Production Networks and Misallocation*

Pravin Krishna†
Johns Hopkins University and NBER

Heiwai Tang‡
Johns Hopkins University and CESIfo

July 1, 2017
(Preliminary. Comments are welcome.)

Abstract

This paper studies aggregate TFP losses due to policy-distorted (mis-)allocation of resources across heterogeneous firms in the presence of intermediate input trade. Different from the existing literature on misallocation, we pay close attention to both allocative efficiency between sectors and amplification through industry input-output (IO) linkages. We extend the model of Hsieh and Klenow (2009) to study firms' decisions to source inputs both domestically and globally from multiple industries. Solving the general equilibrium model with industry linkages reveals that an economy's aggregate TFP loss due to distortions is equal to the geometric mean of sector-level TFP losses, with weights equal to the sector's Domar weights. Using Chinese firm-level and Indian plant-level data over the period of 2000-2007, we find, however, that the estimated aggregate TFP losses due to resource misallocation with IO linkages are lower, relative to the value-added approach adopted by Hsieh and Klenow (2009). Despite the amplification of distortion costs due to input-output linkages, this surprising result is due to the fact that the estimated sectoral TFP losses are on average an order of magnitude smaller than the estimates using the value-added approach, as for both countries, the marginal revenue products of intermediate inputs are much less dispersed across firms within industries, compared to those of labor and capital. Finally, trade liberalization raises the aggregate share of imported intermediates in total intermediates for India, increasing the size of the IO multiplier and thus the macroeconomic costs of firm-level distortions.

Key Words: misallocation, input-output linkages, production networks, industrialization
JEL Classification Numbers: F14, D22, D85, L10, L14, O11, O47, O53

*We thank Hugo Hopenhayn, Chang-Tai Hsieh, Gianmarco Ottaviano, Diego Restuccia, Chad Syverson, Johannes Van Biesebroeck, Xiaodong Zhu, and participants at various seminars and conferences for comments.
†pravin_krishna@jhu.edu
‡hwtang@jhu.edu

1 Introduction

A vast literature studies how resource misallocation across firms, due to policy distortions, results in drastic income and productivity differences across countries.[1] The analysis has largely assessed the effects of misallocation on individual industries and then aggregate up the effects across industries to the national level. However, an economy is an intertwined web of production units, each uses some inputs produced by suppliers in the upstream and sells output to either final-good consumers or other firms in the downstream in the production network. A firm's productivity affects not only its performance, but also its trade partners, and with whom they further trade. As such, there is a recent revival of interest in understanding the potential amplification effect of firms' and industries' input-output (IO) linkages on countries' macroeconomic outcomes, with a particular focus on business-cycle fluctuations (Kiyotaki and Moore, 1997; Ciccone, 2002; Acemoglu, Antras, and Helpman, 2007; Jones, 2011, 2013; Acemoglu et al. 2012; Carvalho and Gabaix, 2013, Baqaee, 2016, and Acemoglu, Ozdaglar and Tahbaz-Salehi, 2017, among others).[2]

It is not hard to appreciate the importance of studying how industry linkages can amplify the costs of resource misallocation from the micro level. Research has suggested that reforms of the state-monopolized banking sector in China (Hachem and Song, 2016) or the centralized electricity markets in India (Hunt, Collard-Wexler, and O'Connell, 2016) can bring about substantial productivity gains. The economic benefits of reforms are obviously beyond the two sectors, as higher efficiency of those sectors will lower the prices and raise the quality of inputs used by downstream firms. Intuitively, the more upstream a sector is, the larger the macroeconomic impact of reforms as more firms and sectors can benefit from the fruits of reforms in downstream sectors of an economy.[3] Despite the obvious reasons, the literature has not incorporated industry linkages in the framework for studying resource misallocation to systematically evaluate their relevance.[4]

This paper sets to close this gap by developing a quantifiable model to analyze firms' input sourcing, which is affected by policy distortions on the one hand but amplify the macroeconomic costs through affecting the domestic input-output (IO) linkages on the other.[5] To this end, we extend the model of Hsieh and Klenow (2009) [HK hereafter] to study firms' sourcing of inputs

[1] This literature includes Banerjee and Duflo (2005), Alfaro et al. (2009), Guner et al. (2008), Restuccia and Rogerson (2008), Hsieh and Klenow (2009), Buera et al. (2011), and Bartelsman et al. (2013), among many others. See Restuccia and Rogerson (2013) and Hopenhayn (2014) for reviews.

[2] An old literature, started by Hirschman (1958), highlights the importance of considering sectors' linkages to other sectors in designing industrial policies. Subsequent studies (e.g., Chenery and Watanabe, 1958; Yotopoulos and Nugent, 1973; Chenery et al., 1986; Jones, 1976; Shultz, 1982) examine the relationship between a country's industry linkages and economic development. A classic study by Long and Plosser (1983) is one of the first to quantitatively examine the importance of domestic input-output linkages in propagating business shocks.

[3] See Liu (2017) for a formal theoretical treatment of the argument in the study of the industrial policies in China.

[4] Consider the macroeconomic effect of the state monopoly in the banking sector in China or the unstable and inefficient power grid in India.

[5] Existing research (e.g., Antràs, Fort, and Tintelnot, 2016 and Bernard, Moxnes, and Saito, 2014) has also examined firms' sourcing decisions and buyer-supplier relationships, but the theoretical models developed usually consider only one layer of upstream-downstream relationships and cannot be extended to study the amplification effects or shocks or misallocations through the realistic industry linkages. Separately, there is an extensive literature that takes the input-output linkages across industries seriously to measure and explain the pattern of value added trade (e.g., Johnson and Noguera, 2012 and Koopman, Wang and Wei 2014).

from both domestic and foreign suppliers. We build explicitly full IO linkages across industries in the domestic economy to study how policy distortions, like taxes on output and inputs, lower firms' output and raise their prices. This, in turn, will raise the input prices of other firms, leading to their suboptimal outsourcing (measured in quantities). Specifically, firms, which are exogenously heterogeneous in productivity and distortions on output and inputs, produce horizontally differentiated varieties of goods by combining primary factors of production (labor and capital) and intermediate inputs from multiple industries. IO linkages alters the costs of policy distortions, thereby lowering firm and sectoral productivity and raising input prices. We solve this circular feedback loop in general equilibrium. Solving the model with industry linkages reveals that an economy's aggregate total factor productivity (TFP) losses due to distortions should be equal to the geometric mean of sector-level TFP losses, with weights equal to the sectors' Domar weights (i.e., the ratio of gross output to total manufacturing value added), rather than their consumption shares as in HK. As pointed out by Jones (2011), when studying the aggregate implications of industry linkages, the details of the IO linkages do not matter– only the sum of the Domar weights across industries does. The sum of the Domar weights (i.e., the IO multiplier) determines how much micro and sector-level distortions get amplified to the aggregate level. A large multiplier may imply a downward bias in the existing estimates of the aggregate TFP losses due to resource misallocation.

Through the lens of such model and using micro data from China and India, we assess how IO linkages amplify the macro cost of resource misallocation across firms; and whether incorporating input sourcing affects the existing estimates of the TFP impact of resource allocation at the industry level. Using detailed establishment data from both countries over the sample period of 2000-2007, we find that the IO multiplier in the manufacturing sector is around 3.6 for China and 1.8 for India.[6] Despite the large multiplier, we find for both countries *smaller* aggregate TFP losses caused by resource misallocation, compared to the estimates of HK, who adopt a value-added approach and ignore IO linkages. Specifically, our gross-output approach with IO linkages reveals that hypothetically removing all distortions raises India's manufacturing TFP by about 86% in 2000, with the potential gain increasing over time to 90% in 2007. Replicating the quantitative exercises of HK shows that the potential gains could be 100% in 2000 but *diminished* to 80% in 2007. For China, our approach shows that removing all distortions in manufacturing can potentially raise the sector's TFP by about 129% in 2000, with the potential gain decreasing gradually over time to 119% in 2007. Estimates based on the value-added approach suggest an opposite and increasing trend of potential gains from 130% in 2000 to 140% in 2007.

The similarity of the estimated costs of resource misallocation between the two approaches, despite significant IO amplification, warrant in-depth investigations. Our model shows that if all firm-level distortions are lognormally and independently distributed, a sector's TFP loss could be lower if the dispersion of firms' marginal revenue products of material (MRPM) is sufficiently lower than those of the primary factors of production (i.e., capital and labor). Indeed, we find that TFP

[6] The computation of the IO mulipliers ignores service and other non-manufacturing sectors because of data limitation. Using World Input-Output Database, which allows us to incorporate non-manufacturing sectors, we find that the IO muliplier for the Chinese economy is around 1.6 for India and 2.1 for China in the sample period.

losses due to resource misallocation at the 4-digit industry level based on the gross-output approach are on average an order of magnitude smaller than those computed using the value-added approach. The within-sector dispersion of firms' MRPM is much smaller than that of firms' marginal revenue product of capital (MRPK) for both countries. In particular, the average standard deviation of the MRPM during the sample period is about 0.6 for material in India, compared to 1.3 for capital. The corresponding numbers for China are 0.3 and 1.1, respectively. The significantly more efficient allocation of materials implies much smaller TFP losses at the sector level. Despite the significant IO multipliers, they are not large enough to push the geometric means of the sectoral losses above the aggregate TFP losses estimated based on the value-added approach.

