NBER WORKING PAPER SERIES

INCREASING RETURNS AND ALL THAT: A VIEW FROM TRADE

Werner Antweiler
Daniel Trefler

Working Paper 7941
http://www.nber.org/papers/w7941

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2000

We have benefited from the comments of Ken Arrow, Don Davis, Mel Fuss, Gene Grossman, Elhanan Helpman, Dale Jorgenson, Paul Krugman, Richard Lipsey, Angelo Melino, Ariel Pakes, Nathan Rosenberg, Scott Taylor, and David Weinstein as well as seminar participants at CIAR, Harvard, MIT, Toronto, UBC, and Yale. Along the 2-year path of data collection we have chalked up debts to Chris Thornberg for older CPS data, to Robert Barro and Jong-Wha Lee for education data, to Roger Theirrien of Statistics Canada for trade data, to Robert Summers and Alan Heston for the Penn World Tables, and to our research assistants Alberto Isgut, Huiwen Lai, Ata Mazaheri, Bridget O'Shaugnessy, and Anindyo Sen. We are particularly indebted to the referees. Research was supported by National Science Foundation grant SBR-95-15107. The views expressed are those of the authors and not necessarily those of the National Bureau of Economic Research.

Increasing Returns and All That: A View From Trade
Werner Antweiler and Daniel Trefler
NBER Working Paper No. 7941
October 2000
JEL No. F11, F12, D2

ABSTRACT

Do scale economies contribute to our understanding of international trade? Do international trade flows encode information about the extent of scale economies? To answer these questions we examine the large class of general equilibrium theories that imply Helpman-Krugman variants of the Vanek factor content prediction. Using an ambitious database on output, trade flows, and factor endowments, we find that scale economies significantly increase our understanding of the sources of comparative advantage. Further, the Helpman-Krugman framework provides a remarkable lens for viewing the general equilibrium scale elasticities encoded in trade flows. In particular, we find that a third of all goods-producing industries are characterized by scale. (The modal range of scale elasticities for this group is 1.10-1.20 and the economy-wide scale elasticity is 1.05.) Implications are drawn for the trade-and-wages debate (skill-biased scale effects) and endogenous growth.

Werner Antweiler
University of British Columbia
Faculty of Commerce and Business Admin.
2053 Main Mall
Vancouver, BC, V6T 1Z2
Canada
Tel: 604-822-8484
Fax: 604-822-8477
werner@pacific.commerce.ubc.ca

Daniel Trefler
University of Toronto
Department of Economics
140 St. George Street
Suite 707
Toronto, ON, M5S 3G6
Canada
Tel: 416-978-1840
Fax: 416-978-5519
trefler@chass.utoronto.ca
and NBER

Over the last 20 years, general equilibrium models of international trade featuring increasing returns to scale have revitalized the international trade research agenda. Yet general equilibrium econometric work remains underdeveloped: it has been scarce, only occasionally well-informed by theory, and almost always devoid of economically-meaningful alternative hypotheses. There are exceptions of course. These include Helpman (1987), Hummels and Levinsohn (1993, 1995), Brainard (1993, 1997), Harrigan (1993, 1996), and Davis and Weinstein (1996). However, this list is as short as the work is hard. The complexity of general equilibrium, increasing returns to scale predictions has deflected empirical research of the kind that is closely aligned with theory.

Surprisingly, one empirically tractable prediction remains overlooked, despite the fact that it is central to the approach of Helpman and Krugman (1985). We are referring to a variant of Vanek's (1968) factor content of trade prediction. In its Helpman-Krugman form the factor content of trade depends critically on the extent of scale returns in each industry. Scale matters because it determines both the pattern of trade and the amount of factors needed to produce observed trade flows. As Helpman and Krugman showed, their variant of the Vanek prediction comes out of a very large set of increasing returns models and so provides a robust way of evaluating these models. Yet remarkably, this increasing returns to scale factor content prediction has not been explored empirically. We know that the Heckscher-Ohlin-Vanek factor content prediction performs poorly e.g., Trefler (1995). Yet there has not been one iota of evidence that the Helpman-Krugman class of models performs better. Exploring this uncharted region is our first goal.

The second and more important goal of this paper is to quantify the extent of increasing returns to scale in the context of a general equilibrium model of international trade. This forces us to part company with the existing scale literature which focuses on estimating scale effects *separately* for each industry. Instead, we must seek a radically different *general equilibrium* strategy. It is as follows. The Helpman-Krugman variant of the Vanek prediction imposes a precise relationship between the elasticities of scale in each industry and a particular set of data on trade, technology, and factor endowments. We search for elasticities of scale that make this relationship fit best. Like cosmologists searching the heavens for imprints of the big bang, we are searching the historical record on trade flows for imprints of scale as a source of comparative advantage.

There are two reasons why 15 years of research has largely escaped exposure to general equilibrium econometric work. As noted, one is the complexity of general equilibrium predictions. The other is the lack of the internationally comparable cost data needed to make inferences about scale. Surprisingly, both problems are resolved by shifting the focus from trade in goods to the factor content of trade. Comparative costs are so obviously the basis of international trade that no amount of evidence to the contrary would dislodge this view. That is, trade flows from low-cost exporters to high-cost importers. It should therefore be possible to use trade flow data to make inferences about international cost differences. What makes this possibility so attractive is that detailed trade flow data are available even where no cost data exist. Factor content calculations transparently structure the problem of inferring costs from trade data. Remember that the factor content of trade is just a set of derived demands for the factor inputs used to produce observed trade flows. By Shephard's lemma, these demands can be integrated back to obtain costs. That is, factor content predictions provide a way of inferring international cost differences from trade flows.

Just as you cannot squeeze water from a stone, you cannot infer costs without extensive data.

For this paper we have constructed a new and remarkable database covering all internationally traded, goods-producing industries (27 in manufacturing and 7 outside of manufacturing) for 71 countries over the period 1972-1992. The database contains bilateral trade and gross output by industry, country, and year as well as factor endowments by country and year. The most difficult part of this project has been the two years spent on database construction.

Our conclusions are as follows. When restricting all industries to have the same scale elasticity our method yields a precisely estimated scale elasticity of 1.05. This is small for mark-up models (Rotemberg and Woodford 1992), arguably large for endogenous growth models, and certainly large for hysteretic models. However, there is considerable heterogeneity among industries. For about a third of all industries the data are not sufficiently informative for making inferences about scale. For another third of all industries, we estimate constant returns to scale. That is, allowing for scale in these industries does not add much to our understanding of the cost basis for trade flows. For the remaining third of all industries, including such industries as pharmaceuticals and machinery, we find strong evidence of increasing returns to scale. For this group, our general equilibrium scale estimates have a modal range of 1.10-1.20. Scale is central to understanding trade for these industries.

At this stage, an important caveat is in order. We will be working with industry-level data, not plant-level data. Thus, we will be finding a relationship between industry output and trade-revealed industry costs. This relationship may be partly induced by underlying plant-level scale economies. However, it may also be induced by industry-level externalities (e.g., Paul and Siegel 1999) and, for the most technologically dynamic industries, by new process technologies that are embodied in an increased scale of operation (e.g., Rosenberg 1982, 1994). To smooth the exposition, we use the term 'scale' to include both the many familiar sources of scale as well as industry-level externalities and scale-biased technical change. Whence the 'All That' of our title. This point is developed in section 3.4.

The reader with no trade interests may want to skip section 1 (a data preview) and skim section 2 (trade theory and econometric identification). The estimating equation appears in section 2.4 (equation 12). The core empirical work appears in section 3, especially tables 3 and 4. Sensitivity analysis appears in section 4.

1. Preliminaries

1.1. The Data

The database constitutes one of the most, if not the most, comprehensive descriptions of the global trading environment ever assembled. Its construction consumed two years of intensive work. The data have four dimensions. (*i*) *Countries*. There are 71 countries covering the entire development spectrum. See appendix table A.1 for a list of countries. (*ii*) *Industries*. There are 34 industries covering virtually the entire tradeables or goods-producing sector. 27 of these are 3-digit ISIC manufacturing industries. The remaining industries are non-manufacturing industries ranging from livestock to electricity generation. A list of industries appears in the tables below. (*iii*) *Factors*. There are 11 factors: capital stock (Summers and Heston 1991), 4 levels of educational attainment (Barro and Lee 1993), 3 energy stocks (coal reserves, oil and gas reserves, and hydroelectric potential), and 3 types of land (cropland, pasture, and forests). (*iv*) *Years*. There are 5 years: 1972, 1977, 1982, 1987, and 1992.

The database contains the following: (*i*) deflated bilateral trade and deflated gross output by country, industry, and year, (*ii*) factor endowments and income by country and year, and (*iii*) double-deflated input-output relations by year for the United States. See appendix A for database details. Finer details will appear as a separate paper when the database is made publicly available.

1.2. Data Preview

The single most important fact supporting the use of increasing returns models is the presence of intra-industry trade, that is, of trade in similar products. Models by Krugman (1979), Helpman (1981), and Ethier (1982) were designed to explain such trade. (Though note that Davis (1995, 1997) provides Ricardian and Heckscher-Ohlin explanations of intra-industry trade.) Figure 1 plots each country's 1992 imports and exports of Instruments against its output of the Instruments industry (ISIC 385). A country's imports from the rest of the world appear as a point above the axis and a country's exports to the rest of the world appear below the axis. Imports, exports, and output are scaled by the country's gross domestic product (gdp). The main point about the picture is that countries with large imports also have large exports. This is intra-industry trade and has been well-documented. What is surprising is how extensive it is both for large and small producers. This may be the graphical counterpart to Hummels and Levinsohn's (1993, 1995) result about how well monopolistic competition models perform for poor countries. Another surprise is the mirror-like quality across the horizontal axis. Some of this is entrepôt trade as in the case of Singapore, but symmetry extends to all countries. One gets no sense of the specialization that lies at the heart of theories of comparative advantage. A possible explanation is offshore sourcing of parts, a phenomenon which we term ‘intra-mediate’ trade.

Does this pattern hold for bilateral trade as well? Figure 2 plots bilateral trade in instruments for cases where the U.S. is the importer (top panel) or exporter (bottom panel). For example, the ‘HKG’ observation in the upper right plots U.S. imports from Hong Kong against Hong Kong output. The data are scaled by the gdp of the (non-U.S.) producer country. Remarkably, symmetry persists even at the bilateral level. Another feature of the data is a point made by Harrigan (1996). The bilateral monopolistic competition model predicts that country i's imports from country j will be proportional to j's output (Helpman and Krugman 1985). Harrigan re-writes this as a log regression of bilateral imports on output, that is, as the regression line plotted in the top panel of figure 2 (though without gdp scaling). He obtains a slope of 1.2 and an R^2 of 0.7. The similar figure 2 statistics - a slope of 1.4 and an R^2 of 0.5 - reinforce how robust Harrigan's results are. An odd feature stemming from symmetry is that the regression line in the bottom panel of figure 2 also fits well (a slope of 0.6 and an R^2 of 0.4). This means that the United States exports instruments to big producers of instruments. This is not a prediction that comes out of the standard monopolistic competition model. It is consistent with multinational sourcing of intermediate parts (Brainard 1993, 1997 and Feenstra and Hanson 1996a, 1996b, 1997).

Increasing returns models often predict extreme patterns of specialization and regional concentration of industry. A measure of regional concentration is the Herfindahl index defined as $\sum_i q_{gi}^2$ where q_{gi} is country i's share of world production of good g. Table 1 reports the Herfindahl index for 1992. Note from the ‘1992-1972’ column that the index has fallen over time in almost all industries. This is the globalization trend that has received so much attention. The most simplistic model of scale returns predicts that increasing returns to scale industries will be geographically specialized and thus appear at the top of the table 1 list. Not surprisingly, Instruments and Machinery appear near the top. Footwear and Leather Products - *a priori* constant returns industries - appear near the bottom of the list as expected. The table reports three commonly used, albeit weak, proxies of scale. These are described in the notes to table 1. There is a weak correlation between these measures and the Herfindahl index. Further inspection of the table reveals unexpected industries near the top and bottom of the list. Coal Mining and Pulp and Paper are natural resource- or endowment-based industries. Their position near the top is consistent with scale returns as well as with the Heckscher-Ohlin model. We also constructed Krugman's (1991) population-scaled Herfindahl index. By this measure natural resource-based industries topped the list of regionally concentrated industries. Clay and Cement Products bottom out the list because of high transport costs. Thus, scale returns in its simplest form explains some of the regional concentration of production, but endowments, trade restrictions, and transport costs also contribute.

2. Theory and Estimating Equation

2.1. The Vanek Prediction with Increasing Returns to Scale

Vanek's factor content prediction transparently structures the problem of inferring unavailable cost and scale-elasticity information from available trade flow data. In this section we review the well-known observation that Vanek's factor content prediction may hold for both constant and increasing returns to scale technologies (Helpman and Krugman 1985). We then follow Trefler (1996) in deriving a definition of the factor content of trade that is theoretically consistent with existing data.

Let $i = 1, \ldots, N$ index countries and $f = 1, \ldots, F$ index factors. Let V_{fi} be the endowment of factor f in country i. In any international cost comparison, which is our main aim here, inputs must be measured in internationally comparable units of quality or productivity. Let π_{fi} be the productivity of factor f in country i relative to the United States. $V^*_{fi} \equiv \pi_{fi}V_{fi}$ is country i's endowment measured in productivity-adjusted units. Let $V^*_{f\text{w}} \equiv \Sigma_i \pi_{fi}V_{fi}$ be the productivity-adjusted world factor endowment. Let s_i be i's share of world income. Country i is said to be abundant in factor f if its share of productivity-adjusted world endowments exceeds its share of world consumption: $V^*_{fi}/V^*_{f\text{w}} > s_i$ or $V^*_{fi} - s_i V^*_{f\text{w}} > 0$.

