East is East and West is West:
A Ricardian-Heckscher-Ohlin Model of Comparative Advantage

Peter M. Morrow
The University of Toronto
June 1st, 2008

Abstract

This paper derives and estimates a unified and tractable model of comparative advantage due to differences in both factor abundance *and* relative productivity differences across industries. It derives conditions when ignoring one force for comparative advantage biases empirical tests of the other. I emphasize two empirical results. First, factor abundance- and relative productivity-based models each possess explanatory power. Second, productivity differences across industries are uncorrelated with the factor intensities of these industries. Therefore, the two models each offer valid partial descriptions of the data and ignoring one force for comparative advantage does not bias empirical tests of the other.

JEL Codes: F10, F11, F12.
Keywords: Heckscher-Ohlin, Ricardian, Increasing Returns to Scale, Omitted Variable Bias.

I thank Mary Amiti, Olivier Coibion, Alan Deardorff, Ann Ferris, Yuriy Gorodnichenko, Juan Carlos Hallak, James Levinsohn, Jagadeesh Sivadasan, Gary Solon, Daniel Trefler, and seminar participants at Columbia University, the Federal Reserve Bank of New York, the University of Michigan, Syracuse University, and the University of Toronto for helpful comments and suggestions. All errors are mine and mine alone. The title of this paper comes from the poem "The Ballad of East and West" by Rudyard Kipling, which begins "Oh, East is East, and West is West, and never the twain shall meet...."

Contact information: Peter M. Morrow, Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON, M5S 3G7, CANADA, email: peter.morrow@utoronto.ca, phone: 416-978-4375

1 Introduction

Production patterns around the world exhibit tremendous heterogeneity and specialization. For example, the United States supplies 35.0% of the world's exports of aircraft while China provides only 0.1%. On the other hand, China supplies 25.8% of the world's export supply of apparel and clothing while the United States only supplies 2.4%.[1] The Ricardian and Heckscher-Ohlin (HO) theories are the two workhorse models used to explain this specialization. The Ricardian model of international trade predicts that countries specialize in goods in which they hold the greatest relative advantage in total factor productivity (TFP). The Heckscher-Ohlin model ignores differences in TFP across industries and assumes that all countries possess the same production function in a given industry. Heckscher-Ohlin asserts that differences in comparative advantage come from differences in factor abundance and in the factor intensity of goods. Specifically, Heckscher-Ohlin predicts that countries will produce relatively more of the goods that use their relatively abundant factors relatively intensively. Neither model, in isolation, offers a unified theory as to why production patterns differ across countries and industries. Consequently, empirical tests of each model can be subject to omitted variable biases associated with ignoring the other.

Such a bias can emerge if countries that possess a relative abundance of a factor also possess levels of relative TFP that are systematically higher (or lower) in industries that use this factor relatively intensively. In trying to explain patterns of skill-biased-technical-change, Acemoglu (1998) suggests that skilled labor-abundant countries will have higher levels of relative TFP in skilled labor-intensive industries than in unskilled labor-intensive industries.[2] Thus, if the mechanisms in his model are pervasive in the data, economists will tend to confound the HO and Ricardian models when one is tested without the other as a meaningful alternate hypothesis. Simply put, it is possible that skilled labor-abundant countries will produce skilled labor-intensive goods both because of their relative abundance of skilled labor and high TFP in skilled labor intensive sectors.[3] There is anecdotal support for this idea. Empirically, Kahn and Lim (1998) find that TFP in the

[1] Data taken from "World Trade Flows" bilateral trade data compiled by Robert Feenstra et al. (2005) for the year 2000. Aircraft is SITC code 792 and Clothing and Apparel is SITC code 84.

[2] However, he also shows that all predictions about *relative* TFP across sectors depend crucially on the enforcement of Northern property rights of technology in the South.

[3] This possibility has also been the subject of conjecture by authors such as Fitzgerald and Hallak (2004), although the modeling techniques have not been developed for empirical examination.

United States in the 1970s increased far more in skill-intensive industries than in industries that use unskilled labor relatively intensively. On the other side, if Ricardian TFP differences influence production patterns in a manner that is inconsistent with HO, this might suggest why HO results sometimes appear to be unstable.[4]

This paper articulates a unified and tractable framework in which comparative advantage exists due to differences in factor abundance *and/or* relative productivity differences across a continuum of monopolistically competitive industries with increasing returns to scale. In this manner, I rely on the quasi-Heckscher-Ohlin market structure of Romalis (2004) while augmenting his model with Ricardian TFP differences. By developing a tractable model that possesses theoretically meaningful nested hypotheses, I can use traditional estimation techniques to dissect patterns of comparative advantage into those driven by Ricardian forces and those driven by HO. I also derive conditions under which tests of the HO model will not suffer from an omitted variable bias if they ignore Ricardian TFP differences.

Empirically, I estimate the model using panel data across 20 developed and developing countries, 24 manufacturing industries, and 11 years (1985-1995). I highlight three important findings. First, both the Ricardian and HO models possess robust explanatory power in determining international patterns of production. Although this has been documented in past work, this is the first to model production and demand in a jointly HO/Ricardian setting where reduced form coefficients can be mapped against structural parameters such as the elasticity of substitution or iceberg transportation costs.

Second, the two models are empirically separable in my broad sample in that the forces that determine comparative advantage in one model are orthogonal to the forces that determine comparative advantage in the other. Specifically, while both productivity differences and the interaction of factor abundance with factor intensity play a role in determining international specialization patterns, I find very little evidence indicating that relative productivity levels are systematically correlated with factor intensity. This suggests that productivity levels that are non-neutral across industries have little influence over whether results consistent with HO appear in the data. Consequently, ignoring one force for comparative advantage will not bias empirical tests of the other.

[4]e.g. Bowen, Leamer and Sveikauskas (1987).

Third, I find that a one standard deviation increase in relative factor abundance is approximately twice as potent in affecting change in the commodity structure of the economy as a one standard deviation change in Ricardian productivity. This suggests that differences in factor abundance are more potent than differences in Ricardian productivity in determining patterns of specialization. The second and third results are new and provide substantial insight into how we can integrate these two important models.

The key to nesting the Ricardian alternate hypotheses involves decomposing industry-level TFP differences into three components: country-level TFP that differs across countries but is identical across industries within any given country, productivity that is correlated with factor intensity and purged of country averages, and productivity that varies across industries but is orthogonal to factor intensity and is purged of country averages. If productivity is correlated with factor intensity, the two models can be confounded easily and tests of a single model will typically suffer from omitted variable bias. If TFP is orthogonal to factor intensity, it is reasonable to model TFP as consisting of a country-specific term that is neutral across industries and an idiosyncratic component that is orthogonal to factor intensity.

An important theoretical contribution of this paper is that when TFP is uncorrelated with factor intensity, HO is valid as a partial description of the data. Consequently, common tests of and the standard comparative statics associated with the HO model (e.g. Rybczynski regressions) are valid because Ricardian TFP differences are not correlated with the factor intensity differences across goods that are the foundation of most of these empirical tests. However, industry-by-industry *level* predictions must take Ricardian differences into account. For example, the change in the commodity structure resulting from a change in number of skilled workers in a country can be estimated from an HO model but the level of production accruing to a certain industry must take HO and Ricardian mechanisms into account. Examining if relative TFP is correlated with factor intensity in other data sets will suggest whether this orthogonality assumption is valid in other cases.

1.1 Relation to the Literature

This paper is related to two distinct strands of literature on the empirical determinants of specialization and trade and their occasional intermingling. The first strand documents the influence of Ricardian TFP on international production patterns. MacDougall (1951,1952) finds early evidence for the Ricardian model using data from the United Kingdom and the United States. Costinot and Komunjer (2007) augment the model of Eaton and Kortum (2002) to include industries and find that relative value-added per worker possesses predictive power in determining patterns of industrial specialization in a broad panel of countries. The second related strand of literature documents the importance of factor abundance and includes Leontief (1954), Baldwin (1971), Davis and Weinstein (1999), Debaere (2003) and Romalis (2004).[5]

This is far from the first paper to examine empirically the interaction of productivity- and HO-based models. However, many prior explorations have been highly restrictive "fixes" that are more concerned with improving the fit of the HO/HOV model than with considering the Ricardian hypothesis on its own merits. For example, Trefler (1993) shows how factor-augmenting technology differences can improve the fit of the HOV model by improving measurements of factor abundance.[6] Trefler (1995) shows how country-specific productivity differences can dramatically improve the results of the HOV model. However, because TFP differences in that paper are country-wide, they are not of the Ricardian nature that I examine here.

Harrigan (1997) is the closest antecedent to this paper. He examines the contributions of TFP and factor abundance in determining specialization in a series of reduced form industry-level studies. He does not examine the conditions under which the omission of Ricardian technology introduces systematic biases in tests of the HO model.[7] This paper contributes to the literature by deriving a condition under which ignoring one force for comparative advantage will or will not bias empirical tests of the other and finds that this condition holds empirically in the data set examined. It also

[5] For thorough surveys of empirical tests of theories of trade, see Deardorff (1985) and Leamer and Levinsohn (1995).

[6] Trefler concedes that it is difficult to untangle pure productivity effects from alternate hypotheses such as that the capital-labor ratio varies across countries (pg. 979-980)

[7] Rather, referring to HO and productivity based forces, he states "these forces must be considered jointly when formulating policies intended to affect the structure of production and trade. (pg. 492)" This implies that omitting consideration of one of these models when examining the other produces results that are incorrect at best or misleading at worst.

articulates a framework within which reduced form coefficients can be mapped against structural parameters unlike previous work integrating the two approaches.

Earlier theoretical work on integrating HO and Ricardian models of comparative advantage include Findlay and Grubert (1959), who were among the first to use a two country, two good, two factor model to consider the effects of Ricardian productivity and factor abundance in jointly determining factor prices and production patterns. Bernard, Schott and Redding (2006) use Melitz's (2003) model of firm TFP heterogeneity with factor abundance differences to derive results consistent with the HO theorem.[8]

1.2 Structure of the Paper

The paper is organized as follows. Section 2 sketches a simple two industry, two country, two factor version of the model. Section 3 extends the framework to a continuum of industries and derives empirically testable expressions. Section 4 describes the data and the construction of the total factor productivity measures used in the paper. Section 5 presents the baseline results. Section 6 presents robustness tests, and Section 7 concludes.

2 Theory: A Simple 2x2x2 Model

I first sketch a simple two country, two factor, two industry model to illustrate the insights of the more general model of Section 3. My model augments the quasi-Heckscher-Ohlin structure of Romalis (2004) with Ricardian TFP differences. This simple model solves for equilibrium factor prices and production as functions of exogenous factor abundance and productivity using two equilibrium conditions to extract the separate contributions of productivity and factor abundance on relative production patterns across industries in a country. I focus on the case where both countries produce in each industry such that intra-industry trade exists. I start by deriving a *goods market clearing condition* that maps relative factor prices to relative production values of goods demanded from skilled and unskilled labor intensive industries. I close the model by deriving a

[8]Their model focuses on the case where firms take productivity draws from the same distribution across industries. Consequently, any differences in average TFP across industries and countries are only endogenous responses to exogenous differences in factor abundance.

factor market clearing condition that assures full employment for each of the two factors. I then illustrate conditions under which Ricardian productivity differences can introduce substantial biases in empirical tests of the HO model.

2.1 Production

This section presents the supply side of the model including the production function and the pricing behavior of a firm. The two factors of production are skilled labor (S) and unskilled labor (U). The wages of these two factors are w_s and w_u, respectively. Let $\omega \equiv \frac{w_s}{w_u}$. For simplicity, define the two countries as the North and the South. All Southern values are indicated by asterisks.

The two industries are indexed by their Cobb-Douglas skilled labor factor cost shares (z) where $z = \frac{w_s S(z)}{w_s S(z) + w_u U(z)}$ and $0 < z < 1$. z_s is the skilled labor cost share of the skilled labor intensive good and z_u is the skilled labor cost share of the unskilled labor intensive good. Consequently, z is both a parameter and an index of industries. Without loss of generality, assume that $z_s > z_u$. Hicks neutral TFP ($A(z)$) augments skilled and unskilled labor in production of a final good ($x(z)$) and coverage of fixed costs such that total cost for a given Northern firm i in industry z takes the following form:

$$TC(z,i) = [x(z,i) + f(z)] \frac{w_s^z w_u^{1-z}}{z^z (1-z)^{1-z} A(z)}. \tag{1}$$

As is common in the literature, I assume that skilled and unskilled labor are used in the same proportion in fixed costs as in marginal costs. Previewing the demand structure, prices are a constant markup over marginal cost. The markup is equal to $\frac{1}{\rho}$ where $0 < \rho < 1$ and $\frac{1}{1-\rho}$ is the elasticity of substitution between varieties within an industry. A zero profit condition solves for output per firm, $x(z) = \frac{\rho f(z)}{1-\rho}$. Assume that the elasticity of substitution and fixed costs are the same in the two countries for a given industry so that output per firm is constant across countries within an industry. I further assume that all firms within an industry and country have access to the same production function and face the same factor prices. Therefore, for a given industry z, the price of a Northern good relative to its Southern equivalent can be expressed as follows where Northern relative to Southern values are indicated by tildes:

$$\tilde{p}(z) = \frac{\tilde{w}_s^z \tilde{w}_u^{1-z}}{\tilde{A}(z)} = \frac{\tilde{\omega}^z \tilde{w}_u}{\tilde{A}(z)}. \tag{2}$$

The following notation introduces Ricardian productivity differences:

$$\tilde{\gamma} \equiv \frac{\tilde{A}(z_s)}{\tilde{A}(z_u)} = \frac{\frac{A(z_s)}{A(z_u)}}{\frac{A^*(z_s)}{A^*(z_u)}}. \tag{3}$$

If $\tilde{\gamma} > 1$, the North is relatively more productive in the skill intensive industry than the unskilled intensive industry. If $\tilde{\gamma} < 1$, the North is relatively more productive in the unskilled labor intensive industry. If $\tilde{\gamma} = 1$, the North is equally relatively productive in the two industries.

