Income Differences, Productivity and Input-Output Networks *

Harald Fadinger† Christian Ghiglino ‡ Mariya Teteryatnikova§

May 2018

Abstract

We study the importance of input-output (IO) linkages and sectoral productivity (TFP) levels in determining cross-country income differences. Using data on IO tables and sectoral TFPs, we find important differences in IO structure and its interaction with TFP levels across countries: while highly connected sectors are more productive than the typical sector in poor countries, the opposite is true in rich ones. To quantitatively assess the role of IO linkages in cross-country income differences, we use tools from network theory to build a multi-sector general equilibrium model. Aggregate income is approximated by a simple function of the first and second moments of the joint distribution capturing interactions of IO linkages and sectoral TFPs. We then structurally estimate country-specific parameters of this distribution and simulate cross-country income differences. Our main finding is that incorporating IO linkages into a model with sectoral TFP differences significantly improves our ability to predict cross-country income variation.

KEY WORDS: input-output structure, networks, productivity, cross-country income differences, development accounting

JEL CLASSIFICATION: O11, O14, O47, C67, D85

*We thank Jean-Noel Barrot, Johannes Boehm, Susanto Basu, Antonio Cabrales, Vasco Carvalho, Antonio Ciccone, Alejandro Cuñat, Manuel García Santana and seminar participants at the Universities of Cambridge, Mannheim, York and Vienna, at the First Worldbank-CEPR Conference on Global Value Chains, Trade and Development, the NBER Summer Institute (Macro-productivity workshop), the SED meeting, the SAET conference, the EEA annual congress and the European Commission ECFIN seminar for useful comments and suggestions. We also thank Susana Parraga Rodriguez for excellent research assistance. Previous versions of this paper were circulated under the titles "Productivity, Networks, and Input-Output Structure" and "Income Differences and Input-Output Structure".

†University of Mannheim and CEPR. Email: harald.fadinger@uni-mannheim.de.

‡University of Essex and GSEM Geneva. Email: cghig@essex.ac.uk.

§National Research University Higher School of Economics, Moscow. Email: mteteryatnikova@hse.ru

1 Introduction

Cross-country differences in income per capita are largely due to differences in aggregate total factor productivity (TFP), which account for at least 50% of income variation.[1] These cross-country differences in aggregate TFP stem from two sources: those due to differences in the technologies used and the efficiency with which they are operated and those due to differences in the so-called input-output (IO) structure of the economies that determines how sectoral TFPs add up at the country level. The role of the first source of aggregate TFP and income differences has been the focus of a large literature on endogenous growth and technology adoption,[2] while the importance of the second has been emphasized by a literature in development economics initiated by Hirschman (1958), with more recent contributions provided by Ciccone (2002) and Jones (2011 a,b). In this paper we contribute to the second literature by establishing systematic and empirically relevant cross-country differences in (i) IO structure and (ii) its interaction with sectoral TFP levels. We then show that these elements are of first-order importance for explaining cross-country income differences.

Countries' IO structure, by means of the linkages between sectors, determines each sector's importance or "weight" in aggregate TFP. It can be effectively summarized using the distribution of sectoral IO multipliers. The (first-order) IO multiplier of a sector depends on the (i) number of sectors to which the sector supplies and (ii) the intensity with which the output of the sector is used as an input by other sectors.[3] It measures by how much aggregate income changes if productivity of a given sector changes by one percent. Thus, TFP levels in sectors with high multipliers have a larger impact on aggregate income compared to sectors with low multipliers.

To quantitatively assess the role of IO linkages and sectoral TFP levels for cross-country income differences, we first build a neoclassical multi-sector model that admits a closed-form solution for aggregate income as a function of the first and second moments of the joint distribution of sectoral IO multipliers and TFP levels.[4] Higher average IO multipliers, higher average sectoral TFP levels and a positive correlation between sectoral IO multipliers and TFP levels all have a positive effect on income per capita.

We then combine data from the World Input-Output Database (Timmer, 2012) and the Global Trade Analysis project (GTAP Version 6) to construct a unique dataset of IO tables and sectoral

[1] See, e.g., Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), Caselli (2005).

[2] See, e.g., Romer (1990), Aghion and Howitt (1992), Comin and Hobijn (2004, 2010).

[3] The intensity of input use is measured by the IO coefficient, which states the cents spent on that input per dollar of output produced. There are also higher-order effects, which depend on the number and the IO coefficients of the sectors to which the sectors that use the initial sector's output as an input supply.

[4] In our baseline model, we take the IO structure as exogenous. Moreover, due to Cobb-Douglas technology sectoral TFP levels are independent of IO structure. In robustness checks we account for possible endogeneity of IO linkages by: (i) allowing for sector-country-specific tax wedges; (ii) introducing CES production functions, which makes IO linkages endogenous to sectoral TFPs.

TFP levels (relative to those of the U.S.) for a large cross section of countries.[5] We observe that the empirical distribution of sectoral multipliers has a fat right tail in all countries, so that the TFP levels of a few high-multiplier sectors can have a large impact on aggregate outcomes. This feature is more pronounced in developing countries than in rich economies. Moreover, in developing countries, sectoral IO multipliers and TFP levels are *positively* correlated, while they are *negatively* correlated in rich economies.

Given the theoretical model and the constructed dataset of IO tables and sectoral TFPs, we then proceed to the empirical analysis. We estimate a set of country-specific model parameters from the joint empirical distribution of sectoral IO multipliers and productivities with Maximum Likelihood and plug the parameter estimates into the structural model to simulate the cross-country income distribution. Our main finding is that the model featuring cross-country differences in the joint distribution of sectoral IO multipliers and TFP levels fares far better in terms of predicting the actual cross-country income variation than restricted versions of the model that: either (i) abstract from linkages and just allow for cross-country differences in sectoral TFP levels; or (ii) abstract from cross-country TFP differences and just allow for differences in IO structure; or (iii) allow for both linkages and sectoral TFP differences but abstract from cross-country variation in the IO structure.

In fact, a simple multi-sector model without IO linkages predicts *too large* cross-country income differences compared to the data given the estimated differences in sector-specific TFP levels. Intuitively, the large sectoral TFP differences in the data are mitigated by countries' IO structures: low-productivity sectors tend to be poorly connected (have low multipliers) in developing countries and are thus not too harmful, while high-productivity sectors have large multipliers and thus boost their income.[6] By contrast, no such tendency exists in rich countries. Thus, if we measured aggregate productivity levels by just averaging sectoral TFPs without accounting for variation in linkages, income levels of developing countries would be significantly lower than they actually are.

Our statistical approach – that considers the moments of the distributions of multipliers and productivities instead of the actual values – has a number of crucial advantages compared to feeding the full set of IO matrices and sectoral productivity levels of each country into a large-scale multi-sector model. First, it allows obtaining analytical results for how sectoral IO structure and TFP levels interact in their impact on the cross-country income distribution. Second, since the whole economic structure can be summarized by a small set of parameters, we can estimate these for each country and project them on per capita GDP. This enables us to obtain income predictions for the full set of countries in the Penn World Tables (155 countries), rather than being constrained to the

[5] Data on sectoral TFPs are available for 36 countries and data on IO tables for up to 65 countries. The full list of countries can be found in the Appendix.

[6] An important exception is agriculture, which, in low-income countries, has a high IO multiplier and a below-average productivity level.

36 countries for which we can actually observe both sectoral TFP levels and IO tables. In doing so, we can compare the model-predicted *world* income distribution with actual data. Finally, this approach enables us to carry out a number of simple counter-factuals by changing the parameters governing the joint distributions of multipliers and TFPs.[7]

The role of linkages and their interaction with sectoral TFPs for income differences is further evaluated by performing a number of counter-factuals. First, we impose the IO structure of the U.S. on all countries. We find that using the dense IO network of the U.S. would significantly reduce income of low- and middle-income countries. For a country at 40% of the U.S. income level (e.g., Mexico) per capita income would decline by around 20% and income reductions would amount to up to 60% for the world's poorest economies (e.g., Congo). Intuitively, imposing the dense IO structure of the U.S. on poor countries makes their typical, low-productivity, sector much more connected to the rest of the economy and thus increases its negative impact on aggregate income. To some extent the sparseness of the IO network in low-income countries is thus good news: in these countries policies that focus on increasing productivity in just a few crucial sectors can have a large effect on aggregate income, while this is not true in rich economies.

Second, we impose that sectoral IO multipliers and productivities are uncorrelated. This scenario would again hurt low-income countries, which would lose up to 10% of their per capita income, because they would no longer have the advantage of having above-average TFP levels in high-multiplier sectors. By contrast, high-income countries would benefit, since for them the correlation between multipliers and TFP levels would no longer be negative.

In our baseline model, differences in IO structure across countries are exogenously given. However, one may be concerned that observed IO linkages are affected by tax wedges. In an extension, we thus identify sector-country-specific tax wedges as deviations of sectoral intermediate input shares from their cross-country average value: a below-average intermediate input share in a given sector identifies a positive implicit tax wedge. We show that poor countries tax their high-multiplier sectors relatively more, while the opposite is the case in rich economies. We find that the distribution of IO multipliers and their correlation with TFP levels are not significantly affected by allowing for wedges. Moreover, introducing wedges does not improve the model's explanatory power in terms of predicting cross-country income levels much. Removing the correlation between wedges and multipliers would also have relatively modest effects. If low-income countries did not have above-average

[7] In the light of Hulten's (1978) results, one may be skeptical whether using a structural general equilibrium model and considering the statistical features of the IO matrices adds much compared to computing aggregate TFP as a weighted average of sectoral TFPs (where the adequate 'Domar' weights correspond to the shares of sectoral gross output in GDP). Absent distortions, Domar weights equal sectoral IO multipliers and summarize the direct and indirect effect of IO linkages. However, such a reduced-form approach does not allow to assess which features of the IO structure matter for aggregate outcomes or to compute counter-factual outcomes due to changes in IO structure, or productivities, as we do. Finally, as Basu and Fernald (2002) show, in the presence of sector-specific distortions (that we consider in an extension) the simple reduced-form connection between sectoral productivities and aggregate TFP breaks down.

tax rates in high-multiplier sectors, they would gain up to 10% of per capita income.[8]

In a further robustness check, we relax the assumption of a unit elasticity of substitution between intermediate inputs, so that IO linkages become endogenous to prices. We show that an elasticity of substitution between intermediate inputs different from unity is hard to reconcile with the data because – depending on whether intermediates are substitutes or complements – it implies that sectoral IO multipliers and TFP levels should either be positively or negatively correlated in *all* countries. Instead, we observe a positive correlation between these variables in poor economies and a negative one in rich countries. Moreover, we extend our baseline model to incorporate cross-country differences in final demand structure and imported intermediate inputs; we also differentiate between skilled and unskilled labor inputs. We find that our results are robust to all of these extensions.

1.1 Literature

We now turn to a discussion of the related literature.

Our work is related to the literature on development accounting, which aims at quantifying the importance of cross-country variation in factor endowments – such as physical, human or natural capital – relative to aggregate productivity differences in explaining disparities in income per capita across countries. This literature typically finds that both are roughly equally important in accounting for cross-country income differences.[9] The approach of development accounting is to specify an aggregate production function for value added (typically Cobb-Douglas) and to back out productivity differences as residual variation that reconciles the observed income differences with those predicted by the model given the observed variation in factor endowments. Thus, this aggregate production function abstracts from cross-country differences in the underlying IO structure and is exactly identified. We contribute to this literature by (i) showing how an aggregate production function for value added can be derived in the presence of IO linkages and (ii) providing an over-identification test for the model, since sectoral TFP estimates are obtained independently. Most importantly, we show that incorporating cross-country variation in IO structure is of first-order importance for explaining cross-country income differences.

The importance of linkages and IO multipliers for aggregate income differences has been highlighted by Fleming (1955), Hirschmann (1958), and, more recently, by Ciccone (2002) and Jones (2011 a,b). The last two authors emphasize that if the intermediate share in gross output is sizable, there exist large multiplier effects: small firm (or industry-level) productivity differences or distortions that lead to misallocation of resources across sectors or plants can add up to large aggregate effects. These authors make this point in a purely theoretical context. While our setup in

[8] In the Appendix we also study optimal taxation and the welfare gains from moving from the current tax wedges to an optimal tax system that keeps tax revenue constant and obtain a similar conclusion.

[9] See, e.g., Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), Caselli (2005), Hsieh and Klenow (2010).

principle allows for a mechanism whereby intermediate linkages amplify small sectoral productivity differences, we find that there is little empirical evidence for this channel: cross-country sectoral TFP differences estimated from the data are even larger than aggregate ones, and the sparse IO structure of low-income countries actually helps to mitigate the impact of very low productivity levels in some sectors on aggregate outcomes.

Our finding that sectoral productivity differences between rich and poor countries are larger than aggregate ones is instead similar to those of the literature on dual economies and sectoral productivity gaps in agriculture.[10] Also closely related to our work is a literature on structural transformation. It emphasizes sectoral productivity gaps and transitions from agriculture to manufacturing and services as a reason for cross-country income differences (see, e.g., Duarte and Restuccia, 2010 for a recent contribution). However, most this literature abstracts from the role of linkages between industries.

In terms of modeling approach, our paper adopts the framework of the multi-sector real business cycle model with IO linkages of Long and Plosser (1983); in addition we model the input-output structure as a network, quite similarly to the setup of Acemoglu, Carvalho and Ozdaglar (2012).[11] In contrast to these studies, which deal with the relationship between sectoral productivity shocks and economic fluctuations, we are interested in the question how sectoral TFP *levels* interact with the IO structure to determine aggregate income *levels* and we provide corresponding structural estimation results.

Other recent related contributions are Oberfield (2013) and Carvalho and Voigtländer (2014), who develop an abstract theory of endogenous input-output network formation, and Boehm (2015), who focuses on the role of contract enforcement on aggregate productivity differences in a quantitative structural model with IO linkages. Differently from these papers, we do not try to model the IO structure as arising endogenously and we take sectoral productivity differences as exogenous. Instead, we aim at understanding how given differences in IO structure and sectoral productivities translate into aggregate income differences.

The number of empirical studies investigating cross-country differences in IO structure is quite limited. In the most comprehensive study up to that date, Chenery, Robinson, and Syrquin (1986) find that the intermediate input share of manufacturing increases with industrialization and – consistent with our evidence – that input-output matrices become less sparse as countries industrialize. Most closely related to our paper is the contemporaneous work by Bartelme and Gorodnichenko (2015). They also collect data on IO tables for many countries and investigate the relationship be-

[10] See, e.g., Caselli (2005), Chanda and Dalgaard (2008), Restuccia, Yang, and Zhu (2008), Vollrath (2009), Gollin et al.(2014).

[11] Related to Acemoglu et al. (2012) empirical work by Barrot and Sauvagnat (2016) provides reduced-form evidence for the short-run propagation of exogenous firm-specific shocks in the production network of U.S. firms.

tween IO linkages and aggregate income.[12] In reduced-form regressions of *aggregate* IO multipliers on income per worker, they find a positive correlation between the two variables. Moreover, they investigate how distortions affect IO linkages and income levels. Differently from the present paper, they neither use data on sectoral productivities nor network theory to represent IO tables. As a consequence, they do not investigate how differences in the distribution of sectoral multipliers and their correlations with productivities impact on aggregate income, which is the focus of our work.

The outline of the paper is as follows. In the next section, we lay out our theoretical model and derive an expression for aggregate GDP in terms of the IO structure and sectoral TFP levels. In the following section, we describe our dataset and present some descriptive statistics. Subsequently, we turn to the structural estimation and model fit. We then present the counter-factual results and a number of robustness checks. The final section presents our conclusions.

2 Theoretical framework

2.1 Model

In this section we present our theoretical framework (based on Jones, 2011b) that will be used in the remainder of our analysis. Consider a static multi-sector economy. n competitive sectors each produce a distinct good that can be used either for final consumption or as an input for production. The technology of sector $i \in 1:n$ is Cobb-Douglas with constant returns to scale. Namely, the output of sector i, denoted by q_i, is

$$q_i = \Lambda_i \left(k_i^{\alpha} l_i^{1-\alpha}\right)^{1-\gamma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}}, \tag{1}$$

where Λ_i is the exogenous total factor productivity of sector i, k_i and l_i are the quantities of capital and labor used by sector i and d_{ji} is the quantity of good j used in production of good i (intermediate good produced by sector j).[13] The exponent $\gamma_{ji} \in [0,1)$ represents the share of good j in the production technology of firms in sector i, and $\gamma_i = \sum_{j=1}^{n} \gamma_{ji} \in (0,1)$ is the total share of intermediate goods in gross output of sector i. Parameters α, $1-\alpha \in (0,1)$ are the shares of capital and labor in the remainder of the inputs (value added).

Given the Cobb-Douglas technology in (1) and competitive factor markets, γ_{ji}s also correspond to the entries of the IO matrix, measuring the value of spending on input j per dollar of production of good i. We denote this IO matrix by $\mathbf{\Gamma}$. The entries of the j'th row of matrix $\mathbf{\Gamma}$ represent the values of spending on a given input j per dollar of production of each sector in the economy. By

[12] Grobovsek (2015) performs a development accounting exercise in a more aggregate structural model with two final and two intermediate sectors.

[13] In section 6 and in the Appendix we consider the case of an open economy, where each sector's production technology employs both domestic and imported intermediate goods.

contrast, the elements of the i'th column of matrix $\mathbf{\Gamma}$ are the values of spending on inputs from each sector in the economy per dollar of production of a given good i.[14]

Output of sector i can be used either for final consumption, y_i, or as an intermediate good:

$$y_i + \sum_{j=1}^{n} d_{ij} = q_i \qquad i = 1:n \tag{2}$$

Final consumption goods are aggregated into a single final good through another Cobb-Douglas production function:

$$Y = y_1^{\frac{1}{n}} \cdot ... \cdot y_n^{\frac{1}{n}}. \tag{3}$$

This aggregate final good is used as households' consumption, C, so that $Y = C$. Note that the symmetry in exponents of the final good production function implies symmetry in consumption demand for all goods. This assumption is useful as it allows us to focus on the interaction effects between the IO structure and sectors' productivities without confounding the role of linkages with the impact of final demand. It is, however, straightforward to introduce asymmetry in consumption demand by defining the vector of demand shares $\boldsymbol{\beta} = (\beta_1, .., \beta_n)$, where $\beta_i \neq \beta_j$ for $i \neq j$ and $\sum_{i=1}^{n} \beta_i = 1$. The corresponding final good production function is then $Y = y_1^{\beta_1} \cdot ... \cdot y_n^{\beta_n}$. This more general framework is analyzed in section 6, where we consider extensions of our benchmark model.

Finally, the total supply of capital and labor in this economy are assumed to be exogenous and fixed at the levels of K and 1, respectively:

$$\sum_{i=1}^{n} k_i = K, \tag{4}$$

$$\sum_{i=1}^{n} l_i = 1. \tag{5}$$

To complete the description of the model, we provide a formal definition of a competitive equilibrium.

Definition A competitive equilibrium is a collection of quantities q_i, k_i, l_i, y_i, d_{ij}, Y, C and prices p_i, p w, and r for $i \in 1:n$ such that

1. y_i solves the profit maximization problem of a representative firm in the perfectly competitive final good's market, taking $\{p_i\}, p$ as given.

2. $\{d_{ij}\}$, k_i, l_i solve the profit maximization problem of a representative firm in the perfectly competitive sector i for $i \in 1:n$, taking $\{p_i\}$ as given (Λ_i is exogenous).

3. Households' budget constraint determines C: $C = w + rK$.

[14] According to our notation, the sum of elements in the i'th column of matrix $\mathbf{\Gamma}$ is equal to γ_i, the total intermediate share of sector i.

4. Markets clear:

   (a) r clears the capital market: $\sum_{i=1}^{n} k_i = K$,

   (b) w clears the labor market: $\sum_{i=1}^{n} l_i = 1$,

   (c) p_i clears sector i's market: $y_i + \sum_{j=1}^{n} d_{ij} = q_i$,

   (d) p clears the final good's market: $Y = C$.

5. Production function for q_i is $q_i = \Lambda_i \left(k_i^{\alpha} l_i^{1-\alpha}\right)^{1-\gamma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}}$.

6. Production function for Y is $Y = y_1^{\frac{1}{n}} \cdot \ldots \cdot y_n^{\frac{1}{n}}$.

7. Numeraire: $p = 1$.

Note that households' consumption is simply determined by the budget constraint, so that there is no decision for the households to make. Moreover, total production of the aggregate final good, Y, which is equal to $\sum_{i=1}^{n} p_i y_i$, represents real GDP (total value added) per capita.

2.2 Equilibrium

The following proposition characterizes the equilibrium value of the logarithm of GDP per capita, which we later refer to equivalently as aggregate output or aggregate income or value added of the economy.

Proposition 1. *There exists a unique competitive equilibrium. In this equilibrium, the logarithm of GDP per capita, $y = \log(Y)$, is given by*

$$y = \sum_{i=1}^{n} \mu_i \lambda_i + \sum_{i=1}^{n} \sum_{j \ s.t. \ \gamma_{ji} \neq 0} \mu_i \gamma_{ji} \log \gamma_{ji} + \sum_{i=1}^{n} \mu_i (1-\gamma_i) \log(1-\gamma_i) - \log n + \alpha \log K, \tag{6}$$

where

$$\boldsymbol{\mu} = \{\mu_i\}_i = \frac{1}{n}[\boldsymbol{I} - \boldsymbol{\Gamma}]^{-1}\mathbf{1}, \qquad n \times 1 \text{ vector of multipliers}$$

$$\boldsymbol{\lambda} = \{\lambda_i\}_i = \{\log \Lambda_i\}_i, \qquad n \times 1 \text{ vector of sectoral log-productivity coefficients}$$

Proof. The proof of Proposition 1 is provided in the Appendix.

Thus, due to the Cobb-Douglas structure of our economy, aggregate per capita GDP can be represented as a log-linear function of (i) terms summarizing the aggregate impact of sectoral productivities via the IO structure and thus, representing aggregate productivity; (ii) terms summarizing the impact of the IO coefficients γ_{ji}s, and (iii) the capital stock per worker weighted by the capital share in GDP, α.

The proposition highlights two important facts. First, aggregate output is an increasing function of sectoral productivity levels. Second, and more importantly, the impact of each sector's productivity on aggregate output is proportional to the value of the sectoral IO multiplier μ_i. This means that the positive effect of higher sectoral productivity on aggregate output is stronger in sectors with larger multipliers.[15]

The vector of sectoral multipliers, in turn, is determined by the features of the IO matrix through the Leontief inverse, $[\boldsymbol{I}-\boldsymbol{\Gamma}]^{-1}$.[16] The interpretation and properties of this matrix as well as a simpler representation of the vector of multipliers are discussed in the next section.

2.3 Intersectoral network. Multipliers as sectors' centrality

The input-output matrix $\boldsymbol{\Gamma}$, where a typical element γ_{ji} captures the value of spending on input j per dollar of production of good i, can be equivalently represented by a directed weighted network on n nodes. Nodes of this network are sectors and directed links indicate the flow of intermediate goods between sectors. Specifically, the link from sector j to sector i with weight γ_{ji} is present if sector j is an input supplier to sector i.

For each sector in the network we define the *weighted in-* and *out-degree*. The weighted in-degree of a sector is the share of intermediate inputs in its production. It is equal to the sum of elements in the corresponding *column* of matrix $\boldsymbol{\Gamma}$; that is, $d_i^{in} = \gamma_i = \sum_{j=1}^{n} \gamma_{ji}$. The weighted out-degree of a sector is the share of its output in the input supply of the entire economy. It is equal to the sum of elements in the corresponding *row* of matrix $\boldsymbol{\Gamma}$; that is, $d_j^{out} = \sum_{i=1}^{n} \gamma_{ji}$.[17]

The interdependence of sectors' production technologies through the network of intersectoral trade helps to obtain some insights into the meaning of the Leontief inverse matrix $[\boldsymbol{I} - \boldsymbol{\Gamma}]^{-1}$ and the vector of sectoral multipliers $\boldsymbol{\mu}$.[18] A typical element l_{ji} of the Leontief inverse can be interpreted as the percentage increase in the output of sector i following a one-percent increase in productivity of sector j. This result takes into account all – direct and indirect – effects at work, such as for example, the effect of raising productivity in sector A that makes sector B more efficient and via this raises the output in sector C. Then multiplying the Leontief inverse matrix by the vector of weights $\frac{1}{n}\mathbf{1}$ adds up the effects of sector j on all the other sectors in the economy, weighting each by its share $\frac{1}{n}$ in GDP. Thus, a typical element of the resulting vector of IO multipliers reveals how a one-percent increase in productivity of sector j affects the overall value added in the economy.[19]

[15] The value of sectoral multipliers is positive due to a simple approximation result (7) in the next section.

[16] See Burress (1994).

[17] Note that if the weights of all links that are present in the network are identical, the in-degree of a given sector is proportional to the number of sectors that supply to it and its out-degree is proportional to the number of sectors to which it is a supplier.

[18] Observe that in this model the Leontief inverse matrix is well-defined since CRS technology of each sector implies that $\gamma_i < 1$ for any $i \in 1 : n$. According to the Frobenius theory of non-negative matrices, this means that the maximal eigenvalue of $\boldsymbol{\Gamma}$ is bounded above by 1, and this, in turn, implies the existence of $[\boldsymbol{I} - \boldsymbol{\Gamma}]^{-1}$.