Both China and India have recently implemented several domestic and external liberalizations, and also embarked on large-scale national infrastructure projects. Our framework can be used to quantify the positive productivity effects of such reforms in each country. As a first pass, we theoretically and quantitatively assess the effects of trade liberalization in India on the cost of resource misallocation. During the sample period, the average firm and aggregate share of imported materials in India rose substantially, while that for China declined.[7] According to our model, a larger share of imports in total intermediates can paradoxically increase the domestic IO multiplier, if a country is a net exporter in the relatively upstream sectors while being a net importer in the relatively downstream sectors. A larger multiplier, given the same levels of distortions in each sector, implies greater macroeconomic costs of resource misallocation. In other words, trade liberalization is not a substitute for reforms. Rather, it raises the urgency of reforms.

The rest of the paper proceeds as follows. Section 2 outlines the theoretical framework. Section 3 describes the accounting framework. Section 4 describes the firm-level for the quantitative exercise, which will be presented in Section 5. The last section concludes with some remarks about future research.

2 Model

We model industry linkages between firms in the presence of policy distortions. The advantage of considering heterogeneous firms is beyond methodological. It permits a more systematic decomposition of the cost of misallocation into various components (e.g., the productivity of sectors, the price indices of input sectors, and the IO multiplier). We will be able to quantify the amplification effects of the domestic IO linkages and conduct a host of counterfactual exercises, including one that sets the IO linkage of an economy back to the level of the initial sample year.

2.1 Environment

In the baseline model, we will first study firms' intermediate input sourcing from domestic suppliers across different sectors, subject to market prices and firm-specific distortions. To focus on the main

[7] The paradoxical downward trend in the share of imports in the Chinese economy during its increasing engagemnt in global trade has been systematically documented and explained in Kee and Tang (2016).

idea related to IO linkages, we assume labor as the only primary factor of production. In the quantitative analysis, we will add capital and imported intermediate inputs in Section 2.5. Our model borrows insights from Fieler, Eslava and Xu (2016) to consider multiple sources of demand for a firm's output – those from domestic final-good consumers and producers across all industries in the immediate downstream. The resulting demand system has a recursive structure. We will consider foreign sourcing in Section 2.5.

2.1.1 Final Good Sectors

There are L consumers in an economy, who also supply labor inelastically to make wages (normalized to 1). They consume goods from J sectors (indexed by j) according to a Cobb-Douglas utility function, each having a cost share equal to β_j that satisfies $\sum_j^J \beta_j = 1$. Consumers consume all available varieties in each sector, with the consumption composite of sector j being a constant elasticity of substitution (CES) aggregate over all available varieties in the sector. The corresponding consumer price index is

$$P = B\prod_{j=1}^{J} P_j^{\beta_j} = B\prod_{j=1}^{J}\left[\int_{\Omega_j} p_j(\omega)^{1-\sigma_j}\, d\omega\right]^{\frac{\beta_j}{1-\sigma_j}}, \tag{1}$$

where $B \equiv \prod_{j=1}^{J} \beta_j^{-\beta_j}$, σ_j is the elasticity of substitution between varieties, and Ω_j denotes the set of available varieties in sector j in the economy.

2.1.2 Production of Intermediate Varieties

Firms use domestically produced intermediate inputs from all sectors. The model features firms with heterogeneous and exogenous core productivity (z), output distortion (τ_y), and labor distortion (τ_l), similar to HK. For simplicity, distortions on intermediate inputs are assumed to be zero, but will be considered in the quantitative exercise. Each firm maximizes profit by choosing the optimal levels of intermediate inputs and labor, as well as optimal output and price, given z, τ_y, τ_l, and its own iso-elastic demand.

A firm with productivity z in sector j has the production function as

$$q_j(z) = z\Lambda_j l_j(z)^{\lambda_j} \prod_{k=1}^{K} m_{jk}(z)^{\gamma_{jk}}, \tag{2}$$

where $l_j(z)$ is labor; λ_j is the cost share of labor; γ_{jk} captures the cost share of sector-k composite inputs to produce sector-j goods. $\Lambda_j \equiv \lambda_j^{-\lambda_j}\prod_{k=1}^{K}\gamma_{jk}^{-\gamma_{jk}}$. Constant returns to scale require that $\sum_{k=1}^{K}\gamma_{jk} = 1-\lambda_j$.

When sourcing intermediates from sector k, a firm will purchase an input composite $m_{jk}(z)$,

which is a sector-specific CES aggregate over all varieties in sector k, with the same σ_k for final-good consumption. The firm's output price is a standard markup over marginal cost, which depends on σ_k, the labor costs, and the price index of intermediates W_j, which equals

$$W_j = \prod_{k=1}^{K} P_k^{\gamma_{jk}}. \tag{3}$$

Given its own iso-elastic demand, firm z's revenue is

$$r_j(z) = \left(\frac{p_j(z)}{P_j}\right)^{1-\sigma_j} R_j, \tag{4}$$

where R_j is the aggregate spending on good j by final-good consumers and all immediate downstream sectors:

$$R_j = \underbrace{\beta_j P Y}_{\text{final goods}} + \underbrace{\sum_{k=1}^{K} \gamma_{jk}\left(1-\sigma_k^{-1}\right) R_k}_{\text{intermediates}}. \tag{5}$$

2.2 Firm Equilibrium

A firm's profit maximization problem is

$$\max_{l_j, m_{jk}\ \forall k} (1-\tau_y) r_j(z) - \sum_{k=1}^{K} P_k m_{jk}(z) - (1+\tau_l) l_j(z), \tag{6}$$

where τ_y and τ_l are the firm-specific output and labor distortions, which can be interpreted as a tax (subsidy), governments' biased implementations of laws, expropriation, and so on. For now, we assume no distortions on intermediate inputs for simplicity.

It aims to choose the optimal levels of labor and intermediate inputs in all K sectors. Solving (6) yields the firm's optimal choices of labor as

$$l_j(z) = \left[\frac{(1-\tau_y) z^{\eta_j}}{(1+\tau_l)^{1-\eta_j(1-\lambda_j)}}\right]^{\frac{1}{1-\eta_j}} \frac{L_j}{T_j}, \tag{7}$$

and intermediate inputs for each input sector k as

$$m_{jk}(z) = \frac{\gamma_{jk}}{\lambda_j P_k}\left[\frac{(1-\tau_y) z^{\eta_j}}{(1+\tau_l)^{\lambda_j \eta_j}}\right]^{\frac{1}{1-\eta_j}} \frac{L_j}{T_j} \quad \forall k \tag{8}$$

with

$$T_j \equiv \iiint \left[\frac{(1-\tau_y) z^{\eta_j}}{(1+\tau_l)^{1-\eta_j(1-\lambda_j)}}\right]^{\frac{1}{1-\eta_j}} dG_y(\tau_y)\, dG_l(\tau_l)\, dH(z). \tag{9}$$

The functions $G_y(\tau_y)$, $G_l(\tau_l)$ and $H(z)$ are the cumulative distribution functions of τ_y, τ_l, and z,

respectively. All $\tau's$ are assumed to be lognormal, with means equal to μ_y and μ_l, and variances equal to υ_y and υ_l, respectively, while z is independently and log-normally distributed with mean and variance equal to μ_z and υ_z.

Plugging (7) and (8) into the (4) yields the firm's revenue in terms of sector-level variables and firm-level exogenous terms:

$$r_j(z) = (\lambda_j W_j)^{-\eta_j} \left[\frac{(1-\tau_y)\, z}{(1+\tau_l)^{\lambda_j}}\right]^{\frac{\eta_j}{1-\eta_j}} \left(\frac{L_j}{T_j}\right)^{\eta_j} P_j^{\eta_j} R_j^{1-\eta_j}. \tag{10}$$

2.3 Industry Equilibrium

To solve for a sector's output, we aggregate revenues of all firms, expressed as (10), in the sector, and then divide the integral by its price index. Sector j's output is

$$Y_j = (\lambda_j W_j)^{-1} \left[\int \left[\frac{(1-\tau_y)\, z}{(1+\tau_l)^{\lambda_j}}\right]^{\frac{\eta_j}{1-\eta_j}}\right]^{\frac{1}{\eta_j}} \int \left[\frac{(1-\tau_y)\, z^{\eta_j}}{(1+\tau_l)^{1-\eta_j(1-\lambda_j)}}\right]^{-\frac{1}{1-\eta_j}} L_j.$$

If there is no distortion (i.e., $\tau_y = \tau_l = 0$ for all firms), the efficient level of output will be

$$Y_j^* = \left(\lambda_j W_j^*\right)^{-1} \left[\int z^{\frac{\eta_j}{1-\eta_j}}\right]^{\frac{1}{\eta_j}} \left[\int z^{\frac{1}{1-\eta_j}}\right]^{-1} L_j.$$

Using the definition that $TFP_j = \dfrac{Y_j}{\Lambda_j L_j^{\lambda_j} \prod_{k=1}^{K} M_{jk}^{\gamma_{jk}}}$, we can express sector j's aggregate (log) TFP as

$$TFP_j = \left[\int \left[\frac{(1-\tau_y)\, z}{(1+\tau_l)^{\lambda_j}}\right]^{\frac{\eta_j}{1-\eta_j}}\right]^{\frac{1}{\eta_j}} \left[\int \left[\frac{(1-\tau_y)\, z^{\eta_j}}{(1+\tau_l)^{\eta_j\lambda_j}}\right]^{\frac{1}{1-\eta_j}}\right]^{-(1-\lambda_j)} T_j^{-\lambda_j} \tag{11}$$

Given that z, $(1-\tau)$, and all $(1+\tau)'s$ are all distributed log-normal, evoking the law of large number implies (see appendix for details):

$$\log TFP_j = \mu_z + \frac{\sigma_j - 1}{2}\upsilon_z - \frac{\sigma_j}{2}\upsilon_y - \frac{\lambda_j + \lambda_j^2(\sigma_j - 1)}{2}\upsilon_l \tag{12}$$

In other words, we can express the (log) ratio of sector j's TFP with distortions to the efficient level of TFP (TFP_j^e) as:

$$\log\left(\frac{TFP_j}{TFP_j^e}\right)_{baseline} = -\frac{\sigma_j}{2}\upsilon_y - \frac{\lambda_j + \lambda_j^2(\sigma_j - 1)}{2}\upsilon_l, \tag{13}$$

where υ_y and υ_l are the variance of τ_y and τ_l, respectively. The extent to which labor distortions can lower sectoral TFP depends on λ_j, the cost share of labor in production.

