table B-4 reports the broad structure of industry classification.

Table B-4: Broad Industries based on India National Industry Classification 2008

Industry	NIC 2008	Section
Agriculture	01-03	Agriculture, forestry and fishing
Manufacture	05-09	Mining of coal and lignite
	10-33	Manufacturing
Construction & Utility	35	Electricity, gas, steam and air conditioning supply
	36-39	Water supply; sewerage, waste management and remediation activities
	41-43	Construction
Service	45-47	Wholesale and retail trade; repair of motor vehicles and motorcycles
	49-53	Transportation and storage
	55-56	Accommodation and Food service activities
	581	Publishing of books, periodicals and other publishing activities
	64-66	Financial and insurance activities
	68	Real estate activities
	69-75	Professional, scientific and technical activities
	77-82	Administrative and support service activities
	86-88	Human health and social work activities
	90-93	Arts, entertainment and recreation
	94-96	Other service activities
	97	Activities of households as employers of domestic personnel
ICT	582-63	Information and communication
Public	84	Public administration and defence; compulsory social security
	85	Education
	99	Activities of extraterritorial organizations and bodies

Note: sector 98 is dropped.

What's more, data in different years use different versions of NIC. We construct a concordance table between 2-digit industries of different versions' NIC based on Official NIC documents and detailed sector descriptions. Sometimes we have to dig into 3 or 4 digit level to get a precise match.

B-4 Bounding regional employment shares

Our model implies that regional CS expenditure shares and hence employment shares are bounded from above by ω_{CS}. Similarly, our model implies that the share of producer service employment ("lawyers") relative to production workers is bounded from above by β. Given our estimated of $\omega_{CS} = 0.607$ and $\beta = 0.7$ reported in Table 3, 7 districts violate this requirement and we drop them from our analysis. Dropping such districts is inconsequential because these districts are very small. In Table B-6 we report the aggregate population share of such districts by year. As can be seen, the aggregate importance of such districts is below 1% in almost all years

Table B-5: Concordance between 2-digit industry Classes of NICs

sector	NIC-1987	NIC-1998 & NIC-2004	NIC-2008
	Agriculture		
Agriculture and hunting	00-04	01	01
Forestry and logging	05	02	02
Fishing and aquaculture	06	05	03
	Manufacture		
Coal, lignite, and peat	10	10	05, 0892
crude petroleum and natural gas	11,19	11	06, 091
Metal ores	12, 13, 14	12,13	07
Other mining and quarrying	15	14	08(except0892), 099
Food products	20,21, 220-224	15	10, 11
Tobacco products	225-229	16	12
Textiles and wearing apparel	23 24	17, 18	13, 14
Leather products	29(except 292)	19	15
Wood products	27(except 276-277)	20	16
Paper products, printing and publishing	28	21, 22	17, 18, 581
Refined petroleum	314-319	23	19
Chemicals	30	24	20, 21
Rubber and plastics products	310-313(except3134)	25	22
Other non-metallic mineral products	32	26	23
Basic metals	33(except338)	27	24
Fabricated metal	34(except342), 352, 391	28, 2927	25, 3311
Machinery and equipment	35-36(except352), 390, 392, 393, 395, 396, 399	29-32 (except2927)	261-264, 268, 27, 28, 3312, 3314, 3319, 332, 9512
Medical, precision and optical instruments	380-382	33	265-267, 325, 3313
Transport equipment	37, 397	34, 35	29, 30, 3315
Furniture	276, 277, 3134, 342	361	31
Other manufacturing	383-389	369	32(except325)
	Construction & Utility		
Electricity, gas, steam supply	40, 41, 43	40	35
Water supply	42	41	36
Sewerage and waste treatment	338, 6892, 91	37,90	37, 38, 39
Construction	50, 51	45	41, 42, 43
	Service		
Wholesale	398, 60-64, 682, 686, 890, 974	50, 51(except51901)	45, 46
Retail	65-68(except682,686,6892)	52(except526,52591)	47
Repair services	97(except974)	526	952
Land transport	70	60	49
Water transport	71	61	50
Air transport	72	62	51
Supporting and auxiliary transport activities	730, 731, 732, 737, 738, 739, 74	63	52, 79
Post and telecommunications	75	64	53, 61
Hotels	691	551	55
Restaurants	690	552	56
Computer and related activities	394, 892, 897	72, 922	582, 62, 63, 9511
Financial service	80	65, 67	64, 66
Insurance and pension	81	66	65
Real estate activities	82	70	68
Legal activities	83	7411	691
Accounting	891	7412	692
Business and management consultancy	893	7413, 7414	70, 732
Architecture and engineering	894, 895	742	71
Research and development	922	73	72
Advertising	896	743	731
Other business activities	898, 899	749	74, 78, 80, 81, 82
Renting	733, 734, 735, 736, 85	71	77
Health and social work	93, 941	85	75, 86, 87, 88
Recreational cultural and sporting activities	95	92(except922)	59, 60, 90, 91, 93
Gambling	84	51901, 52591	92
Membership organizations	94(except941)	91	94
Personal service	96, 99	93, 95	96, 97
goods-producing activities for own use	#N/A	96	981
services-producing activities for own use	#N/A	97	982
	Public		
Public administration and defence	90	75	84
Education	920-921	80	85
Extraterritorial organizations	98	99	99

Total share of population		
Year	CS empl. share >0.56	PS empl. share >0.7
1987	0	0.0031
1999	0.0039	0.0013
2004	0.0068	0.009
2007	0.0085	0.0027
2009	0.0071	0.0002
2011	0.0162	0.0054

Table B-6: Employment Shares of Top-Bottom Code Districts

B-5 Measuring Producer and Consumer Services

This section will start with an introduction on the general methodology of the measurement of CS and PS, followed by detailed explanation and procedures. Since the 'Service Sector in India'(SS) dataset provides the information of major destination of firms, we utilized this information to calculate the approximated share of employment in CS and CS sectors, respectively. This share is calculated industry-wise(at two digit level of NIC04) and is treated as proxies to the share of CS in India. The share from SS were later applied to the Economic Census(EC) dataset to calculate the total share of employment in each district. Every detail step of the procedure can be found in '[12] Prepare share of PS by district-year dataset.do' on 'Dropbox/India/Empirical Analysis/rawdata/Economic Census/dofile'

1. Calculating 2 digits division-firm size level PS share using USS 2005-2006 Information of different kinds(background, employment, etc) in the SS dataset are stored in different STATA files, with an unique key variable 'ID'. Therefore, the first step of the analysis is to merge dataset that contains information about employment with the dataset that contains the key variables of interest—Major destination of output. We then calculate the Producer Service share(PS share) from the constructed dataset. Specific decisions made are listed below:

   (a) Observations with missing value in employment or/and Major destination of output are dropped

   (b) All firms that has 'non-resident'(foreign) as their major destination of output are dropped

   (c) Generate firm size bins according to total employment(1-2, 3-20, 20+)

   (d) Generate the NIC industry indicator that classify industries using the first two-digit of NIC2004 Industrial Classification

   (e) If the 'Major destination of output' of a firm is resident financial enterprises or resident non-financial enterprises, we classified it as Producer Service firm

   (f) Formally, the PS share within a category of interest is the total employment of Producer Service firms within the category, divided by the total employment of that category. Here, we divide the SS dataset into different categories by firm size(see (c)) and industry(see (d)). The multiplier(weight) that comes with the SS dataset can be applied to calculate the PS share in this procedure.

   (g) We calculated the PS share per firm size per industry as stated in (f)

   (h) For some industries, there are not enough data in some firm size bins. For example, there are some industries that has few/no firms with total employment larger than 20. Here, we apply the 'monotonic rule'

   (i) 'monotonic rule': Within an industry, we posit that the larger a firm, the more likely that the firm will serve to producers. Therefore, if within a firm size bin(e.g. 20+) of an industry, there is fewer than 5 observations(5 included), then the PS share of this particular firm size bin should be at least as large as the PS share of the preceding level firm size bin(3-20 in this example)

   (j) 'monotonic rule': Consequently, if there is no data within a firm size bin of an industry, we used the PS share of the preceding level firm size bin as a proxy. If there is less or equal to 5 observations within a firm size bin of an industry and the PS share is smaller to that of the preceding level firm size bin, we use the PS share of the preceding level firm size bin instead.

(k) The Economic Census dataset uses different NIC Industrial Classifications for different years. Since we only use the 'Service Sector in India' of year 2006-2007 dataset and apply the same PS share to Economic Census of different years, we need to convert NIC2004(used by SS 2006-2007) to NIC1987(used by EC1990) and NIC2008(used by EC2013) prior to the PS share calculation. We used the official document to create mapping between different NIC classifications. Check '[12] Prepare share of PS by district-year dataset.do' for detail mapping.

After these procedures, the firm size bin and industry level PS share is produced and will be applied to the Economic Census dataset

2. Applying these shares to corresponding division-firm size bin level employment from Economic Censuses. The specific procedures are:

   (a) generate firm size bins according to total employment(1-2, 3-20, 20+)

   (b) calculate total employment at 2 digits division-firm size bin-region level

   (c) merge division-firm size bin level employment with corresponding PS share information from SS dataset

   (d) fill in PS share for sectors not covered by SS dataset. Some service divisions are not covered by USS and therefore we cannot directly measure their PS shares. We handle these cases one by one to make our estimation as accurate as possible. The following table summarizes how we deal with divisions without PS share information.

Table B-7: Estimating PS share for divisions not covered by USS

NIC2004	Industry	Approach
22	Publishing, printing and reproduction of recorded media	attribute all employment to Producer Service
50	Sale, maintenance and repair of motor vehicles and motorcycles; retail sale of automotive fuel	use the average PS share (at firm size bin level) from other sectors for which we have information
51	Wholesale trade and commission trade, except of motor vehicles and motorcycles	use the average PS share (at firm size bin level) from other sectors for which we have information
52	Retail trade, except of motor vehicles and motorcycles; repair of personal and household goods	attribute all employment to Consumer Service
62	Air transport	attribute all employment to Producer Service

   (e) Calculate Producer Service (Consumer Service) employment by multiplying total employment with PS (1-PS) share

   (f) Aggregate Producer Service and Consumer Service employment into GIS region level

In Figure B-3 we show the positive relationship between firm size and the probability of selling to other firms graphically in the raw data. Our procedure exploits this variation within industries. In Table B-8 we show that the same pattern is presents within two- and three-digit industries and whether or not we use sampling weights.