[19] In particular, for a simple one-sector economy, the multiplier is given by $\frac{1}{1-\gamma}$, where γ is a share of the intermediate

When the elements of the input-output matrix $\mathbf{\Gamma}$ are sufficiently small, the following useful approximation for the vector of multipliers is valid. Suppose that the norm of $\mathbf{\Gamma}$, $\|\mathbf{\Gamma}\|_\infty = \max_{i,j\in 1:n} \gamma_{ji}$, is sufficiently small. Then

$$\boldsymbol{\mu} = \frac{1}{n}[\boldsymbol{I} - \mathbf{\Gamma}]^{-1}\mathbf{1} = \frac{1}{n}\left(\sum_{k=0}^{+\infty}\mathbf{\Gamma}^k\right)\mathbf{1} \approx \frac{1}{n}(\boldsymbol{I} + \mathbf{\Gamma})\mathbf{1} = \frac{1}{n}\mathbf{1} + \frac{1}{n}\mathbf{\Gamma}\mathbf{1} = \frac{1}{n}\mathbf{1} + \frac{1}{n}\boldsymbol{d^{out}}, \tag{7}$$

where $\boldsymbol{d^{out}} = \mathbf{\Gamma}\mathbf{1}$ is the vector of sectoral weighted out-degrees, $\boldsymbol{d^{out}} = \left(d_1^{out}, .., d_n^{out}\right)'$. Thus, larger multipliers correspond to sectors with larger out-degree. In view of Proposition 1, this implies that sectors with the largest out-degree have the most pronounced impact on aggregate value added of the economy.[20] For the sample of countries in our data the approximation of sectoral multipliers by sectoral out-degree (times and plus $1/n$) turns out to be quite good, as demonstrated by Figure A-3 in the Appendix.

2.4 Expected aggregate output

To quantitatively evaluate the model's predictions for cross-country income variation we follow a statistical approach that allows us to represent aggregate income as a simple function of the first and second moments of the joint distribution of sectoral IO multipliers and productivities. The distribution of multipliers captures the properties of the intersectoral network in each country, while the correlation between multipliers and productivities captures the interaction of the IO structure with sectoral TFP levels.

To start with, we assume that the elements γ_{ji} of the IO matrix $\mathbf{\Gamma}$ and the sector-specific productivity levels Λ_i^{rel} are both realizations of random variables. Here indices i, j refer to sectors and $\Lambda_i^{rel} = \frac{\Lambda_i}{\Lambda_i^{US}}$ is the sector-specific productivity level relative to the U.S. one. The randomness of the γ_{ji}s implies that sectoral multipliers μ_j are also random, since by definition in (7), $\mu_j \approx \frac{1}{n} + \frac{1}{n}\sum_{i=1}^n \gamma_{ji}$ for all j.

As we explain in the data section below, the joint empirical distribution of sectoral multipliers and productivities, (μ_i, Λ_i^{rel}), is approximately *log-Normal*, that is, the distribution of logs of these variables is close to Normal.[21] In particular, the fact that the distribution of μ_j is log-Normal

input in the production of that sector. Moreover, $\frac{1}{1-\gamma}$ is also the value of the *aggregate* multiplier in an n-sector economy where only one sector's output is used (in the proportion γ) as an input in the production of all other sectors. Thus, if the share of intermediate inputs in gross output of each sector is, for example, $\frac{1}{2}$ ($\gamma = \frac{1}{2}$), then a one-percent increase in TFP of each sector increases aggregate value added by $\frac{1}{1-\gamma} = 2$ percent. In more extreme cases, the aggregate multiplier – and hence, the effect of sectoral TFP improvements on aggregate value added – becomes infinitely large when $\gamma \to 1$ and it is close to 1 when $\gamma \to 0$. This is consistent with the intuition in Jones (2011b).

[20] Note that the vector of multipliers is closely related to the *Bonacich centrality* vector corresponding to the intersectoral network of the economy. This means that sectors that are more "central" in the network of intersectoral trade have larger multipliers and hence, play a more important role in determining aggregate output.

[21] To be precise, the distribution of $(\log(\mu_i), \log(\Lambda_i^{rel}))$ is a *truncated* bivariate Normal, where $\log(\mu_i)$ is censored from below at a certain $a < 0$. This is taken into account in our empirical analysis. However, the difference from a usual, non-truncated Normal distribution turns out to be inessential.

means that while the largest probability is assigned to relatively low values of a multiplier, a non-negligible weight is assigned to high values, too. In other words, the distribution is positively skewed, or possesses a fat right tail. The log-normality of the multipliers is, in turn, linked to the log-normality of the elements of $\mathbf{\Gamma}$. Indeed, as the distribution of sectoral multipliers μ_j is log-Normal, and $\mu_j \approx \frac{1}{n} + \frac{1}{n}\sum_{i=1}^{n}\gamma_{ji}$, the sum $\sum_{i=1}^{n}\gamma_{ji}$ must be distributed log-Normally. By the Fenton-Wilkinson method (Fenton, 1960), this is consistent with the γ_{ji}s being log-Normal.[22] Note that the assumption of log-normality of the γ_{ji}s imposes few restrictions on the IO matrices and, in particular, allows any given entry of the IO matrix to vary across countries and sectors.

To be more specific, we assume that for all ji, the elements γ_{ji} of the input-output matrix $\mathbf{\Gamma}$ are independent random draws from a log-Normal distribution, and that for each sector i, the pair (μ_i, Λ_i^{rel}) is a random draw from the same bivariate log-Normal distribution (independent of the sector but obviously country specific). Furthermore, in order to express sectoral log productivity coefficients λ_i in terms of the relative productivity Λ_i^{rel}, we use the approximation $\lambda_i = \log(\Lambda_i) \approx \Lambda_i^{rel} + \left(\log(\Lambda_i^{US}) - 1\right)$.

This then allows calculating the *expected aggregate income* $E(y)$ for y defined in (6), as follows:

$$\begin{aligned} E(y) &= n\left(E(\mu)E(\Lambda^{rel}) + cov(\mu, \Lambda^{rel})\right) + E\left[\sum_{i=1}^{n}\sum_{j=1 \text{ s.t. } \gamma_{ji}\neq 0}^{n} \mu_i\gamma_{ji}\log\gamma_{ji}\right] + \\ &\quad + E\left[\sum_{i=1}^{n}\mu_i(1-\gamma_i)\log(1-\gamma_i)\right] - \log n + \alpha\log(K) + E(\mu)\sum_{i=1}^{n}\left(\log\left(\Lambda_i^{US}\right) - 1\right) \end{aligned} \tag{8}$$

Imposing log-Normality, we obtain:

$$E(y) = ne^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} + e^{m_\mu + 1/2\sigma_\mu^2}\sum_{i=1}^{n}\left(\log\left(\Lambda_i^{US}\right) - 1\right) - \log n + \alpha\log(K) + \Psi(m_\mu, \sigma_\mu) \tag{9}$$

Here, m_μ and m_Λ are the means and σ_μ^2, σ_Λ^2 and $\sigma_{\mu,\Lambda}$ are the elements of the variance-covariance matrix of the bivariate Normal distribution of $(\log(\mu_i), \log(\Lambda_i^{rel}))$, so that $E(\mu) = e^{m_\mu + 1/2\sigma_\mu^2}$, $E(\Lambda^{rel}) = e^{m_\Lambda + 1/2\sigma_\Lambda^2}$ and $cov(\mu, \Lambda) = e^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2)} \cdot (e^{\sigma_{\mu,\Lambda}} - 1)$

The term $\Psi(m_\mu, \sigma_\mu)$ is a function of (m_μ, σ_μ) that summarizes the direct impact of the IO structure on aggregate income. It is a complicated function of the underlying distribution of the γ_{ji}s that we provide in Appendix B. Given its complexity, we can make further simplifying assumptions to gain more tractability. Assume that all non-zero elements of $\mathbf{\Gamma}$ are the same, that is, $\gamma_{ji} = \widehat{\gamma}$ for any i and j whenever $\gamma_{ji} > 0$. In addition, consider that the in-degree γ_i is independent of the sector, $\gamma_i = \gamma$ for all i, which is broadly consistent with the empirical homogeneity of intermediate input shares across sectors.[23] The two assumptions together imply that every column of $\mathbf{\Gamma}$ has the same

[22] See Supplementary Appendix for details.

[23] The assumption of constant in-degree is also employed in Acemoglu et al. (2012) and in Carvalho et al.(2010).

number of non-zero elements $\frac{\gamma}{\widehat{\gamma}}$. Instead, the rows of $\mathbf{\Gamma}$ are not restricted in that sense, leading to variation in sectors' out-degrees. These assumptions on γ_{ji} and γ_i then allow us to express expected aggregate output $E(y)$ in a much simpler form:[24]

$$\begin{aligned} E(y) &= ne^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} + (1+\gamma)(\gamma \log(\widehat{\gamma}) - 1) + \\ &+ \log(1-\gamma) - \log n + \alpha \log(K) + e^{m_\mu + 1/2\sigma_\mu^2} \sum_{i=1}^{n} \log\left(\Lambda_i^{US}\right). \end{aligned} \tag{10}$$

From (10) we see that the expected aggregate income increases in the expectation of the sectoral multiplier $E(\mu) = e^{m_\mu + 1/2\sigma_\mu^2}$ and in the expectation of the average productivity level $E(\Lambda^{rel}) = e^{m_\Lambda + 1/2\sigma_\Lambda^2}$. Moreover, a positive covariance between multipliers and productivity also increases aggregate income, while a negative one reduces it. In other words, higher relative productivities have a larger impact if they occur in sectors with higher multipliers. The other terms summarize how the IO coefficients and the capital stock matter for aggregate income, while the term including $\log\left(\Lambda_i^{US}\right)$ appears because productivity levels have been normalized by the ones of the U.S. (see the approximation for λ_i above). In the empirical analysis below we will show that whether we use (9) or the more restrictive version (10) makes hardly any difference for predicting income differences across countries. Then, we prefer the simpler and more intuitive version.

3 Dataset and descriptive analysis

3.1 Data

IO tables measure the flow of intermediate products between different plants, both within and between sectors. The ji'th entry of the IO table is the value of output from establishments in industry j that is purchased by different establishments in industry i for use in production.[25] Dividing the flow of industry j to industry i by gross output of industry i, one obtains the IO coefficient γ_{ji}, which states the cents of industry j's output used in the production of each dollar of industry i's output.

In order to construct a dataset of IO tables for a range of low- and high-income countries, to compute sectoral TFP levels, and to get information on countries' aggregate income and factor endowments, we combine information from three datasets: the World Input-Output Database (WIOD, Timmer, 2012), the Global Trade Analysis Project (GTAP version 6, Dimaranan, 2006), and the Penn World Tables, Version 7.1 (PWT, Heston et al., 2012).

[24] Note that when $\gamma_{ji} = \widehat{\gamma}$ for any i and j where $\gamma_{ji} > 0$ and $\gamma_i = \gamma$ for all i, $\sum_{j \text{ s.t. } \gamma_{ji} \neq 0} \mu_i \gamma_{ji} \log \gamma_{ji} = \mu_i \gamma \log(\widehat{\gamma})$ and $\sum_{i=1}^n \mu_i (1-\gamma_i) \log(1-\gamma_i) = \log(1-\gamma)$ since $\sum_{i=1}^n \mu_i (1-\gamma_i) = \mathbf{1}'[\boldsymbol{I} - \mathbf{\Gamma}] \cdot \frac{1}{n}[\boldsymbol{I} - \mathbf{\Gamma}]^{-1}\mathbf{1} = \frac{1}{n}\mathbf{1}'\mathbf{1} = 1$. Moreover, $\sum_{i=1}^n \mu_i \approx 1 + \gamma$ because from (7) it follows that $\sum_{i=1}^n \mu_i \approx 1 + \frac{\sum_{i=1}^n d_i^{out}}{n} = 1 + \frac{\sum_{i=1}^n d_i^{in}}{n}$ and $d_i^{in} = \gamma_i = \gamma$ for all i.

[25] Note that intermediate outputs must usually be traded between establishments in order to be recorded in the IO tables. Therefore, flows that occur within a given plant are not measured.

The first dataset, WIOD, contains IO data for 36 countries classified into 35 sectors in the year 2005. The list of countries and sectors is provided in Appendix Tables A-1 to A-3. WIOD IO tables are available in current national currency at basic prices.[26] In our main specification, IO coefficients are defined as the value of domestically produced plus imported intermediates divided by the value of gross output at basic prices.[27] Sectoral multipliers are computed as $\boldsymbol{\mu} = \{\mu_i\}_i = \frac{1}{n}[\boldsymbol{I} - \boldsymbol{\Gamma}]^{-1}\mathbf{1}$. As explained in more detail later, the WIOD data also allow us to compute sectoral TFPs.

The second dataset, GTAP version 6, contains data for 65 countries and 37 sectors in the year 2004. We use GTAP data to obtain more information about IO tables of low-income countries. We construct IO coefficients for all 65 countries.[28]

Finally, the third dataset, PWT, includes data on income per capita in PPP, aggregate physical capital stocks (constructed from investment data with the perpetual inventory method) and labor endowments for 155 countries in the year 2005. In our analysis, PWT data are mainly used to make out-of-sample predictions with our model.

3.2 IO structure

To start with, we provide some descriptive analysis of the IO structure of the countries in our data. To this end, we consider the sample of countries from the GTAP database. First, we sum IO multipliers of all sectors to compute the aggregate IO multiplier. While a sectoral multiplier indicates the change in aggregate income caused by a one-percent change in productivity of one specific sector, the aggregate IO multiplier tells us by how much aggregate income changes due to a one-percent change in productivity of all sectors. Figure 1 (left panel) plots aggregate IO multipliers for each country against GDP per capita (relative to the U.S.).

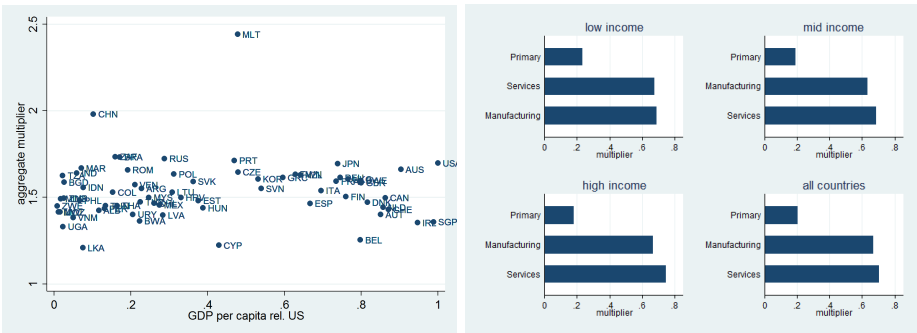

Figure 1: Aggregate IO-multipliers by country (left), sectoral IO-multipliers by income level (right)

[26] Basic prices exclude taxes and transport margins.

[27] In a robustness check, we separate domestically produced from imported intermediates and define domestic IO coefficients as the value of domestically produced intermediates divided by the value of gross output, while IO coefficients for imported intermediates are defined as the value of imported intermediates divided by the value of gross output. We show in the robustness section that this choice does not affect our results.

[28] Compared to the original GTAP classification, we aggregate all agricultural commodities in the GTAP data into a single sector. IO coefficients are computed as payments to intermediates (domestic and foreign) divided by gross output at purchasers' prices. Purchasers' prices include transport costs and net taxes on output (but exclude deductible taxes, such as VAT).

We observe that aggregate multipliers average around 1.6 and are uncorrelated with the level of income. This implies that a one-percent increase in productivity of all sectors raises per-capita income by around 1.6 percent on average.[29]

Next, we separately compute the aggregate IO multipliers for the three major sector categories: primary sectors (which include Agriculture, Coal, Oil and Gas Extraction and Mining), manufacturing and services. Figure 1 (right panel) plots these multipliers by income level. Here, we divide countries into low income (less than 10,000 PPP Dollars of per capita income), middle income (10,000-20,000 PPP Dollars of per capita income) and high income (more than 20,000 PPP Dollars of per capita income).

We find that multipliers are largest in services (around 0.65 on average), slightly lower in manufacturing (around 0.62) and smallest in the primary sector (around 0.2). As before, the level of income does not play an important role in this result: the comparison is similar for countries at all levels of income per capita.[30] We conclude that at the aggregate-economy level or for major sectoral aggregates there are no systematic differences in IO structure across countries.

Let us now look at differences in IO structure at a more disaggregate level. To this end, we compute sectoral IO multipliers separately for each sector and country. Figure 2 presents kernel density plots of the distribution of sectoral multipliers for different levels of income per capita.

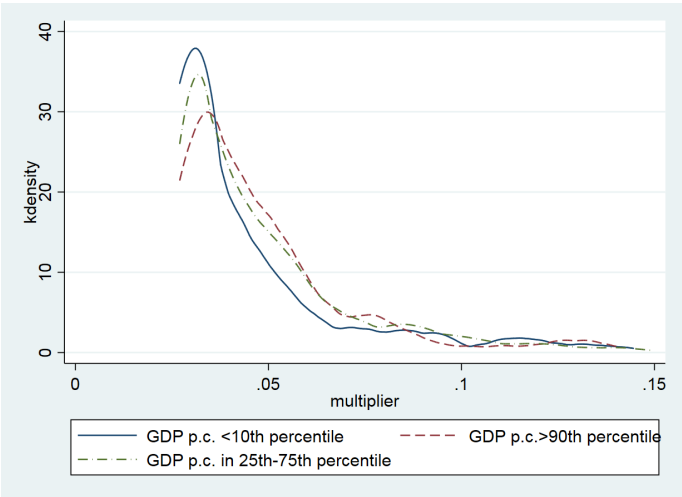

Figure 2: Distribution of sectoral multipliers, GTAP sample.

The following two facts stand out. First, the distribution of sectoral multipliers is *highly skewed*: while most sectors have low multipliers, a few sectors have multipliers way above the average. A typical low-multiplier sector (at the 10th percentile of the distribution of multipliers) has a multiplier of around 0.02 and the median sector has a multiplier of around 0.03. By contrast, a typical high-

[29] Aggregate multipliers for the WIOD sample are somewhat larger (with a mean of around 1.8) and also uncorrelated with the level of per capita income. A simple regression of the aggregate multipliers from the GTAP sample on those from the WIOD data gives a slope coefficient of around 0.8 and the relationship is strongly statistically significant.

[30] Very similar results are obtained for the WIOD sample. The only difference is that primary sectors are somewhat more important in low-income countries compared to others.

multiplier sector (at the 90th percentile of the distribution of multipliers) has a multiplier of around 0.065, while a sector at the 99th percentile has a multiplier of around 0.134.[31]

Second, the distribution of multipliers in low-income countries is *more skewed towards the extremes* than it is in high-income countries. In poor countries, almost all sectors have very low multipliers and a few sectors have very high multipliers. Differently, in rich countries the distribution of sectoral multipliers has significantly more mass in the center.[32]

Finally, we investigate which sectors tend to have the largest multipliers. We thus rank sectors according to the size of their multiplier for each country. The upper panels of Figure 3 plot sectoral multipliers for a few selected countries, which are representative for the whole sample: a very poor African economy (Uganda (UGA)), a large emerging economy (India (IND)) and a large high-income economy (United States (USA)). It is apparent that the distribution of multipliers in Uganda is such that the bulk of sectors have low multipliers, with the exception of Agriculture, Electricity, Trade and Inland Transport. By contrast, a typical sector in the U.S. has a larger multiplier, while the distribution of multipliers in India lies between the one of Uganda and the one of the U.S.[33]

In the lower panels of the same figure we plot sectoral multipliers averaged across countries by income level. Note that while the distributions of multipliers now look quite similar for different levels of income, this is an aggregation bias, which averages out much of the heterogeneity at the country level. From this figure we see that, in low-income countries, the sectors with the highest multipliers are Trade, Electricity, Agriculture, Chemicals, and Inland Transport, while in the set of middle- and high-income countries, the most important sectors in terms of multipliers are Trade, Electricity, Business Services, Inland Transport and Financial Services.

Thus, though in all income groups the sectors with the highest multipliers tend to be services, a notable difference between high-multiplier sectors of rich and poor countries is that the former contain exclusively service sectors, while the latter feature non-service sectors – Agriculture and Chemicals.[34] Moreover, the sectors with the lowest multipliers also differ across income levels and the differences in their composition across income groups are larger than those of the sectors with

[31] These numbers correspond to the GTAP sample. The numbers for the WIOD sample are similar: 10th percentile 0.03; median 0.045; 90th percentile 0.084, 99th percentile 0.153.

[32] A non-parametric Kruskal-Wallis test for equality of the distributions across groups rejects the null of equal distributions across income groups at the one-percent level. We provide more detailed statistical analysis of the shape of the distributions in section 4.1.

[33] One might be concerned that the IO structure in poor countries is mismeasured due to the importance of the informal sector in these countries and that the size of linkages is thus understated (manufacturing census and survey data used to construct IO tables do not include the informal sector). However, the fact that estimated average multipliers do not differ with GDP per capita and that agriculture has strong IO linkages in developing countries, even though most agricultural establishments are in the informal sector, mitigates this concern. In addition, the largest firms in a sector (which operate in the formal economy) typically account for the bulk of sectoral output and inputs and even more so in developing countries (Alfaro et al., 2008), so that the mismeasurement in terms of aggregate output and intermediate input demand is probably small.

[34] Agriculture is a high-multiplier sector in countries with an income level below 10,000 PPP dollars, where agricultural products are an input to many sectors.

Figure 3: Sectoral IO-multipliers by country (top panel)/ income level (bottom panel)

the highest multipliers.[35]

3.3 Productivities

We now explain the construction of a sectoral total factor productivity (TFP) relative to the U.S. and provide some descriptive evidence on sectoral TFPs as well as their correlation with sectoral multipliers. Here, we use the countries in the WIOD sample, because this information is available only for this dataset.

In particular, WIOD contains all the necessary information to compute gross-output-based sectoral total factor productivity: nominal gross output and material use, sectoral capital and labor inputs, sectoral factor payments to labor, capital and inputs for 35 sectors. Crucially, WIOD also provides purchasing power parity (PPP)-deflators (in purchasers' prices) for sector-level gross output that we use to convert nominal values into PPP units and which thus allow us to compute real TFPs at the sector level.[36] These deflators have been constructed by Inklaar and Timmer (2014) and are consistent in methodology and outcome with the latest version of the PWT. They combine expenditure prices and levels collected as part of the International Comparison Program (ICP) with data on industry output, exports and imports and relative prices of exports and imports

[35] In general though, the sectors with the lowest multipliers are also mostly services: Apparel, Air Transport, Water Transport, Gas Distribution and Dwellings (Owner-occupied houses).

[36] WIOD data comprises socio-economic accounts that are defined consistently with the IO tables. We use sector-level data on gross output, physical capital stocks in constant 1995 prices, the price series for investment, and labor inputs in hours. Using the sector-level PPPs for gross output, we convert nominal gross output and inputs into constant 2005 PPP prices. Furthermore, using price series for investment from WIOD and the PPP price index for investment from PWT 7.1, we convert sector-level capital stocks from WIOD into constant 2005 PPP prices.

from Feenstra and Romalis (2014). The authors use export and import values and prices to correct for the problem that the prices of goods consumed or invested domestically do not take into account the prices of exported products, while the prices of imported goods are included. To our knowledge, WIOD combined with these PPP deflators is the best available cross-country dataset for computing sector-level productivities using production data.

Given that we only have information on inputs and outputs in PPPs for a single year, we follow the development/growth accounting literature (e.g. Caselli, 2005; Jorgenson and Stiroh, 2000) and calibrate sector-level production functions. We compute TFP at the sector level relative to the U.S. (measured in constant 2005 PPPs) assuming constant-returns-to-scale Cobb-Douglas sectoral technologies for gross output (see eq. (1)), using average input shares for the output elasticities of inputs:

$$\Lambda_{ic}^{rel} \equiv \frac{\Lambda_{ic}}{\Lambda_{iUS}} = \frac{q_{ic}}{q_{iUS}} \frac{\left(k_{iUS}^{\alpha_i} l_{iUS}^{1-\alpha_i}\right)^{1-\gamma_i} d_{1iUS}^{\gamma_{1i}} d_{2iUS}^{\gamma_{2i}} \cdot \ldots \cdot d_{niUS}^{\gamma_{ni}}}{\left(k_{ic}^{\alpha_i} l_{ic}^{1-\alpha_i}\right)^{1-\gamma_i} d_{1ic}^{\gamma_{1i}} d_{2ic}^{\gamma_{2i}} \cdot \ldots \cdot d_{nic}^{\gamma_{ni}}}, \tag{11}$$

where i is the sector index and c is the country index. Consistently with the notation in the theoretical model, Λ_{ic}^{rel} denotes TFP of sector i normalized relative to the U.S., q_{ic} denotes the gross output of sector i, k_{ic} and l_{ic} are the quantities of capital and labor inputs and d_{ji} is the quantity of intermediate good j used in the production of sector i. $\alpha_i = 1/C \sum_{c=1}^{C} \alpha_{ic}$ and $1 - \alpha_i$ are the empirical sector-specific factor income shares in GDP averaged over the countries in the WIOD sample, $\gamma_{ji} = 1/C \sum_{c=1}^{C} \gamma_{jic} \in [0, 1)$ are the average intermediate input shares in gross output from the WIOD IO tables and $\gamma_i = \sum_{j=1}^{n} \gamma_{ji}$ is the total sector-specific intermediate share in gross output.[37]

In Table 1 we report means and standard deviations of relative productivity levels by income level, as well as the correlation between sectoral multipliers and productivities. To compute the standard deviations and correlations, we consider deviations from country means, so they are to be interpreted as within-country variation.