Let C_{ij} be country i's consumption of final goods produced in country j. Let $C_{\text{w}j} \equiv \Sigma_i C_{ij}$ be the world's consumption of final goods produced in country j. Country i's imports of final goods from j are just C_{ij}. However, it is useful to introduce additional notation for imports: $M^c_{ij} \equiv C_{ij}$ for $i \neq j$ and $M^c_{ii} \equiv 0$. Exports of final goods are $X^c_i \equiv \sum_{j \neq i} C_{ji}$. Let M^y_i and X^y_i be country i's imports and exports of intermediate inputs, respectively. The 'c' and 'y' superscripts distinguish between final consumption and intermediate inputs. All the vectors in this paragraph are $G \times 1$ vectors where G is the number of goods.

Let A_{fi} be the $1 \times G$ choice-of-techniques vector. A typical element gives the amount of factor f required both directly and indirectly (in an input-output or general equilibrium sense) to produce one dollar of a good. Let $A^*_{fi} \equiv \pi_{fi}A_{fi}$ be the productivity-adjusted choice-of-techniques vector.

The Vanek prediction is a prediction of the form $V_{fi} - s_i \Sigma_j V_{fj} = F_{fi}$ where F_{fi} is some measure of the factor content of trade. To move towards such a prediction we assume that factor markets clear nationally. This assumption *alone* implies the following equation:

$$V^*_{fi} - s_i V^*_{f\mathrm{w}} = F^*_{fi} + \varepsilon_{fi} \tag{1}$$

where

$$\varepsilon_{fi} \equiv \Sigma_j A^*_{fj}(C_{ij} - s_i C_{\mathrm{w}j}) \tag{2}$$

$$F^*_{fi} \equiv \left[A^*_{fi}X^c_i - \Sigma_j A^*_{fj}M^c_{ij}\right] + \left[A^*_{fi}(X^y_i - M^y_i) - s_i \Sigma_j A^*_{fj}(X^y_j - M^y_j)\right]. \tag{3}$$

Appendix B provides a simple derivation of equation (1). Aside from using factor market-clearing, the derivation is purely algebraic and based entirely on manipulation of input-output identities. For our purpose, which is to consistently estimate scale elasticities, we do not need to interpret equations (1)-(3). All we will need is that the residuals ε_{fi} satisfy a familiar econometric orthogonality condition. However, equation (1) does have an interpretation. Under the strong assumption that $\varepsilon_{fi} = 0$, equation (1) is the Vanek prediction. By way of an extended aside, we turn to this interpretation.

Under the standard Heckscher-Ohlin-Vanek (HOV) assumptions, the A^*_{fi} are internationally identical and, with identical homothetic preferences and internationally common goods prices, consumption satisfies the condition $\Sigma_j C_{ij} = s_i \Sigma_j C_{\mathrm{w}j}$. Hence, $\varepsilon_{fi} = 0$ and the Vanek prediction holds. Under the assumptions of many increasing returns to scale models, there will be international specialization of production. Even though the A^*_{fi} will not be internationally identical, the usual HOV consumption condition $\Sigma_j C_{ij} = s_i \Sigma_j C_{\mathrm{w}j}$ is again enough to generate $\varepsilon_{fi} = 0$.[1] In-

[1]Let country j' be the only producer. Then C_{ij} vanishes for $j \neq j'$ and the usual HOV consumption condition

ternational specialization of production is associated with scale returns (Helpman and Krugman 1985, chapter 3), exogenous international technology differences (Davis 1995), failure of factor price equalization (Deardorff 1979 and Davis and Weinstein 1998), or a mix of these (Markusen and Venables 1998).

More intriguing is the possibility that the ε_{fi} might vanish even without production specialization or internationally identical A^*_{fi}. In models with taste for variety (Krugman 1979) or ideal-type preferences (Helpman 1981), each country buys every final good from every other country in proportion to its size. Mathematically, $C_{ij} = s_i C_{\mathrm{w}j}\ \forall i, j$. Thus, if two countries produce cheese, all countries buy from both producers, not just one producer. Note that $C_{ij} = s_i C_{\mathrm{w}j}$ applies only to final goods, not intermediate inputs. With $C_{ij} = s_i C_{\mathrm{w}j}$, $\varepsilon_{fi} = 0$ and the Vanek prediction holds. The observation that the Vanek prediction holds for a large class of increasing returns to scale models is one of the central insights in Helpman and Krugman (1985).[2]

There is a potential disconnect between our industry-level empirical work and the models of the previous paragraph which feature internal returns to scale. One issue is whether with internal returns to scale there will be an industry-level relationship between output and output per unit of input. The answer is yes. The only exception is the CES utility case, but as Lancaster (1984) and many others have noted, one does not want to take this case seriously for empirical work. A second issue is whether our industry-level data allow us to distinguish between internal and external

reduces to $C_{ij'} = s_i C_{\mathrm{w}j'}$. Likewise, ε_{fi} reduces to $A^*_{fj'}(C_{ij'} - s_i C_{\mathrm{w}j'}) = 0$.

[2]Note that in the absence of trade in intermediate inputs, the condition $C_{ij} = s_i C_{\mathrm{w}j}$ is just the regression line plotted in the top panel of figure 2 and examined so thoroughly by Harrigan (1996). This follows from the facts that output Q_i equals world demand $C_{\mathrm{w}i}$ and consumption C_{ij} equals imports M_{ij}. Plugging $Q_i = C_{\mathrm{w}i}$ and $C_{ij} = M_{ij}$ into $C_{ij} = s_i C_{\mathrm{w}j}$ yields $M_{ij} = s_i Q_j$. Thus, $C_{ij} = s_i C_{\mathrm{w}j}$ is implicit in much of the literature on monopolistic competition and gravity equations.

returns to scale. Here the answer is no. This is the point of the 'all that' in our title and of the 'What is Scale?' section below. At first glance it seems unusual that we do not need to distinguish between internal and external returns to scale. After all, they are very different in their implications for market structure, the location of production, and trade patterns. The remarkable insight that pervades Helpman and Krugman (1985) is that the form of scale returns has only very modest implications for the factor content of trade. Thus, while predictions about the location of production and the pattern of trade are often complicated, model-dependent, and/or just plain indeterminate in this class of models, predictions about the factor content of trade are relatively straightforward. *This is a key reason for why we have chosen the factor content route.*

To conclude our discussion of the consumption condition, the interpretation of equation (1) as a Vanek prediction requires $\varepsilon_{fi} = 0$. However, consistent estimation of scale elasticities requires a weaker condition, namely, a familiar econometric orthogonality condition involving the 'residual' ε_{fi}. We return to this point below.

We turn next to the interpretation of F^*_{fi} as the factor content of trade. With $\varepsilon_{fi} = 0$, $V^*_{fi} - s_i V^*_{f\mathrm{w}}$ is the productivity-adjusted factor content of i's trade. It follows that so is F^*_{fi}. The question arises as to why the equation (3) expression for F^*_{fi} is so unfamiliar. In the absence of traded intermediates, equation (3) reduces to the usual definition $F^*_{fi} \equiv A^*_{fi} X^c_i - \Sigma_j A^*_{fj} M^c_{ij}$ e.g., Helpman and Krugman (1985), equation 1.11. That is, producer-country choice of techniques are used to calculate the required factor inputs. With traded intermediates there is a subtle problem associated with the measurement of A_{fi}. Current practice by all national statistical agencies is to construct A_{fi} by lumping together the inter-industry purchases of domestically produced and imported intermediate inputs. One needs somehow to net out the imported intermediates. The second term in equation (3) is the theoretically correct way of doing this.

In equation (3) it appears as if we are using net trade for the factor content of intermediates and gross trade for the factor content of final goods. Nothing could be more misleading. The two bracketed terms in equation (3) are not decomposable in this way. Intermediate input terms such as $A^*_{fi}M^y_i$ enter equation (3) for the entirely different reason of netting out imported intermediates. Appendix B develops this point. A more complete discussion appears in Trefler (1996).

2.2. Isolating Scale and Exogenous International Productivity Differences

The Vanek prediction of equation (1) is an implicit relation between trade flows and costs. (By Shephard's lemma the A_{fi} are derivatives of cost functions.) It remains to relate costs to the scale of output. Recall that $A_{fi} \equiv (a_{f1i}, ..., a_{fGi})$ where a_{fgi} is the amount of factor f needed to produce one unit of good g in country i. Since a_{fgi} is derived from cost minimization it depends on intermediate input prices, factor input prices $w_i = (w_{1i}, ..., w_{Fi})$, and industry g output Q_{gi}. The aim is to show how this dependence can be restricted in a way that identifies the role of Q_{gi}. To motivate the analysis we start with the simple case of homothetic production functions and no intermediates. These are reintroduced in the next section.

Homotheticity implies separability of average costs:

$$\mathrm{AC}_{gi}(w_i, Q_{gi}) = c_{gi}(w_i)\ \phi_{gi}(Q_{gi}) \tag{4}$$

where c_{gi} is a constant returns unit cost function and ϕ_{gi} is a decreasing function that captures scale effects.

The productivity-adjusted factor price corresponding to $V^*_{fi} \equiv \pi_{fi} V_{fi}$ is w_{fi}/π_{fi}. To see this by way of example, if Hong Kong workers were half as productive as U.S. workers ($\pi_{L,\text{HK}} = \frac{1}{2}$) then the Hong Kong productivity-adjusted wage would be twice the observed wage ($w_{L,\text{HK}}/\pi_{L,\text{HK}} = 2w_{L,\text{HK}}$). We assume that after adjusting for international factor productivity differences, there are no other sources of international differences in the c_{gi}. That is,[3]

$$c_{gi}(w_{1i}, ..., w_{Fi}) = c_{g,\text{US}}(w_{1i}/\pi_{1i}, ..., w_{Fi}/\pi_{Fi}). \tag{5}$$

Putting equation (5) into equation (4), the implied input demand per unit of output is given by[4]

$$a_{fgi}(w_i, Q_{gi}) = \frac{\partial c_{g,\text{US}}(w_{1i}/\pi_{1i}, \ldots, w_{Fi}/\pi_{Fi})}{\partial (w_{fi}/\pi_{fi})} \frac{\phi_{gi}(Q_{gi})}{\pi_{fi}}. \tag{6}$$

We do not observe the derivative in equation (6). Further, when we move to a general equilibrium interpretation of the a_{fgi}, the a_{fgi} will each depend on cost derivatives not just in industry g, but in all industries. The data requirements will be enormous. We need to cut through this.

It is a commonplace among economists that factor prices are primarily determined by factor productivity. Following Trefler (1993), we take this observation very seriously by assuming that

[3]The proof of the equality is as follows. Let y_i be an input vector, let $f_{gi}(\cdot)$ be a constant returns production function, and define $\Pi_i \equiv \text{diag}(\pi_{1i}, ..., \pi_{Fi})$. Our assumption, couched in terms of f_{gi} rather than c_{gi}, is that $f_{gi}(y_i) = f_{g,\text{US}}(\Pi_i y_i)$. Then $c_{gi}(w_i) \equiv \min_{y_i}\{w_i y_i | f_{gi}(y_i) = 1\} = \min_{y_i}\{w_i \Pi_i^{-1} \Pi_i y_i | f_{g,\text{US}}(\Pi_i y_i) = 1\} = \min_{y_{\text{US}}}\{w_i \Pi_i^{-1} y_{\text{US}} | f_{g,\text{US}}(y_{\text{US}}) = 1\} = c_{g,\text{US}}(w_i \Pi_i^{-1})$ where the second last equality follows from the fact that the change of variable $y_{\text{US}} \equiv \Pi_i y_i$ is an invertible function of y_i.

[4]The proof is as follows. Let $\text{TC}_{gi} = \text{AC}_{gi} Q_{gi}$ be total costs. $a_{fgi} \equiv (\partial \text{TC}_{gi}/\partial w_{fi})/Q_{gi}$ or, using equation (4), $a_{fgi} = \phi_{gi} \partial c_{gi}(w_i)/\partial w_{fi}$. From equation (5), $\partial c_{gi}(w_i)/\partial w_{fi} = \pi_{fi}^{-1} \partial c_{g,\text{US}}(\Pi_i^{-1} w_i)/\partial(w_{fi}/\pi_{fi})$. Equation (6) follows.

factor prices are *completely* determined by factor productivity. That is, $w_{fi}/\pi_{fi} = w_{f,\text{US}}/\pi_{f,\text{US}} = w_{f,\text{US}}$ for all factors f and all countries i. This factor price assumption is not used in the empirical production function literature. Its role here is in allowing us to move from industry-level analyses to the general equilibrium analysis of an economy's factor input requirements.

Substituting $w_{fi}/\pi_{fi} = w_{f,\text{US}}$ into equation (6) and manipulating yields[5]

$$a_{fgi}(w_i, Q_{gi}) = \frac{a_{fg,\text{US}}}{\pi_{fi}} \frac{\phi_{gi}(Q_{gi})}{\phi_{g,\text{US}}(Q_{g,\text{US}})}. \tag{7}$$

where $a_{fg,\text{US}} = a_{fg,\text{US}}(w_{\text{US}}, Q_{g,\text{US}})$. Thus, we have dramatically reduced the amount of international data required to calculate a_{fgi}. Further, one can now see that *we have forced all the international sample variation in factor requirements to operate via the exogenous international productivity term π_{fi} and the scale term $\phi_{gi}(Q_{gi})$.*

2.3. A Generalization

We now introduce intermediate inputs and non-homotheticities. Let d_{fgi} be the amount of primary factor input f demanded per unit of industry g output. Since our empirical results are not sensitive to the choice of functional form for ϕ_{gi}, we avoid excessive generality by introducing the form that appears most frequently in the empirical sections. This is $\phi_{gi} = (Q_{gi})^{-\alpha_g}$. In the empirical work we also sometimes allow α_g to vary across countries and factors. The extra math that stems from letting α_g depend on i and f is not complicated. See Antweiler and Trefler (1997). With

[5] With $\pi_{f,\text{US}} \equiv 1$ and $w_{fi}/\pi_{fi} = w_{f,\text{US}}$, equation (6) becomes $a_{fgi} = (\phi_{gi}/\pi_{fi})\partial c_{g,\text{US}}(w_{\text{US}})/\partial w_{f,\text{US}}$. Setting i = US yields $\partial c_{g,\text{US}}(w_{\text{US}})/\partial w_{f,\text{US}} = a_{fg,\text{US}}/\phi_{g,\text{US}}$. Plugging this back into the last sentence's expression for a_{fgi} yields equation (7).