2.2 Demand

Consumers in each of the two countries have utility (Υ) that is Cobb-Douglas over the two industries but CES across varieties within each of the industries. Although I loosen this assumption in the more general section, the expenditure share for each industry is constant and equal to 0.5. Each firm produces a unique imperfectly substitutable variety so that "firms" and "varieties" are synonymous. For a given industry z, $n(z)$ is the endogenously determined number of Northern firms and $n^*(z)$ is the number of Southern firms and the total number of firms in a given industry is $N(z) = n(z) + n^*(z)$ where i indexes firms within industry z.

$$\Upsilon = C(z_s)^{0.5} C(z_u)^{0.5} \tag{4}$$

$$C(z_k) = \left[\int_0^{N(z_k)} x(z_k, i)^\rho di \right]^{\frac{1}{\rho}} \quad k \in S, U \tag{5}$$

Consumers buying from a foreign firm incur iceberg transportation costs $\tau > 1$ such that if the price of a domestically produced good is $p(z)$ then the price of the same good abroad is $\tau p(z)$. Revenue accruing to a firm is equal to its receipts from domestic and foreign consumers. Appendix A shows how Northern and Southern firms' revenue functions can be used to solve for the number of Northern firms relative to the number of Southern firms in a given industry $\left(\frac{n(z)}{n^*(z)}\right)$ as Romalis (2004):

$$\tilde{n} = \frac{\tau^{2(1-\sigma)}\frac{Y^*}{Y} + 1 - \tau^{1-\sigma}\tilde{p}^{\sigma}(\frac{Y^*}{Y}+1)}{\tilde{p}(\tau^{2(1-\sigma)} + \frac{Y^*}{Y}) - \tilde{p}^{1-\sigma}\tau^{1-\sigma}(\frac{Y^*}{Y}+1)}. \tag{6}$$

Because output per firm is pinned down, aggregate Northern revenue relative to aggregate Southern revenue in industry z is

$$\tilde{R}(\tilde{p}(z)) = \frac{n(z)p(z)x(z)}{n(z)^*p(z)^*x(z)^*} = \frac{\tau^{2(1-\sigma)}\frac{Y^*}{Y} + 1 - \tau^{1-\sigma}\tilde{p}(z)^{\sigma}(\frac{Y^*}{Y}+1)}{\tau^{2(1-\sigma)} + \frac{Y^*}{Y} - \tilde{p}(z)^{-\sigma}\tau^{1-\sigma}(\frac{Y^*}{Y}+1)}. \tag{7}$$

I restrict my attention to the case where $\tilde{R}(z) > 0$ such that both the North and South produce in a given industry.[9] Romalis (2004) derives restrictions on $\tilde{p}(z)$ that provide necessary and sufficient conditions for $\tilde{R}(z) > 0$ and I assume that these conditions hold. Appendix A shows that in a diversified equilibrium firms produce on the elastic portion of their demand curve such that $\frac{\partial \tilde{n}}{\partial \tilde{p}} < 0$ and $\frac{\partial \tilde{R}}{\partial \tilde{p}} < 0$.[10] Finally, it is straightforward to show that the signs of these derivatives ensure that the share of revenue in industry z accruing to the North is also decreasing in $\tilde{p}(z)$ where the share is defined as $v(z) = \frac{R(z)}{R(z)+R^*(z)} = \frac{\tilde{R}(z)}{\tilde{R}(z)+1}$.

2.3 Equilibrium

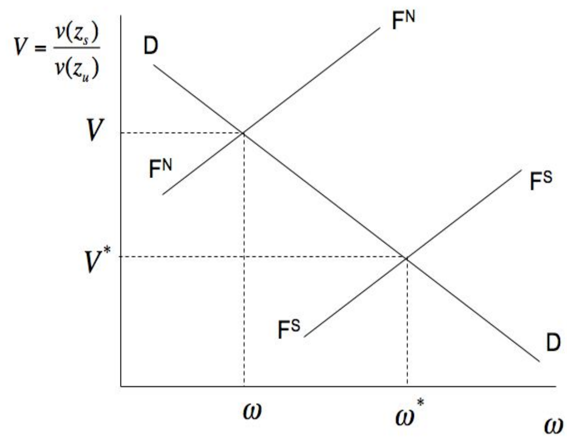
To illlustrate the equilibrium, I start by deriving the *goods market clearing condition.* Factor price equalization fails due to transportation costs in this two country setting. Starting with the simple case where comparative advantage only comes from differences in factor abundance, if $\omega^* > \omega$ then $\frac{v(z_s)}{v(z_u)} > \frac{v^*(z_s)}{v^*(z_u)}$. That is, the relative value of goods demanded in an industry will be *declining* in the relative wage of the factor that is used relatively intensively in that industry. Appendix B derives this rigorously. Figure 1 depicts this *goods market clearing condition* with the line DD.

A set of *factor market clearing conditions* close the model. Define world income as $Y^w = Y + Y^*$. Based on Cobb-Douglas production, the ratio of aggregate payments to skilled labor relative to unskilled labor in the North is

[9] The intuition for the model is unchanged when allowing for specialization although solving for equilibrium production patterns becomes more complex.

[10] As Romalis (2004) notes, as $\sigma \to \infty$ and $\tau = 1$ the model becomes one of perfect competition as in DFS (1977) for the case of comparative advantage from Ricardian productivity and DFS (1980) for the HO case. With transportation costs and perfect competition, there are non-traded goods and no intra-industry trade. With monopolistic competition but no transportation costs, FPE results as long as factor endowments are not too dissimilar, the location of production becomes indeterminate for a given industry and we cannot make industry-by-industry predictions.

Figure 1: Equilibrium: Heckscher-Ohlin Model



$$\frac{0.5\sum_{k\in s,u} v(z_k)z_k Y^w}{0.5\sum_{k\in s,u} v(z_k)(1-z_k)Y^w} = \frac{w_s}{w_u}\frac{S}{U} = \omega\frac{S}{U}. \tag{8}$$

Simple manipulation gives

$$\frac{z_u + \frac{v(z_s)}{v(z_u)}z_s}{(1-z_u) + \frac{v(z_s)}{v(z_u)}(1-z_s)} = \omega\frac{S}{U}. \tag{9}$$

Taking a total derivative of the above expression holding $\frac{S}{U}$ constant gives

$$\Omega d\frac{v(z_s)}{v(z_u)} = \Omega dV = \frac{S}{U}d\omega \tag{10}$$

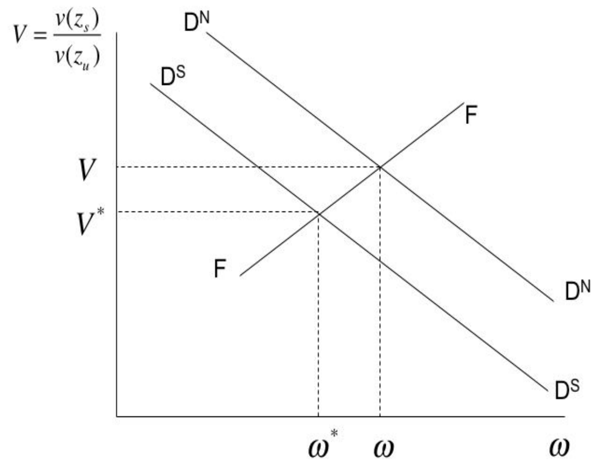
where

$$\Omega = \frac{z_s - z_u}{\left[1 - z_u + \frac{v(z_s)}{v(z_u)}(1-z_s)\right]^2} > 0.$$

Because $z_s > z_u$, $\Omega > 0$ and the relative wage of the factor used relatively intensively in an industry will increase as productive factors are reallocated to that industry. Examining Figure 1, $F^N F^N$ is the factor market clearing condition for the Northern country and the Southern factor market clearing condition $F^S F^S$ is below and to the right of $F^N F^N$. The location of $F^S F^S$ relative to $F^N F^N$ is given by solving for $\frac{d\omega}{d\frac{S}{U}}$ using equation 9.

Figure 1 confirms the intuition of the simplest HO model. The North possesses a relative abun-

Figure 2: Equilibrium: Ricardian Model

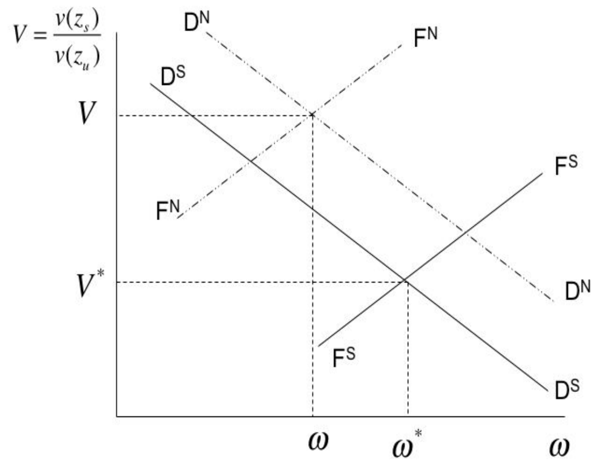


dance of skilled labor and its relative wage of skilled labor is less than in the South. Consequently, the North produces relatively more of the skill intensive good. The South produces relatively more of the unskilled labor intensive good.

Figure 2 illustrates a simple Ricardian model. Suppose that the North possesses the same factor endowments as the South but possesses TFP that is relatively higher in the skilled labor intensive sector than the unskilled labor intensive sector ($\tilde{\gamma} > 1$). If this pattern holds, the Northern goods market clearing condition $D^N D^N$ will be above and to the right of the goods market clearing condition for the South, $D^S D^S$. The fact that $D^N D^N$ lies above and to the right of $D^S D^S$ comes from the fact that for a given $\omega = \omega^*$, $V > V^*$ because TFP is systematically higher in the skilled labor intensive sector in the North. Because factor endowments are the same in each country, they share a common factor market clearing condition, FF. The North produces relatively more of the skill intensive good and the relative wage of skilled labor is bid up as resources are reallocated to the skill intensive industry.

Finally, consider a hybrid of the two models where Northern industry TFP is positively correlated with the skilled labor intensity of goods *and* the North possesses a relative abundance of skilled labor. This hybrid model is portrayed in Figure 3.

Figure 3: Hybrid Model



In this example, omitting productivity from empirical work when factor prices are unobserved will result in a substantial omitted variable bias in interpreting HO tests because the cumulative effect of factor abundance and productivity will be attributed to factor abundance. If we only observe differences in V and V^* and differences in factor abundance, we will confound the effects of high relative productivity and factor abundance when performing tests of HO because we cannot distinguish shifts in the FF curve from shifts in the DD curve.

If relative TFP is negatively correlated with skill intensity in the skill abundant country, HO mechanisms are less likely to appear in the data (i.e. the North produces a lower V than if productivity was distributed identically across industries). In the first case, the unified Ricardian-HO model provides a meaningful alternate hypothesis for a given set of production patterns and a solution to an omitted variable bias. In the second case, it allows for the possibility that HO predictions can be rescued. Finally, if TFP is uncorrelated with factor intensity, we will not expect it to affect HO predictions at all.

3 Theory: A Continuum of Industries

I now generalize my analysis to a continuum of industries as in Dornbusch, Fischer, and Samuelson (1980) and Romalis (2004). Industries with higher values of z use a more skill intensive production technique at a given set of factor prices than those with a lower z. With a continuum of industries, first tier utility (Υ) takes the form:

$$\Upsilon = \int_0^1 b(z) ln[C(z)] dz, \tag{11}$$

$b(z)$ is the exogenous Cobb-Douglas share of expenditures associated with each industry. The consumption aggregator for each industry, $C(z)$, is the same as in the simple model. For a given industry, equation 7 still characterizes the relative value of production in a diversified equilibrium such that $\tilde{R}(\tilde{p}(z)) > 0$. As before, relative revenue in an industry is declining in its relative price:

$$\frac{\partial \tilde{R}(\tilde{p}(z))}{\partial ln(\tilde{p}(z))} = \Gamma(\tilde{p}(z)) = \frac{-\sigma\tau^{1-\sigma}\left(\frac{Y^*}{Y}+1\right)\left[\tilde{p}(z)^{\sigma}\left(\tau^{2(1-\sigma)}+\frac{Y^*}{Y}\right)-2\tau^{1-\sigma}\left(\frac{Y^*}{Y}+1\right)+\tilde{p}(z)^{-\sigma}\left(\tau^{2(1-\sigma)}\frac{Y^*}{Y}+1\right)\right]}{\left[\tau^{2(1-\sigma)}+\frac{Y^*}{Y}-\tilde{p}(z)^{-\sigma}\tau^{1-\sigma}\left(\frac{Y^*}{Y}+1\right)\right]^2} < 0. \tag{12}$$

This derivative is negative when $\sigma > 1$ and iceberg transportation costs exist as in Romalis (2004).[11] Relative prices reflect TFP differences and differences in factor prices

$$\tilde{p}(z) = \tilde{\omega}^z \frac{\tilde{w_u}}{\tilde{A}(z)}. \tag{13}$$

For a given set of relative factor prices, comparative advantage can emerge both because of the interaction of relative factor prices and factor intensity ($\tilde{\omega}^z$) or because of relative differences in TFP, $\tilde{A}(z)$. Because I need to keep track of productivity in many industries, I use a convenient parameterization of productivity as follows where $\tilde{a}(z) = ln\left(\tilde{A}(z)\right)$:

$$\tilde{a}(z) = \tilde{a} + ln\left(\tilde{\gamma}\right) z + \tilde{\epsilon}_{A(z)}; \quad \tilde{\epsilon}_{A(z)} \ i.i.d.(0, \sigma^2_{\tilde{A}(z)}), \tag{14}$$

$$ln(\tilde{\gamma}) = \frac{cov[z, \tilde{a}(z)]}{var(z)}. \tag{15}$$

[11]See Appendix A for a discussion.