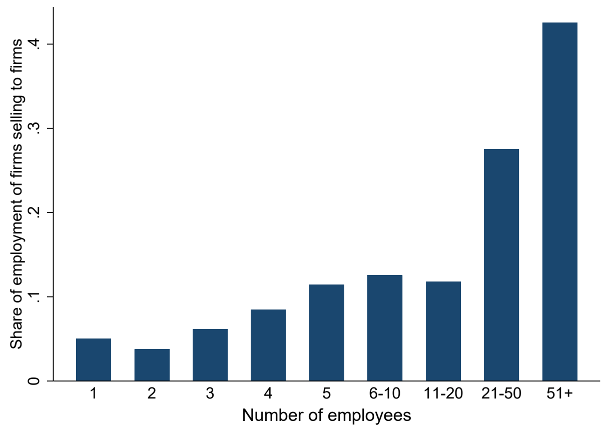

Figure B-3: Producer service share by firm size. The figure shows the share of service firms whose main customer are other firms (as opposed to private individuals) with breakdown by firm size.

	whether enterprises are the major destination of ouput			
2 employees	0.013***	0.014***	0.014***	0.016***
	(0.001)	(0.002)	(0.001)	(0.002)
3 employees	0.030***	0.028***	0.028***	0.029***
	(0.002)	(0.006)	(0.002)	(0.005)
4 employees	0.055***	0.063***	0.049***	0.059***
	(0.004)	(0.011)	(0.004)	(0.011)
5 employees	0.080***	0.074***	0.070***	0.072***
	(0.006)	(0.011)	(0.006)	(0.010)
6-10 employees	0.090***	0.062***	0.080***	0.057***
	(0.005)	(0.007)	(0.005)	(0.007)
11-20 employees	0.085***	0.042***	0.074***	0.039***
	(0.006)	(0.008)	(0.006)	(0.008)
21-50 employees	0.192***	0.106***	0.164***	0.099***
	(0.016)	(0.026)	(0.016)	(0.025)
more than 50 employees	0.345***	0.159***	0.304***	0.137***
	(0.023)	(0.044)	(0.022)	(0.034)
Industry FE (2 digit)	Yes	Yes		
Industry FE (3 digit)			Yes	Yes
Sampling weights	No	Yes	No	Yes
N	173743	173743	173743	173743
R^2	0.100	0.077	0.133	0.104

Table B-8: Corporate Customers and Firm Size. *Notes:* Columns 1 and 2 (3 and 4) control for two (three) digit industry fixed effects. Columns 2 and 4 weigh each observation by the sampling weights. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

B-6 Breakdown of the Construction & Utility sector into PS and CS

This section explains how we split the construction & Utility sector into Consumer Service, Producers Service, and Public sectors. We obtain the information of major destinations for sale of final product/service from the "Informal Non-Agricultural Enterprises Survey 1999-2000" (INAES) dataset. Based on this information, we calculate the approximated share of employment in PS, CS, and Public, respectively at 5-digit industries level. We then apply these shares to NSS datasets to obtain the total PS/CS/Public share in the Construction sectors. For simplicity, we apply the same shares to the Utility sectors.

The following are detailed steps of how we calculate the PS shares in the construction sector.

1: Calculating 5-digits industry level PS/CS/PUBLIC share using INAES 1999-2000

The survey asks the following question: what is the destination agency for sale of final product/service? The choices are: 1-government; 2-cooperative / marketing society; 3-private enterprise; 4-contractor / middleman; 5-private individual / household; 6-no source agency; 9- others. We drop all firms that answer 6 or 9, and attribute firms that answer 1 into Public sector; firms that answer 2, 3 or 4 into PS sector; firms that answer 5 into CS sector. For each 5 digits industry, we calculate the relative employment shares of PS/CS/Public sectors.

2. Adjusting CS/PS/PUBLIC shares

From the description of National Industry Classification, some sectors are clearly for public purposes. We classify 5-digit level industries into Public and Private and the results are reported in table B-9. We assume that the new public share is 1 for Public construction, and is 0 for Private construction. For Private construction sectors, we calculate the revised PS/CS shares based on the relative magnitude of the old PS/CS shares.

Table B-9: National Sample Survey of India: summary statistics

NIC-2004	Description	Public/Private
45101	Site preparation in connection with mining	Public
45102	Site preparation other than in connection with mining	Public
45201	General construction (including alteration, addition, repair and maintenance) of residential buildings.	Private
45202	General construction (including alteration, addition, repair and maintenance) of non-residential buildings.	Private
45203	Construction and maintenance of roads, rail-beds, bridges, tunnels, pipelines, rope-ways, ports, harbours and runways etc.	Public
45204	Construction/erection and maintenance of power, telecommunication and transmission lines	Public
45205	Construction and maintenance of waterways and water reservoirs	Public
45206	Construction and maintenance of hydro-electric projects	Public
45207	Construction and maintenance of power plants, other than hydro-electric power plants	Public
45208	Construction and maintenance of industrial plants other than power plants	Private
45209	Construction n.e.c. including special trade construction	Private
45301	Plumbing and drainage	Private
45302	Installation of heating and air-conditioning systems, antennas, elevators and escalators	Private
45303	Electrical installation work for constructions	Private
45309	"Other building installation n.e.c.	Private
45401	Setting of wall and floor tiles or covering with other materials like parquet, carpets, wall paper etc.	Private
45402	Glazing, plastering, painting and decorating, floor sanding and other similar finishing work	Private
45403	Finish carpentry such as fixing of doors, windows, panels etc. and other building finishing work n.e.c.	Private
45500	Renting of construction or demolition equipment with operator	Private

3. Calculating the total sectoral employment based on revised shares

Finally, we apply the 5-digit industry level PS/CS/PUBLIC share to the employment data from the NSS datasets[33]. After summing across all 5-digit industries, we obtain total PS employment, CS employment, Public employment, and their relative shares. Table B-10 reports the results. We take the average across the shares comes from the NSS survey (round55, round60, round64, round66) weighted by the total construction employment of each round.To be concluded, 12.2% of employment belongs to the PUBLIC. In the remaining employment, 13.1% belong to PS and 86.9% belong to CS. We apply the same rule to the utility sector.

[33](Shengqi: I will check if we can use NSS 43 and 68 with the new method) Conceptually, most residential construction firms belong to CS, and nonresidential construction firms belong to PS and Public, so we need to separate them. However, NSS 43rd round (uses NIC 1970) and NSS 68th round (uses NIC2008) do not distinguish between residential and nonresidential construction. So, we can only use NSS 55th, 60th (use NIC 1987), NSS 64th, and 66th round (use NIC 2004).

Table B-10: Sectoral employment share of the construction sector

	round 55-1999	round 60-2004	round 64-2007	round 66-2009
CS employment share	0.786	0.771	0.782	0.731
PS employment share	0.121	0.118	0.117	0.109
Public employment	0.093	0.111	0.101	0.160
PS/(PS+CS)	0.133	0.133	0.130	0.130
Total Construction employment	9921	5995	15356	17708

B-7 Relative Agricultural Price to Manufacturing

This section introduces how we calculate the relative prices of agricultural goods to manufacturing goods. Ministry of Planning and Program Implementation (MOSPI) of Government of India reports GDP by 2-digit sectors at current prices and constant prices from 1950-2013[34] We construct the price index for agricultural and manufacturing goods respectively by:

$$p_i = \frac{\text{GDP at current price}_i}{\text{GDP at constant price}_i}$$

We normalized both price indexes in the year 2005 to 1. Then we calculate the relative price by:

$$p_{relative} = \frac{p_{agri}}{p_{manu}}$$

To check the validity of our results, we also use two extra data sources to calculate the relative price. The first is the GGDC 10-Sector Database[35], which provides long-run data on sectoral productivity performance in Africa, Asia, and Latin America. Variables covered in the data set are annual series of value-added at current national prices, value-added at constant 2005 national prices, and persons employed for 10 broad sectors. We follow the same procedures to calculate the relative price.

The second is the Wholesale Price Index (WPI) reported by Office of the Economic Advisor[36]. The WPI tracks ex-factory price for manufactured products, agri-market (mandi) price for agricultural commodities. One issue with this is that the base year (and the basket of goods) changes during different time periods. Two series are relevant to our research. The first one is the series with the base year 1993, which is available from 1994 to 2009. The second one is the series with the base year 2004, which is available from 2005 to 2016. Again, we use the same method to calculate the relative price, and normalized the relative prize in the year 2005 to 1.

In Figure B-4 we plot the relative price of agricultural goods to manufacturing goods. Since the pattern from different sources are very similar, we use the results based on MOPSI in the analysis.

B-8 Outliers in the Welfare Analysis

In the counterfactual welfare analysis of Section 5.2.1 the results are somewhat sensitive to outliers. The reason is that a limited number of region have very large swings in the productivity of CS. For this reason, we trim the 2% observation showing the largest changes in both direction. More formally, we calculate $\Delta = \log(A_{CS,1987}) - \log(A_{CS,2011})$, and trim observation below the 2nd percentile and above the 98th percentile of the resulting distribution before calculating the average welfare effects in Figure 4. The trimmed districts are on average small representing altogether 1.5% of the total population in our analysis. Table B-11 below reports the average welfare effects of setting back A_{CS} at different trimming cut-offs.

[34]The data is available at `http://www.mospi.gov.in/data` 1. Summary of macro economic aggregates at current prices, 1950-51 to 2013-14
2. Summary of macro economic aggregates at constant(2004-05) prices, 1950-51 to 2013-14

[35]The data is available at `https://www.rug.nl/ggdc/productivity/10-sector`

[36]The data is available at `https://eaindustry.nic.in/`

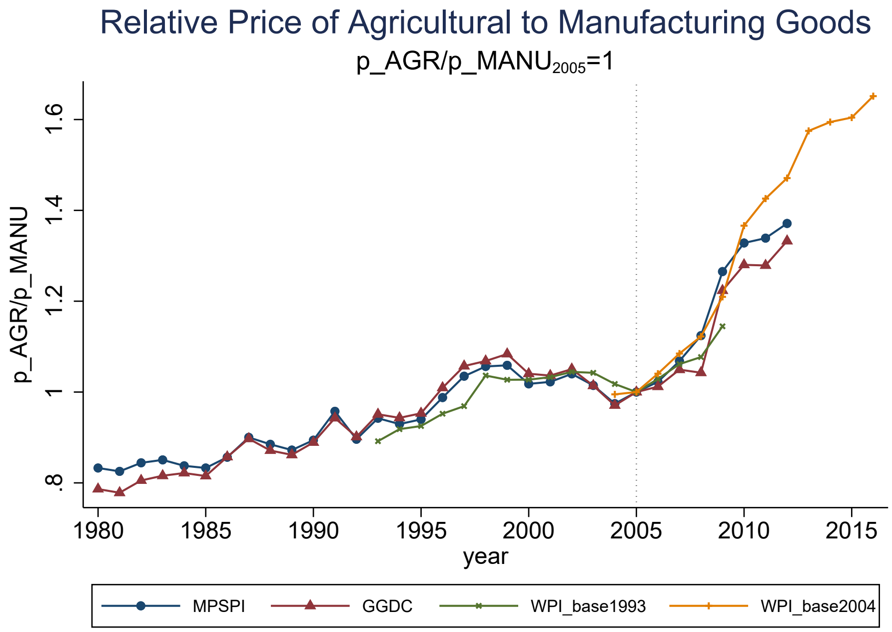

Figure B-4: Relative price of agricultural to manufacturing goods

Trimming Cut-offs	No Trim	1%	2%	3%	4%	5%
Welfare Loss	0.18	0.19	0.25	0.26	0.26	0.27

Table B-11: Counterfactual Welfare Losses with Different Trimming Cut-offs.