The following empirical regularities arise. First, average sectoral TFPs are highly positively correlated with income per capita. Second, the within-country standard deviation is highest for poor countries and lowest for rich countries. This is also apparent from the left panel of Figure 4, which plots histograms of log relative productivities by income level. Thus, low-income countries have much more dispersion in relative productivities across sectors than rich ones. Third, in low-income countries, TFP levels of high-multiplier sectors are above their average productivity level

[37] Applying more sophisticated parametric estimation methods developed for plant-level data to obtain consistent estimates of output elasticities (e.g., Olley and Pakes, 1996) is not necessary in our context. These methods solve the simultaneity bias that may arise when estimating the output elasticities of inputs with regression techniques by taking logs of (11), since unobserved TFP is correlated with input choice. Note, however, that using the empirical intermediate input shares γ_{ji} (as we do) solves this simultaneity problem when the production function is Cobb-Douglas and intermediate inputs are freely adjustable. Under these assumptions the first-order conditions for profit maximization imply that intermediate input shares are independent of (unobserved) TFP.

Sample	N	avg. TFP	std. TFP (within)	corr. TFP, multiplier (within)
low income	236	0.445	0.950	0.189***
mid income	340	0.619	0.667	0.065
high income	745	1.109	0.475	-0.135***
all	1,321	0.891	0.646	-0.026

Table 1: Descriptive statistics for sectoral TFPs and multipliers.*** indicates statistical significance at the 1-percent level.

relative to the U.S., while in richer countries TFP levels in these sectors tend to be below average. This is demonstrated by the examples in the center and right panels of Figure 4. For instance, India (center panel) has productivity levels above its average in the high-multiplier sectors Chemicals, Inland Transport and Refining and Electricity, while its productivity levels in the low-multiplier sectors such as Car Retailing, Telecommunications and Business Services are below average. An exception is India's high-multiplier sector Agriculture, where the productivity level is very low. This confirms the general view that poor countries tend to have particularly low productivity levels in this sector. By contrast, rich European economies, such as Germany (right panel), tend to have below-average productivity levels in high-multiplier sectors such as Financial Services, Business Services and Transport.[38]

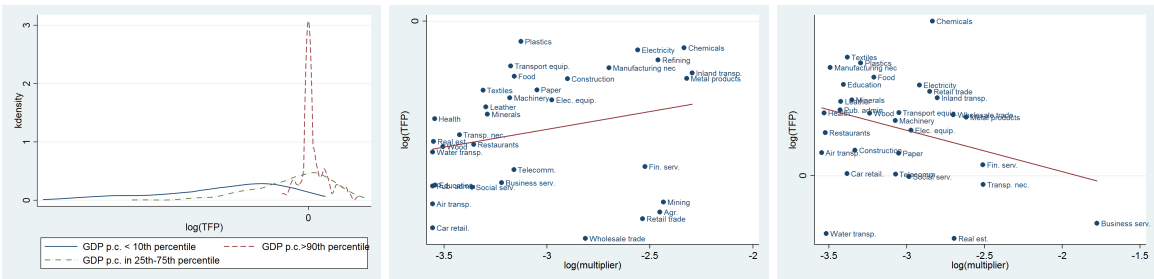

Figure 4: Distribution of sectoral log(TFP) relative to the U.S. (left panel). Correlation between IO-multipliers and productivities: India (middle panel) and Germany (right panel)

4 Empirical analysis

In this section we estimate the parameters of the Normal distribution of $(\log(\mu), \log(\Lambda^{rel}))$ for the sample of countries for which we have data. We allow parameter estimates to vary across countries in order to model the systematic underlying differences in IO structure and productivity that we have discussed in section 3. With the parameter estimates in hand we then use equations (9) and (10) to evaluate the predicted aggregate income in these countries and compare our baseline model with four simple alternatives which abstract from some of the elements present in our model: (i)

[38] While it is beyond the scope of this paper to develop a full economic model that explains why in developing countries productivity levels are above average in high-multiplier sectors (that is, productivity gaps relative to the U.S. are smaller in such sectors) and the opposite is true in industrialized countries, we provide a tentative explanation in section 4.1.

sectoral TFP differences; (ii) IO linkages; (iii) country-specific IO structure. We show that all these elements are important for understanding cross-country income differences.

4.1 Structural estimation

We assume that the vector of log multipliers and log relative productivities $\mathbf{Z} \equiv (\log(\mu), \log(\Lambda^{rel}))$ is drawn from a (truncated) bivariate Normal distribution with country-specific parameters $\mathbf{\Theta} = (\mathbf{m}, \mathbf{\Sigma})$, where $\mathbf{m}$ is the vector of means and $\mathbf{\Sigma}$ denotes the variance-covariance matrix. In order to allow the distributions of log multipliers and productivities to differ across countries, we first estimate the parameters separately for each country using Maximum Likelihood.[39] Observe that in the estimation we do not impose any structure on the data except for assuming joint log-Normality. In a second step, we then regress the estimated country-specific parameters $\hat{\mathbf{\Theta}}$ on (log) per capita income in order to test if the parameters indeed systematically vary with countries' income level, as suggested by the evidence presented in section 3.[40]

We estimate the statistical model using the empirical data for log multipliers and log TFPs constructed from the WIOD dataset (35 sectors, 36 countries). In the panels of Figure 5 we plot the country-specific estimates of all parameters against (log) per capita GDP and in Table 2 we report the corresponding results of regressing each parameter on log per capita GDP. Because the coefficients are Maximum-Likelihood estimates, we report bootstrapped standard errors. We label the set of predicted values from these regressions $\tilde{\mathbf{\Theta}}$.

We find that m_μ does not vary systematically with the income level (column (1)). Instead, σ_μ decreases significantly in log per capita GDP with a slope of -0.076 (column (2)). Thus, in the WIOD sample, poor countries have a distribution of log multipliers with the same average but with more dispersion than rich countries. Average log productivity, m_Λ, increases strongly in log per capita GDP (with a slope of around 1.4, see column (3)), while the standard deviation of log productivity, σ_Λ, is a decreasing function of the same variable (column (4)). This implies that rich countries have much higher average productivity levels and less dispersion in relative productivities across sectors than poor economies. Finally, note that the covariance between log multipliers and log productivity, $\sigma_{\mu,\Lambda}$, has a positive intercept and is a decreasing function of log per capita GDP (column (5)). Hence, poor countries have above-average productivity levels in sectors with higher

[39] The formula for the truncated bivariate Normal, where $\log(\mu)$ is censored from below at a is given by $f(\mathbf{Z}|log(\mu) \geq a) = \frac{1}{\sqrt{(2\Pi)^2|\mathbf{\Sigma}|}} exp[-1/2(\mathbf{Z}-\mathbf{m})'\mathbf{\Sigma}^{-1}(\mathbf{Z}-\mathbf{m})]/(1-F(a))$, where $F(a) = \int_{-\infty}^{a} \frac{1}{\sigma_\mu\sqrt{(2\Pi)}} exp[-1/2(\log(\mu)-m_\mu)^2/\sigma_\mu^2] d\log(\mu)$ is the cumulative marginal distribution of $\log(\mu)$ and where

$$\mathbf{m} = \begin{pmatrix} m_\mu \\ m_\Lambda \end{pmatrix}, \mathbf{\Sigma} = \begin{pmatrix} \sigma_\mu^2 & \sigma_{\mu,\Lambda} \\ \sigma_{\mu,\Lambda} & \sigma_\Lambda^2 \end{pmatrix}. \tag{12}$$

[40] We obtain very similar results by using an alternative, one-step procedure where we pool observations across countries and model coefficients as linear functions of (log) per capita income. Such approach is statistically more efficient than our two-step procedure, but it also imposes more structure on the data ex ante, which we would like to avoid.

multipliers, and the opposite is the case in rich countries.[41,42]

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	WIOD sample					GTAP sample	
	m_μ	σ_μ	m_Λ	σ_Λ	$\sigma_{\mu,\Lambda}$	m_μ	σ_μ
Constant	-5.462***	1.461***	-14.216***	3.606***	2.320***	-8.749***	1.868***
	(1.125)	(0.392)	(2.119)	(0.619)	(0.478)	(2.959)	(0.443)
log(GDP p.c.)	0.168	-0.076*	1.396***	-0.303***	-0.234***	0.368	-0.100**
	(0.112)	(0.039)	(0.209)	(0.061)	(0.047)	(0.300)	(0.046)
R-squared	0.002	0.046	0.590	0.557	0.343	0.012	0.057
Observations	36	36	36	36	36	65	65

Table 2: Regression of estimated country-specific parameters on log per capita GDP. Bootstrapped standard errors in parentheses. Estimates significant at 1% (***), 5% (**), 10% (*) significance level.

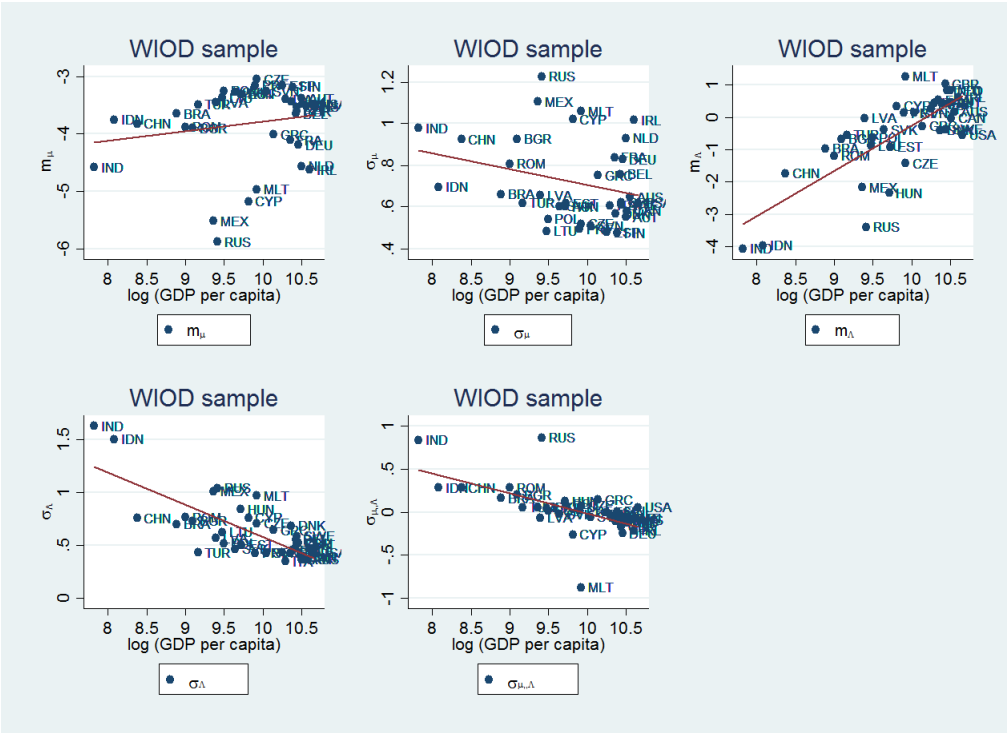

Figure 5: Correlation of country-specific coefficient estimates with log per capita GDP.

To obtain more information on the IO structure of low-income countries, we now redo the estimation using data for the GTAP sample (37 sectors, 65 countries). For this sample, we only have information on IO multipliers but not on productivity levels available. Therefore, we estimate a

[41] The sign of the covariance changes at per capita GDP of approximately 22,026 ($= e^{10}$) PPP Dollars.

[42] While developing a full economic model that explains the difference in covariance signs across countries is beyond the scope of this paper, one potential explanation could be as follows. Consider a model where adoption of technology from a sector-specific frontier is costly, whereas the profitability of adoption depends on market size effects (due to the standard economies-of-scale argument, e.g., Romer, 1990, Comin and Hobijn, 2004; Magalhães and Afonso, 2017) and the IO structure is given. Observe that the country-sector-specific multiplier is a proxy for market size and thus profitability. Then all countries will adopt technologies that are relatively closer to the frontier in sectors with larger multipliers. However, since in industrialized countries the (exogenously given) IO structure is similar to the one in the U.S., their high- and low-productivity sectors will be similar, so that TFP levels *relative* to the U.S. appear uncorrelated with multipliers. Moreover, even given a similar IO structure, distortions that make technology adoption in specific sectors less profitable may induce particularly large productivity gaps of European countries relative to the U.S. in high-multiplier sectors, such as services. This could then even lead to negative correlations between relative TFP levels and multipliers. By contrast, in developing countries, where the IO structure is different from the one in industrialized countries, the set of high-productivity sectors will also be different from those of the U.S., generating a positive correlation between relative TFP levels and multipliers.

univariate (truncated) Normal distribution for m_μ and σ_μ for each country. The results of regressing the country-specific parameter estimates on log per capita GDP are reported in columns (6) and (7) of Table 2. The results are quite similar to those for the WIOD sample: m_μ does not vary significantly with the income level (column (6)), while the standard deviation of log multipliers, σ_μ, is a decreasing function of (log) per capita income with a slope of -0.1 (column (7)). Again, this implies that in poor countries the average sector has the same log multiplier but there is more mass at the extremes of the distribution than in rich countries. We summarize these empirical findings below.

Summary of estimation results:

1. *The estimated distribution of log IO multipliers has a larger variance with more mass at the extremes in poor countries compared to rich ones.*

2. *The estimated distribution of log productivities has a lower mean and a larger variance in poor countries compared to rich ones.*

3. *Log IO multipliers and productivities correlate positively in poor countries and negatively in rich ones.*

4.2 Predicting cross-country income differences

We now plug the predicted values from the regressions of coefficient estimates on log per capita GDP, $\tilde{\mathbf{\Theta}}$, into the expressions for expected per capita GDP to forecast per capita income levels. The expressions for expected per capita GDP are given by equations (9) – for the general model that imposes no restrictions on the IO structure except log-normality – and (10) – for the model where the values of the positive entries in the IO matrix are restricted to be the same but their number and positions are random.[43] The remaining parameters are calibrated as follows. We set $(1-\alpha)$, the labor-income share in GDP, equal to 2/3 and we set n equal to 35, which corresponds to the number of sectors in the WIOD dataset.

4.2.1 Methodology

We compare the above models with four simple alternatives. The first one has no IO structure and no productivity differences, so that $y = E(y) = \alpha log(K)$. The second model, by contrast, features sectoral productivity differences but no IO linkages. It is easy to show that under the assumption that sectoral productivities follow a log-Normal distribution, predicted log income in this model

[43] The expression for $E(y)$ for the *truncated* distribution of (μ_i, Λ_i^{rel}) is somewhat more complicated and less intuitive than (10). However, the results for aggregate income using a truncated normal distribution for μ are very similar to the estimation of (10) and we therefore use the formulas for the non-truncated distribution. The details can be provided by the authors.

is given by $E(y) = e^{m_\Lambda + 1/2\sigma_\Lambda^2} + \alpha \log(K) + \frac{1}{n}\sum_{i=1}^{n}(\log(\Lambda_i^{US})) - 1$.[44] The third alternative model features sectoral productivity differences and IO linkages but keeps the IO structure constant across countries (by restricting the mean and the variance of the distribution of log multipliers and its covariance with log productivities to be constant across countries). Finally, the last model allows for country-specific IO structure but has no productivity differences.

To evaluate model performance, we provide several measures of fit. Our main measure of success in replicating cross-country income variation with the model is given by

$$\textbf{Success} \equiv \frac{\textit{coeff.var.}(Y)}{\textit{coeff.var.(GDP p.c.)}}, \tag{13}$$

where *coeff.var.(Y)* is the coefficient of variation (the standard deviation divided by the mean) of model-predicted income and *coeff.var.(GDP p.c.)* is the coefficient of variation of actual per capita GDP. Observe that the coefficient of variation is a standard scale-less measure of dispersion. **Success** compares the model-predicted variation in Y to the observed variation in GDP per capita, and the closer its value to one, the more successful the model at explaining cross-country income differences.[45] If the model generates less (more) variation in per capita income than is present in the data, **Success** will be smaller (larger) than unity.

Next, as a graphical measure for the goodness of fit, we plot model-predicted income relative to the U.S. against actual relative per capita GDP. Perfect fit would mean that the predicted relative income levels lie exactly on the 45-degree line. Finally, to statistically evaluate this graphical measure of fit, we regress model-predicted income relative to the U.S. on data for actual relative per capita GDP. If the model fits data perfectly, the estimate for the intercept should be zero *and* the regression slope *and* the R-squared should equal unity.

Note that these tests provide over-identification restrictions for our model, since there is no intrinsic reason for the model to fit data on relative per capita income well: we have not matched income data in order to estimate the parameters of the distribution of log IO multipliers and TFPs. Instead, we have just allowed their joint distribution to vary with the level of per capita income in the estimation procedure.

We first predict income levels for the sample of WIOD countries (36 countries), then for the GTAP sample (65 countries) and finally for the PWT sample (155 countries).

[44] $Y = \prod_{i=1}^{n} \Lambda_i^{1/n}(K)^\alpha$, hence $y = \frac{1}{n}\sum_{i=1}^{n}\lambda_i + \alpha\log(K)$. Using our approximation for productivity relative to the U.S., taking expectations and assuming that Λ_i^{rel} follows a log-Normal distribution, we obtain the above formula.

[45] Caselli (2005) instead uses the ratio of variances of log income generated by the model relative to the data as his main measure of success. While the variance of the log is also scale-less, it gives more weight to countries with small income levels. By contrast, we would like to weight observations equally.

4.2.2 Analysis with WIOD sample

The results of **Success** and the regression statistics for the WIOD sample are presented in Table 3. Columns (1), (2), (7) and (8) show the results for the four alternative models discussed in the previous section, while columns (3)-(6) (highlighted in bold type) present the results for our model, with the general and more restricted specification of the IO coefficients. To be more precise, in column (1), we report statistics for the model without TFP differences and IO structure. In column (2), we show the outcomes for the model with productivity differences but no IO structure. In column (3), we present them for our general model described by (9), when taking the parameter estimates for the distribution of (log) multipliers and productivities from the WIOD data.[46] In column (4), we display results for the same model when estimating the distribution of (log) multipliers using the GTAP data.[47] Instead, in columns (5) and (6), we use the more restricted model described by (10), employing parameter estimates from the WIOD and the GTAP data, respectively. In column (7), we force the distribution of multipliers (IO structure) to be the same across countries by restricting m_μ, σ_μ^2 and $\sigma_{\mu,\Lambda}$ to be constant. Finally, in column (8) we report statistics for the model with country-specific IO structure but without productivity differences.

The model without TFP differences and IO structure fails to generate sufficient variation in per capita income (see column (1) of Table 3 and the green squares in the left panel of Figure 6). **Success** is 0.64, which means that this model can explain 64% of income variation in the sample. Not surprisingly, it over-predicts income levels for poor countries. By contrast, the model with productivity differences but no IO linkages (column (2)) generates *more* income variation than is present in the data (**Success** is 1.43). This model predicts many countries to be significantly poorer than they actually are (red triangles in the upper panel of Figure 6). This implies that, when disregarding the role of the IO linkages, the TFP differences estimated from sectoral data are larger than those necessary to generate the observed cross-country income differences.

We next show results for the general model with productivity differences and country-specific IO structure, as estimated from the WIOD data (column (3)) and the GTAP data (column (4)). This model indeed performs much better than the ones without IO structure in terms of predicting cross-country income variation: **Success** for this model is 0.81 when estimating the IO structure from the WIOD data and an impressive 0.96 when estimating it from the GTAP data.[48] Thus, the general model marginally under-predicts cross-country income variation. In columns (5) and (6) we report results for the somewhat more restrictive model described by equation (10). **Success** for this model is 1.10, when using parameter estimates from WIOD and 1.07 when using the GTAP

[46] We use predicted values of the parameters from Table 2, columns (1)-(5).

[47] We use predicted values for the distribution of log multipliers from Table 2, columns (6) and (7).

[48] The GTAP data is more informative about cross-country differences in IO linkages than the WIOD data because it includes a much larger sample of low- and middle-income countries, which allows estimating differences in structure across countries more precisely.

estimates. Hence, the model just slightly over-predicts cross-country income variation.

A visual comparison of actual vs. predicted relative income in the left panel of Figure 6 confirms the substantially better fit of our model with IO linkages and productivity differences (blue circles) compared to the one without IO structure (which under-predicts relative income levels of most countries) and the one without IO structure and productivity differences (which over-predicts relative income levels for virtually all countries).[49] We conclude that allowing for a country-specific IO structure substantially improves model fit. In addition, the most general model performs just slightly better than the more restrictive version that requires the positive values in the IO matrix to be equal, but allows their number and positions to be random. We thus prefer this second, more restrictive model, since (10) is much easier to interpret than (9). For the remainder of this section, we therefore concentrate on this model and refer to it as our *baseline model*.

Next, we test if the inclusion of an IO structure per se or rather the interaction of cross-country differences in IO structure with productivity differences account for the improved model fit. In column (7) we thus restrict the parameters m_μ, σ_μ^2 and $\sigma_{\mu,\Lambda}$ to be the same for all countries. We find that this model fits the data significantly worse than the one with country-specific IO structure and very similarly to the model without IO structure: **Success** is now 1.45. This implies that cross-country variation in IO structure is crucial for predicting differences in income across countries given estimated productivity differences. Finally, in column (8) we report results for the model with a country-specific IO structure but without productivity differences. This model does even worse than the model without TFP differences and IO structure: **Success** goes down to 0.48. Intuitively, poor countries have more dispersion in log multipliers, and hence higher average *levels* of multipliers, than rich countries, which increases their aggregate income levels and exacerbates the problem of models without productivity differences that over-predict income levels of poor countries.[50]

Note that the good fit of the baseline model, which features both country-specific IO structure and TFP differences, does not simply add up these two components but points to complementarities between them. **Success** of the baseline model (columns (5) and (6)) is 1.07 (1.10). This is an improvement of 36 (33) percentage points compared to the model with productivity differences and no IO structure (column (2)) that has a **Success** of 1.43. Of this number, just introducing an IO structure without considering its interaction with sectoral TFPs (transition from column (1) to column (8)) reduces income differences and explains an improved fit of 16 percentage points (=0.64-0.48). The remaining improvement in fit is due to interaction effects between TFP differences and IO structure.

[49]This improved fit is confirmed by the regression statistics: for the baseline model, the intercept is not statistically different from zero, the slope coefficient equals 1.000 and the R-squared is 0.939. By contrast, the model in column (1) has an intercept of 0.371, a slope coefficient of 0.832 and an R-squared of 0.710; the model in column (2) has an intercept of -0.141, a slope coefficient of 0.967 and an R-squared of 0.927.

[50]Note that according to (10), larger σ_μ^2 displayed by poor countries increases predicted income in these countries.

In conclusion, there are two main factors that determine the improved fit of the baseline model with country-specific IO structure compared to the models without IO structure or with constant structure. First, the differences in IO structure between low- and high-income countries: poor countries have only few highly connected sectors and many sectors that are relatively isolated, while rich countries have more intermediately connected sectors. Second, in contrast to rich countries, poor economies have above-average productivity levels in high-multiplier sectors. We will further investigate the impact of each of these factors in the section on counter-factuals.

	WIOD sample							
	(1) no IO structure no TFP diff.	(2) no IO structure TFP diff.	**(3) WIOD est. general IO TFP.diff**	**(4) GTAP est. general IO TFP diff.**	**(5) WIOD est. baseline IO TFP diff.**	**(6) GTAP est. baseline IO TFP diff.**	(7) WIOD est. constant IO TFP diff.	(8) WIOD est. baseline IO no TFP diff.
success	**0.64**	**1.43**	**0.81**	**0.96**	**1.10**	**1.07**	**1.45**	**0.48**
intercept	0.371***	-0.141***	-0.034	0.045*	-0.033	-0.023	-0.145***	0.625***
	(0.060)	(0.029)	(0.054)	(0.025)	(0.023)	(0.021)	(0.029)	(0.082)
slope	0.832***	0.967***	1.014***	1.004***	1.000***	1.040***	0.963***	0.572***
	(0.101)	(0.051)	(0.076)	(0.045)	(0.039)	(0.039)	(0.052)	(0.126)
R-squared	0.710	0.927	0.819	0.923	0.939	0.940	0.925	0.453
Observations	36	36	36	36	36	36	36	36

Table 3: Model fit: WIOD sample. Standard errors in parentheses. Estimates significant at 1% (***), 5% (**), 10% (*) significance level.