This conveniently breaks TFP into three components: country level differences that are neutral across industries ($\tilde{a}$), differences across industries that are correlated with factor intensity ($ln(\tilde{\gamma})z$), and differences across industries that are orthogonal to factor intensity ($\tilde{\epsilon}_{A(z)}$). Country level differences in relative productivity pose the fewest problems for HO theory in that they can easily be modeled as an increase in country size.[12] The component of Ricardian TFP that is correlated with factor intensity is captured by $ln(\tilde{\gamma})z$. $ln(\tilde{\gamma})$ is just the ordinary least squares (OLS) coefficient of a regression of $\tilde{a}(z)$ on skill intensity z under normal OLS assumptions. This poses problems for HO theory because it offers a well articulated hypothesis for why we do or do not find HO production patterns in data.

If $\tilde{\gamma} > 1$, then $cov[z, \tilde{a}(z)] > 0$ and skilled labor intensive industries will *on average* have higher TFP than unskilled labor intensive industries. If $\tilde{\gamma} < 1$, then $cov[z, \tilde{a}(z)] < 0$ and skilled labor intensive industries *on average* have lower TFP than unskilled labor intensive industries. If $\tilde{\gamma} = 1$, then $cov[z, \tilde{a}(z)] = 0$ and productivity is uncorrelated with skill intensity.

TFP that is uncorrelated with factor intensity and purged of country level effects is represented by $\tilde{\epsilon}_{A(z)}$. Because this component of TFP is orthogonal to factor intensity and purged of country effects by assumption, it is part of a model that is empirically separable from HO forces.

I exploit the monotonic relationships between $v(z)$ and $\tilde{R}(z)$ and between $\tilde{R}(z)$ and $\tilde{p}(z)$ and take a first-order linear approximation around the skill labor intensity z_0. Using the implicit function theorem, I simplify $v(z)$ as a linear function of z exploiting the Cobb-Douglas structure of relative prices:

$$v(z) = v(z_0) + \frac{\partial v(z_0)}{\partial ln(\tilde{p}(z_0))} \frac{\partial ln(\tilde{p}(z_0))}{\partial z_0}(z - z_0), \tag{16}$$

$$v(z) = v(z_0) + \frac{\tilde{R}(z_0)}{\left[1 + \tilde{R}(z_0)\right]^2}\Gamma(z_0) \ln\left(\frac{\tilde{\omega}}{\tilde{\gamma}}\right)(z - z_0). \tag{17}$$

The covariance of $v(z)$ with z gives the simple expression where $\Gamma'(z_0) = \frac{\tilde{R}(z_0)}{\left[1+\tilde{R}(z_0)\right]^2}\Gamma(z_0) < 0$:

$$cov[z, v(z)] = \Gamma'(z_0) \ln\left(\frac{\tilde{\omega}}{\tilde{\gamma}}\right) var(z). \tag{18}$$

[12]See Dornbusch, Fischer, and Samuelson (1980) for the simplest example of this.

This is the continuum of industries analog of the goods market clearing condition DD from the two industry model. Although applicable to any two factors of production, this expression shows how a given correlation between skill intensity and production can occur for two reasons. First, if productivity is uncorrelated with factor intensity ($\tilde{\gamma} = 1$), relatively cheap skilled labor ($\tilde{\omega} < 1$) can lead countries to produce more skilled labor intensive goods ($cov[v(z), z] > 0$).[13] Second, even if factor prices do not differ ($\tilde{\omega} = 1$) production can be skewed towards skill intensive industries ($cov[v(z), z] > 0$) because productivity is systematically higher in skilled labor intensive industries ($\tilde{\gamma} > 1$).

I now present the continuum of industries analog of the factor market clearing condition. The following equations are the factor market clearing conditions for the North in skilled and unskilled labor,

$$\int_0^1 b(z)v(z)zY^w dz = w_s S \quad \int_0^1 b(z)v(z)(1-z)Y^w dz = w_u U.$$

Simple division yields

$$\frac{\int_0^1 b(z)zv(z)dz}{\int_0^1 b(z)(1-z)v(z)dz} = \frac{w_s S}{w_u U}. \tag{19}$$

Dividing this by the Southern analog condition gives the following *factor market clearing condition*:

$$\frac{\int_0^1 b(z)zv(z)dz}{\int_0^1 b(z)z(1-v(z))dz} \frac{\int_0^1 b(z)(1-z)(1-v(z))dz}{\int_0^1 b(z)(1-z)v(z)dz} = \tilde{\omega}\frac{\tilde{S}}{\tilde{U}}. \tag{20}$$

Proposition 1 states that when Ricardian productivity differences are uncorrelated with factor intensity, HO forces should be present and should contribute to the relative production structures of the two countries.

3.1 Separability between HO and Ricardian models.

Proposition 1: If productivity is uncorrelated with factor intensity and the relative abundance of factors differs among countries, then the relative wage of a country's abundant factor will be less than in the country where it is a relatively scarce factor. In addition, $cov[v(z), z] > 0$ where z is

[13] Recall that $\Gamma' < 0$.

the Cobb-Douglas cost share of its relatively abundant factor and $cov[v(z'), z'] < 0$ where z' is the Cobb-Douglas cost share of its relatively scarce factor.

Proof: See Appendix C.

This proposition shows that if TFP is uncorrelated with factor intensity, then basic HO results should hold in the data. Intuitively, when relative TFP is uncorrelated with factor intensity, differences in TFP across industries will not cause (nor prevent) empirical tests of Heckscher-Ohlin to find evidence of factor abundance based production and trade. Consequently, differences in the production structure coming from differences in factor abundance (e.g. Rybczynski regressions) or the net exporting position of a given factor (e.g. HOV tests) are unlikely to be affected by differences in relative TFP across industries if TFP is uncorrelated with factor intensity.[14]

When TFP is correlated with factor intensity, any reduced-form relationship between factor intensity, factor abundance and production will likely be due to both factor abundance and Ricardian TFP. It is also possible that relative Ricardian TFP differences will be large enough that a country that possesses a relative abundance of a factor will *not* produce relatively more of the good that uses that factor relatively intensively. For example, the South might have TFP that is systematically high enough in skill intensive industries that it will produce relatively more skilled labor intensive goods than the North. Intuitively, this is most likely to occur when differences in factor abundance are very small and/or differences in γ are very large.

3.2 Empirical Application

I now derive two expressions that test for the contributions of Ricardian and HO forces in production data. I first derive a "restricted expression" that tests whether the relationship between factor intensity, factor abundance and production can be explained by HO and/or Ricardian forces. Unfortunately, it says nothing about the role of Ricardian productivity that is uncorrelated with factor intensity. To assess the role of productivity that is uncorrelated with factor intensity, I then derive an "unrestricted expression."

To derive the restricted expression, I log-linearize the expression for relative revenue in industry z (equation 7) as a function of $ln\,(\tilde{p}(z))$ with the appropriate subscripts for country c relative to

[14] This should not be mistaken for how countrywide differences in TFP affect HOV predictions as in Trefler (1995).

c'.[15] I then take the covariance of this expression with z:

$$cov[z, \tilde{r}(z)_{cc't}] = \Gamma \ln\left(\frac{\tilde{\omega}_{cc't}}{\tilde{\gamma}_{cc't}}\right) var(z). \tag{21}$$

Because sufficiently comparable international factor prices are unavailable and likely to be endogenous themselves, I exploit the fact that equilibrium factor prices are likely to be related to both relative factor abundance (which will lead to lower relative wages of the relatively abundant factor *ceteris paribus*) and the industry structure of TFP (which will lead to higher relative wages for factors used relatively intensively in high relative TFP sectors *ceteris paribus*). For simplicity, I represent this relationship by $ln\left(\tilde{\omega}\right) = \kappa_0 ln\left(\tilde{S}/\tilde{U}\right) + \kappa_1 ln\left(\tilde{\gamma}\right)$ where $\kappa_0 < 0$ and $\kappa_1 > 0$.[16] This allows for equilibrium wages to be determined by both factor abundance and by the industrial structure of productivity differences. This allows for the following expression:

$$cov[z, \tilde{r}(z)_{cc't}] = \kappa_0 \Gamma \ln\left(\frac{\tilde{S}}{\tilde{U}}\right)_{cc't} var(z) - \Gamma\left(1-\kappa_1\right) \ln\left(\tilde{\gamma}_{cc't}\right) var(z). \tag{22}$$

This expression decomposes the covariance of production with skill intensity into that due to factor abundance and that due to Ricardian productivity differences. This expression can then be taken to the data using the following estimation equation where a vector of time fixed effects T allows the results to be invariant to the choice of numeraire:[17]

$$cov[z, r(z)_{ct}] = \beta_0 + \beta_1 \ln\left(\frac{S}{U}\right)_{ct} + \beta_2 \ln\left(\gamma\right)_{ct} + \beta_t' T_t + \zeta_{ct} \tag{23}$$

where $\beta_1 = \kappa_0 \Gamma var(z) > 0$ and $\beta_2 = -\Gamma(1-\kappa_1) var(z) > 0$ if $0 < \kappa_1 < 1$.

This is the "restricted expression." Under the null hypothesis that HO *alone* is responsible for

[15] The use of log revenue and not market share more easily and transparently controls for country and industry fixed effects using country-time and industry-time fixed effects and allows easier interpretation of the regression coefficients.

[16] A similar strategy is taken by Romalis (2004) in his quasi-Rybczynski regressions. Generally, however, the relationship between relative factor prices, endowments, and relative productivity differences is likely to be complex and non-log-linear. In addition, it is well-known (e.g. Jones,1965) that the relative size of a country will play a major role in how autarky factor prices relate to factor prices in a trading equilibrium. I have experimented with the expression $ln\left(\tilde{\omega}\right) = \kappa_0 ln\left(\tilde{S}/\tilde{U}\right) + \kappa_1 ln\left(\tilde{\gamma}\right) + \kappa_2 ln\left(\tilde{S}/\tilde{U}\right) ln\left(G\tilde{D}P\right) + \kappa_3 ln\left(\tilde{\gamma}\right) ln\left(G\tilde{D}P\right) + \kappa_4 ln\left(G\tilde{D}P\right)$ and the results are unchanged. See Appendix D for a deeper discussion on representing unobserved equilibrium factor prices in an empirical framework.

[17] This can be seen by rewriting $cov[z, \tilde{r}(z)_{cc't}]$ as $cov[z, r(z)_{ct}] - cov[z, r(z)_{c't}]$ and $\ln\left(\frac{\tilde{S}}{\tilde{U}}\right)_{cc't}$ as $\ln\left(\frac{S}{U}\right)_{ct} - \ln\left(\frac{S}{U}\right)_{c't}$.

any relationship between factor intensity, factor abundance and production, $\beta_1 > 0$ and $\beta_2 = 0$. Under the null hypothesis that there are no HO forces at work and that any differences in production are due to differences in Ricardian TFP, $\beta_1 = 0$ and $\beta_2 > 0$. If both HO and Ricardian effects explain why specialization occurs, then $\beta_1 > 0$ and $\beta_2 > 0$.

To examine the contribution of TFP that is uncorrelated with factor intensity, I derive the "unrestricted expression" by log-linearizing equation 7 where the linearization occurs at z_0 such that $\tilde{p}(z_0) = 1$:[18]

$$\tilde{r}(z) = \tilde{r}(z_0) + \frac{\partial \tilde{r}(z_0)}{\partial ln(\tilde{p}(z_0))} ln(\tilde{p}(z)). \tag{24}$$

Breaking $ln(\tilde{p}(z))$ into its Cobb-Douglas components gives

$$\tilde{r}(z) = \tilde{r}(z_0) - \frac{\partial \tilde{r}(z_0)}{\partial ln(\tilde{p}(z_0))} \left[\tilde{a} - \tilde{w}_u - ln(\tilde{\omega})z + ln(\tilde{\gamma})z + \epsilon_{\tilde{A}(z)}\right] \tag{25}$$

Revenue depends on country and industry level variables as might be expected. Revenue is increasing in country level productivity ($\tilde{a}$), decreasing in the absolute wage level ($\tilde{w}_u$) and increasing in industry specific relative productivity ($\tilde{\epsilon}_{A(z)}$).[19] If the North possesses relatively cheap skilled labor ($\ln(\tilde{\omega}) < 0$), then relative revenue is systematically increasing in z. If the North has systematically higher relative productivity in skill intensive industries ($\ln(\tilde{\gamma}) > 0$), then relative revenue is also systematically increasing in z. Including fixed effects that make the results insensitive to the choice of numeraire country gives the following expression where ZT is a vector of industry-time fixed effects (e.g. Industry 311 in 1990), CT is a vector of country-time fixed effects (e.g. Japan in 1990), and ζ is an error term that is clustered by country-industry (e.g. Industry 311 in Japan):

$$\tilde{r}(z)_{ct} = \frac{\partial \tilde{r}(z_0)}{\partial ln(\tilde{p}(z_0))} \left[ln(\omega)_{ct}z - ln(\gamma)_{ct}z - \epsilon_{A(z),ct}\right] + \beta_{zt}' ZT_{zt} + \beta_{ct}' CT_{ct} + \zeta_{zct} \tag{26}$$

Again, assuming that the elasticity of relative factor prices with respect to relative endowments and patterns of industry TFP can be expressed as in the "restricted" expression, I can specify the following expression:

[18] Taking the linearization around other relative prices does not affect the result.

[19] Recall that the derivative outside the brackets is negative.

$$\tilde{r}(z)_{ct} = \beta_0 ln\left(\frac{S}{U}\right)_{ct} z + \beta_1 ln(\gamma)_{ct} z + \beta_2 \epsilon_{A(z),ct} + \beta'_{zt} ZT_{zt} + \beta'_{ct} CT_{ct} + \zeta_{zct} \tag{27}$$

where $\beta_0 = \kappa_0 \Gamma > 0$, $\beta_1 = -\Gamma(1-\kappa_1) > 0$, and $\beta_2 = -\Gamma > 0$.

