
1.0 Introduction

One of the most important developments in the growth literature of the last decade is the enhanced appreciation of the role that the misallocation of resources has in helping us understand income differences across countries. Given an economy’s stock of physical capital, labor, human capital, and knowledge, the way in which those aggregate quantities of inputs are allocated across firms and industries — and even potentially within firms — determines the economy’s overall level of production. The best allocation will maximize welfare and, in a sense that can be made precise, output itself in the long run. Other allocations result in lower levels of output and therefore show up in the aggregate as a lower level of total factor productivity (TFP).

In a broad sense, this is an old idea with many antecedents. In the real-business-cycle literature, for example, it is commonly appreciated that tax distortions or regulations may show up as TFP shocks. Chari, Kehoe, and McGrattan (2007) followed in this tradition.

In the literature on growth and development, Restuccia and Rogerson (2008) explicitly analyzed a model of misallocation among heterogeneous plants to quantify the effect on aggregate TFP. Banerjee and Duflo (2005) argued that the marginal product of capital differs widely among firms in India, potentially reducing overall output. Hsieh and Klenow (2009) presented empirical evidence that misallocation across plants within four-digit industries may reduce TFP in manufacturing by a factor of two to three...
in China and India. A large literature surrounding these papers considers various mechanisms through which misallocation can lead to income differences.¹

This chapter provides my own idiosyncratic perspective on misallocation and presents three basic points. First, I begin with an overview of misallocation. A simple toy model illustrates how misallocation can reduce TFP, and I outline several questions related to misallocation that might be considered in future research. Second, I suggest one way in which the effects of misallocation can be amplified: through the input–output structure of the economy. Because the outputs of many firms are used as the inputs of other firms, the effects of misallocation can be amplified. Third, I provide an overview of the input–output structure of the United States and 34 other economies, albeit at a fairly high level of aggregation. In addition to supporting the basic point that the amplification associated with input–output economics can be quantitatively important, this overview suggests what I think is a remarkable similarity in the input–output structures of diverse economies. Understanding whether this actually is the case, why it may be so, and what implications it entails comprise other useful areas for future research.

2.0 Misallocation

This section is an overview of the consequences of the misallocation of resources.

2.1 Misallocation and TFP

We begin by presenting a simple example that illustrates the basic point of the misallocation literature: misallocation reduces TFP. We consider an economy in which the two key produced goods are steel and lattes:

- Production: \( X_{\text{steel}} = L_{\text{steel}}, \quad X_{\text{lattice}} = L_{\text{lattice}} \)
- Resource constraint: \( L_{\text{steel}} + L_{\text{lattice}} = \bar{L} \)
- GDP (aggregation): \( Y = X_{\text{steel}}^{1/2} X_{\text{lattice}}^{1/2} \)

One unit of labor can produce either a unit of steel or a cup of latte. The economy is endowed with \( \bar{L} \) units of labor. We assume that lattes and steel combine in a Cobb–Douglas manner to generate a final good. This last equation could be replaced by a utility function, but then we would need to specify prices to aggregate the two goods; the approach here is simpler.

Obviously, the only allocative decision that must be made is how much labor to employ producing steel versus lattes. We let \( x = L_{\text{steel}} / \bar{L} \) denote the allocation of labor. This could be determined by perfectly competitive markets, by a social planner, by markets distorted by taxes, or in any number of different ways.

Solving for GDP given the allocation yields:

\[
Y = A(x) \bar{L}
\]

where TFP, \( A(x) \), is given by:

\[
A(x) = \sqrt{x (1 - x)}
\]

These two equations summarize simply one of the key points of the recent literature on misallocation: The misallocation of resources reduces TFP. As is clear given the symmetry of the setup, the optimal allocation of labor in this simple economy features \( x^* = 1/2 \). Any departure from this allocation — putting either too little or too much labor into making steel — reduces TFP and therefore GDP. More generally, if the exponents in the production function were \( \sigma \) and \( 1 - \sigma \) instead of \( 1/2 \), then TFP would be simply \( A(x) = x^\sigma (1 - x)^{1-\sigma} \); this expression is useful later. The effects of misallocation are shown graphically in Figure 1.

The figure illustrates another key point: Small departures from the optimal allocation of labor have tiny effects on TFP (an application of the envelope theorem), but significant misallocation can have very large effects. Given the large income differences that we see across countries, this possibility is appealing.

However, more careful consideration of Figure 1 indicates that this simple model has what may be an important limitation: In the presence of significant misallocation, a small improvement in the allocation of resources will have a large impact on TFP.

We contrast this with a hypothetical example like that in Figure 2. The dashed line in the figure repeats the effect of misallocation on TFP from Figure 1, and the new solid line depicts an alternative. It seems that the alternative better captures the world in which we live. As before, small misallocation has small effects and large misallocation has large effects. Now, however, an intermediate degree of misallocation can have large effects as

¹ Examples include Easterly (1995); Parente and Prescott (1999); Caselli and Gennaioli (2005); Lagos (2006); Alfaro, Charlton, and Kanczuk (2008); Buera and Shin (2008); Guner, Ventura, and Xu (2008); La, Porta and Shleifer (2008); Bartelsman, Haltiwanger, and Scarpetta (2009); Epifani and Gancia (2009); Peters (2009); Vollrath (2009); Midrigan and Xu (2010); Moll (2010); and Syverson (2010).
even may help explain the “twin peaks” structure of the world income distribution emphasized by Quah (1996).

One of the challenges going forward in models of misallocation — perhaps — is to ensure that they capture some of the features in Figure 2 rather than some of the limitations in Figure 1. Jones (2011) explored the possibility that the O-ring–style complementarity of Kremer (1993) may help in this regard.

This simple example is useful in illustrating how misallocation reduces TFP, but it fails to capture one of the key points emphasized in the recent literature on misallocation. In the example, misallocation is across sectors: We may have too much or too little steel relative to what is optimal. In contrast, the recent literature often emphasizes misallocation at a more microeconomic level: Within the steel sector, there are some plants that are good at making steel and others that are less good. Misallocation may involve giving the less efficient plants too many resources. This within-sector misallocation was prominently emphasized by Foster, Halliwanger, and Syverson (2008) and Hsieh and Klenow (2009), for example. Clearly, both kinds of misallocation (i.e., across and within sectors) can be important, and an interesting open question in the literature is determining the relative (and absolute) importance of each.

We might even push this insight further: Why are some plants more productive than others? Perhaps because of the misallocation of resources within plants. Organizing a plant and producing output involve an enormous number of decisions, which may be distorted or made incorrectly because of misallocation: Perhaps the plant manager is not the best person for the job (Caselli and Gennaioli 2005), perhaps the most talented workers within the plant are not promoted to the appropriate positions, perhaps the incentives for the workers to produce efficiently are not present (Lazear 2000), perhaps unionization and job protection lead the firm to use too much labor inappropriately (Schmitz 2005), and so on.

### 2.2 Key Questions

At some basic level, there are only two fundamental reasons for income differences across countries in the long run: Either economies have different production possibilities or they have different allocations. In the case of production possibilities, once we have endogenized ideas, the only remaining reason for differences is geographic. Perhaps some parcels of land are more conducive to production than others. Although there probably is something to this explanation, my current understanding is that these effects
are small relative to the large income differences we see across countries. Probably the single most persuasive evidence on this point is the classic Olson (1996) argument: The large income difference that emerged between North and South Korea in the last half-century, for example, surely is not due to geography.²

This leaves differences in the allocation of resources to explain the bulk of income differences across countries. Given the production possibilities, allocations then can differ for two reasons: differences in preferences and in misallocation. Again, there probably is something to the preference story (it seems like a plausible part of the explanation for the difference between the European and the American allocation of resources); however, at some basic level, people are people and any difference in preferences is probably itself an endogenous outcome.