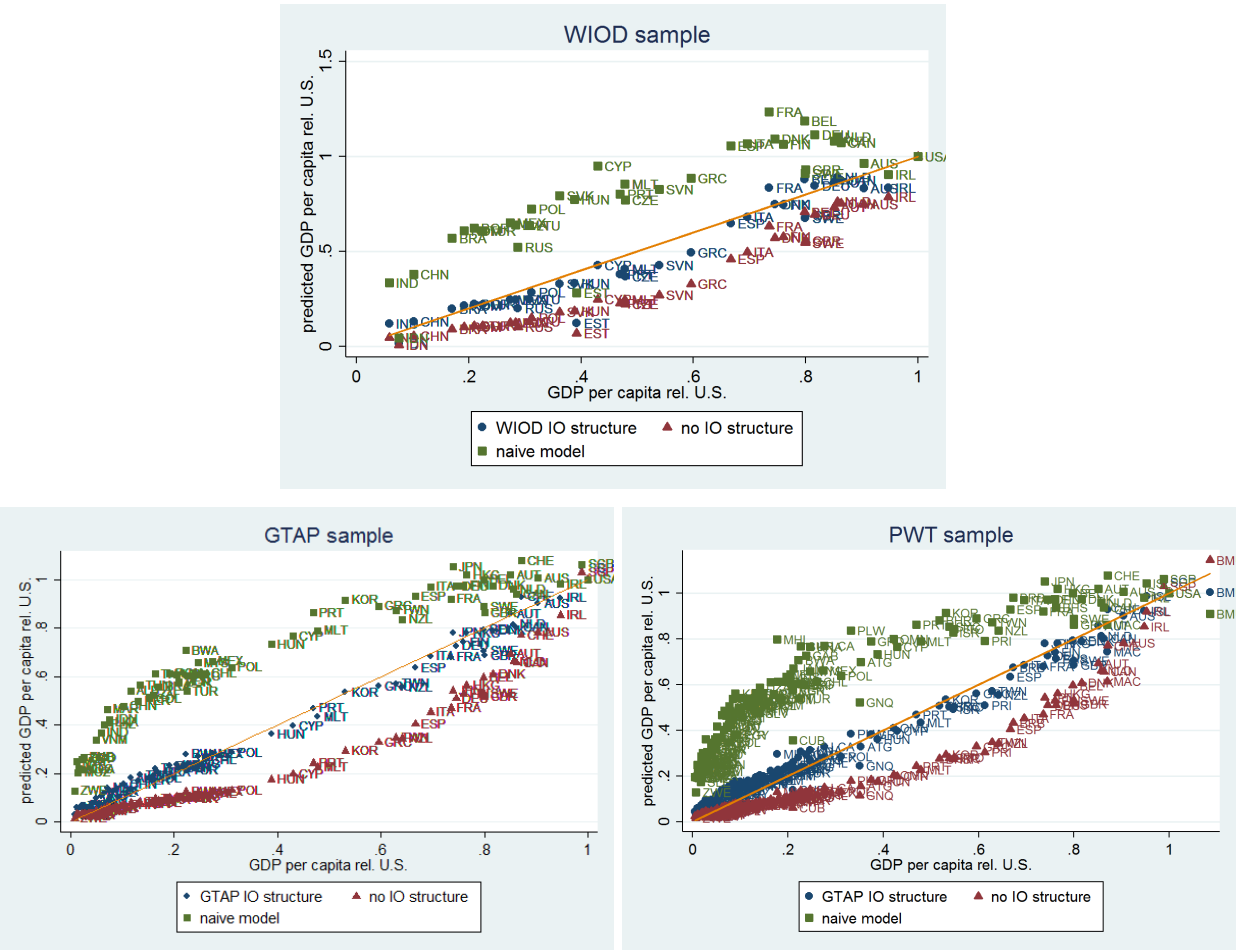

Figure 6: Predicted income per capita: model fit for different samples.

	GTAP sample				PWT sample			
	(1) no IO structure no TFP diff.	(2) no IO structure TFP diff.	**(3) WIOD est. baseline IO TFP diff.**	**(4) GTAP est. baseline IO TFP diff.**	(5) no IO structure no TFP diff.	(6) no IO structure TFP diff.	**(7) WIOD est. baseline IO TFP diff.**	**(8) GTAP est. baseline IO TFP diff.**
success	**0.51**	**1.27**	**0.92**	**0.95**	**0.43**	**1.19**	**0.77**	**0.85**
intercept	0.365***	-0.044***	0.054***	0.042***	0.342***	-0.018***	0.073***	0.051***
	(0.022)	(0.013)	(0.008)	(0.005)	(0.012)	(0.006)	(0.004)	(0.003)
slope	0.779***	0.804***	0.839***	0.918***	0.823***	0.763***	0.802***	0.897***
	(0.039)	(0.043)	(0.025)	(0.016)	(0.034)	(0.038)	(0.020)	(0.013)
R-squared	0.887	0.916	0.968	0.987	0.830	0.910	0.965	0.984
Observations	65	65	65	65	155	155	155	155

Table 4: Model Fit: GTAP and PWT Samples.
Standard errors in parentheses. Estimates significant at 1% (***), 5% (**), 10% (*) significance level.

4.2.3 Analysis with GTAP and PWT samples

Next, we turn to testing model fit in the sample of GTAP countries and the sample of countries in the Penn World Tables. The latter sample is usually employed for development accounting exercises. In Table 4, columns (1)-(4) show results for the GTAP sample. In column (1) we report outcomes for the model without productivity differences and without IO structure, which has a **Success** of 0.51. In column (2) we present results for the model with productivity differences but no IO structure. As before, this model over-predicts income variation across countries, with **Success** equal to 1.27. Next, turning to our baseline model with IO structure and productivity differences, in column (3) we show results using parameter estimates from the WIOD sample. This model performs very well with a **Success** of 0.92. Similarly, the baseline model with parameter estimates from the GTAP sample (column (4)) has a **Success** of 0.95. The increased goodness of fit can also be seen from the left panel of Figure 6, where we plot predicted income against actual income for the baseline model (blue circles), the model without TFP differences and IO structure (green squares) and the model with TFP differences but no IO structure (red triangles). While the model without TFP differences and IO structure considerably over-predicts and the model without IO structure under-predicts relative income levels for most countries, the baseline model with productivity differences and IO structure is extremely close to the 45-degree line. Only for the poorest countries it slightly over-predicts their relative income levels.

Finally, we discuss model fit in the PWT sample (see columns (5)-(8)). This requires to predict not only TFP levels but also IO structure out of sample. As is well known, the performance of the model without productivity differences and IO structure (column (5)) is quite poor, with a **Success** of around 0.43 in this sample, since this model strongly over-predicts income levels for poor countries (green squares in the right panel of Figure 6). In column (6) we report fit for the model with TFP differences but without IO structure. It has a **Succcess** of 1.19 and thus, also in this sample it over-predicts income variation across countries (red triangles in the right panel of Figure 6). Turning to the models with both TFP differences and IO structure, we find that

they somewhat under-predict income variation in this sample. **Success** is 0.77 for the WIOD IO structure (column (7)) and 0.85 for the GTAP IO structure (column (8)). Still, as the right panel of Figure 6 and the regression statistics make clear, this model fits the data better than the other models: most blue circles are extremely close to the 45-degree line. The exception are very poor economies, whose income levels the model with IO structure over-predicts. Here, the extrapolation of IO structure seems to make too extreme predictions for the distribution of log multipliers and their covariance with TFP. Still, we infer that including interaction effects between productivity and IO structure into the model helps to significantly improve model fit. To wrap up, we now present a summary of our findings.

Summary of model fit:

1. *The baseline model with estimated sectoral productivity differences and IO structure performs substantially better in terms of predicting cross-country income levels and their variation than a model without productivity differences (which under-predicts income variation) and a model with productivity differences but without IO structure (which over-predicts income variation).*

2. *The above results hold for three different samples of countries: the WIOD dataset (36 countries), the GTAP dataset (65 countries) and the Penn World Tables dataset (155 countries).*

5 Counter-factual experiments

We now present the results of a number of counter-factual experiments. We first investigate in more detail how differences in IO linkages – as summarized by the distribution of multipliers – matter for cross-country income differences. Thus, in our first counter-factual exercise we set the distribution of log multipliers in all countries equal to the U.S. one by fixing m_μ and σ_μ^2 at the predicted values of a country at the U.S.-level of per capita income.[51] Note that given the Cobb-Douglas structure, our model allows us to separately identify sectoral TFP levels and IO structure and it thus makes sense to vary one of the two factors, while holding the other one fixed.[52] The result of this experiment

[51]The experiment holds m_μ fixed and reduces σ_μ for virtually all countries, since, according to Table 2, σ_μ is a decreasing function of GDP per capita. For a log-normal distribution such a change shifts mass away from the lower and upper tails towards the center of the distribution.

[52]Note that productivity levels are also unaffected by changes in the distribution of IO multipliers even when technologies are not factor-neutral. To see this, note that labor-augmenting or intermediate-augmenting rather than Hicks-neutral technologies would imply:

$$q_i = \left[k_i^\alpha (\Lambda_i l_i)^{1-\alpha}\right]^{1-\gamma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}}, \tag{14}$$
$$q_i = \left(k_i^\alpha l_i^{1-\alpha}\right)^{1-\gamma_i} (\Lambda_i^{\gamma_i}) d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}}$$

Under these assumptions, a change in the γ_{ji}s (reflecting a change in the distribution of multipliers) would potentially also affect measured productivity $\Lambda_i^{(1-\alpha)(1-\gamma_i)}$ or $\Lambda_i^{\gamma_i}$. While this is true in general, given our assumption that the intermediate share $\gamma_i = \sum_{j=1}^N \gamma_{ji}$ is constant across sectors, this is not a concern. Therefore, any change in the underlying IO structure that is implied by a change in the parameters m_μ or σ_μ leaves TFP levels unaffected.

is shown in the left panel of Figure 7. It plots the counter-factual percentage change in income per capita against GDP per capita relative to the U.S. As can be seen from the figure, virtually all countries would lose in terms of income if they had the U.S. IO structure. These losses are decreasing in income per capita and range from negligible levels for countries with income levels close to the U.S. one to more than 60 percent of per capita income for very poor countries such as Congo (ZAR) or Zimbabwe (ZWE).

The reason why most countries lose in this counter-factual experiment is the shape of the distribution of log multipliers in the U.S. compared to the one of low-income countries: the typical sector in the U.S. is intermediately connected (the mode of the distribution is larger than in poor countries) and the distribution of (log) multipliers has less mass in the right tail compared to poor countries. Given the relationship between logs and levels of the distribution of (μ_i, Λ_i^{rel}), this means that assigning the U.S. distribution to other countries reduces both their average (level) multiplier and the absolute value of the correlation between TFP and multipliers. Given the factual positive correlation of TFP and multipliers in low-income countries, they thus perform much worse with their new IO structure: now their average multiplier is lower and so is the correlation between TFP and multipliers, preventing them to benefit from their "super-star" sectors.

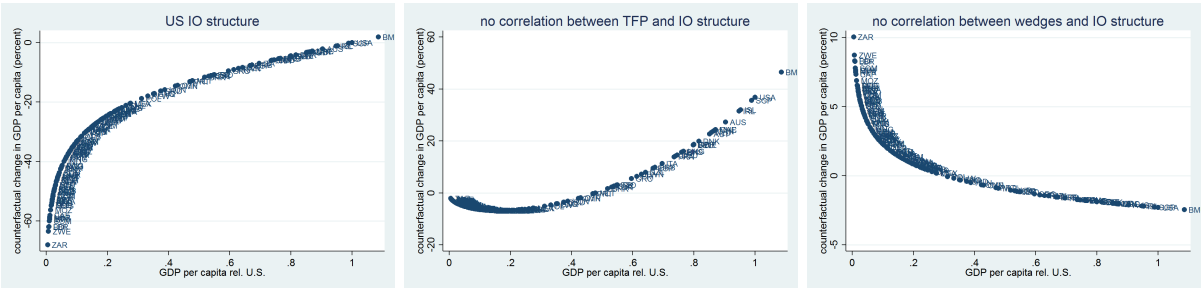

Figure 7: Counter-factuals

In the second counter-factual exercise, we keep the mean and the variance of log multipliers fixed and instead set the covariance between log multipliers and log productivities, $\sigma_{\mu,\Lambda}$, to zero. We can see from the central panel of Figure 7 that poor countries (up to around 40 percent of the U.S. level of income per capita) would lose up to 10 percent in terms of their initial income, while rich countries would gain up to 40 percent from this change. Why is this the case? From our estimates, poor countries have a positive covariance between log multipliers and log TFPs, while rich countries have a negative one. This implies that poor countries are doing relatively well despite their low average productivity levels, because they perform significantly better than average precisely in those sectors that have a large impact on aggregate performance. The opposite is true in rich countries, where the same covariance tends to be negative, so that highly connected sectors perform below average. Eliminating this link improves aggregate outcomes in rich economies further, while hurting

poor countries.[53]

To sum up, recall that in low-income economies just a few sectors, such as Energy, Transport and Trade, provide inputs for most other sectors, while the typical sector provides inputs to only a few sectors. Thus, it suffices to have comparatively high productivity levels in those crucial sectors in order to obtain a relatively satisfactory aggregate outcome. By contrast, in industrialized countries most sectors provide inputs for several other sectors (the IO network is quite dense), but there are hardly any sectors that provide inputs to most others. Thus, with such an IO structure increasing productivity levels in a few selected sectors is no longer enough to achieve a relatively good aggregate performance.

Finally, the last panel of Figure 7 describes the results of the third counter-factual exercise, that is based on a model including sector-specific distortions or tax wedges. We discuss this counter-factual in section 6.1 below.

Summary of counter-factual experiments:

1. *Imposing the dense IO structure of the U.S. on poor economies would reduce their income levels by up to 60 percent because a typical sector, which has a lower productivity level than the high-multiplier sectors in these economies, would become more connected.*

2. *If poor economies did not have above-average productivity levels in high-multiplier sectors, their income levels would be by up to 10 percent lower.*

6 Robustness checks

In this section, we report the results of a number of robustness checks in order to show that our findings do not hinge on the specific restrictions imposed by the baseline model. We consider the following modifications of our benchmark setup. First, we allow IO multipliers to depend on implicit tax wedges. Second, we extend our model to sectoral CES production functions. Third, we generalize the final demand structure by introducing expenditure shares that differ across countries and sectors. Fourth, we explicitly account for imported intermediate inputs. Finally, we allow for skilled and unskilled labor as separate production factors. We show that none of these modifications changes the basic conclusions of the baseline model. The formulas for aggregate income implied by these more general models as well as detailed derivations can be found in the Appendix.

[53]Note that as sectoral productivities are considered *relative* to the U.S., setting $\sigma_{\mu,\Lambda}$ to zero would actually not make any difference for the U.S. (it has zero correlation between multipliers and TFP by construction), but it would make a difference for a country with the U.S. level of GDP per capita (hence, label "U.S." on the figure), such as rich European countries. In these countries negative correlations arise due to particularly large productivity gaps with the U.S. in high-multiplier sectors, such as services (see more on this in section 4.1). Setting $\sigma_{\mu,\Lambda}$ to zero then effectively means bringing European productivity levels in the service sectors to the U.S. level. This would certainly have a large impact on GDP of European countries.

6.1 Wedges

One important concern is that empirically observed IO coefficients do not just reflect technological input requirements but also sector-specific distortions or wedges τ_i in the production of intermediates. To see this, consider the maximization problem of an intermediate producer:

$$\max_{\{d_{ji}\}}(1-\tau_i)p_i\Lambda_i\left(k_i^{\alpha}l_i^{1-\alpha}\right)^{1-\gamma_i}d_{1i}^{\gamma_{1i}}d_{2i}^{\gamma_{2i}}\cdot\ldots\cdot d_{ni}^{\gamma_{ni}}-\sum_{j=1}^{n}p_jd_{ji}-rk_i-wl_i,$$

taking $\{p_i\}$ as given (τ_i and Λ_i are exogenous). Sector-specific wedges are assumed to reduce the value of sector i's production by a factor $(1-\tau_i)$, so that $\tau_i > 0$ means an implicit tax and $\tau_i < 0$ means an implicit subsidy on the production of sector i's output.

The first-order condition w.r.t. d_{ji} is given by

$$(1-\tau_i)\gamma_{ji}=\frac{p_jd_{ji}}{p_iq_i}\quad j\in 1:n$$

Thus, a larger wedge in sector i implies lower observed IO coefficients in this sector since firms in a sector facing larger implicit taxes demand less inputs from all other sectors. Separately identifying wedges τ_i and technological IO coefficients γ_{ji} is an empirical challenge, which requires to impose some additional restrictions on the data. Observe that τ_i is the same for all inputs j demanded by a given sector i. Thus, introducing a country index c and summing across inputs j for a given country, we obtain

$$(1-\tau_{ic})\sum_j\gamma_{jic}\equiv(1-\tau_{ic})\gamma_{ic}=\sum_j\frac{p_{jc}d_{jic}}{p_{ic}q_{ic}}\quad i\in 1:n$$

Now, if we restrict the total technological intermediate share of sector i, γ_{ic}, to be the same across countries for a given sector i, we can identify country-sector specific wedges as

$$(1-\tau_{ic})=\sum_j\frac{p_{jc}d_{jic}}{p_{ic}q_{ic}}\frac{1}{\gamma_i}\quad i\in 1:n \tag{15}$$

Observe that individual IO coefficients γ_{jic} are still allowed to differ across countries in an arbitrary way. According to equation (15), countries with below-average intermediate shares, $\sum_j\frac{p_{jc}d_{jic}}{p_{ic}q_{ic}}$, in a certain sector face an implicit tax in this sector, while countries with above-average intermediate shares receive an implicit subsidy. It is then straightforward to estimate γ_i using regression techniques. Taking logs of equation (15), we obtain:

$$\log\left(\sum_j\frac{p_{jc}d_{jic}}{p_{ic}q_{ic}}\right)=\log(\gamma_i)+\log(1-\tau_{ic}) \tag{16}$$

Given (16), we regress the intermediate input shares of each country-sector pair on a set of sector-specific dummies to obtain estimates of the technological intermediate shares $\log(\gamma_i)$ and then back out $\log(1-\tau_{ic})$ as the residual. The left panel of Figure 8 plots the distribution of intermediate input shares and the right panel plots the distribution of $\log(1-\tau_{ic})$ by income level for the WIOD sample. Average intermediate shares do not vary systematically with per capita income, but poor countries have a larger fraction of sectors with very low intermediate shares and a lower fraction with high intermediate shares. Correspondingly, poor countries have a larger fraction of sectors with relatively high wedges, which corresponds to more mass in the left tail of the distribution of $\log(1-\tau_{ic})$. Given wedges τ_{ic}, we construct IO coefficients adjusted for wedges as $\gamma_{ijc} = \frac{p_{jc}d_{jic}}{p_{ic}q_{ic}}\frac{1}{(1-\tau_{ic})}$. We then recompute sectoral productivities and IO multipliers using these adjusted IO coefficients.

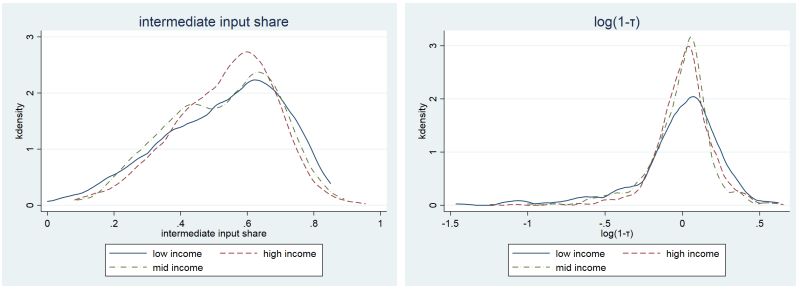

Figure 8: Intermediate input shares (left panel); wedges (right panel).

One can show that in the presence of wedges which are considered as pure waste,[54] and under the same simplifying restrictions used in our baseline model (cf. equation (10)), the expression for aggregate income can be written as:[55]

$$y = \sum_{i=1}^{n} \mu_i \Lambda_i^{rel} + \sum_{i=1}^{n} \mu_i (1-\tau_i) + \sum_{i=1}^{n} \mu_i \gamma \log(\widehat{\gamma}) + \log(1-\gamma) - \log n + \alpha \log K - 2(1+\gamma) + \sum_{i=1}^{n} \mu_i \log(\Lambda_i^{US}).$$

Now, assuming that sectoral multipliers, productivities and $(1-\tau_i)$ are stochastic, we obtain that expected aggregate output, $E(y)$, is given by:

$$\begin{aligned} E(y) &= n\left(E(\mu)E(\Lambda^{rel}) + cov(\mu, \Lambda^{rel}) + E(\mu)E(1-\tau) + cov(\mu, 1-\tau)\right) + (1+\gamma)(\gamma \log(\widehat{\gamma}) - 2) + \\ &+ \log(1-\gamma) - \log n + \alpha \log(K) + E(\mu) \sum_{i=1}^{n} \log\left(\Lambda_i^{US}\right). \end{aligned} \tag{17}$$

Again, this equation has an intuitive interpretation: higher average wedges τ are detrimental to

[54] In an unreported robustness check we verified that considering the revenues from tax wedges and rebating them lump sum to households does not make much difference for the results.

[55] With wedges equation (6) for aggregate income includes in addition the term $\sum_{i=1}^{n} \mu_i \log(1-\tau_i)$, which, for small enough τ_i, can be approximated by $-\sum_{i=1}^{n} \mu_i \tau_i = \sum_{i=1}^{n} \mu_i (1-\tau_i) - \sum_{i=1}^{n} \mu_i$. Then under the same simplifying restrictions as before, $\sum_{i=1}^{n} \mu_i \approx 1+\gamma$, and we obtain an equation very similar to (10).

aggregate income and more so if the average sector has higher multiplier; moreover, the negative impact of high wedges is particularly distorting if wedges co-vary positively with multipliers (i.e., $cov(\mu, 1-\tau) < 0$). If we impose joint log normality on the triple $(\mu, \Lambda_i^{rel}, 1-\tau)$, we obtain:

$$\begin{aligned} E(y) &= n\left(e^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2+\sigma_\Lambda^2)+\sigma_{\mu,\Lambda}} + e^{m_\mu + m_{1-\tau} + 1/2(\sigma_\mu^2+\sigma_{1-\tau}^2)+\sigma_{\mu,1-\tau}}\right) + (1+\gamma)(\gamma\log(\widehat{\gamma}) - 2) + \\ &+ \log(1-\gamma) - \log n + \alpha\log(K) + e^{m_\mu + 1/2\sigma_\mu^2}\sum_{i=1}^{n}\log\left(\Lambda_i^{US}\right), \end{aligned} \tag{18}$$

where m_μ, m_Λ, $m_{1-\tau}$ are the means and σ_μ^2, σ_Λ^2, $\sigma_{1-\tau}^2$, $\sigma_{\mu,\Lambda}$ and $\sigma_{\mu,1-\tau}$ are the elements of the variance-covariance matrix of the Normal distribution of $(\log(\mu), \log(\Lambda_i^{rel}), \log(1-\tau))$.

Given data on $(1-\tau)$, productivities Λ^{rel} and multipliers μ and imposing log-Normality on them, we re-estimate the parameters of their joint distribution separately for each country using Maximum Likelihood. We then regress these country-specific parameter estimates on (log) per capita GDP. Table 5 reports the result.[56] While the point estimates are quantitatively somewhat different from those of the baseline model (compare with Table 2), the qualitative features remain very similar: the average log multiplier, m_μ, does not vary with income, while σ_μ decreases in (log) per capita GDP. Again, this result implies that in poor countries the distribution of log multipliers has more mass at the extremes. Average log productivity, m_Λ, is again strongly increasing in income, while the standard deviation of log productivity, σ_Λ, is decreasing. The mean of the distribution of $\log(1-\tau)$, $m_{1-\tau}$, does not change significantly with the income level, while its standard deviation, $\sigma_{1-\tau}$, decreases in (log) per capita GDP. Moreover, in rich countries wedges tend to be lower ($(1-\tau)$ is larger) in sectors with high multipliers, while the opposite is true in poor countries.[57] Finally, productivity levels correlate positively with log multipliers in poor countries, and the correlation decreases with the income level.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	m_μ	σ_μ	m_Λ	σ_Λ	$m_{1-\tau}$	$\sigma_{1-\tau}$	$\sigma_{\mu,\Lambda}$	$\sigma_{\mu,1-\tau}$
intercept	-2.868***	0.847**	-10.483***	4.002***	-0.116	0.670***	0.607*	-0.105*
	(0.341)	(0.156)	(1.908)	(0.877)	(0.306)	(0.154)	(0.317)	(0.063)
slope	-0.026	-0.035**	1.009***	-0.314***	-0.009	-0.049**	-0.049*	0.012*
	(0.035)	(0.016)	(0.187)	(0.087)	(0.030)	(0.023)	(0.025)	(0.006)
R-squared	0.010	0.151	0.610	0.579	0.008	0.321	0.153	0.287
Observations	31	31	31	31	31	31	31	31

Table 5: Regression of estimated country-specific parameters on log(GDP p.c.).
Bootstrapped standard errors in parentheses. Estimates significant at 1% (***), 5% (**), 10% (*) significance level.

Next, we plug the predicted parameter values into equation (18) to forecast income levels. The results for model fit with this specification are provided in column (1) of Table 6. **Success** of this

[56] Note that we have less observations than in Table 2 (31 instead of 36) because the Maximum Likelihood estimation does not converge for all countries.

[57] The sign of the covariance changes at the level of per capita GDP of approximately 6311 ($= e^{0.105/0.012}$) PPP Dollars.

model is 0.98, which means that the model with wedges predicts cross-country income variation almost perfectly and even better than the baseline model. We thus conclude that introducing wedges in addition to an IO structure helps to improve model fit in the WIOD sample by another 5 percentage points (0.98 instead of 1.07). The reason is that compared to the baseline model, this tends to reduce the income levels of poor economies, where $m_{1-\tau} < 0$ and $\sigma_{\mu,1-\tau} < 0$, which lowers predicted income.

We also check in the following counter-factual experiment if the cross-country variation in the covariance between wedges and log multipliers has important quantitative implications. We thus set this covariance to zero for all countries. The right panel of Figure 7 (see section 5) plots the resulting changes in per capita income (in percent) against GDP relative to the U.S. level. Poor countries – which empirically exhibit a positive covariance between multipliers and wedges – experience an increase in income (up to 10 percentage points for Congo (ZAR)), while rich countries – which empirically have a negative covariance between multipliers and wedges – lose around one to two percentage points of per capita income. This implies that removing the positive covariance between wedges and multipliers in poor economies can lead to significant gains for them. However, cross-country income changes are smaller than those that would be induced by removing the covariance between productivities and multipliers.

In the Appendix we study optimal taxation and the welfare gains from moving from the current tax wedges to an optimal tax system that keeps tax revenue constant. Our results suggest that when the government is concerned with maximizing GDP per capita subject to a given level of tax revenue, the actual distribution of tax rates in rich countries is close to the optimum. By contrast, in poor countries, the mean of the distribution is too low and the variance is too high relative to the optimal values. Furthermore, for a given value of tax variance, a negative correlation of taxes with IO multipliers is optimal, while the actual correlation in poor countries is positive. Overall, we find that the poorest countries in the world could gain up to 10 % in terms of income per capita by moving to an optimal tax system.

