Factor Uses and the Pattern of Specialization

Georg Schaur, Chong Xiang* and Anya Savikhin

Purdue University

December, 2006

Abstract:

We show that in the setting of multiple goods and factors, the factor proportions theory has the following prediction: *across industries*, the impacts of the endowment of a given factor on industry outputs have positive co-variance with the relative uses of this factor. The intuition is that on average, the industries that use a given factor heavily have positive output responses following an increase in the endowment of this factor. This co-variation condition is robust to Hicks-neutral and factor-augmenting productivity differences and constitutes a direct test of the production side of the factor proportions theory. We also show that the co-variation condition finds empirical support. This is evidence consistent with the factor proportions theory.

JEL Classification: F1

Keywords: specialization; factor uses; co-variation condition; relative factor uses; factor proportions theory.

* Corresponding author. Department of Economics, KRAN, Purdue University, 403 West State Street, West Lafayette, IN 47907 (cxiang@purdue.edu, 765-494-4499). We thank David Hummels, as well as seminar participants at Purdue University, Indiana University and the Spring 2006 Midwest Meetings for helpful comments.

Section 1 Introduction

It is important to understand why and how countries specialize. The pattern of specialization is critical to understanding the pattern of trade (Harrigan 2003) and the impacts of globalization on wages (Schott 2003), and may contribute to the dispersion of the world income distribution (Acemoglu and Ventura 2002). It helps shed light on the relation between trade and growth (Redding 2002) and the mechanisms through which countries absorb immigrants (Hanson and Slaughter 2002, Gandal et al. 2004). It may also answer the question of whether developing countries become more agricultural after opening to trade as suggested by the infant industry argument (Dodzin and Vamvakidis 2004). Furthermore, governments all over the world routinely talk about policies that may enhance the competitiveness of domestic industries.[1]

The factor proportions theory suggests that countries specialize according to their comparative advantage and that their comparative advantage depends on their factor endowments. A line of research pioneered by Harrigan (1995) has shown that there is a strong and robust empirical relation between factor endowments and industry outputs in cross-country data[2]. The coefficients of the factor-endowment variables (the Rybczynski coefficients henceforth) are typically statistically significant (Harrigan 2003), the R-squares of the regressions are high (0.7 ~ 0.8) and comparable to the gravity equations (Davis and Weinstein 1996, Bernstein and Weinstein 1998, 2002), and the empirical relation is robust to alternative regression specifications (Harrigan 1997, Fitzgerald and

[1] e.g. the 2006 American Competiveness Initiative by the U.S. (http://www.whitehouse.gov/stateoftheunion/2006/aci/#section2), the 2002 Manufacturing Strategy by the U.K. (http://www.dti.gov.uk/manufacturing/strategy.htm) and the 2005 Integrated Industrial Policy by the EU (http://europa.eu.int/comm/enterprise/enterprise_policy/industry/index_en.htm).

[2] Bernstein and Weinstein (1998, 2002) show that this empirical relation is weak in regional data. Leamer (1984, 1988) shows that there is a strong empirical relation between factor endowments and trade in cross-country data.

Hallak 2004). We refer to this line of research as the literature on cross-country specialization regressions.

However, while the factor proportions theory makes sharp predictions when there are two factors and two goods, it makes no clear prediction about how the endowment of a given factor affects the output of *a given industry* in the case of multiple goods and factors, even if the number of goods equals the number of factors. For example, it is possible that the most capital intensive industry contracts when the capital endowment increases (Harrigan 2003). This creates two problems. One, it is unclear how to interpret the Rybczynski coefficients and whether they are consistent with theory. To quote Davis and Weinstein (1996), "we feel that the [Rybczynski] coefficients cannot be interpreted". Two, it is unclear whether the robust empirical relation between factor endowments and industry outputs constitutes evidence for the factor proportions theory. For example, Fitzgerald and Hallak (2004) argue that this relation can also be due to non-homothetic preferences.

We show that in a setting with multiple goods and factors and Hicks-neutral and factor-augmenting productivity differences, the factor proportions theory makes a clear prediction about how the impacts of the endowment of a given factor *vary across industries*: they have positive co-variance with what we call the "relative uses" of this factor. The relative factor uses compare factor usage intensities across industries. For example, we say that an industry is skilled labor use heavy relative to unskilled labor if the ratio of skilled labor use to unskilled labor use of this industry exceeds the average of all the industries combined. Thus, the intuition for the positive co-variance prediction is that following an increase in, say, the skilled labor endowment, on average, the industries that use skilled labor a lot expand while those that use skilled labor a little bit contract. We call this prediction the “co-variation condition”.[3] Since the co-variation condition incorporates independent information on factor uses, factor endowments and outputs, it constitutes a direct test of the production side of the factor proportions theory, in the sense of Bowen et al. (1987) and Helpman (1999). As noted in the literature (Davis et al. 1997, Harrigan and Zakrajsek 2000), the main intellectual capital of the factor proportions theory is its general-equilibrium model of production.

To see whether the co-variation condition has empirical support, we estimate the Rybczynski coefficients using the regression specifications of Harrigan (1995) and Fitzgerald and Hallak (2004) and data on manufacturing output of 23 industries and factor endowments of 24 OECD countries from 5 years. We show that the Rybczynski coefficients, especially those of skilled and unskilled labor, have positive co-variance with relative factor uses. In other words, the output responses *vary across industries* in accordance with the factor uses by these industries, consistent with the factor proportions theory. As compared to the often disappointing evidence for the factor proportions theory (e.g. Leamer and Levinsohn 1995), our evidence is positive and encouraging.

Ours is not the only paper to have incorporated factor uses into the studies of specialization. Bernstein and Weinstein (2002) test whether the matrix of the Rybczynski coefficients and the matrix of the unit factor requirements multiply to the identify matrix, and fail to find empirical support for this prediction. A key difference of our approach is that the co-variation condition allows for Hicks-neutral and factor-augmenting differences in productivities across countries and factors, which have been found to be an

[3] Ethier (1984) also generalizes the Rybczynski theorem to multiple goods and factors, but this generalization does not incorporate the Rybczynski coefficients.

important feature of the data (e.g. Trefler 1993, 1995, Davis and Weinstein 2001). Thus our co-variation condition is likely to have a better chance of success with data. This may help explain why the co-variation condition finds empirical support while the prediction of Bernstein and Weinstein (2002) does not. On the other hand, Fitzgerald and Hallak (2004) look at the correlations that the Rybczynski coefficients have with the ratios of skilled labor uses to unskilled labor uses and capital-labor ratios. In comparison, we derive the co-variation condition explicitly from theory and take this theoretically appropriate prediction to the data. Finally, the co-variation condition is robust to the biases of the Rybczynski coefficient estimates that are common across industries for the same factor. This suggests that our co-variation condition cannot detect such biases and that we do not have much to add to the discussion of regression specifications (e.g. Harrigan and Zakrajsek 2000).

Recently, several authors (e.g. Schott 2003, Romalis 2004) test the multiple-cone version of the factor proportions theory using the data of both developed and developing countries. Our results show that the single-cone version of the factor proportions theory that accommodates Hicks-neutral and factor-augmenting productivity differences remains useful for understanding the pattern of specialization among OECD countries. Several other authors test very general predictions of the factor proportions theory (e.g. Choi and Krishna 2004, Bernhofen and Brown 2004). Our approach is built on the cross-country specialization regressions that are easy to implement and have the potential for being applied in the studies of specialization patterns across countries and for informing policy discussions.

In what follows, Section 2 shows the co-variation condition and its intuition, Section

3 examines whether the co-variation condition has empirical support, and Section 4 concludes.