While we will let the data to inform us about how incorporating intermediate inputs in the production will affect the estimation of a sector's TFP loss thus the aggregate TFP, some theoretical insights can be obtained here. Consider the one-factor model without intermediate inputs. When TFP is measured based on firms' value added rather than gross output, $\lambda_j = 1$ and the corresponding ratio of distorted TFP to efficient TFP is

$$\log\left(\frac{TFP_j}{TFP_j^e}\right)_{va} = -\frac{\sigma_j}{2}\upsilon_y - \frac{\sigma_j}{2}\upsilon_l. \tag{14}$$

Before considering IO linkages, we notice that even at the sector level, estimated TFP losses due to distortions can be quite different depending on whether we adopt the gross output or value added approach to measure TFP. With an additional factor of production considered, it is possible that the estimated TFP loss at the sector level is lower, if the dispersion of the marginal revenue product of the additional input (i.e., material here) is significantly less dispersed than those of the primary production factors (labor and capital) and output considered in the existing literature on misallocation.

To illustrate the logic, let us assume that firms are also subject to a random tax when purchasing intermediate inputs, which increase the price of intermediate inputs by a multiple of $(1 + \tau_m)$. Suppose the distribution of τ_m across firms is also log-normal with mean and variance equal to μ_m and υ_m, respectively, the TFP ratio of (13) now becomes

$$\log\left(\frac{TFP_j}{TFP_j^e}\right)_{gross} = -\frac{\sigma_j}{2}\upsilon_y - \frac{1}{2}\left(\lambda_j\upsilon_l + (1-\lambda_j)\upsilon_m\right) - \frac{\sigma_j - 1}{2}\left(\lambda_j^2\upsilon_l + (1-\lambda_j)^2\upsilon_m\right).$$

Notice that $\log\left(\frac{TFP_j}{TFP_j^e}\right)_{gross} > \log\left(\frac{TFP_j}{TFP_j^e}\right)_{va}$ if

$$\frac{\upsilon_m}{\upsilon_l} < \frac{\sigma_j + (\sigma_j - 1)\lambda_j}{1 + (\sigma_j - 1)(1 - \lambda_j)}$$

In other words, by considering multiple factors of production, the estimated TFP loss due to resource misallocation at the sector level will be different. If the additional factors considered are somehow subject to sufficiently less distortion compared to those considered in the baseline model, the estimated TFP loss in a more holistic model can be smaller. We will use Chinese and Indian micro data to verify the potential biases in the literature. For now, let us solve for the general equilibrium of the model and derive the expressions to quantify the amplification of the costs of misallocation through IO linkages.

2.4 General Equilibrium with Input-Output Linkages

To close the model in general equilibrium so that we can study the aggregate impact of resource misallocation, we follow Jones (2013) to use the economy's market clearing conditions and con-

sumers' optimality conditions to solve for the endogenous macro variables. The goal is to express price indices (P), aggregate output (Y), and aggregate TFP as functions of labor endowment, L, and variances of the distortion factors. Consumer welfare will naturally be the inverse of the consumer price index P.

Let us start with the resource constraint for sector j:

$$P_j Q_j = \underbrace{P_j C_j}_{\text{final-good consumers}} + \underbrace{\sum_k \gamma_{kj}\eta_k P_k Q_k}_{\text{demand for inputs}},$$

where $P_j C_j$ is the demand from final-good consumers and $\sum_k \gamma_{kj}\eta_k P_k Q_k$ is the intermediate-input demand from firms in all sectors, including sector j itself. Notice that $\eta_k < 1$ is added to "mark down" input demand that depends on downstream sectors' revenue, which includes profits. Replacing P_j by $\beta_j Y/C_j$ and cancelling Y give

$$\frac{\beta_j Q_j}{C_j} = \beta_j + \sum_k \gamma_{kj}\eta_k \frac{\beta_k Q_k}{C_k}, \tag{15}$$

where $Q_k = \left[\int q_k(z)^{\frac{\sigma_k - 1}{\sigma_k}} dG(z)\right]^{\frac{\sigma_k}{\sigma_k - 1}}$.

Let us also define $v_j \equiv \frac{\beta_j Q_j}{C_j}$, similar to Jones (2013), to express (15) in a more compact form. By stacking the equations across all sectors yields a recursive equation in matrix form as

$$\mathbf{v} = \beta + \mathbf{B}\mathbf{v},$$

where $\mathbf{v}$ contains elements of $v_j \equiv \frac{\beta_j Q_j}{C_j}$ and $\mathbf{B}$ is the IO matrix with elements equal to $\gamma_{kj}\eta_k$.

The recursive formula can be solved out as:

$$\delta = (1 - \mathbf{B})^{-1}\beta, \tag{16}$$

where δ is a vector of solutions to $\frac{\beta_j Q_j}{C_j}$. Thanks to Cobb-Douglas preferences, β is vector of consumers' expenditure shares on different goods (components: β_j). Substituting $\beta_j = \frac{P_j C_j}{PY}$ in $\frac{\beta_j Q_j}{C_j}$ gives $\frac{P_j Q_j}{PY}$, which implies that δ is also a vector of sectors' Domar weights (the ratio of a sector's value added to the economy's total value added). As will be obvious immediately, the vector δ plays a crucial role in determining the magnitude of the amplification of TFP losses to the macroeconomic level.

Using the first order conditions for consumers and with some algebra, we can express sector-level demand for intermediate k as[8]

$$M_{jk} = \gamma_{jk}\eta_j \frac{\delta_j}{\delta_k} Q_k$$

[8] Notice that besides paremeters γ_{jk} and η_j, all variables will be ultimately affected by the variance of distortion and average input price indices.

We can also express each sector's labor demand as a function of the economy's labor endowment and distribution of distortion factors and input prices. Plugging these inputs into a sector's production function yields the sector's quantity supplied as a recursive equation of all other sectors' quantity supplied:

$$Q_j = TFP_j \left(l_j L\right)^{\lambda_j} \prod\nolimits_{k=1} \left[\gamma_{jk}\eta_j \frac{\delta_j}{\delta_k} Q_k\right]^{\gamma_{jk}},$$

where l_j is the allocation share of labor to sector j.

Taking log over this equation and stacking $\log Q_j$ up as a vector gives

$$\mathbf{q} = \mathbf{tfp} + \omega_q + \lambda \log L + \mathbf{B}\mathbf{q}, \tag{17}$$

where $\mathbf{tfp}$ is the vector of $\log TFP_j$. λ is a vector of λ_j. ω_q is a vector of

$$\omega_{qj} \equiv \lambda_j \log\left(l_j\right) + \sum\nolimits_{k=1} \gamma_{jk} \left[\log \gamma_{jk}\eta_j + \log \frac{\delta_j}{\delta_k}\right].$$

Rearranging (17) yields

$$\mathbf{q} = \left(1 - \mathbf{B}\right)^{-1} \left(\mathbf{tfp} + \omega_q + \lambda \log L\right). \tag{18}$$

To transform this vector of sectoral gross output to an economy' total value added (GDP), we use the identity that a country's GDP is equal to the weighted sum of final-good consumption across sectors, with weights equal to the cost shares according to the representatives' Cobb-Douglas utility function (i.e., $\log Y = \beta' \log \mathbf{C}$). Together with the outcomes of consumers' optimization: $\log C_j = \log\left(\beta_j Q_j / \delta_j\right)$, we can express the economy's (log) GDP as

$$\log Y = \beta' \left(1 - \mathbf{B}\right)^{-1} \mathbf{tfp} + \beta' \left(1 - \mathbf{B}\right)^{-1} \lambda \log L + \log \varepsilon \tag{19}$$

where ε is a collection of sector-specific constants.[9]. The economy's GDP can be further expressed in levels as

$$Y = \left(\prod\nolimits_j TFP_j^{\delta_j}\right) L\varepsilon, \tag{20}$$

where δ_j is the j-th element of the vector $\beta' \left(1 - \mathbf{B}\right)^{-1}$, which is, according to (16), sector j's Domar weight, $\frac{P_j Q_j}{PY}$.

