No. 2015-08A

OFFICE OF ECONOMICS RESEARCH NOTE
U.S. INTERNATIONAL TRADE COMMISSION

The Impact of Educated Labor on Technology Adoption and Comparative Advantage

Serge Shikher*

U.S. International Trade Commission

August 2015

*The author is with the Office of Economics of the U.S. International Trade Commission. Office of Economics research notes are the result of the ongoing professional research of USITC Staff and are solely meant to represent the opinions and professional research of individual authors. These notes are not meant to represent in any way the views of the U.S. International Trade Commission or any of its individual Commissioners. Research notes are circulated to promote the active exchange of ideas between USITC Staff and recognized experts outside the USITC, and to promote professional development of Office staff by encouraging outside professional critique of staff research.

Address correspondence to:
Office of Economics
U.S. International Trade Commission
Washington, DC 20436 USA

The impact of educated labor on technology adoption and comparative advantage

Serge Shikher[*,†]

Research Division
U.S. International Trade Commission

August 14, 2015

Abstract

Productivity differences across countries and industries play a major role in explaining international trade. But, what is behind these productivity differences? This paper finds that labor with post-secondary education and especially labor with an equivalent of an Associate's degree is the main determinant of productivity because it enables technology adoption. I use a variant of the principal component analysis called Singular Value Decomposition to break down productivity differences into country- and industry-specific components. This approach turns out to be successful empirically, in contrast to the previous evidence on the Heckscher-Ohlin model. I consider several candidates to match principal country- industry-specific components of productivity: physical capital, labor with various levels of education, and institutions. I find that labor with tertiary education is the best match. Analysis of the data on foreign technology licensing also leads to the conclusion that it is educated labor rather than institutions that is the main cause of productivity differences. Evidence on occupations and technology licensing shows that the main function of labor with tertiary education is to enable technology adoption. Based on this evidence, I develop a model in which comparative advantage is endogenous and technology adoption is driven by highly educated labor. I show that this model does a good job explaining comparative advantages.

JEL codes: F1, J24, I2, O4

Keywords: International trade, comparative advantage, specialization, education, human capital, development accounting

1 Introduction

Different countries export different sets of products. The pattern of trade is not random, however, and the focus of this paper is on understanding the pattern of exports across countries. For example, countries with low GDP per capita tend to export products in certain industries, such as textile,

[*] E-mail: serge.shikher@usitc.gov. The views expressed in this paper are strictly those of the author and do not represent the opinions of the U.S. International Trade Commission or of any of its Commissioners.

[†] I thank Darlene Chisholm, Kara Reynolds, seminar participants at the USITC, Washington Area International Trade Symposium (WAITS) at the George Washington University, and the Midwest International Economics Meetings at the Ohio State University for thoughtful comments that helped improve this paper.

basic metals, and food. Countries with high GDP per capita, on the other hand, tend to export certain types of machinery, such as medical equipment.

What explains the pattern of exports? Is it technology, abundance of some factors, or something else? The Ricardian model explains the pattern of trade by productivity differences, which determine comparative advantages. Empirical studies show that productivity differences are some of the most influential, maybe even the most influential, determinants of trade.[1] The goal of this paper is to go beyond the productivity explanation and look for causes of productivity differences across countries and industries.

The paper makes contributions to two literatures. It uses a novel empirical approach to estimate the effects of country and industry characteristics on productivity, thus contributing to the Heckscher-Ohlin empirical literature. This approach does not require taking a stand apriori on which factors to include in the analysis. The paper also contributes to the development accounting literature by adding the industry dimension. It shows that there is additional information contained in the industry dimension that can help us understand the sources of productivity differences across countries.

The paper finds that labor with tertiary education, especially labor with an equivalent of an Associate's degree, is the key determinant of productivity differences across industries and countries. The main contribution of this type of labor is to enable technology adoption.

The first step in the paper is to estimate productivities in each industry and country of the dataset. Comparative advantages are determined by productivities in autarky, which are normally different from the productivities observed with trade. The paper uses the Eaton-Kortum model to estimate productivities in autarky.

When calculating productivities, the paper takes care to account for key factors of production. In addition to physical capital, the paper accounts for three types of labor distinguished by education: labor with primary, secondary, and tertiary education. In order to account for the contributions of these types of labor to production, the paper uses data on wages and employments in a wide set of countries. The paper is the first to the author's knowledge to calculate shares for these types of labor using data for a wide set of countries, not just the U.S.

The paper also accounts for differences in education quality across countries. Recent literature presents evidence on education quality differences from international test scores and earnings of immigrants (Hanushek and Kimko, 2000; Hendricks, 2002; Schoellman, 2012). Consistent with that literature, this paper finds that ignoring education quality differences substantially overestimates productivity gaps across countries.

The next step in the paper is to study the pattern of estimated productivities across industries and countries. The key finding is that productivity gaps between countries are systematically different across industries. Countries that are further away from technological frontier have productivities that are lower in some industries than in others. Since distance to technological frontier has a high negative correlation with GDP per capita, another way of stating the key finding is that productivity gaps between rich and poor countries are systematically greater in some industries than in others.

The next challenge is to parsimoniously describe the observed pattern of productivities. For this purpose, I use a linear combination of country and industry characteristics. This approach can be thought of as a generalized Heckscher-Ohlin approach. However, since we do not know which factors may be affecting productivities, the first question is can a linear combination of any country

[1] See for example Trefler (1995), Harrigan (1997), Davis and Weinstein (2001), and Eaton and Kortum (2012).

and industry characteristics explain productivities. To answer this question, I use a statistical technique called singular value decomposition (SVD), a variant of principal component analysis. The answer to the above question is a resounding “yes”. A linear combination of just one country characteristic and one industry characteristic estimated by SVD can explain 92% of the variation of 742 productivities in 53 countries and 14 industries.

I then proceed to search for real-life counterparts of the country and industry characteristics estimated by SVD. I investigate various factor endowments and measures of institutions. I find that the endowment and intensity of labor with tertiary education have the highest correlations with country and industry characteristics, respectively, estimated by SVD. Since the market cost of labor with tertiary education has already been accounted for when calculating productivities, these results suggest that there is an externality to using highly educated labor. The externality is the ability to use sophisticated and productive technology.

This conclusion is supported by evidence from licensing of foreign technology. In the industries that extensively use highly educated labor, we observe high levels of technology adoption in the countries with high endowments of educated labor and very low levels of technology adoption in the countries with low endowments of this labor. By contrast, in the industries that use little educated labor the levels of foreign technology licensing are equal across all countries. The lack of highly educated labor is one of the main reasons for slow technology adoption (Nelson and Phelps, 1966; Benhabib and Spiegel, 2005).

Based on these results, I create a model of technology adoption in which a country’s ability to adopt technology in a particular industry is a function of that country’s endowment of educated labor and the industry’s need of that labor. The model is able to explain 50% of variation of productivities across both countries and industries.

This paper is related to the extensive literature that searches for the determinants of the pattern of trade and specialization. The Ricardian (1817) model tells us to look at comparative advantages driven by labor productivity differences. The Ricardian model does well empirically: early two-country studies of MacDougall (1951) showed good explanatory power of the Ricardian model and, more recently, the multi-country Ricardian model of Eaton and Kortum (2002) has been shown to fit data well. Eaton and Kortum (2012) provide a review of the recent literature.

However, there is something unsatisfying about the Ricardian model. Labor productivity differences determine the pattern of trade, but what determines the labor productivity differences? Heckscher (1919) and Ohlin (1924) created a model that explained labor productivity differences by differences of factor endowments across countries and differences of factor use across industries. However, studies done until now have shown that factor endowment differences can explain only a fraction of comparative advantages. Productivity differences are still needed to explain the rest (Trefler, 1995; Harrigan, 1997; Davis and Weinstein, 2001).

The search for an explanation of the pattern of trade largely parallels macroeconomics’ search for an explanation of per capita income differences across countries, also known as development accounting. In development accounting, large differences in total factor productivity across countries are needed to explain differences in per capita income (Hall and Jones, 1999; Caselli, 2005). These productivity differences are typically interpreted as differences in technology.

The empirical result that productivity differences play the greatest role in determining comparative advantage is akin to the result that total factor productivity (TFP) differences play the greatest role in explaining per capita income differences across countries. Ricardian productivity differences, just like TFP, are measured as residuals and, therefore, just like TFP, are “measures of our ignorance".[2]

Dissatisfaction with exogenous productivity differences as the explanation for the pattern of income across countries lead to the appearance of the endogenous growth literature. This literature aims to explain the differences in productivities across countries by accounting for additional factors, such as human capital (Mankiw, Romer and Weil, 1992), or by introducing mechanisms for technology production and transfer (Romer, 1990; Nelson and Phelps, 1966; Basu and Weil, 1998; Acemoglu and Zilibotti, 2001).

There is a literature that empirically investigates the effects of human capital on trade and specialization. Romalis (2004) finds that skill-abundant countries specialize in skill-intensive industries. Ciccone and Papaioannou (2009) find that countries with higher initial education levels experienced faster growth in schooling-intensive industries in the 1980s and 1990s. Other papers that studied the relationship between human capital and trade are Keesing (1966), Baldwin (1971), Baldwin (1979), and Harrigan (1997).

There are many papers that investigate the effects of human capital on output (Barro, 1991; Bils and Klenow, 2000; Barro and Lee, 2001; Erosa, Koreshkova and Restuccia, 2007; Manuelli and Seshadri, 2010; Schoellman, 2012).[3]

There is no widely agreed on measure of human capital. Generally, the existing literature has not focused on workers with tertiary education. Typical measures of human capital are average years of schooling, fraction of workers with secondary education, fraction of skilled workers, and fraction of non-production workers. These measures are much less correlated with productivity than the fraction of workers with tertiary education.[4]

This paper also finds that institutions play an important role in determining the pattern of productivities. Therefore, it confirms the findings of several previous papers that looked at the effects on institutions on trade (Nunn, 2007; Levchenko, 2007; Chor, 2010; Costinot, 2009). However, this paper presents evidence that education plays a greater role in determining productivity than institutions.

The results of this paper are useful for applied trade analysis. Understanding causes of productivity differences makes it possible to predict changes in comparative advantages, trade flows, and specialization that will occur in the future. It also improves accuracy with which trade economists can predict the effects of trade policy changes. Policy implications are discussed in the conclusion.

2 Estimation of productivity

I start by estimating country- and industry-specific levels of productivities, which will determine absolute and comparative advantages of countries. The production function in industry j of country

[2] The interpretation of TFP as a "measure of our ignorance" is due to Abramovitz (1956).

[3] Older literature finds the effect of education on output growth to be weak. Several reasons for this finding have been suggested: (a) attenuation due to mismeasured schooling data (Krueger and Lindahl, 2001) and (b) cross-country difference in educational quality (Hanushek and Kimko, 2000; Hendricks, 2002). Once education quality differences are accounted for, the effects of education on output increase significantly (Erosa et al., 2007; Manuelli and Seshadri, 2010; Schoellman, 2012).

[4] Nunn (2007) distinguished between high school graduates and those that did not graduate from high school at the country level and fraction of non-production workers in the U.S. at the industry level. Levchenko (2007) and Chor (2010) looked at a fraction of non-production workers in the U.S. data. Costinot (2009) used average educational attainment for country. At the industry level, he looked at "complexity" of an industry, which is approximately equal to the average educational attainment. He used it as a measure of institutional intensity.

i is

$$Q_{ij} = \tau_{ij} K_{ij}^{\alpha_j} \left(\prod\nolimits_e L_{eij}^{\lambda_{ej}}\right) M_{ij}^{1-\alpha_j-\beta_j} \tag{1}$$

where Q is output, τ is the total factor productivity (TFP), K is physical capital, α is the share of capital, L_{eij} is the quantity of labor with of type e employed in industry j of country i, λ_{ej} is the share of that type of labor in industry j, $\beta_j = \sum_e \lambda_{ej}$ is total labor share, and M is intermediate goods. There are several types of labor differentiated by years of education. These types of labor are imperfect substitutes in production. Productivity can only be measured in relative terms so I choose the U.S. as the base country.