$\phi_{gi} = (Q_{gi})^{-\alpha_g}$, the factor demand counterpart to equation (7) is

$$d_{fgi} = \frac{d_{fg,\text{US}}}{\pi_{fi}} \left(\frac{Q_{gi}}{Q_{g,\text{US}}} \right)^{-\alpha_g} \quad (8)$$

Turning from primary factor inputs to intermediate inputs, let b_{hgi} be the amount of intermediate input h demanded per unit of industry g output. Unlike primary factors, we assume that intermediate inputs are costlessly traded internationally. Thus, each intermediate input has a common, quality-adjusted price internationally. We have extensively examined the empirical possibility of scale effects associated with intermediate inputs. See Antweiler and Trefler (1997). However, this substantially complicates the model without offering any additional empirical insights. As a result, we simplify the exposition by assuming that there are no scale effects associated with intermediate inputs. For example, two cars require twice as many tires as one car. With this, equation (7) becomes

$$b_{hgi} = b_{hg,\text{US}} \quad (9)$$

Note that the system of input demands in equations (8)-(9) is non-homothetic. Duality results in Epstein (1982) ensure that the system is supported by an underlying production function.

Our parameter of interest is the elasticity of scale μ_g. We treat it as being independent of i because we will be estimating it by pooling across countries. μ_g depends on the share of primary factor inputs in total costs θ_g. Specifically,[6]

[6] The proof is as follows. Omitting g and i subscripts, total costs are TC$\equiv \sum_f w_f d_f Q + \sum_h p_h b_h Q$. From equations (8)-(9) together with $w_{fi}/\pi_{fi} = w_{f,\text{US}}$, TC$= \sum_f w_{f,\text{US}} d_{f,\text{US}} Q (Q/Q_{\text{US}})^{-\alpha} + \sum_h p_h b_{h,\text{US}} Q$. Note that $1/\mu \equiv \partial \ln(\text{TC})/\partial \ln(Q) = (1-\alpha) \sum_f w_{f,\text{US}} d_{f,\text{US}} (Q/Q_{\text{US}})^{-\alpha} Q/\text{TC} + \sum_h p_h b_{h,\text{US}} Q/\text{TC} = (1-\alpha) \sum_f w_f d_f Q/\text{TC} + \sum_h p_h b_h Q/\text{TC}$. Further, $\theta \equiv \sum_f w_f d_f Q/\text{TC}$. Hence $1/\mu = (1-\alpha)\theta + (1-\theta) = 1 - \alpha\theta$. Equation (10) follows.

$$\mu_g = (1 - \alpha_g \theta_g)^{-1}. \tag{10}$$

With intermediate inputs, a_{fgi} must be defined as the total factor requirements (direct plus indirect in an input-output sense) needed to bring one unit of good g to final consumers. In matrix notation total factor requirements are defined in the usual input-output way as $A_{fi} \equiv D_{fi}(I - B_i)^{-1}$ where $D_{fi} = (d_{f1i}, ..., d_{fGi})$ and B_i is the $G \times G$-matrix whose (h, g) element is b_{hgi}. We can now state the main result of this section about how A^*_{fi} varies with output and observed data.

Theorem 1. *Input demand equations (8)-(9) imply*

$$A^*_{fi}(\mu) \equiv \pi_{fi} A_{fi} = D_{f,\mathrm{US}} \;\; \Phi(Q_i, Q_{\mathrm{US}}; \mu) \;\; (I - B_{\mathrm{US}})^{-1} \tag{11}$$

where $\mu \equiv (\mu_1, \ldots \mu_G)$ *and* $\Phi(Q_i, Q_{\mathrm{US}}; \mu)$ *is a* $G \times G$ *diagonal matrix whose* g*th element is*

$$\left(\frac{Q_{gi}}{Q_{g,\mathrm{US}}} \right)^{(1-\mu_g)/(\mu_g \theta_g)}.$$

The proof is straightforward.[7] Theorem 1 provides a parameterization of scale effects that is consistent with our general equilibrium trade theories.

[7] Start with $\pi_{fi} A_{fi} \equiv \pi_{fi} D_{fi}(I - B_i)^{-1}$. From equations (8) and (10), $\pi_{fi} d_{fgi} = d_{fg,\mathrm{US}}(Q_{gi}/Q_{g,\mathrm{US}})^{-\alpha_g} = d_{fg,\mathrm{US}}(Q_{gi}/Q_{g,\mathrm{US}})^{(1-\mu_g)/(\mu_g \theta_g)}$. Hence, $\pi_{fi} D_{fi} = D_{f,\mathrm{US}} \Phi$. From equation (9), $(I - B_i)^{-1} = (I - B_{\mathrm{US}})^{-1}$. Hence $\pi_{fi} A_{fi} = D_{f,\mathrm{US}} \Phi (I - B_{\mathrm{US}})^{-1}$ as required.

2.4. The Estimating Equation

At this point we introduce time subscripts t. From equation (1) we are interested in

$$H_{fit}(\mu) \equiv \pi_{fit} V_{fit} - s_{it} \Sigma_j \pi_{fjt} V_{fjt} - F^*_{fit}(\mu) = \varepsilon_{fit} \tag{12}$$

where ε_{fit} is given by equation (2), $F^*_{fit}(\mu)$ is given by equation (3), and A^*_{fit} in equation (3) is given by equation (11). Note that μ does not vary with time. We find no evidence of this empirically (see Antweiler and Trefler 1997) and so forego the additional notation.

$H_{fit}(\mu)$ also depends on the π_{fit}. With 71 countries, 11 factors, 5 years, and the normalization $\pi_{f,\mathrm{US},t} \equiv 1$, there are 3,850 π_{fit}. Rather than estimate them, we use $\pi_{fit} = w_{fit}/w_{f,\mathrm{US},t}$ and plug in data for $w_{fit}/w_{f,\mathrm{US},t}$ wherever π_{fit} appears in equation (12). For labour factors, w_{fit} is the average manufacturing wage (from the same source as the output data). For capital we use the Penn World Table price of capital (Summers and Heston 1991). For cropland we use gdp generated by crops in 1985 per hectare of cropland. For pasture, we use gdp generated by livestock in 1985 per hectare of pasture. Data are from the Food and Agricultural Organization of the United Nations (1992, tables 1.4 and 1.6). For forestry, data limitations force us to assume $\pi_{fit} = 1$. For energy (coal reserves, oil and gas reserves, and hydroelectric potential), endowments and output are either measured in joules in the source data or we have converted them to joules using internationally recognized converters. Since the conversion to joules is country-specific and already takes into account international quality differences, the correct assumption for energy is $\pi_{fit} = 1$.

Three data issues remain before we can examine equation (12). First, following Conway (2000), s_{it} is based on PPP-adjusted income. See appendix A. Second, $D_{f,\mathrm{US}}$ and B_{US} in equation

(11) and θ_g in equation (10) use U.S. data. The source data are described in appendix A. Third, we need to know the share of trade that is intermediate inputs trade. For example, we observe $X^c_{it} + X^y_{it}$, but not X^c_{it} or X^y_{it} separately. Appendix C details our method for allocating trade into its intermediate and final goods components. We were initially concerned that our estimates of μ_g would be sensitive to our allocation method. This turns out to be a misplaced concern. To persuade the reader of this, appendix C also reports estimates of scale for many different allocation methods.

We summarize by noting some differences between our general equilibrium approach and existing partial equilibrium approaches to estimating scale returns in a cross-country setting. On the input side, data on industry-level inputs are notoriously bad or non-existent e.g., labour inputs by educational attainment. While partial equilibrium approaches must use such data, our approach allows us to use national-level aggregates. These aggregates are typically more reliable. In addition, our approach easily allows us to model international differences in input productivities i.e., the π_{fi}. On the output side, we are able to shift some of the burden off of gross output data and onto trade flow data. Real gross output data suffer a number of serious problems, especially for poorer countries. In contrast, trade flow data are more accurate, measured in dollars, and with the approximate assumption of equal prices across countries, correspond more directly to physical quantities. Thus, the general equilibrium approach allows us to exploit alternative and somewhat more reliable data sources.

3. Results

We re-write equation (12) as our final estimating equation

$$\frac{H_{fit}(\mu)}{\sigma_{fit}} = \lambda_i + \eta_{fit} \tag{13}$$

where the following holds. $H_{fit}(\mu)$ is defined in equation (12). The σ_{fit} are generalized least squares (GLS) corrections with factor-year and country-year components. Appendix D describes the σ_{fit} in detail. ε_{fit} of equations (2) and (12) is now written as the sum of two components, $\varepsilon_{fit}/\sigma_{fit} = \lambda_i + \eta_{fit}$. The λ_i are country fixed effects and the η_{fit} are independently and identically distributed disturbances with mean zero. We estimate equation (13) using maximum likelihood (ML), non-linear least squares (NLS), and non-linear two-stage least-squares (NL2SLS) estimators. These are reviewed in the next section.

3.1. Preliminary Estimation

To fix ideas about the specification we start with the strong assumption that all industries exhibit the same degree of scale economies. We will relax this shortly. We pool across all 71 countries, all 11 factors, and all 5 years to give us $3,905$ observations. The 5 years are 1972, 1977, 1982, 1987, and 1992.

Table 2 reports the estimated scale elasticities for a variety of estimators. We start with the ML and NLS estimators. They are distinguished by their use of exogeneity assumptions. In general equilibrium almost everything is endogenous including endowments of physical and human capital. Minimizing endogeneity bias would seem to require treating the most mis-measured variables as left-hand side variables. (This is informally supported by Klepper and Leamer 1984). National endowments and the data used to construct the π_{fit} strike us as fitting the bill. For example, the capital stock data and quality-adjusted measures of education stocks are troublesome. Therefore, for the purposes of ML and NLS estimation, we treat $\pi_{fit}V_{fit} - s_{it}\Sigma_j \pi_{fjt}V_{fjt}$ as a left-hand side variable and $F^*_{fit}(\mu)$ as a right-hand side variable. From row (1) of table 2, the ML estimate is 1.051 and the NLS estimate is 1.050. The other estimated parameters of the model are the variance and GLS parameters. These appear in appendix table A.3. Although not reported in table 2, similar results obtain without the π_{fit} or fixed effects. The $\pi_{fit} = 1$ ML estimate is 1.043 and the no-fixed effect ML estimate is 1.054. See Antweiler and Trefler (1997) for additional specifications.

One null hypothesis is that there exists scale returns (H$_0$: $\mu > 1$) and that the Helpman-Krugman variant of the Vanek factor content prediction is 'true.' An explicit alternative hypothesis is that there are constant returns to scale (H$_0$: $\mu = 1$), in which case our model reduces to the usual Heckscher-Ohlin-Vanek factor content prediction. The ML t-statistic of 13.41 is for the hypothesis $\mu = 1$. *The t-statistic tells us that the data favour the Helpman-Krugman, increasing-returns framework over the Heckscher-Ohlin-Vanek model. This is a novel and important general equilibrium result.*

The ML estimate treats $F^*_{fit}(\mu)$ as exogenous. We examine endogeneity by using instrumental variables methods. The '$Z = F^*_{fi,t-\tau}(\mu)$' columns use a standard Amemiya (1974) NL2SLS estimator with $F^*_{fit}(\mu)$ instrumented by its lagged value $F^*_{fi,t-\tau}(\mu)$. (The NL2SLS minimand appears in appendix E.) The longer is the lag τ, the more likely is the instrument to be orthogonal to the error. Recalling that we have data for 1972, 1977, 1982, 1987, and 1992, the longest possible lag is 20 years. However, the longer the lag, the fewer are the observations left for estimation. Given the poor small-sample properties of IV estimators (Nelson and Startz 1990), we prefer to use a 15-year lag. In this case, $(F^*_{fi,87}, F^*_{fi,92})$ is instrumented by $(F^*_{fi,72}, F^*_{fi,77})$. Row (2) reports the results for 15-year lags. This leaves us with $2 \times 71 \times 11 = 1,562$ observations. The NL2SLS estimate of 1.095 is significantly larger than the NLS estimate of 1.044 in that the estimators' 1% confidence intervals are non-overlapping. This is the basis for a Hausman exogeneity test. While the Hausman statistic is negative, it seems likely that exogeneity would be rejected by a refined test that takes into account the correlation between the two estimators. We have not attempted this.

We also considered a larger instrument set constructed from polynomials of the instrument $F^*_{fi,t-\tau}$. In rows (2)-(4), the instrument set is $\{(F^*_{fi,t-\tau})^k\}_{k=1}^K$ where $K = 1$ in row (2), $K = 3$ in row (3), and $K = 5$ in row (4). See the 'K' column in table 2. As is apparent, the results are insensitive to the size of the instrument set.

Table 2 also reports results for different lag lengths. In rows (5)-(7), we consider the shorter lag length of 10 years ($2,343$ observations). In rows (8)-(10), we consider the longer lag length of 20 years (781 observations). The results are insensitive to the choice of lag length.

NL2SLS with our lagged instrument set is dominated in efficiency terms by NL2SLS with an instrument set based on $\partial \widehat{F}^*_{fit}/\partial\mu$. The $\partial \widehat{F}^*_{fit}/\partial\mu$ are the fitted values of $\partial F^*_{fit}/\partial\mu$ from a preliminary NL2SLS procedure. (See Jorgenson and Laffont 1974 and Amemiya 1975. Details appear in appendix E.) The estimates appear in the columns '$Z = \partial \widehat{F}^*_{fit}/\partial\mu$'. As is apparent, the estimates are the same as for the '$Z = F^*_{fi,t-\tau}(\mu)$' instrument set. Note that for $K = 1$, the two NL2SLS estimates are mathematically equivalent.