B-9 Estimation of the shape of the distribution ζ

We identify the tail of efficient unit distribution ζ from the local expenditure distribution. Specifically, our model assume that the income and expenditure for individual h is given by $e^h_{rt} = q^h w_{rt}$ where q follows a Pareto distribution

$$G_{rt}(q) = 1 - \left(\frac{\underline{q}_{rt}}{q}\right)^{\zeta_{rt}}$$

with $E[q] = 1$.
To estimate ζ_{rt}, note that

$$\ln(1 - G_{rt}(q)) = \zeta_{rt}\ln(\underline{q}_{rt}) - \zeta_{rt}\ln(q)$$

Hence, we estimate ζ_{rt} from a regression of log of the upper tail probability on log efficient unit which calculated as

$$q^h_{rt} = \frac{e^h_{rt}}{w_{rt}}$$

The following table reports the estimated ζ_{rt} by grouping districts according to urbanization quintile and year

Table B-12: Identification of ζ_{rt}

	Pareto Tail					
	All	1st Quintile	2 Quintile	3 Quintile	4 Quintile	5 Quintile
1987	1.45	1.62	1.51	1.42	1.53	1.41
2011	2.24	2.42	1.98	1.91	2.04	1.68

B-10 Estimation of the shape of the productivity distribution λ

We identify the shape parameter λ of the productivity distribution from the tail of the employment distribution. Our model implies that total employment of a firm with productivity z is given by

$$l(z) = L_{PS}(z) + L_{PM}(z) = \frac{\beta+\alpha}{\beta}\varsigma(A_{PS})\frac{z}{z_L}f_0 - \varsigma(A_{PS})f_0.$$

This implies that the employment distribution is given by

$$F_l(l) = P(l(z) \leq l) = 1 - \left(\frac{A_{rMt}\frac{\alpha+\beta}{\beta}\frac{\zeta(A_{PS})f_0}{z_L}}{l + \zeta(A_{PS})f_0}\right)^{\lambda}.$$

Hence, for large firms (i.e. $l \to \infty$), the tail of the employment distribution is exactly equal to λ.

To estimate λ, note that

$$\ln(1 - F_l(l)) = C_0 - \lambda \ln l, \tag{B-1}$$

where $C_0 = \ln\left(A_{rMt}\frac{\alpha+\beta}{\beta}\frac{\zeta f_0}{z_L}\right)^{\lambda}$. Hence, we estimate λ from a regression of the log of the upper tail probability on log employment.

In Figure B-5 we depict the empirical relationship of (B-1) for the different sub samples of tail of the employment distribution. Specifically, we consider a grid of 200 points of the log employment distribution $\ln l$, calculate $F_l(l)$ for these grid points and then plot $\ln(1 - F_l(l))$ against $\ln l$. We consider four samples, namely all firms with employment exceeding the 50%, 70%, 80% and 90% quantiles.

Equation (B-1) implies that the relationship should be linear and that the slope should be equal $-\lambda$. Figure B-5 shows that the employment distribution in India indeed has a pareto tail and that the estimated slope does not depend markedly on the employment cutoff for the employment distribution.

In Table B-13 we report these results in a regression format. The four columns refer to the different subsamples of the tail of the employment distribution. We estimate a pareto tail of 1.42. Reassuringly, the slope is very precisely estimated and the estimates are almost identical across the different samples. We also find the economic magnitude plausible. The firm size distribution in the US is often found to have a tail of around 1.1. Given that the importance of large firms is larger in rich countries, an estimate of 1.42 strikes us as plausible.

Table B-13: Identification of λ

	(1)	(2)	(3)	(4)
	$> 50\%$	$> 70\%$	$> 80\%$	$> 90\%$
log employment	-1.426***	-1.420***	-1.418***	-1.414***
	(0.010)	(0.011	(0.011)	(0.012)
Industry FE	Yes	Yes	Yes	Yes
R^2	0.991	0.989	0.989	0.987

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

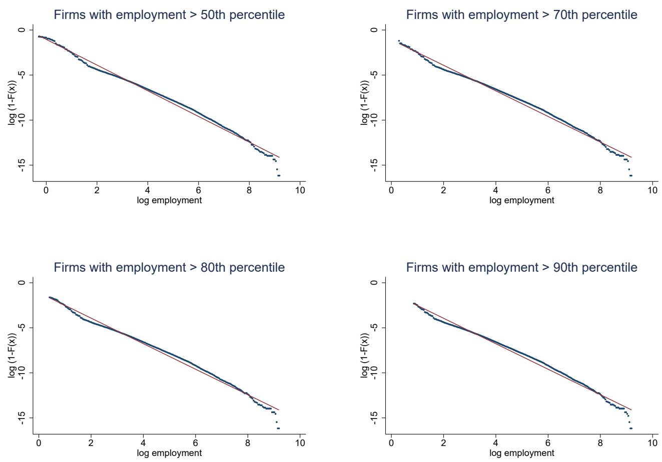

Figure B-5: IDENTIFICATION OF λ. The figures depicts the relationship between $\ln(1 - F_l(x))$ and $\ln x$ for different cutoffs of the employment distribution. We consider all firms with employment exceeding the 50%, 70%, 80% and 90% quantile. For each sample we consider a grid of 200 points the log employment distribution. The slope of the relationship coincides with λ (see (B-1)).

B-11 Details of Robustness Analysis

B-11.1 Generalized PIGL preferences

In Section 6.3 we extended the preference specification in (2) to the more general specification of $D(\mathbf{p})$:

$$D(\mathbf{p}) = \frac{1}{\gamma}\left[\left(\prod_j p_j^{\tilde{\nu}_j}\right)^{\gamma} - 1\right]. \tag{B-2}$$

Our benchmark specification is nested as the special case of $\gamma \to 0$.

We run experiments in which we reduce each of the productivities A_F, A_G, and A_{CS} by 20% in all districts in year 2011 for different values of γ. Figure B-6 summarizes the results.

We find that the welfare effect of productivity growth in CS is increasing in γ. However, the quantitative magnitude is modest and even for values of γ between -2 and 2, we find a significant role for consumer service productivity growth on aggregate welfare. The same qualitative pattern emerges for the agricultural sector. In contrast, the effects for the industrial sector are insensitive to γ. Note that in some ranges of low and high γ (highlighted as dashed lines in Figure B-6) the employment share in the consumer service sector of some districts falls to zero. In these cases, we calculate approximate equilibria as opposed to exact ones. The accuracy of the approximation is very high (details available upon request).

B-11.2 The Model with Imperfect Skill Substitution

Environment

Suppose that technology in sector s in region r is given by

$$Y_{rs} = A_{rs}\left(\left(H_{rs}^{-}\right)^{\frac{\rho-1}{\rho}} + \left(Z_{rs}H_{rs}^{+}\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}},$$

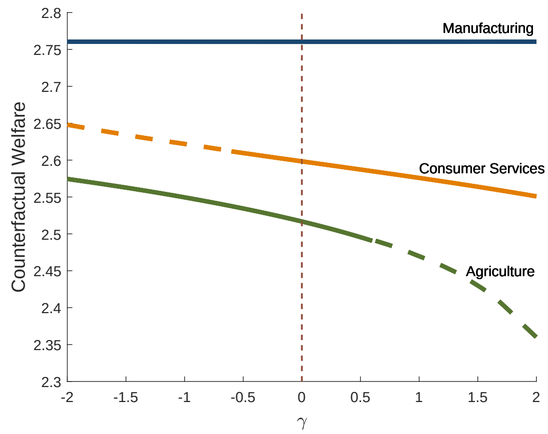

Figure B-6: Generalized PIGL Preferences The figure shows the welfare effect (equivalent variation) of counterfactually reducing by 20% the productivity in agriculture, manufacturing in year 2011 for different values of the income elasticity γ in Equation (16). In the benchmark economy, $\gamma = 0$. A higher counterfactual welfare level means that productivity growth has a smaller welfare effect.

where A_{rs} denotes factor neutral productivity, Z_{rs} denotes the skill bias and H_{rs}^{-} (H_{rs}^{+}) are the quantities of human capital of low (high) skilled individuals. Again we assume that individuals are heterogenous. Specifically, people of skill type $j \in \{-,+\}$ draw their efficiency level from a pareto with the same shape, i.e.

$$P\left(q_i^j \leq k\right) = 1 - \left(\frac{\underline{q}_{rt}^j}{k}\right)^{\zeta} \equiv F_{rt}^j(k).$$

Total income of an individual i of skill type j in region r at time t is therefore given by $y_{rt}^{i,j} = w_{rt}^j q_i^j$, where the skill price w_{rt}^j is now skill-specific. The aggregate expenditure share on goods from sector s goods in region r is then given by

$$\vartheta_{rst} \equiv \frac{L_{rt}^{-}\int \vartheta_s^h\left(qw_{rt}^{-}, p_{rt}\right) qw_{rt}^{-} dF_{rt}^{-}(q) + L_{rt}^{+}\int \vartheta_s^h\left(qw_{rt}^{+}, p_{rt}\right) qw_{rt}^{+} dF_{rt}^{+}(q)}{L_{rt}^{-}\int qw_{rt}^{-} dF_{rt}^{-}(q) + L_{rt}^{+}\int qw_{rt}^{+} dF_{rt}^{+}(q)},$$

where $\vartheta_s^h\left(qw_{rt}^{-}, p_{rt}\right)$ denotes the sectoral expenditure share at the individual level. Substituting the expression for $\vartheta_s^h\left(qw_{rt}^{-}, p_{rt}\right)$ and using the fact that $y_{rt}^{i,j}$ is also pareto distributed yields

$$\vartheta_{rst} = \omega_s + \tilde{\nu}_s \frac{\zeta - 1}{\zeta - (1-\varepsilon)} \left(\frac{1}{\prod_s p_s^{\omega_s}}\right)^{-\varepsilon} \left(s_{rt}^{Y,-}\left(w_{rt}^{-}\underline{q}_{rt}^{-}\right)^{-\varepsilon} + \left(1 - s_{rt}^{Y,+}\right)\left(w_{rt}^{+}\underline{q}_{rt}^{+}\right)^{-\varepsilon}\right).$$

where $s_{rt}^{Y,-} = \frac{L_{rt}^{-} w_{rt}^{-} \underline{q}_{rt}^{-}}{L_{rt}^{-} w_{rt}^{-} \underline{q}_{rt}^{-} + L_{rt}^{+} w_{rt}^{+} \underline{q}_{rt}^{+}}$ is the income share of low skilled individuals in region r at time t. Hence, the sectoral expenditure share is given by

$$\vartheta_{rst} = \vartheta_s\left(\underline{q}_{rt}^{-} w_{rt}^{-}, \underline{q}_{rt}^{+} w_{rt}^{+}, s_{rt}^{Y,-}, \mathbf{p_{rt}}\right),$$

i.e. sectoral spending varies at the regional level because of (i) differences in regional factor prices w_{rt}^{-} and w_{rt}^{+}, (ii) differences in the prices of non-tradable goods p_{rCSt} and (iii) differences in the skill composition $s_{rt}^{Y,-}$.

Equilibrium conditions

The equilibrium is characterized by the following conditions. The CES structure and perfect competition imply that prices are given by

$$p_{rs} = \frac{1}{A_{rs}} \left(\left(w_{rt}^{-}\right)^{1-\rho} + Z_{rs}^{\rho-1} \left(w_{rt}^{+}\right)^{1-\rho} \right)^{\frac{1}{1-\rho}}.$$

The relative skill demand for sector s in region $r$is given by

$$\frac{w_{rt}^{+} H_{rst}^{+}}{w_{rt}^{-} H_{rst}^{-}} = Z_{rt}^{\rho-1} \left(\frac{w_{rt}^{+}}{w_{rt}^{-}} \right)^{1-\rho}.$$

The CES demand system across regional varieties implies the market clearing conditions

$$w_{rt}^{-} H_{rst}^{-} + w_{rt}^{+} H_{rst}^{+} \quad = \quad \left(\frac{p_{rs}}{p_s} \right)^{1-\sigma} \times \sum_{j=1}^{R} \vartheta_s \left(\underline{q}_{jt}^{-} w_{jt}^{-}, \underline{q}_{jt}^{+} w_{jt}^{+}, s_{jt}^{Y,-}, \mathbf{p_{jt}} \right) \overline{w}_{rt} L_{rt},$$

where $\overline{w}_{rt}$ denotes average income and $p_s = \left(\sum_{r=1}^{R} p_{rs}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$. Market clearing condition for non-tradable CS implies

$$w_{rt}^{-} H_{rCSt}^{-} + w_{rt}^{+} H_{rCSt}^{+} = \vartheta_{CS} \left(\underline{q}_{jt}^{-} w_{jt}^{-}, \underline{q}_{jt}^{+} w_{jt}^{+}, s_{jt}^{Y,-}, \mathbf{p_{jt}} \right) \overline{w}_{rt} L_{rt}. \tag{B-3}$$

Finally, labor market clearing implies

$$H_{rF}^{j} + H_{rG}^{j} + H_{rCS}^{j} = H_{r}^{j} \text{ for } j \in \{-, +\}.$$

These equations uniquely determine the regional wages $\left\{ w_{rt}^{-}, w_{rt}^{+} \right\}$ and the sectoral labor allocations $\left\{ H_{rst}^{-}, H_{rst}^{+} \right\}$.