	(1) wedges	(2) demand	(3) open	(4) skill
success	**0.98**	**1.18**	**0.85**	**0.93**
intercept	0.031	-0.031	0.120***	0.092**
	(0.025)	(0.050)	(0.034)	(0.035)
slope	1.018***	0.791***	0.897***	1.030***
	(0.045)	(0.081)	(0.053)	(0.069)
R-squared	0.927	0.826	0.887	0.832
Observations	36	36	36	36

Table 6: Robustness checks

6.2 CES production function

Another potential concern is that sectoral production functions are not Cobb-Douglas, but instead feature an elasticity of substitution between intermediate inputs different from unity. If this were the case, IO coefficients would no longer be sector-country-specific constants γ_{jic} but would instead be endogenous to equilibrium prices, which would reflect the underlying productivities of the upstream sectors. While it has been observed that for the U.S. the IO matrix has been remarkably stable over the last decades despite large shifts in relative prices (Acemoglu et al., 2012) – an indication of a unit elasticity, – in this robustness check we briefly discuss the implications of considering a more general CES sectoral production function. The sectoral production functions are now given by:

$$q_i = \Lambda_i \left(k_i^{\alpha} l_i^{1-\alpha}\right)^{1-\gamma_i} M_i^{\gamma_i}, \tag{19}$$

where $M_i \equiv \left(\sum_{j=1}^{N} \gamma_{ji} d_{ji}^{\frac{(\sigma-1)}{\sigma}}\right)^{\frac{\sigma}{(\sigma-1)}}$. The rest of the model is specified as in section 3.1.

With CES production functions the equilibrium cannot be solved analytically, so one has to rely on numerical solutions. However, it is straightforward to show how IO multipliers are related to sectoral productivities in this case. From the first-order conditions it follows that the relative expenditure of sector i on inputs produced by sector j relative to sector k is given by:

$$\frac{p_j d_{ji}}{p_k d_{ki}} = \left(\frac{p_j}{p_k}\right)^{1-\sigma} \left(\frac{\gamma_{ji}}{\gamma_{ki}}\right) \tag{20}$$

Thus, if $\sigma > 1$ ($\sigma < 1$), each sector i spends relatively more on the inputs provided by sectors that charge lower (higher) prices. These sectors then have higher (lower) multipliers, as multipliers are proportional (up to a shift by $1/n$) to the sector's out-degree $d_j^{out} = \sum_{i=1}^{n} \frac{p_j d_{ji}}{p_i q_i}$ (see equation (7)). Moreover, since prices are inversely proportional to productivities, sectors with higher productivity levels charge lower prices. Consequently, when $\sigma > 1$, sectoral multipliers and productivities should be positively correlated in *all* countries, while when $\sigma < 1$, the opposite should be true. We confirm these results in unreported simulations. Observe that these predictions are not consistent with our empirical finding that multipliers and productivities are positively correlated in low-income countries, while they are negatively correlated in high-income ones. Consequently – unless the elasticity of substitution differs systematically across countries – the data on IO tables and sectoral productivities are difficult to reconcile with CES production functions.

6.3 Cross-country differences in final demand structure

So far we have abstracted from cross-country differences in the final demand structure, which also matter for the values of sectoral multipliers since sectors with higher final-expenditure shares have

a larger impact on GDP. In the next robustness check, we thus consider a more general demand structure. More specifically, we now model the production function for the aggregate final good as $Y = y_1^{\beta_1} \cdot ... \cdot y_n^{\beta_n}$, where β_i is allowed to be country-sector-specific. The advantage of this specification is that it picks up differences in the final demand structure that may have an impact on aggregate income. The drawback is that with this specification multipliers reflect both the IO structure and final demand. Thus, this specification does not allow one to differentiate between the two channels. The vector of sectoral multipliers is now defined as $\boldsymbol{\mu} = \{\mu_i\}_i = [\boldsymbol{I} - \boldsymbol{\Gamma}]^{-1}\boldsymbol{\beta}$, where $\boldsymbol{\beta} = (\beta_1, .., \beta_n)'$. So, holding constant the IO structure $\boldsymbol{\Gamma}$, sectors with larger final-expenditure shares have larger multipliers. The interpretation of IO multipliers is identical to the one before: each sectoral multiplier μ_i reveals how a change in productivity of sector i affects total value added in the economy. Given the new multipliers, we re-estimate their joint distribution and predict income levels using the formula presented in the Appendix.

The fit of this model can be found in column (2) of Table 6. **Success** is now 1.18, which is somewhat worse than the performance of our baseline model (1.07). Like the model without linkages, this model somewhat over-predicts cross-country income differences. This indicates that – within the context of our model – modeling differences in final demand structure does not help to understand differences in aggregate income. The reason is that modeling differences in final demand structure across countries introduces a lot of additional noise in the multiplier data, which makes it harder to estimate the systematic underlying features of the inter-industry linkages.

6.4 Imported intermediates

So far we have abstracted from international trade and we have assumed that all intermediate inputs have to be produced domestically. Imported intermediate inputs may to some extent mitigate low productivity of domestic firms in the upstream sector, by enabling domestic producers to source from foreign suppliers.[58] Here, we allow for both domestically produced and imported intermediates, which are imperfectly substitutable. We thus assume that sectoral production functions are given by:

$$q_i = \Lambda_i \left(k_i^{\alpha} l_i^{1-\alpha}\right)^{1-\gamma_i-\sigma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot ... \cdot d_{ni}^{\gamma_{ni}} \cdot f_{1i}^{\sigma_{1i}} f_{2i}^{\sigma_{2i}} \cdot ... \cdot f_{ni}^{\sigma_{ni}}, \tag{21}$$

where d_{ji} are domestically produced intermediate inputs and f_{ji} are imported intermediate inputs. γ_{ji} and σ_{ji} denote the shares of each domestic and imported intermediate, respectively, in the value of sectoral gross output. We change the construction of the IO tables accordingly by separating domestically produced from imported intermediates. We then re-estimate the joint distributions of IO multipliers and productivities.

[58] See, e.g. Halpern, Koren and Szeidl (2015) for a recent micro-level study on the effect of importing intermediate inputs on the productivity levels of domestic producers.

The results for model fit with this specification are given in column (3) of Table 6. **Success** is now 0.85, which is slightly worse than the fit of the baseline model. The intuition for why results remain similar when considering imported inputs comes from the fact that most high-multiplier sectors tend to be services, which are effectively non-traded. Therefore, allowing for trade does not change the statistical distribution of multipliers and the implied predicted income much. We thus conclude that our results are quite robust to allowing for trade in intermediates.

6.5 Skilled labor

Finally, we split aggregate labor endowments into skilled and unskilled labor. Namely, let the technology of each sector $i \in 1:n$ in every country be described by the following Cobb-Douglas function:

$$q_i = \Lambda_i \left(k_i^{\alpha} u_i^{\delta} s_i^{1-\alpha-\delta}\right)^{1-\gamma_i-\sigma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}}, \tag{22}$$

where s_i and u_i denote the amounts of skilled and unskilled labor used by sector i, $\gamma_i = \sum_{j=1}^{n} \gamma_{ji}$ is the share of intermediate goods in the total input use of sector i and α, δ, $1-\alpha-\delta \in (0,1)$ are the respective shares of capital, unskilled and skilled labor in the remainder of the inputs. The total supply of skilled and unskilled labor in the economy is fixed at the exogenous levels of S and U, respectively. We define skilled labor as the number of hours worked by workers with at least some tertiary education and we define unskilled labor as the number of hours worked by workers with less than tertiary education. Information on skilled and unskilled labor inputs by sector is from WIOD. We recompute productivities Λ^{rel} assuming production-functions as given by (A-16) and then re-estimate all parameter values. We calibrate $\delta = 1/6$ to fit the college skill premium of the U.S. The results for fitting cross-country income variation with this model are provided in column (4) of Table 6. **Success** is now 0.93, which is comparable to the baseline model. This is not suprising: given the great fit of the baseline model, there is little room left for improving the explanatory power of the model by introducing human capital. We conclude that our results are not very sensitive to the definition of labor endowments.

7 Conclusions

In this paper we have studied the role of IO structure and its interaction with sectoral productivity levels in explaining income differences across countries. We have described and formally modeled cross-country differences in IO linkages and shown that they are important for understanding the income differences. Poor countries rely on a few highly connected sectors, which tend to have higher-than-average productivity levels. Their typical, low-productivity sectors are not strongly linked to the rest of the economy, mitigating their impact on aggregate income. By contrast, in rich

countries the typical sector is intermediately connected and the economy is not dominated by a few "super-star" sectors. Thus, while increasing productivity levels in a few sectors can have a large positive impact on aggregate income in poor economies, this is not the case in medium-income and rich countries. In these more densely connected economies the productivity levels of many more sectors need to be sufficiently high in order to guarantee a high income level. These insights have important consequences for the design of development policies.

References

[1] Aghion, P. and P. Howitt (1992), A Model of Growth Through Creative Destruction, *Econometrica* 60, 323-351.

[2] Alfaro, L., Charlton., A. and F. Kanczuk (2008), Plant Size Distribution and Cross-Country Income Differences, in J. Frankel and C. Pissarides, eds.,*NBER International Seminar on Macroeconomics*, University of Chicago Press, 243 - 272.

[3] Arbex, M. and F. S. Perobelli (2010), Solow Meets Leontief: Economic Growth and Energy Consumption, *Energy Economics*, 32 (1), 43-53.

[4] Acemoglu, D., Carvalho V.M., Ozdaglar A., and A. Tahbaz-Salehi (2012), "The Network Origins of Aggregate Fluctuations," *Econometrica*, 80 (5), 1977-2016.

[5] Barrot J. N. and J. Sauvagnat (2016), "Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks, " *Quarterly Journal of Economics*, 131 (3), 1543-1592.

[6] Bartelme, D. and Y. Gorodnichenko (2015), "Linkages and Economic Development", *NBER Working Paper* 21251.

[7] Basu, S. and J. G. Fernald (2002), Aggregate Productivity and Aggregate Technology, *European Economic Review*, 46, 963-991.

[8] Bonacich, P. (1987), "Power and Centrality: A Family of Measures," *American Journal of Sociology*, 92, 1170-1182.

[9] Boehm, J. (2015), The Impact of Contract Enforcement Costs on Outsourcing and Aggregate Productivity, *Sciences Po, unpublished manuscript.*

[10] Burress, D. (1994), "Homeomorphism Between Leontief and Cobb-Douglas Input-Output Models," *Economics Letters*, 44, 49-53.

[11] Carvalho, V.M. (2010), "Aggregate Fluctuations and the Network Structure of Intersectoral Trade," *Working Paper.*

[12] Carvalho, V.M. and N. Voigtländer (2014), "Input diffusion and the evolution of production networks," *National Bureau of Economic Research.*

[13] Caselli, F. (2005), Accounting for Cross-Country Income Differences, in Philippe Aghion and Steven A. Durlauf, eds., *Handbook of Economic Growth, New York: North Holland*, 679-741.

[14] Chenery, H., S. Robinson, and M. Syrquin (1986), Industrialization and Growth: A Comparative Study, *New York, Oxford University Press.*

[15] Ciccone, A. (2002), Input Chains and Industrialization, *Review of Economic Studies*, 69 (3), 565-587.

[16] Comin, D. and B. Hobijn. (2004), Cross-Country Technology Adoption: Making the Theories Face the Facts, *Journal of Monetary Economics*, 51(1), 39-83.

[17] Comin, D. and B. Hobijn. (2010), An Exploration of Technology Diffusion, *The American Economic Review*, 100(5),2031-2059.

[18] Dimaranan, B., ed.(2006), Global Trade, Assistance, and Production: The GTAP 6 Data Base, *Center for Global Trade Analysis, Purdue University*

[19] Feenstra, R .C. and J. Romalis (2014), International Prices and Endogenous Quality, *Quarterly Journal of Economics*, 129(2), 477-527.

[20] Fenton, L.(1960), The Sum of Log-Normal Probability Distributions in Scatter Transmission Systems, *IRE Transactions on Communication Systems*, CS-8, 5767.

[21] Gollin, D., Lagakos D., and M. Waugh (2013), Agricultural Productivity Differences across Countries, *American Economic Review*, 104 (5), 165-170.

[22] Grobovsek, J. (2018), Development Accounting with Intermediate Goods, *B.E. Journal Macroeconomics* 18, (1) 1-27.

[23] Hall, R. and C. I. Jones (1999), Why Do Some Countries Produce So Much More Output Per Worker Than Others?, *Quarterly Journal of Economics*, 114 (1), 83-116.

[24] Halpern, L., M. Koren and A. Szeidl (2015), Imported Inputs and Productivity. *American Economic Review*, 105 (12), 3660-3703.

[25] Heston, A., R. Summers and B. Aten (2012), Penn World Table Version 7.1, *Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania*

[26] Hirschman, A.O. (1958), "The Strategy of Economic Development," *New Haven, CT: Yale University Press.*

[27] Hsieh, C.-T. and P.J. Klenow (2009), Misallocation and Manufacturing TFP in China and India, *Quarterly Journal of Economics*, 124 (4), 1403-1448.

[28] Hulten, C. R. (1978), Growth Accounting with Intermediate Inputs, *Review of Economic Studies*, 45 (3), 511–518.

[29] Inklaar, R. and Timmer, M. (2014), The relative price of services. *Review of Income and Wealth*, 60(4), 727-746.

[30] Jones, I. C. (2011a), "Intermediate Goods and Weak Links in the Theory of Economic Development" *American Economic Journal: Macroeconomics*, 3 (2), 1-28.

[31] Jones, I. C. (2011b), "Misallocation, Input-Output Economics, and Economic Growth" *Version 2.0. Prepared for the 10th World Congress of the Econometric Society meeting in Shanghai.*

[32] Jorgenson, D.W. and K. Stiroh (2000), "U.S. Economic Growth at the Industry Level" *American Economic Review, Papers and Proceedings*, 90 (2), 161-167.

[33] Klenow, P. and A. Rodriguez-Clare (1997), The Neoclassical Revival in Growth Economics: Has It Gone Too Far?, in Ben Bernanke and Julio Rotemberg, eds., *NBER Macroeconomics Annual.*

[34] Long, J. B. and C. I. Plosser (1983), Real Business Cycles, *Journal of Political Economy*, 91 (1), 39-69.

[35] Magalhães, M. and Ò. Afonso (2017), A Multi-Sector Growth Model with Technology Diffusion and Networks, *Research Policy* 46 (7), 1340-1359.

[36] Miller, R. E. and P. D. Blair (2009), Input-Output Analysis: Foundations and Extensions, 2nd edition. *Cambridge University Press.*

[37] Oberfield, E. (2018), A Theory of Input-Output Architecture, *Econometrica*, 86, 2, 559589.

[38] Olley, G. Steven and Pakes, Ariel (1996), The Dynamics of Productivity in the Telecommunications Equipment Industry, *Econometrica*, 64(6), 1263-1297.

[39] Restuccia, D., D. Yang and X. Zhu (2008), Agriculture and aggregate productivity: A quantitative cross-country analysis, *Journal of Monetary Economics*, 55 (2), 234-250.

[40] Romer, P. M. (1990), Endogenous Technological Change, *Journal of Political Economy*, 98 (5), S71-S102.

[41] Timmer, M., ed. (2012), The World Input-Output Database (WIOD): Contents, Sources and Method", *WIOD Working Paper Number 1*

[42] Vollrath, D. (2009), How important are dual economy effects for aggregate productivity?, *Journal of Development Economics*, 88(2), 325-334.

Appendix A: Optimal taxation

The model with wedges employed in section 6.1 considers wedges as exogenously given and wasteful. In this section, we introduce an active role for the government and address the problem of optimal taxation by interpreting wedges τ_i as taxes imposed by the government to finance its expenditures and, possibly, also proceed to redistribution. To do that, we should specify the objective function of the government or social planner that is to be maximized by the choice of tax rates. As there are no other frictions, the redistribution motive is likely to be absent. Then we analyze the problem of optimal taxation for exogenously specified government expenditures. The appealing feature of analyzing such semi-optimal taxation schemes (with exogenously fixed government expenditures) is that they are much less dependent on the specific welfare function. Indeed, as long as welfare increases with individual consumption C, any welfare function would generate the same outcome for exogenously fixed government consumption G. In short, we will designate this analysis as GDP per capita maximization with exogenous G.[59]

A-1 Optimal taxes: setup

To derive characteristics of optimal tax scheme, we use the equilibrium expression for log GDP modified to account for government revenues. The logarithm of GDP per capita, y, is given by

$$\begin{aligned} y &= \sum_{i=1}^{n} \mu_i \lambda_i + \sum_{i=1}^{n} \mu_i \log(1-\tau_i) + \sum_{i=1}^{n} \sum_{j \text{ s.t. } \gamma_{ji} \neq 0} \mu_i \gamma_{ji} \log \gamma_{ji} + \sum_{i=1}^{n} \mu_i (1-\gamma_i) log(1-\gamma_i) - \log n + \\ &+ \log\left(1 + \sum_{i=1}^{n} \tau_i \bar{\mu}_i\right) + \alpha \log K, \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{\tau} &= \{\tau_i\}_i, && n \times 1 \text{ vector of sector-specific taxes} \\ \bar{\boldsymbol{\mu}} &= \{\bar{\mu}_i\}_i = \frac{1}{n}[\boldsymbol{I} - \bar{\boldsymbol{\Gamma}}]^{-1}\mathbf{1}, && n \times 1 \text{ vector of multipliers corresponding to } \bar{\boldsymbol{\Gamma}} \\ \bar{\boldsymbol{\Gamma}} &= \{\bar{\gamma_{ji}}\}_{ji} = \{\frac{\tau_i}{n} + (1-\tau_i)\gamma_{ji}\}_{ji}, && n \times n \text{ input-output matrix adjusted for taxes} \end{aligned}$$

This expression is very similar to the one in (7) of the baseline model but includes two extra terms that capture the effects of taxation: taxes, on the one hand, are distortionary and more so in sectors with larger multipliers, but on the other hand, they also contribute to government expenditures and thereby increase GDP.[60]

We consider the optimization problem in which this expression is maximized subject to a given level of government consumption. To solve that problem, we follow the statistical approach, in line with the rest of the paper. That is, instead of considering actual values of taxes, we focus on the first and second moments of their distribution that generate the highest predicted aggregate output $E(y)$ for a given level of expected tax revenues/government consumption as computed from the data.[61] The expected values of aggregate output and tax revenues/government consumption are computed via a Monte Carlo optimization method under the assumption that sectoral IO multipliers, productivities and $(1-\tau_i)$ follow a trivariate log-Normal distribution. All parameters of this distribution, apart from those that relate to the distribution of taxes, are fixed at the levels of their empirical estimates. Then by varying the mean, variance and covariance of the tax distribution,[62]

[59] In unreported simulations we have considered the case with endogenous government expenditures. There we assumed that government expenditures enter households utility in a Cobb-Douglas fashion. The results were very similar to those of the model that takes government expenditure as given.

[60] The detailed proof is available from the authors.

[61] An analytical solution in terms of actual values of tax rates (that maximize y subject to a given level of tax revenues) appears feasible only under some strong simplifying assumptions, which eventually lead to trivial or corner values of tax rates. We therefore resort to the statistical approach, which is also consistent with our approach in the rest of the paper.

[62] By covariance we mean the covariance between the distribution of taxes and IO multipliers, as the covariance

we derive the features of the optimal tax scheme. The results of this numerical analysis can be briefly summarized as follows.

A-2 Optimal taxes: results

We assume that for each country, government consumption is fixed at the level generated by the estimated distributions. We find that the optimal tax distribution is degenerate with variance $\sigma_\tau^2 \to 0$. The correlation between taxes and IO multipliers is not relevant in the limit. Empirically, the optimal mean tax rate in poor countries is substantially higher than the estimated ones (for some poor countries the optimal mean tax rate can be larger by a factor of 10). For rich countries, the optimal tax rate is only marginally larger. In fact, the estimated distribution of tax rates in rich countries turns out to be close to optimum, featuring low variance and reasonable mean. In poor countries, instead, the variance is high and the estimated mean tax rate is substantially lower than the optimal one. Moreover, there is a large positive correlation between tax rates and sectoral IO multipliers in poor countries, which ensures that high-multiplier sectors are taxed more. The latter is precisely the reason why a given level of tax revenues in poor countries can be reached with a lower mean tax rate than prescribed in optimum. Indeed, under the optimal tax scheme all sectors should be taxed evenly, and then raising the same amount of tax revenues requires a higher mean tax. Still, we find that the distortion loss associated with high (optimal) mean tax is small compared to the loss associated with taxing high-multiplier sectors more. The left panel of Figure A-1 plots welfare gains (in terms of percentage gains in GDP) of moving to a uniform tax rate that generates the same revenue as the current tax system against GDP per capita. The welfare gains are basically zero for all high-income countries but they can rise to up to 10% of GDP for some of the poorest countries in the world.

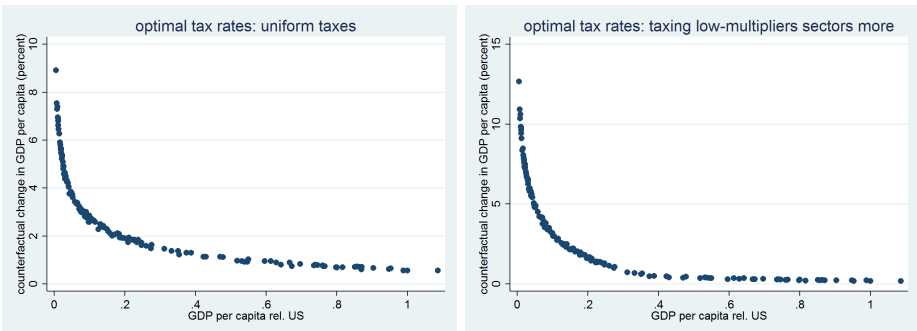

Figure A-1: Optimal taxation

We also perform a more unusual experiment. Indeed, as there might be reasons why tax rates cannot be uniform, we want to explore the role of the covariance between taxes and IO multipliers for a given variation in tax rates. We set the variance of the tax rate distribution to be equal to the estimated value in each country and examine the role of choosing the optimal correlation between the distribution of tax rates and sectoral IO multipliers and the mean tax rate that keeps tax revenue constant. We find that the optimal tax distribution has negative correlation with sectoral IO multipliers, so that consistently with the findings of our empirical analysis, more central sectors should be taxed less. The right panel of Figure A-1 plots the percentage gains in GDP per capita of moving to the optimal correlation between taxes and multipliers that keeps tax revenue constant. Again, welfare gains are substantial for very poor countries. Moreover, moving to a negative correlation between taxes and multipliers and increasing average tax rates would imply gains which are almost as large as those of moving to a uniform tax rate.

between taxes and productivities does not affect the calculated values.

Appendix B: Log-Normally distributed IO coefficients

In the baseline model used for the most of our analysis we imposed the restrictive and unrealistic assumption that all non-zero elements of the input-output matrix $\mathbf{\Gamma}$ are the same, that is, $\gamma_{ji} = \widehat{\gamma}$ for any i and j whenever $\gamma_{ji} > 0$. Here we consider a more general version of the model where γ_{ji}s are independent random draws from a log-Normal distribution and are thus allowed to vary across countries and sectors. As we explained in section 2.4, this distribution is appropriate due to three observations: (i) by equation (7), sectoral multipliers can be approximated by the sum of IO coefficients in the corresponding row of the IO matrix (shifted and multiplied by $1/n$), (ii) sectoral multipliers are log-Normally distributed, and (iii) the sum of independent log-Normal random variables is approximately log-Normal according to the Fenton-Wilkinson method (Fenton, 1960).

When IO coefficients are not constant, the term $\sum_{i=1}^{n} \sum_{j \text{s.t. } \gamma_{ji} \neq 0} \mu_i \gamma_{ji} \log \gamma_{ji}$ in equation (6) is no longer equal to $\sum_{i=1}^{n} \mu_i \gamma \log(\widehat{\gamma})$ and the term $\sum_{i=1}^{n} \mu_i (1-\gamma_i) \log(1-\gamma_i)$ is not equal to $\log(1-\gamma)$. As a result, the expectations of these terms are given by longer and more complex expressions that we derive in the Supplementary Appendix. Both of them are functions of the parameters of the Normal distribution of $\log \gamma_{ji}$, $(\mu_\gamma, \sigma_\gamma^2)$. These parameters, in turn, are related to the parameters of the Normal distribution of $\log(\mu)$, (m_μ, σ_μ^2), due to the relationship established in (7), $\mu_j \approx \frac{1}{n} + \frac{1}{n} \sum_{i=1}^{n} \gamma_{ji}$.