This is the "unrestricted expression." As before, β_0 gauges the validity of the HO models and β_1 assesses the importance of TFP that is correlated with skill intensity. β_2 assesses the importance of Ricardian productivity that is orthogonal to factor intensity in determining production patterns. All country level differences in absolute wages and productivity that are identical across industries in a year are absorbed into the country-time fixed effects. All industry-time characteristics (e.g. average scale of industry) will be absorbed by the industry-time fixed effects.

4 Data

This section outlines the data and variables used to estimate the model. The collected data set covers 24 3 digit ISIC revision 2 industries, 11 years (1985-1995), and the following 20 countries: Austria, Canada, Denmark, Egypt, Finland, Great Britain, Hong Kong, Hungary, Indonesia, India, Ireland, Italy, Japan, Norway, Pakistan, Portugal, South Korea, Spain, Sweden, and the United States. All variables (except those explicitly mentioned) are taken from the World Bank's Trade and Production data set (Nicita and Olarreaga, 2001). All country-years for which complete data exist for at least 15 of the 24 industries in that country and year are kept.[20] Because not all countries have available data in all years, the dataset is an unbalanced panel. The Data Appendix lists the data availability for years and countries. The most binding constraint in assembling this data set is the availability of a continuous time series for investment necessary for creation of capital stock.

[20] There are 28 three digit ISIC manufacturing industries in the Trade and Production dataset. Four industries are excluded from the analysis: 314 (tobacco), 353 (petroleum refineries), 354 (misc. petroleum and coal production), 390 (other manufactures). The first three are excluded because their production values are likely to be substantially influenced by international differences in commodity taxation (Fitzgerald and Hallak, 2004). The last is excluded because its "bag" status makes comparability across countries difficult. All results are invariant to increasing the cutoff to having 18 of the 24 industries although the sample size and power of the empirical tests are obviously smaller.

4.1 Factor Abundance

Although the model is applicable to any set of factors of production, I focus on skilled and unskilled labor as imperfectly substitutable factors of production.[21] As a measure of $\frac{S}{U}$, I examine the ratio of the population that has obtained a tertiary degree to that which does not as found in the Barro and Lee (2001) educational attainment dataset.[22] As might be expected, Canada and the United States have the highest (average) values with $\frac{S}{U} = 0.76$ for the United States and $\frac{S}{U} = 0.70$ for Canada. Pakistan and Indonesia have the lowest (average) values with $\frac{S}{U} = 0.02$ for both countries.

4.2 Skilled Labor Intensity of Industries

Data on the skilled labor cost share (z) for each of the 24 industries come from educational attainment data by worker in the United States Current Population Survey (CPS) dataset where workers are transformed into effective workers using a Mincerian wage regression. The Data Appendix explains the procedure in detail. I examine narrow and broad definitions of skilled labor. The "narrow" definition defines a skilled laborer as a worker with four or more years of college. The "broad" definition defines a skilled laborer as one who has attended any college. Table 1 presents these measures of z.[23] These measures line up with common priors. Among the most skill intensive industries are Scientific Equipment (385), Industrial and Other Chemicals (351,352), and Publishing (342). Among the least skill intensive industries are Textiles (321), Footwear (324) and Wearing Apparel (322).[24] I loosen the assumption of a constant z across countries in a given industry in the robustness section (Section 6).

[21] I select skilled and unskilled labor as the factors of production in this model for two reasons. First, recent work (e.g. Fitzgerald and Hallak (2004)) has shown that skilled and unskilled labor possess more explanatory power in differences in the structure of production than capital. Second, data on skilled labor abundance (as measured by educational attainment rates in Barro and Lee (2001)) is far more comprehensive than the Penn World Tables coverage of capital per worker.

[22] Data are only available at five year intervals. Data for the interim years are interpolated assuming that the growth rate of the variable is constant over the five years. No extrapolations are performed. Results using a broader definition of skilled labor are examined in the robustness section. If $\frac{S}{U} = 1$, then equal proportions of the population are skilled and unskilled.

[23] I assume that z is constant across countries. Similar theoretical results can be derived for CES production functions if skilled and unskilled labor are more substitutable than the Cobb-Douglas case.

[24] UNGISD data on operatives and non-operatives are commonly used to distinguish skilled and unskilled workers within a given country as in Berman, Bound and Machin (1998). However, using it to compare skilled and unskilled workers across countries is highly dubious. For example, the ratio of non-operatives (commonly thought to be "skilled") to operatives (commonly though to be "unskilled") is 0.21 in Indonesia, 0.38 in the United States, 0.85 in Japan, and 0.45 in Italy (U.N., 1995).

Table 1: Industry Skill Intensities

ISIC Code	Description	z_{narrow}	z_{broad}	ISIC Code	Description	z_{narrow}	z_{broad}
311	Food	0.16	0.36	355	Rubber Prod.	0.19	0.44
313	Beverages	0.35	0.57	356	Plastic Prod.	0.19	0.39
321	Textiles	0.13	0.28	361	Pottery, China etc.	0.21	0.50
322	Wearing Apparel	0.10	0.24	362	Glass and Prod.	0.18	0.41
323	Leather Prod.	0.12	0.31	369	Non-Metallic Mineral Prod.	0.20	0.37
324	Footwear	0.16	0.28	371	Iron and Steel	0.15	0.38
331	Wood Prod.	0.13	0.32	372	Non-ferrous Metals	0.19	0.41
332	Furniture	0.13	0.30	381	Fabricated Metal Prod.	0.18	0.40
341	Paper and Prod.	0.21	0.44	382	Machinery (non-elec)	0.20	0.47
342	Printing and Publishing	0.36	0.61	383	Elec. Machinery	0.36	0.60
351	Industry Chemicals	0.42	0.67	384	Transport Equip.	0.29	0.55
352	Other Chemicals	0.45	0.65	385	Prof. and Sci. Equip.	0.37	0.61

Table 2: $Cov[z, r(z)]$ Summary Stats (182 observations)

Measure of z	Mean	Std. Dev	Max	Min
Narrow	0.0284	0.0171	0.0556	-0.0198
Broad	0.0348	0.0246	0.0836	-0.0345

4.3 Production Covariances

I calculate the covariance of (log) revenue with the skill intensity of the industries ($cov[z, r(z)]$) using production value from the Trade and Production dataset. Table 2 presents summary statistics for $cov[z, r(z)]$ based on both the narrow and broad definitions of skill intensity.[25]

[25] These measures are in line with measures from other studies. For example, Fitzgerald and Hallak (2004) use a slightly different measure of skilled labor and examine production in OECD countries. Using their data (table 7), I find that the country that is in the 25th percentile of skilled labor abundance has a covariance of 0.0377 and a country that is in the 75th percentile has a covariance of 0.0698. The values for the 25th and 75th percentile (using the broad definition) for my sample are 0.0206 and 0.0605, respectively. Appendix G contains a list of average $cov[r(z), z]$ by country.

4.4 Factors of Production and TFP

I follow Caves, Christensen and Diewert (1983) and Harrigan (1997) in using the solution to an index number problem to calculate relative productivity levels.[26] This methodology is based on a translog functional form that allows the productivity calculation to be based on any production function up to a second order approximation. Based on this procedure, if capital (K) and homogenous labor (L) are used to produce value added (VA), the TFP productivity level between country a and a multilateral numeraire would be $TFP(z)_{a,t} = \frac{VA(z)_{a,t}}{\overline{VA(z)_t}} \left(\frac{\overline{K(z)_t}}{K(z)_{a,t}}\right)^{\frac{\alpha_{K,a}+\alpha_{K,avg}}{2}} \left(\frac{\overline{L(z)_t}}{L(z)_{a,t}}\right)^{\frac{\alpha_{L,a}+\alpha_{L,avg}}{2}}$, $\alpha_{i,j}$ represents the Cobb-Douglas revenue share of factor i in country j and $\alpha_{i,avg}$ is the average revenue share of factor i across all countries in the given industry.[27] $\overline{K}(z)_t = \frac{1}{N_{z,t}} \sum_c K(z)_{czt}$ and $\overline{L}(z)_t = \frac{1}{N_{z,t}} \sum_c L(z)_{czt}$ where $N_{z,t}$ is the number of countries in the sample in industry z in year t.

4.4.1 Deflators

Very few industry level deflators exist that allow comparison of output or value added across countries. For this reason, I use the *disaggregated* PPP benchmark data provided by the Penn World Tables to deflate the data. These price indexes allow PPP price comparisons across goods and countries and are constructed with an explicit eye toward comparing goods of similar quality. The Data Appendix addresses this in detail.[28]

[26]Basu and Kimball (1997) measure TFP growth taking into account the endogeneity of factor demand and unobservable factor utilization. Unfortunately, there is a lack in this context of strong demand shifting instruments to control for the endogeneity of factor demand across countries. I consider the issue of capacity utilization in the robustness section. I do not to use estimators that are often used in the firm level literature (e.g. Olley and Pakes (1996)) because the assumptions that legitimize their use do not hold in industry-country level analysis. Required assumptions include that all "firms" possess the same demand function for investment or intermediate inputs and the same exogenous factor prices. The assumption that market structure and factor prices are the same across countries is highly questionable.

[27]Because year-to-year measured revenue shares are extremely noisy, I follow Basu, Fernald, and Kimball (2006) and constrain α within a country within an industry (e.g. Indonesia-311) with no time series variation. Labor's factor share of value added is calculated as wages' proportion of value added. Capital's share of value added is one minus labor's share. Observations where the factor share of any input is negative are dropped.

[28]Country level PPP price deflators are incorrect because of the weight that they assign to non-traded goods which leads to a greater dispersion in price indexes than occurs in manufacturing which is highly traded. In addition, any country level output deflators will be differenced out by the country-year fixed effects. See Kravis, Heston and Summers (1982) for a thorough discussion of the process behind the collecting of the data and the preparation of the price indexes that are behind this study and the Penn World Tables. Further, country averages only capture 35% of the variance of relative prices across countries and industries in the disaggregated PWT data. This suggests that using country level price deflators will not capture substantial within-country variation.

4.4.2 Labor and Capital Input

In measuring TFP, I consider differences in the effectiveness of labor across countries because it is not proper to interpret differences in the effectiveness of labor as differences in total factor productivity. Differences in the effectiveness of labor can be modeled as unmeasured differences in the abundance of labor and can be easily written into an HO model. Following Bils and Klenow (2002) and Caselli (2005), I calculate the effectiveness of labor using wage premium and educational attainment data.

Define E as the effectiveness of labor per worker so that EL is the effective labor input. Using the Barro and Lee data on average years of schooling, I normalize the effectiveness of labor with "no schooling" (0 years) to be $E = 1$. Following Caselli (2005), I assume that labor becomes 13% more effective per year for the first four years of schooling, 10% per year for years 4-8, and 7% per year after that. Because the evolution of the skill level of labor in a country is likely to be slow, I use average years of schooling in 1990 for these calculations. The Data Appendix presents measures of E based on this methodology. These measures line up with commonly held priors with Canada, Norway, and the United States possessing the highest levels of the effectiveness of labor and Egypt, India, and Pakistan having the lowest.

Unlike work such as Harrigan (1999) and Keller (2002), I do not consider differences in days or hours worked. Practically, hours worked data that is sufficiently comparable across industries and countries are not available. Harrigan (1999) and Keller (2002) sidestep this issue by imposing measures of hours worked in *aggregate manufacturing* on all sectors within manufacturing. My interest in cross-industry TFP comparisons allows me not to include these measures. This is because hours of labor input will be highly correlated with (if not identical to) hours of capital service. If the value added function is constant returns to scale, then it will also be homogenous of degree one in hours worked. If I use the same measure of hours worked in manufacturing across all manufacturing industries, I will multiply each production function in a given country-year group by the same scalar which will then be differenced out by a country-year fixed effect as derived in section 3.2.

Labor is decomposed into operatives (U) and non-operatives (S) using data from the United

Nations General Industrial Statistical Database.[29] The effectiveness of labor is assumed to augment both operatives and non-operatives. Capital is calculated using the perpetual inventory method.[30] The final (value added) measure of productivity between country a and the multilateral numeraire is then

$$TFP_{a,t} = \frac{VA_{a,t}}{\overline{VA}_t} \left(\frac{\overline{S_t E_t}}{S_{a,t} E_{a,t}} \right)^{\frac{\alpha_{S,a} + \alpha_{S,avg}}{2}} \left(\frac{\overline{U_t E_t}}{U_{a,t} E_{a,t}} \right)^{\frac{\alpha_{U,a} + \alpha_{U,avg}}{2}} \left(\frac{\overline{K}_t}{K_{a,t}} \right)^{\frac{\alpha_{K,a} + \alpha_{K,avg}}{2}} . \tag{28}$$

4.4.3 Plausibility of TFP Measures

Because of the importance of TFP measures in this paper, I check their plausibility. First, I compare my measures to those calculated by others for consistency. Second, I check the correlation of relative TFP across industries with relative revenue. If Ricardo's original insight is fundamentally true, this correlation should be positive and of a reasonable magnitude. Last, I check how much these measures fluctuate over time because large fluctuations would suggest substantial noise in my calculations. My measures meet all of these criteria for desirability.

First, Table 3 presents my estimates of industry level productivity against similar measures calculated by Harrigan (1999) (Table 1). I compare *all* industries and countries for which our calculations of TFP overlap. I calculate TFP of industries 382 (Machinery, non-electric), 383 (Machinery, Electric), 384 (Transport equipment), and 385 (Professional and Scientific Equipment). I then calculate them relative to the average across these four industries and then relative to the United States in that industry and then compare these to similar measures from Harrigan (1999).[31]

Our measures of relative TFP line up broadly. The rank correlation between the two measures

[29] Comprehensive data on operatives and non-operatives are not available from year to year for my broad sample. For this reason, I calculate the average proportions of employment that are operatives and non-operatives for each country-industry. Using the available data, these average measures capture 95% of the year to year variation in a fixed effects regression. I then apply these constant proportions to annual employment data from the Trade and Production dataset to create annual measures of operatives and non-operatives. I follow a similar procedure to decompose wages into those paid to operatives and non-operatives to calculate the measures α_S and α_U.