Measurement and Equilibrium Accounting

As before we use these equations and the observable data to infer the productivity vector $\{A_{rst}, Z_{rst}\}$ for each region-sector pair. To connect our data to the objects in the model, we make the following choices:

1. We classify individuals into high and low skill workers by their years of schooling. We assume workers with more then secondary schooling are high skill workers.

2. As in our baseline model, we assume a Mincerian return $\rho = 5.6\%$ per year of schooling within skill groups. This allows us to measure the aggregate skill supplies H_{rt}^{-} and H_{rt}^{+} for each region.

3. As in our baseline model, we use the observed sectoral earnings shares by skill group to measure sectoral labor supplies. Specifically, for each skill group $j = \{-, +\}$ and sector s, we calculate

   $$H_{rst}^{j} = \frac{\sum_i 1\left[i \in j \text{ and } i \in s\right] w_i}{\sum_i 1\left[i \in j\right] w_i} \times H_{rt}^{j}$$

   where w_i is the wage of individual i.

4. We then calculate the regional skill prices as $w_r^j = \frac{1}{L_{rt}^j} \sum_{i=1}^{L_{rt}^j} y_{rti}^j$ where y_{rti}^j denotes total income of individual i in region $r$at time t in skill group j.

These data is sufficient to uniquely solve for $\{A_{rst}, Z_{rst}\}$ and to perform the counterfactual analysis reported in Section 6.4.

APPENDIX C: FIGURES AND TABLES

In this section, we report additional tables and figures referred to in the text.

C-1 Additional empirical results

Urbanization and Aggregate Growth

In Figure C-1 we report the time-series change in the urbanization rate (panel a) and in income per capita (panel b). The urbanization rate is the share of population living in urban areas according to the definition of the NSS, that defines an urban location in the following way: (i) all locations with a Municipality, Corporation or Cantonment and locations notified as town area, (ii) all other locations that satisfy the following criteria: (a) a minimum population of 5000, (b) at least 75 percent of the male population are employed outside of agriculture, and (c) a density of population of at least 1000 per square mile. This share increased from around 22% in 1987 to 29% in 2010. Income per capita stems from the World Bank. Between 1987 and 2010, income per capita increased by a factor of almost 3.

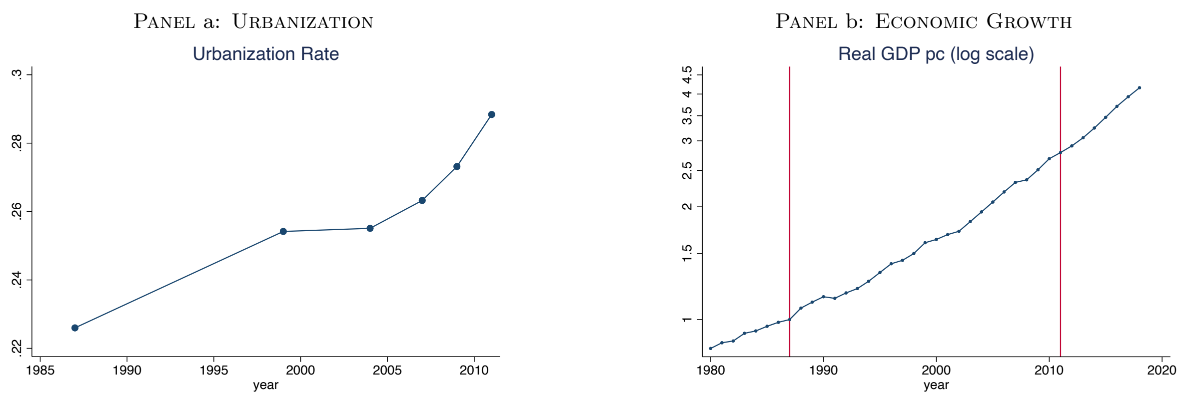

Figure C-1: Structural Change in India: 1987 - 2011. This figure shows the evolution of the urbanization rate (Panel a) and income per capita (Panel b). The urbanization rate is the share of population living in urban areas according to the definition of the NSS. Income per capita stems from World Bank.

Urbanization and Income per Capita

For some of our analysis we choose urbanization as our measure of spatial heterogeneity. We do so as a descriptive device and interpret urbanization as a broad proxy for regional economic development. Figure C-2 shows that there is a strong positive correlation between urbanization and expenditure per capita in the NSS data in 2011.

Spatial Structural Change: Sectoral Income

In Figure C-3 we report sectoral employment shares by urbanization quintiles in 1987 (Panel a) and in 2011 (Panel b).

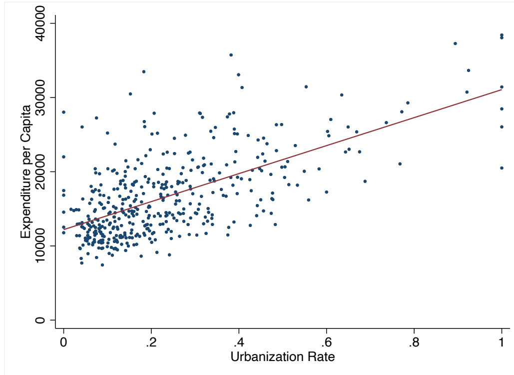

Figure C-2: Expenditure per capita vs. Urbanization. The figure shows a binscatter plot of the average expenditure capita in the NSS data across district-level urbanization rates in 2011.

Panel a: Sectoral Income by Urbanization (1987)

Panel b: Sectoral Income by Urbanization (2011)

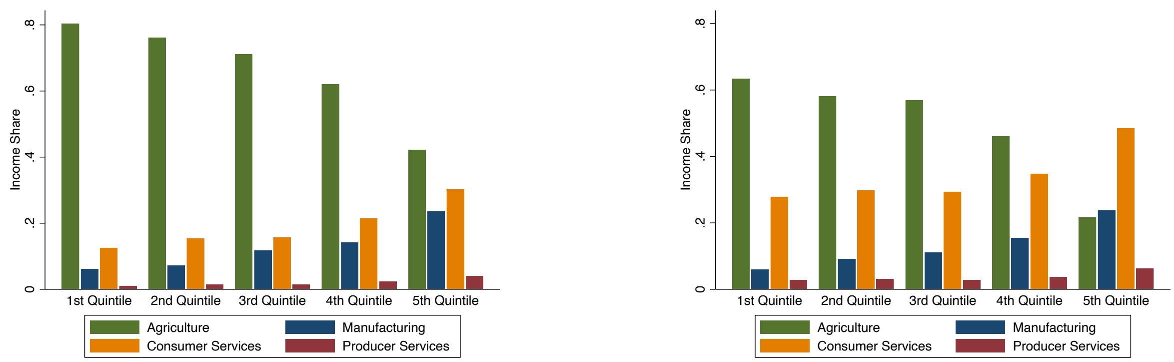

Figure C-3: Spatial Structural Change in India. The figure plots plots the sectoral income shares by urbanization quintile in 1987 and 2011.

Engel Curves in India

In Table 2 in the main text we reported the estimated elasticity of agricultural expenditure shares with respect to expenditure. In Figure C-4 we show the estimated Engel curve for the year 2011 graphically. As implied by the PIGL preference specification, the relationship is approximately linear for a large part of the expenditure distribution.

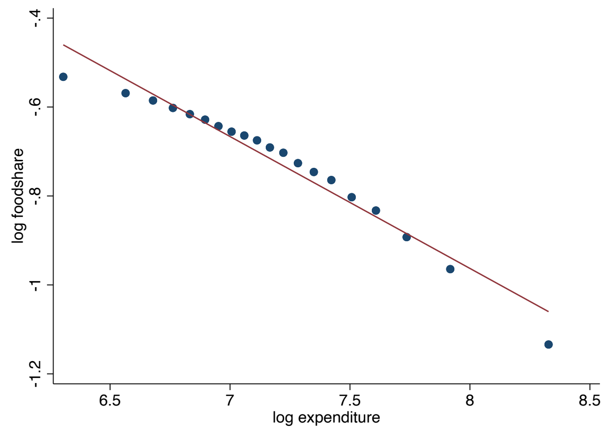

Figure C-4: Engel curves in India. The figure shows a binscatter plot of the log food shares and log expenditure for the year 2011 at the individual household level after absorbing district fixed effects.

Structural Change: Distribution of Productivity Growth in Consumer Services.

Figure C-5 reports the distribution of growth rates of CS productivity between 1987 and 2011 across Indian districts by urbanization.

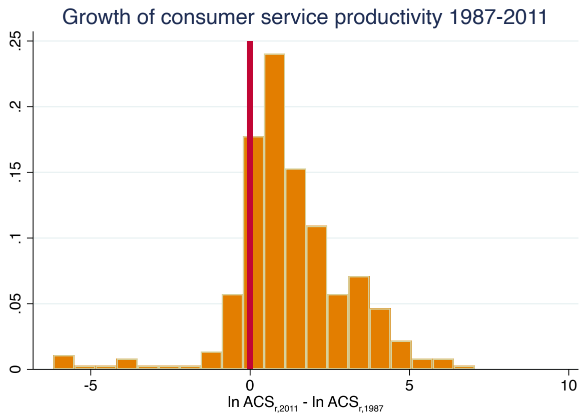

Figure C-5: The Distribution of Productivity Growth in Consumer Services. The figure shows the distribution of estimated consumer service productivity growth ($\ln A_{rCS2011} - \ln A_{rCS1987}$). COULD GO TO APPENDIX

Structural Change: Heterogenous Effects

Figure C-6 disaggregates across districts with different urbanization levels the effects described in Figure 6. We restric attention to partial equilibrium experiments and to a comparison between the joint effect of productivity growth in the traded sectors (panel a) with that in the CS sector (panel a).

Structural Change with Imperfect Substitution in Skills

Figure C-7 is the analogue of Figure 6 in the text for the extension of the model in Section 6.4.