2.5 How may imports of intermediate inputs matter?

We now describe how incorporating imported intermediate inputs in the production function may change the theoretical results. Consider the production function of a firm with productivity z in sector j as

$$q_j\left(z\right) = z\Lambda_j l_j\left(z\right)^{\lambda_j} \prod_{k=1}^{K} m_{jk}\left(z\right)^{\gamma_{jk}} m_{jk}^*\left(z\right)^{\gamma_{jk}^*},$$

[9] $\log \varepsilon = \sum_j \log\left(\beta_j / \delta_j\right)^{\beta_j}$

where $\Lambda_j \equiv \lambda_j^{-\lambda_j} \prod_{k=1}^{K} (\gamma_{jk})^{-\gamma_{jk}} \left(\gamma_{jk}^*\right)^{-\gamma_{jk}^*}$ and $m_{jk}^*(z)$ stands for intermediate inputs of sector k imported by the firm from abroad. Constant returns to scale require that $\sum_k \left(\gamma_{jk} + \gamma_{jk}^*\right) = 1 - \lambda_j$. Solving the model from the beginning as described in the previous section, along with a trade-balance constraint, yields an open-economy version of (19) as

$$\log Y = \beta' (1 - \mathbf{B})^{-1} \mathbf{tfp} + \beta' (1 - \mathbf{B})^{-1} (\lambda \log L + \gamma^* \log Y) + \log \varepsilon \tag{21}$$

Notice that now $\gamma^* \log Y$ appears on the right hand side. The variable γ^* is a vector of imported inputs used by each sector, with element j equal to $\sum_{k=1}^{K} \gamma_{jk}^*$. The new term $\gamma^* \log Y$ is added because a higher level of output requires the use of more domestic and imported intermediates.

Solving (21) yields a country's GDP as

$$Y = \left(\prod\nolimits_j TFP_j^{\delta_j^{adj}} \right) L\varepsilon, \tag{22}$$

where δ_j^{adj} is the j-th element of the vector $\frac{\beta'(1-\mathbf{B})^{-1}}{1-\beta'(1-\mathbf{B})^{-1}\gamma^*}$, equal to

$$\delta_j^{adj} = \frac{\delta_j}{1 - \sum_j^J \delta_j \gamma_j^*}. \tag{23}$$

The conventional wisdom suggests that increased imports in intermediate inputs should create "leakages", which weaken the domestic IO linkages. Instead, the effect of imports on an economy's IO multiplier is actually more subtle. Sector j's Domar weight, δ_j, depends on both domestic and foreign demand (exports). When γ_j^* increases, all else equal in the domestic economy, δ_j will decrease. Since the Domar weights are on average higher in the more upstream sectors, if γ_j^* increases more in the upstream sectors with those in the downstream sectors decreasing, the IO multiplier, which is equal to the sum of δ_j, will decline. Intuitively, when upstream sectors expand relative to other sectors due to larger net exports, the distortions in those sectors will have a larger impact on aggregate TFP. Empirically, whether and how much the multiplier increases will depend on the sectoral pattern of imports, or more specifically the correlation between δ_j and γ_j^*. When the Domar weights and the share of imported intermediate inputs γ_j^* are positively correlated across sectors, the multiplier can increase with the aggregate import share of the economy.

3 Quantitative Assessment

We can estimate time-varying TFP in each sector j, using the Hsieh-Klenow (2009) formula:

$$TFP_j = \left[\sum_{i=1}^{N_j} \left(z_i \frac{\overline{TFPR}_j}{TFPR_i} \right)^{\sigma_j - 1} \right]^{\frac{1}{\sigma_j - 1}}, \tag{24}$$

where z_i is firm i's TFP, $TFPR_i$ is its revenue TFP, and $\overline{TFPR}_j$ is the weighted average of firms' TFPR in the sector, with weights equal to each firm's sales (not value added).

To compute $TFPR_i$, we need to first compute marginal revenue product for each factor (labor, capital, and intermediate inputs) according to the extended production function for sector j that considers capital as another primary factor of production:

$$q_{ij} = z_i \Lambda'_j \left(l_i^{\alpha_j} k_i^{1-\alpha_j} \right)^{\lambda_j} m_i^{1-\lambda_j}.$$

where $m_i = \prod_{k=1}^{K} m_{ik}^{\gamma_k}$ and $\Lambda'_j \equiv \lambda_j^{-\lambda_j} \alpha_j^{-\lambda_j \alpha_j} (1-\alpha_j)^{-\lambda_j(1-\alpha_j)} \left(\prod_{k=1}^{K} \gamma_{jk}^{\gamma_{jk}} \right)^{-(1-\lambda_j)}$. Firm i's TFPR can be expressed exactly as

$$TFPR_i = \frac{\left[(1+\tau_{mi}) P_{mj} \right]^{1-\lambda_j} \left[(1+\tau_{li}) w \right]^{\lambda_j \alpha_j} \left[(1+\tau_{ki}) r \right]^{\lambda_j (1-\alpha_j)}}{\eta_j (1-\tau_{Yi})}, \tag{25}$$

with each of the wedges measured with firm-level data as follows:

$$\begin{aligned} \frac{1+\tau_{ki}}{1+\tau_{li}} &= \frac{1-\alpha_j}{\alpha_j} \frac{w l_i}{r k_i}; \\ \frac{1+\tau_{mi}}{1+\tau_{li}} &= \frac{1-\lambda_j}{\lambda_j \alpha_j} \frac{w l_i}{W_j m_i}; \quad \text{(overall)} \\ \frac{1-\tau_{Yi}}{1+\tau_{li}} &= \frac{1}{\eta_j} \frac{1}{\lambda_j \alpha_j} \frac{w l_i}{R_i}. \end{aligned} \tag{26}$$

The economy's aggregate (manufacturing) TFP loss, relative to the efficient level, can be computed as

$$\frac{TFP^e}{TFP} - 1 = \prod_{j=1}^{J} \left(\frac{TFP_j^e}{TFP_j} \right)^{\delta_j} - 1, \tag{27}$$

according to Section 2.4. Recall that δ_j is the Domar weight of sector j rather than its share in total consumption. Thus, the sum of δ_j across all sectors will be larger than 1, implying that the TFP loss associated with resource misallocation documented in the literature could be biased downward.

[Table 1 here]

In computing firms' z_i (physical TFP) and TFPR, we impose the parameter values described in Table 1 and compute firms' wedges, revenue TFP, and physical TFP according to the following steps.

1. We compute $1+\tau_{li}$, $1+\tau_{ki}$, $1+\tau_{mi}$, and $1-\tau_{Yi}$ based on (26), using firm-level data on wage bills, costs of capital, costs of materials, and total sales.

2. We then compute $TFPR_i$ using the computed firm-level wedges, according to (25).

Finally, we compute firm i's TFP, based on its iso-elastic demand curve, as

$$z_i = \kappa_j \frac{R_i^{\frac{\sigma}{\sigma-1}}}{\left(l_i^{\alpha_j} k_i^{1-\alpha_j}\right)^{\lambda_j} m_{ijk}^{1-\lambda}},$$

where κ_j is a constant, independent of wedges. It will drop out when we compute the ratio of distorted TFP to efficient TFP.

With all the components for computing TFP_j based on (24) available, we can compute the potential welfare gain if the variances of all wedges (τ_{ki}, τ_{mi}, and τ_{yi}) are set to zero.

4 Data

We will use micro data from both China and India to examine predictions at the firm, industry, and aggregate level.

4.1 China

For China, we use two data sets: manufacturing firm panel data from the country's National Bureau of Statistics (NBS) surveys and customs transactions-level trade data. The manufacturing firm data cover all state-owned firms and all non-state-owned firms with sales above 5 million yuan over the 1998-2007 period (about 0.6 million USD during the sample period).

This data set contains detailed balance-sheet information, such as ownership, four-digit industry code (484 categories), the values of output, value added, exports, employment, fixed asset (both original and current values), and intermediate inputs. The firms in the sample account for 57% of total industrial value added in 1998 and 94% in 2007. Moreover, we delete observations with missing values for key variables and those that fail to satisfy some basic error checks.

The cleaned data set provides an unbalanced panel of firms that increases in coverage from 148,685 firms in 1998 to 313,048 in 2007. We use unique numerical IDs to link firms in the sample over time. Firms occasionally receive a new ID as a result of restructuring, merger, or acquisition. Where possible, we aim to track firms as their boundaries or ownership structures change, using information on the firm's name, industry, address, etc., to link them. These other matches are still important as one-sixth of all firms that are observed for more than one year experience a change in their official ID over the period of analysis.

We also use transaction-level trade data from China Customs that cover all transactions of Chinese exporters and importers over 2000-2007. The trade data provide information on import and export values, quantities, and prices (unit values) between China and over 200 destination countries at the HS 6-digit level for each trading firm, by ownership of enterprise, and customs regime (ordinary trade and processing trade).

4.2 India

For India, the main data set is the annual plant-level data from the Annual Survey of Industries (ASI) conducted by the Indian Ministry of Statistics (from 1998-2010). We use the annual data from 1999-2000 to 2007-2008. A few caveats are in order. The sample consists of registered plants employing 10 or more workers using power, or 20 or more workers without power. Plants with over 100 workers are surveyed every year, while all other plants with fewer than 100 workers were randomly sampled. The ASI data provide information on the basic balance-sheet variables, such as the values of output, materials and fuels, capital (measured by the book value of fixed assets), employment, and wages (for both blue-collar workers and all other workers). An advantage of using the ASI data is that each firm in its surveys was asked to report information about quantities and values of different types of intermediates inputs, imported intermediates, and output at the 5-digit ASIS product category level (5204 categories) (See Boehm, Dhingra, Morrow, 2016 for details). Both ASI and NSS data are collected over each fiscal year (April 1 through March 30).

Detailed balance-sheet information, such as ownership, four-digit industry code, the values of output, imports, capital (measured by the book value of fixed assets), employment, and wages (for both blue-collar workers and all other workers), and intermediate inputs are available. Importantly, ASI data provide detailed information about quantities and values of different types of intermediates inputs each surveyed firm imported, produced, and purchased at the 5-digit product category level (5204 categories). We use these information to construct the domestic sourcing measures at the firm level, as well as domestic input prices.