One important issue with differentiating labor by level of education in the production function is that there is no readily available data on labor shares λ_{ej} by industry outside the U.S. This paper is the first to my knowledge to compile such data.[5] I collect the necessary data from a broad set of countries for three types of labor: labor with no more than primary education ($e = 1$), labor with more than primary, but less than tertiary education ($e = 2$), and labor with at least some tertiary education ($e = 3$). I also allow for differences in education quality across countries, as explained later in this section.

2.1 Productivity in autarky

From the production function (1), we derive (in logs) the TFP in industry j of country i relative to the U.S. Since reliable estimates of industry-level capital stocks do not exist for a broad set of countries, I will use the dual expression for TFP.[6] It is given by

$$\log \frac{\tau_{ij}}{\tau_{us,j}} = \alpha_j \log \frac{r_i}{r_{us}} + \sum\nolimits_e \lambda_{ej} \log \frac{w_{ei}}{w_{e,us}} + \left(1 - \alpha_j - \beta_j\right) \log \frac{\rho_{ij}}{\rho_{us,j}}, \tag{2}$$

where r is the cost of capital, w_e is the cost of labor or type e, and ρ is the cost of intermediate goods bundle.[7]

Each industry consists of many products. Trade literature tells us that different products within an industry are produced with different productivities. So we can think of industry productivity τ_{ij} as some average of the productivities of individual producers (and products) within an industry.

Dornbusch, Fischer and Samuelson (1977) describe how trade affects average productivity when there are productivity differences across products. Trade allows countries to produce only a fraction

[5] Previous studies used various measures of skill intensity. Some studies used skilled/unskilled classification of labor reported in some surveys. However, I find there is only weak correlation between skill and education. Skill is typically defined as knowledge of a particular complicated procedure, such as welding. Education, on the other hand, is a much more broad set of knowledge. Some studies measured skill intensity by the proportion of non-production workers in the total labor. I find this statistics also has only a very weak correlation to the level of education of the workforce. For example, in some industries production workers are required to have post-secondary education.

[6] While it is possible to derive industry-level capital stocks from investment time series, the investment data is not reliable for a broad set of countries (see for example Hsieh). In addition, capital services which are used in production can be systematically different from capital stocks because capital utilization rates can vary across countries and industries.

[7] The dual expression for τ is derived as follows $\frac{Q_{ij}}{Q_{us,j}} = \left(\frac{\alpha_j Q_{ij}}{\alpha_j Q_{us,j}}\right)^{\alpha} \left(\frac{\beta_j Q_{ij}}{\beta_j Q_{us,j}}\right)^{\beta} \left(\frac{\left(1-\alpha_j-\beta_j\right) Q_{ij}}{\left(1-\alpha_j-\beta_j\right) Q_{us,j}}\right)^{(1-\alpha-\beta)} = \left(\frac{r_i K_{ij}}{r_{us} K_{us,j}}\right)^{\alpha} \left(\frac{w_i L_{ij}}{w_{us} L_{us,j}}\right)^{\beta} \left(\frac{\rho_i M_{ij}}{\rho_{us} M_{us,j}}\right)^{(1-\alpha-\beta)}$. Taking logs and combining with the primary expression for τ, we obtain the dual expression for relative total factor productivity.

of all goods that they consume. When a country produces a small fraction of all possible products in an industry, it produces the products with the highest productivities. These products have the greatest chance of being successful on domestic and foreign markets. The other products wanted for consumption are imported. When a country produces a large fraction of all products, it produces the products that have high as well as those that have moderate productivities. In autarky, a country is forced to make products with even the lowest productivities. Therefore, trade has an effect on average total factor productivity in an industry.

Product level differences in productivities create product-level comparative advantages. Differences in industry total factor productivities (average across products) create industry-level comparative advantages. For example, one country may have a comparative advantage in making textile products while another country may have a comparative advantage in making electronic components. The goal of this paper is to explain these industry-level comparative advantages.

An industry-level comparative advantage is determined by the average productivity of all goods in an industry, not just those that are currently produced. Therefore, we need to know the average productivity of all goods that can be produced by industry j in country i. In other words, we need to know the average industry productivity in autarky. The total factor productivity (2), on the other hand, gives us the average productivity only of goods that are currently being produced.

I will denote the average productivity of all goods in an industry (even those that are not currently produced) by A_{ij}. I will call this measure "productivity" to distinguish it from total factor productivity τ. The difference between A_{ij} and τ_{ij} will be called R_{ij}. The productivity in industry j of country i relative to the U.S. will be

$$\log \frac{A_{ij}}{A_{us,j}} = \log \frac{\tau_{ij}}{\tau_{us,j}} + \log \frac{R_{ij}}{R_{us,j}}. \tag{3}$$

Parameter A_{ij} determines the set of technologies available for production in industry j of country i.

Plugging (2) into (3) we obtain

$$\log \frac{A_{ij}}{A_{us,j}} = \alpha_j \log \frac{r_i}{r_{us}} + \sum\nolimits_e \lambda_{ej} \log \frac{w_{ei}}{w_{e,us}} + \left(1 - \alpha_j - \beta_j\right) \log \frac{\rho_{ij}}{\rho_{us,j}} + \log \frac{R_{ij}}{R_{us,j}}. \tag{4}$$

The intuition for the relationship between τ_{ij}, R_{ij}, and A_{ij} is the following. Let's say we observe two countries X and Y with the same TFP and trade costs. We also observe that country X has higher exports, which implies that X is producing a greater fraction of products.[8] Since X produces a greater fraction of products, but has the same average TFP as Y, it must have higher A than Y. This is graphically illustrated in a figure in the appendix. Empirically, we will estimate X having greater R than Y, which would result in X having greater estimated A than Y.

2.2 Estimating $\log\left(R_{ij}/R_{us,j}\right)$

What is needed now is a way to estimate $\log\left(R_{ij}/R_{us,j}\right)$. Here I get help from the Eaton and Kortum (2002) model, which is the extension of Dornbusch et al. (1977) to many countries. In the Eaton-Kortum model, productivites of individual products are modeled as draws from the

[8] There is no intensive margin, so greater exports can only come from the extensive margin.

statistical distribution called Fréchet. Using the Eaton-Kortum model, I show that R_{ij} can be estimated using a gravity equation

$$\log \frac{X_{nij}}{X_{nnj}} = -\theta \log d_{nij} + \theta \log R_{ij} - \theta \log R_{nj}, \tag{5}$$

where X_{nij} represents industry j imports from i to n, X_{nnj} represents purchases of domestically-produced goods in n, d_{nij} is the "iceberg" trade cost of delivering goods from i to n in industry j ($d_{nij} \geqslant 1$), and θ is a parameter. Eaton and Kortum call R_{ij} the "competitiveness" of country i's industry j.

Equation (5) is equation (25) in Eaton and Kortum (2002) in logs with $R_{ij}^{\theta} \equiv T_{ij} w_i^{-\theta\beta_j} \rho_{ij}^{-\theta\left(1-\beta_j\right)}$, applied at the industry level. In the Eaton-Kortum model, the price of the intermediate goods bundle is the same as the price of the consumption good p_i. With many industries, the price of the intermediate goods bundle becomes industry-specific ρ_{ij}. Adding physical capital and differentiated labor as factors of production changes the cost of inputs from $w_i^{\beta_j} \rho_{ij}^{\left(1-\beta_j\right)}$ to $r_i^{\alpha_j} \left(\prod_e w_{ei}^{\lambda_{ej}}\right) \rho_{ij}^{\left(1-\alpha_j-\beta_j\right)}$.

In the Eaton-Kortum model, T_{ij} is the parameter of the distribution that describes the productivities of individual products. The mean productivity is $T_{ij}^{1/\theta}$, which I denote by A_{ij}. This is the mean productivity of all products in an industry. If $R_i^{\theta} \equiv r_i^{\alpha_j} \left(\prod_e w_{ei}^{\lambda_{ej}}\right) \rho_{ij}^{\left(1-\alpha_j-\beta_j\right)}$ and $A_{ij} = T_{ij}^{1/\theta}$ then $R_{ij} = A_{ij} / \left(r_i^{\alpha_j} \left(\prod_e w_{ei}^{\lambda_{ej}}\right) \rho_{ij}^{\left(1-\alpha_j-\beta_j\right)} \right)$. Taking logs and rearranging we obtain (4).

To estimate competitiveness measures R_{ij} using (5), I need a trade cost function in order to link d_{nij} with data. As in Eaton and Kortum, I will assume that trade cost d_{nij} is represented by the following trade cost function:

$$\log d_{nij} = d_{kj}^{phys} + b_j + l_j + f_j + m_{nj} + \delta_{nij} \tag{6}$$

where d_{kj} ($k = 1, ..., 6$) is the effect of distance lying in the kth interval, b_j is the effect of common border, l_j is the effect of common language, f_j is the effect of belonging to the same free trade area, m_{nj} is the overall destination effect, and δ_{nij} is the sum of geographic barriers that are due to all other factors.[9] As typical in trade literature, international trade cost is measured relative to domestic trade cost: $\log d_{iij} \equiv 0$.

Combining (5) and (6), I obtain the estimating equation for competitiveness measures R_{ij}:

$$\log \frac{X_{nij}}{X_{nnj}} = -\theta d_{kj}^{phys} - \theta b_j - \theta l_j - \theta f_j + D_{ij}^{exp} + D_{nj}^{imp} + \varepsilon_{nij}, \tag{7}$$

where $D_{ij}^{exp} = \theta \log R_{ij}$ is the exporter fixed effect and $D_{nj}^{imp} = -\theta m_{nj} - \theta \log R_{nj}$ is the importer fixed effect. The error term is $\varepsilon_{nij} = -\theta \delta_{nij}$. The left-hand side of (7) is obtained as follows: X_{nij} is from data and X_{nnj} is calculated as total output minus total exports of industry j in country n. The right-hand side of (7) consists of fixed effects. When estimating (7) the following normalization is used: $D_{us,j}^{exp} = D_{us,j}^{imp} = 0$. Consequently, the estimation produces relative competitiveness measures $R_{ij}/R_{us,j}$.

With estimates of competitiveness measures $R_{ij}/R_{us,j}$ in hand, I can calculate productivities $A_{ij}/A_{us,j}$ using (4). But first I need to obtain input costs and shares. Factor shares and prices are

[9] Note that unlike Eaton and Kortum (2002), trade cost here is industry-specific.

obtained from data. I assume that the intermediate goods bundle is a Cobb-Douglas aggregation of goods from all industries:

$$\rho_{ij} = \prod_{m=1}^{J} p_{im}^{\eta_{jm}}, \tag{8}$$

where p_{im} is the price index in industry m of country i and η_{jm} is the share of industry m in industry j intermediate goods bundle. The price of the intermediate goods bundle in each country and industry can be calculated using the Eaton-Kortum model following Shikher (2012):

$$\log \frac{\rho_{ij}}{\rho_{us,j}} = \frac{1}{\theta} \sum_{m=1}^{J-1} \eta_{jm} \left(\log \frac{X_{iim}/X_{im}}{X_{us,us,m}/X_{us,m}} - \theta \log \frac{R_{im}}{R_{us,m}} \right). \tag{9}$$

2.3 Accounting for differences in education quality

I allow for the possibility that education quality can differ across countries. Measuring relative costs of labor in (4) by $w_{ei}/w_{e,us}$ only makes sense if labor quality is the same across countries. There is a growing body of literature that shows that education quality varies across countries and the variation helps explain GDP per capita differences across countries. The evidence of education quality differences includes international test scores (Hanushek and Kimko, 2000; Kaarsen, 2014) and earnings of immigrants (Hendricks, 2002; Schoellman, 2012). Whether education quality differences can help explain measured cross-industry productivity differences is a question that has not been asked until now.