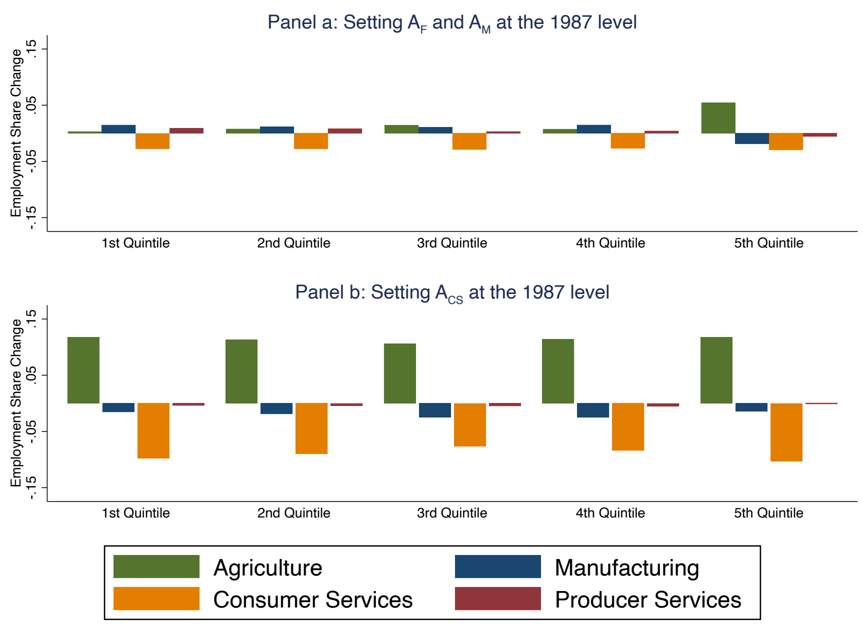

Figure C-6: Sectoral Productivity Growth and Structural Change. Panel a shows the changes in sectoral employment share for five representative districts (by urbanization quintiles) when we counterfactually set both A_F and A_M at the 1987 levels. Panel b shows the same information when we counterfactually set A_{CS} at the 1987 level

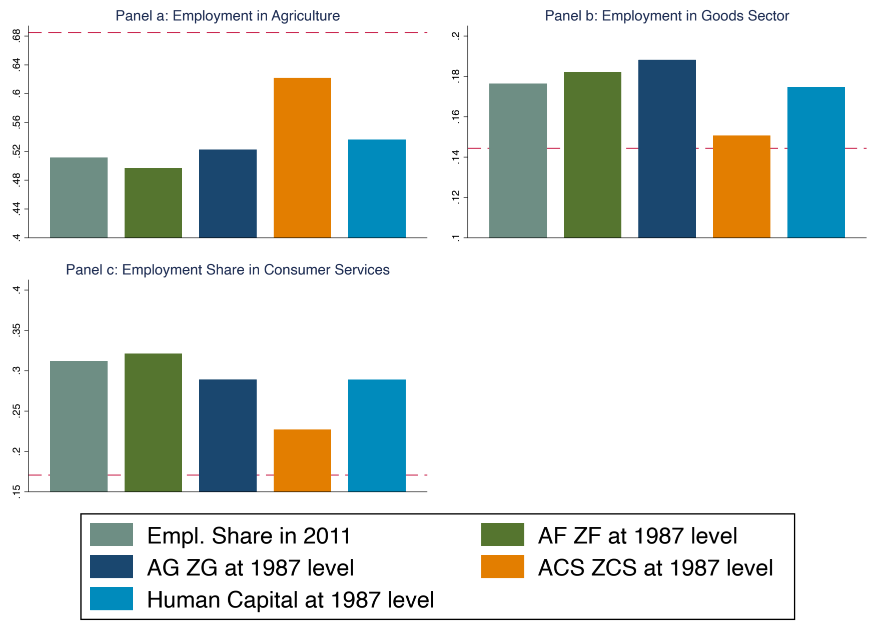

Figure C-7: Sectoral Productivity Growth and Structural Change with Imperfect Substitution in Skills. This figure is the analogue of Figure 6 in the extension allowing for imperfect substitution across skilled groups and for skill-biased technical change.

Online Appendix for
"Service-Led or Service-Biased Growth? Equilibrium Development Accounting Across Indian Districts."

by Tianyu Fan, Michael Peters, and Fabrizio Zilibotti

December 31, 2020

- Not for Publication Unless Requested -

OA-1 Appendix OA: Derivations for the results in Section A-2

OA-1.1 Proof of Proposition 2

Consider the maximization problem in (A-4), i.e.

$$\pi\left(z_i\right)=\max_{H_{PMi},H_{PSi}\geq 0}\left\{p_G z_i^{1-\alpha-\beta}H_{PMi}^{\alpha}\left(A_{PS}H_{PSi}+\kappa\right)^{\beta}-w\left(H_{PMi}+H_{PSi}\right)-f_O w\right\}.$$

The optimality condition for production workers H_{PMi} is given by

$$\alpha z_i^{1-\alpha-\beta}H_{PMi}^{\alpha-1}\left(A_{PS}H_{PSi}+\kappa\right)^{\beta} \quad = \quad \frac{w}{p_G}.$$

This implies that

$$H_{PMi}=\frac{\alpha}{\beta A_{PS}}\left(A_{PS}H_{PSi}+\kappa\right). \tag{OA-1}$$

Substituting H_{PMi} in the objective function yields

$$\max_{H_{PSi}\geq 0} p_G z_i^{(1-\alpha-\beta)}\left(\frac{\alpha}{\beta A_{PS}}\left(A_{PS}H_{PSi}+\kappa\right)\right)^{\alpha}\left(A_{PS}H_{PSi}+\kappa\right)^{\beta}-w\left(\frac{\alpha}{\beta A_{PS}}\left(A_{PS}H_{PSi}+\kappa\right)+H_{PSi}\right)$$

The FOC yields:

$$\left(\alpha+\beta\right)A_{PS}\left(z_i\right)^{(1-\alpha-\beta)}\left(\frac{\alpha}{\beta A_{PS}}\right)^{\alpha}\left(A_{PS}H_{PSi}+\kappa\right)^{\alpha+\beta-1}=\frac{\alpha+\beta}{\beta}\frac{w}{p_G}.$$

Rearranging terms and using the definition of z_L yields

$$H_{PSi}=\max\left\{\frac{\kappa}{A_{PS}}\frac{z_i}{z_L}-\frac{\kappa}{A_{PS}},0\right\}. \tag{OA-2}$$

For all firms with $z_i\geq z_L$, (OA-1) implies that

$$H_{PMi}=\frac{\alpha}{\beta}\left(H_{PSi}+\frac{\kappa}{A_{PS}}\right)=\frac{\alpha}{\beta}\frac{\kappa}{A_{PS}}\frac{z_i}{z_L}=\frac{\alpha}{\beta}\frac{z_i}{z_L}\varsigma f_O. \tag{OA-3}$$

The resulting profits for a firm with $z_i \geq z_L$ are given by

$$\begin{aligned}
\pi(z_i) &= p_G z_i^{1-\alpha-\beta} H_{PMi}^{\alpha} (A_{PS} H_{PSi} + \kappa)^{\beta} - w(H_{PMi} + H_{PSi}) - f_O w \\
&= \frac{1}{\alpha} w H_{PMi} - w\left(H_{PMi} + \frac{\beta}{\alpha} H_{PMi} - \frac{\kappa}{A_{PS}}\right) - f_O w \\
&= \left[\frac{1-\beta-\alpha}{\alpha} H_{PMi} + \varsigma f_O - f_O\right] w \\
&= \left[\left(1 + \frac{1-\beta-\alpha}{\beta} \frac{z_i}{z_L}\right)\varsigma - 1\right] f_O w, \qquad \text{(OA-4)}
\end{aligned}$$

which is the result in (A-9). If $z_i \leq z_L$, we have $H_{PSi} = 0$ and the firm solves

$$\begin{aligned}
\pi(z_i) &= \max_{H_{PMi}} \left\{ p_G z_i^{1-\alpha-\beta} \kappa^{\beta} (H_{PMi})^{\alpha} - w H_{PMi} - f_O w \right\} \\
&= \left(\frac{1-\alpha}{\alpha}\right) w \left(\frac{p_G z_i^{1-\alpha-\beta} \kappa^{\beta} \alpha}{w}\right)^{\frac{1}{1-\alpha}} - f_O w.
\end{aligned}$$

Hence, the productivity cutoff z^*, which is defined by $\pi(z^*) = 0$, is given by

$$\left(\frac{1-\alpha}{\alpha}\right) w \left(\frac{p_G (z^*)^{1-\alpha-\beta} \kappa^{\beta} \alpha}{w}\right)^{\frac{1}{1-\alpha}} = w f_O$$

so that

$$z^* = \left(\frac{w}{p_G} \frac{1}{\kappa^{\beta} \alpha} \left(\frac{\alpha}{1-\alpha} f_O\right)^{1-\alpha}\right)^{\frac{1}{1-\alpha-\beta}}.$$

Note that this implies that

$$\pi(z_i) = \left(\left(\frac{z_i}{z^*}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} - 1\right) f_O w. \qquad \text{(OA-5)}$$

The hiring cutoff for lawyers z_L is therefore defined by (see (OA-4) and (OA-5))

$$\left[\left(1 + \frac{1-\beta-\alpha}{\beta} \frac{z_L}{z_L}\right)\varsigma - 1\right] f_O w = \left(\left(\frac{z_L}{z^*}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} - 1\right) f_O w,$$

which implies that

$$z_L = \left(\left(\frac{1-\alpha}{\beta}\right)\varsigma\right)^{\frac{1-\alpha}{1-\alpha-\beta}} z^*.$$

The optimal number of production workers H_{PMi} is therefore given by

$$H_{PMi} = \left(\alpha \kappa^{\beta} \frac{p_G}{w}\right)^{\frac{1}{1-\alpha}} z_i^{\frac{1-\alpha-\beta}{1-\alpha}} = \frac{\alpha}{1-\alpha} f_O \left(\frac{z_i}{z^*}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} = \frac{\alpha}{\beta} f_O \varsigma \left(\frac{z_i}{z_L}\right)^{\frac{1-\alpha-\beta}{1-\alpha}}.$$

OA-1.2 Proof of Proposition 3

Consider the complementary case in which we assume $\varsigma < \frac{\beta}{1-\alpha}$, so that all firms produce with lawyers. Firm profits are given by

$$\begin{aligned}
\pi(z_i) &= \max_{H_{PMi},H_{PSi}} \left\{ p_G z_i^{1-\alpha-\beta} H_{PMi}^{\alpha} \left(A_{PS}H_{PSi}+\kappa\right)^{\beta} - w\left(H_{PMi}+H_{PSi}\right) - f_O w \right\} \\
&= \max_{H_{PMi},H_{PSi}} \left\{ p_G z_i^{1-\alpha-\beta} H_{PMi}^{\alpha} \left(A_{PS}H_{PSi}+\kappa\right)^{\beta} - wH_{PMi} - \frac{w}{A_{PS}}\left(A_{PS}H_{PSi}+\kappa\right) + \frac{w\kappa}{A_{PS}} - f_O w \right\} \\
&= \max_{H_{PMi},X_{PSi}} \left\{ p_G z_i^{1-\alpha-\beta} A_{PS}^{\beta} H_{PMi}^{\alpha} X_{PSi}^{\beta} - wH_{PMi} - wX_{PSi} \right\} + w f_O\left(\varsigma-1\right) \\
&= \left\{ \left(1-\alpha-\beta\right)\left(\frac{p_G}{w}\right)^{\frac{1}{1-\alpha-\beta}} \alpha^{\frac{\alpha}{1-\alpha-\beta}} \left(\beta A_{PS}\right)^{\frac{\beta}{1-\alpha-\beta}} z_i + \left(\varsigma-1\right) f_O \right\} w
\end{aligned}$$

The productivity cutoff $\tilde{z}$ is defined by $\pi(\tilde{z}) = 0$ and therefore given by[37]

$$\tilde{z} = \frac{1}{1-\alpha-\beta}\left(\frac{w}{p_G}\right)^{\frac{1}{1-\alpha-\beta}}\left(\frac{1}{\alpha}\right)^{\frac{\alpha}{1-\alpha-\beta}}\left(\frac{1}{\beta A_{PS}}\right)^{\frac{\beta}{1-\alpha-\beta}} f_O\left(1-\varsigma\right). \tag{OA-6}$$