[30] See the Data Appendix for more details.

[31] I compare my measure for ISIC 382 to Harrigan's (1999) "Non-electrical machinery", ISIC 383 to his "Electrical machinery", ISIC 384 to his "Motor vehicles" and 385 to his "Radio, TV, & communications Equip." He also calculates TFP for "Office and Computing Equipment" and "Aircraft" but my industrial classification does not allow for easy comparison of these industries. In addition, he also calculates TFP for Australia, Germany and the Netherlands, none of which I calculate because of data constraints. Finally, I drop his 1988 measure for Motor Vehicles in Italy which increases four-fold from the previous year in his measures and is unlikely to reflect true TFP as opposed to short run economic fluctuations.

Table 3: Comparing Relative TFP Measures

					Harrigan (1999)			
		ISIC Industry				ISIC Industry		
Country	382	383	384	385	382	383	384	385
Canada	0.980	1.005	0.681	0.335	0.952	1.663	0.688	0.755
Finland	1.009	0.934	0.244	0.334	0.981	2.325	0.468	0.945
Great Britain	0.613	1.116	0.425	0.321	0.596	1.361	0.294	0.655
Italy	0.782	0.904	0.377	0.221	0.760	1.807	0.477	0.673
Japan	0.861	0.818	0.586	0.287	0.837	1.711	0.826	0.691
Norway	0.811	0.621	0.277	0.326	0.788	1.771	0.339	0.664
United States	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Rank correlation between constructed measures and those of Harrigan (1999): 0.74

based on the 24 observations is 0.74.[32] In addition, although not presented, selected industrial levels line up with other work. For example, Japan is the world leader in TFP for Iron and Steel (371) and Non-Ferrous Metals (372) which is consistent with Dollar and Wolff (1993) and Harrigan (1997). One discrepancy between the calculations here and those of Harrigan (1999) is the lower average TFP level in scientific and professional equipment (ISIC 385) that I calculate relative to his calculations. However, some consolation should be taken from the fact that both calculations find that the United States, Canada and Finland are among the most productive while Italy and Great Britain are among the least productive.

Second, I examine $\frac{cov[\tilde{a}(z),\tilde{r}(z)]}{var(\tilde{a}(z))}$ to gauge the explanatory power of productivity across industries. I calculate this measure for two reasons: first, this statistic should be positive for any non-pathological model. Second, it can be shown that this number should be equal to $-\Gamma$ as defined in equation 12. For 182 observations, each indexed by country-year, the mean is 0.3572 and significantly different from zero at the 1% level of certainty ($t = 3.24$).[33] Because this is a reduced form combination of structural parameters, it is difficult to interpret. However, Anderson and van Wincoop (2004) estimate that international trade barriers impose a total of a 74% ad valorem tax equivalent. I set $Y^*/Y = 30$ based on the fact that the average country in the sample possesses

[32] This excludes the values for the United States that are set equal to 1.00 for normalization.
[33] With standard errors clustered by country.

approximately 3.3% of world GDP. Applying these numbers to the expression for Γ evaluated at $\tilde{p}(z) = 1$ implies a value of $\sigma = 8$. Although this is in the upper range of past estimates for σ, it is within reason.[34]

Third, the measures of total factor productivity are also relatively stable over time. A regression of $\tilde{a}(z)_{ct}$ on a set of country-industry fixed effects (e.g. Indonesia, ISIC-311) explains 91% of its variance. Therefore, although these measures almost surely capture some business cycle fluctuations, this variance is dominated by the larger differences that exist across countries and industries rather than fluctuations over time within a country and industry.

The covariance terms (γ) are then calculated using the skill labor shares (z) and TFP and equation 15 such that $\gamma_{ct} = exp\left[\frac{cov[z,a_{czt}]}{var(z)}\right]$ where $cov[a_{czt}, z]$ is an unweighted covariance of (log) productivity with z. This differences out all country specific effects (e.g. country level business cycles).

Although all comparisons of TFP across countries, industries and time are subject to some difficulties in measurement, the measures presented here are very likely to reflect real differences in TFP based on similarity to previous studies, the positive correlation of productivity and revenue across industries within a country in a given year, and the stability of the estimates over time.

5 Results

I present two sets of results. First, I present a "restricted" version of the model where the dependent variable is $cov[z, r(z)_{ct}]$. This expression asks to what degree a country specializes in the production of skill intensive goods due to HO and Ricardian effects. These results appear in Table 4. Second, I present "unrestricted" results where the dependent variable is $r(z)_{ct}$. This gauges the determinants of revenue industry-by-industry instead of based on country level covariances. It also gauges the importance of TFP that is uncorrelated with skill intensity. These results appear in Table 5. Extensive robustness checks are presented in Section 6.

[34] Broda and Weinstein (2006) estimate σ for 256 industries and find that the 5th and 95th percentiles of the distribution are 1.2 and 9.4, respectively.

Table 4: Restricted Regression

	Narrow z		Broad z	
Variables	(1)	(2)	(3)	(4)
$ln(S/U)_T$	0.0109***	0.0115***	0.0174***	0.0175***
	(0.0033)	(0.0025)	(0.0041)	(0.0036)
$ln(\gamma)$		0.0031		0.0033
		(0.0018)		(0.0028)
Obs	182	182	182	182
Time Fixed Effects	yes	yes	yes	yes
R^2	0.3477	0.3996	0.4230	0.4411

*** estimated at the 1% level, ** estimated at the 5% level. Robust standard errors clustered by country. Observations indexed by country-year. Equation (29) gives the estimation equation for this table. Dependent Variable: $cov[r(z), z]$

5.1 Results: Restricted

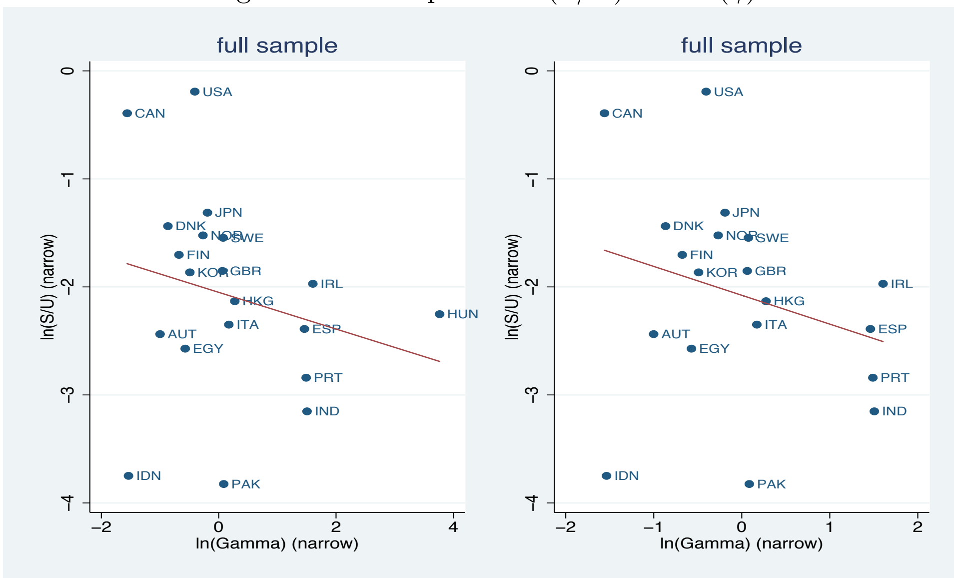
Recall that the "restricted" regression equation is:

$$cov[z, r(z)_{ct}] = \beta_0 + \beta_1 \ln \left(\frac{S}{U} \right)_{ct} + \beta_2 \ln (\gamma)_{ct} + \beta_t' T_t + \zeta_{zct} \tag{29}$$

Column (1) of Table 5 tests the hypothesis that the abundance of skilled labor as measured by the proportion of workers with a tertiary education or higher ($ln(\frac{S}{U})_T$) predicts how skewed productive resources are towards relatively skill intensive industries ($cov[z, r(z)]$). Column (2) includes $ln(\gamma)$ to assess the importance of productivity that is correlated with skill intensity. Columns (3)-(4) perform the same regressions using the broad definition of skilled labor intensity. Robust standard errors are clustered by country and presented in parentheses.

I highlight three results. *First*, each column contains the familiar HO result that countries with a relative abundance of skilled labor produce relatively more skilled intensive goods. As before, because the coefficients are reduced form combinations of structural parameters, it is impossible to identify any of these structural parameters. However, I can gauge their plausibility. For example, the estimate from Column 1 implies an elasticity of substitution (σ) of 5.1 using the algorithm in

Figure 4: Scatterplot of $ln(S/U)$ and $ln(\gamma)$



Section 4.[35]

Second, the inclusion of $ln(\gamma)$ does not substantively change the coefficient on $ln(S/U)$. This suggests that skill abundant (or scarce) countries do not have productivity that is systematically higher in skill intensive (or unskilled intensive) industries. *Third*, the coefficient on $ln(\gamma)$ is statistically indistinguishable from zero. This suggests that Ricardian productivity is relatively uncorrelated with skill intensity. This is confirmed by regressing $ln(\gamma)$ on $ln(S/U)$ which yields a coefficient of -0.1935 with a robust standard error of 0.3494 with clustering by country and inclusion of time fixed effects to control for each annual numeraire. Figure 4 presents scatterplots that present the same information graphically. Because the observation for Hungary is an outlier in the left hand panel, it is excluded in the right hand panel with the same qualitative results. As a whole, these results suggest that TFP that is correlated with factor intensity is unlikely to bias HO results.

[35]This can be calculated by evaluating the expression for $-\kappa_0\Gamma$ at $\tilde{p}(z) = 1$, assigning $\tau = 1.74$, $Y^*/Y = 30$, assigning a value of $-\kappa_0 = 0.95$, and noting that var(z)=0.0106 from table 1, and solving for the σ that is consistent with the coefficient. See Appendix D for a discussion of the decision to set $-\kappa_0 = 0.95$.

5.2 Results: Unrestricted

To examine how Ricardian productivity influences patterns of specialization when relative TFP is uncorrelated with factor intensity, I estimate the "unrestricted" expression where observations are indexed by country-industry-year as below:

$$r(z)_{ct} = \beta_0 ln\left(\frac{S}{U}\right)_{ct} z + \beta_1 ln(\gamma)_{ct} z + \beta_2 \epsilon_{A(z),ct} + \beta'_{zt} ZT_{zt} + \beta'_{ct} CT_{ct} + \zeta_{zct}. \tag{30}$$

A vector of industry-time fixed effects (ZT) controls for all numeraires and a vector of country-time fixed effects (CT) controls for all country-time effects such as aggregate TFP. Recall that $\epsilon_{A(z)}$ is the component of TFP that is uncorrelated with factor intensity and is purged of country averages.

Examining Table 5, I highlight three results. *First*, the coefficient on relative factor abundance is still positive and significant and does not change significantly when productivity measures are included in the regression. These coefficients appear to be larger than those in the restricted regressions. However, in the restricted regressions, $\beta_0 = \kappa_0 \Gamma var(z)$ but in the unrestricted regression $\beta_0 = \kappa_0 \Gamma$. Dividing the coefficient on $zln(S/U)_T$ from Table 4, Column 1 by the variance of z from Table 1 gives a value of 1.0254 which is extremely close to its counterpart in Column 1 of the unrestricted regression (1.0870). *Second*, the inclusion of $ln(\gamma)z$ adds very little explanatory power in terms of its significance and effect on the coefficient on $zln(S/U)_T$.

Third, the residual productivity term, $\epsilon_{A(z)}$, is estimated precisely at the 1% level of certainty, is the expected sign, and changes little over different specifications. Following the algorithm in section 5.1, the coefficients in Column 3 imply a value of $\sigma = 5$ if calculated off of the coefficient on $ln(S/U)z$ or $\sigma = 8.3$ if it is identified off $\epsilon_{A(z),ct}$. Again, either is reasonable based on prior estimates.

This confirms previous findings that Ricardian productivity possesses explanatory power in explaining relative production patterns. However, it offers the new contribution that Ricardian productivity is uncorrelated with factor intensity and explains very little (if any) of why HO results do or do not appear. The forces that determine comparative advantage in the HO model seem to be orthogonal to those that determine comparative advantage in the Ricardian model in my sample. Therefore, it is a reasonable approximation to consider the two models as being separable

Table 5: Extended Regression

	Narrow z			Broad z		
Variable	(1)	(2)	(3)	(4)	(5)	(6)
$zln(S/U)_T$	1.0870***		1.2040***	1.0720***		1.0815***
	(0.3817)		(0.3895)	(0.3060)		(0.2980)
$zln(\gamma)$		0.2074	0.3293		0.2915	0.3179
		(0.2194)	(0.2355)		(0.2375)	(0.2354)
$\epsilon_{A(z),ct}$		0.2966***	0.3079***		0.2999***	0.3001***
		(0.0997)	(0.0984)		(0.0997)	(0.0988)
Obs	4063	4063	4063	4063	4063	4063
Country-Time FE	yes	yes	yes	yes	yes	yes
Industry-Time FE	yes	yes	yes	yes	yes	yes
R^2	0.8883	0.8888	0.8917	0.8895	0.8889	0.8926

*** estimated at the 1% level, ** estimated at the 5% level. Robust standard errors clustered by country-industry. Observations indexed by country-industry-year. Equation 30 gives the estimation equation for this table. Dependent Variable: $r(z)$

and equally valid when contributing to production patterns as stated in Proposition 1.

Table 6 presents standardized coefficients to assess the relative strength of these forces in determining production patterns. It shows that a one standard deviation increase in relative factor abundance is approximately 1.6 (0.1952/0.1223) to 2.4 (0.2806/0.1192) times as potent as a one standard deviation in Ricardian productivity in a given industry. Consequently a one standard deviation change in relative factor endowments is more potent than a one standard deviation change in industry level relative TFP in determining production patterns.