5 Quantitative Results

5.1 A Comparison with Hsieh-Klenow (2009) Estimates of TFP Gains

Figure 1 shows the value of potential TFP gains by hypothetically removing distortions, computed using Indian plant-level data based on (27). There are four curves. The upper two lines represent the gains with all four types of distortions (labor, capital, material, and output wedges) removed. The lower two reveal the gains with all but capital distortions removed. The solid lines show the potential TFP gains based on the gross-output approach with full IO linkages. The dashed lines report the results of the counterfactual exercises, using the value-added approach adopted by HK.

Specifically, our gross-output approach with full IO linkages reveals that hypothetically removing all distortions raises India's manufacturing TFP by about 86% in 2000. The potential gain was increasing over time to 90% in 2007, suggesting that the costs of policy distortions could be rising over time. Replicating the quantitative exercises using the value-added approach, which considers only capital and labor as the primary factors of production, shows that the potential gains were larger – 100% in 2000. Interestingly, the value-added approach suggests that the gains were *diminishing* over time, to only 80% in 2007. We will revisit the different time trends and levels of the estimated TFP gains for India below. For now, through the lens of our model, we can postulate

that the lower levels of the implied TFP gains with IO linkages considered suggest that the distortions on intermediate inputs for India are sufficiently smaller than those on primary factors of production. The different trends, on the other hand, imply increasing distortions on intermediate inputs and/ or the IO multiplier over time.

Another way to show that the two approaches deliver different quantitative results is to estimate the potential TFP gains by removing capital distortions. The lines in the lower part of Figure 1 show the potential gains by removing all but capital distortions. In other words, the associated TFP gains for capital distortions would be the vertical difference between the corresponding upper lines (the solid one for our approach and the dashed one for the value added one). The vertical distance between the two solid lines is much closer for all years, compare to that between the two dashed lines. What it means is that using the value-added approach suggests much larger gains by removing capital misallocation than our approach for India.

[Figure 1 here]

In Figure 2, we show the corresponding results of the same sets of counterfactual exercises using Chinese data. Based on our approach, removing all distortions in the manufacturing sector is estimated to potentially raise the sector's TFP by about 129% in 2000, higher than those for India.[10] Different from India, the potential gain decreased gradually over time to 119% in 2007, suggesting that resources could have become more efficiently allocated across manufacturing firms in China. Estimated TFP gains using the value-added approach suggest an opposite trend– the potential gain of removing all distortions was 130% in 2000, and increased to 140% in 2007. The divergence of the estimated TFP gains between the two approaches is observed since 2002, the first year after China joined the WTO. Again, the changing multiplier, which depends on the cost shares of imported materials across sectors, could be a reason for the divergence. Unlike India, there is no systematic difference in the estimated TFP gains by removing capital distortions between the two approaches for China. Tables 2 and 3 report the estimates of the two sets of counterfactual exercises based on the two approaches for India and China, respectively.

[Figure 2 here]
[Table 2 here]
[Table 3 here]

It is surprising that the TFP losses due to resource misallocation are not larger after considering IO linkages in both economies. Could it be because the IO multiplier is not as large as we expected? Table 4 reports the size of the IO multiplier (i.e., the sum of the Domar weights across industries) for both countries across time. Our model shows that the IO multiplier is $\frac{\beta'(1-\mathbf{B})^{-1}\mathbf{1}}{1-\beta'(1-\mathbf{B})^{-1}\gamma^*}$. According to our model, we can simply compute the IO multiplier using the sum of the Domar weights across

[10] The comparison across countries should be interpreted with the usual cavaets (different measurement errors, sample selections, and so on). Moreover, the Indian ASI 4-digit industry categories are actually more aggregated than the Chinese NBS categories. Imposing the same elasticity of substitution σ on both samples may result in a higher level of TFP losses for China.

industries using our micro data, without using any IO table.[11] The average value of the Chinese manufacturing sector is about 3.57, while that for India is about 1.79. The Chinese IO multiplier appears to be decreasing slightly over time, with that of India fluctuating around an upward trend.

[Table 4 here]

Given the large IO multiplier, especially for China, the only possibility that the aggregate TFP losses due to resource misallocation are lower under the gross-output approach should imply smaller estimated TFP losses at the sector level based on that approach. As we proved in Section 2.3, a significantly lower estimated TFP loss at the sector level could be due to a small $\frac{v_m}{v_l}$ (i.e., the within-sector dispersion in marginal revenue product of intermediate inputs across firms is much smaller than those of labor and capital in both countries.) This possibility is what we will examine next.

5.2 Sectoral TFP Gains

To verify this conjecture, we plot our estimated TFP loss at the sector level against those based on the value-added approach to compute TFP and TFPR at different levels of aggregation. As shown in Figures 3 and 4, while for both countries there is a strongly positive correlation between our estimates and those based on the value-added approach across 4-digit sectors, our numbers are consistently much lower. Specifically, for India, the mean sectoral TFP loss due to resource misallocation is 0.39 based on the gross-output approach, compared to 0.93 based on the value-added approach. For China, these numbers are 0.63 and 1.33, respectively. In other words, the estimated TFP loss is on average an order of magnitude larger using the value-added approach, without considering distortions on intermediate input uses. These findings confirm our conjecture in the previous section and the theoretical prediction that estimated TFP losses can be substantially different between the value-added and gross-output approaches.

[Figure 3 here]
[Figure 4 here]

Furthermore, as expected, the only reason why the estimated sectoral TFP loss associated with resource misallocation is much lower using the gross-output approach is because the dispersion of the marginal revenue product of intermediate inputs across firms is substantially smaller than those of the primary factors of production. In order words, firms face highly distorting policies on employment of labor and capital, but not domestic outsourcing. Figures 5 and 6 illustrate this point for both India and China, respectively, by showing a significant lower dispersion of (log) marginal revenue product of intermediate inputs than that of (log) marginal revenue product of capital for 2005. Table 5 shows the evolution of the standard deviations of these two marginal products for both countries over time. The average standard deviation of (log) marginal revenue product of

[11] Using data from World Input-Output Tables over the same sample period, we find smaller IO multipliers for both countries. The one for China is still much larger than that for India..

materials across the sample years is 0.611 for India, compared to 1.305 for that of capital. Across time, the standard deviation of (log) marginal revenue product of intermediate inputs increases from 0.592 in 2000 to 0.641 in 2007, suggesting that the distortions on domestic sourcing across firms could be increasing over time. For China, the average standard deviation of (log) marginal revenue product of intermediate inputs and capital are 0.282 and 1.132, respectively. Different from India, the distortions on domestic sourcing across firms appear to be decreasing over time.

[Figure 5 here]
[Figure 6 here]
[Table 5 here]

Finally, we report in Tables 6 and 7 the top and bottom 10 industries for India and China, respectively, in terms of the computed Domar weights and the ratio between the computed Domar weights to the final-good consumption share. The latter value can be interpreted as a measure of influence of an industry on other industries. The ratio will be equal to 1 if all outputs of the industry are directly consumed as final goods. In India, the top 3 industries in terms of influence in 2005 are Grain mill products (5.866), Saw milling and planing of wood (5.487), and Processing and preserving of fish and fish products (5.172). These industries tend to produce goods that will be further processed in the downstream round-about supply chains. The 3 industries with the lowest influence factor are glass and glass products (1.234), musical instruments (1.183), and cement, lime and plaster (1.159). In China, the top 3 industries in 2005 in terms of influence are Acrylic Fibres (7.699), Apparel 0.118 Containers (5.441), Steel Rolling 0.096 Synthetic Fibres (5.415). The 3 industries with the lowest influence factor are Other production-use goods (2.117), Aircraft and spacecraft (1.908), Cigarettes (1.754), which tend to be directly consumed by consumers.

[Table 6 here]
[Table 7 here]

6 Conclusion

This paper develops a quantifiable general-equilibrium model to analyze domestic and global sourcing decisions of heterogeneous firms, which are affected by policy distortions on the one hand but amplify their macroeconomic costs through input-output (IO) linkages on the other. Solving the general equilibrium model with industry linkages reveals that an economy's aggregate TFP loss due to distortions can be summarized by the geometric mean of sector-level TFP losses, with weights equal to the sectors' Domar weights. Using Chinese firm-level and Indian plant-level data, we find that the estimated aggregate TFP losses due to resource misallocation with IO linkages are lower and exhibit different trends, relative to the value-added approach adopted by Hsieh and Klenow (2009). The surprising results are due to the fact that estimated sectoral TFP losses are on average an order of magnitude smaller than estimates based on the value-added approach, due to a much

smaller dispersion in the marginal revenue products of intermediate inputs across firms, compared to those of labor and capital, in both countries.

The framework developed in this paper can be extended in several directions. First, it can be used to examine the spatial distribution of economic activities and how they are related to the underlying frictions on trade between firms. For instance, we can quantitatively assess the productivity effects of the national governments' infrastructure investments, which substantially lower transaction costs between firms and mobility of factors.[12] Second, our model can be extended to study the structure, evolution, and related macroeconomic effects of the changing IO linkages across countries. Such analysis would require a deviation from the assumption of Cobb-Douglas functional form of production. Third, measurement errors can be isolated from true distortions that are both behind the dispersion of marginal revenue products of different factors, using a methodology proposed by Bils, Klenow and Ruane (2017). All these extensions will be explored in future research.

[12] In China, the dramatic expansion of the national highway and high-speed rail systems since 2000 have been shown to contribute to the increase in the GDP of the counties that were connected to the systems (Faber, 2014); whereas in India, the construction of the Golden Quadrilateral highway project has been shown to have a significantly positive effect on manufacturing firms' entry and productivity (Ghani, Goswami, and Kerr, 2016).