Hence, we write profits as

$$\pi(z_i) = \left(\frac{z-\tilde{z}}{\tilde{z}}\right) f_O\left(1-\varsigma\right) w.$$

Optimal factor demands satisfy

$$\alpha p_G z_i^{1-\alpha-\beta} H_{PMi}^{\alpha-1}\left(A_{PS}H_{PSi}+\kappa\right)^{\beta} = w$$

and

$$\beta p_G z_i^{1-\alpha-\beta} H_{PMi}^{\alpha}\left(A_{PS}H_{PSi}+\kappa\right)^{\beta-1} A_{PS} = w$$

Hence,

$$\frac{\alpha}{\beta}\frac{\left(A_{PS}H_{PSi}+\kappa\right)}{H_{PMi}A_{PS}} = 1$$

This implies that

$$\begin{aligned}
H_{PMi} &= \alpha^{\frac{1-\beta}{1-\alpha-\beta}}\beta^{\frac{\beta}{1-\alpha-\beta}}\left(\frac{p_G}{w}A_{PS}^{\beta}\right)^{\frac{1}{1-\alpha-\beta}} z_i \\
H_{PSi} &= \alpha^{\frac{\alpha}{1-\alpha-\beta}}\beta^{\frac{1-\alpha}{1-\alpha-\beta}}\left(\frac{p_G}{w}A_{PS}^{\beta}\right)^{\frac{1}{1-\alpha-\beta}} z_i - \frac{\kappa}{A_{PS}}
\end{aligned}$$

Using the definition of $\tilde{z}$ we get that

$$\left(\frac{p_G}{w}A_{PS}^{\beta}\right)^{\frac{1}{1-\alpha-\beta}}\alpha^{\frac{1-\beta}{1-\alpha-\beta}}\beta^{\frac{\beta}{1-\alpha-\beta}} = \frac{1}{\tilde{z}}\frac{\alpha}{1-\alpha-\beta}f_O\left(1-\varsigma\right)$$

Hence,

$$H_{PMi} = \frac{1}{\tilde{z}}\frac{\alpha}{1-\alpha-\beta}f_O\left(1-\varsigma\right)z_i$$

and

$$H_{PSi} = \frac{1}{\tilde{z}}\frac{\beta}{1-\alpha-\beta}f_O\left(1-\varsigma\right)z_i - \frac{\kappa}{A_{PS}}$$

[37]One can show that for $\frac{\kappa}{A_{PS}f_O} = \frac{\beta}{1-\alpha}$ the cutoff z^* coincides with z_L in the other case. Note also that we need to assume that $\varsigma < 1$ to ensure that this equation has a solution.

OA-1.3 Proof of Proposition 4

Consider first the case of $\varsigma \geq \frac{\beta}{1-\alpha}$. Then

$$\begin{aligned}
E[\pi] &= \int_{z^*}^{z_L} \pi^{NL}(z) f(z)\,dz + \int_{z_L}^{\infty} \pi^{L}(z) f(z)\,dz \\
&= f_O w \int_{z^*}^{z_L} \left(\frac{z}{z^*}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} f(z)\,dz + f_O w \varsigma \int_{z_L}^{\infty} \left(1 + \frac{1-\beta-\alpha}{\beta}\frac{z_i}{z_L}\right) f(z)\,dz - f_O w P(z \geq z^*), \\
&= f_O w \left(\frac{1}{z^*}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} \int_{z^*}^{z_L} z^{\frac{1-\alpha-\beta}{1-\alpha}} f(z)\,dz + f_O w \varsigma \left(\int_{z_L}^{\infty} f(z)\,dz + \frac{1-\beta-\alpha}{\beta}\frac{1}{z_L}\int_{z_L}^{\infty} z f(z)\,dz\right) - f_O w P(z \geq z^*)
\end{aligned}$$

where $f(z) = \frac{\lambda A_M^{\lambda}}{z^{\lambda+1}}$. Note that

$$\int_{z_L}^{\infty} f(z)\,dz = P(z \geq z_L) = \left(\frac{A_M}{z_L}\right)^{\lambda}$$

and

$$\int_{z_L}^{\infty} z f(z)\,dz = P(z \geq z_L)\, E[z | z \geq z_L] = \left(\frac{A_M}{z_L}\right)^{\lambda} \frac{\lambda}{\lambda-1} z_L$$

and[38]

$$\begin{aligned}
\int_{z^*}^{z_L} z^{\frac{1-\alpha-\beta}{1-\alpha}} f(z)\,dz &= \int_{z^*}^{\infty} z^{\frac{1-\alpha-\beta}{1-\alpha}} f(z)\,dz - \int_{z_L}^{\infty} z^{\frac{1-\alpha-\beta}{1-\alpha}} f(z)\,dz \\
&= \left(\frac{A_M}{z^*}\right)^{\lambda} \frac{\frac{\lambda(1-\alpha)}{1-\alpha-\beta}}{\frac{\lambda(1-\alpha)}{1-\alpha-\beta} - 1} (z^*)^{\frac{1-\alpha-\beta}{1-\alpha}} - \left(\frac{A_M}{z_L}\right)^{\lambda} \frac{\frac{\lambda(1-\alpha)}{1-\alpha-\beta}}{\frac{\lambda(1-\alpha)}{1-\alpha-\beta} - 1} (z_L)^{\frac{1-\alpha-\beta}{1-\alpha}} \\
&= \frac{\lambda(1-\alpha)}{(\lambda-1)(1-\alpha)+\beta} \left(\frac{A_M}{z_L}\right)^{\lambda} (z_L)^{\frac{1-\alpha-\beta}{1-\alpha}} \left(\left(\left(\frac{1-\alpha}{\beta}\right)\varsigma\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}} - 1\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
& f_O w \left(\tfrac{1}{z^*}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} \int_{z^*}^{z_L} z^{\frac{1-\alpha-\beta}{1-\alpha}} f(z)\,dz \\
= \quad & f_O w \left(\tfrac{1}{z_L}\right)^{\frac{1-\alpha-\beta}{1-\alpha}} \left(\tfrac{1-\alpha}{\beta}\right)\varsigma \tfrac{\lambda(1-\alpha)}{(\lambda-1)(1-\alpha)+\beta} \left(\tfrac{A_M}{z_L}\right)^{\lambda} (z_L)^{\frac{1-\alpha-\beta}{1-\alpha}} \left(\left(\left(\tfrac{1-\alpha}{\beta}\right)\varsigma\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}} - 1\right) \\
= \quad & f_O w \left(\tfrac{1-\alpha}{\beta}\right)\varsigma \tfrac{\lambda(1-\alpha)}{(\lambda-1)(1-\alpha)+\beta} \left(\tfrac{A_M}{z_L}\right)^{\lambda} \left(\left(\left(\tfrac{1-\alpha}{\beta}\right)\varsigma\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}} - 1\right)
\end{aligned}$$

Furthermore,

$$\begin{aligned}
& \int_{z_L}^{\infty} f(z)\,dz + \tfrac{1-\beta-\alpha}{\beta}\tfrac{1}{z_L}\int_{z_L}^{\infty} z f(z)\,dz \\
= \quad & \left(\tfrac{A_M}{z_L}\right)^{\lambda} + \tfrac{1-\beta-\alpha}{\beta}\tfrac{1}{z_L}\left(\tfrac{A_M}{z_L}\right)^{\lambda}\tfrac{\lambda}{\lambda-1} z_L \\
= \quad & \left(\tfrac{A_M}{z_L}\right)^{\lambda}\left(1 + \tfrac{1-\beta-\alpha}{\beta}\tfrac{\lambda}{\lambda-1}\right),
\end{aligned}$$

[38] Suppose $F(x) = 1 - \left(\frac{x_0}{x}\right)^{\zeta}$. Then

$$P(x^{\eta} \leq k) = P\left(x \leq k^{1/\eta}\right) = 1 - \left(\frac{x_0}{k^{1/\eta}}\right)^{\zeta} = 1 - \left(\frac{x_0^{\eta}}{k}\right)^{\zeta/\eta}.$$

so that

$$
\begin{aligned}
E[\pi] &= f_O w\left(\left(\frac{1}{z^*}\right)^{\frac{1-\alpha-\beta}{1-\alpha}}\int_{z^*}^{z_L} z^{\frac{1-\alpha-\beta}{1-\alpha}} f(z)\,dz + \varsigma\left(\int_{z_L}^{\infty} f(z)\,dz + \frac{1-\beta-\alpha}{\beta}\frac{1}{z_L}\int_{z_L}^{\infty} z f(z)\,dz\right) - \left(\frac{A_M}{z^*}\right)^{\lambda}\right) \\
&= f_O w\left(\frac{1-\alpha}{\beta}\right)\varsigma\frac{\lambda(1-\alpha)}{(\lambda-1)(1-\alpha)+\beta}\left(\frac{A_M}{z_L}\right)^{\lambda}\left(\left(\left(\frac{1-\alpha}{\beta}\right)\varsigma\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}} - 1\right) \\
&\quad + f_O w\varsigma\left(\frac{A_M}{z_L}\right)^{\lambda}\left(1+\frac{1-\beta-\alpha}{\beta}\frac{\lambda}{\lambda-1}\right) - \left(\frac{A_M}{z_L}\right)^{\lambda}\left(\left(\frac{1-\alpha}{\beta}\right)\varsigma\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} \\
&= f_O w\left(\frac{A_M}{z_L}\right)^{\lambda}\left(\frac{1-\alpha-\beta}{(\lambda-1)(1-\alpha)+\beta}\right)\left(\left(\frac{1-\alpha}{\beta}\varsigma\right)^{\frac{\lambda(1-\alpha)}{1-\alpha-\beta}} + \frac{\varsigma}{\lambda-1}\right).
\end{aligned}
$$

Hence, the free entry condition requires that

$$
f_O w\left(\frac{A_M}{z_L}\right)^{\lambda}\left(\frac{1-\alpha-\beta}{(\lambda-1)(1-\alpha)+\beta}\right)\left(\left(\frac{1-\alpha}{\beta}\varsigma\right)^{\frac{\lambda(1-\alpha)}{1-\alpha-\beta}} + \frac{\varsigma}{\lambda-1}\right) = w f_E,
$$

which yields (A-12). Now consider the case of $\varsigma < \frac{\beta}{1-\alpha}$. Then

$$
\begin{aligned}
E[\pi] &= \int_{\tilde{z}} \pi^{NL}(z) f(z)\,dz \\
&= f_O(1-\varsigma)w\left(\frac{1}{\tilde{z}}\int_{\tilde{z}} z f(z)\,dz - \int_{\tilde{z}} f(z)\,dz\right) \\
&= f_O(1-\varsigma)w\left(\frac{1}{\tilde{z}}\left(\frac{A_M}{\tilde{z}}\right)^{\lambda}\frac{\lambda}{\lambda-1}\tilde{z} - \left(\frac{A_M}{\tilde{z}}\right)^{\lambda}\right) \\
&= f_O(1-\varsigma)w\left(\frac{A_M}{\tilde{z}}\right)^{\lambda}\frac{1}{\lambda-1}.
\end{aligned}
$$

Hence,

$$
\left(\frac{A_M}{\tilde{z}}\right)^{\lambda} = (\lambda-1)\frac{f_E}{f_O}\frac{1}{1-\varsigma}.
$$