6 Robustness Checks

I explore the robustness of these results in seven ways in Tables 7 and 8. *First*, I use a simple IV regression to consider the role of classical measurement error in the productivity measures. *Second*, I drop countries and years for which exchange rate volatility might induce measurement error. *Third*, I take capacity utilization into account in my TFP calculations. *Fouth*, I show that these results

Table 6: Unrestricted Regression (Standardized Coefficients)

Variable	Narrow z		Broad z	
$zln(S/U)_T$	0.1762	0.1952	0.2782	0.2806
$zln(\gamma)$		0.0548		0.0754
$\epsilon_{A(z),ct}$		0.1223		0.1192

are robust to a broader measure of skilled labor abundance. *Fifth*, I also show that the dynamic correlation of the error term is sufficiently accounted for by standard clustering of the error terms. *Sixth* and *Seventh*, I show that the results are not sensitive to replacing the Cobb-Douglas cost shares with the skill rank of the cost shares both in the U.S. and in each country. I find that these seven factors affect the results very little, if at all. For parsimony, all robustness checks (except that for secondary educational attainment) use the "narrow" definition of skill intensity although the results do not change substantively when the "broad" measure is used.

Table 7, Column 1 starts by using the one year lagged values of $a(z)_{ct}$ and $ln(\gamma_{ct})z$ as instruments for their current values to gauge the importance of classical measurement error in the TFP measures. The estimated coefficient on $\epsilon_{A(z),ct}$ changes very little from the baseline result suggesting that classical measurement error does not play an important role in the baseline results.

Because some countries are vulnerable to large exchange rate movements, this can induce substantial measurement error in measures of inputs (i.e. capital) that will not be differenced out the TFP measures using country-year fixed effects. Columns 2 and 3 drop all country-year observations in which a country experienced a 20% appreciation or depreciation of their nominal exchange rate in the prior twelve months.[36] The results are unchanged.

Basu (1996) shows that incorporating capacity utilization can reduce the spurious correlation between output and "productivity" at the business cycle frequency. It is less obvious that it should matter in this context where the cross section is the primary dimension of identification. I use the

[36] All monthly exchange rate data is from IMF's International Financial Statistics Database $ifs.apdi.net/imf/logon.aspx$.

Table 7: Robustness Check I

Variable	IV	Exchange Rate		Utilization	Secondary	Secondary	Secondary
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$ln(S/U)_T z$	1.2781***	1.0943***	1.2259***	1.2027***			
	(0.3982)	(0.3816)	(0.3877)	(0.3892)			
$ln(S/U)_S z$					0.8451***		0.9315***
					(0.3085)		(0.3049)
$ln(\gamma)z$	0.3516		0.3796			0.2915	0.4541*
	(0.2482)		(0.2357)			(0.2375)	(0.2470)
$ln(\gamma_{util})z$				0.3232			
				(0.2326)			
$\epsilon_{A(z),ct}$	0.3131***		0.3388***			0.2999***	0.3002***
	(0.1081)		(0.1043)			(0.0997)	(0.0992)
$\epsilon_{A(z),util,ct}$				0.3119***			
				(0.0947)			
Obs	3601	3711	3711	4063	4063	4063	4063
Country-Time FE	yes	yes	yes	yes	yes	yes	yes
Industry-Time FE	yes	yes	yes	yes	yes	yes	yes
R^2	0.8911	0.8873	0.8912	0.8919	0.8883	0.8889	0.8919

*** estimated at the 1% level, ** estimated at the 5% level. Robust standard errors clustered by country-industry.

following proxy for capacity utilization $\frac{\frac{M_{czt}}{K_{czt}}}{\left(\frac{M_{cz}}{K_{cz}}\right)_{med}}$ where M is a broad measure of intermediate inputs that is defined as the difference between output and value added. Because the ratio of materials to capital is likely to vary broadly across countries for reasons unrelated to capacity utilization, I divide it by the median value for that country-industry over time. If materials use has increased relative to capital use relative to other years, this can be a signal of an increased workweek of capital and utilization. I multiply the capital stock of the industry-country-year observation by this value. Column 4 shows that does not change the results.

Columns 5, 6 and 7 use the relative abundance of workers with at least a secondary education as defined in the Barro and Lee data set as an alternate measure of skilled labor abundance. I use the broad measure of skilled labor intensity because it is closer in comparability than the narrow measure. In Column 7, it appears that $ln(\gamma)z$ does possess some explanatory power when included with factor abundance. However, Column 6 shows that this is only when it is conditioned on factor abundance and that its explanatory power falls when factor abundance is not included.

The error terms in the panel regressions presented above are undoubtedly correlated. The more substantive question is if the correlation emerges from repeatedly observing a slow moving equilibrium relationship or if the correlation emerges due to a specific dynamic economic structure. Generally, if errors are correlated due to a specific dynamic structure of the underlying economic model, clustering of the standard errors will yield inconsistent point estimates.[37] The first column of Table 8 explores this question. Using data for 1988, I show that nearly all of the variation comes from the cross section and, consequently, this concern is unfounded. I choose this year because it contains the most observations of any single year. The coefficients and standard errors are extremely similar to those in other regressions suggesting that the correlation of the error terms is sufficiently accounted for by clustering of the error terms.[38]

The imposition of a constant z in a given industry is unlikely to be completely true but it is less obvious how severe a bias this introduces. Columns 2-7 in Table 8 address this problem. Columns 2 and 3 replicate Columns 1 and 3 of Table 5 except that they replace all numerical values with

[37] See Maddala (1988), page 200.

[38] Nickell (1981) suggests that other methods such as including a lagged endogenous term are likely to introduce more problems than they solve when the time dimension of the sample is sufficiently short. The same criticism applies to a GLS estimation of the system.

Table 8: Robustness Check II

Variable	1988 (1)	US Rank (2)	US Rank (3)	US Rank (4)	US Rank (5)	Own Rank (6)	Own Rank (7)	Own Rank (8)
$ln(S/U)_T z$	1.1178***							
	(0.3638)							
$rank(S/U)$ $*rank(z)$		0.0127***		0.0124***	0.0113***	0.0046**	0.0042**	0.0039**
		(0.0028)		(0.0028)	(0.0027)	(0.0020)	(0.0021)	(0.0018)
$ln(\gamma)z$	0.1884							
	(0.2278)							
$rank(ln(\gamma))$ $*rank(z)$			0.0017	0.0043*	0.0029		-0.0002	-0.0007
			(0.0024)	(0.0026)	(0.0024)		(0.0021)	(0.0018)
$\epsilon_{A(z),ct}$	0.4098***							
	(0.1253)							
$rank(a(z))$			0.0631***	0.0602***	0.0604***		0.0600***	0.0584***
			(0.0193)	(0.0193)	(0.0186)		(0.0194)	(0.0186)
Obs	454	4063	4063	4063	454	4063	4063	454
Country-Time FE	yes	yes	yes	yes	yes	yes	yes	yes
Industry-Time FE	yes	yes	yes	yes	no	yes	yes	no
Sample	1988	Full	Full	Full	1988	Full	Full	1988

*** estimated at the 1% level of certainty, ** estimated at the 5% level of certainty. Robust standard errors clustered at the country-industry level their rank. Output is replaced by the rank of output in each country industry after it has been purged of country-year and industry-year fixed effects. Educational attainment is replaced by its world rank in that statistic in that year. Each z is replaced by the skill rank of that industry in the United States as measured by the proportion of non-operative wages in total wages in the United Nations General Industrial Statistical Dataset. $a(z)$ is replaced by the TFP rank of that country industry across all countries in that industry in that year after it has been purged of country-year and industry-year fixed effects. $ln(\gamma)$ is replaced by its world rank in that year. Because I am now dealing with rank orderings, OLS is not appropriate and I perform an ordered logit. Although the coefficients are not comparable, the same general patterns of magnitude and significance hold. Again, Column 4 suggests that $ln(\gamma)$ possesses some explanatory power but Column 3 shows that this disappears when it is not conditioned on factor abundance. Column 5 performs the same exercise on cross sectional data from 1988 with the same result.

Columns 6-8 perform the same exercises as 2-4 except that the skill rank of the industry in the United States is now replaced by the rank of the proportion of non-operative wages in total wages of the industry in that country as measured by the United Nations General Industrial Statistical Dataset.[39] Consequently, it is less constrained than columns 2-4. Although the point estimates on factor abundance are now smaller, the same patterns of magnitude hold. The smaller coefficients are possibly due to measurement error in representing skilled labor intensity by the proportion of non-production workers across countries.

7 Conclusion

The Ricardian and Heckscher-Ohlin (HO) theories are the workhorse models of international trade. Neither model, in isolation, offers a complete description of the data, nor does either model offer a unified theory of international trade. This paper presents a unified framework that nests these two models in determining comparative advantage when there is a continuum of industries if countries differ both in factor abundance and relative TFP patterns across industries. In addition, the model's tractability allows me to estimate it easily and to assess the relative contributions of HO

[39] I am not comparing industries across countries but industries within a country so that the objection to using the UNGISD data raised in footnote 23 is not valid.

and Ricardian forces. I highlight three results.

First, both the Ricardian and HO models possess robust explanatory power in determining international patterns of production. Second, these two models are separable in the sense that the forces that determine comparative advantage in one are orthogonal to the forces that determine comparative advantage in the other in my broad sample. Although the first result has been documented in past reduced form estimation, my paper is the first to do so based on a unified model where the estimated coefficients can be mapped against structural parameters. The second result is new and suggests conditions under which the two models are orthogonal in that Ricardian TFP differences do not cause or prevent HO effects holds in the data. Third, I find that a one standard deviation change in relative factor abundance is 1.6 to 2.4 times as potent in changing the structure of an industry in an economy as a one standard deviation change in the relative productivity of that industry.

The theoretical contributions of this paper are twofold. First, if TFP is orthogonal to factor intensity, it is reasonable to model productivity using two components: a country specific term that is neutral across industries and an idiosyncratic component that is orthogonal to factor intensities. Simply examining if relative TFP is relatively more positively or negatively correlated with factor intensity in countries that possess a relative abundance of that factor is a good starting point for assessing if this is likely to be a reasonable assumption.

Second, when trying to make industry by industry *level* predictions, HO models will be misspecified if they omit TFP differences even if TFP is uncorrelated with factor intensity. However, if one is trying to identify coefficients such as in Rybczynski regressions, this will be valid if TFP is uncorrelated with factor intensity but not if TFP is correlated with factor intensity. Although I find that TFP is uncorrelated with factor intensity in my sample, the obvious caveat applies that such a (zero) correlation is ultimately an empirical question that depends on the data set.

References

[1] Acemoglu, Daron, (1998) "Why do New Technologies Complement Skills? Directed Technical Change and Wage Inequality" *The Quarterly Journal of Economics* Vol. 113(4), pg. 1055-1089

[2] Anderson, James E. and van Wincoop, Eric, (2004) “Trade Costs” NBER Working Paper 10480, May 2004.

[3] Baldwin, Robert, (1971) “Determinants of the Commodity Structure of U.S. Trade” *The American Economic Review* Vol. 61(1), pg. 126-146

[4] Barro, Robert J. and Lee, Jong-Wha, ”International Data on Educational Attainment: Updates and Implications” (CID Working Paper No. 42, April 2000)

[5] Basu, Susanto, (1996) ”Procyclical Productivity: Increasing Returns or Cyclical Utilization?” *The Quarterly Journal of Economics* Vol. 111(3), pg. 719-751

[6] Basu, Susanto and Kimball, Miles ,(1997) ”Cyclical Productivity with Unobserved Input Variation” NBER Working Paper 5915, February 1997.

[7] Berman, Eli, Bound, John, and Machin, Steven, (1998) “Implications of Skill Biased Technological Change: International Evidence,” *The Quarterly Journal of Economics* Vol. 113(4), pg. 1245-1279

[8] Bernard, Andrew B. & Peter K. Schott & Redding, Steven, (2006) “Comparative Advantage and Heterogeneous Firms,” *Review of Economic Studies*, Vol. 74(1), pg. 31-66

[9] Bils, Mark & Klenow, Peter J., (2000) ”Does Schooling Cause Growth?” *The American Economic Review*, Vol 90(5), pg. 1160-1183

[10] Bowen, Harvey & Leamer, Edward & and Sveikauskas, Leo, (1997) “Multicountry, Multifactor Tests of Factor Abundance Theory,” *The American Economic Review*, Vol. 77(5), pg. 791-809

[11] Broda, Christian & Weinstein, David, (2006) ”Globalization and the Gains from Variety,” *The Quarterly Journal of Economics*, Vol. 121(2), pages 541-585

[12] Caselli, Francesco, (2005) “Accounting for Cross-Country Income Differences” in *Handbook of Economic Growth, vol. 1*, Edited by Philippe Aghion and Steven Durlauf, Elsevier Science Publishers B.V., 2005

[13] Caves, Douglas W. & Christensen, Lauritis R., & Diewert, W. Erwin, (1997) “Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers” *The Economic Journal* Vol. 92(365), pg. 73-86

[14] Costinot, Arnaud and Komunjer, Ivana (2007) “What Goods Do Countries Trade? New Ricardian Predictions” NBER Working Paper # 13691 December, 2007, National Bureau of Economic Research

[15] Davis, Donald R. & Weinstein, David E. & Bradford, Scott C. & Kazushige, Shimpo, (1997) “Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works” *The American Economic Review* Vol. 87(3), pg. 421-486

[16] Debaere, Peter, (1997) “Relative Factor Abundance and Trade” *The Journal of Political Economy* Vol. 111(3), pg. 589-610

[17] Deardorff, Alan V., (1984) “Testing Trade Theories and Predicting Trade Flows” in *Handbook of International Economics, vol. 1*, Edited by Ronald W. Jones and Peter B. Kenen, Elsevier Science Publishers B.V., 1984

[18] Dollar, David & Wolff, Edward N., *Competitiveness, Convergence, and International Specialization.* Cambridge, MA: MIT Press 1993

[19] Dornbusch, Rudiger & Fischer, Stanley & Samuelson, Paul A, (1977) ”Comparative Advantage, Trade and Payments in a Ricardian Model with a Continuum of Goods” *The American Economic Review* Vol. 67(5), pg. 823-839

[20] Dornbusch, Rudiger & Fischer, Stanley & Samuelson, Paul A, (1980) “Heckscher- Ohlin Trade Theory with a Continuum of Goods” *The Quarterly Journal of Economics* Vol. 95(2), pg. 203-224

[21] Eaton, Jonathan & Kortum, Samuel (1980) “Technology, Geography, and Trade,” *Econometrica* Vol. 70(5), pg. 1741-1779

[22] Feenstra, Robert C.& Lipsey, Robert E.& Deng, Haiyan & Ma, Alyson C. & Mo, Hengyong (2005) “World Trade Flows: 1962-2000” NBER Working Paper # 11040. Available at www.nber.org/data

[23] Fitzgerald, Doireann and Hallak, Juan Carlos, (2004) “Specialization, Factor accumulation and Development” *The Journal of International Economics* Vol. 64(2), pg. 277-302

[24] Hall, Robert E. & Jones, Charles I., (1999) “Why Do Some Countries Produce So Much More Output Per Worker Than Others?” *The Quarterly Journal of Economics* Vol. 114(1), pg. 83-116

[25] Harrigan, James, (1997) “Technology, Factor Supplies, and International Specialization: Estimating the Neoclassical Model” *The American Economic Review* Vol. 87(4), pg. 475-494

[26] Harrigan, James, (1997b) “Cross Country Comparisons of Industry Total Factor Productivity”, Mimeo, Federal Reserve Bank of New York.

