MONOPOLISTIC COMPETITION AND INTERNATIONAL TRADE: RECONSIDERING THE EVIDENCE*

DAVID HUMMELS AND JAMES LEVINSOHN

We test some propositions about international trade flows that are derived from models of monopolistic competition developed by Elhanan Helpman and Paul Krugman. We investigate whether the volume of trade between OECD countries is consistent with the predictions of a model in which all trade is intraindustry trade in differentiated products. We then repeat the test with non-OECD countries. We also investigate whether the share of intraindustry trade is consistent with a more general theoretical model in which some, but not all, trade is intraindustry trade. Our results lead us to question the apparent empirical success of these models.

I. INTRODUCTION

This paper is about testing a relatively new theory of international trade. The life cycle of trade theories seems to progress in the following three steps. In an effort to explain observed trade flows, theorists arrive at a rigorous and logically consistent theory. Next, the theory is subjected to an initial barrage of empirical tests. Frequently, these tests yield conflicting evidence, and seldom is the theory unconditionally accepted or rejected. The third step involves sorting out the puzzles and paradoxes to arrive at a set of facts that international economists consider part of their received wisdom. In the case of traditional endowments-based international trade, Heckscher and Ohlin [1991 translation] exemplify the first step, Leontief [1953] the second step, and Leamer [1984, 1987 with Bowen and Sveikauskas] the third.

If those are the three steps in the life cycle of international trade theories, this paper is part of the second stage. This time around the theory tested is that of monopolistic competition and international trade. In this paper we point out some puzzles and paradoxes. We do not provide many answers. At best, we pave the way for the third stage of the history’s life cycle. At worst, we leave matters confused and unsettled.

There is a long and distinguished literature examining the theory of international trade and monopolistic competition. The first papers were by Krugman [1979, 1981] and Lancaster [1980].

*We are grateful to James Brander, Ronald Cronovich, Alan Deardorff, Jonathan Eaton, Martin Feldstein, Gene Grossman, Elhanan Helpman, Anne Krueger, Paul Krugman, Edward Leamer, and the editors for helpful comments, suggestions, and skepticism.

1. For a recent survey of these tests, see Leamer and Levinsohn [1995].

© 1995 by the President and Fellows of Harvard College and the Massachusetts Institute of Technology.
The Quarterly Journal of Economics, August 1995
This work was further developed and expanded in Helpman [1981], and it is nicely summarized in Helpman and Krugman [1985]. This line of work was in part motivated by the observation that much international trade appears to be in goods that are quite similar. While traditional factor endowments-based explanations of international trade did not explain this empirically relevant component of international trade, Helpman and Krugman showed that a model of monopolistic competition could. There are many models of monopolistic competition and international trade, each with different sets of assumptions. In general, though, these are models in which firms produce differentiated products with an increasing-returns-to-scale technology, while on the consumption side, consumers have utility functions that reward product diversity.

There is also a lengthy literature examining the empirical side of this topic. These studies typically construct an index of intraindustry trade and investigate correlates of that index. While these studies are certainly interesting, their relationship to the theory of monopolistic competition and international trade is often tenuous. This is no surprise, since in many cases the empirical studies preceded Helpman and Krugman’s theoretical work. Nonetheless, the theory is about fifteen years old, and empirical tests closely linked to the theory remain scarce. An important exception to this is a paper by Helpman [1987] in which he developed some simple models of monopolistic competition and tested some hypotheses which were directly motivated by the theory.

Of the many papers that empirically investigate intraindustry trade, this one [Helpman 1987] is especially noteworthy. It constructs two very straightforward theoretical models (drawing heavily on Helpman and Krugman). These models are designed to yield empirically testable hypotheses. Taking the theory on its own terms, Helpman then asks whether the data are consistent with two predictions that come out of the theory. Helpman’s first hypothesis concerns the volume of trade in a model in which all trade is trade in differentiated products. He next asks whether the share of trade that is intraindustry trade is consistent with a model in which some trade is motivated by traditional factors-based

---

2. Recent work by Davis [1992] shows that if intraindustry trade is defined as trade in goods embodying similar factors, traditional (Ricardian) trade theory can indeed explain intraindustry trade.


4. Helpman actually tests three predictions, but for reasons discussed in Section III, only two are relevant to our data.
explanations, while the rest of trade is motivated by a model of monopolistic competition. Using graphical methods and some simple regressions, Helpman finds that the data appear to be consistent with the models tested.

In this paper we revisit Helpman’s tests and reconsider the evidence. We do not substantively amend Helpman’s theoretical models. Rather, we apply a combination of different data and different econometric methods and ask whether the data still support the theory’s specific predictions. In the course of our investigation, we successfully replicate Helpman’s results, pose several new puzzles, and, in the end, find less than overwhelming empirical support for the theory.

The remainder of the paper is organized as follows. In Section II a model in which all trade is motivated by monopolistic competition is presented. This model generates predictions about the volume of trade. Using Helpman’s data set comprised of OECD countries, we retest the model’s predictions. We also test the model using an alternative data set. In Section III a more general model in which some trade is intraindustry while the rest is traditional interindustry (Heckscher-Ohlin) trade is described. We then test the model’s predictions concerning the share of trade that is intraindustry. Section IV concludes by summarizing the puzzles generated by the two tests of the theory. Two appendices are also included. In the first we gather the derivations of the estimating equations, while the second describes our data in detail.

II. MONOPOLISTIC COMPETITION AND THE VOLUME OF TRADE

We begin with the simplest model. Here, all trade between countries is assumed to be intraindustry trade. Firms each produce a different variety of a differentiated product with an increasing returns to scale technology, and monopolistic competition prevails. An important and testable result generated by this theoretical setup is that relative country size determines the volume of trade between countries. This is in contrast to the traditional factor-endowments based explanations for trade in which “differences in relative country size . . . have no particular effect (on the volume of trade).”

5. A summary of some preliminary results using this sort of test is found in Hummels and Levinsohn [1993].
6. See Appendix 1 for a full description of this model.
7. From Helpman [1987], p. 64.
Helpman shows that if countries have identical and homothetic preferences and trade is balanced, then

\[ \frac{V^A}{GDPA} = e_A \left[ 1 - \sum_{j \in A} (e^j_A)^2 \right], \]

where

- \( V^A \) is the volume of trade between countries in group \( A \),
- \( GDPA \) is the GDP of the group of countries comprising group \( A \),
- \( e_A \) is the share of group \( A \)'s GDP in relation to world GDP, and
- \( e^j_A \) is the share of country \( j \)'s GDP in relation to group \( A \)'s GDP.\(^8\)

The right-hand side of (1) is a measure of size dispersion that increases as countries become more similar in size. This particular measure of size dispersion comes directly from Helpman’s theoretical model. Furthermore, theory dictates exactly how relative country size ought to matter. Put another way, (1) is a structural equation from a model of monopolistic competition and international trade; it is not a reduced-form equation. Helpman also amends (1) and shows how the equation is altered in the presence of trade imbalances. Helpman found that correcting for trade imbalances made virtually no difference to his empirical results. We also find this to be true. For expositional simplicity, we present only the model for the balanced trade case, although empirical estimates are for the unbalanced trade case.\(^9\)

Helpman noted that as countries become more similar in size, the volume of trade as a proportion of group GDP should increase. To investigate this hypothesis, he selected a subset of the OECD countries. This seems a judicious choice, for if any group of countries can support the predictions of a model in which all trade is intraindustry, the OECD countries are likely candidates. Using this group of countries, he computed the left-hand side of (1) (the volume of intra-OECD trade relative to OECD GDP) and the right-hand-side index for every year from 1956 to 1981. This yielded 26 points which he then graphed. The resulting graph showed a clear positive correlation between the ratio of intraingroup volume of trade to group GDP and the index of size dispersion.