This argument, then, suggests that all that remains fundamentally to explain differences in incomes across countries is misallocation, working through both traditional inputs like capital and labor and also ideas. Income differences across countries result almost entirely from the misallocation of resources.

Yet, to say misallocation is everything is perhaps not to say very much after all. In particular, the following three additional questions seem pertinent:

1. **What is the nature of the misallocation?** Are certain inputs misallocated more than others? Is the misallocation related to ideas special in any way relative to the misallocation of traditional inputs? How significant is misallocation within or between sectors or even within plants? How much misallocation is there in the richest countries?

2. **How precisely does the misallocation of resources lead to 50-fold income differences?** A simple version of this question is illustrated by the differences we observed in Figures 1 and 2. In the first case, large income differences required extreme forms of misallocation, and small improvements in the allocation of resources would have large effects on income. In the second case, neither point is necessarily true. Why does a given amount of misallocation lead to such large income differences? Why are income differences across countries generally rising over time?

3. **Why is there misallocation, and what can be done about it?** This last question takes us into the realm of political economy. The literature on political economy and on growth and development has been extremely active in the last decade; see Acemoglu, Johnson, and Robinson (2005) for an excellent overview. The state-of-the-art in that literature suggests that misallocation is the equilibrium outcome of a political process interacting with institutions and the distribution of resources (including physical capital, human capital, ideas, and natural resources). Evidently, it is not in the economic interests of the ruling elite to improve the allocation of resources, despite the potentially enormous increase in the size of the economic "pie" that is possible in the long run.

The distinction between Figures 1 and 2 is helpful in this respect and illustrates the important interaction that occurs between the production possibilities of the economy and the political process. In Figure 1, it is more difficult to understand why improvements in the allocation of resources would not take place in the most distorted countries because the immediate gains are so large. Such failures are easier to comprehend in an economy such as in Figure 2, where the immediate gains may be much smaller.

The remainder of this chapter has a much narrower focus and explores one dimension of these key questions about misallocation. In particular, it focuses on the production possibilities and seeks to understand why a given amount of misallocation can lead to large rather than small income differences.

### 3.0 Input–Output Economics

Modern economies involve sophisticated input–output structures. Goods such as electricity, financial services, transportation, information technology, and healthcare are both inputs and outputs. A wide range of intermediate goods is used to produce most goods in the economy, and these goods in turn often are used as intermediates.

Despite our intuitive recognition of this point, standard models of macroeconomics and economic growth typically ignore intermediate goods.³ The conventional wisdom seems to be that as long as we are concerned about overall value-added (GDP) in the economy, we can specify the model entirely in terms of value-added and ignore intermediate goods—hence, the neoclassical growth model. This conventional wisdom is incorrect, and the remainder of this chapter explores implications of the input–output structure of the economy for economic growth and development.

² Obviously, there is a large literature debating this question; see, for example, Gallup, Sachs, and Mellinger (1999) and Acemoglu, Johnson, and Robinson (2002).

³ Of course, there is a significant literature of exceptions, which are discussed herein.
The first insight that emerges from considering intermediate goods is that they are similar to capital. In fact, the only difference between intermediate goods and capital is one of short-run timing: intermediate goods can be installed more quickly than capital and “depreciate” fully during the course of production, whereas capital takes longer to install and only partially depreciates during production. From the point of view of the long run—the perspective relevant in most of this chapter—intermediate goods and capital are essentially the same. In particular, both are produced factors of production.

The key implications of intermediate goods for economic growth, development, and macroeconomics arise from seeing them as another form of capital. It has long been recognized that the share of capital in production is a fundamental determinant of the quantitative predictions of macromodels. When the capital share is 1/3, the intrinsic propagation mechanism of the neoclassical growth model is weak, convergence to the steady state is rapid, and the model generates a small multiplier on changes in productivity or the investment rate. In contrast, when the capital share is higher, such as 2/3, these deficiencies are largely remedied. A fairly large portion of the literature on economic growth can be viewed as an attempt to justify using a (broad) capital share of 2/3 when the data for (narrow) capital loudly proclaim that the correct number is empirically only 1/3.4

As documented carefully below, the intermediate-goods share of gross output is about 1/2 across a large number of countries. The share of capital in value-added is about 1/3, so its share in gross output is 1/6. Combining these two kinds of capital, the share of capital-goods in gross output is our “magic” number, 1/2 + 1/6 = 2/3. Incorporating intermediate goods into macroeconomic models, then, has the potential to help us understand a range of economic phenomona, including the propagation of business cycle shocks and the speed of transition dynamics. These applications are not explored here; instead, the main application in this chapter is to the puzzle of understanding why misallocation leads some countries to be 50-100 times, as opposed to only 10 times, richer than others.


We begin by providing a simple example to illustrate how and why intermediate goods lead to large multipliers. In this example, a single final output good is used as the single intermediate good in the economy, so the input–output structure is simple. Next, we build an N-sector model of economic activity, in which each sector uses the outputs from the other sectors as intermediate goods. This model is similar to the original multisectional business-cycle model of Long and Plosser (1983). The only technological difference is that we include international trade, allowing sectors to import intermediate goods from abroad. The substantive difference is in the application to economic growth and development.

Finally, we connect this model to the wealth of input–output data that exist. Data from 35 countries—including not only the currently rich countries but also Argentina, Brazil, China, and India—allow us to quantify the multiplier associated with the input–output structure of the economy.

Before continuing, it is worth noting that there is an important branch of the economics literature that studies the impact of intermediate goods. Historically, the input–output literature reigned in economics from the 1930s through the 1960s and most commonly is associated with Leontief (1936) and his followers. Hirschman (1958) emphasized the importance of sectoral linkages to economic development, which itself spawned a large literature. Hulten (1978) also is closely related, showing how intermediate goods should be included properly in growth accounting. More recently, the intermediate-goods multiplier shows up most clearly in the economic fluctuations literature; see Long and Plosser (1983); Basu (1995); Horvath (1998); Dupor (1999); Conley and Dupor (2003); Gabaix (2005); and Basu, Fernald, Fisher, and Kimball (2010). In the international-trade context, Yi (2003) argued that tariffs can multiply in much the same way that goods are traded multiple times during the stages of production. Ciccone (2002) was the first modern-growth paper I know of to develop this insight, deriving a multiplier formula for a triangular input–output structure. Jones (2011) also emphasized the importance of the intermediate-goods multiplier, albeit for a relatively restrictive input–output structure.5

3.1 A Simple Example

A simple example is helpful for understanding how intermediate goods generate a multiplier. We suppose that gross output $Q_t$ is produced using

5 Other recent examples incorporating an input–output structure include Acemoglu, Ozdaglar, and Tahbaz-Salehi (2010) and Arbex and Perobelli (2010).
capital $K_t$, labor $L_t$, and intermediate goods $X_t$:

$$Q_t = \tilde{A} \left( K_t^\rho L_t^{1-\alpha} \right)^{1-\sigma} X_t^\sigma$$  \hspace{1cm} (1)

Gross output can be used for consumption or investment or it can be carried over to the next period and used as an intermediate good. To keep things simple, we assume that a constant fraction $\bar{x}$ is used as an intermediate good:

$$X_{t+1} = \bar{x} Q_t$$  \hspace{1cm} (2)

GDP in this economy is gross output net of spending on intermediate goods: $Y_t = (1 - \bar{x}) Q_t$. In a steady state with no growth, it is easy to show that GDP will be given by:

$$Y_t = TFP \cdot K_t^\rho L_t^{1-\alpha}$$  \hspace{1cm} (3)

where

$$TPF = (\tilde{A} \bar{x}^\sigma (1 - \bar{x})^{1-\sigma})^{\frac{1}{1-\alpha}}$$  \hspace{1cm} (4)

TFP depends on the allocation of resources to intermediate goods. It will be maximized when $\bar{x} = \sigma$, which is the optimal spending share on intermediates. For any other spending share, however, TFP will be lower, and this effect will be amplified the higher is the intermediate-goods share.