[27] Harrigan, James, (1999) “Estimation of cross-country differences in industry production functions,” *The Journal of International Economics* Vol. 47, pg. 267-293

[28] Harrigan, James and Zakrajzek, Egon (2000) “Factor Supplies and Specialization in the World Economy” NBER Working Paper # 7848 August, 2000, National Bureau of Economic Research

[29] Jones, Ronald (1965) “The Structure of Simple General Equilibrium Models” *The Journal of Political Economy* Vol. 73, pg. 557-572

[30] Kahn, James A. & Lim, Jong-Soo (1998) “Skilled Labor-Augmenting Technical Progress in U.S. Manufacturing” *The Quarterly Journal of Economics* Vol. 113(4), pg. 1281-1308

[31] Keller, Wolfgang (2002) “Geographic Localization of International Technology Diffusion” *The American Economic Review* Vol. 92(1), pg. 120-142

[32] Kravis, Irving B. & Heston, Alan & and Summers, Robert, (1982) World Product and Income: International Comparisons of Real Gross Product Produced by the Statistical Office of the United Nations and the World Bank ; Baltimore : Published for the World Bank [by] the Johns Hopkins University Press, c1982

[33] Krugman, Paul (1995), “Growing World Trade: Causes and Consequences”, Brookings Papers on Economic Activity, Volume 1.

[34] Krugman, Paul (2008), “Trade and Wages, Reconsidered”, Brookings Papers on Economic Activity, *forthcoming.*

[35] Leamer, Edward E. “A Combined Ricardian and Heckscher-Ohlin Model of Comparative Advantage” in Quiet Pioneering: Robert M. Stern and His International Legacy, edited by Keith E. Maskus, Peter M. Hooper, Edward E. Leamer, and J. David Richardson, Ann Arbor, The University of Michigan Press, 1997, pg. 11-34

[36] Leamer, Edward E. & Levinsohn, James “International Trade Theory: The Evidence” in Handbook of International Economics, edited by Gene Grossman, Ken Rogoff, Vol. III, pg. 1339-1394

[37] Leontief, Wassily (1954) ”Domestic Production and Foreign Trade: The American Capital Position Reexamined,” *Economia Internatiozionale* 7, (February), pp. 3-32.

[38] MacDougall, G.D.A. (1951) ”British and American Exports: A Study Suggested by the Theory of Comparative Costs, part I,” *Economic Journal* Vol. 61(244), pp. 697-724

[39] MacDougall, G.D.A. (1952) ”British and American Exports: A Study Suggested by the Theory of Comparative Costs, part II,” *Economic Journal* Vol. 62(247), pp. 487-521

[40] Maddala, G.S., *Introduction to Econometrics*. New York, NY: Macmillan Publishing Company 1988

[41] Melitz, Marc, (2003) "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, Vol. 71(6), pg. 1695-1725

[42] Nickell, Steven J. (1981) "Biases in Dynamic Models with Fixed Effects," *Econometrica*, Vol. 49(6), pg. 1417-1426

[43] Nicita, Alessandro and Olarreaga, Marcelo (2001) "Trade and Production, 1976-1999," Dataset and Documentation World Bank

[44] Olley, Steven and Pakes, Ariel (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica* Vol. 64(6), pp. 1263-1298

[45] Romalis, John, (2004) "Factor Proportions and the Commodity Structure of Trade" *The American Economic Review* Vol. 94(1), pg. 67-97

[46] Stern, Robert M. (1962) "British and American Productivity and Comparative Costs in International Trade," *Oxford Economic Papers* Vol. 14(3), pp. 275-296

[47] Trefler, Daniel, (1993) "International Factor Price Differences: Leontief Was Right!" *The Journal of Political Economy* Vol. 101(6), pg. 961-987

[48] Trefler, Daniel, (1995) "The Case of the Missing Trade and Other Mysteries" *The American Economic Review* Vol. 85(5), pg. 1029-1046

[49] United Nations (1995), General Industrial Statistics Dataset and Documentation.

A Relative Number of Firms

Romalis (2004) (equation 14) solves for the relative number of firms/varieties produced in the North relative to the South in a given industry z. He starts with the fact that firms' income in the North and South equal revenue from Northern and Southern consumers as reflected in the below equations:

$$p(z)x(z) = \frac{1}{2}\left(\frac{p(z)}{P(z)}\right)^{1-\sigma} Y + \frac{1}{2}\left(\frac{p(z)\tau}{P^*(z)}\right)^{1-\sigma} Y^*$$

$$p^*(z)x^*(z) = \frac{1}{2}\left(\frac{p(z)^*\tau}{P(z)}\right)^{1-\sigma} Y + \frac{1}{2}\left(\frac{p^*(z)}{P^*(z)}\right)^{1-\sigma} Y^*$$

Using the fact that $P(z)^{1-\sigma} = n(z)p(z)^{1-\sigma} + n^*(z)\left(p^*(z)\tau\right)^{1-\sigma}$ and an analogous expression for $P^*(z)$, he solves for $\tilde{n}(z) = \frac{n(z)}{n^*(z)}$:

$$\tilde{n} = \frac{\tau^{2(1-\sigma)}\frac{Y^*}{Y} + 1 - \tau^{1-\sigma}\tilde{p}^{\sigma}(\frac{Y^*}{Y}+1)}{\tilde{p}(\tau^{2(1-\sigma)} + \frac{Y^*}{Y}) - \tilde{p}^{1-\sigma}\tau^{1-\sigma}(\frac{Y^*}{Y}+1)}. \tag{31}$$

Romalis emphasizes that the above expression is not guaranteed to be positive. When it is positive, intra-industry trade exists, otherwise specialization occurs with only one country producing in the industry. I examine the case of intra-industry trade in this paper. A necessary and sufficient condition for this to hold is that $\tilde{p}(z)_{lower} < \tilde{p}(z) < \tilde{p}(z)_{upper}$ where

$$\tilde{p}_{upper} = \left[\frac{\tau^{2(1-\sigma)}\frac{Y^*}{Y} + 1}{\tau^{1-\sigma}\left(\frac{Y^*}{Y}+1\right)}\right]^{\frac{1}{\sigma}} > 1 \tag{32}$$

$$\tilde{p}_{lower} = \left[\frac{\tau^{1-\sigma}(\frac{Y^*}{Y}+1)}{\tau^{2(1-\sigma)} + \frac{Y^*}{Y}}\right]^{\frac{1}{\sigma}} < \tilde{p}_{upper}. \tag{33}$$

Romalis also proves that the derivative of the number of firms with respect to relative prices is negative. For a more in depth discussion of this, the interested reader is directed to the Technical Appendix of Romalis (2004). Dropping the z notation, the derivative is as follows:

$$\frac{d\tilde{R}(\tilde{p})}{d\ln(\tilde{p}(z))} = \frac{-\sigma\tau^{1-\sigma}\left(\frac{1-\pi}{\pi}+1\right)\left[\tilde{p}_0^{\sigma}\left(\tau^{2(1-\sigma)}+\frac{1-\pi}{\pi}\right)-2\tau^{1-\sigma}\left(\frac{1-\pi}{\pi}+1\right)+\tilde{p}_0^{-\sigma}\left(\tau^{2(1-\sigma)}\frac{1-\pi}{\pi}+1\right)\right]}{\left[\tau^{2(1-\sigma)}+\frac{1-\pi}{\pi}-\tilde{p}_0^{-\sigma}\tau^{1-\sigma}\left(\frac{1-\pi}{\pi}+1\right)\right]^2}$$

B Derivation of Goods Market Clearing Condition

To show that the goods market clearing condition is downward sloping in $\tilde{\omega}-V$ space, I simply show that if $\omega < \omega^*$, then $V > V^*$. Start by noting that $V > V^*$ if and only if $\tilde{R}(z_s) > \tilde{R}(z_u)$. Therefore, It is sufficient to show that if $\omega < \omega^*$, then $\tilde{R}(z_s) > \tilde{R}(z_u)$ or simply that $\tilde{R}(z)$ is increasing in z if and only if $\omega < \omega^*$. Taking the derivative of $\tilde{R}(z)$ with respect to z yields the following expression

$$\frac{\partial\tilde{R}(z)}{\partial z} = \frac{-\sigma\tau^{1-\sigma}\left(\frac{Y^*}{Y}+1\right)\left[\tilde{p}(z)^{sigma}\left(\tau^{2(1-\sigma)}+\frac{Y^*}{Y}\right)-2\tau^{1-\sigma}\left(\frac{Y^*}{Y}+1\right)+\tilde{p}(z)^{sigma}\left(\tau^{2(1-\sigma)}\frac{Y^*}{Y}+1\right)\right]}{\left[\tau^{2(1-\sigma)}+\frac{Y^*}{Y}-\tau^{1-\sigma}\tilde{p}(z)^{-\sigma}\left(\frac{Y^*}{Y}+1\right)\right]^2} ln(\tilde{\omega})$$

The large fraction is unambiguously negative as noted in Appendix A and Romalis (2004), therefore $\tilde{R}(z)$ is increasing in z if and only if $\omega < \omega^*$.

C Proof of Proposition 1

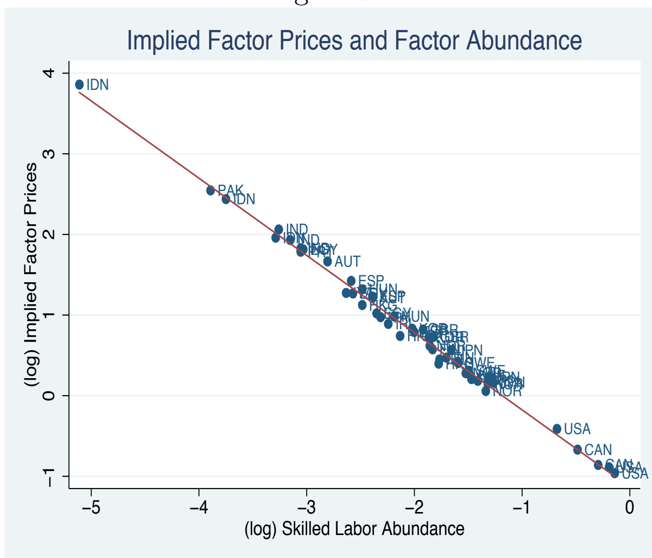
Proposition 1 If productivity is uncorrelated with factor intensity and the relative abundance of factors differs among countries, then the relative wage of a country's abundant factor will be less than in the country where it is a relatively scarce factor. In addition, $cov[v(z), z] > 0$ where z is the Cobb-Douglas cost share of its relatively abundant factor and $cov[v(z'), z'] < 0$ where z' is the Cobb-Douglas cost share of its relatively scarce factor.

Proof of Proposition 1: This proof proceeds in two steps and closely resembles a similar proof in Romalis (2004). First, I show that FPE breaks down. Second, I show that the North possesses relatively cheap skilled labor. First, based on the expression for $\tilde{p}(z)$, if FPE results $v(z)$ is constant across sectors and uncorrelated with $b(z)$ and z. Consequently, the left hand side of equation 20 is equal to unity and the right hand side is greater than unity. This is a contradiction. Consequently, FPE breaks down. Second, given assumptions about factor abundance, full employment of each factor implies that the North either i) has a larger share of relatively skilled labor intensive industries or ii) use more skilled labor intensive techniques. If TFP is uncorrelated with z, the first statement requires that $\tilde{\omega} < 1$ based on equation 18. Second, based on Cobb-Douglas production, the use of more skilled labor intensive techniques in each industry also requires that $\tilde{\omega} < 1$. Consequently, $\tilde{\omega} < 1$ and $cov[z, v(z)] > 0$ by equation 18. $cov[z', v(z')] < 0$ follows trivially.

D Empirical Representation of Equilibrium Factor Prices

This section deals with the relationship between relative factor prices, relative endowments, and the industrial structure of productivity differences and the question of if $ln\left(\tilde{\omega}\right) = \kappa_0 ln\left(\tilde{S}/\tilde{U}\right) + \kappa_1 ln\left(\tilde{\gamma}\right)$ is a reasonable empirical representation. Specifically, I combine production data, relative endowments, and the factor market clearing condition (equation 20) and calculate the equilibrium factor prices that are implied by these patterns. I then show that the vast majority of the cross sectional variation in relative implied factor prices is accounted for by relative endowments, thus validating the approach in the text.