Section 2 The Theory and Intuition of the Co-variation Condition

Let there be J industries and F factors in each country. Assume that factors are perfectly mobile across industries within each country but immobile across country borders and that international trade in goods is free. When factor markets clear in each country:

$$V^c = A^c y^c \tag{1}$$

where c indexes countries, V^c is an F x 1 vector of factor endowments, y^c a J x 1 vector of industry outputs, and A^c an F x J matrix of unit factor requirements. An element of A^c, a_{fj}^c, is the amount of factor f required to produce one unit of output in industry j for country c. Equation (1) says that factor supply (V^c) equals demand ($A^c y^c$). Assume that the production technology has constant returns to scale at the *industry* level. This is consistent with both the Heckscher-Ohlin model in which production is constant returns to scale and the monopolistic competition model (Krugman 1979) in which production has scale economies at the *firm* level. It follows that the unit factor requirements A^c depend only on factor prices and not on the industry outputs y^c.

Assume factor price equalization (FPE) across countries.[4] Then

$$A^c = A \text{ for all c.} \tag{2}$$

On the other hand, the cross-country specialization regressions (e.g. Harrigan 1995)

[4] For the purpose of regression (3), a weaker condition, Local Factor Price Insensitivity (Leamer and Levinsohn 1995), is sufficient (Hanson and Slaughter 2002). This condition says that small changes in factor endowments do not change factor prices.

estimate the following relation:[5]

$$y^c = BV^c, \tag{3}$$

where B is a J x F matrix. The Rybczynski Theorem shows that under the assumption of FPE, an element of B, b_{jf}, is the (total) response of industry j's output to a small change of the endowment of factor f; i.e. the changes in factor endowments can be fully absorbed by the changes in outputs without any changes in factor prices. From now on we will refer to the b_{jf}'s as the Rybczynski coefficients.

Plug equations (2) and (3) into equation (1):

$$V^c = (AB)\, V^c \quad \Rightarrow \quad AB = I, \tag{4}$$

where I is the F x F identity matrix. This is consistent with Bernstein and Weinstein (2002). Re-writing equation (4) as the equalities between the elements of the matrix AB and the matrix I,

$$\Sigma_j a_{fj} b_{jf} = 1 \qquad \text{for } f = 1, 2, \ldots F, \tag{5}$$

$$\Sigma_j a_{kj} b_{jf} = 0 \qquad \text{for } f, k = 1, 2, \ldots F \text{ and } k \neq f, \tag{6}$$

where j indexes industries and f and k index factors. Intuitively, equation (5) says that across industries, the Rybczynski coefficients of a given factor f, b_{jf}, have positive co-variation with the unit requirements of the factor f, a_{fj}, while equation (6) says that across industries, the Rybczynski coefficients b_{jf} have zero co-variation with the unit requirements of any other factor $k \neq f$, a_{kj}.

Divide both sides of equation (5) by $\Sigma_j a_{fj}$ and those of equation (6) by $\Sigma_j a_{kj}$ and then take the difference

[5] One may argue that theoretically, when J > F, the industry outputs are not deterministic so that equation (3) is not well specified; i.e. we need to assume that J = F. However, one can also argue that theoretically, even tiny disturbances (e.g. infinitesimally small transport costs) resolve the indeterminacy issue (see, e.g. Harrigan 2003) so that the J=F assumption is unnecessary.

$$R_{fk} \equiv \sum_j \left(\frac{a_{fj}}{\sum_j a_{fj}} - \frac{a_{kj}}{\sum_j a_{kj}}\right) b_{fj} = \frac{1}{\sum_j a_{fj}} \quad \text{for all f, k = 1, 2, ... F and f} \neq \text{k,} \qquad (7)$$

or

$$R_{fk} \equiv \sum_j \left(\frac{a_{fj}}{\sum_j a_{fj}} - \frac{a_{kj}}{\sum_j a_{kj}}\right) b_{fj} > 0 \quad \text{for all f, k = 1, 2, ... F and f} \neq \text{k.} \qquad (8)$$

Definition 1 Relative factor use: Let $r_{fkj} \equiv \frac{a_{fj}}{\sum_j a_{fj}} - \frac{a_{kj}}{\sum_j a_{kj}}$ be the relative use of factor f with respect to factor k for industry j. Factor k is the "reference factor". Industry j is "factor-f-use heavy relative to factor k" if $r_{fkj} > 0$.

To better understand the intuition of equation (8), note that:

$$r_{fkj} > 0 \Leftrightarrow \frac{a_{fj}}{a_{kj}} > \frac{\sum_j a_{fj}}{\sum_j a_{kj}} \qquad (9)$$

The left-hand side of equation (9) is the ratio of the unit requirement for factor f by industry j to the unit requirement for factor k, or the factor-f-usage intensity of industry j, using factor k as the reference. The right-hand side of equation (9) is the ratio of the sum of the unit requirements for factor f across industries to the sum of the unit requirements for factor k, and can be thought of as the average factor-f-usage intensity of all the industries combined, again using factor k as the reference. Then intuitively, industry j is factor-f-use heavy (relative to factor k) if and only if its factor-f-usage intensity (using factor k as the reference) exceeds the average factor-f-usage intensity.

For a given pair of factors f and k, the relative factor uses sum to 0 across industries:

$$\Sigma_j r_{fkj} = \sum_j \frac{a_{fj}}{\sum_j a_{fj}} - \sum_j \frac{a_{kj}}{\sum_j a_{kj}} = 0 \qquad (10)$$

Then the left-hand side of equation (8), R_{fk}, equals the co-variance between the

Rybczynski coefficients b_{fj} and the relative factor uses r_{fkj}. Therefore:

Proposition 1 The co-variation condition: across industries, the Rybczynski coefficients of a given factor have positive co-variance with the relative uses of this factor, no matter which other factor serves as the reference factor.

Proposition 1 provides a way to assess whether the variations of the Rybczynski coefficients across industries are consistent with the factor proportions theory. If the co-variation condition holds, intuitively, the output responses vary across industries in accordance with the factor uses by these industries: for the co-variance to be positive, on average, the industries that are, say, skilled labor use heavy, must have positive output responses following an increase in the skilled labor endowment; the industries that are not skilled labor use heavy must on average have negative output responses.

At this moment, it is useful to compare Proposition 1 and equation (8) with equation (7), a variant of the approach taken in Bernstein and Weinstein (2002). Suppose that the Rybczynski coefficients b_{jf} are known; we will discuss the implications of having to estimate b_{jf} in Corollary 2. A key difference is that Proposition 1 allows for Hicks-neutral and factor-augmenting differences in productivities across countries and factors whereas equation (7) does not. Such differences are an important feature of the data (e.g. Trefler 1993, 1995, Davis and Weinstein 2001) so that the physical units of factor endowments (e.g. the number of workers) may not accurately reflect their contributions to industry outputs. Therefore, a test based on Proposition 1 has a better chance with data than one based on equation (7).