Continuing, we assume that a constant fraction $\delta$ of GDP is invested:

$$K_{t+1} = \delta Y_t + (1 - \delta) K_t$$

$$= \delta (1 - \bar{x}) Q_t + (1 - \delta) K_t$$  \hspace{1cm} (5)

We assume labor is exogenous and constant.

This model features a steady state, where the level of GDP per worker $y_t = Y_t / L_t$ is:

$$y^* = \frac{Y}{L} = \left( \frac{\tilde{A} \bar{x}^\sigma (1 - \bar{x})^{1-\sigma}}{(\delta \delta)} \right)^{\frac{\sigma}{1-\alpha}}$$  \hspace{1cm} (6)

A key implication of this result is that the effects of misallocation or basic productivity differences are multiplied. In particular, we consider the simple allocation term $\bar{x}^\sigma (1 - \bar{x})^{1-\sigma}$, familiar from our steel-and-latte economy. Now that the misallocation applies to a produced good, its effects are amplified: There is an exponent of $\frac{1}{(1-\alpha)(1-\sigma)} > 1$ that applies to misallocation.

To see this more simply, we observe that a 1 percent increase in productivity $\tilde{A}$ increases output by more than 1 percent because of the multiplier $\frac{1}{(1-\alpha)(1-\sigma)}$. In the absence of intermediate goods ($\sigma = 0$), this multiplier is simply the familiar $\frac{1}{1-\alpha}$: An increase in productivity raises output, which leads to more capital, which leads to more output, and so on. The cumulative of this virtuous circle is $1 + \alpha + \alpha^2 = \frac{1}{1-\alpha}$.

In the presence of intermediate goods, there is an additional multiplier: Higher output leads to more intermediate goods, which raises output (and capital), and so on. The overall multiplier therefore is $\frac{1}{1-\bar{x}^\sigma (1-\bar{x})^{1-\sigma}}$. In fact, this multiplier also can be written as $\frac{1}{1-\bar{x}}$, where $\beta = \sigma + \alpha(1-\sigma)$ is the total factor share of produced goods in gross output, capital, and intermediates.

Quantitatively, the addition of intermediate goods has a large effect. For example, we consider the multipliers using conventional parameter values, a capital exponent of $\alpha = 1/3$, and an intermediate-goods share of gross output of $\sigma = 1/2$.

In the absence of intermediate goods, the multiplier is $\frac{1}{1-\alpha} = 3/2$, and a doubling of $\tilde{A}$ raises output by a factor of $2^{1/3} = 2.8$. However, with intermediate goods, the multiplier is $\frac{1}{1-\bar{x}^\sigma (1-\bar{x})^{1-\sigma}} = \frac{1}{2} \cdot 2 = 3$, and a doubling of $\tilde{A}$ raises output by a factor of $2^3 = 8$. As discussed in Jones (2011), if we think of the standard neoclassical factors (such as $\delta$ in the example) as generating a 4-fold difference in incomes across rich and poor countries, then this 2-fold difference in TFP leads to an 11.3-fold difference in the model with no intermediate goods but to a 32-fold difference once intermediate goods are taken into account—close to what we see in the data.\(^6\)

The deeper question in this chapter is whether this multiplier carries over into a model with a rich and realistic input–output structure. Perhaps the input–output structure in practice does not lead to these large feedback effects. Or, perhaps importing intermediate goods dilutes the multiplier substantially in practice. In fact, the remainder of this chapter shows that these concerns are not important in practice. The simple “one-over-one-minus-the-intermediate-goods-share” formula suggested by this example turns out to be a good approximation to the true input–output multiplier in modern economies. Moreover, we explicitly introduce distortions and show how this same multiplier applies to misallocation.

\(^6\) An implication of this reasoning is that it is worthy of further exploration is related to transition dynamics. A puzzle in the growth literature is why speeds of convergence are so slow, on the order of 2 percent per year, see Hauk and Wacziarg (2004) for a summary of the evidence. The standard neoclassical growth model with a capital share of one third leads to a speed of convergence of about 7 percent per year. The presence of intermediate goods would slow down this rate, just as it raises the multiplier. (A difficulty in quantifying this effect is the question of how long it takes to produce and use intermediate goods: one week, one month, or one year? That is, how long is a period?)
4.0 Preliminary Exploration of a Full Input–Output Model

We assume that the economy consists of $N$ sectors. Each sector uses capital, labor, domestic intermediate goods, and imported intermediate goods to produce gross output. In turn, this output can be used for final consumption or as an intermediate good in production.

Given this general picture, we specialize to a particular structure with two goals: (1) analytic tractability, and, (2) obtaining a model that can be connected closely to the rich input–output data. To these ends, the model augments the original Long and Plosser (1983) business-cycle model, based on Cobb–Douglas production functions, by embedding it in a model with trade.

We begin by describing the economic environment and then allocating resources using a competitive equilibrium with distortions.

4.1 The Economic Environment

Each of the $N$ sectors produces with the following Cobb–Douglas technology:

$$Q_i = A_i \left( K_i^{\alpha_i} H_i^{1-\alpha_i} \right)^{1-\gamma_i} \prod_{j=1}^{N} d_{ij}^{\sigma_{ij}} m_{ij}^{\lambda_{ij}}$$

where $i$ indexes the sector, $A_i$ is an exogenous productivity term, which is the product of aggregate productivity $A$ and sectoral productivity $\eta_i$. $A_i = A \eta_i$. $K_i$ and $H_i$ are the quantities of physical and human capital used in sector $i$. Two kinds of intermediate goods are used in production: $d_{ij}$ is the quantity of domestic good $j$ used by sector $i$, and $m_{ij}$ is the quantity of the imported intermediate good $j$ used by sector $i$. (We assume that imported intermediate goods are different so that they are not perfect substitutes; this fits with the empirical fact that countries both import and produce intermediate goods in narrow six-digit categories.) We abuse notation by assuming that there are $N$ different intermediate goods that can be imported and by indexing these by $j$ as well. The parameter values in this production function satisfy $\sigma_{ij} = \sum_{i=1}^{N} \sigma_{ij}$ and $\lambda_{ij} = \sum_{i=1}^{N} \lambda_{ij}$ and $0 < \alpha_i < 1$, so the production function features constant-returns-to-scale.

Each domestically produced good can be used for final consumption, $c_j$, or as an intermediate good:

$$c_j + \sum_{i=1}^{N} d_{ij} = Q_j, \quad j = 1, \ldots, N$$

Rather than specifying a utility function over the $N$ different consumption goods and performing a formal national income-accounting exercise, it is more convenient to aggregate these final consumption goods into a single final good through another log-linear production function:

$$Y = c_1^{p_1} \cdots c_N^{p_N}$$

where $\sum_{i=1}^{N} p_i = 1$.

