Comparative Advantage, Complexity, and Volatility*

Pravin Krishna
Johns Hopkins University
and NBER

Andrei A. Levchenko
University of Michigan
and IMF

May 18, 2009

Abstract

Less developed countries tend to experience higher output volatility, a fact that is, in part, explained by their specialization in more volatile sectors. This paper proposes theoretical explanations for this pattern of specialization – with the complexity of the goods playing a central role. Specifically, less developed countries with low levels of human capital, or alternately, with lower institutional ability to enforce contracts, will specialize in less complex goods which are also characterized by higher levels of output volatility. We provide novel empirical evidence that less complex industries are indeed more volatile.

JEL Classifications: F40

Keywords: Product Complexity, Comparative Advantage, Volatility

*We are grateful to Arnaud Costinot, Miklós Koren and workshop participants at the University of Michigan and the 2009 AEA Meetings for helpful suggestions. The views expressed in this paper are those of the authors and should not be attributed to the International Monetary Fund, its Executive Board, or its management. Email: Pravin_Krishna@jhu.edu, alev@umich.edu.

1 Introduction

Understanding the sources of volatility is an important quest in economics. In a seminal paper, Lucas (1988) observed that, over long horizons, fluctuations in rates of growth are likely to be more substantial in less developed countries, suggesting a link between a country's level of economic development and its volatility. Indeed, as Figure 1 illustrates, the negative relationship between a country's level of development and aggregate volatility is quite pronounced. Analyzing the sources of this differential aggregate volatility across countries, Koren and Tenreyro (2007) show that an important explanation for the higher output volatility in developing countries is their production specialization in more volatile sectors. Figure 2 depicts the weighted average volatility of the sectors in which a country specializes against the level of development. More developed countries tend to specialize, on average, in less volatile sectors.

At the same time, several recent empirical studies have suggested that openness to international trade plays an important role in determining economic volatility (see, e.g., Rodrik 1998, Krebs, Krishna and Maloney 2008, di Giovanni and Levchenko 2007, di Giovanni and Levchenko 2008). While the empirical literature has variously suggested links between a country's level of development, pattern of production specialization, trade, and economic volatility, the reasons for these patterns are not well understood – that is, a coherent theoretical explanation linking these factors is, as yet, lacking.

This paper develops a theoretical framework in which openness to international trade leads to specialization in more volatile sectors in poorer countries – consistent with the empirical findings of Koren and Tenreyro (2007) mentioned above. In our framework, the central concept driving the linkage between trade openness, specialization, and volatility is the complexity of goods being produced. Complexity is defined as the number of different inputs required for the production of one unit of the good (as in Becker and Murphy 1992). We show that sectoral output volatility depends on the complexity of goods produced in that sector. This is because when individual inputs to production are subject to shocks, the volatility of output will depend on how many such inputs there are. In particular, the more complex goods are less volatile, as the production in a sector that uses many inputs will be less affected, on average, by shocks to any particular input (a point also emphasized by Koren and Tenreyro 2008). By contrast, the volatility of a good that uses very few inputs will be more affected by the shocks to each individual input.

Starting from this technological characterization of industries in terms of their product complexity, we model two mechanisms through which less developed countries come to exhibit comparative advantage in the less complex − and therefore more volatile − goods. The first, building on Blanchard and Kremer (1997) and Levchenko (2007), relies on differences in the quality of contract enforcement. The more complex the production process, the greater is the number of parties to production, and the greater is the number of contracts that it requires. This implies that the relative loss of output due to imperfect contract enforcement is greater in the more complex sectors in countries with worse institutions, generating comparative advantage.

The second approach, building on Costinot (2009), relies on the differences in human capital endowments across countries and the optimal division of labor in production. In this second mechanism, the scope of the division of labor in production is determined endogenously as a function of the complexity of goods. Countries with higher levels of human capital per worker have a comparative advantage in the more complex goods because higher human capital allows each worker to learn more of the necessary production tasks (as we discuss in detail below).[1]

Thus, openness to international trade moves less developed countries towards the production of less complex and more volatile goods. This is the main theoretical result obtained in

[1] While we use the theoretical frameworks of Blanchard and Kremer (1997) and Costinot (2009) as building blocks of our analysis, it is worth noting that neither of these important papers is concerned with the question of economic volatility, which is the focus of the present analysis.

the paper. The relationship between economic development and volatility in our framework is then driven by two mechanisms: the specialization in less complex goods by less developed countries and the greater volatility of goods with lower complexity. The first theoretical prediction – that less developed economies will specialize in less complex goods – is supported by several recent empirical studies. For instance, Levchenko (2007) has shown that countries with worse institutions have relatively higher export shares in goods with low product complexity − with complexity measured as the number of intermediates required for production in each sector. Similarly, Costinot (2009) has found that less developed countries specialize in less complex goods, with complexity measured as the average learning cost that a worker must pay in each sector before becoming productive. Finally, Nunn (2007) has demonstrated that less developed countries specialize in industries requiring less "relationship-specific" investments in their production – which could also be interpreted as industries with a lower degree of product complexity. By way of illustration of these results, Figure 3 shows that there is a pronounced positive relationship between the average complexity of a country's specialization pattern and the level of development: richer contries tend to specialize in more complex goods.

The second theoretical mechanism on which our paper relies − that less complex goods are characterized by greater volatility − has not previously been analyzed empirically in the literature. In this paper, we provide evidence regarding this relationship. Using data on sectoral production data from the NBER Productivity Database, we calculate industry-level volatility measures for some 460 4-digit SIC87 sectors over the period 1970-1997. We combine the volatility data with empirical measures of product complexity computed from the U.S. Input-Output tables. Our results demonstrate that there is a strong negative relationship between complexity and volatility, with complexity alone explaining some 18% of the variation in the actual volatility found in the data. The results are robust to a number of controls, such as factor intensity and sector-level elasticity of substition.

In sum, this paper contributes to the literature on economic development and international trade, by linking the patterns of comparative advantage with volatility. In our framework, production specialization in more volatile sectors takes place in poorer countries and emerges naturally from differences in complexity of goods and in the productivity of input factors across countries. The theoretical predictions are consistent with stylized empirical facts.

The rest of the paper is organized as follows. Section 2 describes the theoretical framework and derives the main results. Section 3 provides empirical evidence on the complexity-volatility link. Section 4 concludes.

2 Modeling Complexity, Volatility, and Comparative Advantage

We present below two theoretical mechanisms through which less developed countries come to exhibit comparative advantage in the less complex − and therefore more volatile − goods. The first, building on Blanchard and Kremer (1997) and Levchenko (2007), relies on differences in the quality of contract enforcement. The second approach, building on Costinot (2009), relies on the differences in human capital endowments across countries and the optimal division of labor. We now consider each of these mechanisms in turn.

2.1 Intermediate Inputs and the Contracting Environment

Consider an economy with a large number of industries, each characterized by the number of intermediates z required for production, $z \in (0, \overline{z}]$. For simplicity, we assume that the final output in industry z is produced with a Leontief production function

$$q_z = \min(q(1), .., q(s), .., q(z)), \tag{1}$$

where $q(s)$ is the quantity of intermediate good s that goes into production of the final good, $s = 1, ..., z$. There is one factor of production, L, and a large number of ex-ante identical potential intermediate goods producers. These hire labor to produce intermediates with a linear production function $q(s) = l$. Both the intermediate and the final goods sectors are competitive.