References

[1] Acemoglu D., P. Antràs, and E. Helpman (2007) "Contracts and Technology Adoption." *American Economic Review* 97 (3): 916-43.

[2] Acemoglu, Daron, Vasco Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi (2012) "The Network Origins of Aggregate Fluctuations." *Econometrica*.

[3] Antràs, P., T.C. Fort, and F. Tintelnot (2016) "The Margins of Global Sourcing: Theory and Evidence from U.S. Firms", Harvard University manuscript.

[4] Asker, J., A. Collard-Wexler, and J. De Loecker (2014) "Dynamic Inputs and Resource (Mis)Allocation", *Journal of Political Economy*.

[5] Baqaee, David (2016) "Cascading Failures in Production Networks." LSE Working Paper.

[6] Banerjee, A. and E. Duflo (2005) "Growth Theory Through the Lens of Development Economics." *Handbook of Economic Growth*.

[7] Bernard, A., A. Moxnes and Y. Saito (2015) "Production Networks, Geography and Firm Performance." Dartmouth University Working Paper.

[8] Bils, M., P. Klenow and C. Ruane (2017) "Misallocation or Mismeasurement?", Stanford University Working Paper.

[9] Blaum, J., M. Peters and C. Lelarge (2016) "The Gains from Input Trade in Firm-Based Models of Importing", Yale University Working Paper.

[10] Boehm, J. (2015). The Impact of Contract Enforcement Costs on Outsourcing and Aggregate Productivity. Science Po Working Paper.

[11] Boehm, J., S. Dhingra, and J. Morrow (2016) "Swimming Upstream: Input-Output Linkages and the Direction of Product Adoption", CEP Discussion Paper No 1407.

[12] Bollard, A., P. Klenow, and G. Sharma (2013) "India's Mysterious Manufacturing Miracle", *Review of Economic Dynamics* 16, 59-85.

[13] Brandt, L., J. Van Biesebroeck, Y. Zhang (2013) "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing" *Journal of Development Economics*, 97(2), pp. 339-351.

[14] Carvalho V. and X. Gabaix (2013) "The Great Diversification and Its Undoing" *American Economic Review*, vol. 103(5), p. 1697-1727.

[15] David, J. M., H. A., Hopenhayn and V. Venkateswaran (2016) "Information, misallocation and aggregate productivity" forthcoming *Quarterly Journal of Economics*.

[16] Faber, B. (2014) "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System" *Review of Economic Studies*, 2014, Volume 81(3), pages 1046-1070.

[17] Fieler, A.C., M. Eslava, D. Xu (2016) "Trade, Skills, and Quality Upgrading: A Theory with Evidence from Colombia" University of Pennsylvania Working Paper.

[18] Ghani, E., A.G. Goswami, and W.R. Kerr (2016) "The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing." *Economic Journal*, 126, 591: 317–357

[19] Hopenhayn, H. (2014) "Firms, Misallocation, and Aggregate Productivity: A Review" *Annual Review of Economics* Vol. 6: 735-770, August 2014.

[20] Hsieh, C. and P. Klenow (2009) "Misallocation and Manufacturing TFP in China and India" *Quarterly Journal of Economics*, 124, November 2009, 1403-1448.

[21] Hsieh, C. and P. Klenow (2014) "The Life Cycle of Plants in India and Mexico", *Quarterly Journal of Economics,* 129, August 2014, 1035-1084.

[22] Hunt, A., A. Collard-Wexler, and S.D. O'Connell (2016) ""

[23] Jia, P. (2008) "What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry." Econometrica, 76(6), 1263-1316.

[24] Jones, C. (2011) "Intermediate Goods and Weak Links in the Theory of Economic Development" *American Economic Journal: Macroeconomics*, April 2011.

[25] Jones, C. (2013) "Misallocation, Economic Growth, and Input-Output Economics" in *Advances in Economics and Econometrics*, Tenth World Congress, Volume II, Cambridge University Press, 2013.

[26] Kee H. and H. Tang (2016) "Domestic Value Added in Exports: Theory and Firm Evidence from China," *American Economic Review*, 106(6), pp. 1402-1436.

[27] Khandelwal, A., P. Schott and S.J. Wei. (2013) "Trade Liberalization and Embedded Institutional Reform: Evidence from Chinese Exporter", *American Economic Review,* 103(6):2169-95.

[28] Hachem, K. and M. Song (2016) "Liquidity Regulation and Unintended Financial Transformation in China", University of Chicago Working Paper.

[29] Kiyotaki, N., and Moore, J. (1997) "Credit cycles", *Journal of Political Economy*, 105(2), 211-248.

[30] Liu, E. (2017) "Industrial Policies and Economic Development, MIT Working Paper.

[31] Restuccia, D. and R. Rogerson (2008) "Policy distortions and aggregate productivity with heterogeneous establishments", *Review of Economic Dynamics*, Volume 11, Issue 4, 707-720.

[32] Restuccia, D. and R. Rogerson (2013) "Misallocation and Productivity", *Review of Economic Dynamics*, Vol 16 (1), pp. 1-10.

7 Appendix

7.1 Construction of Price Indices

We construct sector-level input prices using the Tornqvist method. For each industry j, let I_{kt} be the set of common sub-industries in two consecutive years, t and $t-1$. Denote sub-industry s's (or firm s's) output price index in year t and $t-1$ by P_{st} and $P_{s,t-1}$; and denote the share of sub-industry s's sales in industry $k's$ total sales in year t and $t-1$ by ω_{skt} and $\omega_{sk,t-1}$. Data on output price indices at the 4-digit sector level (based on India's NIC or China's CIC classifications). The industry-specific rate of output price inflation in year t is defined as

$$\triangle \ln \widetilde{P}_{kt} = \sum_{s \in I_{jt}} \frac{1}{2} \left(\omega_{skt} + \omega_{sk,t-1}\right) \left(\ln P_{kt} - \ln P_{k,t-1}\right).$$

Using this weighted average of inflation rates, the sector-specific output price level is defined as

$$\widetilde{P}_{kt} = \widetilde{P}_{k,t-1} \exp\left(\triangle \ln \widetilde{P}_{kt}\right),$$

with $\widetilde{P}_{kt}$ normalized to 1 in 2000.

Next, we compute each Indian firm's input price indices based on their reported input purchases on the market. For inputs that the firm produces in-house (which we do not see they purchase on the market), we do not observe the costs and need to assume that their prices are higher than those that are available on the market. In-house input prices are assumed to be 1 in the model but will need to set them equal to the maximum of all input prices that we observe in the data. Then, we use $\widetilde{P}_{kt}$ to construct the input price for each firm in India, and each narrow sector in China. We compute the weighted average of $\widetilde{P}_{jt}$, with weights equal to the coefficients from each country's IO table for the initial year of the sample. The goal is to compute the average domestic prices facing processing firms in industry j. Specifically, for each industry j, the weighted average of input prices is $\log P_{jt}^{D} = \sum_{k=1}^{K} \widehat{\gamma}_{jk} \log \widetilde{P}_{kt}$, where $\widehat{\gamma}_{jk}$ is the share of industry j goods in total material costs for production of a unit of industry k goods, according to IO tables. Notice that P_{jt}^{D} varies across time purely due to the variation in $\widetilde{P}_{kt}$, since $\widehat{\gamma}_{jk}$ is fixed throughout the sample. A similar price index can be computed for imported input prices.

Finally, we the average input price indices across firms for sector j to obtain $\int \prod_{\mathbf{K}(z)} P_k^{\gamma_{jk}} dz$. We can either simulate the first best average prices by assuming that all $\tau = 0$ or conduct a few counterfactual exercises to see how TFP loss would be if prices (and their underlying distortions) remain the same as the base year, and so forth.

7.2 Theoretical Expression of the TFP loss due to Resource Misallocation

Given that both z, $(1-\tau)$, and all $(1+\tau)'s$ are log-normal, evoking the law of large number implies that each term of the log of TFP_j (expressed in (11) can be expressed as

$$tfp_{j1} \equiv \log \int \left(\frac{(1-\tau_y)\, z}{(1+\tau_l)^{\lambda_j}} \right)^{\frac{\eta_j}{1-\eta_j}} = \frac{\eta_j}{1-\eta_j} \left[\mu_y + \mu_z - \lambda_j \mu_l + \frac{1}{2} \frac{\eta_j}{1-\eta_j} \left(\upsilon_y + \upsilon_z + \lambda_j^2 \upsilon_l \right) \right]$$

$$tfp_{j2} \equiv \log \int \left[\frac{(1-\tau_y)\, z^{\eta_j}}{(1+\tau_l)^{1-\eta_j(1-\lambda_j)}} \right]^{\frac{1}{1-\eta_j}} = \frac{1}{1-\eta_j} \left[\begin{array}{c} \mu_y + \eta_j \mu_z - \left[1 - \eta_j \left(1 - \lambda_j \right) \right] \mu_l + \\ \frac{1}{2} \frac{1}{1-\eta_j} \left(\upsilon_y + \eta_j^2 \upsilon_z + \left[1 - \eta_j \left(1 - \lambda_j \right) \right]^2 \upsilon_l \right) \end{array} \right]$$

$$tfp_{j3} \equiv \log \int \left[\frac{(1-\tau_y)\, z^{\eta_j}}{(1+\tau_l)^{\eta_j \lambda_j}} \right]^{\frac{1}{1-\eta_j}} = \frac{1}{1-\eta_j} \left[\begin{array}{c} \mu_y + \eta_j \mu_z - \eta_j \lambda_j \mu_l + \\ \frac{1}{2} \frac{1}{1-\eta_j} \left(\upsilon_y + \eta_j^2 \upsilon_z + \left(\eta_j \lambda_j \right)^2 \upsilon_l \right) \end{array} \right]$$

Using these extended terms we can express $\log TFP_j$ as

$$\begin{aligned} \log TFP_j &= \frac{1}{\eta_j} tfp_{j1} - \lambda_j tfp_{j2} - \left(1 - \lambda_j \right) tfp_{j3} \\ &= \mu_z + \frac{\sigma_j - 1}{2} \upsilon_z - \frac{\sigma_j}{2} \upsilon_y - \frac{\lambda_j + \lambda_j^2 \left(\sigma_j - 1 \right)}{2} \upsilon_l, \end{aligned}$$

which is the same expression as (12).