This aggregate final good can be used in one of two ways – that is, as consumption or exported to the rest of the world:

$$C + X = Y$$

It is these exports that pay for the imported intermediate goods. We think of this (static) model as describing the long-run steady state of a model, so we impose balanced trade:

$$X = \sum_{i=1}^{N} \sum_{j=1}^{N} \tilde{p}_j m_{ij}$$

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

Finally, we assume fixed, exogenous supplies of physical and human capital; the effects of endogenizing physical capital in the usual way are well understood:

$$\sum_{i=1}^{N} K_i = K$$

$$\sum_{i=1}^{N} H_i = H$$

4.2 A Competitive Equilibrium with Misallocation

To allocate resources in this economy, we focus on a competitive equilibrium with distortions. As in Chari, Kehoe, and McGrattan (2007), Hsieh and Klenow (2009), Lagos (2006), and Restuccia and Rogerson (2008), distortions at the micro (here, sectoral) level can aggregate to provide differences in TFP. For simplicity, we model these distortions as sector-specific reductions in revenue, denoted $\tau_i$. These distortions literally could be taxes, but they are better thought of as representing any kind of policy that favors one sector over another (e.g., regulations, special consideration for credit). The introduction of these exogenous sector-specific distortions should be understood as a reduced-form shortcut. Ideally, we would build a
full microeconomic model in which the distortions formally correspond to explicit regulations, capital-market imperfections, market power, and theft. The insight developed here is that the misallocation associated with these distortions — whatever they are — becomes amplified by the input–output structure of the economy.

**Definition:** A competitive equilibrium with misallocation in this environment is a collection of quantities \( C, Y, X, Q_i, K_i, H_i, c_i, d_{ij}, m_{ij}, \) and prices \( p_{ij}, w, \) and \( r \) for \( i = 1, \ldots, N \) and \( j = 1, \ldots, N \) such that:

1. \( \{c_i\} \) solves the profit-maximization problem of a representative firm in the perfectly competitive final-goods market:

   \[
   \max_{\{c_i\}} \sum_{i=1}^{N} p_i c_i
   \]

   taking \( \{p_i\} \) as given.

2. \( \{d_{ij}, m_{ij}\}, K_i, H_i \) solve the profit-maximization problem of a representative firm in the perfectly competitive sector \( i \) for \( i = 1, \ldots, N \):

   \[
   \max_{\{d_{ij}, m_{ij}\}, K_i, H_i} \left( 1 - \tau_i \right) p_i A_i \left( K_i^{\alpha_i} H_i^{1-\alpha_i} \right)^{1-\delta_i} \sum_{i=1}^{N} d_{i1} \cdots d_{i2} \cdots ...
   \]

   \[
   \sum_{j=1}^{N} p_j d_{ij} - \sum_{j=1}^{N} p_j m_{ij} - r K_i - w H_i
   \]

   taking \( \{p_j\} \) as given (\( \tau_i, A_i, \) and \( p_j \) are exogenous).

3. Markets clear:
   (a) \( r \) clears the capital market: \( \sum_{i=1}^{N} K_i = K \)
   (b) \( w \) clears the labor market: \( \sum_{i=1}^{N} H_i = H \)
   (c) \( p_j \) clears the sector \( j \) market: \( c_j + \sum_{i=1}^{N} d_{ij} = Q_j \)

4. Balanced trade determines \( X \):

   \[
   X = \sum_{i=1}^{N} \sum_{j=1}^{N} \hat{p}_j m_{ij}
   \]

5. Production functions for \( Q_i \) and \( Y \):

   \[
   Q_i = A_i \left( K_i^{\alpha_i} H_i^{1-\alpha_i} \right)^{1-\delta_i} \sum_{i=1}^{N} d_{i1} \cdots d_{i2} \cdots ...
   \]

   \[
   \sum_{j=1}^{N} m_{ij} = \sum_{j=1}^{N} \hat{p}_j m_{ij}
   \]

   \[
   Y = c_1^{\delta_1} \cdots c_N^{\delta_N}
   \]

6. Consumption is the residual:

   \[
   C + X = Y
   \]

Counting loosely, there are 12 equilibrium objects to be determined and 12 equations implicit in this equilibrium definition. Concealed behind the last equation is the fact that the revenues from distortions are assumed to be lump-sum rebates to households. Because of balanced trade, however, there is no decision for households to make regarding final consumption \( C \), and it is determined simply as the residual of final output less exports.\(^7\)

### 4.3 Solving

In solving for the equilibrium of the model, it is useful to define some notation involving linear algebra, which is summarized in Table 1. Then, the following proposition characterizes the equilibrium (all proofs are given in the Appendix).

**Proposition 1 (Solution for \( Y \) and \( C \))**: In the competitive equilibrium with misallocation, the solution for total production of the aggregate final good is:

\[
Y = A^\delta K^\alpha H^{1-\alpha} \quad \text{(14)}
\]

where the following notation applies:

\[
\mu' = \frac{\beta(1-\beta)^{-1}}{1-\beta^{1-(1-\beta)}a_k}, \quad (N \times 1 \text{ vector of multipliers})
\]

\[
\bar{\alpha} = \mu' \bar{\beta}
\]

\[
\omega = \frac{\beta(1-\beta)^{-1}a_k}{1-\beta^{1-(1-\beta)}a_k}
\]

\[
\log \epsilon = \omega + \mu' \bar{\eta}
\]

Moreover, because trade is balanced, GDP for this economy is given by \( C \), which equals:

\[
C = Y \left( 1 - \sum_{i=1}^{N} \sum_{j=1}^{N} (1-\tau_i)Y_i \lambda_{ij} \right) \quad \text{(15)}
\]

There are several points of this proposition that merit discussion. First, and not surprisingly, our \( N \)-sector Cobb–Douglas model aggregates to yield a Cobb–Douglas aggregate-production function. More interesting, aggregate TFP depends on both sectoral TFPs and the underlying distortions. The latter point requires digging into the \( \epsilon \) term, where distortions then

\(^7\) Given Walras's law, this equation could be replaced by \( C = w H + r K + T \), where \( T \) is the lump-sum rebate of the revenue obtained by the distortion. We notice that all production either is consumed or exported.
The matrix $L = (I - B)^{-1}$ is known as the Leontief inverse. The typical element $\ell_{ij}$ of this matrix can be interpreted in the following way: Ignoring trade for the moment, a 1 percent increase in productivity in sector $j$ raises output in sector $i$ by $\ell_{ij}$ percent. This result takes into account all of the indirect effects at work in the model. For example, raising productivity in the electricity sector makes banking more efficient and this, in turn, raises output in the construction industry. The Leontief inverse incorporates these indirect effects. (Notice that it is the matrix equivalent of $1/(1 - \sigma_i)$.)

Multiplying this matrix by the vector of value-added weights in $\beta$ leads to $\beta'(I - B)^{-1} = \sum_{i=1}^{N} \beta_i \ell_{ij}$. That is, we add up the effects of sector $j$ on all of the other sectors in the economy, weighting by their shares of value-added. The typical element of this multiplier matrix then reveals how a change in productivity in sector $j$ affects overall value-added in the economy.

This would be precisely correct if $\lambda_{ij}$ were zero—that is, in the absence of trade. In the presence of trade, this multiplier is adjusted by the factor $1/(1 - \beta'(I - B)^{-1}\lambda)$. We discuss this factor in more detail below, but it is enough to note that this factor is larger than 1: Trade strengthens rather than attenuates the multiplier.