Final good producers in a sector requiring z inputs contract with z intermediate goods producers to deliver the inputs. The country's contracting environment is imperfect. In particular, after the intermediates have been contracted for by the final goods producer, each intermediates producer reneges on the project with some probability $1 - \rho$. When this happens, the entire project yields the output of zero. The value of ρ captures the level of institutional quality in the country. The higher it is, the more unlikely an intermediate input producer is to renege. As a result, for a given level of investment into each intermediate good production, the final output is given by

$$\min\left(l(1), ..., l(z)\right). \tag{2}$$

with probability ρ^z, and zero with probability $1 - \rho^z$.

Since all intermediate goods producers are ex ante identical and enter symmetrically into the production function, the final goods producer contracts for the same level of investment/employment for each intermediate. Therefore, the expected final output per worker is given by $\frac{\rho^z}{z}$ in industry z. Thus, output per worker depends on the complexity, z, of the good, and the quality of institutions ρ. Note that expected output per worker decreases in z at any level of institutional quality ρ. This is a result in the spirit of the O-ring theory of Kremer (1993) and the complementarity model of Jones (2008): the more steps the good requires, the higher is the chance that something will go wrong. If the inputs are "essential" as they are with a Leontief production function, the lower is the expected output. Note that in a model with multiple goods, this also implies that more complex goods command higher prices. While we are not aware of empirical studies of such a relationship in the data, the prediction appears prima facie sensible and intuitive.[2]

[2]It is well documented that more developed countries export goods with higher unit values. Since both

In this setting, less complex goods – that is, goods with lower z – are more volatile. The overall output per worker is $\frac{1}{z}$ with probability ρ^z and zero with probability $1-\rho^z$. Therefore, the variance of output per worker is given by $\frac{1}{z^2}\rho^z(1-\rho^z)$. We state the following Lemma, proved formally in Appendix B:

Lemma 1 ***(Complexity and Volatility in the Imperfect Contracting Model)*** *The variance of output per worker decreases in z:* $\frac{d}{dz}Var\left(\frac{\rho^z}{z}\right) < 0$.

This result is intuitive: as complexity of the good increases, the higher is the chance that output will be zero, and the lower will be output per worker if it is positive.[3] The two effects combine to deliver the negative relationship that we formalize in the Lemma.[4]

2.2 Imperfect Contracting and Comparative Advantage

Now suppose there are two countries, North and South. While we do not model contract enforcement explicitly, we assume that a better contracting environment in a country implies that the probability that someone reneges $(1-\rho)$ is lower there. Without loss of generality, let us assume that the North has a more efficient contracting environment. Thus, $\rho_N > \rho_S$.

We can map this setting into the Ricardian model of Dornbusch, Fischer and Samuelson (1977). Denote by $a^N(z) = \frac{z}{\rho_N^z}$ the unit labor requirement for good z in the North, and similarly in the South. Then, the relative unit labor requirement for good z in the two countries is:

$$A(z) = \frac{a^S(z)}{a^N(z)} = \left(\frac{\rho_N}{\rho_S}\right)^z. \tag{3}$$

in the theoretical model and in the data more developed countries export the more complex goods, in this respect the model appears consistent with the data.

[3]Note that the assumption of zero output in case of default is not important for the results. Alternatively, we could assume that even when the supplier reneges successfully, the final goods producer can force it to deliver a fraction δ of the contracted quantity of the intermediate good. In that case, the total output per worker is $\frac{\delta}{z}$ when no supplier defaults, and the variance of output is simply $\frac{1}{z^2}\rho^z(1-\rho^z)(1-\delta)^2$. It is clear that all the results carry over to this case.

[4]Notice that the only shock, and thus the only volatility in this model comes from the possibility that an input supplier reneges on the delivery of the good. This assumption is not crucial: the result above extends to the case in which there is both a reneging shock and a genuine productivity shock, as long as the two are uncorrelated for each supplier.

Does an efficient contracting environment create comparative advantage? That is, is the North relatively more productive in goods with higher z? The derivative of this ratio of productivities with respect to z is:

$$\left(\frac{\rho_N}{\rho_S}\right)^z \log\left(\frac{\rho_N}{\rho_S}\right) > 0. \tag{4}$$

We have just proved the following result:

Lemma 2 ***(Comparative Advantage in the Imperfect Contracting Model)*** *The North has comparative advantage in goods with higher z: $\frac{dA(z)}{dz} > 0$.*

The North is indeed relatively more productive in high-z goods. Though we imposed a Leontief production function at the outset, notice that this comparative advantage result does not depend on the functional form of the production function q_z.[5]

The preceding theoretical framework, while admittedly highly stylized, serves to illustrate two key ideas. First, output volatility is driven by product complexity. And second, better (worse) contracting environments can therefore generate comparative advantage in less (more) volatile goods. If we expect the efficiency of the contracting environment to improve with development, we obtain specialization in more volatile goods in less developed countries – consistent with the empirical findings of Koren and Tenreyro (2007) discussed above.

2.3 Human Capital and the Division of Labor

Our second modeling approach, based on Costinot (2009), relies on the differences in human capital endowments across countries and the optimal division of labor in the production of final goods. As in the previous section, more complex final goods require a larger number of intermediate inputs to be supplied or, as interpreted here, a larger number of different tasks to be performed. It is assumed that each of the tasks necessary for the production of the

[5]This is of course notwithstanding the issue that our modeling approach to the contracting frictions has a strong "Leontief" flavor.

final good requires some fixed labor costs to be incurred. As in Costinot (2009), dividing up the tasks among a larger number of workers generates gains from specialization – fewer tasks taken on by a single worker implies lower fixed costs incurred per worker, thereby raising output per worker. On the other hand, since workers are subject to random productivity shocks, complementarity in production implies that expected level of output is lower with a larger number of workers.[6] In the analysis that follows, we see how the trade-off between these two forces determines optimal team size used in production (and thus unit production costs) as a function of complexity. Then we will examine how, in this context, countries with high human capital workers have comparative advantage in the production of more complex goods.

Once again, consider an economy with many goods indexed by $z \in (0, \bar{z}]$. Each good is produced with a Leontief technology requiring z tasks to be performed. Let $s \in (0, z]$ denote a particular task that must be performed in order to produce good z, and let $q(s)$ be the quantity of task s. Then, the total output of good z, q_z, is given by

$$q_z = \min_{s \in (0,z)} q(s). \tag{5}$$

The economy is populated by L workers, each with productivity h. There are fixed costs associated with performing each task s. In particular, a worker must first spend 1 unit of labor learning to perform each task. Let N be the team size that characterizes production of a good with complexity z. The first question we ask is what is the team size that maximizes output per worker in sector z.