---

8. The derivation of (1) is provided in Appendix 1.
9. See Appendix 1 for details.
MONOPOLISTIC COMPETITION AND TRADE

803

That is, as country size became more similar, intragroup trade volume rose, hence confirming the theory’s prediction.

It is important to note that Helpman (and Krugman) were not the first to suggest a relationship between the volume of trade and some combined measure of trading partners’ incomes, nor is their model of monopolistic competition the only way to generate equation (1). Indeed, (1) fits the general form of the “gravity equation” given by \( V T_{ij} = f(Y_i, Y_j, Z) \), where \( V T_{ij} \) is the volume of trade between countries \( i \) and \( j \), \( Y_i \) is a measure of income from country \( i \), and \( Z \) may include various measures of trade resistance. While there are many variants of gravity models of international trade, variables often included in \( Z \) are measures of distance, trade barriers, and even variables reflecting common language or culture. It is well-known that such an equation fits the data remarkable well, but for years this has been an empirical regularity in search of a theoretical foundation. Anderson [1979] shows how to derive such a result based on the properties of Cobb-Douglas expenditure systems when each good is produced in only one country. Following years of empirical work that included variables like distance and trade barriers, Anderson, and later Bergstrand [1985], showed why variables such as these might be theoretically justified.

The contribution of monopolistic competition models is not so much the observation that trade volumes are related to GDPs (we already knew that!), but rather in presenting a coherent theory of why product differentiation occurs. That is, the role of monopolistic competition in these models is to insure that all goods are produced in only one country.

Having each good produced in only one country is crucial to deriving (1), and monopolistic competition is one way of generating this outcome. There are other ways. As noted by Leamer [1993] and as reiterated in Leamer and Levinsohn [1995], one can derive (1) without intraindustry trade if each good is produced in one country and tastes are identical and homothetic. For example, this outcome is generated absent monopolistic competition if one adopts an Armington Assumption. Such an assumption posits that consumers view goods as differentiated by country of origin, and this demand side phenomenon will also give rise to the functional form employed in (1).

With this background in mind, are Helpman’s results surprising? On the one hand, they are surprising. The theory that generated the estimating equation seems quite restrictive: every
good is produced in only one country, and all countries have identical homothetic preferences. Whether this structure is the outcome of monopolistic competition or of an Armington Assumption does not alter our surprise. We do not view either of these sets of underlying assumptions as particularly plausible. Nonetheless, the data appear consistent with the theories. On the other hand, since the theory’s prediction closely mimics that of gravity equation’s (which we know work well empirically), the results are not surprising.

What can one hope to learn by further testing such a model if one already knows that it will perform well empirically? First, gravity equation estimation is almost invariably performed in a cross section, while Helpman’s study uses a time-series approach. Using variation over time allows him to better address deeper questions. For example, is the rising trade-to-income ratio observed in postwar data due to increasing size similarity? Second, further testing can help put the relevance of monopolistic competition into a more reasonable perspective. Many studies motivate the use of the gravity model by appealing to the underlying framework of monopolistic competition as if it were a foregone conclusion that monopolistic competition is the true source of their results. One of our purposes here is to see whether gravity results pertain even in cases where we think the monopolistic competition framework is inappropriate.

We revisit Helpman’s first test and apply more standard econometric methods. Helpman’s original graph of 26 points, while a prudent methodology given the small sample size, did not allow him to conduct standard hypothesis tests. The theory holds for country groups of any size. Rather than aggregating over the entire OECD sample, we opt to treat each country-pair in each year as an observation. This yields 91 country-pair observations for each of the 22 years for which we have OECD data (1962–1983). This gives a total of 2002 observations.

There are several reasons why, even if the underlying theoretical model is correct, the model might not fit the data exactly in every year for every country-pair. For example, border trade, seasonal trade, trade restrictions that vary across country-pairs, language, and cultural ties may encourage or discourage international trade. Each of these is basically an explanation of trade that is unique to pairs of countries, but orthogonal to GDP. (Of these, trade restrictions are the example for which this assumption is most questionable. We will return to this issue below.) Because
these factors are country-pair specific, they can be accurately modeled as a country-pair fixed effect. There are also idiosyncratic reasons why the model might not fit exactly even if the underlying theory is correct. Prominent among these is measurement error in the volume of trade. Indexing country-pairs by \( i \) and years by \( t \) and taking logs of (1), rearranging yields

\[
(2) \quad \ln (V_{it}) = \alpha_1 \ln \left[ GDP_{it} (1 - (e_{it}^1)^2 - (e_{it}^2)^2) \right] + \nu_i + \epsilon_{it},
\]

where

- \( e_{it}^1 \) is the first country’s share of country-pair \( i \)’s GDP,
- \( e_{it}^2 \) is the second country’s share of country-pair \( i \)’s GDP,
- \( \nu_i = \mu_i + \ln (e_i) \) is the country-pair fixed effect,\(^{10}\)
- and \( \epsilon_{it} \) is the idiosyncratic component of the disturbance term.

The term \( \mu_i \) is capturing the effect of the myriad influences on trade flows that are orthogonal to the included right-hand-side variable. For example, one might expect the \( \mu_i \) for the Japan-Austria pair to be quite small or negative; whereas the \( \mu_i \) for the Austria-Germany country-pair might be quite high. That is, for reasons that have nothing to do with country size, Austria and Germany trade a lot with one another while Austria and Japan do not. These country-pair fixed effects may be thought of as taking the place of the trade resistance variables that frequently entered the cross-section gravity model results.

Prior to estimating (2), we first plot the right-hand-side variable against the left-hand-side variable. This is our analog to Helpman’s graphical test of the hypothesis. The plot, using mean-differenced data to capture the fixed effects, is given in Figure I. This plot of over 2000 country-pair-years shows a clear positive correlation between a measure of trade volume and country size.

We next estimate (2).\(^{11}\) Our base-case estimates are for the fixed-effects estimator. The results are given in the first column of Table I. The results confirm the simple plot of the data as well as Helpman’s initial findings. With a \( t \)-statistic of 110.8, there is little doubt that the particular measure of country size dispersion dictated by the theory is quite important in explaining trade volumes. Indeed, inclusive of the fixed effects, about 98 percent of the variation in trade volume is explained by the model.