The elasticity of final output with respect to aggregate TFP is $\bar{\mu} = \mu'$. That is, we sum all of the multipliers in $\mu$ because an increase in aggregate TFP affects not only sector $j$ but also all of the sectors.

A final remark about Proposition 1 concerns the capital exponent in the aggregate-production function, $\delta = \mu' \delta_k$. Recall that $\delta_k$ is the vector of capital exponents $\alpha_i (1 - \sigma_i - \lambda_i)$. The aggregate exponent therefore is a weighted average of the sectoral-capital shares, where the weights depend on the intermediate-goods share. This remark makes more sense after the next proposition.

5.0 Special Cases, to Build Intuition

5.1 The Multiplier in a Special Case

The linear-algebra formula is a useful theoretical result and proves convenient when we apply the model to the rich input-output data that exist. However, analyzing a special case can be helpful in obtaining intuition for how the model works.

We consider the following special case. Suppose that all sectors have the same cumulative elasticities of output with respect to domestic and imported intermediate goods, although the composition across sectors is allowed to vary. For example, one sector may use a lot of electricity and steel
whereas another uses a lot of financial services and information technology. The composition can vary across sectors, but suppose that each sector spends 50 percent of its revenue on intermediate goods. What does the multiplier look like in a case such as this?

The following proposition provides the answer. In fact, it allows for imported intermediate goods as well (where the overall share spent on these goods is the same in each sector).

**Proposition 2 (Multiplier in a Special Case):** Assume \( \sigma_i = \sum_{j=1}^{N} \sigma_{ij} = \hat{\sigma} \) and \( \lambda_i = \sum_{j=1}^{N} \lambda_{ij} = \hat{\lambda} \) for all \( i \), where \( \hat{\sigma} \) and \( \hat{\lambda} \) are positive scalars whose sum is less than 1, and define \( \hat{\sigma} = \hat{\sigma} + \hat{\lambda} \) to be the total intermediate-goods share. Then:

\[
\frac{\partial \log Y}{\partial \log A} = \mu'1 = \frac{\beta'(I - B)^{-1}1}{1 - \beta'(I - B)^{-1}\hat{\lambda}} = \frac{1}{1 - \hat{\sigma}}
\]

This special case makes two general points about the model. First, the “sparseness” of the input–output matrix \( B \) is not especially important. For example, our special case includes a “clock” structure, in which every sector uses as an input only the good produced by the sector above it. It also includes the case in which every sector uses only its own output. In both cases, the input–output matrix is very sparse, with zeros almost everywhere. Yet, the overall multiplier remains equal to “one-over-one-minus-the-intermediate-goods-share.” This special case suggests that if the overall intermediate-goods share is about 1/2, we should not be surprised to find a multiplier of about 2. This intuition is confirmed in the next section, when we turn to quantitative results.

The second key point made in this proposition is that the intuition that imports would dilute the multiplier is a “red herring.” In fact, there is no dilution at all: In the proposition, it is the overall intermediate-goods share \( \hat{\sigma} \equiv \hat{\sigma} + \hat{\lambda} \) that matters for the multiplier, and the composition between domestic and imported goods is completely irrelevant.

Why is this the case? The answer is that we have imposed balanced trade in our (long-run) model. Therefore, exports are used to “produce” imports. A higher productivity in the domestic computer-chip sector raises overall exports, which in turn increases imports; therefore, the virtuous circle is not broken by the presence of trade.\(^8\)

\(^8\) This assumption of balanced trade is the key difference that makes the intuition from the Keynesian business-cycle model inappropriate. In the business-cycle context, an increase in exports leads to a trade surplus and does not increase imports.

![Figure 3. Consumption versus the average distortion, \( \hat{\tau} \). Note: This example is drawn for \( \hat{\sigma} + \hat{\lambda} = 1/2 \). Notice the similarity to Figure 1.](image)

### 5.2 Symmetry and Distortions

The second special case allows us to study the multiplier associated with distortions. First, we consider a world where the intermediate-goods share of production is the same in every sector and there is a symmetric distortion at rate \( \tau_i = \hat{\tau} \). In this case, GDP in the economy is given by the following proposition.

**Proposition 3 (Symmetry and Distortions):** Suppose that \( \sigma_{ij} = \hat{\sigma} / N \), \( \lambda_{ij} = \hat{\lambda} / N \), \( \beta_i = 1 / N \), \( \alpha_i = \alpha \), and \( \tau_i = \hat{\tau} \). Then:

\[
\log C = \text{Constant} + \frac{\hat{\sigma}}{1 - \hat{\sigma}} \log(1 - \hat{\tau}) + \log(1 - \hat{\sigma}(1 - \hat{\tau}))
\]

where \( \hat{\sigma} = (\hat{\sigma} + \hat{\lambda}) \) denotes the total intermediate-goods share and Constant is a collection of terms that does not depend on \( \hat{\tau} \). Moreover, consumption is an inverse-U-shaped function of the distortion rate, with a peak that occurs at \( \hat{\tau} = 0 \).

An example of this proposition is shown in Figure 3. Notice that the effect of a change in the distortion rate on GDP essentially depends on \( \hat{\sigma} \). If there are no intermediate goods in this economy, output distortions have no effect. This is because the distortions here represent a violation of the Diamond and Mirrlees (1971) dictum of “no taxation of intermediate goods.” In our
(current) setup, \( K \) and \( H \) are nonproduced factors, so a symmetric tax does not distort the allocation of capital.\(^9\)

The key distortion is between consumption and intermediate goods. A good that is consumed suffers the distortion only once when the good is produced; a good that is used as an intermediate is distorted when it is first produced and then again when it is used as an intermediate. Because a constant fraction of output is consumed and the rest is used as an intermediate good, this process suffers from the vicious cycle of the multiplier.

Symmetric distortions affect GDP through the two terms in Equation (17). The first term is the direct effect, in which distortions enter the model much like productivity: We recall that both \( 1 - \tau_j \) and \( A_i \) are subject to the multiplier effect through the \( \epsilon \) term in Proposition 1. The second term somewhat mitigates this effect and captures the indirect effect, whereby higher distortions raise consumption (i.e., by reducing the purchase of intermediate goods).

5.3 Symmetry with Random Distortions
The final special case allows us to consider variation in distortions across sectors. We suppose that everything in the model other than distortions is symmetric, and we allow distortions to be a log-normally distributed random variable.

**Proposition 4 (Symmetry with Random Distortions):** We suppose that \( \sigma_{ij} = \frac{\theta_i}{N}, \lambda_{ij} = \frac{\lambda_j}{N}, \) and \( \beta_i = 1/N \), and we let \( \delta = \delta + \lambda \). We assume \( \log(1 - \tau_j) \sim N(\theta,\nu^2) \) and let \( 1 - \tau = e^{\theta + \frac{1}{2}\nu^2} \) reflect the average distortion. Then:

\[
\text{plim}_{N \to \infty} \text{log } C = \text{Constant} + \frac{\delta}{1 - \delta} \cdot (1 - \tau)
+ \log(1 - \delta(1 - \tau)) - \frac{1}{1 - \delta} \cdot \frac{1}{2} \cdot \nu^2
\]

where Constant is a collection of terms that do not depend on \( \theta \) or \( \nu^2 \). Moreover, consumption is maximized when there are no distortions.