To be specific, let π^c and π_f represent Hicks-neutral and factor-augmenting productivity differences. Then the relative factor uses are un-affected by π^c and π_f

because

$$r_{fkj}{}^{c} = \frac{\pi^{c}\pi_{f}a_{fj}}{\sum_{j}\pi^{c}\pi_{f}a_{fj}} - \frac{\pi^{c}\pi_{k}a_{kj}}{\sum_{j}\pi^{c}\pi_{k}a_{kj}} = \frac{a_{fj}}{\sum_{j}a_{fj}} - \frac{a_{kj}}{\sum_{j}a_{kj}}. \tag{11}$$

Thus equation (8) and Proposition 1 are also un-affected. Equation (7), however, is affected because it becomes

$$\sum_{j}\left(\frac{\pi^{c}\pi_{f}a_{fj}}{\sum_{j}\pi^{c}\pi_{f}a_{fj}} - \frac{\pi^{c}\pi_{k}a_{kj}}{\sum_{j}\pi^{c}\pi_{k}a_{kj}}\right)b_{fj} = \frac{1}{\sum_{j}\pi^{c}\pi_{f}a_{fj}}$$

$$\Leftrightarrow \sum_{j}\left(\frac{a_{fj}}{\sum_{j}a_{fj}} - \frac{a_{kj}}{\sum_{j}a_{kj}}\right)b_{fj} = \frac{1}{\sum_{j}\pi^{c}\pi_{f}a_{fj}}. \tag{12}$$

Equations (11) and (12) imply that whereas a test based on equation (7) needs to explicitly estimate Hicks-neutral and factor-augmenting differences in productivities, π^c and π_f, a test based on Proposition 1 is robust to such productivity differences. We summarize this advantage of Proposition 1 as follows.

Corollary 1 Suppose the co-variation condition holds for the Rybczynski coefficients b_{jf} and the unit factor requirements a_{fj}. Then it also holds for $\pi^c\pi_f a_{fj}$ and b_{fj}, where π^c and π_f are positive numbers.

Corollary 1 also alleviates the measurement errors in the data on unit factor requirements because π^c and π_f can also represent country- and factor-specific measurement errors that take the multiplicative form.

Proposition 1 has another corollary:

Corollary 2 Suppose the co-variation condition holds for the Rybczynski coefficients b_{jf} and the unit factor requirements a_{fj}. Then it also holds for $b_{jf} + d_f$ and a_{fj}, where d_f is a positive or negative number.

Corollary 2 holds because d_f does not vary across industries and so affects only the mean of b_{jf} and not its co-variance with the relative factor uses r_{fkj}. Thus Corollary 2 says that the co-variation condition cares only about the variation of the Rybczynski coefficients and not about their levels and so is robust to the bias of the estimated Rybczynski coefficients that may differ for different factors but is common across industries for the same factor. However, Corollary 2 also shows the limitation of the co-variation condition: if the Rybczynski coefficient estimates contain a common bias, the co-variation condition cannot detect it. In other words, the co-variation condition cannot assess whether a given regression specification produces consistent estimates of the Rybczynski coefficients.

Section 3 Data, Estimation and Results

3.1 Data

To test the co-variation condition, we need estimates of the Rybczynski coefficients and data on the unit factor requirements. For the former, we have output (value added) data for 23 2-digit ISIC revision 3 manufacturing industries of 24 OECD countries from UNIDO[6] and data on GDP and the endowments of capital, skilled labor, unskilled labor and arable land of these countries from OECD and the Penn World Tables. Both sets of data are for 1991, 1995 and 1998 ~ 2000. For the latter, we have the U.S. data on the total uses of capital, skilled labor and unskilled labor by industry put together by Feenstra and Hanson (2000). This dataset covers 371 U.S. manufacturing industries and is constructed using the 1982 US Input-Output Table and the NBER-CES Manufacturing Industry

[6] We focus on manufacturing industries because factor use data on non-manufacturing industries is difficult to come by. For the OECD countries the UNIDO data comes from OECD.

Database.[7] Following the literature (e.g. Berman, Bound and Griliches 1994) we classify production workers as unskilled labor and non-production workers as skilled labor.[8] Table 1 shows the summary statistics of our variables. Table A1 shows the industry descriptions and the Data Appendix describes our data sources and variable construction.

3.2 Regression Specifications

To estimate the Rybczynski coefficients we follow the literature and estimate the following two regressions for each industry j:

$$y_{ctj} = \alpha_j + \beta_{Kj}K_{ct} + \beta_{Sj}SL_{ct} + \beta_{Uj}UL_{ct} + \beta_{Dj}D + \varepsilon_{ctj} \tag{13}$$

$$\frac{y_{ctj}}{GDP_{ct}} = \alpha_j + \beta'_{Kj}\frac{K_{ct}}{L_{ct}} + \beta'_{Sj}\frac{SL_{ct}}{L_{ct}} + \beta'_{Dj}\frac{D_{ct}}{L_{ct}} + \varepsilon_{ctj} \tag{14}$$

where y stands for industry output (value added), K, SL, UL, D and L are the capital, skilled labor, unskilled labor, arable land and total labor (skilled plus unskilled labor) endowments, c indexes countries and t indexes years. Regression (13) is motivated by equation (3) and has been widely used in the literature (e.g. Harrigan 1995) while regression (14) is proposed by Fitzgerald and Hallak (2004) as a reduced form relation between industry outputs and countries' factor endowments. The β's and β''s are the Rybczynski coefficient estimates.[9] Fitzgerald and Hallak (2004) argue that (13) fails to

[7] Hakura (2001) has put together total-factor-use data of four European countries for 1970 and 1980. We use the Feenstra and Hanson (2000) data because it is readily available to us. The Davis and Weinstein (2001) total-factor-use data, on the other hand, has only two factors, capital and labor.

[8] Berman, Bound, and Griliches (1994) show that the production/non-production worker distinction closely mirrors the distinction between blue- and white-collar occupations, which, in turn, closely reflect an educational classification of high school/college. Krueger (1997) shows that the raw correlation between average education and the share of production workers is –0.61 at the 1980 three-digit Census Industry Classification level.

[9] Fitzgerald and Hallak (2004) argue that the β's of equation (14) have Rybczynski interpretations. As it is unclear how to convert these β's into Rybczynski coefficients, we assume that they vary across industries in the same way as the Rybczynski coefficients.

take cross-country Hicks-neutral productivity differences into account and might produce biased Rybczynski coefficient estimates, and that (14) addresses this issue. Since the co-variation condition is silent about the consistency of the Rybczynski coefficient estimates (Corollary 2) we do not have much to add to the discussion of regression specifications;[10] rather, we use the estimates from both (13) and (14). For both specifications we use the White heteroscedasticity-consistent standard errors.

3.3 Results

We first estimate regression (13) for each of our 23 industries, and this produces many Rybczynski coefficient estimates. Rather than reporting them all, we summarize their pattern in Table 2 and compare this pattern with the literature. In the comparison we focus on Fitzgerald and Hallak (2004) since their results are representative of the literature. The most consistent finding in the literature is that most capital coefficients are positive (to quote Harrigan 2003, "capital is manufacturing's friend"), and this pattern shows up in our results, too: among the 23 capital coefficients, 20, or 87%, are positive, and among the 21 capital coefficients that are statistically significant, 19, or 90%, are positive. As shown in columns 1 and 2 of Table 2, this pattern closely matches Fitzgerald and Hallak (2004).[11] We also find that most skilled labor coefficients are positive (87% of the cases) and most unskilled labor coefficients are negative (74% of the cases), and again, this pattern is consistent with Fitzgerald and Hallak (2004). However, unlike

[10] We have not used the specification of Harrigan and Zakrajsek (2000) because a key part of their specification is the use of country fixed effects and this is hard for us to implement since we have only 5 years of data.
[11] A larger fraction of our Rybczynski coefficient estimates are statistically significant than in Fitzgerald and Hallak (2004) since we have more observations.

Fitzgerald and Hallak (2004) most of our arable land coefficients are positive (65% of the cases).