Figure 5:



Specifically, I use the Trade and Production data to calculate market shares for each country in each industry ($v(z)$) and to calculate world expenditure shares by industry ($b(z)$). I combine these data with the data on skill intensity from Table 1 and the relative endowments data with equation 20 and calculate ω. Graph 5 below plots the implied values of $ln(\omega)$ against $ln(\frac{S}{U})$ for 1985, 1990, and 1995 pooled together.

Table 9 shows the estimated coefficients from the following simple regressions to judge the empirical validity of the linear expression proposed in the text in columns (1)-(3) respectively:

$$ln(\omega_{ct}) = \beta_0 ln(S_{ct}/U_{ct}) + \beta_T T' + \epsilon_{ct}$$

$$ln(\omega_{ct}) = \beta_0 ln(S_{ct}/U_{ct}) + \beta_1 ln(\gamma_{ct}) + \beta_T T' + \epsilon_{ct}$$

$$ln(\omega_{ct}) = \beta_0 ln(S_{ct}/U_{ct}) + \beta_1 ln(\gamma_{ct}) + \beta_2 ln(GDP) ln(S_{ct}/U_{ct}) + \beta_3 ln(GDP) ln(\gamma_{ct}) + \beta_4 ln(GDP) + \beta_T T' + \epsilon_{ct}$$

Based on the extremely high R^2 in the first column and the only marginal improvements in the second and third columns. A simple log-linear specification with endowments alone is likely to be reasonable as a reduced form characterization of factor prices. It is useful to define two special cases that bound the estimates of κ_0. First, if factor abundance has no effect on production patterns such that the left hand side of equation 20 is constant across countries, the estimated coefficient should be unity. If we observe FPE due to strong HO effects, we would expect to observe a coefficient close to or equal to zero. The value of $\kappa_0 = -0.95$ suggests that trade has little effect on relative factor prices consistent with Krugman (1995,2008).

E Data Appendix: Sample

See Tables 10 and 11.

Table 9: Determinants of Equilibrium Implied Factor Prices

Variable	(1)	(2)	(3)
$ln(S/U)$	-0.9540^{***}	-0.9563^{***}	-0.9562^{***}
	(0.0128)	(0.0108)	(0.0100)
$ln(\gamma)$		-0.0018	
		(0.0151)	
$ln(GDP)$		$1.41e-11$	$1.14e-11^{**}$
		$(1.05e-11)$	$(4.65e-12)$
$ln(S/U)ln(GDP)$		$-1.79e-11$	$-1.99e-11^{*}$
		$(1.43e-11)$	$(1.05e-11)$
$ln(\gamma)ln(GDP)$		$5.58e-11$	
		$(1.84e-11)$	
Time Fixed Effects	Yes	Yes	Yes
R^2	0.9932	0.9950	0.9950

*** estimated at the 1% level, ** estimated at the 5% level
* estimated at the 10% level. Standard errors clustered by country.

Table 10: Sample

Country	Years Available	Obs	Country	Years Available	Obs
Austria	1985-1994	239	Ireland	1985-1991	154
Canada	1985-1990	138	Italy	1985-1994	190
Denmark	1985-1991	168	Japan	1985-1995	264
Egypt	1985-1995	247	Korea	1985-1995	264
Finland	1985-1995	261	Norway	1985-1995	246
Great Britain	1985-1992	192	Pakistan	1985-1988	96
Hong Kong	1985-1995	189	Portugal	1985-1989	120
Hungary	1985-1993	216	Spain	1985-1995	176
Indonesia	1985-1995	242	Sweden	1985-1990	138
India	1985-1995	264	United States	1985-1995	264

Table 11: Sample by Year

Year	Observations	Year	Observations
1985	451	1991	359
1986	452	1992	312
1987	451	1993	282
1988	454	1994	257
1989	430	1995	210
1990	405		

F Data Appendix: Calculating the Cobb-Douglas Cost Share of Skilled Labor

I calculate Cobb-Douglas factor cost shares of the total wage bill for skilled and unskilled labor. Suppose that s indexes the different types of skilled labor. For any level of skill s, its Cobb-Douglas factor cost share of wage will be:

$$z_s = \frac{w_s L_s}{\sum_{s'} w_{s'} L_{s'}} \tag{34}$$

To calculate this value, I estimate a Mincerian wage regression of the form

$$ln(w_{it}) = \beta_0 + \beta_1 age_{it} + \beta_2 age_{it}^2 + \beta'_{edu} EDU_{it} + \beta' T_{it} + \epsilon_{it} \tag{35}$$

w is the hourly wage based on data on income, weeks worked, and average work week. *Age* is the age of the worker. EDU_{it} is a vector of dummy variables indicating education attainment of different levels. T is a series of time fixed effects. All data comes from the March U.S. Current Population Surveys for the years 1988-1992. The regression itself is run on data that is pooled over industries and years. The data is available for download from $http://www.ipums.umn.edu/usa/data.html$.Wage and salary income is *incwage*. Weeks worked is $wkswork1$. Average work week is $uhrswork1$. Age is *age*. The variable *educrec* indicates the highest education level of the worker in the survey. The levels of educational attainment indicated are: 1) None of Preschool, 2) Grades 1-4, 3) Grades 5-8, 4)Grade 9, 5) Grade 10, 6) Grade 11, 7) Grade 12, 8) 1-3 years of college, and 9) 4 or more years of college.

When running the wage regression, a vector of coefficients will be returned that give the skill premium for different levels of educational attainment. Because they are dummy variables, they will state the wage of a person of that educational attainment relative to the omitted level. I use the variable *educrec* and define four types of labor: 0-11 grades of school completed, 12th grade completed, 1-3 years of college, and 4 or more years of college). Applying this to the definition of z given above, this is equivalent to dividing the numerator and denominator by a given (omitted) wage level:

$$z_{s''} = \frac{\frac{w_{s''}}{w_s} L_{s''}}{\sum_{s' \neq s} \frac{w_{s'}}{w_s} L_{s'} + L_s} \tag{36}$$

By dividing through by a numeraire wage, the physical workers are converted to effective workers. Although this will be invariant to the omitted skill level, I do need to take a stand on what comprises skilled and unskilled labor. Suppose that the factor share of "skilled labor" is the sum of the factor shares of the types of labor deemed to be skilled:

Table 12: Country Level Covariances

Country	Cov (narrow)	Cov (broad)	Country	Cov (narrow)	Cov (broad)
Austria	0.0183	0.0206	Ireland	0.0366	0.0438
Canada	0.0267	0.0436	Italy	-0.0006	0.0059
Denmark	0.0450	0.0500	Japan	0.0421	0.0605
Egypt	0.0252	0.0258	Korea	0.0152	0.0166
Finland	0.0278	0.0383	Norway	0.0425	0.0636
Great Britain	0.0397	0.0508	Pakistan	0.0245	0.0206
Hong Kong	0.0300	0.0356	Portugal	0.0031	-0.0051
Hungary	0.0262	0.0309	Spain	0.0334	0.0246
Indonesia	-0.0070	-0.0165	Sweden	0.0473	0.0713
India	0.0381	0.0464	United States	0.0520	0.0659

$$z_{skill} = \sum_{s \in skilled} z_s \tag{37}$$

I use two measures of skilled labor: a "narrow" measure that only counts those in the final category as skilled labor and a "broad" measure for those who have any college and fall into the last two groups. If I divide the numerator and denominator by the amount of total labor employed in a industry and define $\alpha_s = \frac{L_s}{L_{total}}$, the skilled labor share is:

$$z_{s''} = \frac{\frac{w_{s''}}{w_s}\alpha_{s''}}{\sum_{s' \neq s} \frac{w_{s'}}{w_s}\alpha_{s'} + \alpha_s} \tag{38}$$

Therefore, I can calculate skilled labor intensity using data on the proportion of workers in a given industry of differing education levels and the coefficients from the wage regression.

G Country Level Covariances

See Table 12.

H Data Appendix: Deflators from the Penn World Tables Disaggregated Benchmark Data

I use the Penn World Tables benchmark data to deflate value added across countries. This is obviously not a first best outcome but it represents a substantial improvement on the literature. The benchmark data is available at *http://pwt.econ.upenn.edu/Downloads/benchmark/benchmark.html*. This data was collected by examining very narrowly defined goods across a number of countries with specific attention paid to the quality of goods across countries. See Kravis, Heston and Summers (1982) for a thorough explanation of the process of creating the price indexes. Because of substantially finer disaggregation across goods, I use the benchmark data from 1985 instead of 1996. I also use the 1980 data to fill in missing observations for Indonesia. I also assume that all prices increase at the same rate as the PPP GDP price deflator which allows me to fill in observations for other years. Because all country-year level price differences are differenced out through the use of logs, this filling in of the interim years assumes that the relative prices across industries in a country in 1985 (and 1980 in Indonesia) persist throughout the sample.

As noted in Harrigan (1997b), these measures are subject to the following criticisms as to why they might not truly reflect country-industry level deflators. First, these prices include import prices and exclude export prices. Second, these prices include transport and distribution margins. Third, they include indirect taxes and exclude subsidies. Finally, fourth, these prices only refer to final output and not intermediate goods. For these reasons, these deflators should only be taken as approximations to actual deflators. For this reason, he constructs actual deflators from the OECD national accounts data. Because of the severe limitations that this places on the data, I choose to use the ICP data and compare my results to his. As shown in Table 5, this appears to be a reasonable approximation.

The original data was collected via the United Nations International Comparison Programme (ICP) classification level which is available at *http://unstats.un.org/unsd/methods/icp/ipc8_htm.htm*. Because there is no clean concordance between this classification and the ISIC classification used in the Trade and Production dataset, I created a concordance that is available on my website. The only departures from this process were Iron and Steel (ISIC 371) and Non-Ferrous Metals (ISIC 372). These goods have no convenient analog in the ICP project and they are relatively homogenous and highly traded. Therefore, I assume that the appropriate cross country deflator for these industries is unity.

Unlike other authors (e.g. Dollar and Wolff), I do not use the country level PPP price levels because this is highly influenced by the non-traded industries. This will lead to output being deflated "too much" in poor countries which will understate their productivity levels. In addition, even if a researcher possesses a PPP deflator for traded goods, there is substantial heterogeneity in the PPP price deflator across ICP industries. A simple fixed effects regression of all logged PPP deflators across industries and countries on a series of country level fixed effects only captures 35% of the variation in estimations that I have carried out.

I Data Appendix: Effective Labor

The employment measure L does not differentiate between skilled and unskilled labor. However, I follow Caselli (2005) and Bils and Klenow (2002) and use educational attainment and wage premium data to construct measures of the effectiveness of labor. The most basic specification would be a log-linear structure in which the effectiveness of a measured unit of labor (E) is affected by years of schooling (s) according to the semi-elasticity ϕ. Table 13 presents estimates.

$$ln(E) = \phi_0 + \phi_1 s \tag{39}$$

The parameter ϕ_1 is taken to be the coefficient on years of schooling in a Mincerian wage regression. However, country level data on ϕ are likely to be incomparable for two reasons. First, the samples from which these estimates are drawn might differ even controlling for the level of development in the country. Secondly, even if the economic relationship is stable across countries, ϕ_1 is likely to be higher for less developed countries due to the relative paucity of skilled workers. This is confirmed by examining the data presented in Psacharopoulos (1994). For this reason, I follow Casellli (2005) and assume that each additional year of education makes a worker 13% more effective for the first four years of schooling, 10% for years 4-8, and 7% a year after that. In addition to having published data on the educational attainment rates for different levels of education, Barro and Lee also possess average years of schooling. This data is available at *http://www.cid.harvard.edu/ciddata/ciddata.html*.

J Data Appendix: Capital Stock Calculation

Capital is calculated using the perpetual inventory method where investment is deflated across countries using the Penn World Tables PPP investment price deflator and the United States implicit price deflator for non-residential investment from the Bureau of Economic Analysis to achieve comparability across time.

Table 13: Effective Labor Across Countries

Country	E	Country	E
Austria	2.55	Ireland	2.60
Canada	2.99	Italy	1.96
Denmark	2.91	Japan	2.73
Egypt	1.59	Korea	2.74
Finland	2.78	Norway	3.06
Great Britain	2.64	Pakistan	1.35
Hong Kong	2.58	Portugal	1.72
Hungary	2.64	Spain	1.95
Indonesia	1.54	Sweden	2.80
India	1.61	United States	3.32

To attain the least sensitivity, I merge the Trade and Production dataset with the United Nations General Industrial Statistical Dataset used by Berman, Bound and Machin (1998). All data begins in 1976 for the Trade and Production dataset, however, merging it with the UNGISD database gives earlier initial years. The following data gives the average initial capital stock remaining in 1985 (from its initial year) and the initial year from which the capital stock calculations is made (t_0): Austria (0.238,1967), Canada (0.106,1967), Denmark (0.114,1967), Egypt (0.044,1967), Finland (0.067,1967), Great Britain (0.151,1968), Hong Kong (0.22,1973), Hungary (0.140,1970), Indonesia (0.199,1970), India (0.279,1977), Ireland (0.124,1969), Italy (0.084,1967), Japan (0.106,1967), South Korea (0.019,1967), Norway (0.114,1967), Pakistan (0.61,1976), Portugal (0.151,1971), Spain (0.10,1967), Sweden (0.126,1967), United States (0.010,1967). Initial Capital stock is calculated as follows:

$$K(z)_{c,t_0} = \frac{I_{cz,t_0}}{\delta + g} \tag{40}$$

where g is the median growth of gross investment over the available sample for a country and $\delta = 0.125$.[40] Starting from this point, I calculate the capital stock as the sum of flow investment net depreciation as below.

$$K(z)_{c,t+1} = (1-\delta)K(z)_{c,t} + I(z)_{c,t} \tag{41}$$

[40] This is similar to the approach taken by Hall and Jones (1998). In some cases, the growth rate of the gross investment over the sample was negative enough to result in estimates of the starting value of the capital stock being negative. In these cases, I set $g = 0$