With a team of size N, each team member specializes in $\frac{z}{N}$ tasks, and allocates her endowment of labor equally to each of them. Therefore, after paying the fixed cost to learn these tasks, each worker has $h - \frac{z}{N}$ units of labor to spend on production. Hence, each worker is able to dedicate $\frac{h-\frac{z}{N}}{\frac{z}{N}}$ units of labor to each task.

[6]This point has been emphasized recently by Jones (2008).

After paying the fixed costs, each worker receives a productivity shock ε, affecting her performance of each task equally. Therefore, the worker's actual output is

$$q(s) = \left(\frac{hN}{z} - 1\right)\varepsilon \tag{6}$$

of each task s. Plugging equation (6) into (5), it is immediate that the total output of this team is $q_z = \left(\frac{hN}{z} - 1\right)\min_{n=1,\dots,N}\varepsilon_n$, while the output per worker is

$$\frac{q_z}{N} = \left(\frac{h}{z} - \frac{1}{N}\right)\min_{n=1,\dots,N}\varepsilon_n. \tag{7}$$

Note that, holding team size fixed, output per worker is higher, the greater the human capital level h and the lower the complexity of the good being produced z. This is as is expected. Furthermore, output per worker is a function of the random productivity shocks faced by workers. What is the team size that maximizes output per worker in this setting? Assume that the shocks are uncorrelated across workers. The expected output per worker is then equal to

$$E\left(\frac{q_z}{N}\right) = \left(\frac{h}{z} - \frac{1}{N}\right)E(\varepsilon_{(1)}), \tag{8}$$

where $\varepsilon_{(1)} \equiv \min_{n=1,\dots,N}\varepsilon_n$ is the first order statistic associated with the sample of N outcomes of a random variable ε_n across the workers in a team (see Appendix A).

For the sake of tractability, assume for now that the shocks to workers are distributed $\varepsilon \sim$Uniform$(0,1)$. This assumption has the advantage of leading to a simple closed-form solution for the optimal team size. In particular, the expected output per worker with team size N is equal to (see eq. A.4): $\left(\frac{h}{z} - \frac{1}{N}\right)\frac{1}{N+1}$. Following Costinot (2009), this expression can be used to find the optimal team size N_z in a sector with complexity z:

$$N_z = \text{argmax}_N \left(\frac{h}{z} - \frac{1}{N}\right)\frac{1}{N+1}. \tag{9}$$

The first-order condition is given by:

$$\frac{1}{N^2}\frac{1}{N+1} + \left(\frac{h}{z} - \frac{1}{N}\right)\left(\frac{1}{(N+1)^2}\right) = 0. \tag{10}$$

Straightforward manipulation gives the optimal team size in sector z of:

$$N_z = \frac{z}{h}\left(1 + \sqrt{1 + \frac{h}{z}}\right). \tag{11}$$

The optimal team size increases in the complexity of the good, z, and decreases in the worker productivity h. Optimal team size increases with complexity due to the gains from specialization that are obtained when the necessary tasks are divided up among a larger number of workers. The higher is the level of human capital, the costlier a low productivity draw becomes, and thus optimal team size falls in h. As we will see, the relationship we have established between optimal team size, complexity and human capital will be important in determining the pattern of comparative advantage.

Though the model of the division of labor and team size follows Costinot (2009), the key tension that pins down the optimal team size is different in our paper. In Costinot (2009), the tension is between greater division of labor and the resulting higher per worker productivity on the one hand, and imperfect contract enforcement: the more workers are in a team, the greater is the probability that at least one of them reneges. In our setup, the tension is between division of labor and the greater possibility of an adverse productivity shock that an individual worker may experience, in a production setting characterized by strong complementarities, a mechanism inspired by Jones (2008). Note also that though we choose to follow Costinot's terminology and call the team members "workers," the model will not change if we think of N as intermediate inputs suppliers instead.

The first result we would like to establish is that in this setting more complex goods are also less volatile. Going back to the expression for output per worker (7), it is immediate that the volatility of output per worker is given by:

$$Var\left(\frac{q_z}{N}\right) = \left(\frac{h}{z} - \frac{1}{N}\right)^2 Var(\varepsilon_{(1)}). \tag{12}$$

We state the counterpart of Lemma 1 for this model:

Lemma 3 ***(Complexity and Volatility in the Division of Labor Model)*** *The variance of output per worker decreases in z:* $\frac{d}{dz}Var\left(\frac{q_z}{N}\right) < 0$.

The proof is provided in Appendix B. We should note that the result in Lemma 3 depends in an important way on the property that the variance of the first order statistic (the minimum of a random sample) decreases in the sample size. Though this property appears intuitive, there are no finite sample general results in statistics about how the variance of the first order statistic behaves as the sample size increases. However, it can be confirmed using direct calculation that this variance indeed decreases in the sample size for some important distributions such as the uniform (as in this paper), exponential, Pareto, and Fréchet. This gives us some confidence that our main results are not excessively driven by the particular distributional assumptions that we adopt.

A related result is that *in each sector* z, a country with lower productivity of workers experiences lower volatility. This is because higher productivity implies lower team size, which in turn increases the volatility of output.

2.4 Human Capital Differences and Comparative Advantage

Suppose now that there are two countries, North and South. The only difference between them is that the North's workers are more productive: $h^N > h^S$. Following Costinot (2009), we map this model into the Ricardian framework of Dornbusch et al. (1977), by considering the unit labor requirements in each good z in the two countries. The average labor requirement of a unit of the good z in the North is:

$$a^N(z) = \frac{h^N}{\left(\frac{h^N}{z} - \frac{1}{N_z^N}\right)\frac{1}{N_z^N+1}} = \frac{zh^N N_z^N(N_z^N+1)}{(hN_z^N - z)}, \tag{13}$$

and similarly in the South. Therefore, the ratio of relative unit labor requirements is given by:

$$A(z) = \frac{a^S(z)}{a^N(z)} = \frac{h^S N_z^S(N_z^S+1)(hN_z^N - z)}{h^N N_z^N(N_z^N+1)(hN_z^S - z)}. \tag{14}$$

In order to establish the direction of comparative advantage, we must ascertain whether the schedule $A(z)$ is increasing or decreasing. Taking the derivative with respect to z, and applying the envelope theorem, we obtain:

$$A'(z) = \frac{\frac{\partial a^S}{\partial z} a^N - a^S \frac{\partial a^N}{\partial z}}{\left(a^N\right)^2}. \tag{15}$$

Evaluating the partial derivatives with respect to z based on equation (13) and simplifying, $A'(z)$ becomes:

$$A'(z) = \frac{h^S N^S (N^S + 1)}{(h^S N^S - z)^2 h^N N^N (N^N + 1)} \left(h^N N^N - h^S N^S\right). \tag{16}$$

Therefore, the sign of this derivative is the same as the sign of $\left(h^N N^N - h^S N^S\right)$. Using equation (11), it is immediate that $\left(h^N N^N - h^S N^S\right) > 0$, and therefore the North has a comparative advantage in the more complex goods, as expected. We summarize the discussion above in the following Lemma:

Lemma 4 ***(Comparative Advantage in the Division of Labor Model)*** *The North has comparative advantage in goods with higher z:* $\frac{dA(z)}{dz} > 0$.