\(^{10}\) \( \epsilon_i \) is considered to be a constant because we assume, like Helpman, that group GDP, as a fraction of world GDP, is constant over time.

\(^{11}\) Details concerning the data used in this estimation are gathered in Appendix 2.
There are several reasons, though, why the fixed-effects estimate of (2) might be misspecified. For example, we are treating the $\mu_i$'s as fixed when in fact they may be random. The second column of Table I gives the estimates of (2) when a random-effects estimator is employed, and it makes no difference to the results.

**TABLE I**

**EQUATION (2) ESTIMATES**

OECD DATA (1962–1983)

<table>
<thead>
<tr>
<th></th>
<th>Fixed effects</th>
<th>Random effects</th>
<th>Fixed effects (instrumental variables)</th>
<th>Fixed effects (detrended data)</th>
<th>OLS (detrended data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_1$</td>
<td>1.405</td>
<td>1.403</td>
<td>1.268</td>
<td>1.094</td>
<td>1.170</td>
</tr>
<tr>
<td>$t$-statistic</td>
<td>110.8</td>
<td>110.8</td>
<td>47.7</td>
<td>33.7</td>
<td>44.8</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.865</td>
<td>.860</td>
<td>—</td>
<td>.373</td>
<td>.501</td>
</tr>
<tr>
<td>(w/dummies)</td>
<td>.981</td>
<td></td>
<td></td>
<td>.980</td>
<td></td>
</tr>
</tbody>
</table>
Another possible problem is that the disturbance term $\epsilon_{it}$ may be correlated with the included regressor. That is, if exports receive a positive shock, trade volume rises, but by an accounting identity, so does GDP. Since a function of GDP appears as a regressor, we have an endogeneity problem. The standard solution to this is to employ an instrumental variables estimator. In this particular case, economic theory suggests some appropriate instruments. Following the strategy used by Harrigan [1992], we use countries' factor endowments as instruments. These are likely to be correlated with GDP and are orthogonal to idiosyncratic trade shocks. The results with the fixed effects instrumental variables estimator are given in the third column of Table I. We find that factor endowments are excellent instruments, for variation in factor endowments explains a very large share of the variation in GDP. While careful accounting suggests that the measure of size dispersion may be correlated with the idiosyncratic disturbance term, correcting for this makes very little difference. We noted above that if country-pair-specific trade barriers are an important component of the fixed effect, the fixed effect may be correlated with GDP. (This would be the case if high trade barriers lead to low GDP.) In this case, the fixed-effects estimates remain consistent, but inefficient, while the random-effects estimates are biased. Hence, the true standard error may be smaller than the reported standard error for the fixed-effects estimator, but since estimates are still very precisely estimated, this is not a cause for much concern.

Another potential explanation for the remarkable fit of (2) is that the volume of trade and group GDP may be trending upward over the period spanned by the sample. This might be the case, for example, as trade barriers fell in the European Community. We investigate how robust are our estimates to this concern by estimating (2) using (deterministically) detrended data. The results are given in the fourth column of Table I. Even after sweeping out trends and all country-pair fixed effects, the results are still strong, as the coefficient on the measure of size dispersion is still quite precisely measured. In the final column of Table I, we report the estimates that result from simple OLS on the detrended

12. The factors used as instruments are labor, divided into two skill levels, and capital stock.
13. The data were detrended with a linear trend common to all countries.
data. The message of this table is that even controlling for trends and country-pair fixed effects, our regressions strongly support Helpman’s original finding.14

We began with a simple model in which all trade is trade in differentiated products, and everyone has identical and homothetic tastes. This model implied a very specific estimating equation in which a very particular index of size dispersion was predicted to explain trade volume. And it all worked! Is the world really so simple?

To address this question, the model is reestimated using a data set that we believe, ex ante, is grossly inappropriate for a model of monopolistic competition and international trade. Instead of using the OECD countries, we create a data set comprised of Brazil, Cameroon, Colombia, Congo, Greece, Ivory Coast, South Korea, Nigeria, Norway, Pakistan, Paraguay, Peru, Philippines, and Thailand. This group of countries is referred to as the NOECD countries.

Equation (2) is reestimated using slightly different data definitions. The sample stops in 1977 instead of 1983 because several countries stopped reporting trade data in 1978. Also, we estimate the equation in levels instead of logs. This is because 142 of the 1456 observations of the dependent variable are zero, while others are close to zero. Column one of Table II reports the fixed effects estimates. Even for this sample of countries, the particular measure of size dispersion suggested by the theory matters, and it matters in a precisely measured way. The t-statistic drops to 24, but by any conventional standard, this is remarkably significant. Furthermore, the result is robust. When (2) is estimated with NOECD data using random effects, fixed-effects instrumental variables, and detrended data (fixed effects and OLS), the results

14. We also estimated the equations reported in Tables I and II using slightly different specifications in order to test the robustness of the relationship between the volume of trade and a measure of size dispersion. We estimated two alternative specifications based on (1). These are as follows. (i) Estimating \( \ln(V_{it}) = \alpha_1 \ln(Total \ GDP_{it}) + \alpha_2 Size_{it} \) to investigate whether it is the total GDP variable or the size dispersion index that is driving the correlation. (ii) Estimating \( \ln(V_{it}/Total \ GDP_{it}) = \alpha_1 \ln(Size_{it}) \) for the OECD data set, including total GDP separately, as in (i), yields \( \ln(V_{it}/Total \ GDP_{it}) = 1.290 \ln(Total \ GDP_{it}) + 0.976 \ln(Size_{it}) \), with a t-statistic on the size variable of 28 and of 50 on the GDP variable. Estimating (ii) gave a coefficient of 3.13 with a t-statistic of 25. Estimates with the NOECD data set are quite similar. We interpret all this to mean that, while total GDP is important (rather sensibly, the trade volume between two big countries is greater than the trade volume between two small countries), the size dispersion index is also very important in these regressions.
do not vary much. While using the NOECD data set does not explain as much of the variation in the volume of trade as was the case with OECD data (the $R^2$ falls from .98 to .67 in one specification), the results of Table II still provide strong support for the theory. Put another way, if Table II had been presented prior to Table I, most would agree that the model fit the data well.

We set out to see whether a monopolistic condition model's predictions about the volume of trade provided insights into the theoretical source of the gravity equation. We find that while the model explains trade volumes well, it may actually explain them too well. The predictions hold for a set of countries that we feel fit the assumptions of the monopolistic competition model—goods are differentiated, demands are identical and homothetic—reasonably well. But the predictions also hold for a set of countries that we feel are not appropriately characterized by differentiated goods trade or by identical homothetic demands. The estimated equation works extremely well in both cases, and this causes us to question whether monopolistic competition is the right theoretical justification for it.