Figure 1. Aggregate TFP Gains by Removing Distortions (India)

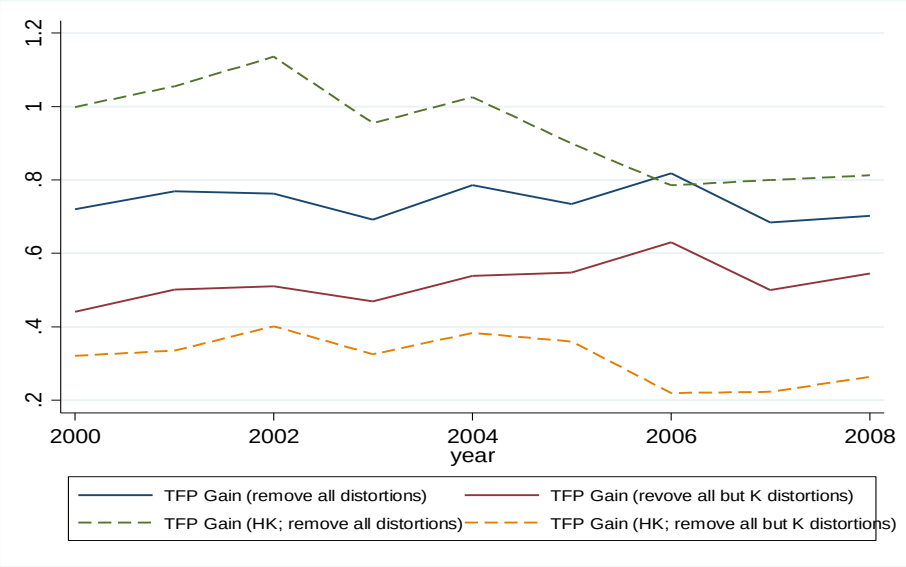


Figure 2. Aggregate TFP Gains by Removing Distortions (China)

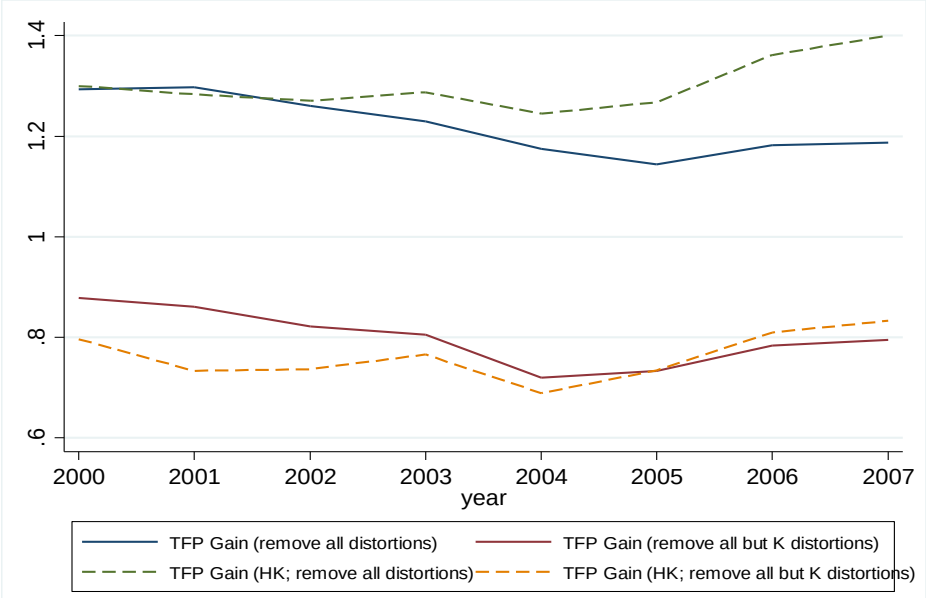


Fig 3. TFP losses due to distotions (our estimates versus Hsieh and Klenow's); India

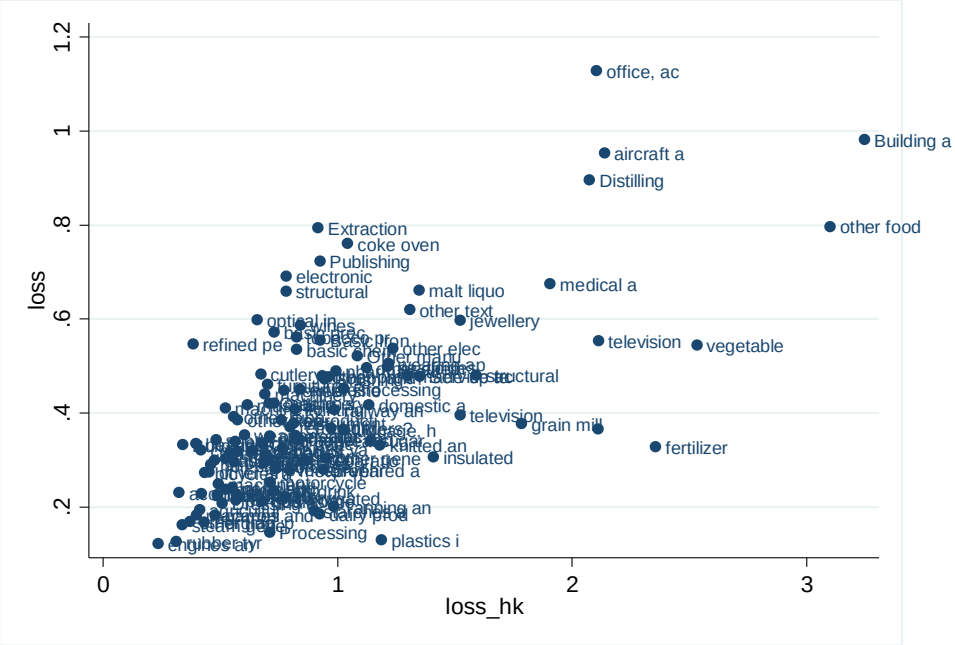


Fig 4. TFP losses due to distotions (our estimates versus Hsieh and Klenow's); China

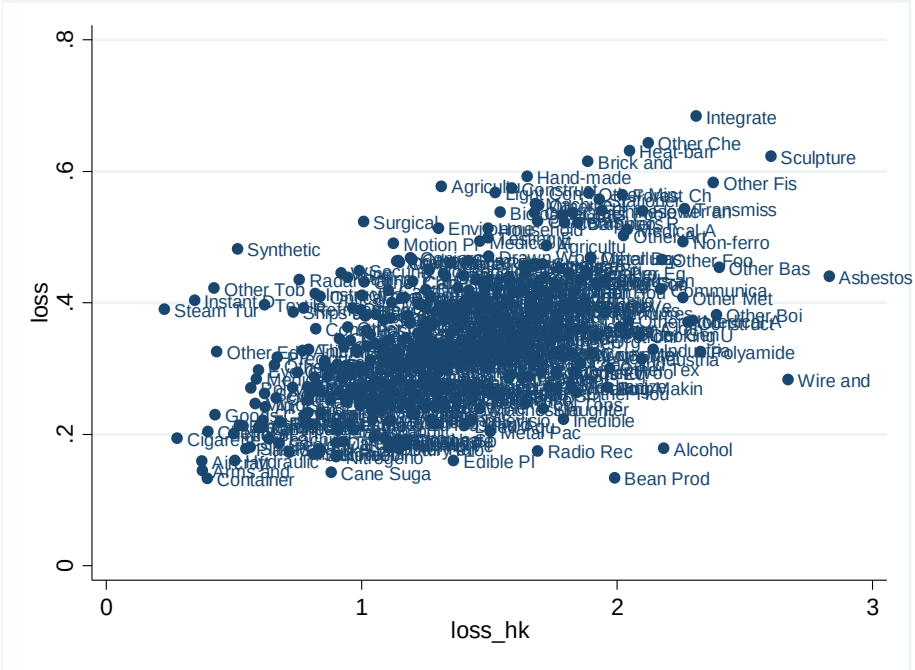


Figure 5. Dispersion of Marginal Revenue Products of Intermediate Inputs and Capital (India)