The bottom line from table 2 is that there is evidence of modest scale economies at the aggregate level no matter how one tackles estimation. The elasticity estimate of 1.051 is well within the bounds of what has been reported in the U.S. time series literature (e.g., Basu 1995 and Basu and Fernald 1997). Against our conclusion of statistically significant scale returns must be weighed the economically small size of the estimated μ. A 1% rise in output is associated with a 0.05% fall in average costs. Further, a country operating at a tenth of U.S. levels has only 14% higher average costs.[8] Of course, in dynamic models such as those displaying endogenous growth and especially hysteresis, $\mu = 1.051$ can have important consequences.

The factor demand implications of $\mu = 1.051$ are much larger than the average cost implications. From equation (8), the elasticity of factor demand per unit of output is α. From the last footnote, $\mu = 1.051$ implies $\alpha = 0.18$. That is, a 1% rise in output leads to a 0.18% fall in demand for primary inputs per unit of output. Further, a country operating at a tenth of U.S. output levels uses 51% $(= .10^{-.18} - 1)$ more productivity-adjusted factor inputs per unit of output. *Thus, from the perspective of factor endowments theory even this small scale estimate is very important.* We will see that it has some implications for the 'mystery of missing trade' (Trefler 1995).

[8] Dropping g and i subscripts, 0.05% follows from the fact that $\partial \ln \text{AC}/\partial \ln Q = \partial \ln \text{TC}/\partial \ln Q - 1 = 1/\mu - 1 = -0.05$. 14% is calculated as follows. From $\text{AC}(Q) = \Sigma_f w_f d_f + \Sigma_h p_h b_h$ and equations (8)-(9), $\text{AC}(Q) = (Q/Q_{\text{US}})^{-\alpha} \Sigma_f w_{f,\text{US}} d_{f,\text{US}} + \Sigma_h p_h b_{h,\text{US}}$. Using $\theta_{\text{US}} = \Sigma_f w_{f,\text{US}} d_{f,\text{US}} / \text{AC}(Q_{\text{US}})$ yields $\text{AC}(Q)/\text{AC}(Q_{\text{US}}) = (Q/Q_{\text{US}})^{-\alpha}\theta + (1-\theta)$. With $\theta = 0.274$ (see appendix A) and $\mu = 1.051$, equation (10) implies $\alpha = 0.18$. Hence $\text{AC}(Q)/\text{AC}(Q_{\text{US}}) = 1.14$.

3.2. Industry-Level Analysis

General equilibrium estimation is complicated. We have found it computationally infeasible to simultaneously estimate separate scale elasticities for all 34 industries. Fortunately, there are other paths to interesting results. Consider the following iterative three-step procedure:

1. Rank industries by the size of their scale elasticities. In the first iteration this ranking is based on external information. In subsequent iterations it is taken from the output of the previous iteration.

2. Pick a particular industry g, place all industries higher in the ranking in one group and all industries lower in the ranking in another group. Industry g sits in a separate, third group. Use ML to estimate equation (13) subject to the restriction that all industries within a group share a common scale elasticity. This yields scale estimates $\widehat{\mu}_{HIGH}$, $\widehat{\mu}_{LOW}$, and $\widehat{\mu}(g)$ for the high, low, and g groups, respectively.[9]

3. Repeat step 2 for each g. This yields a set of scale estimates $\{\widehat{\mu}(g)\}_{g=1}^{34}$. Return to step 1 using $\{\widehat{\mu}(g)\}_{g=1}^{34}$ to rank industries.

Since the choice of initial rank does not matter, we defer description of this choice to the 'Sensitivity Analysis' section. In that section we also note that there are no algorithm convergence issues.

Table 3 reports the $\widehat{\mu}(g)$ for the specification with 71 countries, 11 factors, and 5 years. In the table industries are classified into four groups. The increasing returns to scale (IRS) group contains those industries with $\widehat{\mu}(g)$ that are significantly greater than unity at the 1% level. (Instruments is

[9] Since the lowest ranked industry can never be in the g group, its elasticity is given by μ_{LOW} for the case where both the low and g groups each have only 1 industry. A similar detail applies to the highest ranked industry.

an exception, but we include it because it is significant in every other specification examined in the 'Sensitivity Analysis' section.) We distinguish between manufacturing and natural resources because the natural resource scale estimates may reflect not only scale, but also Heckscher-Ohlin misspecification. The constant returns to scale (CRS) group contains those industries with $\widehat{\mu}(g)$ that are insignificantly different from unity at the 1% level. In a few cases, such as electricity where $\widehat{\mu}(g) = 1.04$ and $t = 0.67$, being in the CRS group really means that scale returns are imprecisely estimated. The 'Non-Robust' group collects industries for which scale is not robustly estimated. Robustness will be made precise in the 'Sensitivity Section' section below.

Table 3 shows a number of striking patterns. First, the constant returns to scale industries are all sensible. These include Apparel, Leather, and Footwear. Second, the industries estimated to display scale returns are also all sensible. These include Pharmaceuticals, Electric and Electronic Machinery, and Non-Electrical Machinery.

To get a handle on magnitudes, consider a scale elasticity of 1.15. This implies that a 1% rise in output leads to a 0.13% fall in average cost. Further, a country operating at a tenth of U.S. levels faces 55% higher average costs. These are large numbers.

How do our results compare to existing partial equilibrium production function-based estimates? Tybout (2000) surveyed studies for many developing countries and found little evidence of scale returns. This is a key finding. He partly attributed it to the fact that "small firms in developing countries tend not to locate in those industries where they would be at substantial cost disadvantage relative to larger incumbents" (Tybout 2000, page 19). This squares readily with our finding of constant returns to scale in all low-end manufacturing industries. The evidence of scale from middle-income countries is less clear cut. Using Chilean plant-level data for 8 industries, Levinsohn and Petrin (1999) estimated scale returns from value-added production functions that are never less than 1.20 and reach as high 1.44. Using Mexican plant-level data, Tybout and Westbrook (1995) obtained almost no evidence of scale for large plants. It is not immediately clear that our results are inconsistent with Tybout and Westbrook: we find constant returns for many industries and increasing returns primarily for industries not examined by them. For rich countries, Harrigan (1999) found no evidence of scale in cross-country regressions. However, many other OECD country studies point clearly to the existence of scale returns. Paul and Siegel (1999) estimated industry-level scale returns in the range of 1.30 for many U.S. manufacturing industries. Estimates closer to our mode of 1.15 are common in OECD studies e.g., Fuss and Gupta (1981) for Canada and Griliches and Ringstad (1971) for Norway. Thus, our general equilibrium results are consistent with some, though not all of the existing partial equilibrium benchmarks.

3.3. Grouped Results

A drawback to the approach of the last section is that we did not simultaneously estimate each industry's scale elasticity. The estimation algorithm is only partially simultaneous. Further, we did not deal with endogeneity. To address these issues we follow table 3 in classifying industries into three groups: IRS, CRS, and Non-Robust.[10] We then estimate the model under the assumption that scale elasticities are the same *within* each group, but different *between* groups. Note that the IRS group includes both manufacturing and natural resource-based IRS industries. Disaggregation

[10]The classification of industries into these 3 groups is *slightly* different from that reported in table 3. This is a minor point; however, we cannot properly explain it until after we have laid out the criteria for being in the 'Non-Robust' group. See the 'Sensitivity Analysis' section below.

of the IRS group into manufacturing and natural resources adds little to the analysis.

The top panel of table 4 reports the results for the usual specification with 71 countries and 11 factors. The NL2SLS estimator uses the preferred instrument set from row 3 of table 2. That is, it uses the $F^*_{fi,t-\tau}(\mu)$ instrument set with a $K = 3$ polynomial order and a 15-year lag length. This leaves us with only the years 1987 and 1992. We do not report any of the other specifications that appeared in table 2 since this would be too repetitive. Instead, we report the results for the sample consisting only of our 23 OECD members. (See table A.1 for a list of these countries and the end of appendix A for a discussion of how the world is defined with 23 countries.) From the bottom panel of table 4, the 23-country results are very similar to the 71-country results, though somewhat less significant. The reduced significance arises from having eliminated an important source of sample variation, namely, trade with non-OECD countries such as Hong Kong and Singapore. For 23 countries, the Hausman χ^2_1 test statistic of 490 strongly supports the hypothesis of endogeneity.

The conclusions from table 4 are clear. The IRS group is always estimated to have significant scale returns. As in table 2, the NL2SLS estimate is larger than its ML counterpart which assume exogeneity. The CRS group is always estimated to have a scale elasticity that is insignificantly different from unity. For the Non-Robust group, almost by definition, the conclusion depends on the estimation method. Overall, the table 4 results are wholly consistent with the estimates reported in tables 2-3. *By implication we must abandon empirical models that treat all industries as if they were subject exclusively to either constant or increasing returns. Both play an important role for understanding the sources of comparative advantage*.

3.4. What is Scale?

As discussed in the introduction, we are using scale (and 'all that') to mean something more than just plant-level economies. For one, scale likely includes industry-level externalities. By way of example, growth of the Petroleum Refining industry was accompanied by the development of a host of specialized inputs including specialized engineering firms, industry-sponsored institutes, and specialized machinery manufacturers. Paul and Siegel (1999) argued that more than half of the relationship between cost and scale in U.S. manufacturing is due to such industry-level externalities. To the extent that this holds worldwide, it implies that a portion of what we are calling 'scale' is actually industry-level externalities.

Also, scale likely includes aspects of dynamic international technology differences. In table 3, the industries with the largest scale estimates are often those where technical change has been most rapid e.g., Pharmaceuticals. New goods often engender new process technologies and these new process technologies are often embodied in larger plants. As Rosenberg (1982, 1994) has tirelessly argued:

> "In this respect it is much more common than it ought to be to assume that the exploitation of the benefits of large-scale production is a separate phenomenon independent of technological change. In fact, larger plants typically incorporate a number of technological improvements ..." (Rosenberg 1994, page 199)

While it is important conceptually to distinguish between scale and dynamic technical change, the fact that technical change is the hand maiden of scale makes this distinction empirically problematic. Even with the unusually detailed McKinsey data, Baily and Gersbach (1995) were forced to lump international technology differences together with scale as inseparable sources of global competition. In short, for the most dynamic industries, available data do not allow us to distinguish between scale and scale-biased technical change. This is a critical area for future research.

Is it possible that our scale estimates also capture more traditional or *static* international technology differences? At the most obvious level the answer is no. Our π_{fi} explicitly capture the most important static international technology differences. A more subtle approach to answering the question is as follows. Suppose that country i is particularly efficient at producing a good in the sense of having a large ratio of output to employment. By Ricardian comparative advantage, the country will expand output of the industry. Looking across countries for a single industry, the larger is the level of industry output, the larger is the ratio of output to employment. The careless researcher will incorrectly attribute this correlation to increasing returns to scale. Obviously we have not been careless. However, to dispel any possible confusion we make the following observations.

The first observation draws on an analogy between our general equilibrium econometric work and the econometrics of production functions.[11] Static international technology differences can be thought of as unobserved productivity shocks that effect decisions about both inputs and outputs. They therefore induce endogeneity bias in regressions of output on inputs (e.g., Griliches and Mairesse 1995). The techniques commonly used to control for unobserved productivity shocks are fixed effects and IV estimators. These are exactly the econometric techniques we have used throughout this paper. *These techniques provided no evidence that our scale estimates are arti-*

[11]We hasten to add that the analogy with partial equilibrium production function estimation only goes so far. In our general equilibrium setting, observations have a factor dimension and data from all industries enter into each observation. See equation (13).

facts of unobserved international productivity differences.[12]

The second observation about the correct interpretation of our results comes from the cross-industry distribution of our scale estimates. In the Ricardian interpretation, *all* relatively low-cost producers are large producers. That is, static international technology differences will masquerade as scale effects in *all* industries, be it Pharmaceuticals or Apparel. If so, we should have estimated significant scale effects in *every* industry. We found nothing of the sort. To the contrary, we found strong evidence of constant returns to scale in many industries e.g., Apparel. We thus find the Ricardian interpretation untenable.

To summarize this section, our scale estimates likely capture plant-level scale, industry-level externalities, and scale-biased technical change. However, the idea that our estimates of increasing returns to scale are really static international technology differences in disguise has about as much life in it as old Ricardo himself.

4. Sensitivity Analysis

In assigning industries to the IRS and CRS groups, we have been using a stringent set of criteria: the inference of increasing or constant returns was required to be robust across a wide variety of specifications. Industries that did not meet the criteria were unceremoniously dumped into the Non-Robust group. This section reviews the sensitivity analysis underlying our robustness criteria. In

[12]To the contrary, the IV estimates were sometimes larger than the ML estimates. This IV-ML ranking deserves further consideration. In a production function setting, endogeneity bias is usually taken to mean that the coefficient on labour is upward-biased and the coefficient on less variable inputs such as capital are downward biased (Olley and Pakes 1996). It need *not* imply that the sum of the coefficients (i.e., the scale elasticity) is upward biased. Indeed, Levinsohn and Petrin (1999) found that a modified Olley-Pakes solution to endogeneity bias *raises* the estimate of scale in 3 of 8 Chilean industries examined. Thus, there is nothing terribly unusual about our IV-ML ranking.

the interests of space, the discussion is terse.

There are three issues to be clarified before robustness can be fully defined. The first deals with whether the results depend on the initial ranking of industries used to kick off the section 3.2 estimation algorithm. Consider table 5. Results in the 'Initial Rank: Baseline' column with 71 countries and 11 factors reproduce the table 3 results. A single asterisk indicates that $\widehat{\mu}(g)$ is greater than unity at the 1% level, two asterisks indicate $t > 5$ and three asterisks indicate $t > 10$. The baseline initial rank uses an initial ranking that is an average of industries' capital-intensity and skill-intensity ranks. See the last column of table 1. Results in the 'Initial Rank: Alternative' column use an initial ranking based on the scale parameters reported in Paul and Siegel (1999).[13] The baseline and alternative initial ranks are very different. The correlation is only 0.13. It is therefore re-assuring that the scale estimates produced by the two initial ranks are almost identical. That is, the table 3 results are insensitive to the initial ranking.