This leaves several possible explanations unaddressed. First, it is possible that our ex ante beliefs about the nature of production in the NOECD country set are wrong. That is, perhaps it really is appropriate to think of goods from these countries as also being differentiated. Or perhaps the Armington Assumption really is

15. If the NOECD data set is applied to (2) in logs rather than levels, the magnitudes of the coefficients are similar to those reported for the OECD data set. In particular, $\ln V_{it} = 2.08 * SIZE$ with a $t$-statistic of 20.8 and an $R^2$ of .792.
correct. Second, perhaps our understanding of the role of country size in a traditional Heckscher-Ohlin model is incomplete. It remains an intriguing topic for future research to determine whether one can generate gravity-like results without the assumption of complete specialization in production.

III. A More General Approach

In Section II we used a model of monopolistic competition to show how the presence of intraindustry trade results in a specific and testable, but not unique, hypothesis about the bilateral volume of trade. One of the underlying assumptions of that section was that all trade was intraindustry. In this section we relax that assumption and assume that some trade is intraindustry and some interindustruy. We then examine how the fraction of trade that is intraindustry varies between countries and over time.16

Theoretical research into the causes of intraindustry trade can be divided into "small numbers" and "large numbers" explanations, with the label referring to the number of firms. Small numbers models involve intraindustry trade in oligopolistic industries. These models come in many flavors as assumptions concerning homogeneity of product, the firms' strategic variable, and entry conditions vary. It is well-known that the results derived in these small numbers models are often not robust to these varying assumptions. These models, therefore, are of limited use in constructing general country characteristic hypotheses about intraindustry trade.

Large numbers explanations model free entry by firms into increasing returns to scale industries. We turn again to the model we find most convincing, Helpman [1987], where he shows that the bilateral share of intraindustry trade increases as two countries become more similar in factor composition.17

The intuition for this is as follows. In a model with homogeneous and differentiated goods, some interindustruy trade will be

16. Our study will focus on country characteristic explanations for intraindustry trade (IIT), that is, how differences across countries explain IIT. There is another, extensive literature on how intraindustry trade varies across industries within countries. The model of monopolistic competition in this paper assumes two types of industries, homogeneous and differentiated goods. Within each type, industries are identical so it makes little sense to test intraindustry trade variation across them.

17. See Appendix 1 for a formal statement of the $2 \times 2 \times 2$ case. It is important to note that this result may not hold in a model with many countries, many goods, or many factors.
motivated by relative factor abundance, and some intraindustry trade will be motivated by the exchange of varieties of differentiated goods. A standard measure of intraindustry trade is the Grubel-Lloyd [1975] index. The share of intraindustry trade between countries $j$ and $k$ in industry $i$ is given by

$$IIT_{ijk} = \frac{2 \min (X_{ijk}, X_{ikj})}{(X_{ijk} + X_{ikj})},$$

where $X_{ijk}$ are exports of industry $i$ from country $j$ to country $k$. The share of intraindustry trade between country $j$ and $k$, over all industries, is given by

$$IIT_{jk} = \frac{2 \sum_i \min (X_{ijk}, X_{ikj})}{\sum_i (X_{ijk} + X_{ikj})}.$$

The numerator captures two-way trade within industries, and the denominator is the total volume of trade. More transparently, we can think of this index as

$$IIT_{jk} = \frac{INTRA}{INTRA + INTER}.$$

In a two-country, two-factor model with one homogeneous goods sector and one differentiated goods sector, allow both countries to have identical capital-to-labor ratios. Then no trade is motivated by relative factor abundance. That is, $INTER = 0$, and the intraindustry trade index ($IIT_{jk}$) equals one. Now, perturb the capital-to-labor ratios, holding relative size constant. $INTER$ increases because there is now a reason for trade motivated by factor differences. $INTRA$ will decrease, and the above index will decrease as well.

We were careful to note that the reallocation of capital and labor must occur holding relative size constant. We know from Section II that relative size can have an important effect on the volume of trade in differentiated products. A reallocation of capital and labor that widened factor differences and also changed relative size (for example, making the two countries more equal), may actually increase intraindustry trade.

Finally, note that this relationship between the similarity of capital-labor ratios and intraindustry trade has as much, and perhaps more, to do with traditional explanations for trade as it does with monopolistic competition models. Put another way, if we

18. See Appendix 1.
are to find empirical evidence of the hypothesized relationship between intraindustry trade and factor differences, it must be that trading patterns are sensitive to factor differences in a way suggested by the Heckscher-Ohlin model.

To test the relationship between factor differences and the share of intraindustry trade, Helpman estimated equation (3) on a cross section of 91 country-pairs, using separate regressions for each year from 1970–1981:

\[
IIT_{jk} = \alpha_0 + \alpha_1 \log \left| \frac{GDP^j}{N^j} - \frac{GDP^k}{N^k} \right| \\
+ \alpha_2 \min (\log GDP^j, \log GDP^k) \\
+ \alpha_3 \max (\log GDP^j, \log GDP^k) + \epsilon_{jk},
\]

where \(IIT_{jk}\) is the Grubel-Lloyd [1975] index for the bilateral trade of a country-pair consisting of countries \(j\) and \(k\), \(N^j\) is the population of country \(j\), and an industry is defined as a four-digit SITC group. Per capita GDP is used to proxy factor composition. \(MINGDP\) and \(MAXGDP\) are included to control for relative size effects. The model predicts that \(\alpha_1 < 0, \alpha_2 > 0,\) and \(\alpha_3 < 0.\)

Helpman found that the data supported these predictions. In particular, he found a negative and significant correlation between factor differences and the \(IIT_{jk}\) index, although it weakened toward the end of his sample.\(^{19}\) There are, however, two potential problems with his approach. One, Helpman uses per capita income as a proxy for factor composition. Two, he does not exploit the panel nature of his data.

Two problems are posed by the use of per capita income as a proxy for factor composition. First, it is an appropriate proxy if there are only two factors of production and all goods are traded. As this is probably not the case, we would like to know to what degree a better measure of factor composition might alter the results.

Second, this approach runs afoul of a long-standing debate on whether per capita income is proxying factor endowments or consumer tastes. Linder [1961] hypothesized that manufactured products must first be developed for home markets before they can be exported successfully. Countries with similar demand structures would develop similar goods for home use and later export. If

\(^{19}\) Specifically, the coefficient on his factor differences variable is negative and significant in the first seven years, but becomes insignificant thereafter. Also, the \(R^2\) in the regression drops steadily from .266 in 1970 to .039 by 1981.
per capita income is a good gauge of demand, then two countries with similar per capita income will have similar demand, and will produce and export similar goods. Krugman [1980] and Bergstrand [1990] have subsequently demonstrated the importance of taste differences in more rigorous models of monopolistic competition with nonhomothetic demand. The empirical literature has generally interpreted differences in per capita income as a demand side phenomenon, and found good support for a negative relationship between per capita income and intraindustry trade. This leads to some confusion as to whether the difference in per capita income is proxying differences in factor composition, as posited by Helpman, or demand structure, as posited by Linder. To address these potential problems with the proxy variable, we alternately employ per capita income and actual factor data to measure differences in factor composition.