To account for cross-country differences in education quality I use methodology from Schoellman (2012) and educational quality measures from Kaarsen (2014). The wage of workers with education level e in country i is given by $w_{ei} = \overline{w}_i h_{ei}$, where $\overline{w}_i$ is the base wage in country i and h_{ei} is the human capital of labor with education level e in country i. Human capital is a function of years and quality of education: $h_{ei} = e^{\phi(s_e,q_i)}$, where s_e is the number of year of education of level e and q_i is the quality of education in country i. Function ϕ is given by $\phi(s_e, q_i) = \frac{\theta}{\eta}(s_e q_i)^{\eta}$, where θ and η are parameters. With this specification, labor is differentiated by the level of schooling and also the quality of this schooling. So workers with secondary education in the United States and Brazil may have different levels of human capital for use in production.

The wages observed in data are w_{ei}^{i}, where the superscript represents the quality of education that workers received. In order to compare wages across countries, the wages need to be for workers with the same level and quality of education. So while we observe $w_{ei}^{i}/w_{e,us}^{us}$, we need $w_{ei}^{us}/w_{e,us}^{us}$. In the second ratio both wages are for workers with the U.S. quality of education.[10] We can derive the second ratio from the first using methodology of Schoellman (2012) and Kaarsen (2014) and educational quality estimates from Kaarsen (2014):

$$\begin{aligned}\log \frac{w_{ei}^{us}}{w_{e,us}^{us}} &= \log \frac{w_{ei}^{i}}{w_{e,us}^{us}} + \log \frac{w_{ei}^{us}}{w_{ei}^{i}} = \log \frac{w_{ei}^{i}}{w_{e,us}^{us}} + \log \frac{\overline{w}_i h_{e,us}}{\overline{w}_i h_{ei}} = \log \frac{w_{ei}^{i}}{w_{e,us}^{us}} + \phi(s_e, q_{us}) - \phi(s_e, q_i) = \\ &= \log \frac{w_{ei}^{i}}{w_{e,us}^{us}} + \frac{\theta}{\eta}\left[(s_e q_{us})^{\eta} - (s_e q_i)^{\eta}\right] = \log \frac{w_{ei}^{i}}{w_{e,us}^{us}} + \frac{\theta s_e^{\eta}}{\eta}\left[(q_{us})^{\eta} - (q_i)^{\eta}\right] \end{aligned} \tag{10}$$

[10] Wages w_{ei}^{us} are the wages that a U.S.-educated worker from i would earn in i. There is no data on such wages, other than a few anecdotal observations, therefore they need to be estimated.

Combining the above expression with (4) the expression for relative productivity becomes

$$
\begin{aligned}
\log \frac{A_{ij}}{A_{us,j}} \quad = \quad & \log \frac{R_{ij}}{R_{us,j}} + \alpha_j \log \frac{r_i}{r_{us}} + \lambda_{1j} \log \frac{w_{1i}^{us}}{w_{1,us}^{us}} + \lambda_{2j} \log \frac{w_{2i}^{us}}{w_{2,us}^{us}} + \lambda_{3j} \log \frac{w_{3i}^{us}}{w_{3,us}^{us}} + \\
& + \left(1 - \alpha_j - \beta_j\right) \log \frac{\rho_{ij}}{\rho_{us,j}} + \frac{\theta}{\eta} \left((q_{us})^{\eta} - (q_i)^{\eta}\right) \left(\lambda_{1j} s_1^{\eta} + \lambda_{2j} s_2^{\eta} + \lambda_{3j} s_3^{\eta}\right) \qquad (11)
\end{aligned}
$$

Schoellman (2012) and Kaarsen (2014) provide a range of estimates of parameters θ and η. I use the following numbers of years of education: $s_1 = 3$ for primary, $s_2 = 9$ for secondary, and $s_3 = 15$ for tertiary.

The education quality in most countries is lower than that of the U.S., so their quality-adjusted wages are higher than non-quality-adjusted wages. Adjusting wages for quality in most countries leads to lower measured productivity differences with the United States in all industries. In other words, it helps explain productivity differences across countries. However, there are not enough differences in quality adjustment terms across industries to help explain the pattern of productivity differences across industries, i.e. the pattern of comparative advantages.

To summarize, the procedure for obtaining relative productivities $A_{ij}/A_{us,j}$ is to first estimate (7) in order to obtain the relative competitiveness measures $R_{ij}/R_{us,j}$. The second step is to calculate $A_{ij}/A_{us,j}$ using (11).

3 Data

I estimate country- and industry-specific productivities A_{ij} for 15 manufacturing industries in 53 countries in 2005. These productivities will inform us about the comparative advantages of countries. The countries include both rich and poor ones. For example, there are 30 countries with per capita GDP less than 20% of the U.S. and 10 countries with GDP per capita less than 5% of the U.S.

The bilateral trade data needed to estimate (7) was obtained from COMTRADE and concorded to 15 2-digit ISIC.[11] The data is for 53 countries in 2005. Imports from home X_{nnj} are calculated as output minus exports. Output data is originally from INDSTAT2-2010. The data on physical distance, common border, common language, and free-trade agreements is originally from the Gravity Database by CEPII. As in Eaton and Kortum, physical distance is divided into 6 intervals: [0,375), [375,750), [750,1500), [1500,3000), [3000,6000), and [6000,maximum).

Capital shares α_j, labor shares β_j, and intermediate inputs shares η_{jm} are calculated as the average shares of 43 countries in the input-output tables collected by the OECD.[12] Rates of return to physical capital are calculated in two different ways using two different assumptions. Under the first assumption, rates of return are assumed to be equal in all countries (meaning that capital is assumed to be internationally mobile, subject to transport costs, and economy is in a long-run equilibrium). Under the second assumption, rates of return are given by $r = \alpha Y/K$, where α is the capital share in the economy, equal to 0.3, Y is GDP, and K is capital stock, obtained from Penn World Tables 8 (Feenstra, Inklaar and Timmer, 2013). The choice of the rate of return measure

[11] The results of estimating (7) were originally presented in Yaylaci and Shikher (2014), which takes data from Yaylaci (2013).

[12] In the data, in addition to intermediate and final goods, there are also investment goods. Since there is no investment in the model, investment goods are treated as intermediate goods.

has little effect on the results and conclusions of this paper. Results presented in the rest of the paper are obtained using the first assumption.

3.1 Data on the earnings of three types of labor

In order to operationalize (11) I need to know the earnings by country and labor type, w_{ei}, and income shares by labor type in every industry, λ_{ej}. The earnings are obtained from data. The income shares are calculated from earnings w_{ei} and data on employment by labor type, industry, and country, L_{eij}. I use multiple data sources for earnings and employment that sometimes supplement each other and sometimes serve to cross-verify each other. What follows is fairly brief exposition of data sources. A much more detailed review is presented in the Data Appendix, available upon request.

There are three data sources for earnings. The first is the Freeman-Oostendorp's Occupational Wages Around the World (OWW) database, which takes its data from the ILO's October Inquiry. It has data for 1983-2008 and 44 countries out of 53 countries in my dataset. For each country, it reports earnings for up to 161 occupations. Each occupation (coded according to the ISCO-88 standard) is related to an industry (ISIC) and level of education (ISCED). For example, occupation number 52 in OWW is a Chemical Engineer employed in the Manufacture of Industrial Chemicals industry who has tertiary education.

To obtain average earnings for a given level of education in a country, I take an average of earnings of all occupations with that level of education in the country. While OWW has many occupations, it does not cover all occupations and does not represent a random sample. Therefore, to check how accurate the average earnings produced by OWW are, I use data from Eurostat's Structure of Earnings Survey (SES). It has 2006 data for 22 out of 53 countries in my dataset. For 15 of those countries, there is earnings data in both OWW and SES datasets. The earnings for each level of education and country are similar in the two datasets with correlation being 0.92.

Two countries in my dataset have no earnings data in either OWW or SES dataset. In addition, data for five countries in OWW is suspect or missing. For these seven countries, I obtain earnings data from country-specific studies. There is a large literature that uses microdata to estimate returns to education, also known as Mincerian returns. These returns are slopes from the regression of the log of earnings on the number of years of education (Mincer, 1974). In addition to the seven countries already mentioned, I calculate earnings by education from Mincerian returns for two more (randomly chosen) countries to see how similar the calculated earnings are to those in OWW. Altogether, I have six countries for which I calculated earnings from Mincerian returns and have earnings data from OWW. The correlation between the earnings obtained from the two sources is 0.9.

Combining all the sources of earnings information, I obtain earnings for each of the 53 countries in my datasets and 3 levels of education. As expected, earnings vary significantly across countries. The cross-country variation in hourly earnings is highly correlated with GDP per capita. Within each country, earnings increase with education ("education premia").[13] The cross-country average premium for having secondary education is 34%. This number is not adjusted for differences in education quality. The average earnings premium of workers with tertiary education over those with secondary education is 84%. The average earnings premium of workers with tertiary education

[13] The only exception is Ukraine where the average worker with secondary education earns a little less than the average worker with primary or less education.

over those with primary or no education is 149%. Therefore, having additional education, especially college education, significantly improves ones's standard of living.

In order to adjust wages for differences in education quality, as explained in Section 2.3, I use two sources of estimates of education quality: Schoellman (2012) and Kaarsen (2014). Schoellman (2012) estimates quality of education from earning of immigrants while Kaarsen (2014) estimates them from international science and math test results. The choice of the source for the estimate of education quality makes only small difference in the results presented in this paper and does not affect the conclusions. When estimating productivities, I only present the results obtained using Kaarsen's estimates of q. When looking at the effects of country-specific determinants on productivities in Section 5.1 I show results obtained using both Schoellman's and Kaarsen's estimates. I use each author's corresponding estimates of parameters η and θ. Schoellman estimates $\eta = 0.5$ while setting $\theta = 1$. Kaarsen estimates $\eta = 0.35$ and $\theta = 0.46$.

3.2 Data on the employment and shares of three types of labor

The main source of data for the employment by country, industry, and level of education, L_{eij}, is the World Bank Enterprise Surveys (WBES). The surveys were conducted during 2002-05 and have data on 6,000 enterprises from 21 countries out of 53 studied in this paper.[14] Half of these 21 countries are low and low-middle income countries. In addition, World Management Survey (WMS) data is used to check the WBES data. WMS was conducted during 2004-2010 and has data on 10,000 enterprises in 20 countries. It only collected employment data on workers with tertiary education, which can be compared to the data from WBES. The correlation is 0.89.

Using data on earnings w_{ei} and employment L_{eij} we calculate shares of each type of labor in total labor income, $w_{ei}L_{eij}/(w_{1i}L_{1ij} + w_{2i}L_{2ij} + w_{3i}L_{3ij})$. The average of these shares across countries is equal to λ_{ej}/β_j from which we can back out λ_{ej} using data on β_j (described previously).