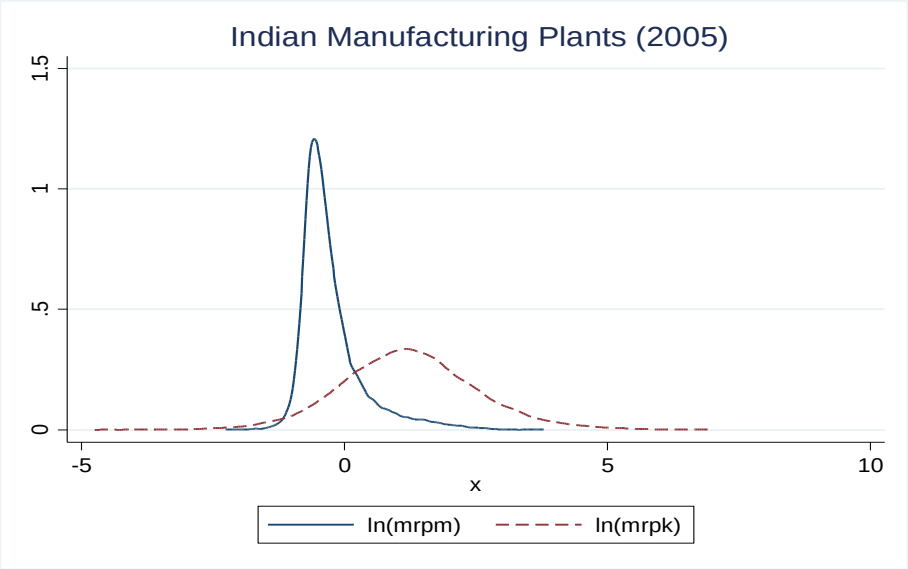


Figure 6. Dispersion of Marginal Revenue Products of Intermediate Inputs and Capital (China)

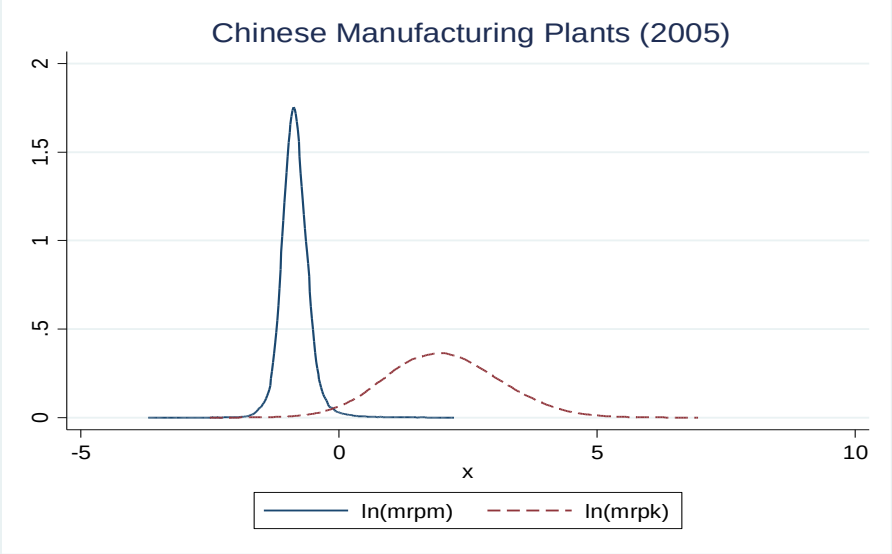


Table 1: Parameter Values

σ_j	3 (Hsieh and Klenow, 2009)
capital, labor, material shares	US 6-digit NAICS (average 1990-2005)
r	0.2 (China: Song and Wu, 2016)
	0.2 (India: Bils, Klenow, and Ruane, 2017)
w	1 (normalized)
Intermediate Input Deflators: P_{mj}	China: Brandt, Van Biesebroeck, and Zhang (2013)
	India: Allcott, Collard-Wexler, and O'Connell (2016)
Sectors' Domar weights: δ_j	Computed using each country's firm data

Table 2: Aggregate TFP Gain by Removing Distortions (India)

Scenerio	remove all distortions		remove all but capital distortions	
Model	Gross Output Approach with IO Linkages	VA Approach w/o IO Linkages (i.e., HK)	Gross Output Approach with IO Linkages	VA Approach w/o IO Linkages (i.e., HK)
2000	0.863	0.999	0.529	0.321
2001	0.941	1.055	0.614	0.335
2002	0.935	1.136	0.624	0.402
2003	0.847	0.955	0.575	0.326
2004	0.976	1.025	0.668	0.384
2005	0.930	0.899	0.694	0.360
2006	1.076	0.786	0.828	0.220
2007	0.901	0.800	0.659	0.223
Average	0.934	0.957	0.649	0.321

Source: Authors' calculation based on China's NBS and India's ASI data

Table 3: Aggregate TFP Gain by Removing Distortions (China)

Scenerio	remove all distortions		remove all but capital distortions	
Model	Gross Output Approach with IO Linkages	VA Approach w/o IO Linkages (i.e., HK)	Gross Output Approach with IO Linkages	VA Approach w/o IO Linkages (i.e., HK)
2000	1.293	1.300	0.878	0.796
2001	1.297	1.283	0.861	0.733
2002	1.261	1.271	0.822	0.737
2003	1.229	1.287	0.805	0.766
2004	1.175	1.245	0.720	0.689
2005	1.144	1.268	0.733	0.734
2006	1.182	1.362	0.784	0.810
2007	1.187	1.400	0.795	0.833
Average	1.221	1.302	0.800	0.762

Source: Authors' calculation based on China's NBS and India's ASI data

Table 4: Input-Output Multiplier

year	China	India
2000	3.608	1.765
2001	3.572	1.794
2002	3.560	1.751
2003	3.582	1.770
2004	3.612	1.828
2005	3.551	1.831
2006	3.532	1.814
2007	3.502	1.781
Average	3.565	1.792

Source: Authors' calculation based on China's NBS and India's ASI data

Table 5: Dispersions of log Marginal Revenue Products (MRP)

Year	India		China	
	sd(ln(mrpk))	sd(ln(mrpm))	sd(ln(mrpk))	sd(ln(mrpm))
2000	1.356	0.592	1.191	0.306
2001	1.325	0.592	1.187	0.288
2002	1.323	0.628	1.161	0.287
2003	1.298	0.612	1.133	0.278
2004	1.304	0.613	1.130	0.277
2005	1.292	0.603	1.091	0.275
2006	1.269	0.608	1.082	0.272
2007	1.270	0.641	1.079	0.273
Average	1.305	0.611	1.132	0.282

Source: Authors' calculation based on China's NBS and India's ASI data

Table 6: Domar Weights and Doma/ Consumption Share by Industry - India (2005)

Top 10 industries based on Domar weights	Domar Weight	Top 10 industries based on Domar/ Consumption Share	Domar/Beta
Refined petroleum products	0.256	Grain mill products	5.866
Motor vehicles	0.135	Saw milling and planing of wood	5.487
Basic Iron & Steel	0.111	Processing and preserving of fish and fish products	5.172
Preparation and spinning of textile fiber	0.100	Dairy product	4.520
Sugar	0.059	Bicycles and invalid carriages	3.575
Fertilizers and nitrogen compounds	0.059	prepared animal feeds	3.556
Cement, lime and plaster	0.058	Manufacturing of wooden containers	3.529
Plastics in primary forms and of synthetic rubber	0.057	Vegetable and animal oils and fats	3.289
Parts and accessories for motor vehicles and their engines	0.055	Dressing and dyeing of fur; articles of fur	3.225
Pharmaceuticals, medicinal chemicals and botanical products	0.055	Bodies (coach work) for motor vehicles; trailers and semi-trailers	3.098
Bottom 10 industries based on Domar weights		Bottom 10 industries based on Domar/ Consumption Share	
Builders' carpentry and joinery	0.000	non-structural non-refractory ceramic ware	1.294
Dressing and dyeing of fur; articles of fur	0.000	office, accounting and computing machinery	1.286
Saw milling and planing of wood	0.000	Building and repairing of ships	1.285
Wooden containers	0.000	structural non-refractory clay and ceramic products	1.268
Other products of wood, articles of cork, straw & plaiting mat	0.000	wines	1.256
Service activities related to printing	0.000	malt liquors and malt	1.242
Macaroni, noodles, conscious and similar farinaceous products	0.000	tobacco products	1.235
Aircraft and spacecraft	0.000	glass and glass products	1.234
Musical instruments	0.000	musical instruments	1.183
Building and repairing of pleasure and sporting boats	0.000	cement, lime and plaster	1.159

Source: Authors' calculation and India's Annual Survey of Industries (ASI) Data.

Table 7: Domar Weights and Doma/ Consumption Share by Industry - China (2005)

Top 10 industries based on Domar weights	Domar Weight	Top 10 industries based on Domar/ Consumption Share	Domar/Beta
Cotton Spinning	0.132	Acrylic Fibres	7.699
Apparel	0.118	Containers	5.441
Steel Rolling	0.096	Synthetic Fibres	5.415
Automobile Parts	0.091	Mining Vehicles	5.413
Cement	0.074	Tin Smelting	5.350
Wire and Cable	0.056	Television Sets	5.256
Electronic Components	0.053	Bicycles	5.230
Light Non-ferrous Metal Mangling	0.049	Copiers	5.184
Machine-made Paper and Paperboard	0.046	Goods Carriage	4.956
Feed processing	0.043	Refrigerators	4.896
Bottom 10 industries based on Domar weights		**Bottom 10 industries based on Domar/ Consumption Share**	
Hand-made Paper	0.000	Beer	2.386
Animal Husbandry Machinery	0.000	Potash Fertilizer	2.380
Lumber Equipment	0.000	White Spirits	2.339
Stationery Products	0.000	Wine	2.304
Tram manufacturing	0.000	Diving and underwater rescue equipment	2.274
Artificial Crude Oil Production	0.000	Biological Products	2.236
Nuclear and Nuclear Radiation Measurement Instruments	0.000	Chinese Medicines and Phamaceuticals	2.171
Other Production-Use Goods	0.000	Other Production-Use Goods	2.117
Diving and underwater rescue equipment	0.000	Other Aircraft and Spacecraft	1.908
Other Aquatic Products	0.000	Cigarettes	1.754

Source: Authors' calculation and China's National Bureau of Statistics Data.