It is straightforward to use these Rybczynski coefficient estimates and our data on factor uses to calculate the co-variance between the Rybczynski coefficients and the relative factor uses, R_{fk} (i.e. the left-hand side of equation 8). The unit factor requirements are calculated as the ratios of total factor uses to value added for 1991. Since R_{fk} is a linear combination of the Rybczynski coefficient estimates, its standard error can be calculated using the variance-covariance matrix of these estimates.[12] Since we have factor use data for three factors, capital, skilled labor and unskilled labor, equation (8) suggests that we calculate six co-variances, one for each pair of factors.

Table 3 reports these co-variances between the Rybczinski coefficients and the relative factor uses. The co-variance between the unskilled labor coefficients, β_U, and the relative unskilled labor uses is positive and statistically significant (37796.4 with a standard error of 8003.86) when skilled labor is the reference factor. This pattern of positive co-variance can also be seen in Figure 1, which plots β_U against the relative unskilled labor uses with respect to skilled labor. That is to say, the industries that use unskilled labor heavily (relative to skilled labor) tend to have positive output responses following an increase in the unskilled-labor endowment. This is consistent with Proposition 1. Similarly, the the capital coefficients, β_K, have a positive and significant co-variance with the relative capital uses with respect to unskilled labor, and the skilled-

[12] The industry-by-industry OLS we run does not produce the estimates of the off-diagonal elements of the variance-covariance matrix. Ideally we would obtain this information by estimating all 23 industries as a system of seemingly-unrelated regressions (SUR), yet doing so requires more degrees of freedom than we have in our data. Therefore for industries j and m and factor f, we obtain the estimates of the variance of b_{jf} and b_{mf} from industry-by-industry OLS and the estimate of the co-variance between b_{jf} and b_{mf} from the SUR with industries j and m. Because we have the same regressors for industries j and m, the SUR and industry-by-industry OLS produce identical estimates (e.g. Greene 1997).

labor coefficients, β_S, have positive and significant co-variances with the relative skilled-labor uses when either capital or unskilled labor is the reference factor. Again, these are consistent with Proposition 1.

On the other hand, the covariance between β_U and the relative unskilled-labor uses with respect to capital is positive but insignificant, and the co-variance is negative and insignificant for β_K and the relative capital uses with respect to skilled labor. Figures 2 and 3 show the culprit for these results. Figure 2 plots β_U against the relative unskilled labor uses with respect to capital, and Figure 3 plots β_K against the relative capital uses with respect to skilled labor. In both cases, industry 23, coke, petroleum and nuclear fuel, is the clear outlier to the otherwise positive co-variation between the Rybczynski coefficients and the relative factor uses. Since industry 23 is especially resource intensive and we do not have data for natural resources, we drop industry 23 from our sample. Doing so results in positive and significant co-variances for all the six pairs of factors.

In summary, the results using regression (13) suggest positive co-variances between the Rybczynski coefficients and the relative factor uses. This is consistent with the factor proportions theory.

We also estimate regression (14) for each of our 23 industries, and columns 3 and 4 of Table 2 compare the pattern of our results with Fitzgerald and Hallak (2004). Our pattern matches theirs. Unlike the results of regression (13), now most capital coefficients are negative (14 out of 23 cases, or 61%). Similar to the results of regression (13), most skilled labor coefficients are positive (15 out of 23, or 65%). Furthermore, the pattern of our arable land coefficients matches Fitzgerald and Hallak (2004) better. On the other hand, as Table 3 shows, now the co-variation condition holds for the skilled labor coefficients $\beta_s^{'}$ when either capital (see also Figure 4) or unskilled labor is the reference factor. For the capital coefficients $\beta_K^{'}$, however, the co-variance is negative and insignificant whether the reference factor is skilled labor or unskilled labor. Finally, to get the unskilled labor coefficients that regression (14) implies, note that L = SL + UL so that (14) can be re-written as:

$$\frac{y_{ct}}{GDP_{ct}} = (\alpha + \beta_S^{'}) + \beta_K^{'} \frac{K_{ct}}{L_{ct}} - \beta_S^{'} \frac{UL_{ct}}{L_{ct}} + \varepsilon_{ct}$$

Thus the unskilled labor coefficients are $\beta_U^{'} = -\beta_S^{'}$. As shown in Table 3, $\beta_U^{'}$ has a positive co-variance with relative unskilled labor uses that is significant when skilled labor is the reference factor but insignificant when capital is the reference factor. Figure 5 plots $\beta_U^{'}$ against the relative unskilled labor uses with respect to capital and shows that the insignificance of the co-variance is again driven by industry 23, like in the case of regression (13). Not surprisingly, this insignificance goes away when we drop industry 23, although the co-variances between $\beta_K^{'}$ and the relative capital uses remain negative and insignificant. Overall, the results using regression (14) also suggest positive co-variances between the Rybczynski coefficients and the relative factor uses. This is consistent with the factor proportions theory.

We have also done the following robustness exercises for regression (13); to save space we do not report these results and they are available from us upon request.[13] First, our data is an unbalanced panel and we have much better data coverage for 1998~2000 (see Table A2). When we use only data from 1998~2000 we obtain similar results. In

[13] We have also done the first exercise for regression (14) and got similar results. On the other hand, regression (14) has industry outputs divided by GDP as the dependent variable and so is robust to the issues addressed by the second and the last exercises.

particular, when we drop industry 23 the co-variances are positive and significant for all the six pairs of factors. Second, we follow the literature (e.g. Harrigan 1995) and weigh each observation by the inverse of GDP to better control for size-related heteroscedasticity. We run the regression both with and without 23 in the sample. The results are similar. Finally, industry outputs may fluctuate over time due to business cycles and such fluctuations may not be related to changes in factor endowments. To control for this issue we introduce time fixed effects.[14] Again, the results are similar.

Section 4 Conclusion

We show that in the case of multiple goods and factors, the factor proportions theory predicts the following co-variation condition: across industries, the impacts of a given factor on industry outputs have positive co-variance with the relative uses of this factor. The intuition is that on average, the industries that use a given factor heavily have positive output responses following an increase in the endowment of this factor. A key innovation of our approach is that the co-variation condition accommodates Hicks-neutral and factor-augmenting productivity differences. Since the co-variation condition incorporates independent information on factor uses, factor endowments and outputs, it constitutes a direct test of the production side of the factor proportions theory.

We then estimate the impacts of factor endowments on industry outputs using cross-country specialization regressions. We show that the positive co-variance predicted by theory finds empirical support, especially for unskilled labor and skilled labor coefficients. This is evidence that the factor proportions theory is at work behind the

[14] Since we only have 5 years of data we do not consider country fixed effects. See also note 10.

robust empirical relation between industry outputs and factor endowments in cross-country data.

Reference:

Acemoglu, Daron and Jaume Ventura, 2002. “The World Income Distribution”, Quarterly Journal of Economics, May 2002, 659-694.

Bartelsman, Eric J. and Wayne Gray. 1996. "The NBER Manufacturing Productivity Database." NBER Technical Paper #205.

Berman, Eli, John Bound, and Zvi Griliches. 1994. "Changes in the Demand for Skilled Labor with U.S. Manufacturing: Evidence from the Annual Survey of Manufactures." Quarterly Journal of Economics, 109, pp. 367-97.

Bernhofen, Daniel M. and John C. Brown. 2004. “A Direct Test of the Theory of Comparative Advantage: the Case of Japan”, Journal of Political Economy, 112(1), 48 – 67.

Berstein, Jeffrey R. and David E. Weinstein. 1998. "Do Endowments Predict the Location of Production? Evidence from National and International Data.". NBER Working Papers 6815.