Table 1 shows factor shares in output. We see that Nonmetals, Chemicals, and Paper industries are the most capital intensive while Textile, Other Machinery, and Transport industries are the least capital intensive industries.[15] The share of capital in the most capital intensive industry, Nonmetals, is 1.84 times higher than the share in the least capital intensive industry, Transport.

Looking at the total shares of labor, we see that Medical, Metal Product, and Textile industries are the most labor intensive while Metals, Food, and Petroleum Products are the least labor intensive. The share of labor in the most labor intensive industry, Medical, is nearly five times higher than in the least labor intensive industry, Petroleum products. It is 1.93 times higher than in the second least labor intensive industry, Food.[16]

We can also look at the shares of each type of labor. It is interesting, for example, to compare Textile and Medical industries. Both are very labor intensive. However, they use different types of labor. The share of labor with primary or less education (L_1) is 1.65 times higher in Textile industry. At the same time, the share of labor with some tertiary education (L_3) is 2.45 times higher in Medical industry. In addition to Medical, Other Machinery and Paper industries use highly educated labor intensively. Textile, Wood, and Nonmetals industries use least educated labor intensively.

[14] This means that the shares are calculated for 21 countries, not for all 53 countries. The model assumes that the shares are the same in all countries and the average shares across the 21 countries are used the analysis.

[15] The Paper industry is dominated by the Printing and Publishing (sub)industry (ISIC 22). Other Machinery industry includes office and computing machinery industries (ISIC 29 and 30).

[16] Petroleum Products industry is often an outlyer and is omitted from much of the analyis done in this paper.

Table 1: Factor shares in output

Code	Industry	Capital	Lab-Tot	Lab-Pri	Lab-Sec	Lab-Ter
1	Food	0.123	0.127	0.015	0.076	0.036
2	Textile	0.110	0.211	0.022	0.148	0.042
3	Wood	0.136	0.184	0.021	0.123	0.040
4	Paper	0.156	0.195	0.010	0.115	0.070
5	Petroleum products	0.114	0.052	0.000	0.025	0.026
6	Chemicals	0.162	0.139	0.005	0.074	0.059
7	Rubber	0.126	0.193	0.013	0.121	0.059
8	Nonmetals	0.173	0.203	0.026	0.128	0.049
9	Metals	0.115	0.133	0.014	0.087	0.032
10	Metal products	0.130	0.226	0.019	0.140	0.068
11	Machinery, other	0.108	0.207	0.012	0.117	0.079
12	Machinery, e&c	0.118	0.182	0.011	0.105	0.066
13	Medical	0.150	0.246	0.013	0.129	0.104
14	Transport	0.094	0.172	0.007	0.113	0.053
15	Other	0.142	0.210	0.018	0.135	0.057

Note: Share of capital is α_j, share of labor is β_j, share of labor with level of education e is λ_{ej}

3.3 Data for sixteen types of labor from the United States

I consider the shares obtained from the international data (described in the previous section) to be my primary source of information on the use of different types of labor, but I also use data from the United States to have a more detailed information on the use of different types of labor in manufacturing. The U.S. data provides information on 16 types of labor differentiated by level of education. The source of the U.S. data is the American Community Survey (ACS). I use microdata from that survey, which provides detailed information on about 3 million people. In addition to educational attainment, this survey collects data on employment status, industry of employment by 3 to 5 digit NAICS classification (which I concord to my 15 industries), salary/wages, and occupation by SOC code (465 non-military occupations).

For each industry, I calculate shares of each type of labor in total labor earnings λ_{ej}/β_j. If I aggregate these shares into the three types of labor described in the previous section, then the shares from the ACS are correlated (across industries) with the shares from international data. The correlation is 0.78 for the first type of labor, 0.75 for the second, and 0.94 for the third.

4 What do estimated relative productivities tell us?

I estimate relative productivities $\log(A_{ij}/A_{us,j})$ using equations (7), (11), and data on 15 industries in 53 countries. The estimates of productivities for select countries are shown in Table 2. The estimates for all industries and countries are presented in the appendix.

Looking at the industry-level productivities $\log(A_{ij}/A_{us,j})$ we can make several observations. First is that some countries have higher productivities than others in all industries. For example,

Table 2: Productivities in select countries

	China	Ethiopia	Germany	Korea	Mexico	Turkey	Vietnam
Food	0.66	0.48	0.88	0.60	0.57	0.67	0.56
Textile	0.80	0.42	0.96	0.91	0.59	0.78	0.54
Wood	0.74	0.33	0.98	0.60	0.47	0.53	0.46
Paper	0.59	0.25	0.95	0.73	0.49	0.49	0.33
Chemicals	0.66	0.39	0.91	0.72	0.62	0.59	0.37
Rubber	0.60	0.27	0.94	0.93	0.52	0.60	0.40
Nonmetals	0.71	0.30	1.01	0.77	0.53	0.66	0.40
Metals	0.77	0.46	0.98	0.90	0.62	0.69	0.42
Metal products	0.62	0.25	0.98	0.76	0.53	0.59	0.35
Machinery, other	0.58	0.21	0.97	0.75	0.54	0.54	0.31
Machinery, e&c	0.67	0.22	0.97	0.89	0.59	0.60	0.38
Medical	0.50	0.20	0.95	0.66	0.47	0.41	0.24
Transport	0.58	0.28	0.98	0.89	0.55	0.63	0.40
Other	0.67	0.26	0.91	0.76	0.54	0.59	0.41
AVERAGE	0.65	0.31	0.95	0.78	0.54	0.60	0.40

productivity in Germany is higher than productivity in Ethiopia in all industries. The cross-industry average productivity in Germany is about 3 times higher than the average in Ethiopia.

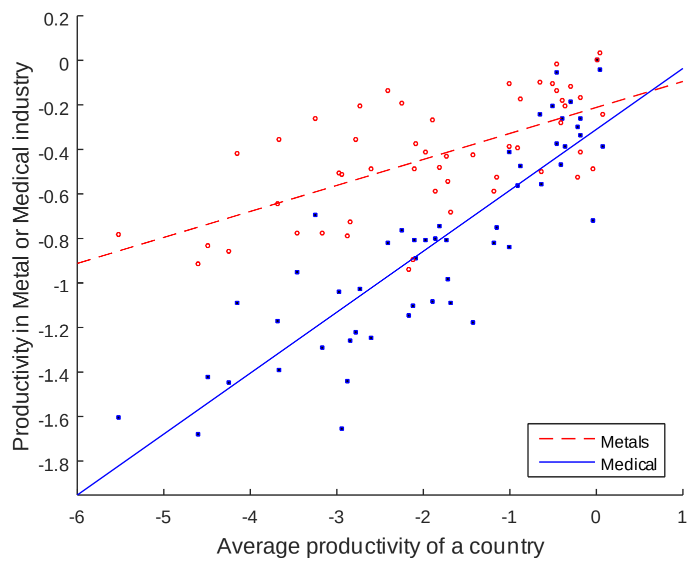
The second observation is that in each country relative productivities vary significantly across industries. For example in Vietnam the Food industry is 56% as productive as the Food industry in the U.S. while the Metal Products industry is only 35% as productive as the Metal Products industry in the U.S. This within-country cross-industry variation represents industry-level comparative advantages enjoyed by each country.

The third observation, which is key to this paper, is that as the overall productivity of a country declines the productivities of individual industries decline at different rates. The productivity declines quickly in some industries and slowly in others. Figure 1 illustrates this phenomenon by plotting productivities in two industries, Metals and Medical, against a simple average productivity for each country, $\overline{A}_i = (1/J) \sum_j \log(A_{ij}/A_{us,j})$. As average productivity declines, the relative productivity falls much faster in Medical than in Metals industry. Productivity differences between these two industries are small in rich countries, but become obvious in middle-income countries. They are very large in poor countries. Clearly, the productivity-driven comparative advantage of poor countries lies much more in Metals industry than in Medical.[17] This implies that there is a pattern of productivity differences across countries.

This pattern is more clearly seen on Figure 2 which shows the productivities in all industries of all countries. The countries are sorted by the average (across industries) country productivity while the industries are sorted by the average (across countries) industry productivity. We will now

[17] We can quantify how fast the technological gap grows as GDP per capita declines by the slope of the regression $\log(A_{ij}/A_{us,j}) = \mu_{0j} + \mu_{1j}\log(Y_i/Y_{us}) + \varepsilon_{ij}$, where Y_i is the GDP per capita of country i. This slope is the elasticity of relative productivity with respect to GDP per capita. Table XX shows industries in the dataset ranked according to the estimated elasticity μ_{1j}. Food and Metals industries have the lower estimated elasticities while Metal Products and Medical have the highest. The regression R^2 increases together with the slope (elasticity).

Figure 1: Productivity in two industries, in logs



proceed to characterize the pattern of productivities more formally.

Let's see if we can decompose productivity differences into industry-specific components γ_j^k, $k = 1, ..., M$ and country-specific components ϕ_i^k, $k = 1, ..., M$, where M is the number of components. This approach is in spirit of the Heckscher-Ohlin model in which labor productivity in an industry is a function of an industry-specific capital intensity and country-specific capital endowment.[18]

I will assume a specific functional relationship between productivity differences, industry, and country components:

$$\frac{A_{ij}}{A_{us,j}} = f\left(\frac{\phi_i^1}{\phi_{US}^1}, ..., \frac{\phi_i^M}{\phi_{US}^M}, \gamma_j^1, ..., \gamma_j^M\right) = -\prod\nolimits_{k=1}^{M} \left(\frac{\phi_i^k}{\phi_{US}^k}\right)^{\gamma_j^k} \tag{12}$$

or in logs

$$\log \frac{A_{ij}}{A_{us,j}} = -\sum_{k=1}^{M} \gamma_j^k \log \frac{\phi_i^k}{\phi_{US}^k} \tag{13}$$

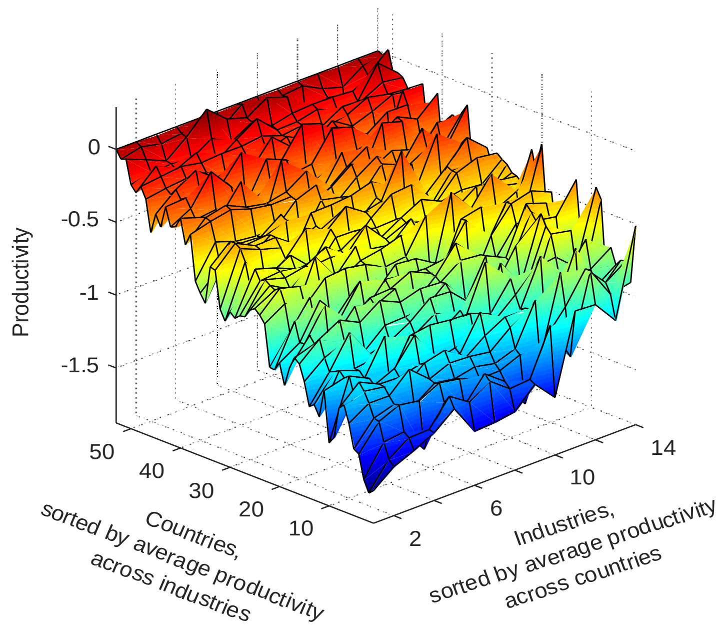
or in matrix form

$$\underset{N\times J}{\mathbf{A}} = \underset{N\times M}{\mathbf{U}} \cdot \underset{J\times M}{\mathbf{V}}^{\mathrm{T}}, \tag{14}$$

where each row of $\mathbf{U}$ contains the prices of M factors in country i, and each row of $\mathbf{V}$ contains shares of M factors in industry j. We can use a statistical technique called Singular Value Decomposition (SVD) to decompose $\mathbf{A}$ into $\mathbf{U}$ and $\mathbf{V}$. More precisely, SVD decomposes $\mathbf{A}$ (in the least squared

[18] Recent literature that looks at the impact of institutions on trade models the effect of institutions on trade as a product of country-specific endowment of institutions and industry-specific reliance on institutions ("institutional intensity") (Nunn, 2007; Levchenko, 2007; Costinot, 2009). The main differences between the previous literature and this paper are that this paper does not take a stand apriori on what γ_j^k and ϕ_i^k are and uses a new estimation procedure.