The intuition for this result is straightforward: When workers have higher human capital, they spend a smaller fraction of their time learning, and so unit labor requirements are lower. Importantly, this reduction is not uniform across goods. In the more complex sectors, learning costs are more important and the decrease in unit labor requirements is larger. As a result, the country with workers with greater human capital is relatively more efficient in the more complex industries.

2.5 Trade Equilibrium

We now specified the pattern of comparative advantage $A(z)$ in two ways: by relying on contract enforcement (equation 3), and human capital differences (equation 14). In order to close the model, we must specify agents' preferences. Assume, following Dornbusch et al.

(1977), that all agents have identical Cobb-Douglas preferences, so that each good receives a constant share of expenditure. Let $\omega = \frac{w^N}{w^S}$ be the relative wage between the two countries. There exists a cutoff $\tilde{z}$, such that

$$\omega = A(\tilde{z}). \tag{17}$$

Let $S(\tilde{z})$ be the share of income spent on Southern goods. Then, the trade balance condition is given by

$$\omega = \frac{h^S L^S \left[1 - S(\tilde{z})\right]}{h^N L^N S(\tilde{z})}. \tag{18}$$

The equilibrium specialization pattern is illustrated in Figure 4. Equations (17) and (18) jointly determine the equilibrium pair $(\omega, \tilde{z})$. It is immediate that the South produces goods $(0, \tilde{z})$, while the North produces goods $(\tilde{z}, \bar{z})$. As such, the South ends up in the less complex industries in which production is the most volatile for each firm.

3 Empirical Evidence

There are two crucial pieces of evidence that we must bring to bear to support the theory proposed above. The first is that poorer countries do indeed specialize in less complex goods. This result has been established recently in a series of studies. Levchenko (2007) shows that countries with worse institutions – which are essentially the less developed countries – have relatively higher export shares in goods with low product complexity. In that study, measures of product complexity at sector level are constructed using the Input-Output tables for the United States, and by examining how many intermediates each sector requires to produce. Costinot (2009) provides similar results using an alternative measure, which is the average learning cost that a worker must pay in each sector before she becomes productive. Finally, Nunn (2007) constructs a measure of contract intensity by combining the U.S. Input-Output table data with a classification of intermediate goods industries into those that require relationship-specific investments and those that do not. Nunn finds that less developed countries specialize in industries that do not rely on relationship-specific investments, which could be another way of capturing industries with a low z in the model above.

The second crucial element is the negative relationship between complexity and volatility at sector level. On this score, we are not aware of any existing empirical evidence. In this section, we use data on the actual complexity and volatility of the U.S. manufacturing sectors to demonstrate that complexity is a robust and highly significant predictor of volatility.

3.1 Data

Industry-level data on volatility come from the NBER Productivity Database that reports information on 459 manufacturing sectors at the 4-digit SIC87 classification. We compute output per worker using data on total shipments and employment in each sector. Total output is deflated using sector-specific deflators provided in the database, ensuring that we capture the volatility of quantities. Because the level of real output per worker exhibits a trend, we compute the time series of the growth rate of sales per worker for each sector, and take the standard deviation over time for the period 1970-1997. Taking growth rates is the simplest way of detrending the data. To check robustness of the results, we also HP-filter the output per worker series in each sector, and compute the volatility of the deviations from the HP-filtered trend. Following the recommendation of Ravn and Uhlig (2002), we set the HP filter parameter to 6.25, since the data are at the annual frequency. Output per worker data may be contaminated by the time variation in the use of inputs or other factors of production. Thus, we compute the volatility of two alternative series: value added per worker, and Total Factor Productivity (TFP). The sector-specific TFP series is available in the same database. For both of these, we compute the standard deviation of the growth rate of the series, though the HP-filtering procedure delivers the same results.

Data on product complexity come from the U.S. Input-Output Tables for 1992, and have been previously used by Cowan and Neut (2007) and Levchenko (2007). In particular, in this exercise we use the total number of intermediates in production as a proxy for product complexity z in the model above. It turns out that the number of intermediates ranges from 16 to 160, a tenfold difference. Table 1 reports the summary statistics for both complexity and the actual volatility (standard deviation of output per worker growth) of the sectors in our data. Table 2 reports the top 10 most and least complex sectors, according to the total number of intermediates used.

Using the variation in actual product complexity in place of z in the model, we can compute the optimal team size N from equation (11), and as a result the volatility in each sector from equation (B.2). The resulting standard deviation of output as a function of product complexity z is depicted in Figure 5. Volatility is decreasing in complexity.[7]

Is the standard deviation of a sector as implied by its complexity a robust predictor of the actual volatility in that sector? Figure 6 presents the scatter plot of the standard deviation of output per worker growth against the implied volatility of output per worker constructed based on our model. There is a robust positive relationship between the two variables.

Table 3 presents the regression results. All throughout, we report the standardized beta coefficients, obtained by first demeaning all the variables and normalizing each to have a standard deviation of 1. Thus, the regression coefficients correspond to the number of standard deviations change in the left-hand side variable that would be due to a one standard deviation change in the corresponding independent variable. The four panels differ only in the measure of actual volatility used on the left-hand side. Panel A uses standard deviation of output per worker growth; Panel B, the volatility of deviations from HP trend; Panel C, volatility of value added per worker; Panel D, standard deviation of TFP growth. Column 1 reports the results of a bivariate regression of the actual on the implied volatility. The

[7] The relationship between complexity and volatility would be similar if we instead computed implied volatility using the imperfect contracting model of section 2.1. All of the results are virtually unchanged under this alternative approach, so we do not report them to avoid unnecessary repetition. To compute the variance, we choose the value of $h = 20$. We checked the robustness using all values of h between 1 and 200, and while h affects the level of the implied variance of output, the statistical significance of the results is unchanged.

positive relationship is very pronounced: with the exception of the deviations from the HP trend series, the t-statistics on the coefficient on the implied volatility are in the range of 5-7, and the R^2's of the bivariate regressions are as high as 0.18.

Column 2 controls for other sector characteristics, such as raw materials intensity, capital intensity, and skill intensity, constructed based on Romalis (2004). As we can see, after controlling for other sector characteristics, the coefficient of interest in Panel B goes from being insignificant to significant at the 1% level, while the rest of the results are virtually unchanged. Finally, column 3 removes the outliers in terms of actual volatility, and still finds that the relationship of interest is quite strong and statistically significant.[8] Finally, it may be that what we are picking up are differences in the elasticity of substitution across goods. For instance, Kraay and Ventura (2007) argue that developing countries are more volatile because they specialize in goods that have a higher elasticity of substitution. We use data from Broda and Weinstein (2006) to check whether sectoral volatility is systematically correlated with elasticity. Column 4 in each panel reports the results. Because the Broda-Weinstein data are in a different industrial classification, we lose 10 of the the sectors due to an imperfect concordance. Controlling for it the elasticity of substitution leaves the main results completely unchanged. The coefficient on the elasticity of substition is positive, and significant in two out of four specifications. Plausibly, sectors with higher elasticity of substitution are also more volatile.