In terms of the mean effect of distortions, this result is identical to the previous result. Now, however, we have an additional result related to the variance of distortions across sectors. In particular, a higher variance of distortions reduces GDP, even in the absence of intermediate goods, because random distortions will distort the allocation of capital and labor across sectors. However, the variance term is subject to the now-familiar multiplier effect associated with \( 1/(1 - \delta) \). A higher variance of distortions is more costly in an economy with intermediate goods. This makes sense: The first best in this economy is to have no distortions. Either a constant or a random tax distorts the allocation of resources and reduces GDP. The magnitude of the distortion depends on the Diamond–Mirrlees effect—that is, how important intermediate goods are in production. An example is illustrated in Figure 4.

5.4 Random Distortions with a General Input–Output Structure
So far, I lack the correct combination of time, talent, and insight to derive a more general result for the effect of log-normal distortions in the presence of the full input–output structure, although I suspect that good results are possible along these lines. Intuition from the previous propositions strongly suggests that something like the general multiplier \( \mu \) will continue to have a crucial role. For the empirical applications that follow, we therefore focus on this general multiplier.
6.0 Quantitative Analysis

We now turn to the rich input–output data that exist for both the United States and many other countries. These data allow us to calculate aggregate and sectoral multipliers and to study the effect of sectoral distortions on aggregate GDP. First, we use the six-digit-level data available from the Bureau of Economic Analysis (BEA) for the United States in 1997. Then, we turn to the Organisation for Economic Co-operation and Development (OECD) Input–Output Database, which contains data for 48 industries and 35 countries.

6.1 The U.S. Input–Output Data, 480 Industries

Figure 5 is close to the $B$ matrix for the United States, using the 480 commodities in the BEA’s 1997 benchmark input–output data. Actually, the figure plots the matrix of $\sigma_{ij} + \lambda_{ij}$ instead, so that the entries show the overall exponents on intermediate goods used in producing each of the 480 goods. A contour-plot method is used, showing only those shares greater than 2, 4, and 8 percent.

Three points stand out in the figure: (1) there is a strong diagonal; (2) the matrix is relatively sparse; and (3) there are a few exceptions to this sparseness: a few key goods are used by numerous industries in a significant way, including wholesale trade, trucking, management of companies, real estate, paperboard products, and iron and steel mills.

Table 2 reports basic statistics of the U.S. input–output matrix that help place these visual conclusions in context. Although the diagonal elements were important visually, the table points out that these elements typically are small: the mean is only 3.3 percent and the median is only 1.0 percent. This is true despite the fact that a typical industry pays a large share...
of its gross output to intermediate goods: 56.4 percent at the mean. The
industry at the 75th percentile pays out about two thirds of its revenue to
intermediate goods, whereas even the industry at the 25th percentile pays
nearly half. Along these lines, it is worth noting that even though only
0.13 percent of the elements of the input–output matrix exceed 10 per-
cent, this is still 288 elements overall – on average, about once every two
sectors. Similarly, 83 of the entries are greater than 20 percent. As shown
at the bottom of the table, the overall intermediate-goods share for the
U.S. economy is about 43.4 percent; service industries are more impor-
tant as a share of value-added and they have lower intermediate-goods
shares.

The last part of the table computes the aggregate multiplier using the
six-digit input–output data. The overall multiplier is 1.65. This number
is the product of a domestic multiplier of 1.61 (which would obtain if
no intermediate goods were imported) and an import multiplier of 1.03.
Imports are relatively unimportant in the multiplier.

To what extent is the simple \( \frac{1}{1-x} \) formula accurate? The multiplier of
1.65 would result from this formula if the intermediate-goods share were
0.394. In fact, the intermediate-goods share using this six-digit data is 0.434.
This simple aggregate formula appears to give a good approximation to the
result found by computing the \( 480 \times 480 \) Leontief inverse, although there is
a small degree of dilution: Applying the formula to the 0.434 share suggests
a multiplier that overstates the truth by about 10 percent.

6.2 The OECD Input–Output Data, 48 Industries

The 2006 edition of the OECD Input–Output Database contains input–
output data for 35 countries and 48 industries, typically for 2000. In addition
to covering OECD countries, the data include poor and middle-income
countries such as China, India, Argentina, Brazil, and Russia.

Figure 6 shows the input–output matrix for the United States at this higher
level of aggregation. The pattern at the more detailed level of aggregation
of a sparse matrix with a strong diagonal and only a few goods that are used
widely is repeated at this higher level of aggregation.

A nice feature of the OECD data is that we can consider the question
of how much the input–output structure of an economy differs across
countries. The general and perhaps surprising answer that we obtain is “not
much.” Figure 7 shows the input–output matrix for two countries, Japan
and China, as an example.

![Figure 6. The U.S. input–output matrix, 2000 (48 industries). Note: The plot shows the
matrix \([A_{ij} + \lambda I]\), that is, the matrix of intermediate-goods shares for 48 industries.
A contour-plot method is used, showing only those shares greater than 2, 5, 10 and
20 percent. Source: OECD 2006 database.]

The matrix for Japan appears similar to the matrix for the United
States. This is true more generally, especially for the richer countries in
the dataset, but it is true even for the poorer countries. The input–output
matrix for China is perhaps most different from the United States, but
the overall structure still is similar. Electricity appears as being noticeably
more important, and other business activities (e.g., advertising, accounting,
and legal services) as somewhat less important. These are the main
differences.

The first column of Table 3 makes these comparisons more systematically.
It shows the fraction of elements in the input–output matrix that differ by
more than 0.02 from the corresponding elements in the U.S. input–output
matrix. Slightly more than 16 percent of the elements exceed this difference
in China’s input–output matrix, whereas the corresponding number for
Japan is about 9 percent. For this level of the cutoff, the average across the
35 countries is 11 percent. If we lower the cutoff to 0.01, the typical country
has differences of this magnitude in slightly more than 20 percent of the
Figure 7. Input–output matrix in Japan and China (48 industries).
cells. If we raise the cutoff to 0.05, the average across countries is 3.9 percent of cells.

Figure 8 shows the aggregate multipliers, $\tilde{\mu}$, for the 35 countries in our sample. The average value for the multiplier in this sample is about 1.9. It ranges from a high of 2.53 in China to lows of 1.51 and 1.59 in Greece and India respectively. It is interesting that China and India are two of the poorest countries in the sample, and they have widely different multipliers. The multiplier for the United States using this data works out to be 1.77, slightly higher than what we found in the six-digit data.

Table 3 shows these multipliers in more detail, including the contribution from imported intermediate-goods as well as the aggregate intermediate goods share and the “as-if” share that corresponds to the multiplier computed using the Leontief inverse. The simple approximation of “one-over-one-minus-the-intermediate-goods-share” approximates well the true multiplier.

6.3 Take-Away from the Input–Output Data

What do we learn from the input–output data? First, the common $1/1-\sigma$ formula that emerges from simple models of intermediate goods is remarkably robust: More careful analysis with full input–output structures across a range of economies suggests that the basic multiplier from simple models carries over quite well. Working with simpler models, then, may be appropriate.

Second, there is a surprising degree of similarity in these matrices across countries. This is surprising in that we might have expected significant differences for technological reasons as well as reasons related to misallocation. Technologically, countries at different levels of development presumably produce with different technologies, and we might have expected to see this more strongly in the input–output structure of these economies. This is particularly true given the specialization arguments associated with international trade.

