Supplementary Appendix for Income Differences and Input-Output Structure

September 11, 2016

Appendix B: Extensions of the benchmark model

0.1 Log-Normally distributed IO coefficients

Consider a more general 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 (9) 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).

When (non-zero) IO coefficients are not all equal to $\widehat{\gamma}$, the term $\sum_{i=1}^{n}\sum_{j\text{s.t. }\gamma_{ji}\neq 0}\mu_i\gamma_{ji}\log\gamma_{ji}$ in equation (7) is no longer equal to $\sum_{i=1}^{n}\mu_i\gamma\log(\widehat{\gamma})$ (as in (10)). Instead, we can express it using the approximation of μ_i in (9) 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):

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

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

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

To employ this in our estimation, we need to calculate the expectation of this expression. Given the assumption that all IO coefficients are distributed independently, we obtain that

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

Then it remains to calculate the expectations $E\left[\gamma_{ij}\right]$, $E\left[\gamma_{ji}\log\gamma_{ji}\right]$ and $E\left[\gamma_{ii}^2\log\gamma_{ii}\right]$. First, let us denote by $(\mu_\gamma,\sigma_\gamma)$ the mean and variance of the Normal distribution of $\log(\gamma_{ij})$. $E\left[\gamma_{ij}\right]$ 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:

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

Now, it remains to relate the distribution of γ_{ji}'s to the distribution of sectoral multipliers μ_j, so as to express $E\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\mu_i\gamma_{ji}\log\gamma_{ji}\right]$ in terms of earlier estimated parameters (m_μ,σ_μ^2). This relationship is provided by equation (9) 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,[1] μ_{sum}, σ_{sum}^2 can be expressed in terms of $(\mu_\gamma,\sigma_\gamma^2)$ by means of the Fenton-Wilkinson method. This then provides us with the sought-after relationship between parameters $(\mu_\gamma,\sigma_\gamma^2)$ and (m_μ,σ_μ^2).

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

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

$$\mu_{sum} = \log\left(ne^{\mu_\gamma}\right)+\frac{1}{2}\left(\sigma_\gamma^2-\sigma_{sum}^2\right)=\log\left(ne^{\mu_\gamma}\right)+\frac{1}{2}\left(\sigma_\gamma^2-\log\left(\frac{\left(e^{\sigma_\gamma^2}\right)-1}{n+1}\right)\right). \tag{3}$$

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 (2) – (3), equations $E(\mu)=\frac{1}{n}+\frac{1}{n}\mu_{sum}$, $var(\mu)=\frac{1}{n^2}\sigma_{sum}^2$, and the expressions for $E(\mu)$, $var(\mu)$

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

in footnote 1, we derive:

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

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

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

Now we can substitute this for $E\left[\sum_{i=1}^n \sum_{j=1}^n \mu_i \gamma_{ji} \log \gamma_{ji}\right]$ in the expression for the expected aggregate income, and we arrive at

$$\begin{aligned}
E(y) &= ne^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} - (1+\gamma) + E\left[\sum_{i=1}^n \sum_{j=1}^n \mu_i \gamma_{ji} \log \gamma_{ji}\right] + \\
&+ \log(1-\gamma) - \log n + \alpha \log(K) + e^{m_\mu + 1/2\sigma_\mu^2} \sum_{i=1}^n \log\left(\Lambda_i^{US}\right) = \\
&= ne^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} - (1+\gamma) + \\
&+ x^{\frac{1}{2}} z [n + x^{\frac{1}{2}} z (n^2 - 1)] (\log(x) + \log(z)) + x^2 z^2 (\log(z) + 2\log(x)) + \\
&+ \log(1-\gamma) - \log n + \alpha \log(K) + e^{m_\mu + 1/2\sigma_\mu^2} \sum_{i=1}^n \log\left(\Lambda_i^{US}\right). \qquad (4)
\end{aligned}$$

This is the expression for the expected aggregate income in terms of parameter estimates used in the benchmark model (analogue of equation (13)). We bring it to estimation and predict cross-country income differences in the setting with asymmetric IO linkages.

0.2 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 (27) 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:

$$\begin{aligned} y &= \sum_{i=1}^{n} \mu_i \Lambda_i^{rel} + \sum_{i=1}^{n} \mu_i \gamma \log(\widehat{\gamma}) + \log(1-\gamma) + \sum_{i=1}^{n} \beta_i \log(\beta_i) + \alpha \log(K) - \\ &- (1+\gamma) + \sum_{i=1}^{n} \mu_i \log(\Lambda_i^{US}). \end{aligned} \tag{5}$$

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.[2] Then, using the estimates of the parameters of this distribution, $\boldsymbol{m}$ and $\boldsymbol{\Sigma}$, together with the equations (12) (see footnote 32), we find the predicted aggregate output $E(y)$ as a function of these parameters:[3]

$$\begin{aligned} E(y) &= n e^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} + (1+\gamma)(\gamma \log(\widehat{\gamma}) - 1) + \log(1-\gamma) + \\ &+ \sum_{i=1}^{n} \beta_i \log(\beta_i) + \alpha \log(K) + e^{m_\mu + 1/2\sigma_\mu^2} \sum_{i=1}^{n} \log\left(\Lambda_i^{US}\right). \end{aligned} \tag{6}$$

The resulting expression for $E(y)$ is similar to (13) in our benchmark model.