We begin by estimating equations similar to (3) for our OECD sample separately for each year from 1962 to 1983. In the first estimation we use per worker GDP. In the second, we use actual capital-to-labor and land-to-labor ratios. The estimating equations, then, are given by

\[ IIT_{jk} = \alpha_0 + \alpha_1 \log \left| \frac{GDP^j}{L^j} - \frac{GDP^k}{L^k} \right| \]

(4) 
\[ + \alpha_2 \min (\log GDP^j, \log GDP^k) + \alpha_3 \max (\log GDP^j, \log GDP^k) + \epsilon_{jk}, \]

\[ IIT_{jk} = \alpha_0 + \alpha_1 \log \left| \frac{K^j}{L^j} - \frac{K^k}{L^k} \right| + \alpha_2 \log \left| \frac{T^j}{L^j} - \frac{T^k}{L^k} \right| \]

(5) 
\[ + \alpha_3 \min (\log GDP^j, \log GDP^k) + \alpha_4 \max (\log GDP^j, \log GDP^k) + \epsilon_{jk}, \]

where \(L^j\) is the working population of country \(j\), \(T^j\) is \(j\)'s land endowment, and \(K^j\) is \(j\)'s capital stock. We label log \(|(GDP^j)|

21. Unlike the test in Section II, we do not replicate this test using NOECD data. This is because the NOECD set, by construction, contains virtually no intraindustry trade and would therefore be of little use in studying cross-country variation in an IIT index.
22. We use per worker GDP instead of per capita GDP, since the former seems more consistent with the underlying theory.
Li) - (GDPk/Lk)], which gives differences in income per worker, YDIF. Analogously, KLDIF will refer to the differences in capital per worker, and TLDIF will refer to differences in land per worker (as in (5)).

GDP, K (constructed capital stock), and L (labor force) come directly from, or are constructed from, Penn World Tables, Mark V data. GDP and K are measured in constant 1985 international prices. T (land) is constructed from land estimates contained in Leamer [1984]. See Appendix 2 for details.

Equations (4) and (5) are estimated with ordinary least squares (OLS). IITjk is an index varying between zero and one. We apply a logistic transformation\textsuperscript{23} to IIT so that OLS using the transformed variable is appropriate. The results are reported in Tables III and IV.

Table III reports the results of estimating (4). The results are quite similar to Helpman's.\textsuperscript{24} The coefficient on YDIF is negative in each sample year, but is only significant through roughly half of the sample. The coefficients on MINGDP and MAXGDP are consistent with theory, but only MINGDP is significant. Finally, like Helpman, the explanatory power of the regression drops steadily over time.

Just as in Helpman's study, the relationship between the share of intraindustry trade and differences in factor composition is strongly negative in early years of the sample, but breaks down in later years. Having replicated Helpman's results, we turn to the estimation of equation (5), where per worker income as a proxy for factor composition is replaced with actual factor data.

In Table IV we see that TLDIF is negative and highly significant in all sample years. KLDIF is negative and significant initially, but in later years becomes positive and significant. Also, the explanatory power of the estimates in Table IV is roughly double that of the estimates in Table III. These results would seem to support the hypothesis that factor differences, especially the difference in land-to-labor ratios, are important in explaining the share of intraindustry trade. This may be because there is very little IIT in agricultural products. Countries with much arable land

\textsuperscript{23} The transform is given by \( IIT = \ln[IIT/(1 - IIT)] \).

\textsuperscript{24} The variables employed here differ in three ways from Helpman's. First, we use per worker income rather than per capita income. Second, YDIF is measured in constant 1985 dollars. Helpman's study employed per capita income measured in current dollars. When we use a current dollar measure, we obtain regression results very similar to Helpman's, but that differ slightly from constant dollar measures. Third, we apply the logistic transform to IIT.
TABLE III
EQUATION (3) ESTIMATES WITH GDP PER WORKER INSTEAD OF GDP PER CAPITA
(1962–1983)

<table>
<thead>
<tr>
<th>Year</th>
<th>YDIF</th>
<th>MINGDP</th>
<th>MAXGDP</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1962</td>
<td>-0.275**</td>
<td>0.389**</td>
<td>-0.012</td>
<td>.209</td>
</tr>
<tr>
<td>1963</td>
<td>-0.316**</td>
<td>0.377**</td>
<td>-0.009</td>
<td>.239</td>
</tr>
<tr>
<td>1964</td>
<td>-0.227**</td>
<td>0.371**</td>
<td>-0.013</td>
<td>.216</td>
</tr>
<tr>
<td>1965</td>
<td>-0.255**</td>
<td>0.366**</td>
<td>0.022</td>
<td>.239</td>
</tr>
<tr>
<td>1966</td>
<td>-0.290**</td>
<td>0.334**</td>
<td>0.011</td>
<td>.242</td>
</tr>
<tr>
<td>1967</td>
<td>-0.217**</td>
<td>0.349**</td>
<td>0.023</td>
<td>.230</td>
</tr>
<tr>
<td>1968</td>
<td>-0.234**</td>
<td>0.309**</td>
<td>0.007</td>
<td>.213</td>
</tr>
<tr>
<td>1969</td>
<td>-0.206**</td>
<td>0.332**</td>
<td>-0.009</td>
<td>.214</td>
</tr>
<tr>
<td>1970</td>
<td>-0.156*</td>
<td>0.355**</td>
<td>-0.062</td>
<td>.171</td>
</tr>
<tr>
<td>1971</td>
<td>-0.196*</td>
<td>0.335**</td>
<td>-0.085</td>
<td>.176</td>
</tr>
<tr>
<td>1972</td>
<td>-0.222**</td>
<td>0.336**</td>
<td>-0.040</td>
<td>.188</td>
</tr>
<tr>
<td>1973</td>
<td>-0.233*</td>
<td>0.253**</td>
<td>-0.031</td>
<td>.147</td>
</tr>
<tr>
<td>1974</td>
<td>-0.109</td>
<td>0.235**</td>
<td>-0.058</td>
<td>.074</td>
</tr>
<tr>
<td>1975</td>
<td>-0.189</td>
<td>0.226*</td>
<td>-0.045</td>
<td>.089</td>
</tr>
<tr>
<td>1976</td>
<td>-0.159</td>
<td>0.181*</td>
<td>-0.050</td>
<td>.056</td>
</tr>
<tr>
<td>1977</td>
<td>-0.260**</td>
<td>0.152</td>
<td>-0.043</td>
<td>.095</td>
</tr>
<tr>
<td>1978</td>
<td>-0.214*</td>
<td>0.111</td>
<td>0.004</td>
<td>.067</td>
</tr>
<tr>
<td>1979</td>
<td>-0.126</td>
<td>0.130</td>
<td>-0.057</td>
<td>.028</td>
</tr>
<tr>
<td>1980</td>
<td>-0.074</td>
<td>0.127</td>
<td>-0.090</td>
<td>.008</td>
</tr>
<tr>
<td>1981</td>
<td>-0.078</td>
<td>0.116</td>
<td>-0.106</td>
<td>.011</td>
</tr>
<tr>
<td>1982</td>
<td>-0.064</td>
<td>0.103</td>
<td>-0.105</td>
<td>.011</td>
</tr>
<tr>
<td>1983</td>
<td>-0.091</td>
<td>0.054</td>
<td>-0.064</td>
<td>.003</td>
</tr>
</tbody>
</table>