Berstein, Jeffrey R. and David E. Weinstein. 2002. "Do Endowments Predict the Location of Production? Evidence from National and International Data." Journal of International Economics, 56:1, pp. 55-76.

Bowen, Harry P., Edward E. Leamer and Leo Sveikauskas, “Multicountry, Multifactor Tests of the Factor Abundance Theory”, American Economic Review LXXVII (1987), 791 – 809.

Choi, Yong-Seok and Pravin Krishna. 2004. “The Factor Content of Bilateral Trade: An Empirical Test”. Journal of Political Economy, 112(4), 887 – 914.

Davis, Donald and David E. Weinstein. 1996. "Does Economic Geography Matter for International Specialization?" NBER Working Paper # 5706.

Davis, Donald and David E. Weinstein. 2001. "An Account of Global Factor Trade." American Economic Review, 91:5, pp. 1423 - 53.

Davis, Donald R., David E. Weinstein, Scott C. Bradford and Kazushige Shimpo. “Using International and Japanese Regional Data to Determine When the Factor Abundance Theory of Trade Works”, American Economic Review 87 (1997), 421 – 446.

Dodzin, Sergei and Athanasios Vamvakidis. 2004. “Trade and Industrialization in Developing Economies”. Journal of Development Economics, October 2004, 319-328.

Ethier, Wilfred J.,1984. “Higher Dimensional Issues in Trade Theory” in Handbook of International Economics, 1984, pp. 131-84, Volume 1. Handbooks in Economics series, no. 3 New York; Amsterdam and Oxford: North-Holland; distributed in U.S. and Canada by Elsevier Science, New York

Feenstra, Robert and Gordon Hanson. 2000. "Aggregation Bias in the Factor Content of Trade: Evidence from U.S. Manufacturing." American Economic Review, 90:2, pp. 155-60.

Fitzgerald, Doireann and Juan Carlos Hallak. 2004. "Specialization, Factor Accumulation and Development." Journal of International Economics, 64:2, pp. 277 - 302.

Gandal, Neil, Gordon H. Hanson and Matthew J. Slaughter, 2004. “Technology, Trade, and Adjustment to Immigration in Israel”. European Economic Review 48, 403 – 428.

Greene, William H. 1997. Econometric Analysis. Upper Saddle River: Prentice-Hall, Inc.

Hakura, Dalia S. 2001. "Why Does HOV Fail? The Role of Technological Differences within the EC." Journal of International Economics, 54, pp. 361-82.

Hakura, Dalia S. 2002. "Erratum to "Why Does HOV Fail? The Role of Technological Differences within the EC" [Journal of International Economics 54 (2001) 361-382]." Journal of International Economics, 56, pp. 509-11.

Hanson, Gordon and Matthew Slaughter. 2002. "Labor-Market Adjustment in Open Economies: Evidence from US States." Journal of International Economics, 57:1, pp. 3-29.

Harrigan, James. 1995. "Factor Endowments and the International Location of Production: Econometric Evidence from the OECD, 1970 - 1985." Journal of International Economics, 39:1-2, pp. 123-41.

Harrigan, James. 1997. "Technology, Factor Supplies, and International Specialization: Estimating the Neoclassical Model." American Economic Review, 87:4, pp. 475-94.

Harrigan, James. 2003. "Specicalization and the Volume of Trade: Do the Data Obey the Laws?," in Handbook of International Trade. James Harrigan and E. Kwan Choi eds. Malden, MA: Blackwell Publishing Ltd., pp. 85 - 118.

Harrigan, James and Egon Zakrajsek. 2000. "Factor Supplies and Specialization in the World Economy." Board of Governors of the Federal Reserve System (U.S.), Finance and Economics Discussion Series: 2000-43.

Helpman, Elhanan. 1999. “The Structure of Foreign Trade”. Journal of Economic Perspectives, Spring 1999, v. 13, iss. 2, pp. 121-44

Krueger, Alan B. 1997. "Labor Market Shifts and the Price Puzzle Revisited." Industrial Relations Section Working Paper no. 375, Princeton University.

Leamer, Edward E. 1984. Sources of International Comparative Advantage: Theory and Evidence. MIT Press, Cambridge.

Leamer, Edward E. 1988. “Cross-section Estimation of the Effects of Trade Barriers”, in Feenstra, Robert C. (Ed.), Empirical Methods for International Trade. MIT Press, Cambridge, pp. 51-82.

Leamer, Edward E. and James Levinsohn. 1995. "International Trade Theory: The Evidence," in Handbook of International Economics. Gene M. Grossman and Kenneth Rogoff eds. The Netherlands: North-Holland, pp. 1339-94.

Redding, Stephen. 2002. “Specialization Dynamics”. Journal of International Economics, December 2002, v. 58, iss. 2, pp. 299-334

Romalis, John. "Factor Proportions and the Structure of Commodity Trade." American Economic Review, March 2004, 94(1), 67 – 97.

OECD. “Classifying Educational Programmes: Manual for ISCED-97 Implementation in OECD countries”, 1999 Edition.

Schott, Peter K. 2003. "One Size Fits All? Theory, Evidence and Implications of Cones of Diversification." American Economic Review, 93:3, pp. 686-708.

Trefler, Daniel. 1993. “International Factor Price Differences: Leontief Was Right!” Journal of Political Economy 101 (December), 961-987.

Trefler, Daniel. 1995. “The Case of the Missing Trade and Other Mysteries”. American Economic Review 85 (December), 1029-1046.

United Nations. “Statistical Yearbook.” Various issues.

US Census Bureau. "Economic Census 1992-Disc 1J (CD-EC92-1J)."

US Census Bureau. 1993. "Current Industrial Report." Various Issues.

Data Appendix

A. The Industry Variables

The value added data of 4-digit ISIC revision 3 manufacturing industries for 24 OECD countries comes from UNIDO (INDSTAT4 2004). We aggregate the data to 2-digit levels so that it is consistent with our factor use data. This leaves us with 23 industries. Our choice of the years 1991, 1995, 98 ~ 2000 is dictated by the availability of the skilled and unskilled labor endowment data. Table A1 shows the list of our industries and Table A2 shows our data coverage by tabulating the number of observations a given country has for a given year.

The data on the total factor uses of 371 US manufacturing industries comes from Feenstra and Hanson (2000). This classification (FH2000 henceforth) is based on 1972 4-digit US SIC but at a more aggregated level. We construct the concordance from FH2000 to 2-digit ISIC revision 3 using the concordance from 1972 4-digit US SIC to 1987 4-digit US SIC (source: http://www.nber.org), the concordance from 1972 4-digit US SIC to FH2000 (source: Feenstra and Hanson 2000), and the concordance from 4-digit 1987 US SIC to 2-digit ISIC revision 3 discussed below.

The data on the 4-digit 1987 US SIC manufacturing industries comes from the NBER-CES Manufacturing Database. The standard concordance from 4-digit 1987 US SIC to 4-digit ISIC revision 3 (source: http://www.haveman.org) is difficult to work with because a 4-digit 1987 US SIC industry is often matched to many 4-digit ISIC revision 3 industries (e.g. industry 3559 is matched to 11 ISIC industries and industry 3999 matched to 18) and the concordance does not provide the shares of the ISIC industries. Therefore we first aggregate the 4-digit ISIC industries to 2 digits. Next, when the standard concordance maps a SIC industry into 2 or more 2-digit ISIC industries, we match all the 7-digit 1987 US SIC products of the SIC industry to the 2-digit ISIC industries and then use the 2-digit ISIC industries' output shares for concordance. A 2-digit ISIC industry's output share equals the sum of the outputs of all the products that are matched to it divided by the aggregate output of the SIC industry, and our product-level output data comes from 1992 Census of Manufactures and 1992 Current Industrial Survey. We note that the 4-digit 1987 US SIC industries 2999 "Petroleum and Coal Products n.e.c." and 2411 "Logging" are mapped into the *non-manufacturing* ISIC industries 10 "Mining of coal and lignite; extraction of peat" and 02 "Forestry, logging and related service activities" in the standard concordance. Thus their ISIC industry shares do not sum to one in our concordance. This concordance is available from us upon request.