0.3 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

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

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

following production function:

$$q_i = \Lambda_i \left(k_i^{\alpha} l_i^{1-\alpha}\right)^{1-\gamma_i-\sigma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}} \cdot f_{1i}^{\sigma_{1i}} f_{2i}^{\sigma_{2i}} \cdot \ldots \cdot f_{ni}^{\sigma_{ni}}, \tag{7}$$

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.[4] 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 intermediate goods and are defined by the balanced trade condition:

$$X = \sum_{i=1}^{n} \sum_{j=1}^{n} \overline{p}_j f_{ji}, \tag{8}$$

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 (27) of Proposition 2, adopted to our framework here:

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

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.[5]

In the empirical analysis we use an approximate representation of aggregate output, where a range

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

[5] 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 $\sigma_i + \gamma_i < 1$: $\sum_{j=1}^{n} \left(\frac{1}{n}\sigma_i + \gamma_{ji}\}_{ji}\right) = \sigma_i + \gamma_i < 1$.

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 $\mathbf{\Gamma}$. 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.[6] 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:

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

Now, using equations (12) (see footnote 32) for the parameters of the bivariate log-Normal distribution of (μ_i, Λ_i^{rel}), we can derive the predicted aggregate output $E(y)$:

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

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 $\{\overline{p}_j\}_{j=1}^{n}$.

0.4 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

[6] This allows approximating $\log\left(1+\sum_{i=1}^{n}\sigma_i\bar{\mu_i}\right)$ with $\sigma\sum_{i=1}^{n}\bar{\mu_i} = \sigma\,(1+\gamma+\sigma)$, 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 $\frac{1}{n}\left(\sum_{k=0}^{+\infty}\bar{\mathbf{\Gamma}}^k\right)\mathbf{1}$ and the analogous approximation for $\{\mu_i\}_{i=1}^{n}$ (see section 3.3).

can be described by the following Cobb-Douglas function:

$$q_i = \Lambda_i \left(k_i^{\alpha} u_i^{\delta} s_i^{1-\alpha-\delta}\right)^{1-\gamma_i-\sigma_i} d_{1i}^{\gamma_{1i}} d_{2i}^{\gamma_{2i}} \cdot \ldots \cdot d_{ni}^{\gamma_{ni}}, \tag{9}$$

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 (27) 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)):

$$\begin{aligned} y &= \sum_{i=1}^{n} \mu_i \Lambda_i^{rel} + \sum_{i=1}^{n} \mu_i \gamma \log(\widehat{\gamma}) + \log(1-\gamma) - \log n + \alpha \log(K) + \\ &+ \delta \log U + (1-\alpha-\delta) \log S - \log(U+S) - (1+\gamma) + \sum_{i=1}^{n} \mu_i \log(\Lambda_i^{US}), \end{aligned} \tag{10}$$

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 in (12) (see footnote 32), we derive the expression for the predicted aggregate output $E(y)$ in terms of the estimated parameters:

$$\begin{aligned} E(y) &= n e^{m_\mu + m_\Lambda + 1/2(\sigma_\mu^2 + \sigma_\Lambda^2) + \sigma_{\mu,\Lambda}} + (1+\gamma)(\gamma \log(\widehat{\gamma}) - 1) + \log(1-\gamma) - \log n + \alpha \log(K) + \\ &+ \delta \log U + (1-\alpha-\delta) \log S - \log(U+S) + e^{m_\mu + 1/2\sigma_\mu^2} \sum_{i=1}^{n} \log\left(\Lambda_i^{US}\right). \end{aligned} \tag{11}$$

This equation for the predicted aggregate output is analogous to the equation (13) that we employed in our estimation of the benchmark model.

Appendix D: Additional Figures and Tables

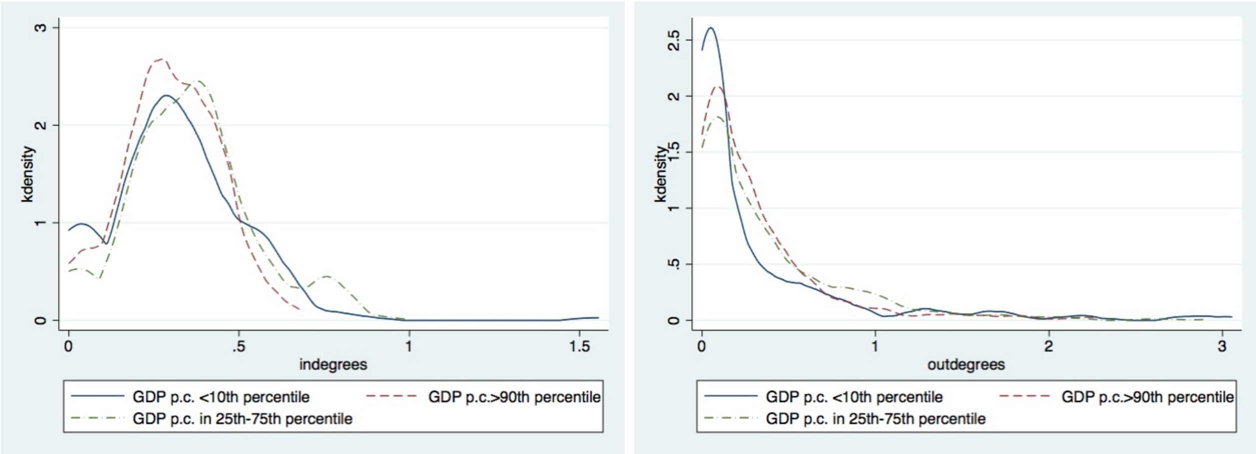

Figure A-1: Distribution of sectoral in-degrees (left) and out-degrees (right) (GTAP sample)

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