Using (A-10), this implies that

$$
\frac{w}{p_G} = \left(\left(\frac{1}{(\lambda-1)f_E}\right)^{\frac{1}{\lambda}}\left(\frac{1}{f_O}\frac{1}{1-\varsigma}\right)^{\frac{\lambda-1}{\lambda}}(1-\alpha-\beta)\right)^{1-\alpha-\beta}(\beta A_{PS})^{\beta}\alpha^{\alpha}A_M^{1-\alpha-\beta}.
$$

OA-1.4 Proof of Proposition 5

Consider first the case of $\varsigma \geq \frac{\beta}{1-\alpha}$. The aggregate demand for lawyer services is given from Proposition 2 as

$$\begin{aligned} H_{PS} &= M\int_{z_L}^{\infty} H_{PS}(z) f(z)\, dz = M\varsigma f_0 \left(\frac{1}{z_L}\int_{z_L}^{\infty} z_i f(z)\, dz - \int_{z_L}^{\infty} f(z)\, dz\right) \\ &= M\varsigma f_0 \left(\left(\frac{A_M}{z_L}\right)^{\lambda}\frac{\lambda}{\lambda-1} - \left(\frac{A_M}{z_L}\right)^{\lambda}\right) \\ &= M\varsigma f_0 \frac{1}{\lambda-1}\left(\frac{A_M}{z_L}\right)^{\lambda}. \end{aligned}$$

Similarly, the aggregate demand for production workers is

$$\begin{aligned} H_M &= M\int_{z^*}^{\infty} H_M(z) f(z)\, dz = M\int_{z^*}^{z_L} H_M(z) f(z)\, dz + M\int_{z_L}^{\infty} H_M(z) f(z)\, dz \\ &= M\frac{\alpha}{\beta} f_O \varsigma \left(\left(\frac{1}{z_L}\right)^{\frac{1-\alpha-\beta}{1-\alpha}}\left(\int_{z^*}^{\infty} z^{\frac{1-\alpha-\beta}{1-\alpha}} f(z)\, dz - \int_{z_L}^{\infty} z^{\frac{1-\alpha-\beta}{1-\alpha}} f(z)\, dz\right) + \frac{1}{z_L}\int_{z_L}^{\infty} z_i f(z)\, dz\right) \\ &= M\frac{\alpha}{\beta} f_O \varsigma \left(\left(\frac{1}{z_L}\right)^{\frac{1-\alpha-\beta}{1-\alpha}}\left(\left(\frac{A_M}{z^*}\right)^{\lambda}\frac{\frac{\lambda(1-\alpha)}{1-\alpha-\beta}}{\frac{\lambda(1-\alpha)}{1-\alpha-\beta}-1}(z^*)^{\frac{1-\alpha-\beta}{1-\alpha}} - \left(\frac{A_M}{z_L}\right)^{\lambda}\frac{\frac{\lambda(1-\alpha)}{1-\alpha-\beta}}{\frac{\lambda(1-\alpha)}{1-\alpha-\beta}-1}(z_L)^{\frac{1-\alpha-\beta}{1-\alpha}}\right) + \left(\frac{A_M}{z_L}\right)^{\lambda}\frac{\lambda}{\lambda-1}\right) \\ &= M\frac{\alpha}{\beta} f_O \varsigma \left(\frac{A_M}{z_L}\right)^{\lambda}\left(\frac{\lambda(1-\alpha)}{(\lambda-1)(1-\alpha)+\beta}\left(\left(\frac{z_L}{z^*}\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha}} - 1\right) + \frac{\lambda}{\lambda-1}\right) \\ &= M\frac{\alpha}{\beta} f_O \varsigma \left(\frac{A_M}{z_L}\right)^{\lambda}\left(\frac{\lambda(1-\alpha)}{(\lambda-1)(1-\alpha)+\beta}\left(\left(\frac{1-\alpha}{\beta}\varsigma\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}} - 1\right) + \frac{\lambda}{\lambda-1}\right). \end{aligned}$$

The total number of people employed as overhead workers are

$$H_{OM} = f_0 M\left(\frac{A_M}{z^*}\right)^{\lambda}.$$

Finally, the total number of people employed for entry activities are

$$H_{EM} = f_E M.$$

Now note from (A-12) that

$$\left(\frac{z_L}{A_M}\right)^{\lambda} = \frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left[\left(\varsigma\frac{1-\alpha}{\beta}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} + \frac{\varsigma}{\lambda-1}\right]\frac{f_O}{f_E}.$$

Hence,

$$\begin{aligned}
H_M &= \frac{M\frac{\alpha}{\beta}f_O\varsigma\left(\frac{\lambda(1-\alpha)}{(\lambda-1)(1-\alpha)+\beta}\left(\left(\frac{1-\alpha}{\beta}\varsigma\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}}-1\right)+\frac{\lambda}{\lambda-1}\right)}{\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left[\left(\varsigma\frac{1-\alpha}{\beta}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}}+\frac{\varsigma}{\lambda-1}\right]\frac{f_O}{f_E}} \\
&= Mf_E\frac{\alpha}{\beta}\frac{\frac{\lambda(1-\alpha)}{(\lambda-1)(1-\alpha)+\beta}\varsigma\left(\frac{1-\alpha}{\beta}\varsigma\right)^{\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}}-\frac{\lambda(1-\alpha)}{(\lambda-1)(1-\alpha)+\beta}\varsigma+\varsigma\frac{\lambda}{\lambda-1}}{\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left(\varsigma\frac{1-\alpha}{\beta}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}}+\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\frac{\varsigma}{\lambda-1}} \\
&= Mf_E\frac{\alpha}{\beta}\frac{\frac{\lambda\beta}{(\lambda-1)(1-\alpha)+\beta}\left(\frac{1-\alpha}{\beta}\varsigma\right)^{\lambda\frac{(1-\alpha)}{1-\alpha-\beta}}+\varsigma\frac{\lambda}{\lambda-1}\frac{\beta}{(\lambda-1)(1-\alpha)+\beta}}{\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left(\varsigma\frac{1-\alpha}{\beta}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}}+\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\frac{\varsigma}{\lambda-1}} \\
&= Mf_E\frac{\lambda\alpha}{(1-\alpha-\beta)}
\end{aligned}$$

Note also that

$$\begin{aligned}
H_{PS}+H_{OM} &= M\varsigma f_0\frac{1}{\lambda-1}\left(\frac{A_M}{z_L}\right)^{\lambda}+f_0M\left(\frac{A_M}{z^*}\right)^{\lambda} \\
&= M\left(\frac{A_M}{z^*}\right)^{\lambda}f_0\left(\varsigma\frac{1}{\lambda-1}\left(\frac{z^*}{z_L}\right)^{\lambda}+1\right) \\
&= M\left(\frac{A_M}{z^*}\right)^{\lambda}f_0\left(\varsigma\frac{1}{\lambda-1}\left(\frac{1-\alpha}{\beta}\varsigma\right)^{-\lambda\frac{1-\alpha}{1-\alpha-\beta}}+1\right) \\
&= M\left(\frac{A_M}{z^*}\right)^{\lambda}f_0\left(\frac{\beta}{1-\alpha}\frac{1}{\lambda-1}\left(\frac{1-\alpha}{\beta}\varsigma\right)^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}+1\right) \\
&= M\left(\frac{A_M}{z^*}\right)^{\lambda}f_0\left(\frac{\beta}{1-\alpha}\frac{1}{\lambda-1}\left(\frac{1-\alpha}{\beta}\varsigma\right)^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}+1\right)
\end{aligned}$$

From (A-12) we have that that

$$\left(\frac{z^*}{A_M}\right)^{\lambda}=\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left[1+\frac{1}{\lambda-1}\left(\frac{\beta}{1-\alpha}\right)^{\frac{(1-\alpha)\lambda}{1-\alpha-\beta}}\varsigma^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}\right]\frac{f_O}{f_E}$$

Hence,

$$\begin{aligned}
H_{PS}+H_{OM} &= M\frac{f_0\left(\frac{\beta}{1-\alpha}\frac{1}{\lambda-1}\left(\frac{1-\alpha}{\beta}\varsigma\right)^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}+1\right)}{\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left[1+\frac{1}{\lambda-1}\left(\frac{\beta}{1-\alpha}\right)^{\frac{(1-\alpha)\lambda}{1-\alpha-\beta}}\varsigma^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}\right]\frac{f_O}{f_E}} \\
&= Mf_E\frac{\left(\frac{\beta}{1-\alpha}\frac{1}{\lambda-1}\left(\frac{1-\alpha}{\beta}\varsigma\right)^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}+1\right)}{\frac{(1-\alpha-\beta)}{\beta+(1-\alpha)(\lambda-1)}\left[1+\frac{1}{\lambda-1}\frac{\beta}{1-\alpha}\left(\frac{1-\alpha}{\beta}\varsigma\right)^{-\frac{(1-\alpha)(\lambda-1)+\beta}{1-\alpha-\beta}}\right]} \\
&= Mf_E\frac{\beta+(1-\alpha)(\lambda-1)}{1-\alpha-\beta}. \qquad \text{(OA-7)}
\end{aligned}$$

This implies that

$$\begin{aligned}
H_G &= Mf_E\frac{\lambda\alpha}{(1-\alpha-\beta)}+f_EM+Mf_E\frac{\beta+(1-\alpha)(\lambda-1)}{1-\alpha-\beta} \\
&= Mf_E\frac{\lambda}{1-\alpha-\beta}.
\end{aligned}$$

Hence,

$$M=\frac{1-\alpha-\beta}{\lambda}\frac{H_G}{f_E}.$$

Now consider $\varsigma<\frac{\beta}{1-\alpha}$. Using the results in Proposition 3, we get that total demand for production workers is given by

$$\begin{aligned}
H_M &= M\int_{\tilde{z}}H_M(z)f(z)\,dz \\
&= M\frac{\alpha}{1-\alpha-\beta}f_O(1-\varsigma)\frac{1}{\tilde{z}}\int_{\tilde{z}}z_if(z)\,dz \\
&= M\frac{\alpha}{1-\alpha-\beta}f_O(1-\varsigma)\left(\frac{A_M}{\tilde{z}}\right)^{\lambda}\frac{\lambda}{\lambda-1}.
\end{aligned}$$

Using the expression for $\left(\frac{\tilde{z}}{A_M}\right)^{\lambda}$ in (A-16) yields

$$H_M=M\frac{\alpha}{1-\alpha-\beta}f_O(1-\varsigma)\frac{\frac{\lambda}{\lambda-1}}{\frac{1}{\lambda-1}\frac{f_O}{f_E}(1-\varsigma)}=M\frac{\alpha\lambda}{1-\alpha-\beta}f_E.$$

The total demand for lawyers is given by

$$\begin{aligned}
H_{PS} &= M \int_{\tilde{z}} H_{PS}(z) f(z)\, dz \\
&= M \frac{\beta}{1-\alpha-\beta} f_O (1-\varsigma) \left(\frac{A_M}{\tilde{z}}\right)^{\lambda} \frac{\lambda}{\lambda-1} - M \frac{\kappa}{A_{PS}} \left(\frac{A_M}{\tilde{z}}\right)^{\lambda} \\
&= M \left(\frac{\beta}{1-\alpha-\beta} f_E \lambda - \frac{\kappa}{A_{PS}} \frac{f_E}{f_O} \frac{\lambda-1}{(1-\varsigma)}\right) \\
&= M f_E \left(\frac{\beta\lambda}{1-\alpha-\beta} - \varsigma \frac{\lambda-1}{(1-\varsigma)}\right).
\end{aligned}$$

The total demand for workers used for overhead is

$$\begin{aligned}
H_{OM} &= M f_O \left(\frac{A_M}{\tilde{z}}\right)^{\lambda} = M f_O \frac{1}{\frac{1}{\lambda-1}\frac{f_O}{f_E}(1-\varsigma)} \\
&= M f_E \frac{\lambda-1}{1-\varsigma},
\end{aligned}$$

so that again

$$\begin{aligned}
H_{PS} + H_{OM} &= M f_E \left(\frac{\beta\lambda}{1-\alpha-\beta} - \varsigma \frac{\lambda-1}{(1-\varsigma)}\right) + M f_E \frac{\lambda-1}{1-\varsigma} \\
&= M f_E \left(\frac{(\lambda-1)(1-\alpha)+\beta}{1-\alpha-\beta}\right).
\end{aligned}$$

This is the same expression as (OA-7).