Rather than use the data on the number of intermediate inputs to compute sectoral volatility as implied by the model, we can also assess whether actual volatility is positively correlated with measures of product complexity directly. Table 4 presents the results of estimating the relationship between actual volatility and various indicators of product complexity. Following Cowan and Neut (2007) and Levchenko (2007), we use a number of variables, all constructed using the 1992 Benchmark Input-Output Table for the United

[8]More precisely, we drop the top 5% most volatile sectors, according to each corresponding measure of volatility.

States. Column 1 regresses volatility on the number of intermediates used in production. Column 2 used the Herfindahl index of intermediate goods shares; Column 3, the Gini coefficient of intermediate use, columns 4 and 5 the shares of the 10 and 20 largest intermediate inputs in the total input use. Note that complexity increases in the number of intermediates, but decreases in all the other indicators. Thus, in columns 2 through 5 we should expect positive coefficients. We can see that with the exception of column 2, all the coefficients are significant at the 1% level. The coefficient on the Herfindahl index is not significant, but nonetheless enters with the expected sign.

We conclude that in a large sample of sectors, variation in complexity does play a significant role in explaining sectoral volatility, which is a key building block of our theory.

4 Conclusion

Recent literature has made important advances in understanding the patterns of macroeconomic volatility across countries. It is well known that poorer countries experience higher volatility. Koren and Tenreyro (2007) demonstrate that part of the higher volatility in developing countries can be accounted for by the fact that they produce on average in more volatile sectors.

How can we explain this puzzling observation? In this paper, we argue that international trade plays an important role. In particular, recent literature emphasized that poorer countries tend to export goods that are less complex (Levchenko 2007, Costinot 2009). Since these goods use fewer intermediates, shocks to each intermediate input are more important for production (a point also emphasized by Koren and Tenreyro 2008). Therefore, less complex goods tend to be more volatile. Comparative advantage in less complex goods, that could arise from institutional quality or productivity differences, drives specialization in more volatile industries by developing countries.

There is one aspect of our argument for which no empirical evidence currently exists.

Namely, it has not been demonstrated previously that less complex goods are indeed more volatile. In the last section of the paper we use data on the actual complexity of sectors in the United States to construct the volatility of each industry based on our model. We then relate this implied volatility to the actual volatilities of sectors observed in the data, and show that there is a robustly significant relationship: less complex industries are indeed more volatile.

Appendix A Order statistics

Suppose that $\varepsilon_1, ..., \varepsilon_N$ is a random sample of size N drawn from a distribution with pdf f_ε and cdf F_ε. The first order statistic is defined as $\varepsilon_{(1)} \equiv \min_{n=1,...,N} \varepsilon_n$, that is, it is the minimum value in this random sample. The distribution of $\varepsilon_{(1)}$ can be derived as follows. The cdf of this variable is given by:

$$F_{\varepsilon_{(1)}}(x) = P(\min_{n=1,\ldots,N} \varepsilon_n < x) = 1 - P(\min_{n=1,\ldots,N} \varepsilon_n > x) = 1 - (1 - F_\varepsilon(x))^N. \tag{A.1}$$

Correspondingly, the pdf of $\varepsilon_{(1)}$ is obtained by differentiating the cdf:

$$f_{\varepsilon_{(1)}}(x) = N(1 - F_\varepsilon(x))^{N-1} f_\varepsilon(x). \tag{A.2}$$

As an example, suppose that $\varepsilon \sim$Uniform$(0,1)$. The pdf of ε is $f_\varepsilon(x) = 1$, and the cdf is $F_\varepsilon(x) = x$. Then, the pdf of the first order statistic is

$$f_{\varepsilon_{(1)}}(x) = N(1 - x)^{N-1}. \tag{A.3}$$

Using integration by parts, it is straightforward to establish that in this case, the expectation and the variance of $\varepsilon_{(1)}$ are given by:

$$E(\varepsilon_{(1)}) = \frac{1}{N+1} \tag{A.4}$$

$$Var(\varepsilon_{(1)}) = \frac{N}{(N+1)^2(N+2)}. \tag{A.5}$$

These results will be useful in the main text.

Appendix B Proofs

B.1 Proof of Lemma 1

Proof: Taking this derivative directly,

$$\begin{aligned} \frac{d}{dz} Var\left(\frac{1}{z}\varepsilon_{(1)}\right) &= -\frac{2}{z^3}\rho^z (1-\rho^z) + \rho^z (1 - 2\rho^z) \ln \rho \frac{1}{z^2} \\ &= \frac{\rho^z}{z^2}\left[\ln \rho (1 - 2\rho^z) - \frac{2}{z}(1-\rho^z)\right]. \end{aligned}$$

Rearranging, the Proposition holds if and only if:

$$\frac{1-2\rho^z}{1-\rho^z}\ln\rho^z < 2.$$

When $\rho \in (0,1)$ and $z \geq 1$, it is always the case that $\rho^z \in (0,1)$. Therefore, the result nesessary for the proposition obtains if

$$\frac{1-2\rho}{1-\rho}\ln\rho < 2 \qquad \forall\rho \in (0,1)\,.$$

We now show that this is condition holds by proceeding in two steps. First, we show that the function $f(\rho) = \frac{1-2\rho}{1-\rho}\ln\rho$ is monotonically increasing throughout the interval $\rho \in (0,1)$. And second, we show that the supremum of this function, which obtains when $\rho \to 1$ is less than 2, satisfying this required condition.

Differentiating $f(\rho)$:

$$\begin{aligned}
\frac{d}{d\rho}\left[\frac{1-2\rho}{1-\rho}\ln\rho\right] &= \frac{1-2\rho}{1-\rho}\frac{1}{\rho} + \ln\rho\left[\frac{-2(1-\rho)-(-1)(1-\rho)}{(1-\rho)^2}\right] \\
&= \frac{1-2\rho}{1-\rho}\frac{1}{\rho} + \frac{\ln\rho}{(1-\rho)^2} \\
&= \frac{(1-2\rho)(1-\rho)-\rho\ln\rho}{(1-\rho)^2\rho} \\
&= \frac{(1-\rho)^2-\rho(1-\rho)-\rho\ln\rho}{(1-\rho)^2\rho} \\
&= \frac{1}{\rho} - \frac{(1-\rho)+\ln\rho}{(1-\rho)^2}.
\end{aligned}$$

Thus, this derivative is positive if $(1-\rho)+\ln\rho < 0$. It is immediate that $\lim_{\rho\to 0}(1-\rho+\ln\rho) = -\infty$ and $\lim_{\rho\to 1}(1-\rho+\ln\rho) = 0$. Therefore, if this function is monotonic for $\rho \in (0,1)$, it is everywhere less than 0, as required. Taking the derivative of this function,

$$\frac{d}{d\rho}\left[1-\rho+\ln\rho\right] = -1+\frac{1}{\rho} > 0 \qquad \forall\rho \in (0,1).$$

This establishes that $f(\rho) = \frac{1-2\rho}{1-\rho}\ln\rho$ is monotonically increasing in the interval $(0,1)$. We now show that its supremum is less than 2. The supremum obtains as $\rho \to 1$.