Figure 2: Pattern of industry productivities, in logs



sense) into $\mathbf{A} = \mathbf{USV}^{\mathrm{T}}$, where $\mathbf{S}$ is a diagonal $M \times M$ matrix with each diagonal element showing the importance or weight of each factor. SVD tries to explain as much as possible of $\mathbf{A}$ by the first factor, then uses other factors to tweak the fit.

Table 3 shows the estimated diagonal elements of $\mathbf{S}$. We immediately notice the very large explanatory power of the first factor. This means that the elements of $\mathbf{A}$ are not random, but have a structure.[19] The results also imply that there is one component with a very large explanatory power for both cross-industry and cross-country variation of relative productivities. The R^2 of the fit with only the first factor is 0.92. Figure 3 plots the fitted vs. actual productivities in all industries and countries, where the fitted productivities are obtained using only the first factor. Table 4 shows the ranking of the industries according to the estimated first industry component, γ_j^1.

If we find some industry characteristic that is highly correlated with γ_j^1 and some country characteristic that is highly correlated with ϕ_i^k/ϕ_{US}^k then we would be able to explain most of the variation in productivities $A_{ij}/A_{us,j}$. In the next sections we will be looking for such industry and country characteristics.[20]

[19] If the elements of $\mathbf{A}$ were random, the estimated diagonal elements of $\mathbf{S}$ would have been slowly declining.

[20] We have assumed that θ is the same in all countries and industries. If θ were different across industries and countries, could its variation explain the pattern of competitiveness that we observe? In order to explain the pattern described in this section, θ would have to be the same across industries in the richer countries. It would need to be lower in Metals than Medical in poor countries. Parameter θ is related to the variable of the productivity distribution of the firms in the Eaton-Kortum model. Lower θ leads to high variance of the distribution. There is no reason for θ to vary in such a way across countries and industries.

Table 3: Singular value decomposition results

1	20.79
2	1.45
3	1.30
4	1.05
5	0.93
6	0.66
7	0.53
8	0.48
9	0.45
10	0.38
11	0.36
12	0.34
13	0.29
14	0.20

Figure 3: Fitted vs. actual productivities

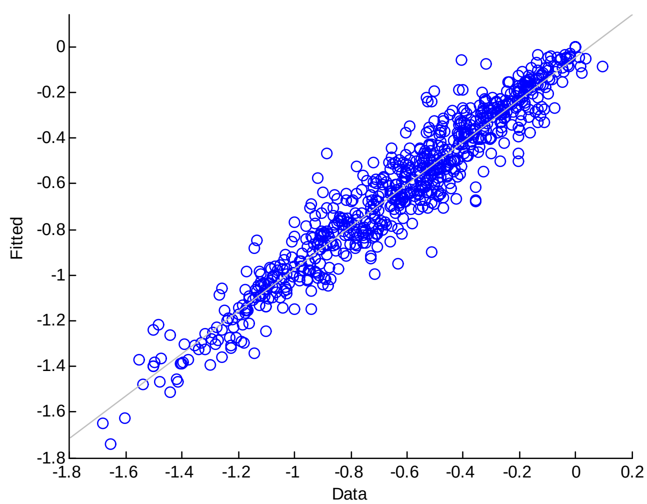


Table 4: Ranking of industries according to the first industry-specific factor γ_j^1

Metals
Food
Textile
Chemicals
Wood
Machinery, e&c
Rubber
Nonmetals
Transport
Other
Paper
Machinery, other
Metal products
Medical

5 In search of key determinants of productivity

In this section, we will search for country- and industry-level determinants of productivity. The industry determinant(s) should be highly correlated with γ_j^1 while the country determinant(s) should be highly correlated with ϕ_i^k/ϕ_{US}^k.

Literature, reviewed in the Introduction, suggests several possible causes of productivity differences across countries and industries. A combination of factor endowment differences across countries and factor intensity differences across industries can lead to productivity differences. Since factor prices and shares have been accounted for when calculating productivities, any remaining effect of factor endowments and intensities must be externalities. The are many examples of externalities coming from factor accumulation in the literature, so we will check if they play a role here.

Productivity can also be cased by differences in institutions. Countries may differ in the quality of their institutions and industries may differ in the degree to which they rely on institutions. Finally, productivity differences can come from differences in technology production or technology adoption. This section will present evidence, building on the results of the previous section, to help us evaluate these and other alternative explanations. The section will conclude with a discussion of the evidence. Section 6 will present a model based on the results of the discussion.

5.1 Country-level determinants

In this section I will look for country characteristics that are correlated with ϕ_i^1/ϕ_{US}^1. The most obvious country characteristic that is commonly used in macroeconomics is GDP per capita. The correlation between $\log(y_i/y_{US})$, where y_i/y_{US} is the GDP per capita of country i relative to the U.S., is 0.8.

5.1.1 Factor endowments

I check if factor endowments are correlated with ϕ_i^1/ϕ_{US}^1. I consider physical capital, labor with primary education, secondary education, and tertiary education. Data on physical capital per person is from Penn World Tables 8 (Feenstra et al., 2013). The correlation between physical capital per capita, $\log(k_i/k_{us})$, where k_i is capital per capita in country i, and ϕ_i^1/ϕ_{US}^1 is 0.75. All correlations are summarized in Table 6.

Table 5: Quality and endowments of labor with tertiary education in select countries

	China	Ethiopia	Germany	Slovakia	Norway	Turkey	Vietnam
Quality of education: human capital per person with tertiary education, relative to the U.S., in logs	-0.462	-0.930	-0.059	-0.206	0.013	-0.564	-1.280
Endowments: fraction of population over 25 with tertiary education in 2005							
Data from from IIASA	0.052	0.014	0.219	0.130	0.267	0.089	0.042
Relative to the U.S., in logs	-1.633	-2.933	-0.201	-0.723	-0.002	-1.097	-1.843
Adjusted for education quality, relative to the U.S., in logs	-2.094	-3.864	-0.260	-0.929	0.011	-1.662	-3.123

Table 6: Correlations between various country-level determinants and ϕ_i^1/ϕ_{US}^1

GDP per capita	0.8
Capital stock per capita	0.75
Labor with primary education	(-0.23)-(-0.09)
Labor with primary education (outliers removed)	(-0.30)-(-0.18)
Labor with secondary education	0.48-0.55
Labor with secondary education (outliers removed)	0.56-0.69
Labor with tertiary education	0.55-0.65
Labor with tertiary education (outliers removed)	0.67-0.76
Rule of law	0.69
Quality of legal system	0.65
WB Doing Business Overall Distance To Frontier 2010	0.7
WB Doing Business Distance To Frontier 2006	0.26-0.61

Note: For labor with tertiary education, Russia, Ukraine, Kazakhstan, and Bulgaria are outliers

I consider two measures of educational attainment, one from Barro and Lee (2013) and the other from IIASA/VID[21]. As before, measures of skilled labor endowments need to be adjusted for

[21] Described in Lutz, Goujon, K.C. and Sanderson (2007).

cross-country differences in education quality. I use estimates of education quality from Schoellman (2012) and Kaarsen (2014) as follows. Let L^i_{3i} be the quantity of labor with tertiary education in country i. Subscript denotes location of labor while superscript denotes quality of education. Each worker has human capital h^i_3 so the total human capital embodied in L^i_{3i} is $H_{3i} = h^i_3 L^i_{3i}$. The same human capital can be embodied in L^{us}_{3i} workers with U.S.-quality education: $H_{3i} = h^{us}_3 L^{us}_{3i}$. While L^i_{3i} is observed in the data, I need L^{us}_{3i} to make meaningful comparisons of skilled labor endowments across countries. It is obtained as $L^{us}_{3i} = L^i_{3i}\left(h^i_3/h^{us}_3\right)$, where following Schoellman (2012) and Kaarsen (2014) $h^i_3/h^{us}_3 = \exp\left(\left(\theta/\eta\right) s^{\eta}_3 \left(q^{\eta}_i - q^{\eta}_{us}\right)\right)$, where q_i is the quality of education in country i.

I look at the correlations between ϕ^1_i/ϕ^1_{US} and $\log\left(l^{us}_{ei}/l^{us}_{e,us}\right)$, where labor endowments l are measured in per capita terms. There are two sources of information on educational attainment (Barro-Lee and IIASA/VID) and two measures of educational quality (Schoellman and Kaarsen).[22] Table 5 shows endowments of labor with tertiary education for select countries using IIASA/VID data and Schoellman's measure. The first row shows $\log\left(h^i_3/h^{us}_3\right)$, the second row l^i_{3i}, the third row $\log\left(l^i_{ei}/l^{us}_{e,us}\right)$, and the last row $\log\left(l^{us}_{ei}/l^{us}_{e,us}\right)$. We can see that differences in education quality amplify gaps in effective endowments of educated labor between rich and poor countries.

Table 6 shows the range of correlations for all possible combinations of data sources, 0.55-0.65. Barro-Lee measures produce lower correlations than IIASA/VID measures. There are several countries that are outliers in terms of the relationship between ϕ^1_i/ϕ^1_{US} and $\log\left(l^{us}_{3i}/l^{us}_{3,us}\right)$. They are the former Soviet republics in my dataset (Russia, Ukraine, and Kazakhstan), and Bulgaria. These countries have high educational attainment, but relatively low ϕ^1_i/ϕ^1_{US} (and low GDP per capita).[23] Dropping these four countries raises the correlations to 0.67-0.76.

5.1.2 Role of institutions

I check how country institutional quality correlates with ϕ^1_i/ϕ^1_{US}. I use a measure of the rule of law in 1998 from Kaufmann, Kraay and Mastruzzi (2003) and the measure of the quality of legal system in 1995 from Gwartney and Lawson (2003). Both of these measures were used in Nunn (2007). I also use several measures from the World Bank's Doing Business report. One is the "overall distance to the frontier" in 2010. It is a score between 0 and 100 with 100 being the highest. Ideally, I would use data for 2005 since this is the year for which productivities were estimated. However, the overall distance to the frontier is not available for years prior to 2010 so I use distances to frontier in 9 different Doing Business report topics for 2006.[24]

The first three measures of institutions (Kaufmann, Kraay, and Mastruzzi's, Gwartney and Lawson's, and World Bank's overall distance to frontier) have similar correlations with ϕ^1_i/ϕ^1_{US}, between 0.65 and 0.7. The correlations between 9 different distances to frontier and ϕ^1_i/ϕ^1_{US} vary between 0.26 ("paying taxes") and 0.61 ("trading across borders").

Of all the variables reviewed in this section, GDP per capita has the highest correlation with ϕ^1_i/ϕ^1_{US}, so higher overall productivity is associated with higher GDP per capita. Physical capital,

[22] Both Barro-Leee and IIASA/VID educational attainment datasets combine census and other data with estimates.

[23] Schoellman's educational quality measure, which is based on earnings of immigrants, produces lower measures quality of education does a better job of accounting for lower quality of education in those four countries than Kaarsen's measure, which is derived from science and math test results. However, Schoellman's measure produces slightly lower correlations between ϕ^1_i/ϕ^1_{US} and $\log\left(l^{us}_{ei}/l^{us}_{e,us}\right)$ when outliers are dropped.