On the misallocation front, we appreciate that many distortions that might be present would show up by changing observed factor shares, even if the underlying technologies were the same. One way to see that is to recall the first-order condition in a simple neoclassical growth model with Cobb–Douglas production: $(1-\tau)\alpha K/Y = r$. Firms rent capital until the postdistortion marginal product falls to equal the rental rate. However, in this case, $r K/Y = \alpha (1-\tau)$; therefore, the observed capital share will differ from the technological parameter by the distortion rate.

This, in turn, has important implications. There is a fundamental identification problem: We see data on observed intermediate-goods shares and we do not know how to decompose that data into distortions and differences in technologies. This identification problem is not solved in anything presented here. Instead, I simply show that the observed spending shares are remarkably similar across countries.

My tentative conclusion given this fact is that the misallocation across four-digit sectors is not particularly large in this sample of countries. Without solving the basic identification problem, however, this conclusion must remain tentative. A useful way to check this is to assume that the U.S. input–output structure measures the true technology for all countries and to use observed spending shares on intermediate goods to measure the distortions that apply, on average, across the four-digit sectors. This would be a valuable exercise. Of course, we could question the assumption that the underlying technologies in all countries are the U.S. factor shares. Moreover, this approach would not measure the distortions that apply within each sector, which may be important in practice. Redoing the Hsieh and Klenow (2009) analysis using gross output and intermediate goods within sectors for China and India (and other countries) also would be valuable.
7.0 Conclusion

One of the most exciting directions in the growth literature in recent years is the recognition that the misallocation of resources at the microlevel can aggregate to look like differences in TFP. Quantifying these effects in novel ways – two examples being the extensive use of firm-level data and the exploration of input–output tables – is yielding new insights on why some countries are so much richer than others, and it likely has a promising future.

APPENDIX: PROOFS OF THE PROPOSITIONS

Proofs of the Propositions 1 (Solving for $Y$ and $C$): We begin by considering the profit-maximization problems for the final-goods and intermediate-goods firms. For the final-goods firms, we recall that the problem is:

$$\max_{\{\beta_i\}} \beta_1 c_1 \cdots \beta_N c_N - \sum_{i=1}^N p_i c_i$$

This yields a first order condition (FOC) where the spending shares are the exponents:

$$\frac{p_i c_i}{Y} = \beta_i$$  \hspace{1cm} (18)

Next, we consider the intermediate firms:

$$\max_{d_{ij}, m_{ij}, k_i, l_i} (1 - \tau_i) p_i A_i \left( K_i^{-\alpha_i} - K_{i-1}^{\alpha_i} \right) \cdots d_{i1} \cdots d_{iN} m_{i1} \cdots m_{iN}$$

$$- \sum_{j=1}^N p_j d_{ij} - \sum_{i=1}^N p_i m_{ij} - r K_i - w H_i$$

The first-order conditions for this problem are:

$$(1 - \tau_i) \alpha_i (1 - \sigma_i - \lambda_i) \frac{p_i Q_i}{K_i} = r$$  \hspace{1cm} (19)

$$(1 - \tau_i) (1 - \alpha_i) (1 - \sigma_i - \lambda_i) \frac{p_i Q_i}{H_i} = w$$  \hspace{1cm} (20)

$$\frac{1 - \tau_i}{\sigma_{ij}} \frac{p_i Q_i}{d_{ij}} = p_j$$  \hspace{1cm} (21)

$$\frac{1 - \tau_i}{\lambda_{ij}} \frac{p_i Q_i}{m_{ij}} = p_j$$  \hspace{1cm} (22)

Now we are ready to use these FOCs to solve for some allocations. We begin with the resource constraint for sector $j$:

$$c_j + \sum_{i=1}^N d_{ij} = Q_j$$

We now use Equation (21) to delete $d_{ij}$ and rearrange slightly to obtain:

$$\frac{p_j c_j + \sum_{i=1}^N (1 - \tau_i) \sigma_{ij} p_i Q_i}{c_j} = p_j Q_j$$

Finally, from the first-order condition for the final-goods firm, $p_j = \beta_j Y / c_j$. Using this expression for $p_j$ and canceling $Y$ from both sides of the equation gives:

$$\beta_j + \sum_{i=1}^N (1 - \tau_i) \sigma_{ij} \frac{\beta_i Q_i}{c_i} = \beta_j Q_j$$  \hspace{1cm} (23)

Now, we define $v_j \equiv \frac{\beta_j Q_j}{c_j}$ and we let $v$ denote the $N \times 1$ vector of $v_j$.

Then, we can stack the $N$ equations in Equation (23) to obtain an equation involving vectors and a matrix:

$$\beta + \tilde{B}^T v = v$$  \hspace{1cm} (24)

where $\beta$ is the $N \times 1$ vector of final-goods exponents and $\tilde{B}$ is the $N \times N$ matrix of intermediate-goods shares adjusted for taxes; a typical element is $(1 - \tau_i) \sigma_{ij}$. This equation solves easily to give:

$$v^* = (I - \tilde{B}^T)^{-1} \beta \equiv \gamma$$  \hspace{1cm} (25)

We notice that this defines the solution for $\beta_j Q_j / c_j$ as $\gamma_j$. It is easy to show, in fact, that these elements are also the solution for $p_j Q_j / Y = \gamma_j$, so that the $\gamma_j$ terms are the “Domar” weights – the ratio of total spending on intermediate good $j$ to $Y$.

At this point, we can use this solution to obtain useful expressions for $d_{ij}$ and $m_{ij}$ as well. Equation (18) implies that $\frac{p_i}{p_j} = \frac{\beta_i}{\beta_j} \cdot \frac{Q_i}{c_i}$. Substituting this into the FOC for $d_{ij}$ in Equations (21) leads to:

$$d_{ij} = (1 - \tau_i) \sigma_{ij} \gamma_j \frac{Y}{\gamma_j} \cdot Q_j$$  \hspace{1cm} (26)

Similarly, $m_{ij}$ satisfies:

$$m_{ij} = (1 - \tau_i) \lambda_{ij} \gamma_i Y / \beta_j$$  \hspace{1cm} (27)
The FOCs for $K_i$ and $H_i$ similarly yield:

$$
\frac{K_i}{K} = \frac{(1 - \tau_i)(1 - \sigma_i - \lambda_i)\omega_i y_i}{\sum_i(1 - \tau_i)(1 - \sigma_i - \lambda_i)\alpha_i y_i} \equiv \bar{\theta}_i
$$

(28)

and

$$
\frac{K_i}{H_i} = \frac{(1 - \tau_i)(1 - \sigma_i - \lambda_i)(1 - \alpha_i)\gamma_i}{\sum_i(1 - \tau_i)(1 - \sigma_i - \lambda_i)(1 - \alpha_i)\gamma_i} \equiv \bar{\theta}_i
$$

(29)

In this expression, the “bars” over the $\bar{\theta}$s denote that these expressions include the $(1 - \tau_i)$ terms in the numerator; this is useful in the subsequent discussion.

Now we can substitute the expressions in Equations (26) through (29) in the main production function for intermediate goods. Factoring out the $1 - \tau_i$ terms (and letting $\bar{\theta}_K = \bar{\theta}_{K_i}(1 - \tau_i)$) gives:

$$
Q_i = A_i(1 - \tau_i)^{(\bar{\theta}_K,K)^{\bar{\sigma}}(\bar{\theta}_H,H)^{\bar{\lambda}} - \bar{\sigma}_i - \bar{\omega}_i}
\cdot \prod_{j=1}^{N}(\sigma_{ij}Q_jy_j/y_j)^{\bar{\sigma}_{ij}} \cdot \prod_{j=1}^{N}(\lambda_{ij}y_j/y_j)^{\bar{\lambda}_{ij}}
$$

(30)

Taking logs of this expression, stacking into a vector, and using much of the notation in Table 1 gives:

$$
q = \vec{a} + \vec{\omega}_q + \delta_K \log K + \delta_H \log H + Bq + \lambda \log Y
$$

(31)

where $q$ is the vector with typical element $\log Q_i$, $\vec{a}$ is a vector with typical element $\log A_i(1 - \tau_i)$, and $B$ is the matrix of $\sigma_{ij}$ (see the table for the other notation).