The second robustness issue deals with parameter stability across a variety of specifications. Table 3 reports results for the specification with 71 countries, 11 factors, and 5 years. For robustness, we require similar results from a specification using only the 23 OECD countries in our sample. We also require similar results from a specification that places more weight on capital and labour and less weight on land and energy i.e., a specification with 7 factors (aggregate land, aggregate energy, capital, and 4 types of labour). The last two columns of table 5 provide the scale estimates for these two specifications.

[13] Paul and Siegel (1999) work at a higher level of aggregation (19 manufacturing industries). We therefore repeated their scale elasticity values at the disaggregated level (27 manufacturing industries) where necessary. For industries not covered by Paul and Siegel (7 non-manufacturing industries), we set the scale elasticities to unity. In the case of ties we used information from the baseline initial rank.

Looking across the four table 4 specifications, the following emerges. There is considerable stability across specifications for the 'IRS - Manufacturing' industries and somewhat less stability for the 'IRS - Natural Resources' industries. The least amount of stability is for the 'Non-Robust' category. It is composed of industries that display significant increasing returns in one specification and constant returns in another. By definition, the Non-Robust group displays the greatest instability. Table 5 over-represents instability by omitting the CRS industries listed in table 3. By definition of the CRS group, its member industries always display constant returns (insignificant $\widehat{\mu}(g)$). Hence scale estimates for this group are very stable across specifications.

Before turning to our last robustness issue we will need to comment on the convergence of our section 3.2 algorithm. All the results reported in this paper are based on 10 iterations of the algorithm. To examine convergence, at least for the table 3 specification, we allowed the algorithm to run the extra two weeks needed to complete 20 iterations. The results were unchanged from those reported in table 3. Thus, we have restricted ourselves to 10 iterations of the algorithm.

The third and last robustness issue arises from the fact that the estimation algorithm converges to a cycling pattern for some industries. For example, in the baseline initial rank specification with 71 countries and 11 factors, the Basic Chemicals and Plastic Products industries either both display increasing returns (as in odd iterations of our algorithm) or both display constant returns (as in even iterations). Cycling could never happen in traditional industry-by-industry production function estimation and captures in an obvious way the fact that we are estimating scale returns in a general equilibrium setting. In tables 3 and 5, the cycling industries are added to the Non-Robust group. In the absence of cycling it does not matter whether we report the $9th$ or $10th$ iterations. Thus, for the IRS category we only report the $10th$ iteration. For those Non-Robust industries that cycle, it does matter which iteration we report. In order to keep the table manageable, for the Non-Robust group we report the average value of $\widehat{\mu}(g)$ across the $9th$ and $10th$ iterations.

The classification of industries is based on robustness *across* specifications. However, for a *given* specification some of the Non-Robust industries properly belong in the CRS or IRS groups. In table 5, these CRS and IRS industries are indicated by the absence of an entry and by a †, respectively. These Non-Robust industries are classified accordingly in table 4. For example, in the table 4 specification with 71 countries, Furniture and Fixtures is included in the CRS group.

We can now explain exactly how industries are classified. For each industry consider the 8 estimates of $\mu(g)$ that come from 4 specifications (the 4 column headings in table 5) and 2 iterations of our algorithm (the $9th$ and $10th$ iterations). If all 8 of these $\widehat{\mu}(g)$ are insignificant at the 1% level then the industry is classified as CRS. If all 8 of these $\widehat{\mu}(g)$ always exceed unity and do so significantly at least once, then the industry is classified as IRS. If the industry fails to meet either criteria then it is classified as Non-Robust.[14] *This scheme ensures that only the most robust results appear in the CRS and IRS groups*.

To conclude, table 5 pointedly illustrates two important features of our work. First, for about a third of all industries, the data are simply not informative about scale economies. These are the industries that appear in the Non-Robust group. Second, our criteria for classifying an industry as CRS or IRS is that the inference about scale is the same across many different specifications. For

[14]This classification criterion is for manufacturing industries. For natural resource-based industries the criterion is the same except that we omit from consideration the specification with 71 countries and 7 factors. With energy and land endowments so heavily aggregated, the 7-factor data convey very little information about natural resource-based industries. This can be seen from table 5.

these industries, the inference about scale is remarkably robust.

5. Trade and Wages

We have been assuming that scale effects are the same across factors. This masks non-homotheticities such as skill-biased scale effects. With factor non-homotheticities scale is more difficult to estimate, but the elasticity of unit factor demand $\alpha_{fg} \equiv -\partial \ln a_{fgi}/\partial \ln Q_{gi}$ remains a straightforward concept.[15] To examine factor non-homotheticities we drop our prevailing assumption that α_{fg} is independent of f and instead assume that it is independent of g.

Table 6 reports the ML estimates of α_f for the specification with 71 countries and 5 years. There are only 355 observations. We do not report the instrumental variables estimates since in all cases the Hausman test easily rejects endogeneity. The table evidences substantial non-homotheticities across factors. A 1% rise in output leads to a fall of 0.21% in the demand for high school dropouts per unit of output and a rise of 0.21% in the demand for high school graduates. That is, the ratio of skilled to unskilled labour rises by an enormous $0.42\% = (0.21 - (-0.21))$. When further disaggregated we see the same pattern magnified. Larger output is associated with skill-biased demand. In their detailed study of the telecommunications industry, Denny and Fuss (1983) also found skilled-biased output effects.

Skill-biased output effects have important implications for the trade-and-wages debate. To the extent that trade leads to specialization and greater output, it also leads to relatively greater demand

[15] Dropping g subscripts, the difficulty is that equation (10) must be replaced by $\mu = 1/(1 - \sum_f \alpha_f \theta_f)$ where θ_f is the share of factor f in total costs. The complication is that whereas $\Sigma_f \theta_f = 0.274$ can be read straight off of input-output tables, θ_f is not available for factors individually.

for skilled labour and hence to rising earnings inequality. Our skill-biased result also has implications for the rise of ‘intra-mediate’ trade (see section 1.2) since it is consistent with the Feenstra and Hanson (1996a,b) picture of offshore sourcing as a cause of rising earnings inequality.

6. The Case of the Missing Trade and Other Mysteries

The fully simultaneous model that we have presented is sparsely parameterized: it has only three scale parameters. It is thus not surprising that we do not dramatically revitalize the HOV model. Consider the specification with 71 countries, 11 factors, and 5 years from the first row of table 4. We use three statistics to evaluate it: (*i*) Trefler’s (1995) ‘missing trade’ statistic; (*ii*) the correlation between the factor content of trade and its endowment predictor; and, (*iii*) the Bowen *et al.* (1987) sign test. For the unmodified model (constant returns and $\pi_{fit} = 1$), the three statistics are 0.003, 0.14, and 0.67, respectively. For the modified model with three groups and $\pi_{fit} \neq 1$ (row 1 of table 4) the statistics are 0.117, 0.40, and 0.66, respectively. That is, there is a significant improvement in the missing trade and correlation statistics. Much of the improvement in the correlation statistic comes from introducing the π_{fit}. The improvement in the missing trade statistic comes from both the π_{fit} and the scale parameters. It is clear why there is less missing trade. As noted at the end of section 3.1, even small scale estimates translate into large international differences in the factor inputs needed per unit of output. It is interesting that the missing trade statistic improves most for skilled labour (secondary and post-secondary education) and not at all for workers with no education.

7. Conclusions

Do scale economies contribute to our understanding of the factor content of trade? To answer this question we examined Helpman and Krugman's variant of the Vanek factor content prediction. The prediction is particularly interesting in that it arises from a large class of general equilibrium, increasing returns to scale models. To examine the prediction, we proposed an estimating equation that is tightly allied to the theory. The equation embodies three determinants of trade: endowments, exogenous international productivity differences (π_{fi}), and increasing returns to scale. Admittedly, this list is incomplete. However, in our quest for truly general equilibrium estimates of scale, the requirement of tractability forced us to limit our consideration to only these three determinants of trade. This said, our empirical results strikingly demonstrate that scale economies must figure prominently for any understanding of the factor content of trade.

We also asked a second question. *What does international trade reveal about the extent of scale economies*? Much less of our formal structure was needed to answer this second question. Consistent estimation of scale economies primarily requires a standard orthogonality condition on the ε_{fi} of equation (2). While the economics of this condition are not entirely clear, it is no surprise that less structure is needed for estimator consistency. After all, the guiding empirical insight of this paper is that trade moves from low-cost exporters to high-cost importers. By implication, trade data must convey information about costs and scale elasticities.

Just as you cannot squeeze water from a stone, you cannot infer scale without extensive data. To this end, we constructed a unique and comprehensive database on trade flows, output, and factor endowments for 71 countries over the 1972-92 period. Even with this detailed data, we were unable to completely disentangle the factors underlying our industry-level scale estimates. The remaining entangled factors are plant-level scale, industry-level externalities and, for the most dynamic industries, scale-biased technical change. This list is the 'all that' of our title and a primer for future research.

Our main conclusions are as follows. (*i*) We estimated that output expansion is strongly skill-biased. This has important and hitherto unnoticed implications for rising wage inequality. (*ii*) Pooling across industries, we precisely estimated moderate scale elasticities ($\mu = 1.05$). This has important implications for endogenous growth, particularly since our sample of countries spans the entire development spectrum. (*iii*) At least a third of all goods-producing industries display constant returns to scale. For this group, scale does not contribute to our understanding of international trade. Another third of all goods-producing industries are characterized by increasing returns. The modal range of scale elasticities for this group is 1.10–1.20. For this group scale is central to our understanding of the factor content of trade.

Our results point to the importance of integrating constant and increasing returns to scale industries within a single general equilibrium framework. We showed how to implement this empirically. Further, we found that the Helpman-Krugman framework provides a remarkable lens for viewing the scale elasticities encoded in trade flows. Finally, our results highlight the importance of scale and 'all that' as a source of comparative advantage.

Appendix

A. Data and Their Sources

Data on endowments are constructed as follows. Capital stocks are from Summers and Heston (1991). When available, we used the capital stock series. Otherwise, we used Leamer's (1984) double declining balance method applied to investment. Educational-attainment in the population is from Barro and Lee (1993). We updated it to 1992. Energy is the sum of joule-equivalent reserves of hard coal, soft coal, crude oil, natural gas, and hydroelectric potential. Data are from the World Resources Institute diskette and other minor sources. Land is the sum of cropland, pasture, and forests as reported in FAO diskettes. All endowments are stocks, not flows.

For each country, industry, and year we have assembled data on international trade and production. Table A.1 lists the countries in the database. (South Korea was inadvertently omitted from our group of 23 OECD countries.) Trade data are from the Statistics Canada "World Trade Database." Production data are from the UNIDO INDSTAT production data base, from the UN General Industrial Statistics (including its earlier incarnations), and from other minor sources. Our 34 industries completely cover the tradeables sector except for metal mining, non-metal mining, and miscellaneous manufacturing. These three have no sensible production data because of aggregation problems. As a result they are included in the analysis, but are treated as constant returns to scale industries $(\mu = 1)$ so that the estimating equations are independent of their data. We also applied the same $\mu = 1$ constraint to the Crude Petroleum and Natural Gas industry for the 23-country results in table 4. This makes little difference to the results, but seemed the sensible way to deal with trade flows given that (*i*) the 23-country group omits the major oil producers and (*ii*) the 23-country $\widehat{\mu}(g)$ for Crude Petroleum and Natural Gas is $\mu = 1.01$. Manufacturing output was converted into 1987 U.S. dollars using IFS and PWT exchange rate data, PWT PPPs, and BEA industry-level price indexes. Agricultural output was converted using FAO data. Other non-manufacturing output was in physical units.

The U.S. choice-of-techniques matrices are from various sources. $b_{hg,\mathrm{US}}$ in equation (9) is from the benchmark input-output tables (1972, 1977, 1982, 1987, and 1992). The $d_{fg,\mathrm{US}}$ in equation (8) are from many sources. For labour they come from the U.S. 'employment and earnings' series combined with Current Population Surveys. For capital they primarily come from the U.S. 'fixed reproducible tangible wealth' series. For all other factors they come from the usual unique-use assumption. For example, the coal mining industry is the only industry that uses coal reserves. The $A_{f,\mathrm{US}}$ were deflated by the theoretically correct method of double-deflating using BEA price indexes. This is equivalent to putting both the input-output tables and the trade data into 1987 U.S. dollars.

For θ_g in equation (10), we use the 1982 U.S. value of θ_g calculated across our 34+3 industries. From the U.S. input-output table, this value is $\theta_g = 0.274$.

s_i is defined as follows. Let RGDPC$_i$ be real GDP per capita in constant dollars using the chain index (1985 international prices in the Penn World Tables). Let POP$_i$ be population from the Penn World Tables. Let $Y_i \equiv$RGDPC$_i\times$POP$_i$. Let PM_{gij} be the value of i's imports of good g from country j. Let $TB_i \equiv \Sigma_g \Sigma_j (PM_{gji} - PM_{gij})$ be i's trade balance. Then $s_i \equiv (Y_i - TB_i)/\Sigma_j(Y_j - TB_j)$. Note that the summation over countries in the definition of TB_i is the sum over the $N = 23$ or $N = 71$ countries in the database. The same point applies to all the summations in the paper, including summations implicit in X_j^c, X_j^y, and M_j^y.