[24] Those are Starting a Business, Dealing with Construction Permits, Registering Property, Getting Credit, Protecting Minority Investors, Paying Taxes, Trading Across Borders, Enforcing Contracts, and Resolving Insolvency.

skilled labor, and institutional endowments have similar correlations with ϕ_i^1/ϕ_{US}^1, around 0.7. These three measures are also highly correlated with each other.

5.1.3 Altogether

The first principal component ϕ_i^1/ϕ_{US}^1 can actually be a linear combination of several country-level determinants. To see if this is the case, I regress it on factor endowments and a measure of institutions. Table 7 shows the results.

Table 7: Regression of ϕ_i^1/ϕ_{US}^1 on various country-level determinants

Source of educational quality data	Schoellman	Shoellman	Kaarsen
Source of institutions data	DTF10	DTF10	qc
Constant	-0.044 (0.000)	-0.044 (0.000)	-0.053 (0.000)
Physical capital per capita	0.026 (0.002)	0.030 (0.000)	0.024 (0.024)
Fraction of population with tertiary education, quality adjusted	0.014 (0.087)	0.021 (0.007)	0.018 (0.035)
Institutions	0.104 (0.032)	0.023 (0.686)	0.033 (0.197)
R^2	0.64	0.71	0.70
N	53	49	47

p-values in parentheses
Source of educational attainment data is IIASA
DTF10 is Overall distance to frontier in 2010 from WB
qc is Quality of Legal System in 1995 from Gwartney and Lawson (2003)

When physical capital is included in a regression, it is always statistically significant. Fraction of labor force with tertiary education is also statistically significant in all regressions, regardless of the source of data or measure of education quality used. Institutions are significant when all observations are included, but become insignificant when the four outlier countries are omitted. This finding is robust to the measure of institutions used. It seems that physical capital and labor with tertiary education can explain average productivity differences across countries, except in the four outlier countries. In those countries, poor institutions help explain low average productivity.

5.2 Industry-level determinants

This section looks at several sets of industry characteristics that can be correlated with γ_j^1. It will look at intensities with which industries use capital and labor. I have accounted for capital and labor costs when calculating productivity (4), but there could be effects of these factors on productivity not accounted for in the production function ("externalities"). I discuss in the next section what these effects may be.

The section will also look at the effects of institutions. Several recent papers found that industries vary in the degree to which they rely on institutions (Nunn, 2007; Levchenko, 2007; Costinot, 2009; Chor, 2010; Nunn and Trefler, 2015). These differences in institutional reliance, called institutional intensities, can lead to productivity differences.

5.2.1 Factor intensities

I start by checking if intensities of some factor of production correlate with the estimated γ_j^1. Table 8 shows the correlations between factor shares α_j, λ_1, λ_2, λ_3 and γ_j^1. The correlation between α_j and γ_j^1 is very low at 0.27. The correlation between γ_j^1 and shares of labor with primary education, λ_{1j}, is close to zero, so this type of labor is not a significant determinant of the pattern of productivity differences between rich and poor countries. The correlation between γ_j^1 and shares of labor with secondary education, λ_{2j}, is 0.51, so it is positive, but not very strong. This means that this factor of production can help explain the pattern of productivity differences, but its explanatory power is weak.

Table 8: Correlations between various industry-level determinants and γ_j^1

Capital	0.27
Labor, primary	-0.08
Labor, secondary	0.51
Labor, tertiary	0.88
Contract intensity, z^{rs1}	0.65-0.69
Contract intensity, z^{rs2}	0.76-0.78
Input concentration	0.67

z^{rs1}: fraction of inputs not sold on exchange and not reference priced
z^{rs2}: fraction of inputs not sold on exchange
Input concentration: one minus the Herfindahl index of intermediate input use

The correlation between γ_j^1 and shares of labor with at least some tertiary education, λ_{3j}, is 0.88, so it is positive and high. Since γ_j^1 is closely related to λ_{3j} and ϕ_i^1/ϕ_{US}^1 is closely related to GDP per capita, we can say that as GDP per capita decreases relative productivity falls faster in the education-intensive industries.

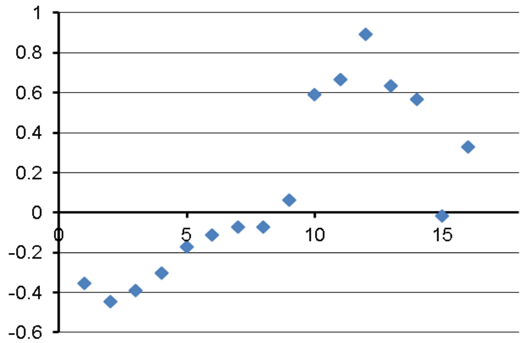
The high correlation between λ_{3j} and γ_j^1 may be surprising to some. It is important to remember that tertiary education includes many types of post-secondary schooling. Workers with technical-school education and Associate's degrees constitute a large portion of the labor force in many industries. For example, aircraft assembly typically requires workers to have an post-secondary education.

I will now use the U.S. data on labor shares to learn more precisely which type of labor is the key to the pattern of productivity. While using U.S. data to proxy for international data is not ideal, I have shown that labor shares obtained from the U.S. data are closely correlated with the shares obtained from international data, especially for the labor with tertiary education, which is our focus.

The U.S. data provides us with shares of 16 types of output shown on Figure 4. For each type of labor, I calculate the correlation between shares and γ_j^1. Figure 4 plots the correlations with the type of labor on the horizontal axis and correlation between labor shares and γ_j^1 on the vertical.

The figure shows a very clear pattern. The correlation is negative for low levels of education until $e = 8$ (12th grade, no diploma). There is a big jump between levels 9 (high school graduate) and 10

Figure 4: Correlations between γ_j^1 and shares of 16 types of labor (from the U.S. data)



1	No school completed
2	Nursery school to grade 4
3	Grade 5 or grade 6
4	Grade 7 or grade 8
5	Grade 9
6	Grade 10
7	Grade 11
8	Grade 12 no diploma
9	High school graduate
10	Some college, but less than 1 year
11	One or more years of college, no degree
12	Associate's degree
13	Bachelor's degree
14	Master's degree
15	Professional school degree
16	Doctorate degree

(some college, but less than one year). The correlation peaks at level 12 (Associate's degree) and drops after that. The correlation is close to zero for level 15 (professional degrees). The correlation for level 16 (doctorate degree) is higher than for level 9 (high school graduate), but lower than for level 10 (some college). The correlation for labor with Associate's degrees is high, 0.89.

The U.S. data also makes it possible to break down shares by educational attainment and occupation. The data shows that people with Associate's degrees work in many occupations: management, office support, production, maintenance, engineering, technicians, and others. The data also shows that industries that are more education-intensive use more educated workers in all occupations. In other words, administrators, engineeers, maintenance workers, production workers, technicians, and sales people are all more educated in the education-intensive industries.

The evidence presented in Figure 4 supports the idea that labor with Associate's degrees is key for the U.S. and other developed countries' competitiveness in manufacturing. The manufacturing operations that exist in developed countries are highly computerized and use sophisticated equipment that requires labor with specialized technical education. This education is provided by technical schools, community colleges and other institutions. In the U.S., this education can also be obtained in the armed forces.

5.2.2 Role of institutions

I also explore if γ_j^1 is correlated with some measure of institutional intensity or institutional dependence of industries. I use two industry-level measures of institutional intensity used in Nunn (2007): a measure of contract intensity and one minus the Herfindahl index of intermediate input use. Chor (2010) also uses both of these measures and Levchenko (2007) uses the latter measure.

Both of these measures try to quantify how easy it is for a producer to source intermediate goods from alternative suppliers. The first measure, contract intensity or relationship specificity, is based on Rauch's (1999) classification of goods into those sold on exchange, reference priced, or neither. Nunn (2007) combines this information with the U.S. I-O use table to calculate, for each industry, the fraction of inputs not sold on organized exchange, denoted z^{rs2}, and the fraction of inputs not sold on organized exchange or reference priced, denoted z^{rs1}. Greater z^{rs1} (or z^{rs2}) implies greater dependence on institutions. Since Rauch created two classifications of goods, conservative and liberal, depending on how he treated ambiguous cases, there are actually two measures, conservative and liberal, of z^{rs1} and two measures of z^{rs2}.

Another measure of institutional intensity is the Herfindahl index of intermediate input use, which tells us how concentrated (across industries) are the intermediate goods used by an industry. An industry that sources many of its intermediate goods from other industries will have a high value of (1 − Herfindahl index) and will depend more on institutions.

Table 8 shows the correlations between various measures of institutional intensity and γ_j^1. The correlations range between 0.65 and 0.78. The fraction of inputs not sold on organized exchange, z^{rs2}, has a fairly high correlation with γ_j^1, 0.76 or 0.78, depending on whether we use the conservative or liberal measure of z^{rs2}. Other measures of institutional intensity, z^{rs1} and one minus the Herfindahl index, have lower correlations with γ_j^1, between 0.65 and 0.69.

5.2.3 Altogether

Similarly to the first country-specific principal component, ϕ_i^1/ϕ_{US}^1, the first industry-specific principal component γ_j^1 can be a linear combination of industry-specific variables. Therefore, I regress γ_j^1 on several variables that I suspect may be important. I consider intensities of several factors of production: physical capital, labor with primary education, secondary education, and tertiary education. I also consider several measures of institutional reliance of industries: contract intensity and input concentration. Since the conservative measure of z^{rs2} has the highest correlation with γ_j^1, I concentrate on this measure of contract intensity.

Table 9: Regression of γ_j^1 on various industry-level determinants

Constant	-0.018 (0.796)
Share of physical capital	0.107 (0.763)
Share of labor with primary education	0.439 (0.855)
Share of labor with secondary education	0.228 (0.743)
Share of labor with tertiary education	1.634 (0.008)
Contract intensity*	0.150 (0.228)
R^2	0.89
N	14

p-values in parentheses
*Measure of contract intensity is z^{rs2} (conservative)

Table 9 summarizes the results. The first column shows the results of a regression of γ_j^1 on four factor intensities and a contract intensity. It shows that this model has high explanatory power ($R^2 = 0.89$), but only the share of labor with tertiary education is statistically significant.

5.3 Discussion of the results so far

Sections 4, 5.1, and 5.2 have presented many pieces of evidence. What did we learn from them? We learned that industry productivities are not random, but have a structure. As average productivity of a country declines, productivities in some industries decline faster than in others. To characterize the structure of industry productivities, we decomposed the matrix of productivities into industry and country components. The next challenge was to identify the real-world determinants of productivity that match the principal components.

Looking at the first country principal component, we found that many variables are correlated with it. These variables include physical capital, educated labor, and institutions. However, when all determinants are included jointly in a regression, only physical capital and labor with tertiary education are robust consistently statistically significant determinants of the first country principal component. Institutions are only statistically significant when the four Eastern European countries are included in the regression. The effects of physical capital and labor with tertiary education are robust to the list of countries in the regression.[25]

Looking at the first industry principal component, we found that the share of labor with tertiary education has the highest correlation with it. Measures of institutional dependence and share of labor with secondary education have lower correlations. However, only the share of labor with tertiary education is significant when all determinants are included together in a regression.[26]

Putting together the above results, labor with tertiary education is the most robust and significant determinant of both country and industry principal components. Why does labor with tertiary education have such a strong effect on industry productivities? It cannot be the differences in cost of this labor because the cost has been taken into account when calculating productivities. In other words, the direct effect of human capital, which can be called Heckscher-Ohlin or Becker-Mincer effect, has been already accounted for by including human capital into the production function.