This equation can be solved to yield:

$$
q = (I - B)^{-1}(\vec{a} + \vec{\omega}_q + \delta_K \log K + \delta_H \log H + \lambda \log Y)
$$

(32)

We use this expression shortly. First, however, we return to the final-goods production function. We recall that $c_i = \beta_i Q_i/y_i$. Taking logs and stacking into a vector, gives:

$$
c = \omega_c + q
$$

(33)

where $\omega_c$ is a vector with typical element $\log \beta_i/y_i$ and $c$ denotes the vector of log $c_i$.

Then, from the final-goods production function and using the vector notation, $\log Y = \beta'c = \beta'\omega_c + \beta'q$.

We now are at the last step. We substitute (32), Equation in this last expression to obtain:

$$
\log Y = \beta'\omega_c + \beta'(I - B)^{-1}(\vec{a} + \vec{\omega}_q + \delta_K \log K + \delta_H \log H + \lambda \log Y)
$$

(34)

We let $A_i = A_{\bar{\theta}_{K_i}}$. This last equation then can be solved for $\log Y$ to yield:

$$
\log Y = \bar{\mu} \log A + \bar{\alpha} \log K + (1 - \bar{\alpha}) \log H + \log \epsilon
$$

(35)

where the notation used is that in Table 1.

The expression for consumption comes from using balanced trade and the expression for $m_{ij}$ given in Equation (22).

Proofs of the Propositions 2 (The Multiplier in a Special Case): In matrix notation, the assumption that all sectors have a cumulative domestic intermediate-goods share of $\bar{\sigma}$ is simply $B1 = \bar{\sigma}1$. This implies the following:

$$(I - B)1 = (I - \bar{\sigma})1
1 = (I - B)^{-1}1 \cdot (1 - \bar{\sigma})
1 = \beta1 = \beta'(I - B)^{-1}1 \cdot (1 - \bar{\sigma})
\Rightarrow \beta'(I - B)^{-1}1 = \frac{1}{1 - \bar{\sigma}}
$$

Similarly, $\beta'(I - B)^{-1}1 = \frac{1}{1 - \bar{\sigma}}$. Therefore:

$$
\mu1 = \frac{\beta'(I - B)^{-1}1}{1 - \beta'(I - B)^{-1}1} = \frac{1}{1 - (\bar{\sigma} + \lambda)} = \frac{1}{1 - \bar{\sigma}}
$$

Proofs of the Propositions 3 (Symmetry and Distortions): The key step in solving the model is to use the same general result as in the previous proposition: If a matrix $X$ has rows that sum to the same value, $\tilde{x}$, then $(I - X)^{-1}1 = 1 \cdot \frac{1}{1 - \tilde{x}}$. In this case, this result is used to compute $\gamma = (I - B)^{-1}\beta$, where $\beta_i = 1/N$. Because of the symmetry, we obtain:

$$
\gamma_i = \gamma = \frac{1}{1 - \sigma(1 - \tilde{x})}
$$

Substituting this result into the FOCs, we obtain:

$$
d_{ij} = (1 - \tau_i)\sigma Q_j/N
m_{ij} = (1 - \tau_i)\hat{\lambda}_i/N\gamma y_j/\tilde{p}_j
$$

as well as $K_i/K = H_i/H = 1/N$.

Using these conditions to compute the allocation terms, we obtain:

$$
\omega_{qi} = -(1 - \sigma - \hat{\lambda})\log(1 - \tilde{x}) - \hat{\lambda}\log(1 - \sigma(1 - \tilde{x})) + \text{Constant}
$$

where the constant does not depend on $\tau$ or the sector $i$. In addition, $\beta'\omega_c = \log(1 - \sigma(1 - \tilde{x}))$. 
Following the definitions in Table 1, these expressions can be combined to yield the allocation term:

$$\omega = \log(1 - \hat{\sigma} (1 - \hat{\tau})) - \log(1 - \hat{\tau}) + \text{Constant}$$

for another constant that does not depend on \(\hat{\tau}\). Substituting this into the definition of \(\epsilon\) gives:

$$\log \epsilon = \log(1 - \hat{\sigma} (1 - \hat{\tau})) + \frac{\hat{\sigma} + \hat{\lambda}}{1 - (\hat{\sigma} + \hat{\lambda})} \log(1 - \hat{\tau}) + \text{Constant}$$

Furthermore, \(\log Y\) will have the same form.

Finally, we need the expression for GDP, \(C\). We recall that because of balanced trade:

$$C = Y \left(1 - \sum_{i=1}^{N} \sum_{j=1}^{N} (1 - \tau_i) Y_{ij} \lambda_{ij}\right)$$

Similar arguments give:

$$C = Y \left(\frac{1 - (\hat{\sigma} + \hat{\lambda})(1 - \hat{\tau})}{1 - \hat{\sigma} (1 - \hat{\tau})}\right)$$

Substituting this into the expression for \(Y\) (or rather \(\log \epsilon\)) completes the proof.

**Proofs of the Propositions 4** (Symmetry with Random Distortions): This proof follows fairly closely that for Proposition 3. The main exception is in obtaining the solution for \(\mu' \ell\), where \(\ell\) denotes the \(N \times 1\) vector with typical element \(\log(1 - \tau_i)\). This proceeds as follows. Using the fact that \(\beta_i = 1/N:\)

$$\mu' \ell = \frac{1 - \hat{\sigma}}{1 - (\hat{\sigma} + \hat{\lambda})} \cdot \frac{1}{N} \cdot 1'(I - B)^{-1} \ell$$

The key step then is to show that \(1'(I - B)^{-1} = \frac{1}{1 - \hat{\sigma}} \cdot 1'\), which can be done as follows. We let \(\sigma\) denote an \(N \times 1\) vector where each element is \(\hat{\sigma}\). Then, because \(B\) is an entire matrix of \(\hat{\sigma}\), it follows that:

$$1' B = \sigma'$$

Then:

$$1' - 1' B = 1' - \sigma'$$

and, therefore:

$$1'(I - B) = (1 - \hat{\sigma}) 1'$$

Postmultiplying by \((I - B)^{-1}\) and rearranging gives the result we want:

$$1'(I - B)^{-1} = \frac{1}{1 - \hat{\sigma}} \cdot 1'\). When this is substituted in Equation (36), we obtain the expected result that:

$$\mu' \ell = \frac{1}{1 - (\hat{\sigma} + \hat{\lambda})} \cdot \frac{1}{N} \sum \log(1 - \tau_i)$$

Following the same arguments used in proving Proposition 3, we obtain:

$$\log C = \log(1 - (\hat{\sigma} + \hat{\lambda})(1 - \hat{\tau})) - \log(1 - \hat{\tau})$$

$$+ \frac{1}{1 - (\hat{\sigma} + \hat{\lambda})} \cdot \frac{1}{N} \sum \log(1 - \tau_i) + \text{Constant}$$

The remainder of the proposition follows naturally from the log-normal assumptions on \(1 - \tau_i\).

**References**