B. Proof of Equation (1)

Recall that M_i^y is country i's imports of foreign-produced intermediate inputs. Let Y_{ii} be i's use of domestically produced intermediate inputs. Let $Y_i \equiv Y_{ii} + M_i^y$. The inter-industry shipments matrix B_i is constructed to identically satisfy $Y_i \equiv B_i Q_i$ so that

$$B_i Q_i \equiv Y_{ii} + M_i^y. \tag{14}$$

In input-output analysis, gross output is defined as $Q_i \equiv (C_{ii} + X_i^c) + (Y_{ii} + X_i^y)$. Substituting equation (14) into this yields $(I - B_i)Q_i = C_{ii} + X_i^c + X_i^y - M_i^y$. Premultiplying this last equation by $A_{fi}^* \equiv \pi_{fi} D_{fi} (I - B_i)^{-1}$, simplifying with $\pi_{fi} D_{fi} Q_i = \pi_{fi} V_{fi} \equiv V_{fi}^*$, noting that $M_{ii}^c \equiv 0$, and adding to the result $\Sigma_{j\neq i} A^*_{fj} C_{ij} - \Sigma_j A^*_{fj} M^c_{ij} = 0$ yields

$$V^*_{fi} = F^c_{fi} + A^*_{fi}(X^y_i - M^y_i) + \Sigma_j A^*_{fj} C_{ij} \tag{15}$$

where $F^c_{fi} \equiv A^*_{fi} X^c_i - \Sigma_j A^*_{fj} M^c_{ij}$. Summing equation (15) across countries i and multiplying by s_i yields

$$s_i V^*_{f\mathrm{w}} = s_i \Sigma_j F^c_{fj} + s_i \Sigma_j A^*_{fj}(X^y_j - M^y_j) + s_i \Sigma_j A^*_{fj} C_{\mathrm{w}j}. \tag{16}$$

Subtracting equation (16) from equation (15) and using $\Sigma_j F^c_{fj} = 0$ yields equation (1).

By the usual input-output logic, the delivery of one unit of final demand requires $(I - B_i)^{-1}$ units of gross output. Since B_i includes imported intermediates (see equation 14), the required gross output includes *foreign-produced* intermediates. $A_{fi} \equiv D_{fi}(I - B_i)^{-1}$ is thus the amount of factor f that would have been used to produce the required gross output had the required output all been produced domestically. Since, roughly speaking, M^y_i of it was produced abroad, using A_{fi} inflates the factor content of trade by an amount $A_{fi} M^y_i$. The equation (3) factor content expression corrects for this by including terms like $-A_{fi} M^y_i$.

C. Trade in Intermediate Inputs

In this appendix we investigate the role of our assumption about the share of trade that is trade in intermediate inputs. Let M^c_{gijt} and M^y_{gijt} be country i's imports of good g from country j in year t for final goods and intermediate inputs, respectively. We observe $M_{gijt} \equiv M^c_{gijt} + M^y_{gijt}$, but not M^c_{gijt} or M^y_{gijt}. The problem is to choose an appropriate value for the share of intermediate inputs in total trade, $\gamma_{gijt} \equiv M^y_{gijt} / M_{gijt}$.

We begin by noting that the bulk of world trade is trade in intermediate inputs so that we might conservatively expect the γ_{gijt} to exceed one half. Consider Canada-U.S. bilateral trade since we are most familiar with these data. Careful examination of these data reveals that *at least* two thirds of it is intermediates trade.[16] This conclusion is consistent with the aggregate numbers. For the

[16]Government of Canada (1997) reports the average of imports and exports for the following categories. Consumer goods account for 7% of trade. None of this is intermediate inputs. Machinery and industrial goods account for 42%

United States in 1987, intermediate inputs plus private investment as a percentage of gross output was 53%. The percentage is higher for the goods-producing tradeable sector i.e., for the industries that concern us. Thus, we expect γ_{gijt} to exceed one half on average. We emphasize that one half is a conservative number.

We next turn to our choice of γ_{gijt}. Let $Q_{g\mathrm{US}t}$, $Y_{g\mathrm{US}t}$, and $I_{g\mathrm{US}t}$ be U.S. gross output, intermediate inputs, and investment in industry g in year t. We assume that γ_{gijt} equals $\gamma^*_{gijt} \equiv (Y_{g\mathrm{US}t} + I_{g\mathrm{US}t})/Q_{g\mathrm{US}t}$. γ^*_{gijt} ranges from 0.28 to 1.00 across our 34 industries and 5 years.

Table A.2 explores the sensitivity of our results to different choices of γ_{gijt}. The first row of the table reports the ML estimate $\widehat{\mu}$ for the model with 71 countries, 11 factors, 5 years, and our baseline assumption $\gamma_{gijt} = \gamma^*_{gijt}$. It is the same $\widehat{\mu}$ as in the first row of table 2. The second row uses a γ_{gijt} that is halfway between γ^*_{gijt} and unity. The third row uses a γ_{gijt} that is halfway between γ^*_{gijt} and zero. *It is clear that such radical changes in the choice of γ_{gijt} make little difference to our estimates of scale.*

The remaining rows of table A.2 set γ_{gijt} to a constant. It is our strong contention that γ_{gijt} averaged across industries should exceed one half. From table A.2, allowing γ_{gijt} to fall to one half makes no difference to $\widehat{\mu}$. Further, the loglikelihood value declines monotonically as γ_{gijt} moves from unity to zero. This suggests, as expected, that the model fits the data better for large values of γ_{gijt}. *To conclude, our results are robust to a wide range of assumptions about the allocation of trade between intermediate inputs and final goods.*

D. GLS Variance Parameters

The estimating equation (13) involves a GLS correction σ_{fit}. We parameterize σ_{fit} as $\sigma_{fit} \equiv s^{\omega}_{it}\sigma_{ft}$ where s^{ω}_{it} is a control for country size and σ_{ft} is a control for factor size. ω is a parameter to be

of trade. All of this is intermediate inputs. Automotive goods account for 29% of trade. From the Canadian 1988 input-output tables, under half of this or 12% is trade in parts. Parts are intermediate inputs. Assuming conservatively that no finished automotive goods are intermediate inputs (e.g., treating trucks as final goods), only 12% of automotive trade is intermediate inputs. 'Other' goods account for 23% of trade. 'Other' goods is dominated by forestry products which in turn is primarily used as intermediates in the construction industry. We conservatively assume that only half or 11% of 'Other' goods trade is intermediate inputs. The sum of 42% plus 12% plus 11% is 65%. That is, at least two-thirds of Canada-U.S. trade is intermediates trade.

estimated. σ_{ft} is defined as the cross-country variance of $H_{fit}(\mu = 1)/s_{it}^{\omega_0}$ where H_{fit} is defined in equation (12). ω_0 was chosen by an iterative procedure: pick an initial ω_0, estimate the model to obtain an $\widehat{\omega}$, set ω_0 equal to $\widehat{\omega}$, and repeat until ω_0 converges to $\widehat{\omega}$. It quickly became clear that the $\widehat{\omega}$ always lie in the tight interval $(0.82, 0.93)$. As a result, we set ω_0 equal to 0.9 for all the estimates reported in the paper. (Setting ω_0 to 0.80 or 1.00 makes no difference.) The $\widehat{\omega}$ for the ML specifications in tables 2 and 4 appear in table A.3. All of them are close to 0.90. The $\widehat{\omega}$ underlying the table 3 estimates vary between 0.85 and 0.93 across the 34 industries.

For NLS and NL2SLS, one typically replaces ω with a consistent estimate. We chose $\omega = 0.9$.

Allowing the variance to vary additionally across factors and years or introducing factor fixed effects makes no difference to the estimates. For one, the effect of scaling by σ_{ft} is that factors have similar variances in each year. For another, $H_{fit}(\mu = 1)$ has a zero mean for each factor.

E. Details of the NL2SLS Estimators

Let $t = 1, \ldots, t_0$ be the years lost as a result of instrumenting using lagged regressors. We are interested in 'within' estimators. Let x_{fit} be any variable and let ∇ be an operator that transforms x_{fit} by subtracting off its mean across all factors $f = 1, ..., F$ and years $t = t_0 + 1, ..., T$. That is, $\nabla x_{fit} \equiv x_{fit} - \Sigma_{f'} \Sigma_{t' > t_0} x_{f'it'} / (F(T - t_0))$.

The NL2SLS estimator is based on Amemiya (1974). Consider the '$Z = F^*_{fi,t-\tau}(\mu)$' estimator. From equation (13), let $\nabla\eta(\mu)$ be the $FN(T - t_0) \times 1$ vector formed by stacking the $\nabla\eta_{fit}(\mu) = \nabla(H_{fit}(\mu)/\sigma_{fit})$. Let $Z^K_{fit}(\mu) \equiv [\nabla(F^*_{fi,t-t_0}(\mu)/\sigma_{fit}) \cdots \left(\nabla(F^*_{fi,t-t_0}(\mu)/\sigma_{fit})\right)^K]$ be the $1 \times K$ vector of instruments. Let $Z^K(\mu)$ be the corresponding $FN(T - t_0) \times K$ matrix of stacked instruments. The '$Z = F^*_{fi,t-\tau}(\mu)$' estimator minimizes $\nabla\eta(\mu)' P^K_Z(\mu) \nabla\eta(\mu)$ where $P^K_Z(\mu) \equiv Z^K (Z^{K\prime} Z^K)^{-1} Z^{K\prime}$. Let μ^K_0 be the minimizer. Since $Z^K(\mu)$ is a function of μ, the instruments must be simultaneously estimated along with the rest of the model. That is, NL2SLS is a 1-stage estimator, not a 2-stage one.

Consider the '$\partial\widehat{F}^*_{fit}/\partial\mu$' estimator. Let k be the dimension of μ, define the gradient $G_{fit}(\mu) \equiv \nabla((\partial F^*_{fit}/\partial\mu)/\sigma_{fit})$, and let $G(\mu)$ be the $FN(T - t_0) \times k$ matrix of stacked $G_{fit}(\mu)$. In the first stage, instrument $G(\mu^K_0)$ by $Z^K(\mu^K_0)$ to obtain $\widehat{G} \equiv P^K_Z(\mu^K_0) G(\mu^K_0)$. In the second stage, minimize

$\nabla\eta(\mu)' P_G(\mu_0^K)\nabla\eta(\mu)$ where $P_G(\mu_0^K) \equiv \widehat{G}(\widehat{G}'\widehat{G})^{-1}\widehat{G}'$. Denote the minimizer by μ_1^K. In table 2, μ_1^K depends on K only via μ_0^K. Also, the table 2 conclusions are robust to expanding the instrument set $\widehat{G}$ to include polynomials and cross-products of $\widehat{G}$.[17]

t-statistics for the two NL2SLS estimators are based on the covariance matrices $\widehat{\sigma}^2 \left(G(\mu_0^K)' P_Z^K(\mu_0^K)G(\mu_0^K)\right)^{-1}$ and $\widehat{\sigma}^2 \left(G(\mu_1^K)' P_G(\mu_0^K)G(\mu_1^K)\right)^{-1}$ where $\widehat{\sigma}$ is the model standard error (based on unprojected residuals, of course). The Hausman test is based on the difference between the NLS and NL2SLS estimates.

[17]In table 2, the standard errors of the two NL2SLS estimators are similar and occasionally the '$\partial\widehat{F}^*_{fit}/\partial\mu$' standard errors are (unexpectedly) larger. When squares and cross-products of $\widehat{G}$ are included in the instrument set, the '$\partial\widehat{F}^*_{fit}/\partial\mu$' standard errors are always lower, but not significantly so.

References

Amemiya, Takeshi (1974) "The Nonlinear Two-Stage Least-Squares Estimator," *Journal of Econometrics, 2* (2), 105–110.

—— (1975) "The Nonlinear Limited-Information Maximum-Likelihood Estimator and the Modified Nonlinear Two-Stage Least-Squares Estimator," *Journal of Econometrics, 3* (4), 375–386.

Antweiler, Werner and Daniel Trefler (1997) "Increasing Returns and All That: A View From Trade." Mimeo, University of Toronto.

Baily, Martin Neil and Hans Gersbach (1995) "Efficiency in Manufacturing and the Need for Global Competition," *Brookings Papers on Economic Activity, Microeconomics*, 307–358.

Barro, Robert and Jong-Wha Lee (1993) "International Comparisons of Educational Attainment," *Journal of Monetary Economics, 32* (3), 363–394.

Basu, Susanto (1995) "Procyclical Productivity: Increasing Returns or Cyclical Utilization?," *Quarterly Journal of Economics, 111* (3), 719–751.

—— **and John G. Fernald** (1997) "Returns to Scale in U.S. Production: Estimates and Implications," *Journal of Political Economy, 105* (2), 249–279.

Bowen, Harry P., Edward E. Leamer, and Leo Sveikauskas (1987) "Multicountry, Multifactor Tests of the Factor Abundance Theory," *American Economic Review, 77* (5), 791–809.

Brainard, S. Lael (1993) "An Empirical Assessment of the Factor Proportions Explanation of Multinational Sales." NBER Working Paper #4580.

—— (1997) "An Empirical Assessment of the Proximity/Concentration Tradeoff between Multinational Sales and Trade," *American Economic Review, 87* (4), 520–540.

Conway, Patrick J. (2000) "The Case of the Missing Trade and other Mysteries: Comment." Mimeo, University of North Carolina.

Davis, Donald R. (1995) "Intra-industry Trade: A Heckscher-Ohlin-Ricardo Approach," *Journal of International Economics, 39* (3-4), 201–226.

—— (1997) "Critical Evidence on Comparative Advantage? North-North Trade in a Multilateral World," *Journal of Political Economy, 105* (5), 1051–1060.

—— **and David E. Weinstein** (1996) "Does Economic Geography Matter for International Specialization?" NBER Working Paper #5706.

—— **and** —— (1998) "An Account of the Global Factor Content of Trade." NBER Working Paper #6785.

Deardorff, Alan V. (1979) "Weak Links in the Chain of Comparative Advantage," *Journal of International Economics*, *9* (2), 197–209.

Denny, Michael and Melvyn Fuss (1983) "The Effects of Factor Prices and Technological Change on the Occupational Demand for Labor: Evidence from Canadian Telecommunications," *Journal of Human Resources*, *18* (2), 161–176.

Epstein, Larry G. (1982) "Integrability of Incomplete Systems of Demand Functions," *Review of Economic Studies*, *49* (3), 411–425.

Ethier, Wilfred J. (1982) "National and International Returns to Scale in the Modern Theory of International Trade," *American Economic Review*, *72* (3), 389–405.

Feenstra, Robert C. and Gordon H. Hanson (1996a) "Foreign Investment, Outsourcing, and Relative Wages," in Robert C. Feenstra, Gene M. Grossman, and Douglas A. Irwin, eds., *The Political Economy of Trade Policy: Papers in Honor of Jagdish Bhagwati*, Cambridge MA: MIT Press.