$$\begin{aligned}
\lim_{\rho\to 1}\left[\frac{1-2\rho}{1-\rho}\ln\rho\right] &= \lim_{\rho\to 1}(1-2\rho)\lim_{\rho\to 1}\left[\frac{\ln\rho}{1-\rho}\right] \\
&= (-1)\lim_{\rho\to 1}\frac{\ln\rho}{1-\rho} = -\lim_{\rho\to 1}\frac{\frac{1}{\rho}}{-1} = 1 < 2,
\end{aligned}$$

where the last equality comes from applying l'Hôpital's Rule. This completes the proof. ■

B.2 Proof of Lemma 3

Proof: Using the optimal value of N in equation (11), the term in parentheses simplifies to:

$$\left(\frac{h}{z} - \frac{1}{N}\right)^2 = \frac{1 + \frac{h}{z}}{N^2}. \tag{B.1}$$

Using equation (A.5) from the Appendix, the variance becomes:

$$Var\left(\frac{q_z}{N}\right) = \left(1 + \frac{h}{z}\right) \frac{1}{N(N+1)^2(N+2)}. \tag{B.2}$$

To establish that the variance decreases in good complexity z, evaluate its derivative with respect to z:

$$\frac{d}{dz} Var\left(\frac{q_z}{N}\right) = \frac{\partial}{\partial z} Var\left(\frac{q_z}{N}\right) + \frac{\partial}{\partial N} Var\left(\frac{q_z}{N}\right) \frac{dN}{dz}. \tag{B.3}$$

Evaluating each of these subcomponents separately, it is indeed the case that $\frac{d}{dz} Var\left(\frac{q_z}{N}\right) < 0$.

■

References

Becker, Gary and Kevin Murphy, "The Division of Labor, Coordination Costs, and Knowledge," *Quarterly Journal of Economics*, 1992, *107*, 1137–1160.

Blanchard, Olivier and Michael Kremer, "Disorganization," *Quarterly Journal of Economics*, 1997, *112*, 1091–1126.

Broda, Christian and David Weinstein, "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, May 2006, *121* (2), 541–85.

Costinot, Arnaud, "On the Origins of Comparative Advantage," *Journal of International Economics*, April 2009, *77* (2), 255–264.

Cowan, Kevin and Alejandro Neut, "Intermediate Goods, Institutions, and Output Per Worker," 2007. Central Bank of Chile Working Paper No. 420.

di Giovanni, Julian and Andrei A. Levchenko, "Trade Openness and Volatility," 2007. Forthcoming, *Review of Economics and Statistics*.

—— and —— , "The Risk Content of Exports: a Portfolio View of International Trade," 2008. Mimeo, IMF.

Dornbusch, Rudiger, Stanley Fischer, and Paul Samuelson, "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods," *American Economic Review*, 1977, *67*, 823–39.

Jones, Charles, "Intermediate Goods and Weak Links: A Theory of Economic Development," 2008. mimeo, UC Berkeley.

Koren, Miklós and Silvana Tenreyro, "Volatility and Development," *Quarterly Journal of Economics*, February 2007, *122* (1), 243–87.

—— and —— , "Technological Diversification," March 2008. mimeo, Central European University and London School of Economics.

Kraay, Aart and Jaume Ventura, "Comparative Advantage and the Cross-Section of Business Cycles," *Journal of European Economic Association*, December 2007, *5* (6), 1300–1333.

Krebs, Tom, Pravin Krishna, and William Maloney, "Trade Policy, Income Risk, and Welfare," 2008. Forthcoming, *Review of Economics and Statistics*.

Kremer, Michael, "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics*, August 1993, *108* (3), 551–75.

Levchenko, Andrei A., "Institutional Quality and International Trade," *Review of Economic Studies*, 2007, *74* (3), 791–819.

Lucas, Robert E., Jr., "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 1988, *22* (3-42).

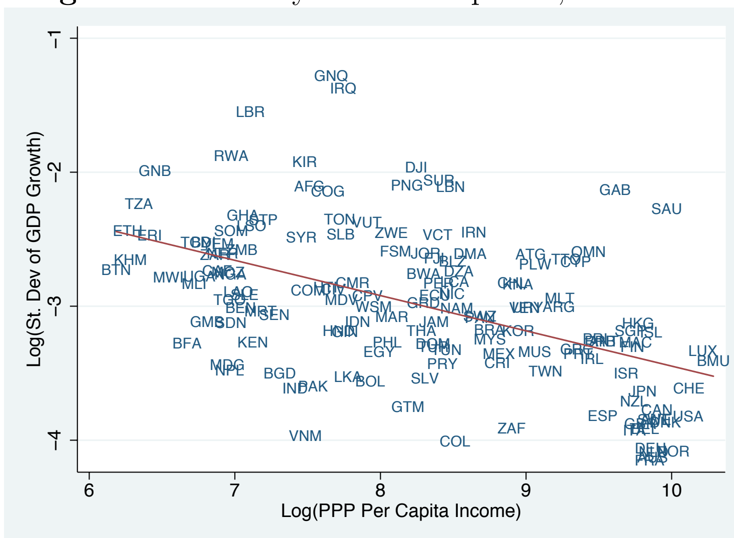
Nunn, Nathan, "Relationship-Specificity, Incomplete Contracts and the Pattern of Trade," *Quarterly Journal of Economics*, 2007, *122*, 569–600.

Ravn, Morten O. and Harald Uhlig, "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *Review of Economics and Statistics*, May 2002, *84* (2), 371–376.

Rodrik, Dani, "Why Do More Open Economies Have Bigger Governments?," *Journal of Political Economy*, October 1998, *106* (5), 997–1032.

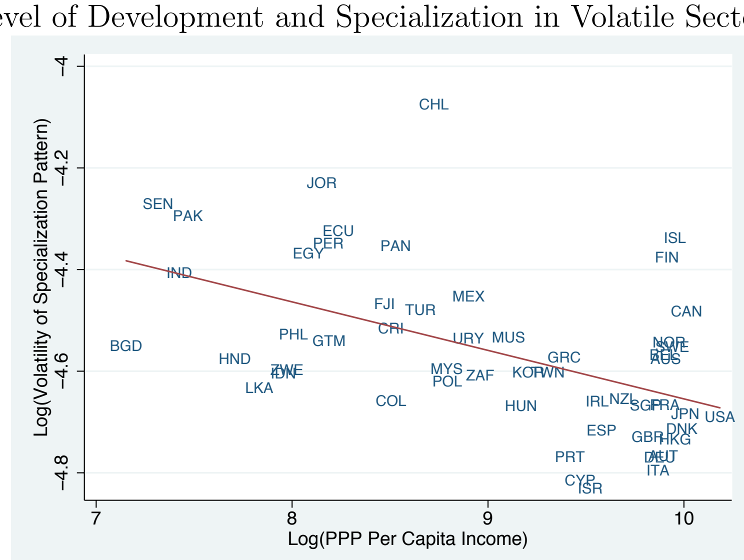
Romalis, John, "Factor Proportions and the Structure of Commodity Trade," *American Economic Review*, 2004, *94* (67-97).