This then leads to the expression for the expected aggregate income that was given in (9):

$$
\begin{aligned}
E(y)) &= n\left(E(\mu)E(\Lambda^{rel}) + cov(\mu, \Lambda^{rel})\right) + E\left[\sum_{i=1}^{n} \sum_{j=1 \text{ s.t. } \gamma_{ji} \neq 0}^{n} \mu_i \gamma_{ji} \log \gamma_{ji}\right] + \\
&\quad + E\left[\sum_{i=1}^{n} \mu_i (1-\gamma_i) log(1-\gamma_i)\right] - \log n + \alpha \log(K) + E(\mu) \sum_{i=1}^{n} \left(\log\left(\Lambda_i^{US}\right) - 1\right) \\
&= n e^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} + e^{m_\mu + 1/2\sigma_\mu^2} \sum_{i=1}^{n} \left(\log\left(\Lambda_i^{US}\right) - 1\right) - \log n + \alpha \log(K) + \Psi(m_\mu, \sigma_\mu),
\end{aligned}
$$

with

$$
\begin{aligned}
\Psi(m_\mu, \sigma_\mu) &= x^{\frac{1}{2}} z \Big[\left(n + x^{\frac{1}{2}} z (n^2 - 1)\right) (\log(x) + \log(z)) + \frac{1}{2n} z^2 x^4 - 1 \\
&\quad + x^2 z^2 \left(\frac{n}{2} + 1 - \frac{3}{2n}\right) + x z^2 \left(\frac{n^2}{2} - \frac{n}{2} - 1 + \frac{1}{n}\right) \Big] \\
&\quad + \frac{1}{2n} x z^2 \left[2nx(\log(z) + 2\log(x)) + x(n-2) - n^2 - n + 2\right],
\end{aligned}
$$

where x and z are functions of (m_μ, σ_μ^2), which are provided in the Supplementary Appendix.

This expression for aggregate income depends only on the parameter estimates used in the baseline model without imposing any symmetry on the IO coefficients. It is similar to the one of the baseline model but includes additional terms that capture the effect of asymmetric linkages. We use this expression to predict cross-country income differences in this more general setting. The results are presented in columns (3) - (4) of Table 3.

Appendix C: Proofs for the benchmark model and its extensions

Proposition 1 and formulae for aggregate output in the main text are particular cases of Proposition 2 that applies in a generic setting – with imported intermediates, division of labor into skilled and unskilled labor inputs and unequal demand shares. A brief description of this economy, as well as Proposition 2 and its proof are provided below.

- The technology of each of n competitive sectors is Cobb-Douglas with constant returns to

scale. Namely, the output of sector i, denoted by q_i, is

$$q_i = \Lambda_i \left(k_i^{\alpha} u_i^{\delta} s_i^{1-\alpha-\delta}\right)^{1-\gamma_i-\sigma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}} \cdot f_{1i}^{\sigma_{1i}} f_{2i}^{\sigma_{2i}} \cdot \ldots \cdot f_{ni}^{\sigma_{ni}},$$

where s_i and u_i are the amounts of skilled and unskilled labor, d_{ji} is the quantity of the domestic good j and f_{ji} is the quantity of the imported good j used by sector i. $\gamma_i = \sum_{j=1}^{n} \gamma_{ji}$ and $\sigma_i = \sum_{j=1}^{n} \sigma_{ji}$ are the respective shares of domestic and imported intermediate goods in the total input use of sector i and α, δ, $1-\alpha-\delta$ are the respective shares of capital, unskilled and skilled labor in the remainder of the inputs.

- A good produced by sector i can be used for final consumption, y_i, or as an intermediate good:

$$y_i + \sum_{j=1}^{n} d_{ij} = q_i \qquad i = 1:n$$

- Final consumption goods are aggregated into a single final good through another Cobb-Douglas production function:

$$Y = y_1^{\beta_1} \cdot \ldots \cdot y_n^{\beta_n},$$

where $\beta_i \geq 0$ for all i and $\sum_{i=1}^{n} \beta_i = 1$.

- This aggregate final good can itself be used in one of two ways, as households' consumption or export to the rest of the world:

$$Y = C + X.$$

- Exports pay for the imported intermediate goods, and we impose a balanced trade condition:

$$X = \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{p}_j f_{ji},$$

where $\overline{p}_j$ is the exogenous world price of the imported intermediate goods.

- Households finance their consumption through income:

$$C = w_U U + w_S S + rK.$$

- The total supply of physical capital, unskilled and skilled labor are fixed at the exogenous levels of K, U and S, respectively:

$$\begin{aligned} \sum_{i=1}^{n} k_i &= K, \\ \sum_{i=1}^{n} u_i &= U, \\ \sum_{i=1}^{n} s_i &= S. \end{aligned}$$

For this "generic" economy, the competitive equilibrium is defined by analogy with the definition in section 3.1. The solution is described by Proposition 2.

Proposition 2. *There exists a unique competitive equilibrium. In this equilibrium, the logarithm*

of GDP per capita, $y = \log(Y/(U+S))$, is given by

$$
\begin{aligned}
y \;=\; & \frac{1}{\sum_{i=1}^{n}\mu_i(1-\gamma_i-\sigma_i)}\Bigg[\sum_{i=1}^{n}\mu_i\lambda_i + \sum_{i=1}^{n}\sum_{j\ s.t.\gamma_{ij}\neq 0}\mu_i\gamma_{ji}\log\gamma_{ji} + \sum_{i=1}^{n}\sum_{j\ s.t.\sigma_{ij}\neq 0}\mu_i\sigma_{ji}\log\sigma_{ji} - \\
& -\sum_{i=1}^{n}\sum_{j=1}^{n}\mu_i\sigma_{ji}\log\bar{p_j} + \sum_{i=1}^{n}\beta_i log\beta_i + \sum_{i=1}^{n}\mu_i(1-\gamma_i-\sigma_i)log(1-\gamma_i-\sigma_i)\Bigg] + \log\left(1+\sum_{i=1}^{n}\sigma_i\bar{\mu_i}\right) + \\
& +\alpha\log K + \delta logU + (1-\alpha-\delta)logS - \log(U+S). \qquad \text{(A-1)}
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{\mu} &= \{\mu_i\}_i = [\boldsymbol{I}-\boldsymbol{\Gamma}]^{-1}\boldsymbol{\beta}, && n\times 1 \text{ vector of multipliers} \\
\boldsymbol{\lambda} &= \{\lambda_i\}_i = \{\log\Lambda_i\}_i, && n\times 1 \text{ vector of sectoral log-productivity coefficients} \\
\bar{\boldsymbol{\mu}} &= \{\bar{\mu_i}\}_i = [\boldsymbol{I}-\bar{\boldsymbol{\Gamma}}]^{-1}\boldsymbol{\beta}, && n\times 1 \text{ vector of multipliers corresponding to } \bar{\boldsymbol{\Gamma}} \\
\bar{\boldsymbol{\Gamma}} &= \{\bar{\gamma_{ji}}\}_{ji} = \{\beta_j\sigma_i+\gamma_{ji}\}_{ji}, && n\times n \text{ input-output matrix adjusted for trade}
\end{aligned}
$$

Proof. Part I: Calculation of $\log w_U$.
Consider the profit maximization problems of a representative firm in the final goods market and in each sector. For a representative firm in the final goods market the FOCs allocate to each good a spending share that is proportional to the good's demand share β_i:

$$
p_i y_i = \beta_i Y = \beta_i(C+X) = \beta_i(w_U U + w_S S + rK) + \beta_i\sum_{i=1}^{n}\sum_{j=1}^{n}\bar{p}_j m_{ji} \quad \forall i\in 1:n
$$

where the price of the final good is normalized to 1, $p=1$. For a firm in sector i the FOCs are:

$$
\begin{aligned}
\alpha(1-\gamma_i-\sigma_i)\frac{p_i q_i}{r} &= k_i && \text{(A-2)} \\
\delta(1-\gamma_i-\sigma_i)\frac{p_i q_i}{w_U} &= u_i && \text{(A-3)} \\
(1-\alpha-\delta)(1-\gamma_i-\sigma_i)\frac{p_i q_i}{w_S} &= s_i && \text{(A-4)} \\
\gamma_{ji}\frac{p_i q_i}{p_j} &= d_{ji} \quad j\in 1:n && \text{(A-5)} \\
\sigma_{ji}\frac{p_i q_i}{\bar{p_j}} &= f_{ji} \quad j\in 1:n && \text{(A-6)}
\end{aligned}
$$

Substituting the left-hand side of these equations for the values of k_i, u_i, s_i, $\{d_{ji}\}$ and $\{f_{ji}\}$ in firm i's log-production technology and simplifying the obtained expression, we derive:

$$
\begin{aligned}
\delta\log w_U \;=\; & \frac{1}{1-\gamma_i-\sigma_i}\Big(\lambda_i + \log p_i - \sum_{j=1}^{n}\gamma_{ji}\log p_j + \sum_{j\text{ s.t. }\gamma_{ji}\neq 0}\gamma_{ji}\log\gamma_{ji} - \\
& -\sum_{j=1}^{n}\sigma_{ji}\log\bar{p_j} + \sum_{j\text{ s.t. }\sigma_{ji}\neq 0}\sigma_{ji}\log\sigma_{ji}\Big) - \alpha\log r - (1-\alpha-\delta)\log(w_S) + \\
& +\log(1-\gamma_i-\sigma_i) + \alpha\log(\alpha) + \delta\log\delta + (1-\alpha-\delta)\log(1-\alpha-\delta) \qquad \text{(A-7)}
\end{aligned}
$$

Next, we use FOCs (A-2) – (A-6) and market clearing conditions for labor and capital to express

r and w_S in terms of w_U:

$$w_U = \frac{1}{U}\delta\sum_{i=1}^{n}(1-\gamma_i-\sigma_i)(p_iq_i) \tag{A-8}$$

$$w_S = \frac{1}{S}(1-\alpha-\delta)\sum_{i=1}^{n}(1-\gamma_i-\sigma_i)(p_iq_i) = \frac{w_UU}{S}\frac{1-\alpha-\delta}{\delta} \tag{A-9}$$

$$r = \frac{1}{K}\alpha\sum_{i=1}^{n}(1-\gamma_i-\sigma_i)(p_iq_i) = \frac{w_uU}{K}\frac{\alpha}{\delta} \tag{A-10}$$

Substituting these values of r and w_S in (A-7) we obtain:

$$\begin{aligned}\log w_U = {} & \frac{1}{1-\gamma_i-\sigma_i}\Big(\lambda_i+\log p_i-\sum_{j=1}^{n}\gamma_{ji}\log p_j+\sum_{j \text{ s.t. } \gamma_{ji}\neq 0}\gamma_{ji}\log\gamma_{ji}-\sum_{j=1}^{n}\sigma_{ji}\log\bar{p}_j+\\ & +\sum_{j \text{ s.t. } \sigma_{ji}\neq 0}\sigma_{ji}\log\sigma_{ji}\Big)+\alpha\log K-(1-\delta)\log U+(1-\alpha-\delta)\log S+\log(1-\gamma_i-\sigma_i)+\log\delta\end{aligned}$$

Multiplying this equation by the ith element of the vector $\boldsymbol{\mu'Z} = \boldsymbol{\beta'1'}[\boldsymbol{I}-\boldsymbol{\Gamma'}]^{-1}\cdot\boldsymbol{Z}$, where $\boldsymbol{Z}$ is a diagonal matrix with $Z_{ii} = 1-\gamma_i-\sigma_i$, and summing over all sectors i gives

$$\begin{aligned}&\sum_{i=1}^{n}\mu_i(1-\gamma_i-\sigma_i)\log w_U = \sum_{i=1}^{n}\mu_i\lambda_i+\sum_{i=1}^{n}\beta_i\log p_i+\sum_{i=1}^{n}\sum_{j \text{ s.t. } \gamma_{ji}\neq 0}\mu_i\gamma_{ji}\log\gamma_{ji}-\\ &-\sum_{j=1}^{n}\mu_i\sigma_{ji}\log\bar{p}_j+\sum_{j \text{ s.t. } \sigma_{ji}\neq 0}\mu_i\sigma_{ji}\log\sigma_{ji}+\sum_{i=1}^{n}\mu_i(1-\gamma_i-\sigma_i)\log(1-\gamma_i-\sigma_i)+\\ &+\sum_{i=1}^{n}\mu_i(1-\gamma_i-\sigma_i)\left(\alpha\log K-(1-\delta)\log U+(1-\alpha-\delta)\log S+\log\delta\right)\end{aligned}$$

Next, we use the relation between the price of the final good p (normalized to 1) and prices of each sector goods, derived from a profit maximization of the final good firm that has Cobb-Douglas technology.[63] This relation implies that $\prod_{i=1}^{n}(p_i)^{\beta_i} = \prod_{i=1}^{n}(\beta_i)^{\beta_i}$, so that $\sum_{i=1}^{n}\beta_i\log p_i = \sum_{i=1}^{n}\beta_i\log\beta_i$, and the above equation becomes:

$$\begin{aligned}\log w_U = {} & \frac{1}{\sum_{i=1}^{n}\mu_i(1-\gamma_i-\sigma_i)}\Big[\sum_{i=1}^{n}\mu_i\lambda_i+\sum_{i=1}^{n}\beta_i\log\beta_i+\sum_{i=1}^{n}\sum_{j \text{ s.t. } \gamma_{ji}\neq 0}\mu_i\gamma_{ji}\log\gamma_{ji}-\\ & -\sum_{j=1}^{n}\mu_i\sigma_{ji}\log\bar{p}_j+\sum_{j \text{ s.t. } \sigma_{ji}\neq 0}\mu_i\sigma_{ji}\log\sigma_{ji}+\sum_{i=1}^{n}\mu_i(1-\gamma_i-\sigma_i)\log(1-\gamma_i-\sigma_i)\Big]+\\ & +\alpha\log K-(1-\delta)\log U+(1-\alpha-\delta)\log S+\log\delta\end{aligned} \tag{A-11}$$

Part II: Calculation of y.

Recall that our ultimate goal is to find $y = \log(Y/(U+S)) = \log(C+X)-\log(U+S)$. From the households' budget constraint and from the balanced trade condition, $C+X = w_UU+w_SS+rK+\sum_{i=1}^{n}\sum_{j=1}^{n}\bar{p}_jf_{ji}$, where in the last term, $\bar{p}_jf_{ji} = \sigma_{ji}p_iq_i$ (cf. (A-6)). Below we show that p_iq_i can be expressed as a product of w_UU+w_SS+rK and another term that involves structural

[63] Profit maximization of the final good's firm implies that $\frac{\partial Y}{\partial y_i} = \frac{p_i}{p}$. On the other hand, since $Y = y_1^{\beta_1}\cdot\ldots\cdot y_n^{\beta_n}$, we have $\frac{\partial Y}{\partial y_i} = \beta_i\frac{Y}{y_i}$. Hence, $\beta_i\frac{Y}{y_i} = \frac{p_i}{p}$, or $y_i = \beta_i\frac{pY}{p_i}$. Substituting this in the production technology of the firm in final good market, we obtain:

$$Y = \prod_{i=1}^{n}\left(\beta_i\frac{pY}{p_i}\right)^{\beta_i} = pY\prod_{i=1}^{n}\left(\beta_i\frac{1}{p_i}\right)^{\beta_i}.$$

So, $p\prod_{i=1}^{n}\left(\beta_i\frac{1}{p_i}\right)^{\beta_i} = 1$. Now, since we used the normalization $p = 1$, it must be that $\prod_{i=1}^{n}(p_i)^{\beta_i} = \prod_{i=1}^{n}(\beta_i)^{\beta_i}$.

characteristics. Then using (A-9) and (A-10), we obtain the representation of $C + X$ as a product of w_U and another term determined by exogenous variables. This representation, together with (A-11), will then allow us to solve for y.

Consider the resource constraint for sector j, with both sides multiplied by p_j:

$$p_j y_j + \sum_{i=1}^{n} p_j d_{ji} = p_j q_j$$

Using FOCs of the profit maximization problem of the final good's firm and a firm in sector i, this can be written as:

$$\beta_j Y + \sum_{i=1}^{n} \gamma_{ji} p_i q_i = p_j q_j$$

or

$$\beta_j (w_U U + w_S S + rK) + \sum_{i=1}^{n} \gamma_{ji} p_i q_i + \beta_j \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ji} p_i q_i = p_j q_j.$$

Using the fact that $\sum_{j=1}^{n} \sigma_{ji} = \sigma_i$ and combining terms, we obtain:

$$\beta_j (w_U U + w_S S + rK) + \sum_{i=1}^{n} \left[\beta_j \sigma_i + (1 - \tau_i)\gamma_{ji}\right] p_i q_i = p_j q_j.$$

Denote by $a_j = p_j q_j$ and by $\bar{\gamma}_{ji} = \beta_j \sigma_i + \gamma_{ji}$. Then the above equation in the matrix form is:

$$(w_U U + w_S S + rK)\, \boldsymbol{\beta} + \bar{\boldsymbol{\Gamma}} \boldsymbol{a} = \boldsymbol{a}$$

where $\boldsymbol{\beta} = (\beta_1, .., \beta_n)'$, $\bar{\boldsymbol{\Gamma}} = \{\bar{\gamma}_{ji}\}_{ji}$ and $\boldsymbol{a} = \{a_j\}_j$. Hence,

$$\boldsymbol{a} = (\boldsymbol{I} - \bar{\boldsymbol{\Gamma}})^{-1} (w_U U + w_S S + rK)\, \boldsymbol{\beta} = (w_U U + w_S S + rK) \bar{\boldsymbol{\mu}}$$

where $\bar{\boldsymbol{\mu}} = \left(\boldsymbol{I} - \bar{\boldsymbol{\Gamma}}\right)^{-1} \boldsymbol{\beta}$.[64] So, $a_i = p_i q_i = (w_U U + w_S S + rK)\bar{\mu}_i$ and therefore,

$$\begin{aligned} Y &= C + X = w_U U + w_S S + rK + \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ji} p_i q_i = \\ &= (w_U U + w_S S + rK) \left(1 + \sum_{i=1}^{n} \sigma_i \bar{\mu}_i\right) \end{aligned}$$

Using (A-9) and (A-10), this leads to

$$Y = \frac{w_U U}{\delta} \left(1 + \sum_{i=1}^{n} \sigma_i \bar{\mu}_i\right).$$

so that

$$y = \log Y - \log(U + S) = \log w_U + \log U + \log \left(1 + \sum_{i=1}^{n} \sigma_i \bar{\mu}_i\right) - \log \delta - \log(U + S).$$

[64] Notice that $\left(\boldsymbol{I} - \bar{\boldsymbol{\Gamma}}\right)^{-1}$ exists because the sum of elements in each column of $\bar{\boldsymbol{\Gamma}}$ is less than 1 for any $\sigma_i + \gamma_i < 1$: $\sum_{j=1}^{n} (\beta_j \sigma_i + \gamma_{ji}) = \sigma_i + \gamma_i < 1$.

Finally, substituting $\log w_U$ with (A-11) yields our result:

$$\begin{aligned} y &= \frac{1}{\sum_{i=1}^n \mu_i(1-\gamma_i-\sigma_i)}\Bigg[\sum_{i=1}^n \mu_i\lambda_i + \sum_{i=1}^n \sum_{j \text{ s.t.} \gamma_{ij}\neq 0} \mu_i\gamma_{ji}\log\gamma_{ji} + \sum_{i=1}^n \sum_{j \text{ s.t.} \sigma_{ij}\neq 0} \mu_i\sigma_{ji}\log\sigma_{ji} - \\ &- \sum_{i=1}^n\sum_{j=1}^n \mu_i\sigma_{ji}\log\bar{p_j} + \sum_{i=1}^n \mu_i(1-\gamma_i-\sigma_i)log(1-\gamma_i-\sigma_i) + \sum_{i=1}^n \beta_i log\beta_i\Bigg] + \log\left(1+\sum_{i=1}^n \sigma_i\bar{\mu_i}\right) + \\ &+ \alpha\log K + \delta logU + (1-\alpha-\delta)logS - \log(U+S). \end{aligned}$$

This completes the proof. □

Application of Proposition 2 to the case of the benchmark economy:

Proof. (Proposition 1) In case of our benchmark economy, we assume that: i) there is no distinction between skilled and unskilled labor, so that $\delta = 1-\alpha$ and the total supply of labor is normalized to 1; ii) demand shares for all final goods are the same, that is, $\beta_i = \frac{1}{n}$ for all i; iii) the economies are closed, so that no imported intermediate goods are used in sectors' production, that is, $\sigma_{ji} = 0$ for all $i, j \in 1:n$ and $\sigma_i = 0$ for all i. This simplifies the expression for y in Proposition 2 and produces:

$$y = \frac{1}{\sum_{i=1}^n \mu_i(1-\gamma_i)}\left(\sum_{i=1}^n \mu_i\lambda_i + \sum_{i=1}^n \sum_{j \text{ s.t. } \gamma_{ji}\neq 0} \mu_i\gamma_{ji}\log\gamma_{ji} + \sum_{i=1}^n \mu_i(1-\gamma_i)log(1-\gamma_i) - \log n\right) + \alpha\log K.$$

Now, observe that $\sum_{i=1}^n \mu_i(1-\gamma_i) = \mathbf{1}'[\boldsymbol{I}-\boldsymbol{\Gamma}]\cdot\frac{1}{n}[\boldsymbol{I}-\boldsymbol{\Gamma}]^{-1}\mathbf{1} = \frac{1}{n}\mathbf{1}'\mathbf{1} = 1$. Then the expression simplifies even further and leads to the result of Proposition 1:

$$y = \sum_{i=1}^n \mu_i\lambda_i + \sum_{i=1}^n \sum_{j \text{ s.t. } \gamma_{ji}\neq 0} \mu_i\gamma_{ji}\log\gamma_{ji} + \sum_{i=1}^n \mu_i(1-\gamma_i)log(1-\gamma_i) - \log n + \alpha\log K,$$

where

$$\begin{aligned} \boldsymbol{\mu} &= \{\mu_i\}_i = \frac{1}{n}[\boldsymbol{I}-\boldsymbol{\Gamma}]^{-1}\mathbf{1}, && n\times 1 \text{ vector of multipliers} \\ \boldsymbol{\lambda} &= \{\lambda_i\}_i = \{\log\Lambda_i\}_i, && n\times 1 \text{ vector of sectoral log-productivity coefficients.} \end{aligned}$$

□

Appendix D: Extensions of the benchmark model

A-3 Cross-country differences in final demand structure

Consider now the economy that is identical to our benchmark economy in all but demand shares for final goods. Namely, let us generalize the production function for the aggregate final good to accommodate arbitrary, country-sector-specific demand shares:

$$Y = y_1^{\beta_1}\cdot\ldots\cdot y_n^{\beta_n},$$

where $\beta_i \geq 0$ for all i and $\sum_{i=1}^n \beta_i = 1$. As before, suppose that this aggregate final good is fully allocated to households' consumption, that is, $Y = C$.

Using the generic expression for aggregate output (A-1) of Proposition 2 and adopting this expression to the case of our economy here, we obtain the following formula for y:

$$\begin{aligned} y &= \sum_{i=1}^n \mu_i\lambda_i + \sum_{i=1}^n \sum_{j \text{ s.t. } \gamma_{ji}\neq 0} \mu_i\gamma_{ji}\log\gamma_{ji} + \sum_{i=1}^n \mu_i(1-\gamma_i)log(1-\gamma_i) + \\ &+ \sum_{i=1}^n \beta_i\log(\beta_i) + \alpha\log K. \end{aligned}$$

In this formula the vector of sectoral multipliers is defined differently than before, to account for the arbitrary demand shares. The new vector of multipliers is $\boldsymbol{\mu} = \{\mu_i\}_i = [\boldsymbol{I} - \boldsymbol{\Gamma}]^{-1}\boldsymbol{\beta}$. Its interpretation, however, is identical to the one before: each sectoral multiplier μ_i reveals how a change in productivity (or distortion) of sector i affects the overall value added in the economy.

Given this expression for y, we now derive the approximate representation of the aggregate output to be used in our empirical analysis. For this purpose, we employ the same set of simplifying assumptions as before, which results in:

$$\begin{aligned} y &= \sum_{i=1}^{n} \mu_i \Lambda_i^{rel} + \sum_{i=1}^{n} \mu_i \gamma \log(\widehat{\gamma}) + \log(1-\gamma) + \sum_{i=1}^{n} \beta_i \log(\beta_i) + \alpha \log(K) - \\ &- (1+\gamma) + \sum_{i=1}^{n} \mu_i \log(\Lambda_i^{US}). \end{aligned} \tag{A-12}$$

Following the same procedure as earlier, we use this expression to find the predicted value of y. First, we estimate the distribution of (μ_i, Λ_i^{rel}) in every country. We find that even though the definition of sectoral multipliers is now different from the one in our benchmark model, the distribution of the pair (μ_i, Λ_i^{rel}) is still log-Normal.[65] Then, using the estimates of the parameters of this distribution, $\boldsymbol{m}$ and $\boldsymbol{\Sigma}$, together with the equations describing the relationship between Normal and log-Normal distributions (see eq. (9)), we find the predicted aggregate output $E(y)$ as a function of these parameters:[66]

$$\begin{aligned} E(y) &= n e^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} + (1+\gamma)(\gamma \log(\widehat{\gamma}) - 1) + \log(1-\gamma) + \\ &+ \sum_{i=1}^{n} \beta_i \log(\beta_i) + \alpha \log(K) + e^{m_\mu + 1/2\sigma_\mu^2} \sum_{i=1}^{n} \log\left(\Lambda_i^{US}\right). \end{aligned} \tag{A-13}$$

The resulting expression for $E(y)$ is similar to (10) in our benchmark model.

A-4 Imported intermediates

Another extension of the benchmark model allows for trade between countries. The traded goods are used as inputs in production of the n competitive sectors, so that both domestic and imported intermediate goods are employed in sectors' production technology. Then the output of sector i is determined by the following production function:

$$q_i = \Lambda_i \left(k_i^{\alpha} l_i^{1-\alpha}\right)^{1-\gamma_i-\sigma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}} \cdot f_{1i}^{\sigma_{1i}} f_{2i}^{\sigma_{2i}} \cdot \ldots \cdot f_{ni}^{\sigma_{ni}}, \tag{A-14}$$

where d_{ji} is the quantity of the domestic good j used by sector i, and f_{ji} is the quantity of the imported intermediate good j used by sector i. The imported intermediate goods are assumed to be different, so that domestic and imported goods are not perfect substitutes. Also, with a slight abuse of notation, we assume that there are n different intermediate goods that can be imported.[67] The exponents γ_{ji}, $\sigma_{ji} \in [0,1)$ represent the respective shares of domestic and imported good j in the technology of firms in sector i, and $\gamma_i = \sum_{j=1}^{n} \gamma_{ji}$, $\sigma_i = \sum_{j=1}^{n} \sigma_{ji} \in (0,1)$ are the total shares of domestic and imported intermediate goods, respectively.