Therefore, there must be an externality associated with this type of labor. Several models in macroeconomics (Nelson-Phelps and others) have previously suggested that the main role of human capital is to enable technology adoption. In those models, countries with high stocks of educated labor are able to adopt the latest technologies while other countries are not.

5.4 Evidence on technology adoption from licensing

In this section I review evidence on technology adoption through licensing. This evidence comes from data on the use of licensed foreign technology collected by the World Bank Enterprise Surveys. For each country and industry I calculate the percentage of plants that report usage of foreign licensed technology. The average percentage of plants in the data (across all industries and countries) that report using foreign licensed technology is about 16%. This percentage varies across industries. The correlation, across industries, between the fraction of plants which report usage of foreign technology licensing and share of workers with tertiary education is 0.6. Food and Metals industries have 13.3% and 14.6% of plants using licensed foreign technology, while Medical and Other Machinery have 24.6% and 27.1%. Therefore, education-intensive industries have much more foreign technology adoption through licensing.

However, if we decompose this data by country income we see that greater licensing in education-intensive industries occurs only in richer countries. The average percentage of plants that use foreign technology licensing is 21.6% in the upper middle income countries, 13.8% in the lower

[25] Even when all countries are included in the regression, some measures of institutions do not have statistically significant effects on the first country principal component.

[26] Therefore, this paper finds that education of the labor force is a more important determinant of productivity than institutions. Intuitively, before contract disputes can even arise, there needs to be a highly educated labor force that can read and implement complex blueprints. In modern world, contractual relationships can be substituted by networks and vertical integration (see Nunn and Trefler (2015) for discussion). Many of supplier relationships are international and disputes are increasingly resolved through international arbitration rather than domestic court systems, thus decreasing the importance of domestic institutions.

Table 10: Pattern of foreign technology licensing

	Country income		
	Upper middle income	Lower middle income	Low income
High education intensity	34%	13%	13%
Low education intensity	15%	12%	14%
Correlation between education intensity and use of licensed foreign technology			
	0.84	-0.07	0.06
Number of countries reporting data			
	14	21	11

middle income countries, and 12.8% in the low income countries.[27] The correlation between the fraction of plants which report using licensed foreign technology and share of workers with tertiary education is 0.84 in the upper middle income countries and about zero in the lower middle income and low income countries. This information is summarized in Table 10. Richer countries have more foreign technology licensing in most of the industries. However, the difference in the prevalence of foreign technology licensing between rich and poor countries is much greater in the industries with high shares of workers with tertiary education. For example, 40.5% and 50.0% of plants in Other Machinery and Medical industries of the upper middle income countries report using technology licensed from a foreign-owned company. These numbers for the low middle income countries are 17.1% and 8%.

These numbers tell us that there is much more technology diffusion through licensing in rich countries. The difference in licensing of foreign technology between rich and poor countries is much greater in the education-intensive industries. There are two possible explanations of these observations. First is that for some reason, most likely bad institutions in poor countries, innovators do not want to license technology to firms in poor countries. Since they cannot license the latest technology, poor countries cannot develop comparative advantages in the industries with high rates of innovation.

The second explanation is that poor countries have comparative disadvantage in education-intensive industries and, therefore, have much less demand for foreign-licensed technology in those industries. Poor countries cannot use the latest technology of the education-intensive industries because they do not have the pool of educated workers to use it.

Can we distinguish between these two explanations using available data? Table 10 summarizes the incidence of foreign licensed technology in different types of industries and countries. We can see that in the industries with low education intensity, the incidence of foreign licensed technology does not drop as country income drops. In the industries with high education intensity, it drop significantly, decreasing by half, as we go from upper-middle income countries to low-middle income countries. It seems reasonable to think that if bad institutions were to blame for lower incidence of foreign licensed technology use in poor countries, we would see lower incidence in all industries,

[27]The surveys in high income countries did not ask the question about use of technology licensed from a foreign-owned company.

not just the education-intensive ones. The observed pattern of licensing is caused by the pattern of the demand for licensed technologies which, in turn, is driven by countries' abilities to implement these technologies.[28]

The pattern of foreign technology licensing implies that the countries with low GDP per capita (and low endowment of labor with tertiary education, since the two measures are highly correlated) are not able to absorb the latest technology in the education-intensive industries. The following picture emerges: productivity differences are due to differences in technology adoption; differences in technology adoption are in turn driven by differences in education requirements across industries, and availability of educated labor. In the next Section, I will create a model that incorporates these features.

6 Model of technology adoption

Based on the evidence presented in the previous section, this section develops a model in which technology adoption is enabled by labor with tertiary education. This effect of educated labor is an externality, not captured by labor's market wage.

The model explains why mean productivity in industry j, A_{ij}, varies across countries. So the focus on the model is on the average productivity of an industry, rather than productivities of individual goods. In the context of the Eaton-Kortum model, the productivity of individual products within an industry will be given by draws from a statistical distribution with mean A_{ij}. In autarky, two countries i and n with $A_{ij} = A_{nj}$ will have the same average productivity across all goods in industry j, even if they may have different productivities for individual goods.

I denote technological frontier by A_j^*. It represents the best possible state of technology (productivity) for industry j. Therefore, technological frontier is delineated at the industry, not product level. A country with $A_{ij} = A_j^*$ has, on average, the best available technology in industry j. Technologies for individual products within j may be less efficient than the best available in the world.

The model takes technological frontier A_j^* as given. It does not specify how new technologies arrive.[29] Instead, it focuses on how technology is adopted around the world. I assume that frontier technology is available to all countries around the world through various means, such as licensing, foreign direct investment, import of capital goods, and publicly available information. However, countries vary in their abilities to use technologies. Each technology requires a particular quantity of educated labor to use it. Post-secondary education is key to the ability to use technology. More

[28] Licensing is one of the channels through which technology diffuses across countries. Another channel is computer use since a large fraction of productivity-enhancing innovations in recent years require computer use. I look for evidence on computer use across industries. The WBES ask for the percent of the workforce that regularly uses a computer in their jobs. This percent ranges from 13.3 and 13.5 in the Food and Metals industries to 17.8 and 27.1 in the Other Machinery and Medical industries. The correlation between computer use and share of workers with tertiary education is 0.62, so the education-intensive industries are characterized by higher use of computers. The incidence of computer use goes up with country income, from 11.6% for low income to 17.2% for high income countries. The most education-intensive industries have lower incidence of computer use in lower income countries, while the least education-intensive industries have about the same incidence. This evidence is similar to the evidence on licensing, but, unfortunately, only a few countries collected data on computer use, so this evidence should be taken with a grain of salt (there were 1 low income, 7 lower middle, and 2 upper middle income countries reporting).

[29] An example of a model of innovation is Eaton and Kortum (2001). Technolody (productivity) in their model is developed in each country by scientists through R&D.

sophisticated technology requires more workers with post-secondary education.[30]

Educated labor requirements vary across industries. Some industries have technologies that require more educated labor to use. The ability of producers in industry j of country i to use technology will depend on country i availability of educated labor and industry j requirements for educated labor. The average productivity in industry j of country i relative to the technology frontier in industry j will be a function of (a) the stock of educated labor in country i, H_i, relative to the stock of educated labor required by the frontier technology, H^*, and (b) industry j requirements for educated labor. Algebraically,

$$\frac{A_{ij}}{A_j^*} = \mu \left(\frac{H_i}{H^*} \right)^{\psi \lambda_{3j}}, \tag{15}$$

where μ and ψ are (scaling) parameters.[31]

An implication of this model is that countries with higher stocks of labor with tertiary education will have comparative advantages in more education-intensive industries. This is what was found in the previous sections of the paper. In the empirical analysis, I used the U.S. to proxy for technological frontier. The next section will evaluate how well this model fits the data.

6.1 Model fit

This section will evaluate how well the model described in the previous section can explain the pattern of productivities across countries and industries. As a benchmark, I will use the model of Section 4, given by equation (13), with only the first principal component. With $M = 1$ equation (13) becomes

$$\log \frac{A_{ij}}{A_{us,j}} = -\gamma_j^1 \log \frac{\phi_i^1}{\phi_{US}^1}, \tag{16}$$

which is equivalent to equation (15) in logs with $A_j^* = A_{us,j}$, $\mu = 1$, $H_i/H^* = \phi_i^1/\phi_{US}^1$, and $\psi \lambda_{3j} = -\gamma_j^1$.

As was reported in Section 4, the R^2 of this model is 0.92. This means the the first principal component can explain 92% of the variance of productivities across countries and industries. There are 53*14=742 productivities $A_{ij}/A_{us,j}$ on the left-hand side of (16). They are explained by the first principal component vectors ϕ_i^1 and γ_j^1 that have 53+14=67 elements in total.

The stock of educated labor in (15) will be measured by the stock of labor with tertiary education $l_{3i}^{us}/l_{3,us}^{us}$ described in Section 5.1.1. Data on educational attainment is from IIASA/VID and educational quality is from Schoellman. Taking logs of (15) we have

$$\log \frac{A_{ij}}{A_{us,j}} = \log \mu + \psi \lambda_{3j} \log \frac{l_{3i}^{us}}{l_{3,us}^{us}} \tag{17}$$

The R^2 for this model is 0.41. If the four outlier countries (Russia, Ukraine, Kazakhstan, and Bulgaria) are dropped, the R^2 increases to 0.50. So the model can explain 50% of the variation in the productivities. Full results of this regression are shown in Table 11.

[30] Simple technologies may require only the manager of the plant to have post-secondary education. More sophisticated technologies may require line managers to have post-secondary education. Even more sophisticated technology requires even line workers and support personnel to have post-secondary education in order to operate computerized machinery, follow complex blueprints, and provide support.

[31] This is a static model so time indices are dropped for convenience.

Table 11: Regression of
$\log A_{ij}/A_{us,j}$ on $\log l_{3i}^{us}/l_{3,us}^{us}$

$\log \mu$	-0.300 (0.000)
ψ	3.812 (0.000)
R^2	0.5
N	686

p-values in parentheses

Table 12: Accounting for productivity gaps between the most and least productive countries in each industry, $\max_i A_{ij}/\min_i A_{ij}$

	Data			Model	
	Capital and one type of labor	Capital and three types of labor	Labor adjusted for differences in education quality	SVD model	Model of technology adoption
Food	3.27	3.29	3.10	2.66	2.03
Textile	3.30	3.21	3.10	2.81	2.30
Wood	4.40	4.60	4.40	3.38	2.21
Paper	5.39	5.06	4.40	4.33	3.99
Chemicals	3.98	4.07	3.55	3.15	3.21
Rubber	4.92	4.66	4.45	3.93	3.23
Nonmetals	4.45	4.70	4.52	3.98	2.62
Metals	2.73	2.90	2.65	2.45	1.87
Metal products	5.71	5.57	4.67	4.40	3.79
Machinery, other	6.52	6.11	4.72	4.33	4.75
Machinery, e&c	5.95	5.61	4.49	3.77	3.65
Medical	7.28	6.83	5.37	5.70	7.74
Transport	5.00	4.86	4.03	4.04	2.83
Other	5.34	4.99	4.50	4.06	3.08

One of the goals of the models presented here, (15) and (16), is the explain the productivity gaps across countries. Table 12 shows productivity gaps between the most productive and least productive countries in every industry, $\max_i A_{ij}/\min_i A_{ij}$. The first three columns show productivity gaps estimated from data using different production functions. The first column uses the production function with capital and labor, measured by the number of workers. The labor in this case is not disaggregated by education, as in the rest of this paper. Stocks of labor are not adjusted for educational attainment, either. This is the most basic approach to incorporating labor into the production function when calculating productivities. The second column uses the production function with three types of labor, but without accounting for differences in education quality across countries. The third column uses three types of labor and accounts for education quality differences. This is the approach taken in this paper to calculate productivities.