—— and —— (1996b) "Globalization, Outsourcing, and Wage Inequality," *American Economic Review Papers and Proceedings*, *86* (2), 240–245.

—— and —— (1997) "Foreign Direct Investment and Relative Wages: Evidence from Mexico's Maquiladoras," *Journal of International Economics*, *42* (3-4), 371–193.

Food and Agricultural Organization of the United Nations (1992) "Intercountry Comparisons of Agricultural Output and Productivity: Revised Methodology and Empirical Results," *FAO Quarterly Bulletin of Statistics*, *5* (4), iii–vi.

Fuss, Melvyn A. and V. K. Gupta (1981) "A Cost Function Approach to the Estimation of Minimum Efficient Scale, Returns to Scale, and Suboptimal Capacity," *European Economic Review*, *15* (2), 123–135.

Government of Canada (1997) "NAFTA: A Partnership at Work." Department of Foreign Affairs and International Trade, http://www.dfait-maeci.gc.ca/nafta-alena/what2-e.asp.

Griliches, Zvi and Jacques Mairesse (1995) "Production Functions: The Search for Identification." NBER Working Paper #5067.

—— and Vidar Ringstad (1971) *Economies of Scale and the Form of the Production Function: An Econometric Study of Norwegian Manufacturing Establishment Data*, Amsterdam: North Holland.

Harrigan, James (1993) "OECD Imports and Trade Barriers in 1983," *Journal of International Economics*, *35* (1-2), 91–111.

—— (1996) "Openness to Trade in Manufactures in the OECD," *Journal of International Economics*, *40* (1-2), 23–39.

—— (1999) "Estimation of Cross-Country Differences in Industry Production Functions," *Journal of International Economics*, *47* (2), 267–93.

Helpman, Elhanan (1981) "International Trade in the Presence of Product Differentiation, Economies of Scale and Monopolistic Competition: A Chamberlin-Heckscher-Ohlin Approach," *Journal of International Economics*, *11* (3), 305–340.

—— (1987) "Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries," *Journal of the Japanese and International Economies*, *1* (1), 62–81.

—— **and Paul R. Krugman** (1985) *Market Structure and Foreign Trade*, Cambridge, MA: MIT Press.

Hummels, David and James Levinsohn (1993) "Product Differentiation as a Source of Comparative Advantage?," *American Economic Review Papers and Proceedings*, *83* (2), 445–449.

—— **and** —— (1995) "International Trade and Monopolistic Competition: Reconsidering the Evidence," *Quarterly Journal of Economics*, *110* (3), 799–836.

Jorgenson, Dale W. and Jean-Jacques Laffont (1974) "Efficient Estimation of Nonlinear Simultaneous Equations with Additive Disturbances," *Annals of Economic and Social Measurement*, *3* (4), 615–640.

Klepper, Steven and Edward E. Leamer (1984) "Consistent Sets of Estimates for Regressions with Errors in All Variables," *Econometrica*, *52* (1), 163–183.

Krugman, Paul R. (1979) "Increasing Returns, Monopolistic Competition, and International Trade," *Journal of International Economics*, *9* (4), 469–479.

—— (1991) *Geography and Trade*, Cambridge, MA: MIT Press.

Lancaster, Kelvin (1984) "Protection and Product Differentiation," in Henryk Kierzkowski, ed., *Monopolistic Competition and International Trade*, Oxford: Oxford University Press.

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

Levinsohn, James and Amil Petrin (1999) "When Industries Become More Productive, Do Firms? Investigating Productivity Dynamics." Mimeo, University of Michigan.

Markusen, James R. and Anthony J. Venables (1998) "Multinational Firms and the New Trade Theory," *Journal of International Economics*, *46* (2), 183–203.

Nelson, Charles R. and Richard Startz (1990) "Some Further Results on the Exact Small Sample Properties of the Instrumental Variables Estimator," *Econometrica*, *58* (4), 967–976.

Olley, G. Steven and Ariel Pakes (1996) "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, *64* (6), 1263–1297.

Paul, Catherine J. Morrison and Donald S. Siegel (1999) “Scale Economies and Industry Agglomeration Externalities: A Dynamic Cost Function Approach,” *American Economic Review*, *89* (2), 272–290.

Rosenberg, Nathan (1982) *Inside the Black Box: Technology and Economics*, Cambridge: Cambridge University Press.

—— (1994) *Exploring the Black Box: Technology, Economics, and History*, Cambridge: Cambridge University Press.

Rotemberg, Julio and Michael Woodford (1992) “Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity,” *Journal of Political Economy*, *100* (6), 1153–1207.

Summers, Robert and Alan Heston (1991) “The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988,” *Quarterly Journal of Economics*, *106* (2), 327–68.

Trefler, Daniel (1993) “International Factor Price Differences: Leontief was Right!,” *Journal of Political Economy*, *101* (6), 961–987.

—— (1995) “The Case of the Missing Trade and Other Mysteries,” *American Economic Review*, *85* (5), 1029–1046.

—— (1996) “The Structure of Factor Content Predictions.” Mimeo, University of Chicago.

Tybout, James R. (2000) “Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?,” *Journal of Economic Literature*, *38* (1), 11–44.

—— **and M. Daniel Westbrook** (1995) “Trade Liberalization and the Dimensions of Efficiency Change in Mexican Manufacturing Industries,” *Journal of International Economics*, *39* (1-2), 53–78.

Vanek, Jaroslav (1968) “The Factor Proportions Theory: The N-Factor Case,” *Kyklos*, *21*, 749–56.

Figure 1. Trade in Instruments (1992)

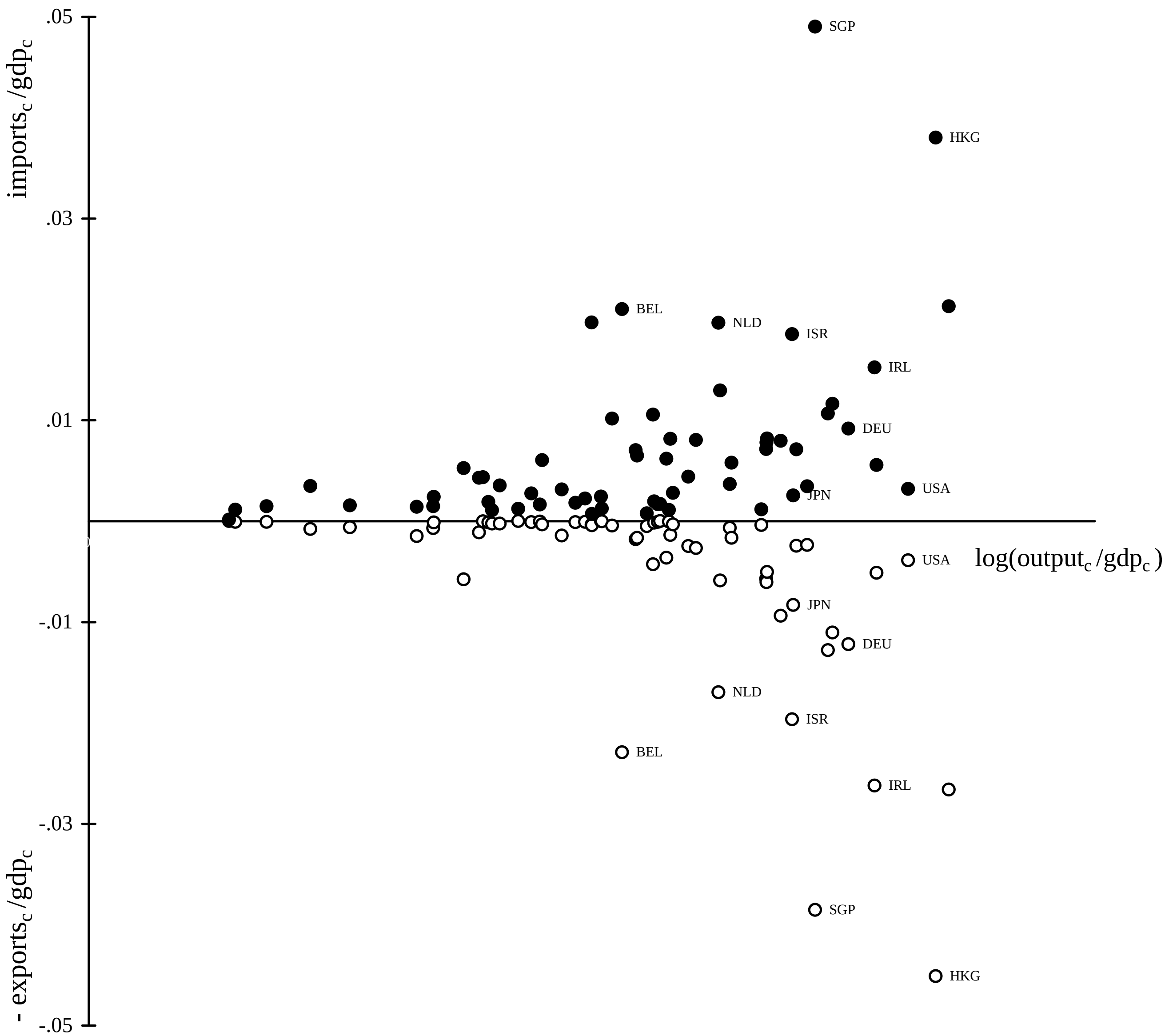

Figure 2. U.S. Bilateral Trade in Instruments (1992)

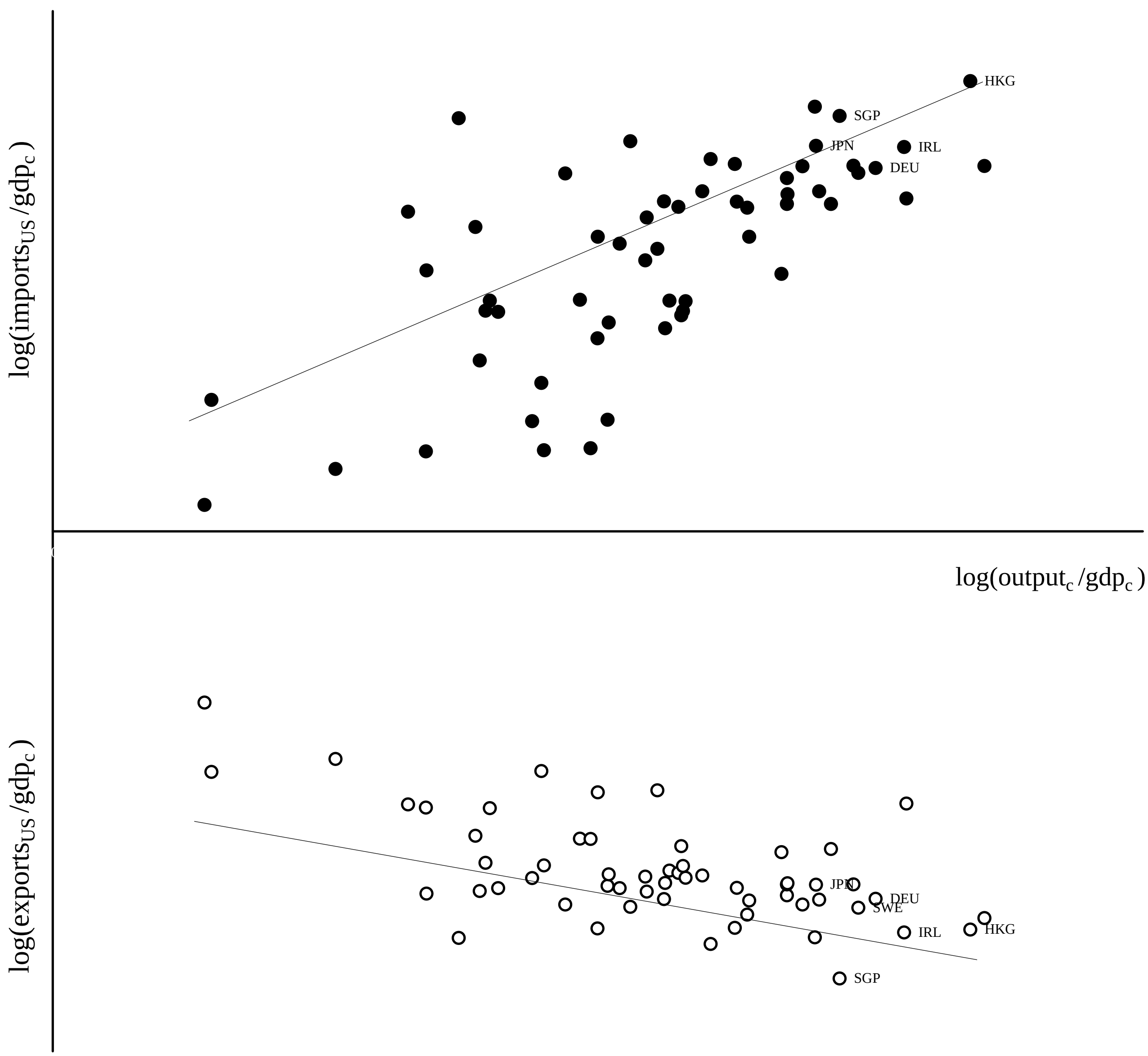

Table 1. The Geographical Concentration of Production

		Herfindahl Index		Measures of Scale		
ISIC	Industry	1992	1992-1972	Capital Intensity	Skill Intensity	Average Rank
385	Instruments	.38	+.07	.04	.42	22
210	Coal Mining	.23	–.03	.15	.18	24
342	Printing and Publishing	.17	–.03	.03	.43	21
220	Oil and Gas Mining	.15	–.20	.34	.31	30
382	Machinery (Non-Electrical)	.15	–.04	.05	.31	22
341	Pulp and Paper	.13	–.03	.11	.21	23
384	Vehicles	.13	–.07	.06	.30	22
356	Plastic Products	.12	–.01	.04	.16	14
383	Electrical-Electronic Mach.	.12	–.05	.06	.34	24
381	Metal Products	.11	–.04	.04	.15	13
331	Sawmill Products	.10	–.05	.03	.12	10
332	Furniture and Fixtures	.10	–.04	.02	.11	7
351	Basic Chemicals	.10	–.03	.14	.28	27
352	Pharmaceuticals	.10	–.02	.08	.49	28
354	Petro. and Coal Products	.10	–.03	.12	.42	29
	GDP	**.10**	**–.02**			
020	Crops	.09	+.01	.07	.06	12
314	Tobacco	.09	–.01	.14	.36	29
362	Glass Products	.09	–.05	.06	.16	19
411	Electricity	.09	–.03	1.00	1.00	35
010	Livestock	.08	–.01	.18	.11	21
031	Forestry	.08	+.00	.02	.11	5
322	Apparel	.08	–.07	.01	.06	2
353	Petroleum Refineries	.08	–.04	.51	.62	34
353	Rubber Products	.08	–.05	.05	.22	19
371	Iron & Steel Basic Indus.	.08	–.04	.13	.19	23
372	Non-Ferrous Metal Prod.	.08	–.07	.09	.23	23
311	Food	.07	–.04	.06	.11	14
313	Liquors	.07	–.02	.10	.25	24
321	Textiles	.07	–.01	.03	.10	7
361	Pottery and China	.07	–.02	.02	.21	14
032	Fishing	.06	–.02	.03	.10	7
323	Leather	.06	–.01	.02	.11	5
324	Footwear	.06	–.03	.01	.10	4
369	Clay & Cement Products	.06	–.05	.05	.15	14
Correlation with 1992 Herfindahl						
	Pearson			–.03	.23	.27
	Rank			.26	.51	.43

Notes: The Herfindahl index is a 1992 index of the cross-country dispersion in the location of output. Smaller values mean less regional concentration. Capital intensity is 1992 U.S. data on each industry's capital-labour ratio. Skill intensity is 1992 U.S. data on the ratio of skilled workers (completed high school) to unskilled workers (did not complete high school). The capital- and skill-intensity columns are scaled so that electricity is unity. Average rank is the average of the ranks of capital intensity and skill intensity.