The estimated regression is

\[ IIT_{jk} = \alpha_0 + \alpha_1 \log \left( \frac{GDPI_j - GDPI_k}{L_j - L_k} \right) + \alpha_2 \min (\log GDPI_j, \log GDPI_k) + \alpha_3 \max (\log GDPI_j, \log GDPI_k) + \epsilon_{jk} \]

*indicates statistical significance at the 95 percent level.
**indicates statistical significance at the 99 percent level.

tend to exchange agricultural products for manufactures and therefore engage in less IIT. Using actual factor data instead of the proxy yields results more consistent with the theory’s predictions.25

However, there are several ways in which the three factors, land, labor, and capital, can be combined into factor difference measures. The theory underlying equation (5) gives little guidance as to which of these factor difference measures to include, and an

25. In earlier versions of this paper we included only two factor variables, capital and labor, and found that KLDIF has little effect. We have included land in the list of factors at the suggestion of referees, and note that its inclusion seems quite important.
TABLE IV


<table>
<thead>
<tr>
<th>Year</th>
<th>KLDIF</th>
<th>TLDIF</th>
<th>MINGDP</th>
<th>MAXGDP</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1962</td>
<td>-0.289**</td>
<td>-0.184**</td>
<td>0.392**</td>
<td>-0.058</td>
<td>.357</td>
</tr>
<tr>
<td>1963</td>
<td>-0.261**</td>
<td>-0.190**</td>
<td>0.395**</td>
<td>-0.074</td>
<td>.390</td>
</tr>
<tr>
<td>1964</td>
<td>-0.201*</td>
<td>-0.201**</td>
<td>0.359**</td>
<td>-0.042</td>
<td>.392</td>
</tr>
<tr>
<td>1965</td>
<td>-0.139*</td>
<td>-0.203**</td>
<td>0.350**</td>
<td>-0.010</td>
<td>.385</td>
</tr>
<tr>
<td>1966</td>
<td>-0.196*</td>
<td>-0.209**</td>
<td>0.319**</td>
<td>-0.041</td>
<td>.416</td>
</tr>
<tr>
<td>1967</td>
<td>-0.183*</td>
<td>-0.208**</td>
<td>0.324**</td>
<td>-0.011</td>
<td>.426</td>
</tr>
<tr>
<td>1968</td>
<td>-0.157*</td>
<td>-0.223**</td>
<td>0.291**</td>
<td>-0.049</td>
<td>.442</td>
</tr>
<tr>
<td>1969</td>
<td>-0.130</td>
<td>-0.217**</td>
<td>0.318**</td>
<td>-0.048</td>
<td>.435</td>
</tr>
<tr>
<td>1970</td>
<td>-0.112</td>
<td>-0.231**</td>
<td>0.316**</td>
<td>-0.073</td>
<td>.442</td>
</tr>
<tr>
<td>1971</td>
<td>-0.071</td>
<td>-0.231**</td>
<td>0.293**</td>
<td>-0.093</td>
<td>.421</td>
</tr>
<tr>
<td>1972</td>
<td>-0.051</td>
<td>-0.224**</td>
<td>0.300**</td>
<td>-0.056</td>
<td>.393</td>
</tr>
<tr>
<td>1973</td>
<td>-0.033</td>
<td>-0.209**</td>
<td>0.238**</td>
<td>-0.058</td>
<td>.354</td>
</tr>
<tr>
<td>1974</td>
<td>-0.015</td>
<td>-0.228**</td>
<td>0.194**</td>
<td>-0.045</td>
<td>.395</td>
</tr>
<tr>
<td>1975</td>
<td>-0.016</td>
<td>-0.208**</td>
<td>0.204**</td>
<td>-0.047</td>
<td>.313</td>
</tr>
<tr>
<td>1976</td>
<td>0.048</td>
<td>-0.202**</td>
<td>0.169*</td>
<td>-0.051</td>
<td>.277</td>
</tr>
<tr>
<td>1977</td>
<td>0.037</td>
<td>-0.207**</td>
<td>0.160</td>
<td>-0.066</td>
<td>.260</td>
</tr>
<tr>
<td>1978</td>
<td>0.052</td>
<td>-0.190**</td>
<td>0.116</td>
<td>-0.016</td>
<td>.225</td>
</tr>
<tr>
<td>1979</td>
<td>0.069</td>
<td>-0.172**</td>
<td>0.130</td>
<td>-0.056</td>
<td>.203</td>
</tr>
<tr>
<td>1980</td>
<td>0.095</td>
<td>-0.170**</td>
<td>0.126</td>
<td>-0.073</td>
<td>.197</td>
</tr>
<tr>
<td>1981</td>
<td>0.123*</td>
<td>-0.168**</td>
<td>0.120</td>
<td>-0.089</td>
<td>.209</td>
</tr>
<tr>
<td>1982</td>
<td>0.166*</td>
<td>-0.148**</td>
<td>0.122</td>
<td>-0.091</td>
<td>.216</td>
</tr>
<tr>
<td>1983</td>
<td>0.221**</td>
<td>-0.159**</td>
<td>0.078</td>
<td>-0.049</td>
<td>.237</td>
</tr>
</tbody>
</table>

The estimated regression is

$$IIT_{jk} = \alpha_0 + \alpha_1 \log \left| \frac{K_j}{L_j} - \frac{K_k}{L_k} \right| + \alpha_2 \log \left| \frac{T_j}{L_j} - \frac{T_k}{L_k} \right| + \alpha_3 \min \left( \log GDP'_j, \log GDP'_k \right) + \alpha_4 \max \left( \log GDP'_j, \log GDP'_k \right) + \epsilon_{jk}.$$  

*indicates statistical significance at the 95 percent level.

**indicates statistical significance at the 99 percent level.

exhaustive listing of results with all possible combinations is not possible in this space. We generally find that factor differences which account for land tend to be negatively correlated with $IIT$.

We noted above that we saw two ways in which one might improve upon Helpman’s approach. The first was to use actual factor data, rather than a proxy. This change in specification changed the results in important ways. The second potential improvement is to take advantage of the panel nature of the data.