Figure 1. Volatility and Development, 1970-2000



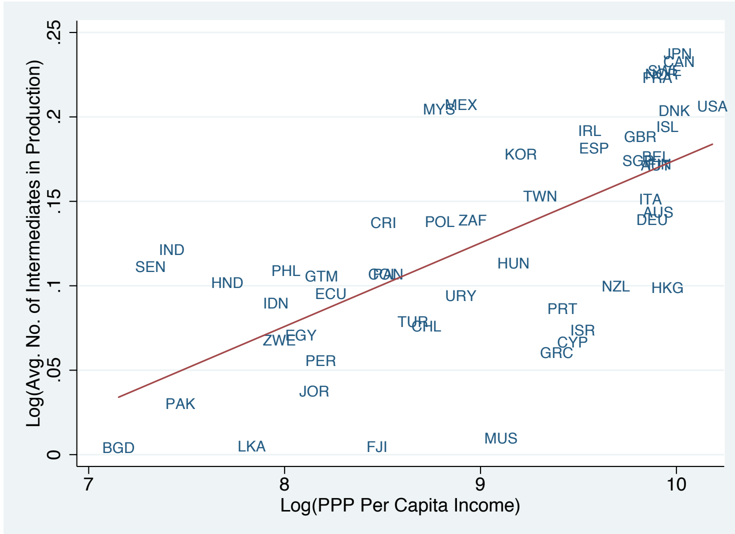
Notes: This figure displays the relationship between per capita income and the standard deviation of per capita GDP growth, in natural logs. Source: Penn World Tables.

Figure 2. Level of Development and Specialization in Volatile Sectors, 1970-2000



Notes: This figure displays the relationship between per capita income and the weighted average variance of the specialization pattern constructed following the methodology of Koren and Tenreyro (2007), in natural logs. Source: Penn World Tables and UNIDO.

Figure 3. Level of Development and Specialization in Complex Sectors, 1970-2000



Notes: This figure displays the relationship between per capita income and the weighted average number of intermediates used in production, with the weights equal to output shares. Source: Penn World Tables and UNIDO.

Figure 4. Pattern of Production and Trade

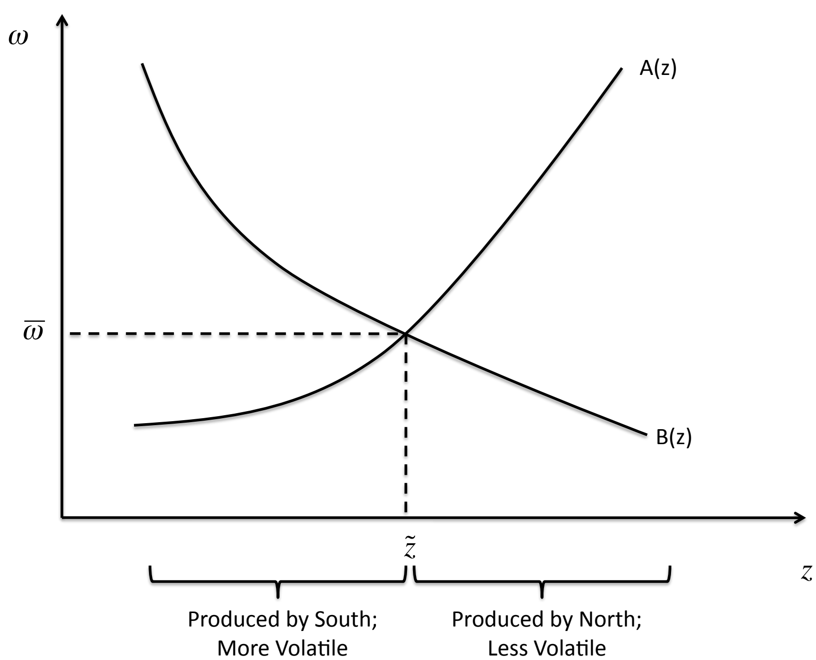
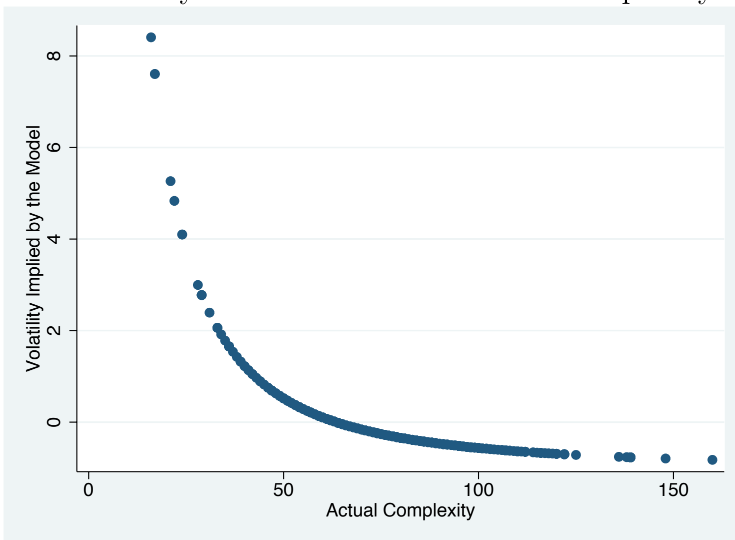
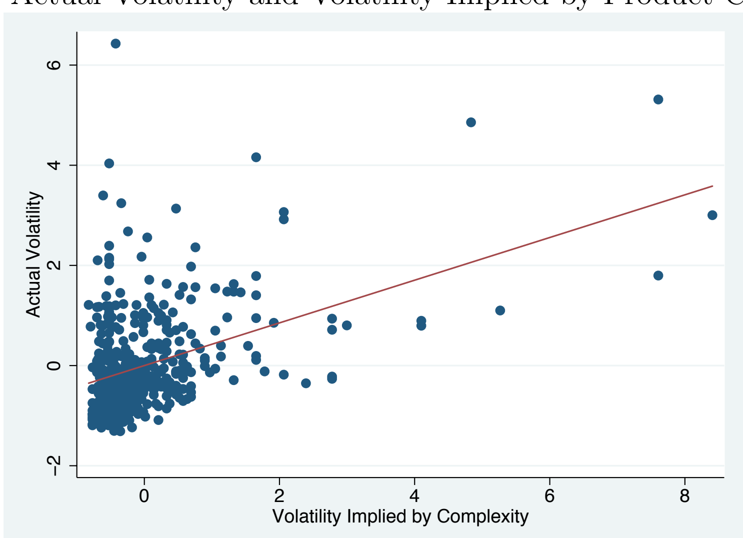


Figure 5. Volatility as a Function of Product Complexity: Model



Notes: This figure displays the relationship between the number of intermediate inputs in production (z) and volatility of output per worker, as implied by theory. *Actual Complexity* is the number of intermediate inputs used in a 4-digit SIC sector, calculated from the 1992 U.S. Input-Output Tables. It ranges from 1 to 160.