Table 2. The Elasticity of Scale for Different Estimators: Pooling Across Industries

		ML		NLS		NL2SLS $Z = F^*_{fi,t-\tau}(\mu)$				NL2SLS $Z = \partial \widehat{F}^*_{fit}/\partial\mu$	
Row	Years	μ	t	μ	t	Lag Length	K	μ	t	μ	t
(1)	1972-92	1.051	13.41	1.050	12.56						
(2)	1987-92	1.045	6.75	1.044	6.11	15 Years	1	1.095	14.03	1.095	14.03
(3)							3	1.098	15.05	1.095	14.13
(4)							5	1.098	15.23	1.095	14.24
(5)	1982-92	1.048	9.47	1.047	8.67	10 Years	1	1.088	16.67	1.088	16.67
(6)							3	1.091	17.91	1.088	16.89
(7)							5	1.090	17.82	1.088	17.05
(8)	1992	1.037	3.58	1.037	3.21	20 Years	1	1.098	9.00	1.098	9.00
(9)							3	1.098	9.65	1.098	9.66
(10)							5	1.065	4.80	1.084	7.59

Notes: a. This table reports estimates of the elasticity of scale μ for the case where all industries are assumed to share a common μ. *t* is the *t*-statistic for the hypothesis of constant returns to scale (μ = 1).
b. Each estimate uses all 71 countries and all 11 factors. The 11 factors are capital, 4 types of labour (no education, primary education, secondary education, and post-secondary education), 3 types of land (cropland, pasture, and forest), and 3 types of energy (coal reserves, oil and gas reserves, and hydroelectric potential). The number of years T used varies across specifications. We use every fifth year of data over the period indicated in the 'Years' column. The number of observations is $71 \times 11 \times T = 781T$.
c. The NL2SLS columns report the instrumental variables results. The $Z = F^*_{fi,t-\tau}(\mu)$ instrument set uses lags of order τ given by the 'Lag Length' column. The instrument set also includes polynomials of the $F^*_{fi,t-\tau}(\mu)$ where the polynomial order is given by the 'K' column. The more complicated $Z = \partial \widehat{F}^*_{fit}/\partial\mu$ instrument set consists of fitted values of the $\partial F^*_{fit}/\partial\mu$ obtained from a preliminary NL2SLS procedure. See appendix E for details.

Table 3. Elasticities of Scale by Industry

	$\mu(g)$	t
IRS - Manufacturing		
Petroleum and Coal Products	1.403	35.70
Pharmaceuticals	1.306	17.52
Electric and Electronic Machinery	1.197	12.12
Petroleum Refineries	1.192	4.08
Iron and Steel Basic Industries	1.146	5.59
Instruments	1.124	1.69
Machinery (Non-Electrical)	1.113	6.14
IRS - Natural Resources		
Forestry	1.181	11.57
Livestock	1.075	9.54
Crude Petroleum and Natural Gas	1.050	6.52
Coal Mining	1.049	6.84
Constant Returns to Scale		
Apparel, Leather, Footwear, Food, Liquors, Sawmill Products, Fishing, Agricultural Crops, Textiles, and Electricity.		
Non-Robust		
Vehicles, Basic Chemicals, Pulp and Paper, Printing and Publishing, Plastic Products, Non-Ferrous Metal Products, Metal Products, Rubber Products, Clay and Cement Products, Glass Products, Pottery and China, Furniture and Fixtures, and Tobacco.		

Notes: This table reports the ML scale estimates for the specification with 71 countries, 11 factors (listed in the note to table 2), and 5 years. The $\mu(g)$ column reports the $\widehat{\mu}(g)$ scale elasticities. t-statistics are for the hypothesis of constant returns ($\mu(g) = 1$).

Table 4. Scale Elasticities for Groups of Industries

		IRS		Non-Robust		CRS	
		μ	t	μ	t	μ	t
71 Countries							
1972-92	ML	1.062	15.75	1.072	6.57	0.888	–1.67
1987-92	ML	1.060	9.06	1.075	4.10	0.775	–0.80
	NL2SLS	1.116	6.54	0.897	–0.02	0.700	-0.00
23 Countries							
1972-92	ML	1.066	10.76	1.108	4.34	0.846	–1.98
1987-92	ML	1.057	6.16	1.081	1.50	0.824	–1.33
	NL2SLS	1.101	5.18	0.876	–0.01	0.813	–0.02

Notes: *a*. All specifications use the 11 factors listed in the notes to table 2. The 23 countries are the OECD countries in our sample. See appendix table A.1 for a list of these countries. *b*. ML is maximum likelihood. NL2SLS is non-linear 2-stage least squares as reported in row 3 of table 2. That is, it uses the $Z = F^*_{fi,t-\tau}(\mu)$ instrument set with a $K = 3$ polynomial order and a 15-year lag. With a 15-year lag, only the years 1987 and 1992 are left.

Table 5. Sensitivity Analysis: $\mu(g)$ for Different Specifications

Countries	N=71	N=71	N=23	N=71
Factors	F=11	F=11	F=11	F=7
Initial Rank	Baseline	Alternative	Baseline	Baseline
IRS - Manufacturing				
Petroleum and Coal Products	1.40***	1.40***	1.38***	1.38***
Pharmaceuticals	1.31***	1.30***	1.33***	1.28***
Electric and Electronic Machinery	1.20***	1.20***	1.18*	1.19***
Petroleum Refineries	1.19*	1.19*	1.20	1.20*
Iron and Steel Basic Industries	1.15**	1.08	1.00	1.14*
Instruments	1.12	1.15*	1.24**	1.22***
Machinery (Non-Electrical)	1.11**	1.10*	1.15**	1.12**
IRS - Natural Resources				
Forestry	1.18***	1.18***	1.05*	1.18***
Livestock	1.08**	1.08**	1.15***	1.02
Crude Petroleum & Natural Gas	1.05**	1.05**	1.01	1.06**
Coal Mining	1.05**	1.05**	1.05**	0.97
Non-Robust				
Vehicles	1.02	1.02	1.09	0.97
Basic Chemicals	1.06	1.06	1.20†	1.02
Pulp and Paper	1.15	1.12		1.19†
Printing and Publishing	1.00	1.01		
Plastic Products	1.11	1.08	1.28†	1.06
Non-Ferrous Metal Products			1.24†	
Metal Products	1.02	1.00	1.02	
Rubber Products	1.23†	0.99†	1.13	1.24†
Clay and Cement Products	1.07	1.03	1.03	
Glass Products	1.02	1.02	1.00	1.02
Pottery and China	0.83	0.83	1.01	
Furniture and Fixtures			0.76	
Tobacco			1.08	

Notes: *a*. This table reports the ML scale estimates $\widehat{\mu}(g)$. There are 71 countries in the sample, 23 of which are OECD members. The 11 factors are listed in the note to table 2. The 7 factors aggregate the 3 land factors and aggregate the 3 energy factors. All 5 years of data are used.

b. The 'Baseline' initial rank columns initiate the estimation algorithm with a ranking based on industries' capital and skill intensities. The 'Alternative' initial rank column initiates the algorithm with a ranking based on scale elasticities from Paul and Siegel (1999).

c. A * indicates significance at the 1% level ($t > 2.58$). ** and *** indicate $t > 5$ and $t > 10$, respectively.

d. For the IRS groups, the data are from the 10*th* iteration. For the Non-Robust group, the data are averages of the 9*th* and 10*th* iterations. No *t*-statistics are reported for the Non-Robust group.

e. For the Non-Robust group an empty cell denotes constant returns (insignificant $\mu(g)$) in both iterations. A † denotes increasing returns (significant $\mu(g)$) in both iterations.

Table 6. The Elasticity of Factor Demand

Factor	α_f	t
High School Not Completed	0.21	1.40
High School Completed	–0.21	–2.08
No Education	0.83	9.65
Primary Education	–0.08	–0.28
Secondary Education	–0.05	–0.67
Post-Secondary Education	–0.27	–1.34

Notes: This table reports the estimates of $\alpha_f = \partial \ln a_{fgi} / \partial \ln Q_{gi}$ i.e., the output elasticity of unit factor demand. The estimates are ML results for a specification with 71 countries, 5 years, and the single factor indicated in the first column. There are $71 \times 5 = 355$ observations.

Table A.1: Countries in the Database

	Country	GDP per capita		Country	GDP per capita
*	United States	1.00		Brazil	0.22
	Hong Kong	0.92	*	Turkey	0.21
*	Canada	0.91		Costa Rica	0.20
*	Norway	0.87		Fiji	0.19
*	Japan	0.84		Colombia	0.19
*	Germany	0.82		Panama	0.19
*	Australia	0.81		Tunisia	0.17
*	Denmark	0.79		South Africa	0.17
*	Sweden	0.78		Ecuador	0.16
*	France	0.78		Suriname	0.14
*	Belgium	0.75		Jamaica	0.14
*	Netherlands	0.74		Dominican Rep.	0.13
*	Austria	0.72		Guatemala	0.13
*	United Kingdom	0.71		Sri Lanka	0.12
*	Italy	0.71		Morocco	0.12
	Singapore	0.71		Indonesia	0.12
*	Iceland	0.70		Peru	0.12
*	Finland	0.67		El Salvador	0.11
*	New Zealand	0.63		Egypt	0.10
	Israel	0.55		Bolivia	0.10
*	Spain	0.55		Philippines	0.09
*	Ireland	0.54		Papua New Guinea	0.09
	Barbados	0.43		Bangladesh	0.08
	Venezuela	0.40		Pakistan	0.08
*	Greece	0.37		Honduras	0.08
	Malta	0.35		India	0.07
*	Mexico	0.35		Zimbabwe	0.07
*	Portugal	0.34		Cameroon	0.06
	Mauritius	0.34		Nigeria	0.05
	South Korea	0.32		Ghana	0.05
	Malaysia	0.32		Zambia	0.04
	Argentina	0.30		Madagascar	0.03
	Uruguay	0.29		Tanzania	0.03
	Chile	0.27		Malawi	0.03
	Syria	0.24		Ethiopia	0.02
	Thailand	0.22			

Notes: A * indicates that the country is a member of the OECD. There are 23 members in our sample. GDP per capita is expressed relative to the United States. Most data are for 1992. Where these were not available, 1988 data were used or, in the case of Ethiopia, 1986 data.

Table A.2 Sensitivity to Assumptions About Intermediate Inputs

γ_{gijt}	μ	t	likelihood
γ^*_{gijt}	1.051	13.41	-35,461
$(\gamma^*_{gijt}+1)/2$	1.052	14.72	-35,453
$(\gamma^*_{gijt}+0)/2$	1.039	8.42	-35,488
1.0	1.052	15.63	-35,447
0.9	1.053	15.01	-35,452
0.7	1.051	13.04	-35,464
0.5	1.045	10.28	-35,479
0.3	1.035	7.35	-35,493
0.1	1.024	4.79	-35,503
0.0	1.024	4.79	-35,503

Notes: This table reports the ML estimates for the specification with 71 countries, 11 factors, and 5 years. γ_{gijt} is the share of trade that is intermediate inputs trade. The baseline choice of γ_{gijt} throughout the paper is γ^*_{gijt}. It is described in appendix C. t is the t-statistic for the hypothesis of constant returns to scale ($\mu = 1$). 'Likelihood' is the loglikelihood value.

Table A.3. GLS and Variance Parameter Estimates

	ω	std.err.	σ	std.err.
Table 2				
row 1	0.85	0.0061	0.68	0.025
row 2	0.84	0.0094	0.64	0.036
row 5	0.84	0.0078	0.67	0.031
row 8	0.82	0.0134	0.55	0.045
Table 4				
$N = 71$, $F = 11$, 1972-92	0.85	0.0061	0.69	0.025
$N = 71$, $F = 11$, 1987-92	0.84	0.0094	0.64	0.037
$N = 23$, $F = 11$, 1972-92	0.90	0.0138	0.92	0.056
$N = 23$, $F = 11$, 1987-92	0.91	0.0228	0.95	0.095

Notes: This table reports the estimated GLS parameter ω and the estimated variance parameter σ for the specifications reported in tables 2 and 4. ω and σ are described in appendix D. N and F are the number of countries and factors, respectively. 'std. err.' is the standard error of the estimate.