As in our benchmark economy, each domestically produced good can be used for final consumption, y_i, or as an intermediate good, and all final consumption goods are aggregated into a single final good through a Cobb-Douglas production function, $Y = y_1^{\frac{1}{n}} \cdot \ldots \cdot y_n^{\frac{1}{n}}$. Now, in case of an open economy considered here, the aggregate final good is used not only for households' consumption but also for export to the rest of the world; that is, $Y = C + X$. The exports pay for the imported

[65] In fact, differently from the benchmark model, the distribution is "exactly" log-Normal and not *truncated* log-Normal as it was before.

[66] As before, we also assume for simplicity that all other variables on the right-hand side of (A-12) are non-random.

[67] This is consistent with the specification of input-output tables in our data.

intermediate goods and are defined by the balanced trade condition:

$$X = \sum_{i=1}^{n}\sum_{j=1}^{n} \overline{p}_j f_{ji}, \tag{A-15}$$

where $\overline{p}_j$ is the exogenous world price of the imported intermediate goods. Note that the balanced trade condition is reasonable to impose if we consider our static model as describing the steady state of the model.

Aggregate output y is determined by equation (A-1) of Proposition 2, adopted to our framework here:

$$\begin{aligned} y &= \frac{1}{\sum_{i=1}^{n}\mu_i(1-\gamma_i-\sigma_i)}\Big(\sum_{i=1}^{n}\mu_i\lambda_i + \sum_{i=1}^{n}\sum_{j \text{ s.t. } \gamma_{ji}\neq 0}\mu_i\gamma_{ji}\log\gamma_{ji} + \sum_{i=1}^{n}\sum_{j \text{ s.t.} \sigma_{ji}\neq 0}\mu_i\sigma_{ji}\log\sigma_{ji} - \\ &- \sum_{i=1}^{n}\sum_{j=1}^{n}\mu_i\sigma_{ji}\log\bar{p}_j + \sum_{i=1}^{n}\mu_i(1-\gamma_i-\sigma_i)log(1-\gamma_i-\sigma_i) - \log n\Big) + \\ &+ \log\left(1+\sum_{i=1}^{n}\sigma_i\bar{\mu}_i\right) + \alpha\log K, \end{aligned}$$

where vector $\{\bar{\mu}_i\}_i = \frac{1}{n}[\boldsymbol{I}-\bar{\boldsymbol{\Gamma}}]^{-1}\mathbf{1}$ is a vector of multipliers corresponding to $\bar{\boldsymbol{\Gamma}}$ and $\bar{\boldsymbol{\Gamma}} = \{\bar{\gamma_{ji}}\}_{ji} = \{\frac{1}{n}\sigma_i + \gamma_{ji}\}_{ji}$ is an input-output matrix adjusted for shares of imported intermediate goods.[68]

In the empirical analysis we use an approximate representation of aggregate output, where a range of simplifying assumptions is imposed. First, to be able to compare the results with the results of the benchmark model, we employ the same assumptions on in-degree and elements of matrix $\boldsymbol{\Gamma}$. Second, in the new framework with imported intermediates we also impose some conditions on imports. We assume that the total share of imported intermediate goods used by any sector of a country is sufficiently small and identical across sectors, that is, $\sigma_i = \sigma$ for any sector i.[69] We also regard any non-zero elements of the vector of import shares of sector i as the same, equal to $\widehat{\sigma}_i$ (such that $\sum_{j \text{ s.t.} \sigma_{ji}\neq 0}\widehat{\sigma}_i = \sigma$). Then we obtain the following approximation for the aggregate output y:

$$\begin{aligned} y &= \frac{1}{(1-\sigma(1+\gamma))}\Big(\sum_{i=1}^{n}\mu_i\Lambda_i^{rel} + \sum_{i=1}^{n}\mu_i\gamma\log\widehat{\gamma} + \sum_{i=1}^{n}\mu_i\sigma\log\widehat{\sigma}_i - \\ &- \sum_{i=1}^{n}\mu_i\widehat{\sigma}_i\sum_{j \text{ s.t.} \sigma_{ji}\neq 0}\log\bar{p}_j - \log n\Big) + \log(1-\gamma-\sigma) + \sigma\left(1+\gamma+\sigma\right) + \alpha\log K - \\ &- \frac{1+\gamma}{(1-\sigma(1+\gamma))} + \frac{1}{(1-\sigma(1+\gamma))}\sum_{i=1}^{n}\mu_i\log(\Lambda_i^{US}). \end{aligned}$$

Now, using the relationship between Normal and log-Normal distributions (see eq. (9)), we can derive the predicted aggregate output $E(y)$ in terms of the parameters of the bivariate log-Normal

[68] Observe that $\left(\boldsymbol{I}-\bar{\boldsymbol{\Gamma}}\right)^{-1}$ exists because the maximal eigenvalue of $\bar{\boldsymbol{\Gamma}}$ is bounded above by 1. The latter is implied by the Frobenius theory of non-negative matrices, that says that the maximal eigenvalue of $\bar{\boldsymbol{\Gamma}}$ is bounded above by the largest column sum of $\bar{\boldsymbol{\Gamma}}$, which in our case is smaller than 1 as soon as $\sigma_i+\gamma_i < 1$: $\sum_{j=1}^{n}\left(\frac{1}{n}\sigma_i + \gamma_{ji}\}_{ji}\right) = \sigma_i+\gamma_i < 1$.

[69] This allows approximating $\log\left(1+\sum_{i=1}^{n}\sigma_i\bar{\mu}_i\right)$ with $\sigma\sum_{i=1}^{n}\bar{\mu}_i = \sigma\left(1+\gamma+\sigma\right)$, where the equality follows from $\bar{\mu}_i \approx \mu_i + \frac{1}{n}\sum_{j=1}^{n}\frac{1}{n}\sigma_j$. The latter, in turn, is a result of the approximation of $\{\bar{\mu}_i\}_i$ by the first elements of the convergent power series $\frac{1}{n}\left(\sum_{k=0}^{+\infty}\bar{\boldsymbol{\Gamma}}^k\right)\mathbf{1}$ and the analogous approximation for $\{\mu_i\}_{i=1}^{n}$ (see section 2.3).

distribution of (μ_i, Λ_i^{rel}):

$$\begin{aligned} E(y) &= \frac{n}{(1-\sigma(1+\gamma))} e^{m_\mu + m_\Lambda 1/2(\sigma_\mu^2+\sigma_\Lambda^2)+\sigma_{\mu,\Lambda}} + \\ &+ \frac{1}{(1-\sigma(1+\gamma))} \sum_{i=1}^{n} \left(\sigma \log \widehat{\sigma}_i - \widehat{\sigma}_i \sum_{j=1, j \text{ s.t. } \sigma_{ji}\neq 0}^{n} \log \bar{p}_j + \log(\Lambda_i^{US}) \right) e^{m_\mu + 1/2\sigma_\mu^2} + \\ &+ \frac{(1+\gamma)\gamma \log \widehat{\gamma}}{(1-\sigma(1+\gamma))} - \frac{\log n}{(1-\sigma(1+\gamma))} + \log(1-\gamma-\sigma) + \sigma\,(1+\gamma+\sigma) + \alpha \log(K) - \frac{1+\gamma}{(1-\sigma(1+\gamma))}. \end{aligned}$$

We bring this expression to data and evaluate predicted output in all countries of our data sample. We note, however, that the vector of world prices of the imported intermediates $\{\overline{p}_j\}_{j=1}^n$ is not provided in the data. Then to make the comparison of aggregate income in different countries possible, we assume that for any sector i, the value of $\widehat{\sigma}_i \sum_{j=1, j \text{ s.t. } \sigma_{ji}\neq 0}^{n} \log \bar{p}_j$ is the same across countries, so that this term cancels out when the difference in countries' predicted output is considered. For this purpose we assume that in all countries, the vector of shares of the imported intermediate goods used by sector i is the same and that all countries face the same vector of prices of the imported intermediate goods $\{\overline{p}_j\}_{j=1}^n$.

A-5 Skilled labor

Consider the economy of our benchmark model where we introduce the distinction between skilled and unskilled labor. This distinction implies that the technology of each sector $i \in 1:n$ in every country can be described by the following Cobb-Douglas function:

$$q_i = \Lambda_i \left(k_i^\alpha u_i^\delta s_i^{1-\alpha-\delta} \right)^{1-\gamma_i-\sigma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}}, \tag{A-16}$$

where s_i and u_i denote the amounts of skilled and unskilled labor used by sector i, $\gamma_i = \sum_{j=1}^n \gamma_{ji}$ is the share of intermediate goods in the total input use of sector i and α, δ, $1-\alpha-\delta \in (0,1)$ are the respective shares of capital, unskilled and skilled labor in the remainder of the inputs. The total supply of skilled and unskilled labor in the economy is fixed at the exogenous levels of S and U, respectively.

In this case, the logarithm of the value added per capita, $y = \log(Y/(U+S))$, is given by the expression (A-1) of Proposition 2, adopted to our framework here. In fact, it is only slightly different from the expression for y in our benchmark model (cf. Proposition 1), where $\delta = 0$ and the total supply of labor is normalized to 1. With skilled and unskilled labor, the aggregate output per capita is given by:

$$\begin{aligned} y &= \sum_{i=1}^{n} \mu_i \lambda_i + \sum_{i=1}^{n} \sum_{j \text{ s.t. } \gamma_{ji}\neq 0} \mu_i \gamma_{ji} \log \gamma_{ji} + \sum_{i=1}^{n} \mu_i (1-\gamma_i) log(1-\gamma_i) - \log n + \\ &+ \alpha \log K + \delta \log U + (1-\alpha-\delta) \log S - \log(U+S). \end{aligned}$$

Then the approximate representation of y is also similar to the corresponding representation of y in the benchmark model (cf. (10)):

$$\begin{aligned} y &= \sum_{i=1}^{n} \mu_i \Lambda_i^{rel} + \sum_{i=1}^{n} \mu_i \gamma \log(\widehat{\gamma}) + \log(1-\gamma) - \log n + \alpha \log(K) + \\ &+ \delta \log U + (1-\alpha-\delta) \log S - \log(U+S) - (1+\gamma) + \sum_{i=1}^{n} \mu_i \log(\Lambda_i^{US}), \end{aligned} \tag{A-17}$$

where the same assumptions and notation as before apply.

We now employ this representation of y to find the predicted value of aggregate output $E(y)$. Note that since the new framework, with skilled and unskilled labor, does not modify the definition

of the sectoral multipliers, the distribution of the pair (μ_i, Λ_i^{rel}) in every country remains the same. It is a bivariate log-Normal distribution with parameters $\boldsymbol{m}$ and $\boldsymbol{\Sigma}$ that have been estimated for our benchmark model. Using these parameters, together with the equations describing the relationship between Normal and log-Normal distributions (see eq. (9)), we derive the expression for the predicted aggregate output $E(y)$ in terms of the estimated parameters:

$$\begin{aligned} E(y) &= ne^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} + (1+\gamma)(\gamma \log(\widehat{\gamma}) - 1) + \log(1-\gamma) - \log n + \alpha \log(K) + \\ &+ \delta \log U + (1-\alpha-\delta)\log S - \log(U+S) + e^{m_\mu + 1/2\sigma_\mu^2} \sum_{i=1}^{n} \log\left(\Lambda_i^{US}\right). \end{aligned} \tag{A-18}$$

This equation for the predicted aggregate output is analogous to the equation (10) that we employed in our estimation of the benchmark model.

Appendix E: Additional Figures and Tables

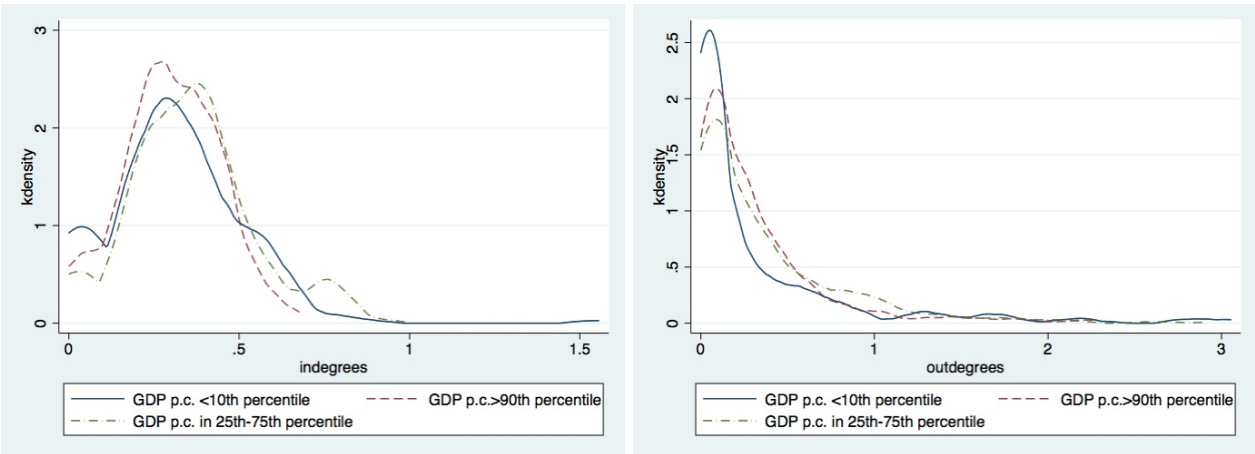

Figure A-2: Distribution of sectoral in-degrees (left) and out-degrees (right) (GTAP sample)

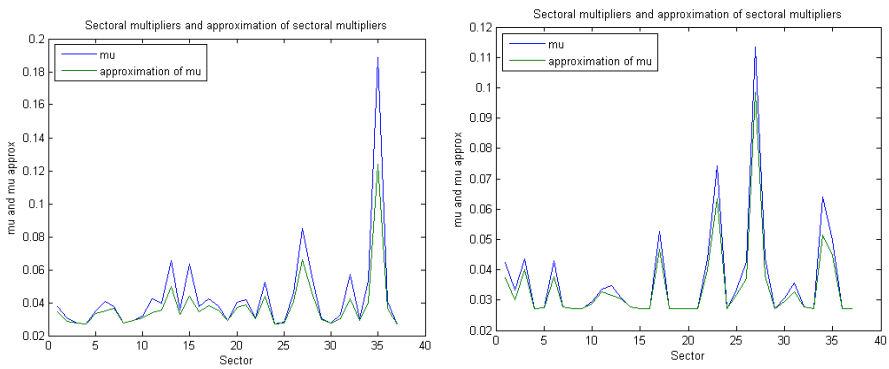

Figure A-3: Sectoral multipliers and their approximation by eq. (7) in Germany (left) and Botswana (right). GTAP sample.

Table A-1: Countries: WIOD Sample

countries	
AUS	IDN
AUT	IND
BEL	IRL
BGR	ITA
BRA	LTU
CAN	LVA
CHN	MEX
CYP	MLT
CZE	NLD
DEU	POL
DNK	PRT
ESP	ROM
EST	RUS
FIN	SVK
FRA	SVN
GBR	SWE
GRC	TUR
HUN	USA

Table A-2: Countries: GTAP Sample

countries	
ALB	LTU
ARG	LUX
AUS	LVA
AUT	MDG
BEL	MEX
BGD	MLT
BGR	MOZ
BRA	MWI
BWA	MYS
CAN	NLD
CHE	NZL
CHL	PER
CHN	PHL
COL	POL
CYP	PRT
CZE	ROM
DEU	RUS
DNK	SGP
ESP	SVK
EST	SVN
FIN	SWE
FRA	THA
GBR	TUN
GRC	TUR
HKG	TWN
HRV	TZA
HUN	UGA
IDN	URY
IND	USA
IRL	VEN
ITA	VNM
JPN	ZAF
KOR	ZMB
LKA	ZWE

Table A-3: Sector List

WIOD sectors		GTAP sectors	
1	Agriculture	1	Agriculture
2	Mining	2	Coal
3	Food	3	Oil
4	Textiles	4	Gas
5	Leather	5	Mining
6	Wood	6	Food
7	Paper	7	Textiles
8	Refining	8	Apparel
9	Chemicals	9	Leather
10	Plastics	10	Wood
11	Minerals	11	Paper
12	Metal products	12	Refining
13	Machinery	13	Chemicals
14	Elec. equip.	14	Minerals
15	Transport equip.	15	Iron
16	Manufacturing nec	16	Oth. metals
17	Electricity	17	Metal products
18	Construction	18	Cars
19	Car retail.	19	Transport equip.
20	Wholesale trade	20	Electric equip.
21	Retail trade	21	Oth. Machinery
22	Restaurants	22	Manuf. nec
23	Inland transp.	23	Electricity
24	Water transp.	24	Gas Distr.
25	Air transp.	25	Water Distr.
26	Transp. nec.	26	Construction
27	Telecomm.	27	Trade
28	Fin. serv.	28	Inland transp.
29	Real est.	29	Water transp.
30	Business serv.	30	Air transp.
31	Pub. admin.	31	Telecomm.
32	Education	32	Financial serv.
33	Health	33	Insurance
34	Social serv.	34	Business serv.
35	Household empl.	35	Recreation
		36	Education, Health
		37	Dwellings

Supplementary Appendix for
Income Differences, Productivity and Input-Output Networks
For Online Publication

Log-Normally distributed IO coefficients

In this Supplementary Appendix we provide the derivations for Appendix B. Consider a version of the model, where the elements γ_{ji}'s of the input-output matrix $\mathbf{\Gamma}$ are *independent random draws from a log-Normal distribution* and are thus allowed to vary across countries and sectors. As we explain in more detail later, a log-Normal distribution is an appropriate choice due to (i) equation (7) of the main text establishing that sectoral multipliers can be approximated by the sum of IO coefficients in the corresponding row of the IO matrix (shifted and multiplied by $1/n$), (ii) the fact that sectoral multipliers are log-Normally distributed, and (iii) the sum of independent log-Normal random variables is approximately log-Normal according to the Fenton-Wilkinson method (Fenton, 1960).

The general expression for y given in Proposition 1 is

$$y = \sum_{i=1}^{n} \mu_i \lambda_i + \sum_{i=1}^{n} \sum_{j \text{ s.t. } \gamma_{ji} \neq 0} \mu_i \gamma_{ji} \log \gamma_{ji} + \sum_{i=1}^{n} \mu_i (1-\gamma_i) log(1-\gamma_i) - \log n + \alpha \log K,$$

To employ this in our estimation, we need to calculate the expectation of this expression. The expectation of the first sum is simple and given by the same expression as in our main model:

$$E\left[\sum_{i=1}^{n} \mu_i \lambda_i\right] = n e^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} + e^{m_\mu + 1/2\sigma_\mu^2} \sum_{i=1}^{n} \left(\log\left(\Lambda_i^{US}\right) - 1\right).$$

The expectations of the other two sums, $E\left[\sum_{i=1}^{n} \sum_{j \text{ s.t. } \gamma_{ji} \neq 0} \mu_i \gamma_{ji} \log \gamma_{ji}\right]$ and $E\left[\sum_{i=1}^{n} \mu_i (1-\gamma_i) log(1-\gamma_i)\right]$, are more complex in case when γ_{ji} are treated as random. In what follows we derive both of these expectations. First, we obtain them as functions of the parameters of the log-Normal distribution of γ_{ji}. Then, we establish a relationship between these parameters and the parameters of the log-Normal distribution of the sectoral multipliers, μ_j, that we have estimated earlier. Finally, we use this relationship to express both of the computed expectations and the whole expression for $E(y)$ in terms of (m_μ, σ_μ^2).

Let as start with the first sum. Note that we can express it using the approximation of μ_i in (7) and extending the function $\gamma_{ji} \log \gamma_{ji}$ by continuity to $\gamma_{ji} = 0$ (for which in the limit it takes the value of 0):

$$\begin{aligned}
&\sum_{i=1}^{n} \sum_{j=1}^{n} \mu_i \gamma_{ji} \log \gamma_{ji} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left(1 + \sum_{s=1}^{n} \gamma_{is}\right) \gamma_{ji} \log \gamma_{ji} = \\
= \; & \frac{1}{n} \sum_{i=1}^{n} \sum_{j \neq i}^{n} \left(1 + \sum_{s=1}^{n} \gamma_{is}\right) \gamma_{ji} \log \gamma_{ji} + \frac{1}{n} \sum_{i=1}^{n} \left(1 + \sum_{s=1}^{n} \gamma_{is}\right) \gamma_{ii} \log \gamma_{ii} = \\
= \; & \frac{1}{n} \sum_{i=1}^{n} \sum_{j \neq i}^{n} \left(1 + \sum_{s=1}^{n} \gamma_{is}\right) \gamma_{ji} \log \gamma_{ji} + \frac{1}{n} \sum_{i=1}^{n} \left(1 + \sum_{s \neq i}^{n} \gamma_{is}\right) \gamma_{ii} \log \gamma_{ii} + \frac{1}{n} \sum_{i=1}^{n} \gamma_{ii}^2 \log \gamma_{ii}.
\end{aligned}$$

Given this expression and employing the assumption that all IO coefficients are distributed inde-

pendently, we obtain that

$$E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\mu_i\gamma_{ji}\log\gamma_{ji}\right]=\frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\left(1+\sum_{s=1}^{n}E\left[\gamma_{is}\right]\right)E\left[\gamma_{ji}\log\gamma_{ji}\right]+$$
$$\frac{1}{n}\sum_{i=1}^{n}\left(1+\sum_{s\neq i}^{n}E\left[\gamma_{is}\right]\right)E\left[\gamma_{ii}\log\gamma_{ii}\right]+\frac{1}{n}\sum_{i=1}^{n}E\left[\gamma_{ii}^2\log\gamma_{ii}\right].$$

Now, to calculate the expectations $E\left[\gamma_{ij}\right]$, $E\left[\gamma_{ji}\log\gamma_{ji}\right]$ and $E\left[\gamma_{ii}^2\log\gamma_{ii}\right]$, let us first denote by $(\mu_\gamma,\sigma_\gamma)$ the mean and the variance of the Normal distribution of $\log(\gamma_{ij})$. Then $E\left[\gamma_{ij}\right]$ can be expressed in terms of these parameters using the relationship between the Normal and log-Normal distributions:

$$E\left[\gamma_{ij}\right]=e^{\mu_\gamma+\frac{1}{2}\sigma_\gamma^2}.$$

The expressions for $E\left[\gamma_{ji}\log\gamma_{ji}\right]$ and $E\left[\gamma_{ii}^2\log\gamma_{ii}\right]$ are less straightforward. They are established by the following claim.