Figure 6. Actual Volatility and Volatility Implied by Product Complexity



Notes: *Actual Volatility* is the standard deviation of output per worker growth of a 4-digit SIC manufacturing sector in the United States over the period 1970-1997, sourced from the NBER Productivity database. *Volatility Implied by Complexity* is the standard deviation of output per worker implied by the theory, given the number of intermediate inputs used in that sector. The number of intermediates used in each 4-digit SIC sector is computed using the 1992 U.S. Input-Output Tables.

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min	Max	Obs.
Complexity (z)	77.0	25.2	16	160	459
Actual volatility	0.080	0.040	0.027	0.339	459

Notes: Complexity is the number of intermediates used in production, calculated based on the US I-O matrix. Actual volatility is the standard deviation of real output per worker growth, 1970-1997, calculated based on the NBER Productivity Database.

Table 2: Most and Least Complex Sectors

SIC Code	Sector Name	Number of Intermediates
	Least Complex Sectors	
2429	Special product sawmills, n.e.c.	16
3263	Semivitreous table and kitchenware	17
3151	Leather gloves and mittens	17
3131	Footwear cut stock	21
3292	Asbestos products	22
3142	House slippers	24
2397	Schiffli machine embroideries	24
3259	Structural clay products, n.e.c.	28
2441	Nailed wood boxes and shook	29
2121	Cigars	29
	Most Complex Sectors	
3728	Aircraft parts and equipment, n.e.c.	120
2865	Cyclic crudes and intermediates	122
281	Industrial Inorganic Chemicals	122
3585	Refrigeration and heating equipment	122
3731	Ship building and repairing	125
3812	Search and navigation equipment	136
3721	Aircraft	138
308	Miscellaneous Plastics Products	139
3714	Motor vehicle parts and accessories	148
3711	Motor vehicles and car bodies	160

Notes: Complexity is the number of intermediates used in production, calculated based on the US I-O matrix. Actual volatility is the standard deviation of real output per worker growth, 1970-1997, calculated based on the NBER Productivity Database.

Table 3: Actual and Implied Volatility

	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	Panel A: Dep. Var.: Standard Deviation of Output per Worker Growth				*Panel B: (B): Dep. Var.: Standard Deviation of Output per Worker Deviation from HP Trend*			
Implied Volatility	0.426***	0.447***	0.343***	0.445***	0.043	0.145***	0.092***	0.143***
	(0.064)	(0.071)	(0.044)	(0.071)	(0.040)	(0.039)	(0.024)	(0.039)
Raw Materials Intensity		0.261***	0.243***	0.219**		1.105***	0.631***	1.071***
		(0.100)	(0.067)	(0.102)		(0.114)	(0.052)	(0.118)
Capital Intensity		0.084	0.116**	0.069		0.661***	0.438***	0.650***
		(0.071)	(0.052)	(0.072)		(0.076)	(0.043)	(0.076)
Skill Intensity		0.118*	0.045	0.088		0.336***	0.139***	0.306***
		(0.069)	(0.051)	(0.069)		(0.074)	(0.028)	(0.079)
Elasticity				0.085*				0.096**
				(0.047)				(0.044)
Observations	459	459	436	449	459	459	436	449
R-squared	0.18	0.2	0.16	0.21	0.00	0.32	0.41	0.33
Sample	Full	Full	No Outliers	Full	Full	Full	No Outliers	Full
	Panel C: Dep. Var.: Standard Deviation of Value Added per Worker Growth				*Panel D: Dep. Var.: Standard Deviation of TFP Growth*			
Implied Volatility	0.300***	0.330***	0.234***	0.328***	0.426***	0.489***	0.316***	0.489***
	(0.059)	(0.072)	(0.029)	(0.072)	(0.084)	(0.092)	(0.049)	(0.093)
Raw Materials Intensity		0.642***	0.366***	0.622***		0.466***	0.197***	0.448***
		(0.111)	(0.051)	(0.119)		(0.103)	(0.065)	(0.106)
Capital Intensity		0.184***	0.138***	0.175***		0.303***	0.178***	0.302***
		(0.055)	(0.037)	(0.055)		(0.070)	(0.050)	(0.073)
Skill Intensity		0.181***	0.097**	0.165***		0.245***	0.091*	0.226***
		(0.053)	(0.041)	(0.055)		(0.072)	(0.051)	(0.066)
Elasticity				0.039				0.063
				(0.039)				(0.055)
Observations	459	459	436	449	459	459	436	449
R-squared	0.09	0.26	0.26	0.26	0.18	0.23	0.15	0.24
Sample	Full	Full	No Outliers	Full	Full	Full	No Outliers	Full

Notes: Standardized beta coefficients reported throughout. Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variables are standard deviations computed over the period 1970-1997. *Implied Volatility* is the standard deviation of a sector implied by its complexity as in equation (12), where complexity is measured as the number of intermediates used by a sector, from the US Input-Output matrix. *raw material intensity*=(value of raw material inputs)/(value of raw material inputs+value added); *capital intensity*=[1-(total compensation)/(value added)]*(1-raw material intensity); *skill intensity*=[(nonproduction workers)/(total employment)]*(1-capital intensity)*(1-raw material intensity), all computed based on the NBER Productivity Database. *Elasticity* is the elasticity of substitution between varieties in a given SIC sector (source: Broda and Weinstein, 2006).

Table 4: Alternative Measures of Complexity

	(1)	(2)	(3)	(4)	(5)
Dep. Var.: Standard Deviation of Output per Worker Growth					
Number of Intermediates	-0.300***				
	(0.056)				
Herfindahl Index		0.121			
		(0.087)			
Gini Coefficient			0.319***		
			(0.057)		
Share of 10 Largest Intermediates				0.253***	
				(0.056)	
Share of 20 Largest Intermediates					0.358***
					(0.065)
Raw Materials Intensity	0.161	-0.009	-0.04	-0.045	-0.039
	(0.116)	(0.141)	(0.125)	(0.128)	(0.123)
Capital Intensity	0.012	-0.083	-0.079	-0.089	-0.082
	(0.081)	(0.083)	(0.082)	(0.084)	(0.080)
Skill Intensity	0.073	-0.025	0.075	0.037	0.098
	(0.072)	(0.073)	(0.068)	(0.071)	(0.067)
Observations	459	459	459	459	459
R-squared	0.1	0.03	0.09	0.06	0.11

Notes: Standardized beta coefficients reported throughout. Robust standard errors in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. The dependent variable is the standard deviation of output per worker growth computed over the period 1970-1997. *Number of Intermediates* is the number of intermediates used in production; *Herfindahl Index* is the Herfindahl index of intermediate input use; *Gini Coefficient* is the Gini coefficient of the intermediate input use. *Share of 10 and 20 Largest Intermediates* are shares in of the top 10 or 20 intermediate inputs in the total intermediate input use. *raw material intensity*=(value of raw material inputs)/(value of raw material inputs+value added); *capital intensity*=[1-(total compensation)/(value added)]*(1-raw material intensity); *skill intensity*=[(nonproduction workers)/(total employment)]*(1-capital intensity)*(1-raw material intensity), all computed based on the NBER Productivity Database.