B. The Country Variables

We calculate the real capital stocks by the perpetual inventory method, which cumulates past and present gross investment series assuming a fixed depreciation rate. Following Hummels and Levinsohn (1995) we assume that the initial capital stock for each country equals 250 percent of its GDP and adopt the constant depreciation rate of 13.3 percent.

We set the benchmark year to 1970. The data on base year GDP and gross investment comes from the Penn World Tables (version 6.1).

The data on GDP and skilled and unskilled labor endowments comes from the Penn World Tables (version 6.1) and OECD. We construct the skilled and unskilled labor endowment data in the following way. Let e_{itc} denote the fraction of the population with education level i in year t for country c and f_{itc} denote the fraction of the population who have education level i but are *not* in the labor force in year t for country c. Then the share of the labor force with education level i in year t for country c, s_{ict}, equals $l_{ict}/\sum_i l_{ict}$, where $l_{itc} = e_{itc}(1 - f_{itc})$. Let L_{tc} denote the labor force of country c in year t, and H and U denote the sets of education levels that correspond to skilled and unskilled labor. Then the skilled and unskilled labor endowments of country c in year t equal $L_{ct}\sum_{i\in H} s_{ict}$ and $L_{ct}\sum_{i\in U} s_{ict}$, respectively. We obtain e_{itc} and f_{itc} from OECD's Education at a Glance 2004 (Table A3.4a and Table A10.2c), and we obtain L_{tc} from OECD's Labor Market Statistics. The population is those between ages 25 and 64. The data on e_{itc} and f_{itc} is available for 1991, 1995 and 1998 ~ 2000 and distinguishes three levels of education: A. below-upper-secondary education (ISCED 0, pre-primary education, ISCED 1, primary education, and ISCED 2, lower secondary education); B. upper-secondary-and-post-secondary-but-non-tertiary education (ISCED 3, upper secondary education, and ISCED 4, post-secondary non-tertiary education, including vocational or technical education); C. tertiary education (ISCED 5 and 6, first and second stages of tertiary education). We choose the unskilled labor set U to include all those with level A plus half of those with level B. We choose the skilled labor set H to include half of those with level B plus all those with level C.

The arable land data comes from the United Nations' Statistical Yearbook.

Table 1 Summary Statistics

Variables	Mean	Std. Deviation	Min.	Max.
The Industry variables				
Value added	7.19E+09	1.89E+10	0	2.32E+11
Total capital use	85577.69	60909.37	168.935	243257
Total skilled-labor use	361.6058	240.8719	0.59448	967.7794
Total unskilled-labor use	856.7736	491.3138	1.98445	1980.313
The Country variables				
GDP	8.10E+11	1.37E+12	5.05E+10	8.77E+12
Capital endowment	1.25E+09	2.01E+09	6.05E+07	1.17E+10
Skilled labor endowment	7104.894	12042.16	340.193	72756.43
Unskilled labor endowment	8119.9	8465.709	676.607	40900.11
Arable land endowment	14588.1	28453.14	413	177232

Notes: Value added and GDP are in US dollars. Capital endowment is in thousands of US dollars and capital use is in millions of US dollars. Skilled-labor and unskilled labor uses are in thousands of non-production and production workers. Skilled and unskilled labor endowments are in thousands of employees. Arable land endowment is in thousands of hectares.

Table 2 Patterns of Rybczynski Coefficients

	Regression (13)		Regression (14)	
	Our Results	FH (2004)	Our Results	FH (2004)
	1	2	3	4
Capital				
Positive	87% (20 of 23)	88% (22 of 25)	39% (9 of 23)	20% (5 of 25)
Positive of all significant	90% (19 of 21)	100% (10 of 10)	33% (2 of 6)	0% (0 of 5)
Skilled Labor				
Positive	87% (20 of 23)	64% (16 of 25)	65% (15 of 23)	76% (19 of 25)
Positive of all significant	100% (9 of 9)	75% (3 of 4)	71% (10 of 14)	100% (8 of 8)
Unskilled Labor				
Positive	26% (6 of 23)	40% (10 of 25)	-	-
Positive of all significant	8% (1 of 13)	29% (2 of 7)	-	-
Arable Land				
Positive	65% (15 of 23)	8% (2 of 25)	26% (6 of 23)	36% (9 of 25)
Positive of all significant	71% (10 of 14)	0% (0 of 4)	23% (3 of 13)	0% (0 of 0)

Notes: FH (2004) refers to Fitzgerald and Hallak (2004). "Positive" is the fraction of positive coefficients among all the coefficients, and "positive of all significant" is the fraction of positive coefficients among all the coefficients that are statistically significant.

Table 3 Co-Variance between Rybczynski Coefficients and Relative Factor Uses

Ryb. Coefficients	β_U or β'_U		β_K or β'_K		β_S or β'_S	
	1	2	3	4	5	6
Reference Factor	capital	skilled labor	skilled labor	unskilled labor	unskilled labor	capital
Reg. (13)						
Co-variance	12792.1	**37796.4**	-0.107	**0.33**	**84216.6**	**20188.5**
Std. Error	8223.43	**8003.9**	0.146	**0.14**	**23531.5**	**2453.9**
Reg. (14)						
Co-variance	1.02	**2.76**	-0.0065	-0.0107	**2.76**	**1.74**
Std. Error	0.78	**0.81**	0.005	0.0058	**0.81**	**0.64**
Reg. (13) without 23						
Co-variance	**38841.43**	**42498.74**	**0.48**	**1.01**	**82895.99**	**53625.3**
Std. Error	**8239.93**	**7922.78**	**0.12**	**0.14**	**24146.29**	**18610.38**
Reg. (14) without 23						
Co-variance	**1.46**	**2.82**	-0.0012	-0.0046	**2.82**	**1.36**
Std. Error	**0.51**	**0.65**	0.0029	0.0035	**0.65**	**0.38**

Notes: β_U, β_K and β_S are the estimated Rybczynski coefficients of unskilled labor, capital and skilled labor. The "co-variance" cells report the co-variance between the Rybczynski coefficients and relative factor uses, $R_{fk} = \Sigma_j r_{fkj}\beta_{fj}$ (i.e. the left hand side of equation 9) for f, k = unskilled labor, capital and skilled labor and $f \neq k$ (j = 1, 2, ... J indexes industries). The "Std. Error" cells report the standard errors of R_{fk}, $\sqrt{r_{fk}' B_f r_{fk}}$, where B_f is the J x J variance-covariance matrix of the Rybczynski coefficient estimates (β_{f1}, β_{f2}, ... β_{fJ}) and r_{fk} is the J x 1 vector (r_{fk1}, r_{fk2}, ... r_{fkJ})'. The R_{fk}'s that are positive and statistically significant are in boldface.