OA-1.5 Proof that w/p_G is increasing in A_{PS}

The crux is to establish that

$$\frac{\partial \log(z_L)}{\partial A_{PS}} = -\psi(A_{PS}) \frac{1}{A_{PS}} \tag{OA-8}$$

where $\psi(A_{PS})$ is defined in the text. To prove that this is the case, note that (A-12) implies that:

$$\log(z_L) = \log A_M - \frac{1}{\lambda} \log f_E + \frac{1}{\lambda} \log Q(A_{PS}),$$

where[39]

$$Q(A_{PS}) \equiv (1-\alpha)^2 \frac{\kappa}{\beta A_{PS}} \lambda \frac{\left(\frac{1}{\Psi}\right)^{\lambda - \frac{1-\alpha-\beta}{1-\alpha}} - 1}{\beta + (1-\alpha)(\lambda-1)} + \left(1 + \frac{1-\alpha-\beta}{\beta}\frac{\lambda}{\lambda-1}\right)\frac{\kappa}{A_{PS}} - \frac{f_O}{\Psi^{\lambda}}$$
$$= \frac{1-\alpha-\beta}{\beta + (\lambda-1)(1-\alpha)} f_O \left[\left(\frac{\kappa(1-\alpha)}{\beta A_{PS} f_O}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} + \frac{\kappa}{A_{PS} f_O (\lambda-1)}\right] > 0.$$

Hence,

$$\frac{\partial \log(z_L)}{\partial A_{PS}} = \frac{1}{\lambda}\frac{\partial \log(Q(A_{PS}))}{\partial A_{PS}} = \frac{1}{\lambda}\frac{Q'(A_{PS})}{Q(A_{PS})}$$

Differentiating the expression of Q above yields

$$Q'(A_{PS}) = -\frac{1-\alpha-\beta}{\beta + (\lambda-1)(1-\alpha)} \times \frac{1}{A_{PS}}\left[\lambda \frac{1-\alpha}{1-\alpha-\beta} f_O \left(\frac{\kappa(1-\alpha)}{\beta A_{PS} f_O}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} + \frac{\kappa}{A_{PS}(\lambda-1)}\right],$$

and

$$\frac{1}{\lambda}\frac{Q'(A_{PS})}{Q(A_{PS})} = -\frac{\frac{1-\alpha}{1-\alpha-\beta}\lambda f_O \left(\frac{\kappa(1-\alpha)}{\beta A_{PS} f_O}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} + \frac{\kappa}{A_{PS}(\lambda-1)}}{\lambda f_O \left(\frac{\kappa(1-\alpha)}{\beta A_{PS} f_O}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} + \lambda\frac{\kappa}{A_{PS}(\lambda-1)}} \times \frac{1}{A_{PS}} \equiv -\psi(A_{PS})\frac{1}{A_{PS}}$$

where $\psi(A_{PS}) \in (0,1)$. Hence,

$$\frac{\partial \log(z_L)}{\partial A_{PS}} = \frac{1}{\lambda}\frac{Q'(A_{PS})}{Q(A_{PS})} = -\psi(A_{PS})\frac{1}{A_{PS}},$$

which establishes the result.

OA-2 The Computational Algorithm

Step 1: Clean Data: For each year perform the following steps:

1. Normalize the size of the population to unity

$$1 = \sum_r L_{rt}$$

[39]The algebra leading to the simplified expression of $Q(A_{PS})$ is the following:

$$Q(A_{PS}) = \left[\begin{array}{c} (1-\alpha)^2 \frac{\kappa}{\beta A_{PS}} \lambda \frac{\left(\frac{\kappa(1-\alpha)}{\beta A_{PS} f_O}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}-1} - 1}{\beta + (1-\alpha)(\lambda-1)} \\ + \left(1 + \frac{1-\alpha-\beta}{\beta}\frac{\lambda}{\lambda-1}\right)\frac{\kappa}{A_{PS}} - f_O\left(\frac{\kappa(1-\alpha)}{\beta A_{PS} f_O}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} \end{array}\right]$$
$$= \left[\begin{array}{c} (1-\alpha)\lambda \frac{\left(\frac{\kappa(1-\alpha)}{\beta A_{PS} f_O}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} f_O}{\beta + (1-\alpha)(\lambda-1)} - f_O\left(\frac{\kappa(1-\alpha)}{\beta A_{PS} f_O}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} \\ + \left(1 + \frac{1-\alpha-\beta}{\beta}\frac{\lambda}{\lambda-1}\right)\frac{\kappa}{A_{PS}} - (1-\alpha)^2 \frac{\kappa}{\beta A_{PS}} \lambda \frac{1}{\beta + (1-\alpha)(\lambda-1)} \end{array}\right]$$
$$= \frac{1-\alpha-\beta}{\beta + (\lambda-1)(1-\alpha)}\left[f_O\left(\frac{\kappa(1-\alpha)}{\beta A_{PS} f_O}\right)^{\lambda\frac{1-\alpha}{1-\alpha-\beta}} + \frac{\kappa}{A_{PS}(\lambda-1)}\right]$$

2. Normalize the level of wages to unity

$$1 = \sum w_{rt} L_{rt} \tag{OA-9}$$

Step 2: Calculate human capital levels H_{rst} and H_{rt} To calculate the total supply of human capital in region r at time t we use the information on education attainment. Let ρ denote the yearly return to education. Suppose we have G groups of education attainment and group g has s_g years of schooling. Let l_{rgt} denote the share of people in region r at time t in education group g. We then calculate average human capital in region r at time t as

$$h_{rt} = \sum_{g=1}^{G} \exp(\rho s_g) l_{rgt}. \tag{OA-10}$$

The aggregate supply of human capital is then given

$$H_{rt} = h_{rt} L_{rt}. \tag{OA-11}$$

To calculate the distribution of human capital units across sectors *within* a location, we rely sectoral earnings shares, which in our theory are proportional to human capital units. Hence, we calculate H_{rst} as

$$H_{rst} = \frac{\sum_{i \in s} w_{rt}^i}{\sum_i w_{rt}^i} \times H_{rt}. \tag{OA-12}$$

Here, w_{rt}^i denotes total earnings individual i in region r at time t observed in the micro data. Hence, $\sum_{i \in s} w_{rt}^i$ are aggregate earnings in sector s in region r and $\sum_i w_{rt}^i$ are aggregate earnings in region r. Hence, we use the information in human capital (schooling) to measure human capital differences across space and time and information on relative earnings to measure human capital differences across sectors within a location. For the remainder we treat H_{rst} and H_{rt} as known.

Step 3: Pick structural parameters For now set the following parameters

$$\begin{aligned} \varepsilon &= 0.297 \\ \omega_A &= 0.01 \\ \tilde{\nu}_{CS} &= -1 \end{aligned}$$

Step 4: Calibrate remaining structural parameters to ensure market clearing in 1987 and 2011 Consider the market clearing condition in (A-28). Using (OA-9), this can be written as

$$\begin{aligned} \sum_r w_{rt} H_{rFt} &= \omega_A \sum_{r=1}^{R} w_{rt} H_{rt} + \tilde{\nu}_A \sum_{r=1}^{R} \left(\omega_{CS} - \frac{H_{rCSt}}{H_{rt}} \right) w_{rt} H_{rt} \\ &= \omega_A + \tilde{\nu}_A \omega_{CS} - \tilde{\nu}_A \sum_{r=1}^{R} H_{rCSt} w_{rt}. \end{aligned}$$

Given ω_A, chose $\tilde{\nu}_A$ and ω_{CS} to solve the two equations

$$\begin{aligned} \sum_r w_{r1987} H_{rF1987} &= \omega_A + \tilde{\nu}_A \omega_{CS} - \tilde{\nu}_A \sum_{r=1}^{R} H_{rCS1987} w_{r1987} \\ \sum_r w_{r2011} H_{rF2011} &= \omega_A + \tilde{\nu}_A \omega_{CS} - \tilde{\nu}_A \sum_{r=1}^{R} H_{rCS2011} w_{r2011} \end{aligned}$$

Step 5: Calibrate relative productivities a_{rFt} and a_{rGt} Equations (A-19) and (A-20) allows us measure a_{rFt} and a_{rGt} directly from the data. We replicate these equations here for convenience:

$$a_{rF} = \left(\frac{L_{rF} w_r^{\sigma}}{\sum_{r=1}^{R} \left(L_{rF}\right) w_r^{\sigma}} \right)^{\frac{1}{\sigma-1}}$$

$$a_{rG} = \left(\frac{L_{rG} w_r^{\sigma}}{\sum_{r=1}^{R} \left(L_{rG}\right) w_r^{\sigma}} \right)^{\frac{1}{\sigma-1}}$$

Note that these are insensitive to the scale of w_{rt}

Step 6: Calculate the level of productivity in manufacturing and agriculture in 1987 Given the price normalization $p_{G1987} = 1$ and the observed relative price p_{1987}^{AG}, calculate A_{G1987} and A_{F1987} as

$$A_{F1987} = \frac{1}{p_{1987}^{AG}} \left(\sum_r \left(\frac{w_{r1987}}{a_{rF1987}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

$$A_{G1987} = \left(\sum_r \left(\frac{w_{r1987}}{a_{rG1987}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

Step 7: Calculate the level of productivity in CS in 1987 From (A-26) we get that $A_{rCS1987}$ solves the equation

$$\frac{H_{rCS1987}}{H_r} = \omega_{CS} - \left(p_{1987}^{AG}\right)^{\varepsilon\omega_A} \left(\frac{w_{r1987}}{A_{rCS1987}} \right)^{\varepsilon\omega_{CS}} \left(E_{r1987}[q] w_{r1987}\right)^{-\varepsilon}$$

Step 8: Calculate physical quantities in 1987 The physical quantities of consumption in 1987 are given by

$$y_{rCS1987} = A_{rCS1987} L_{rCS1987}$$

$$y_{F1987} = \frac{\text{Spending on agricultural goods}}{P_{A1987}} = \frac{\text{Income in agriculture}}{P_{A1987}} = \frac{\sum_r w_{r1987} L_{rF1987}}{p_{1987}^{AG}}$$

$$y_{G1987} = \sum_r w_{r1987} L_{rG1987}$$

Step 9: Find the right scale in 1999

- Pick a scalar λ^{1999}.
- Calculate A_{G1999} and A_{F1999} as

$$A_{F1999} = \frac{1}{p_{1999}^{AG}} \left(\sum_r \left(\frac{\lambda^{1999} w_{r1999}}{a_{rF1999}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} = \frac{\lambda^{1999}}{p_{1999}^{AG}} \left(\sum_r \left(\frac{w_{r1999}}{a_{rF1999}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

$$A_{G1999} = \left(\sum_r \left(\frac{\lambda^{1999} w_{r1999}}{a_{rG1999}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} = \lambda^{1999} \left(\sum_r \left(\frac{w_{r1999}}{a_{rG1999}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

- Calculate