By estimating equations (4) and (5) year by year, we ignore the possibility that the reason the model does not fit exactly may be
correlated over time for a given country-pair. That is, for reasons outside of the model and resulting specification, intraindustry trade between Japan and the United Kingdom might always be quite low relative to the sample as a whole. Here, the theory provides some guidance. The comparative statics exercise in question takes two countries, and, holding other things constant, perturbs their relative factor endowments. The natural experiment this suggests is to examine the relationship between intraindustry trade and factor differences as they change over time for a given country-pair. By looking only at cross-sectional variation, the "holding other things constant" assumption is far less tenable. This approach may be especially important if much of observed intraindustry trade is due to idiosyncratic differences between country-pairs that do not change much over time. Examples of such time-stationary idiosyncratic differences might include geography, seasonal trade, cultural and language ties, and trade barriers.\textsuperscript{26} For example, in the cross section we try to ascribe the variability in intraindustry trade between Germany-Austria and intraindustry trade between Japan–United Kingdom to differences in their relative factor endowments. If Germany and Austria are more similarly endowed than are Japan and the United Kingdom, we expect them to have more intraindustry trade. However, it may be that the "similar factor" effect is swamped by the fact that Germany and Austria are next door to one another while Japan and the United Kingdom are thousands of miles apart, or that Germany and Austria belong to a customs union.

Unfortunately, this approach poses a problem: factor difference measures that include land show very little variation over time for a given country-pair.\textsuperscript{27} As a result, even though the land measures were quite important in explaining cross-sectional variation in intraindustry trade, we cannot use them directly to explain its time series variation. Put another way, land endowments themselves are perfectly collinear with, and factor difference measures like TLDIF are highly collinear with, country-pair dummies. In spite of this, we feel that there is much to be learned by taking advantage of the panel nature of the data. We return to the role of land endowments later in this section.

\textsuperscript{26} Previous cross-sectional studies (see Loertscher and Wolter [1980]) have tried to capture these effects with dummy variables, and consistently found them to be significant.

\textsuperscript{27} That is, even though capital endowments may be changing over time, our data show that differences in capital-to-land ratios are highly stable.
To examine the relationship between intraindustry trade and factor differences over time, we pool our 22 years into a single panel. We first estimate a panel data version of (5) in order to pick up both cross-sectional and time series variation in $IIT_{jk,t}$. The estimating equation becomes

$$IIT_{jk,t} = \alpha_0 + \alpha_1 \log \left| \frac{K^j_t}{L^j_t} - \frac{K^k_t}{L^k_t} \right|$$

$$+ \alpha_2 \min (\log GDP^j_t, \log GDP^k_t)$$

$$+ \alpha_3 \max (\log GDP^j_t, \log GDP^k_t) + \epsilon_{jk,t}.$$  

where $jk$ indexes a country-pair as before and $t$ now indexes time.

We also estimate a variant of (6) which includes a vector of country-pair-specific fixed effects, $v_{jk}$, thereby sweeping out all of the cross-sectional variation. Hence we have

$$IIT_{jk,t} = \alpha_1 \log \left| \frac{K^j_t}{L^j_t} - \frac{K^k_t}{L^k_t} \right|$$

$$+ \alpha_2 \min (\log GDP^j_t, \log GDP^k_t)$$

$$+ \alpha_3 \max (\log GDP^j_t, \log GDP^k_t) + v_{jk} + \epsilon_{jk,t}.$$  

Finally, we estimate the equations using per worker income differences instead of capital per worker differences. (See the equations in Table V.) The OLS results using either income per worker or capital per worker as a regressor are reported in the first two columns of Table V. The results differ considerably depending on which regressor is included. The income per worker variable is negative and highly significant, while the capital per worker variable is not significantly different from zero. These results are generally consistent with those reported in Tables III and IV.\(^{28}\) For both OLS regressions the coefficient on $MINGDP$ is consistent with theory and precisely estimated, while the coefficient on $MAXGDP$ is not precisely estimated in the regressions using income differences.

Fixed-effects estimators are presented in the third and fourth columns of Table V. Recall that these estimates sweep out all country-pair specific effects. The coefficient on the income differences variable, $YDIF$, is now positive and quite significant, whereas

\(^{28}\) That is, the coefficient on $YDIF$ is negative and significant throughout, while the coefficient on $KLDIF$ goes from negative to positive in later years, resulting in an ambiguous effect over the whole sample.
before it was negative and very significant. In the regression using capital per worker, the factor difference variable, $KLDIF$, is both positive and significant. For both regressions, $MINGDP$ and $MAXGDP$ are as before, and the explanatory power of the regressions increases substantially. It is also interesting to note that when country dummies are employed in the regressions, the $R^2$ jumps to around .96.

We speculated above that there may be reasons why the model does not fit exactly that are correlated over time for a given country-pair. Further, we noted that this could be especially important if much of the variability in intratrade was explained by idiosyncratic differences between country-pairs. The fixed-effect regression results appear to bear this out. Country-pair dummies seem to explain a tremendous proportion of the variation in our intratrade index. Further, when country-pair effects are swept out, the coefficient on one measure of factor differences goes from being insignificantly different from zero to being significantly positive, while the coefficient on the other before it was negative and very significant. In the regression using capital per worker, the factor difference variable, $KLDIF$, is both positive and significant. For both regressions, $MINGDP$ and $MAXGDP$ are as before, and the explanatory power of the regressions increases substantially. It is also interesting to note that when country dummies are employed in the regressions, the $R^2$ jumps to around .96.

We speculated above that there may be reasons why the model does not fit exactly that are correlated over time for a given country-pair. Further, we noted that this could be especially important if much of the variability in intratrade was explained by idiosyncratic differences between country-pairs. The fixed-effect regression results appear to bear this out. Country-pair dummies seem to explain a tremendous proportion of the variation in our intratrade index. Further, when country-pair effects are swept out, the coefficient on one measure of factor differences goes from being insignificantly different from zero to being significantly positive, while the coefficient on the other

29. One interpretation of this result is simply that the fixed effects are picking up differences in land endowments. We return to this issue below.
measure goes from being a precisely measured negative estimate to a quite significant positive estimate. Hence, accounting for fixed effects yields precisely estimated results exactly counter to those implied by the theory.

Fixed-effects estimation treats the $u_{tk}$'s as fixed constants over time. If they are random variables instead, a random-effects estimator is appropriate. The results for the random-effects estimates are reported in the final columns of Table V. The random-effects estimator can be thought of as lying between the within and between estimators, and hence makes use of variation both between country-pairs and within country-pairs over time. The random-effects regression results are similar to the fixed-effects results. Coefficients on the factor differences variables are (still) positive and significant in both regressions, although the explanatory power of the regressions drops a small amount. The basic message of the fixed-effects estimates—that country-pair effects drastically change the empirical role of factor differences—also comes through clearly with random-effects estimates.

Because sensitivity analyses are important, we investigate how robust our results are to reasonable alternative specifications. Whereas the test described in Section II revolved around a structural equation, this test employed a reduced-form regression. That is, the theory does not dictate the appropriate specification. It only informs one of the variables that ought to enter the specification. While we have followed Helpman in estimating Table III—V using a semi-log specification, there is no theoretical justification for this particular specification; hence we experiment. We begin by estimating (6) and (7) in levels, and these results are reported in the first two columns of Table VI. Estimating in levels does not appear to change the punch line. $MINGD$ and $MAXGD$ are included largely as size effect controls. Since we do not know how they covary with the factor differences variable, we want to see how the coefficients on $KLDIF$ and $YDIF$ change when $MIN/MAXGD$ are omitted. Dropping $MINGD$ and $MAXGD$ reduces the significance of the factor differences variable, and the $R^2$ drops to about zero. Hence, while $MINGD$ and $MAXGD$ may be of secondary importance in the underlying theory, they are very important in the empirical work.