Table A1 List of Industries

Industries	Short Descriptions
15	Food and beverages
16	Tobacco
17	Textiles
18	Apparel and Fur
19	Leather, luggage and footwear
20	Wood, straw and plaiting materials
21	Paper
22	Recorded Media
23	Coke, petroleum and nuclear fuel
24	Chemicals and related products
25	Rubber and plastics
26	Non-metallic minerals
27	Basic metals
28	Fabricated metal products
29	Miscellaneous machinery
30	Office, accounting and computing machinery
31	Electrical machinery
32	Radio, TV and communication equipments
33	Medical and optical instruments
34	Motor vehicles
35	Other transportation equipments
36	Furniture and miscellaneous manufacturing
37	Recycling

Table A2 Data Coverage

country	1991	1995	1998	1999	2000
Australia	0	18	18	18	18
Austria	0	21	21	20	20
Belgium-Luxembourg	0	0	0	23	0
Canada	21	21	21	21	21
Denmark	0	21	20	20	21
Finland	23	21	22	22	22
France	20	20	21	21	21
Germany	0	0	22	22	22
Greece	0	23	23	0	0
Hungary	0	0	23	23	23
Ireland	21	21	18	18	0
Italy	0	23	23	23	23
Japan	0	0	23	23	23
Korea	23	23	23	23	23
Mexico	0	22	22	22	22
Netherlands	0	18	17	21	0
Norway	19	22	22	22	22
Portugal	0	0	23	23	23
Spain	0	23	23	23	23
Sweden	23	22	23	22	18
Switzerland	0	0	20	20	19
Turkey	0	21	22	22	22
United Kingdom	0	23	22	22	22
United States	0	0	22	22	0
Total	150	363	494	496	408

Notes: The cells report the numbers of available observations by country by year.