Comparing columns 1 and 2, we can see that disaggregating labor into three types makes little difference to the productivity gaps, which are under 3 in the Metals industry and about 7 in the Medical industry.[32] Accounting for differences in education quality makes a noticeable reduction in productivity gaps. The gap in the Medical industry falls from 7.28 (calculated using the "basic approach") to 5.37. The average reduction going from column 1 to column 3 is 16%.

Columns 4 and 5 show productivity gaps predicted by the SVD model (16) and the model of technology adoption (17). The gaps predicted by the SVD model track fairly closely those estimated from the data, shown in column 3.

7 Conclusion

Productivity determines the comparative advantages of countries, but productivity is calculated as a residual and, therefore, is a "measure of our ignorance". The goal of this paper is to endogenize the industry-level productivities that determine comparative advantages.

The approach of this paper is different from the existing literature. I start by estimating productivity in autarky for each industry and country. Then I look for a pattern in these productivities across industries and countries, without making assumptions regarding the determinants of productivities.

In the spirit of the Heckscher-Ohlin model, I decompose productivities into industry and country-specific components using a statistical technique called singular value decomposition. This approach turns out to be very successful empirically, in contrast to the previous evidence on the Hechscher-Ohlin model, which is mixed. The main departure from the previous literature is that I do not take a stand apriori on what the industry and country determinants of productivity are. I find that the interaction of the first principal industry and country components can explain the vast majority of variation in the productivity matrix. In fact, one country-specific component and one industry-specific component can explain 92% of the variation in the productivities of 14 industries and 53 countries.

Having decomposed productivities in this manner, I look for country- and industry-specific variables that correlate highly with the country- an industry-specific components produced by the singular value decomposition. I consider physical capital, labor with three different levels of education, and various measures of institutions. I find that among all of these variables, the endowment of labor with tertiary education has the highest correlation with the country-specific

[32] It increases productivity gaps in the industries with low gaps and decreases productivity gaps in the industries with high gaps.

component while the intensity of labor with tertiary education has the highest correlation with the industry-specific component.

Several important conclusions emerge from the analysis. First, countries with high average productivity have comparative advantages in the industries that use highly educated labor more intensively. These countries also have high GDP per capita. Second, highly educated labor is a key determinant of productivity across both countries and industries. Previous macro literature found that human capital explains the pattern of productivities across countries. This paper finds that human capital also explains the pattern of productivities across industries.

The fourth conclusion is that industry matters as a unit of analysis. The productivity differences across industries are not random, but contain important information. Previous literature has found that capital and labor intensities have little explanatory power for the pattern of trade. Those results could lead one to conclude that the industry dimension is not important. This paper finds that while some types of labor have little explanatory power for the pattern of trade, labor with tertiary education can explain a significant portion of the pattern of trade in manufactures. Classification of labor by education is the one most suited for the analysis of productivity. Classification of labor into production/non-production and skilled/unskilled are much less relevant.

Even though highly educated labor is an important determinant of trade, the effect of this labor is not through its wage. In other words, differences in marginal product of labor across different levels of education and countries are not big enough to explain productivity differences. Therefore, there is an externality associated with highly educated labor.

Existing literature tells us that educated labor is not just a factor of production, but also a factor that facilitates technology adoption. This has important policy implications because the benefits of educated labor extend beyond its marginal product. This function of educated labor was modeled by Nelson and Phelps, Acemoglu and Zilibotti, and others for the whole economy. I create a simple model of technology adoption at the industry level, parametrize it, and show that it fits the data well.

I further study technology adoption across industries by looking at the foreign technology licensing data. While there are several ways that technology can diffuse across countries, licensing is the most direct route. I find that foreign technology licensing is the most prevalent in the education-intensive industries of the rich countries. While poor countries license as much foreign technology as the rich countries in the non-education-intensive industries, they license much less in the education-intensive ones. This happens because poor countries have little educated labor and are not able to adopt the latest technology in the education-intensive industries, which requires more educated labor.

After parametrizing the model of technology adoption, I find that it can explain 50% of the variation in productivities across countries and industries. I also find that there are four countries in Eastern Europe for which this model does work well and where institutional deficiencies seem to play an important role.

There are several important implications of the model of technology adoption. First, since the endowments of educated labor change slowly, so does the pattern of comparative advantages. Second, governments that wish to change the comparative advantage of their countries should focus on growing the pool of labor with tertiary education and improving the quality of education. More specifically, they should focus on increasing the number of workers with an equivalent of an Associate's degree. This level of education provides the necessary skills to operate and maintain the sophisticated computerized machinery used in modern manufacturing.

References

Abramovitz, M. (1956). Resource and output trends in the United States since 1870, *American Economic Review* **46**(2): 5–23.

Acemoglu, D. and Zilibotti, F. (2001). Productivity differences, *Quarterly Journal of Economics* **116**(2): 563–606.

Baldwin, R. E. (1971). Determinants of the commodity structure of U.S. trade, *The American Economic Review* **61**(1): 126–146.

Baldwin, R. E. (1979). Determinants of trade and foreign investment: Further evidence, *The Review of Economics and Statistics* **61**(1): 40–48.

Barro, R. (1991). Economic growth in a cross section of countries, *Quarterly Journal of Economics* **106**: 407–444.

Barro, R. J. and Lee, J.-W. (2001). International data on educational attainment: Updates and implications, *Oxford Economic Papers* **53**: 541–563.

Barro, R. J. and Lee, J.-W. (2013). A new data set of educational attainment in the world, 1950-2010, *Journal of Development Economics* **104**: 184–198.

Basu, S. and Weil, D. (1998). Appropriate technology and growth, *The Quarterly Journal of Economics* .

Benhabib, J. and Spiegel, M. M. (2005). Human capital and technology diffusion, *in* P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, Vol. 1A, Elsevier B.V., chapter 13.

Bils, M. and Klenow, P. J. (2000). Does schooling cause growth?, *American Economic Review* **90**(5): 1160–1183.

Caselli, F. (2005). Accounting for cross-country income differences, *in* P. Aghion and S. N. Durlauf (eds), *Handbook of Economic Growth*, Vol. 1A, Elsevier B.V., chapter 9.

Chor, D. (2010). Unpacking sources of comparative advantage: A quantitative approach, *Journal of International Economics* **82**(2): 152–167.

Ciccone, A. and Papaioannou, E. (2009). Human capital, the structure of production, and growth, *The Review of Economics and Statistics* **91**(1): 66–82.

Costinot, A. (2009). On the origins of comparative advantage, *Journal of International Economics* **77**: 255–264.

Davis, D. R. and Weinstein, D. E. (2001). An account of global factor trade, *The American Economic Review* **91**(5): 1423–1453.

Dornbusch, R., Fischer, S. and Samuelson, P. (1977). Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods, *American Economic Review* **67**: 823–839.

Eaton, J. and Kortum, S. (2001). Technology, trade, and growth: A unified framework, *European Economic Review* **45**: 742–755.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade, *Econometrica* **70**(5): 1741–1779.

Eaton, J. and Kortum, S. (2012). Putting Ricardo to work, *Journal of Economic Perspectives* **26**(2): 65–90.

Erosa, A., Koreshkova, T. and Restuccia, D. (2007). How important is human capital? A quantitative theory assessment of world income inequality, *University of Toronto mimeo* .

Feenstra, R. C., Inklaar, R. and Timmer, M. P. (2013). The next generation of the Penn World Table, *Available for download at www.ggdc.net/pwt* .

Gwartney, J. and Lawson, R. (2003). Economic freedom of the world: 2003 annual report, *Fraser Institute Mimeograph* .

Hall, R. E. and Jones, C. I. (1999). Why do some countries produce so much more output per worker than others?, *The Quarterly Journal of Economics* .

Hanushek, E. A. and Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations, *American Economic Review* **90**(5): 1184–1208.

Harrigan, J. (1997). Technology, factor supplies, and international specialization: Estimating the neoclassical model, *The American Economic Review* **87**(4): 475–494.

Heckscher, E. (1919). The effects of foreign trade on the distribution of income, *Ekonomisk Tidskrift* **21**: 497–512.

Hendricks, L. (2002). How important is human capital for development? Evidence from immigrant earnings, *American Economic Review* **92**(1): 198–219.

Kaarsen, N. (2014). Cross-country differences in the quality of schooling, *Journal of Development Economics* **107**: 215–224.

Kaufmann, D., Kraay, A. and Mastruzzi, M. (2003). Governance matters iii: Governance indicators for 1996Ű2002, *World Bank Policy Research Working Paper No. 3106* .

Keesing, D. B. (1966). Labor skills and comparative advantage, *The American Economic Review* **56**(1/2): 249–258.

Krueger, A. B. and Lindahl, M. (2001). Education for growth: Why and for whom?, *Journal of Economic Literature* **39**(4): 1101–1136.

Levchenko, A. A. (2007). Institutional quality and international trade, *Review of Economic Studies* **74**(3): 791–819.

Lutz, W., Goujon, A., K.C., S. and Sanderson, W. (2007). Reconstruction of populations by age, sex and level of educational attainment for 120 countries for 1970-2000, *Vienna Yearbook of Population Research* **5**: 193–235.

MacDougall, G. D. A. (1951). British and American exports: A study suggested by the theory of comparative costs, *The Economic Journal* **61**(244).

Mankiw, G. N., Romer, D. and Weil, D. N. (1992). A contributionn to the empirics of economic growth, *Quarterly Journal of Economics* **107**: 407–437.

Manuelli, R. E. and Seshadri, A. (2010). Human capital and the wealth of nations, *University of Wisconsin mimeo* .

Mincer, J. A. (1974). *Schooling, Experience, and Earnings*, NBER Books, National Bureau of Economic Research.

Nelson, R. R. and Phelps, E. S. (1966). Investment in humans, technical diffusion, and economic growth, *American Economic Review* **56**(2): 69–75.

Nunn, N. (2007). Relationship-specificity, incomplete contracts, and the pattern of trade, *Quarterly Journal of Economics* **122**(2): 569–600.

Nunn, N. and Trefler, D. (2015). Domestic institutions as a source of comparative advantage, *in* E. Helpman, K. Rogoff and G. Gopinath (eds), *Handbook of International Economics*, Vol. 4, Elsevier, pp. 263–315.

Ohlin, B. (1924). *Handelns Teori*, AB Nordiska Bokhandeln.

Rauch, J. E. (1999). Networks versus markets in international trade, *Journal of International Economics* **48**: 7–35.

Ricardo, D. (1817). *The Principles of Political Economy and Taxation*, John Murray, London.

Romalis, J. (2004). Factor proportions and the structure of commodity trade, *American Economic Review* **94**(1).

Romer, P. M. (1990). Endogenous technological change, *Journal of Political Economy* **98**: S71–S102.

Schoellman, T. (2012). Education quality and development accounting, *Review of Economic Studies* **79**: 388–417.

Shikher, S. (2012). Putting industries into the Eaton-Kortum model, *Journal of International Trade and Economic Development* **21**(6): 807–837.

Trefler, D. (1995). The case of the missing trade and other mysteries, *The American Economic Review* **85**(5): 1029–1046.

Yaylaci, O. (2013). Evolution of trade costs, *Suffolk University mimeo* .

Yaylaci, O. and Shikher, S. (2014). What would Korea-US free-trade agreement bring?, *International Economic Journal* **28**(1): 161–182.