Claim If x $\sim$ log-Normal with parameters of the corresponding Normal distribution $(\mu_\gamma,\sigma_\gamma)$, then

$$E\left[x\log x\right]=e^{\mu_\gamma+\frac{\sigma_\gamma^2}{2}}\left(\mu_\gamma+\sigma_\gamma^2\right)\text{ and }E\left[x^2\log x\right]=e^{2\mu_\gamma+2\sigma_\gamma^2}\left(\mu_\gamma+2\sigma_\gamma^2\right).$$

Proof.

$$E\left[x\log x\right]=\int_0^\infty x\log x\frac{1}{x\sqrt{2\pi}\sigma_\gamma}e^{-\frac{(\log x-\mu_\gamma)^2}{2\sigma_\gamma^2}}dx$$

Let $\log x=y$, so that $dy=\frac{dx}{x}$. Then

$$\begin{aligned}E\left[x\log x\right]&=E\left[e^y y\right]=\int_{-\infty}^{\infty}e^y y\frac{1}{\sqrt{2\pi}\sigma_\gamma}e^{-\frac{(y-\mu_\gamma)^2}{2\sigma_\gamma^2}}dy=\frac{1}{\sqrt{2\pi}\sigma_\gamma}\int_{-\infty}^{\infty}ye^{-\frac{(y-\mu_\gamma)^2}{2\sigma_\gamma^2}+y}dy=\\&=\frac{1}{\sqrt{2\pi}\sigma_\gamma}\int_{-\infty}^{\infty}ye^{-\frac{y^2+\mu_\gamma^2-2y\mu_\gamma-2\sigma_\gamma^2 y}{2\sigma_\gamma^2}}dy=\frac{1}{\sqrt{2\pi}\sigma_\gamma}\int_{-\infty}^{\infty}ye^{-\frac{\left[y-(\mu_\gamma+\sigma_\gamma^2)\right]^2}{2\sigma_\gamma^2}}e^{\frac{(\mu_\gamma+\sigma_\gamma^2)^2-\mu_\gamma^2}{2\sigma_\gamma^2}}dy=\\&=e^{\frac{2\mu_\gamma\sigma_\gamma^2+\sigma_\gamma^4}{2\sigma_\gamma^2}}\frac{1}{\sqrt{2\pi}\sigma_\gamma}\int_{-\infty}^{\infty}ye^{-\frac{\left[y-(\mu_\gamma+\sigma_\gamma^2)\right]^2}{2\sigma_\gamma^2}}dy=e^{\mu_\gamma+\frac{\sigma_\gamma^2}{2}}\left(\mu_\gamma+\sigma_\gamma^2\right).\end{aligned}$$

Similarly,

$$\begin{aligned}E\left[x^2\log x\right]&=E\left[e^{2y}y\right]=\int_{-\infty}^{\infty}e^{2y}y\frac{1}{\sqrt{2\pi}\sigma_\gamma}e^{-\frac{(y-\mu_\gamma)^2}{2\sigma_\gamma^2}}dy=\frac{1}{\sqrt{2\pi}\sigma_\gamma}\int_{-\infty}^{\infty}ye^{-\frac{(y-\mu_\gamma)^2}{2\sigma_\gamma^2}+2y}dy=\\&=\frac{1}{\sqrt{2\pi}\sigma_\gamma}\int_{-\infty}^{\infty}ye^{-\frac{y^2+\mu_\gamma^2-2y\mu_\gamma-4\sigma_\gamma^2 y}{2\sigma_\gamma^2}}dy=\frac{1}{\sqrt{2\pi}\sigma_\gamma}\int_{-\infty}^{\infty}ye^{-\frac{\left[y-(\mu_\gamma+2\sigma_\gamma^2)\right]^2}{2\sigma_\gamma^2}}e^{\frac{(\mu_\gamma+2\sigma_\gamma^2)^2-\mu_\gamma^2}{2\sigma_\gamma^2}}dy=\\&=e^{\frac{4\mu_\gamma\sigma_\gamma^2+4\sigma_\gamma^4}{2\sigma_\gamma^2}}\frac{1}{\sqrt{2\pi}\sigma_\gamma}\int_{-\infty}^{\infty}ye^{-\frac{\left[y-(\mu_\gamma+2\sigma_\gamma^2)\right]^2}{2\sigma_\gamma^2}}=e^{2\mu_\gamma+2\sigma_\gamma^2}\left(\mu_\gamma+2\sigma_\gamma^2\right).\end{aligned}$$

□

Collecting the terms, we obtain:

$$E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\mu_i\gamma_{ji}\log\gamma_{ji}\right]=\frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\left(1+\sum_{s=1}^{n}E\left[\gamma_{is}\right]\right)E\left[\gamma_{ji}\log\gamma_{ji}\right]+$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(1+\sum_{s\neq i}^{n}E\left[\gamma_{is}\right]\right)E\left[\gamma_{ii}\log\gamma_{ii}\right]+\frac{1}{n}\sum_{i=1}^{n}E\left[\gamma_{ii}^2\log\gamma_{ii}\right]=\frac{1}{n}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\left(1+\sum_{s=1}^{n}E\left[\gamma_{is}\right]\right)E\left[\gamma_{ji}\log\gamma_{ji}\right]+$$

$$+\frac{1}{n}\sum_{i=1}^{n}E\left[\gamma_{ii}\log\gamma_{ii}\right]+\frac{1}{n}\sum_{i=1}^{n}E\left[\gamma_{ii}\log\gamma_{ii}\right]\left(\sum_{s\neq i}^{n}E\left[\gamma_{is}\right]\right)+n\frac{1}{n}E\left[\gamma_{ii}^2\log\gamma_{ii}\right]=$$

$$\left(1+ne^{\mu_\gamma+\frac{\sigma_\gamma^2}{2}}\right)(n-1)e^{\mu_\gamma+\frac{\sigma_\gamma^2}{2}}\left(\mu_\gamma+\sigma_\gamma^2\right)+e^{\mu_\gamma+\frac{\sigma_\gamma^2}{2}}\left(\mu_\gamma+\sigma_\gamma^2\right)+(n-1)e^{\mu_\gamma+\frac{\sigma_\gamma^2}{2}}e^{\mu_\gamma+\frac{\sigma_\gamma^2}{2}}\left(\mu_\gamma+\sigma_\gamma^2\right)+$$

$$+e^{2\mu_\gamma+2\sigma_\gamma^2}\left(\mu_\gamma+2\sigma_\gamma^2\right)=\left[e^{\frac{1}{2}\sigma_\gamma^2+\mu_\gamma}n+e^{\sigma_\gamma^2+2\mu_\gamma}\left(n^2-1\right)\right]\left(\mu_\gamma+\sigma_\gamma^2\right)+e^{2\sigma_\gamma^2+2\mu_\gamma}\left(\mu_\gamma+2\sigma_\gamma^2\right)=$$

$$=e^{\frac{1}{2}\sigma_\gamma^2+\mu_\gamma}\left[n+\left(n^2-1\right)e^{\frac{1}{2}\sigma_\gamma^2+\mu_\gamma}\right]\left(\mu_\gamma+\sigma_\gamma^2\right)+e^{2\sigma_\gamma^2+2\mu_\gamma}\left(\mu_\gamma+2\sigma_\gamma^2\right). \tag{SA-1}$$

Next, let us consider the expectation of the second sum, $E\left[\sum_{i=1}^{n}\mu_i(1-\gamma_i)log(1-\gamma_i)\right]$. The main difficulty with evaluating this expectation comes from the term $log(1-\gamma_i)$. Recall however that by definition, γ_i is the total share of intermediate goods in gross output of sector i. Therefore, a realization of γ_i is typically well below unity, not only theoretically but also empirically. This then allows adopting a quadratic approximation of $(1-\gamma_i)log(1-\gamma_i)$ by means of the second-order Taylor expansion:

$$(1-z)\log(1-z)\simeq -z+z^2/2,$$

so that

$$E\left[\mu_i(1-\gamma_i)log(1-\gamma_i)\right]=E\left[\mu_i(-\gamma_i+\frac{1}{2}\gamma_i^2)\right]+R,$$

where R is an error term due to the second-order approximation. Thus, we obtain

$$E\left[\mu_i(-\gamma_i+\frac{1}{2}\gamma_i^2)\right]=E\left[\left(\frac{1}{n}+\frac{1}{n}\sum_{s=1}^{n}\gamma_{is}\right)\left(-\sum_{p=1}^{n}\gamma_{pi}+\frac{1}{2}\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right)\right]$$

$$=\frac{1}{n}E\left[\left(1+\sum_{s=1}^{n}\gamma_{is}\right)\left(-\sum_{p=1}^{n}\gamma_{pi}\right)\right]+\frac{1}{2n}E\left[\left(1+\sum_{s=1}^{n}\gamma_{is}\right)\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right]$$

$$=-\frac{1}{n}E\left[\sum_{p=1}^{n}\gamma_{pi}\right]-E\left[\left(\sum_{s=1}^{n}\gamma_{is}\right)\left(\sum_{p=1}^{n}\gamma_{pi}\right)\right]+\frac{1}{2n}E\left[\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right]$$

$$+\frac{1}{2n}E\left[\left(\sum_{s=1}^{n}\gamma_{is}\right)\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right].$$

However,

$$
\begin{aligned}
& E\left[\left(\sum_{s=1}^{n}\gamma_{is}\right)\left(\sum_{p=1}^{n}\gamma_{pi}\right)\right] = E\left[\left(\sum_{s=1\neq i}^{n}\gamma_{is}+\gamma_{ii}\right)\left(\sum_{p=1}^{n}\gamma_{pi}\right)\right] \\
= \; & E\left[\sum_{s=1\neq i}^{n}\gamma_{is}\right]E\left[\sum_{p=1}^{n}\gamma_{pi}\right]+E\left[\gamma_{ii}^{2}\right]+E[\gamma_{ii}]E\left[\sum_{p=1\neq i}^{n}\gamma_{pi}\right] \\
= \; & \left[\sum_{s=1\neq i}^{n}E[\gamma_{is}]\right]\left[\sum_{p=1}^{n}E[\gamma_{pi}]\right]+E\left[\gamma_{ii}^{2}\right]+E[\gamma_{ii}]\left[\sum_{p=1\neq i}^{n}E[\gamma_{pi}]\right] \\
= \; & (n-1)n\left(E[\gamma_{ij}]\right)^{2}+E\left[\gamma_{ij}^{2}\right]+(n-1)\left(E[\gamma_{ij}]\right)^{2},
\end{aligned}
$$

and

$$
\begin{aligned}
& E\left[\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right] = E\left[\sum_{q=1}^{n}\gamma_{qi}^{2}+\sum_{q,r=1,q\neq r}^{n}\gamma_{qi}\gamma_{ri}\right] \\
= \; & \sum_{q=1}^{n}E[\gamma_{qi}^{2}]+\sum_{q,r=1,q\neq r}^{n}E[\gamma_{qi}]E[\gamma_{ri}] = nE[\gamma_{ij}^{2}]+\left(n^{2}-n\right)\left(E[\gamma_{ij}]\right)^{2},
\end{aligned}
$$

and finaly,

$$
\begin{aligned}
& E\left[\left(\sum_{s=1}^{n}\gamma_{is}\right)\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right] = E\left[\left(\gamma_{ii}+\sum_{s=1\neq i}^{n}\gamma_{is}\right)\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right] \\
= \; & E\left[\gamma_{ii}\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right]+E\left[\sum_{s=1\neq i}^{n}\gamma_{is}\right]E\left[\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right] \\
= \; & E\left[\gamma_{ii}\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right]+(n-1)E[\gamma_{ij}]\left[nE[\gamma_{ij}^{2}]+\left(n^{2}-n\right)\left(E[\gamma_{ij}]\right)^{2}\right],
\end{aligned}
$$

where

$$
\begin{aligned}
& E\left[\gamma_{ii}\left(\sum_{q=1}^{n}\gamma_{qi}\right)\left(\sum_{r=1}^{n}\gamma_{ri}\right)\right] \\
= \; & E\left[\gamma_{ii}\left(\gamma_{ii}+\sum_{q=1\neq i}^{n}\gamma_{qi}\right)\left(\gamma_{ii}+\sum_{r=1\neq i}^{n}\gamma_{ri}\right)\right] \\
= \; & E\left[\left(\gamma_{ii}^{2}+\gamma_{ii}\sum_{q=1\neq i}^{n}\gamma_{qi}\right)\left(\gamma_{ii}+\sum_{r=1\neq i}^{n}\gamma_{ri}\right)\right] \\
= \; & E\left[\left(\gamma_{ii}^{3}+\gamma_{ii}^{2}\sum_{q=1\neq i}^{n}\gamma_{qi}\right)+\left(\gamma_{ii}^{2}\sum_{q=1\neq i}^{n}\gamma_{qi}\right)+\gamma_{ii}\sum_{q=1\neq i}^{n}\gamma_{qi}\sum_{r=1\neq i}^{n}\gamma_{ri}\right] \\
= \; & E[\gamma_{ii}^{3}]+E[\gamma_{ii}^{2}](n-1)E[\gamma_{ij}]+E[\gamma_{ii}^{2}](n-1)E[\gamma_{ij}] \\
& +E[\gamma_{ij}](n-1)\left[E[\gamma_{ij}^{2}]+(n-2)\left(E[\gamma_{ij}]\right)^{2}\right],
\end{aligned}
$$

with

$$E\left[\sum_{q=1\neq i}^{n}\gamma_{qi}\sum_{r=1\neq i}^{n}\gamma_{ri}\right] = (n-1)\,E[\gamma_{ij}^2] + \Big((n-1)^2-(n-1)\Big)\,(E[\gamma_{ij}])^2$$
$$= \quad (n-1)\left[E[\gamma_{ij}^2] + (n-2)\,(E[\gamma_{ij}])^2\right].$$

For these derivations we employed the assumption that γ_{ji} are all *independently* and *identically* distributed. Combining all terms, we obtain:

$$\begin{aligned}
& E\left[\mu_i(-\gamma_i+\frac{1}{2}\gamma_i^2)\right] \\
= \; & -E\left[\gamma_{ij}\right] - \frac{1}{n}\left[(n-1)n\,(E[\gamma_{ij}])^2 + E\left[\gamma_{ij}^2\right] + (n-1)\,(E[\gamma_{ij}])^2\right] \\
& +\frac{1}{2n}\left[nE[\gamma_{ij}^2] + \left(n^2-n\right)(E[\gamma_{ij}])^2\right] \\
& +\frac{1}{2n}\left[E[\gamma_{ij}^3] + E[\gamma_{ij}^2](n-1)E[\gamma_{ij}] + E[\gamma_{ij}^2](n-1)E[\gamma_{ij}]\right] \\
& +\frac{1}{2n}\left[E[\gamma_{ij}](n-1)\left(E[\gamma_{ij}^2] + (n-2)\,(E[\gamma_{ij}])^2\right)\right] \\
& +\frac{1}{2n}\left[(n-1)E[\gamma_{ij}]\left(nE[\gamma_{ij}^2] + \left(n^2-n\right)(E[\gamma_{ij}])^2\right)\right] \\
= \; & \frac{1}{2n}E[\gamma_{ij}^3] + \frac{1}{n}E[\gamma_{ij}^2]\left(-1+\frac{n}{2}+E\left[\gamma_{ij}\right]\left[\frac{3}{2}(n-1)+\frac{1}{2}(n)(n-1)\right]\right) \\
& +\frac{1}{n}\,(E[\gamma_{ij}])^2\left[-(n-1)-(n-1)n+\frac{n^2-n}{2}+\frac{1}{2}E\left[\gamma_{ij}\right]\left[(n-1)(n-2)+(n-1)(n^2-n)\right]\right] - E\left[\gamma_{ij}\right] \\
= \; & \frac{1}{2n}E[\gamma_{ij}^3] + E[\gamma_{ij}^2]\left(\frac{n-2}{2n}+E\left[\gamma_{ij}\right]\left[\frac{1}{2}n+1-\frac{3}{2n}\right]\right) \\
& +(E[\gamma_{ij}])^2\left[-\frac{1}{2}n-\frac{1}{2}+\frac{1}{n}+\frac{1}{2}E\left[\gamma_{ij}\right]\left(n^2-n-2+\frac{2}{n}\right)\right] - E\left[\gamma_{ij}\right].
\end{aligned}$$

The fact that γ_{ij} is log-Normal (so $\log\gamma_{ij}$ is Normal) implies that γ_{ij}^2 and γ_{ij}^3 are also log-Normal because $\log(\gamma_{ij}^2) = 2\log\gamma_{ij}$ and $\log(\gamma_{ij}^3) = 3\log\gamma_{ij}$. Then we know that the expectations of all these variables can be written as

$$\begin{aligned}
E\left[\gamma_{ij}\right] &= e^{\mu_\gamma+\frac{1}{2}\sigma_\gamma^2} \\
E\left[\gamma_{ij}^2\right] &= e^{2\mu_\gamma+2\sigma_\gamma^2} \\
E\left[\gamma_{ij}^3\right] &= e^{3\mu_\gamma+\frac{9}{2}\sigma_\gamma^2}
\end{aligned}$$

Therefore,

$$\begin{aligned}
E\left[\mu_i(-\gamma_i+\frac{1}{2}\gamma_i^2)\right] \;=\; & \frac{1}{2n}e^{3\mu_\gamma+\frac{9}{2}\sigma_\gamma^2} + e^{2\mu_\gamma+2\sigma_\gamma^2}\left(\frac{n-2}{2n}+e^{\mu_\gamma+\frac{1}{2}\sigma_\gamma^2}\left[\frac{1}{2}n+1-\frac{3}{2n}\right]\right) \\
& +e^{2\mu_\gamma+\sigma_\gamma^2}\left[-\frac{1}{2}n-\frac{1}{2}+\frac{1}{n}+\frac{1}{2}e^{\mu_\gamma+\frac{1}{2}\sigma_\gamma^2}\left(n^2-n-2+\frac{2}{n}\right)\right] - e^{\mu_\gamma+\frac{1}{2}\sigma_\gamma^2} \qquad \text{(SA-2)}
\end{aligned}$$

Now, it remains to relate the distribution of γ_{ji}'s to the distribution of sectoral multipliers μ_j, so as to express $E\left[\sum_{i=1}^n\sum_{j=1}^n\ \mu_i\gamma_{ji}\log\gamma_{ji}\right]$ and $E\left[\mu_i(-\gamma_i+\frac{1}{2}\gamma_i^2)\right]$ in terms of the earlier estimated parameters (m_μ,σ_μ^2). This relationship is provided by equation (7) according to which $\mu_j \approx \frac{1}{n} + \frac{1}{n}\sum_{i=1}^n\gamma_{ji}$. From this equation it follows that $E(\mu) = \frac{1}{n}+\frac{1}{n}\mu_{sum}$ and $var(\mu) = \frac{1}{n^2}\sigma_{sum}^2$, where μ_{sum}, σ_{sum}^2 are the mean and the variance of the distribution of the sum $\sum_{i=1}^n\gamma_{ji}$. Now, while $E(\mu)$, $var(\mu)$ can be expressed in terms of (m_μ,σ_μ^2) by means of the relationship between the Normal and

log-Normal distributions,[70] μ_{sum}, σ^2_{sum} can be expressed in terms of $(\mu_\gamma, \sigma^2_\gamma)$ by means of the Fenton-Wilkinson method. This then provides us with the sought-after relationship between parameters $(\mu_\gamma, \sigma^2_\gamma)$ and (m_μ, σ^2_μ).

The Fenton-Wilkinson method implies that the distribution of the sum $\sum_{i=1}^n \gamma_{ji}$ of the independent log-Normally distributed random variables is approximately log-Normal with

$$\sigma^2_{sum} = \log\left(\frac{\left(e^{\sigma^2_\gamma}\right)-1}{n+1}\right), \tag{SA-3}$$

$$\mu_{sum} = \log\left(ne^{\mu_\gamma}\right) + \frac{1}{2}\left(\sigma^2_\gamma - \sigma^2_{sum}\right) = \log\left(ne^{\mu_\gamma}\right) + \frac{1}{2}\left(\sigma^2_\gamma - \log\left(\frac{\left(e^{\sigma^2_\gamma}\right)-1}{n+1}\right)\right). \tag{SA-4}$$

Note that it is this method, in the first place, that justifies our assumption that IO coefficients γ_{ji}'s are log-Normally distributed. Indeed, as the distribution of sectoral multipliers μ_j has been shown to be log-Normal, and $\mu_j \approx \frac{1}{n} + \frac{1}{n}\sum_{i=1}^n \gamma_{ji}$, the sum $\sum_{i=1}^n \gamma_{ji}$ must be distributed log-Normally. By Fenton-Wilkinson method, this is consistent with γ_{ji}'s being log-Normal.

Using (SA-3) – (SA-4), equations $E(\mu) = \frac{1}{n} + \frac{1}{n}\mu_{sum}$, $var(\mu) = \frac{1}{n^2}\sigma^2_{sum}$, and the expressions for $E(\mu)$, $var(\mu)$ in footnote 70, we derive:

$$e^{\sigma^2_\gamma} = (n+1)\,e^{\sigma^2_{sum}} + 1 = (n+1)e^{n^2 var(\mu)} + 1 = (n+1)e^{n^2 e^{2m_\mu + \sigma^2_\mu}\cdot[e^{\sigma^2_\mu}-1]} + 1, \tag{SA-5}$$

$$\begin{aligned} e^{\mu_\gamma} &= \frac{e^{\mu_{sum}}}{n}\left(n+1+e^{-\sigma^2_{sum}}\right)^{-\frac{1}{2}} = \frac{e^{nE(\mu)-1}}{n}\left(n+1+e^{-n^2 var(\mu)}\right)^{-\frac{1}{2}} = \\ &= \frac{e^{ne^{m_\mu + 1/2\sigma^2_\mu}-1}}{n}\left(n+1+e^{-n^2 e^{2m_\mu+\sigma^2_\mu}\cdot[e^{\sigma^2_\mu}-1]}\right)^{-\frac{1}{2}}. \end{aligned} \tag{SA-6}$$

This is the relationship between $(\mu_\gamma, \sigma^2_\gamma)$ and (m_μ, σ^2_μ). Let us denote the expression for $e^{\sigma^2_\gamma}$ by x and the expression for e^{μ_γ} by z. Then using this in (SA-1), we obtain:

$$\begin{aligned} E\left[\sum_{i=1}^n\sum_{j=1}^n \mu_i\gamma_{ji}\log\gamma_{ji}\right] &= e^{\frac{1}{2}\sigma^2_\gamma+\mu_\gamma}\left[n+\left(n^2-1\right)e^{\frac{1}{2}\sigma^2_\gamma+\mu_\gamma}\right]\left(\mu_\gamma+\sigma^2_\gamma\right) + e^{2\sigma^2_\gamma+2\mu_\gamma}\left(\mu_\gamma+2\sigma^2_\gamma\right) = \\ &= x^{\frac{1}{2}}z[n+(n^2-1)x^{\frac{1}{2}}z](\log(x)+\log(z)) + x^2z^2(\log(z)+2\log(x)). \end{aligned}$$

Similarly, using this in (SA-2), we obtain:

$$\begin{aligned} E\left[\mu_i(-\gamma_i+\frac{1}{2}\gamma_i^2)\right] &= \frac{1}{2n}e^{3\mu_\gamma+\frac{9}{2}\sigma^2_\gamma} + e^{2\mu_\gamma+2\sigma^2_\gamma}\left(\frac{n-2}{2n} + e^{\mu_\gamma+\frac{1}{2}\sigma^2_\gamma}\left[\frac{1}{2}n+1-\frac{3}{2n}\right]\right) \\ &\quad + e^{2\mu_\gamma+\sigma^2_\gamma}\left[-\frac{1}{2}n-\frac{1}{2}+\frac{1}{n}+\frac{1}{2}e^{\mu_\gamma+\frac{1}{2}\sigma^2_\gamma}\left(n^2-n-2+\frac{2}{n}\right)\right] - e^{\mu_\gamma+\frac{1}{2}\sigma^2_\gamma} \\ &= \frac{1}{2n}z^3x^{\frac{9}{2}} + z^2x^2\left(\frac{n-2}{2n} + zx^{\frac{1}{2}}\left[\frac{1}{2}n+1-\frac{3}{2n}\right]\right) \\ &\quad + z^2x\left[-\frac{1}{2}n-\frac{1}{2}+\frac{1}{n}+\frac{1}{2}zx^{\frac{1}{2}}\left(n^2-n-2+\frac{2}{n}\right)\right] - zx^{\frac{1}{2}}. \end{aligned}$$

Now we can substitute these expressions for $E\left[\sum_{i=1}^n\sum_{j=1}^n \mu_i\gamma_{ji}\log\gamma_{ji}\right]$ and $E\left[\sum_{i=1}^n \mu_i(1-\gamma_i)log(1-\gamma_i)\right]$

[70] $E(\mu) = e^{m_\mu + 1/2\sigma^2_\mu}$, $var(\mu) = e^{2m_\mu+\sigma^2_\mu}\cdot[e^{\sigma^2_\mu}-1]$

in the expression for the expected aggregate income, and we arrive at

$$
\begin{aligned}
E(y) &= ne^{m_\mu+m_\Lambda+1/2(\sigma_\mu^2+\sigma_\Lambda^2)+\sigma_{\mu,\Lambda}} + E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\mu_i\gamma_{ji}\log\gamma_{ji}\right] + E\left[\sum_{i=1}^{n}\mu_i(1-\gamma_i)log(1-\gamma_i)\right] \\
&\quad -\log n + \alpha\log(K) + e^{m_\mu+1/2\sigma_\mu^2}\sum_{i=1}^{n}\left(\log\left(\Lambda_i^{US}\right)-1\right) \\
&= ne^{m_\mu+m_\Lambda+1/2(\sigma_\mu^2+\sigma_\Lambda^2)+\sigma_{\mu,\Lambda}} \\
&\quad +x^{\frac{1}{2}}z[n+x^{\frac{1}{2}}z(n^2-1)](\log(x)+\log(z)) + x^2z^2(\log(z)+2\log(x)) \\
&\quad +\frac{1}{2n}z^3x^{\frac{9}{2}} + z^2x^2\left(\frac{n-2}{2n} + zx^{\frac{1}{2}}\left[\frac{1}{2}n+1-\frac{3}{2n}\right]\right) \\
&\quad +z^2x\left[-\frac{1}{2}n-\frac{1}{2}+\frac{1}{n}+\frac{1}{2}zx^{\frac{1}{2}}\left(n^2-n-2+\frac{2}{n}\right)\right] - zx^{\frac{1}{2}} \\
&\quad -\log n + \alpha\log(K) + e^{m_\mu+1/2\sigma_\mu^2}\sum_{i=1}^{n}\left(\log\left(\Lambda_i^{US}\right)-1\right) \\
&= ne^{m_\mu+m_\Lambda+1/2(\sigma_\mu^2+\sigma_\Lambda^2)+\sigma_{\mu,\Lambda}} - \log n + \alpha\log(K) + e^{m_\mu+1/2\sigma_\mu^2}\sum_{i=1}^{n}\left(\log\left(\Lambda_i^{US}\right)-1\right) + \Psi(m_\mu,\sigma_\mu),
\end{aligned}
$$

where

$$
\begin{aligned}
\Psi(m_\mu,\sigma_\mu) &= x^{\frac{1}{2}}z\Big[\left(n+x^{\frac{1}{2}}z(n^2-1)\right)(\log(x)+\log(z)) + \frac{1}{2n}z^2x^4 - 1 \\
&\quad +x^2z^2\left(\frac{n}{2}+1-\frac{3}{2n}\right) + xz^2\left(\frac{n^2}{2}-\frac{n}{2}-1+\frac{1}{n}\right)\Big] \\
&\quad +\frac{1}{2n}xz^2\left[2nx(\log(z)+2\log(x)) + x(n-2) - n^2 - n + 2\right].
\end{aligned}
$$

Note that $\Psi(m_\mu,\sigma_\mu)$ is a function of (m_μ,σ_μ) via the definition $x = e^{\sigma_\gamma^2}, z = e^{\mu_\gamma}$ and the relationship between $(\mu_\gamma,\sigma_\gamma^2)$ and (m_μ,σ_μ^2) in (SA-5) - (SA-6). This is the expression for the expected aggregate income in terms of the parameter estimates used in the benchmark model (analogue of equation (9)). We bring it to estimation and predict cross-country income differences in the setting with asymmetric IO coefficients.