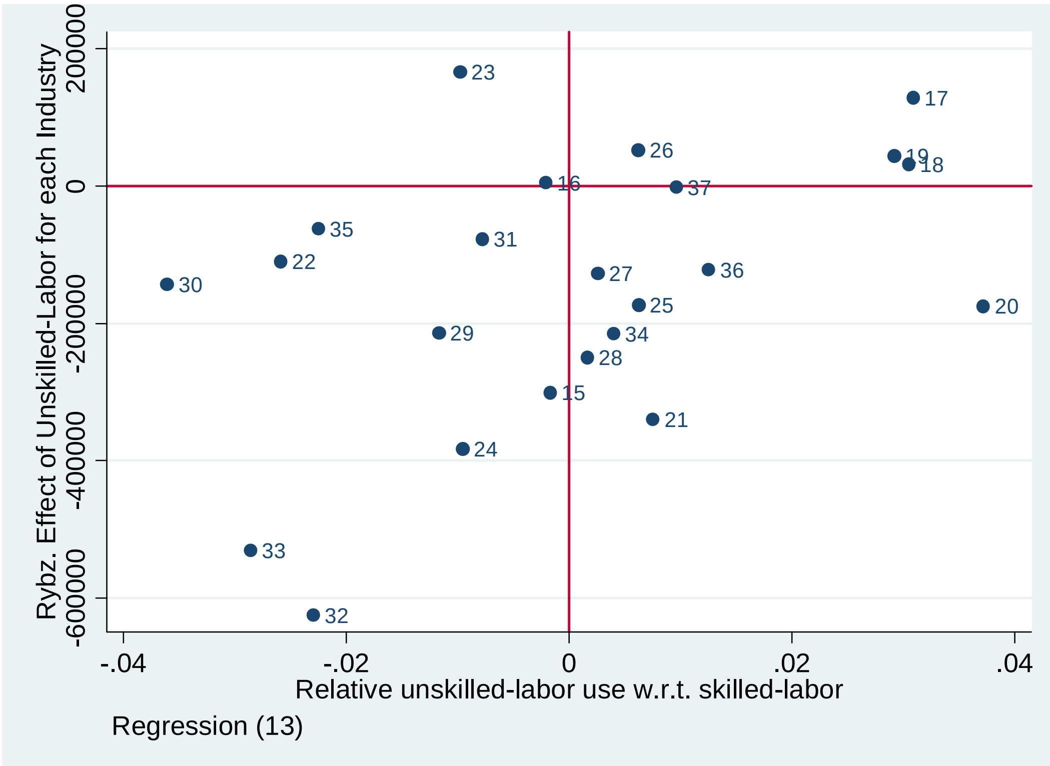


Figure 1

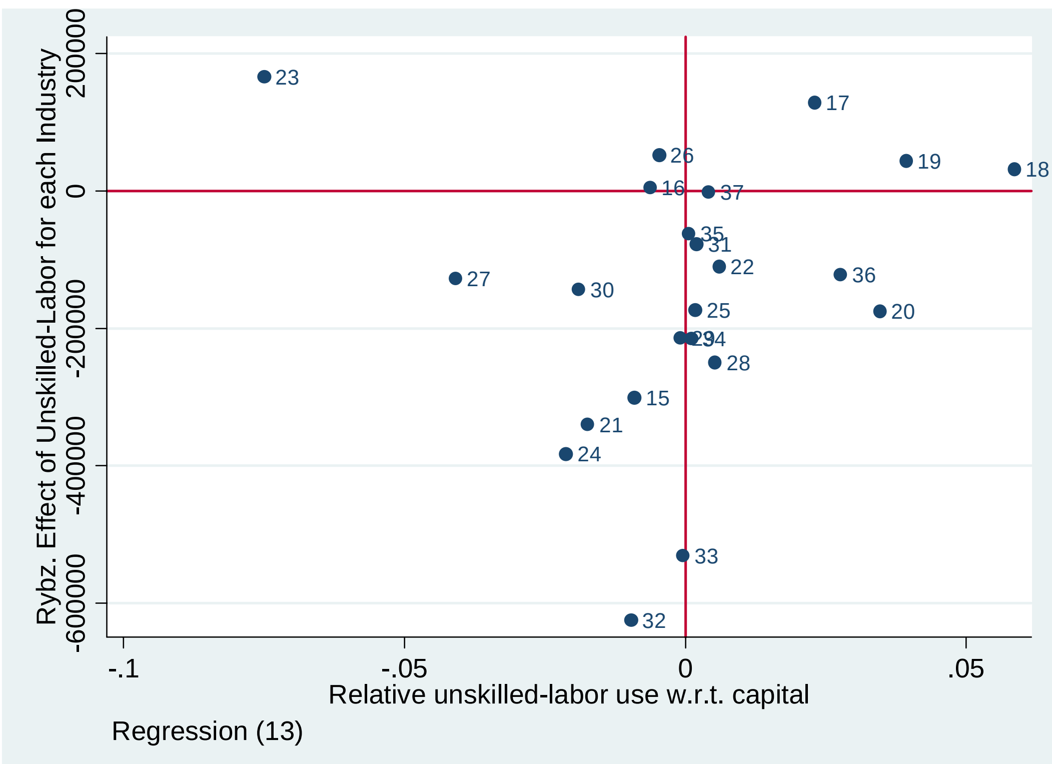


Figure 2

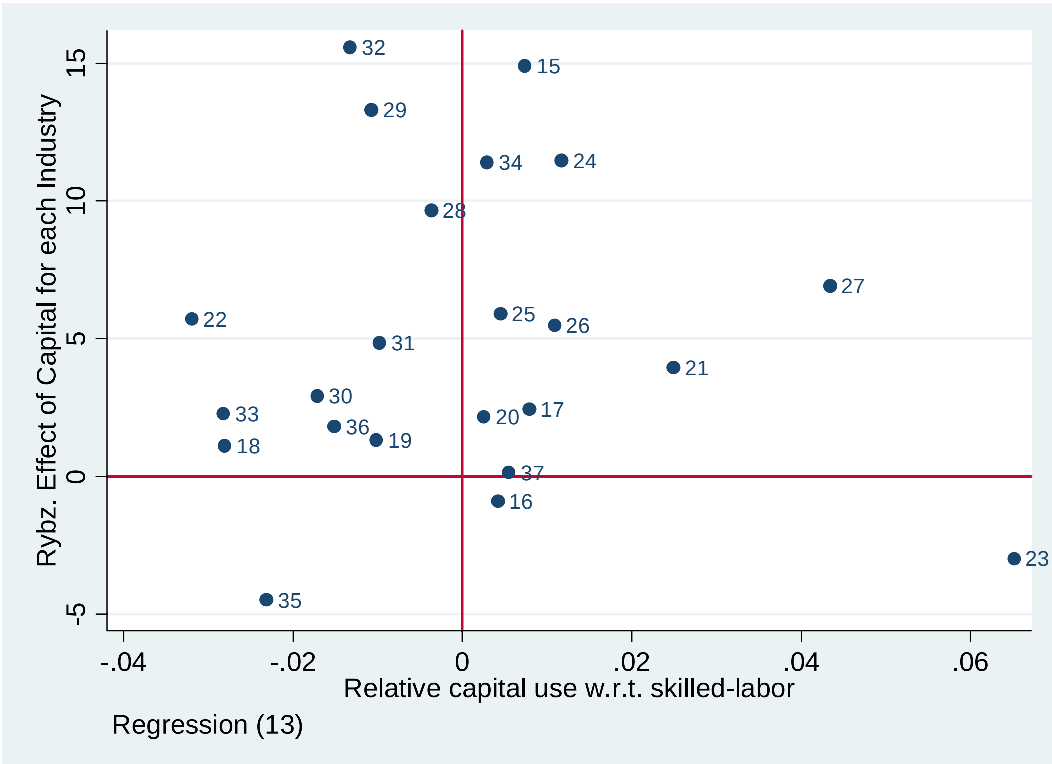


Figure 3

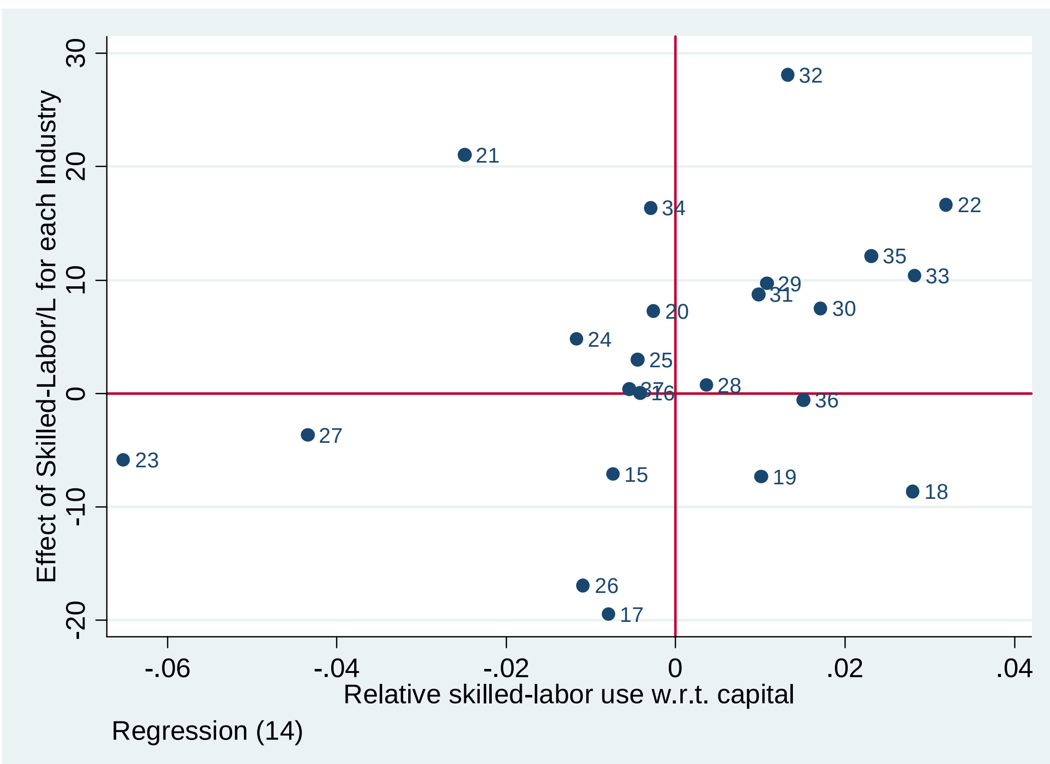
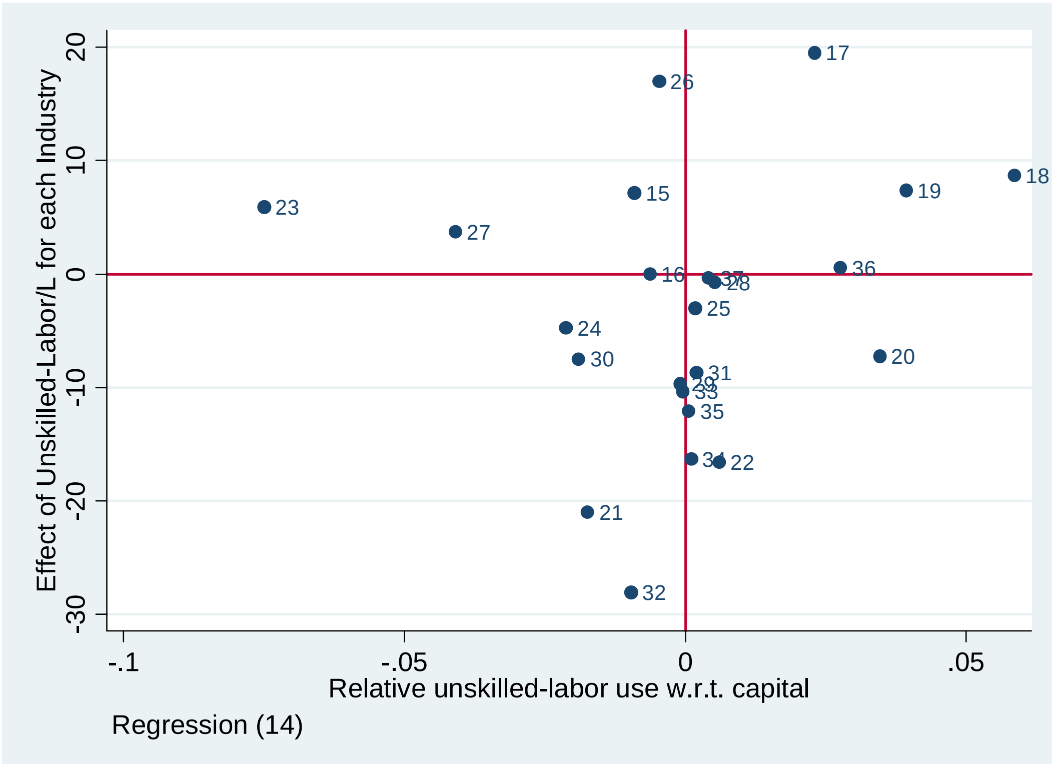


Figure 4



Effect of Unskilled-Labor/L for each Industry
-30
-20
-10
0
10
20
-.1
-.05
0
.05
Relative unskilled-labor use w.r.t. capital
Regression (14)
17
26
18
19
15
23
27
36
16
37
28
25
24
20
30
31
29
33
35
34
22
21
32

Figure 5