It may be the case that cross-sectional estimates which impose a linear relationship between $KLDIF$ and $IIT$ fit less well in later years because the relationship is, in fact, nonlinear. To begin to investigate this, we include a quadratic term for $KLDIF$ and for
MONOPOLISTIC COMPETITION AND TRADE

TABLE VI
SENSITIVITY ANALYSIS OF EQUATION (7) ESTIMATES (1962–1983)

Same specification as equation (7) except:

<table>
<thead>
<tr>
<th>Variable</th>
<th>In levels, not logs</th>
<th>Drop MIN/MAXGDP</th>
<th>Add DIF$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>KLDIF</td>
<td>3.33E-06</td>
<td>-0.033</td>
<td>-0.361</td>
</tr>
<tr>
<td></td>
<td>(1.080)</td>
<td>(-1.849)</td>
<td>(-2.422)</td>
</tr>
<tr>
<td>MINGDP</td>
<td>1.11E-09</td>
<td>0.314</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(8.926)</td>
<td>(15.386)</td>
<td>(2.391)</td>
</tr>
<tr>
<td>MAXGDP</td>
<td>-1.50E-11</td>
<td>-0.048</td>
<td>-0.083</td>
</tr>
<tr>
<td></td>
<td>(-0.643)</td>
<td>(-2.532)</td>
<td>(-8.928)</td>
</tr>
<tr>
<td>KLDIF$^2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R$^2$</td>
<td>.064</td>
<td>.001</td>
<td>.117</td>
</tr>
<tr>
<td>YDIF</td>
<td>-8.62E-05</td>
<td>-0.211</td>
<td>1.032</td>
</tr>
<tr>
<td></td>
<td>(-17.142)</td>
<td>(-11.42)</td>
<td>(7.456)</td>
</tr>
<tr>
<td>MINGDP</td>
<td>8.91E-10</td>
<td>0.261</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>(7.620)</td>
<td>(13.246)</td>
<td>(-8.928)</td>
</tr>
<tr>
<td>MAXGDP</td>
<td>1.52E-10</td>
<td>0.027</td>
<td>0.196</td>
</tr>
<tr>
<td></td>
<td>(6.370)</td>
<td>(1.434)</td>
<td>(.196)</td>
</tr>
<tr>
<td>YDIF$^2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R$^2$</td>
<td>.164</td>
<td>.061</td>
<td>.523</td>
</tr>
</tbody>
</table>

$t$-statistics are in parentheses. The reported $R^2$ in the fixed effects models is that for the regression using mean-differenced data.

YDIF in equations (6) and (7). For KLDIF, we find that the OLS coefficient on the linear term is negative while that on the quadratic term is positive. Evaluating the net effect of factor differences or GDP per worker differences on IIT in the neighborhood of the data indicates that IIT covaries negatively with the factor differences. However, the fixed-effects estimates show that both the linear and quadratic KLDIF terms are about zero. When income differences proxy for factor differences, the OLS coefficient on the linear term is positive and that on the quadratic term is negative. The sign pattern is exactly opposite in the fixed-effects case. However, evaluating the net effect in the neighborhood of the data indicates that per worker income differences covary positively with IIT with both the OLS and fixed-effects estimators.

It appears that the results presented in Table V are robust to some other reasonable specifications. In the year-by-year cross-
sectonal regressions, and in the OLS regressions with pooled years, our measures were either negative and significant (YDIF), or insignificant (KLDIF). When we estimate country-pair dummies and remove all the cross-sectional variation, the coefficient for both measures becomes positive and significant. Why is this?

One explanation might be that we have very little time series variation in the right-hand-side variables, KLDIF and YDIF. That is, relative capital-labor ratios for a given country-pair do not change much over time, so that when we sweep out cross-sectional variation, there is nothing left for IIT to vary against. However, an analysis of variance (ANOVA) shows that 58 percent of the total variation in KLDIF is between country-pairs (cross-sectional variation), and 42 percent is within country-pairs (time series variation). The ANOVA for YDIF shows that 65 percent of the variation is between, and 35 percent within. In both cases, it would appear that there remains sufficient variation after mean-differencing to give interesting results.

The second explanation is that the industry classifications in the trade data are far noisier than are supposed in the simple theoretical model. Thus far, we have uncritically accepted the SITC categories as appropriate definitions for industries. There is some danger that, by measuring intraindustry trade with SITC classifications, our results are subject to an aggregation problem (see Finger [1975]). For example, SITC categories sometimes group goods with similar consumption uses, but different factor inputs. Trade within this “industry” would be measured as intraindustry, when in fact it is motivated by relative factor abundance. The reverse is true when SITC categories fail to group goods that ought properly be considered an industry, i.e., SITC 7361 (metal cutting machine tools) and SITC 7362 (metal-forming machine tools). When SITC classifications fail to capture appropriate industry definitions, the sign on the factor differences variable becomes ambiguous. The difficulty with this explanation is that there is no necessary reason why factor differences and intraindustry trade should be negatively correlated in cross section, and positively correlated in time series. Put another way, if the classification problem were to bias our estimates, the bias should not vary depending on whether our variation is cross-sectional or time series. Indeed, this offers another plausible reason for preferring a fixed-effects estimator. If the bias in the data due to inappropriate aggregation is constant over time, it will be swept out when we mean-difference the data.

A third possible explanation emphasizes the role of geography.
There are several ways in which geography might play a significant role in intraindustry trade. First, countries sharing a border may see two-way trade in homogeneous goods, and such trade will appear in the data as intraindustry trade. This is more likely to be important for country-pairs that share a long border like the United States and Canada. Second, one can imagine a model in which distance (modeled with transport costs) may have a larger negative effect on intraindustry trade than on interindustry trade. (Simply adding transport costs to a model of international trade will tend to decrease all types of trade.) In a model with monopolistic competition, consumers will, on the margin, trade off the increased utility of another variety with increased transport costs. If the elasticity of substitution between varieties of a differentiated product is greater than the elasticity of substitution between homogeneous goods, a decline in distance will have a larger (positive) effect on the volume of intraindustry trade than it does on the volume of interindustry trade.

If proximate countries have similar per capita (or per worker) income, we may see a spurious correlation between factor differences and the $IIT_{jk}$ index in cross section. That is, nearby countries may have similar incomes for some unspecified reason, and they may have much intraindustry trade because of low transport costs. By estimating country-pair dummies in equation (7), we sweep out the constant effect of geography on intraindustry trade. Only the correlation between intraindustry trade and factor differences, independent of geography, remains, and it is no longer negative as predicted by theory.

One can begin to evaluate the relevance of some of these explanations by examining the magnitude of the estimated fixed effects from (7). In Table VII we report some normalized country-pair intercepts. The left panels of the table show country-pairs with large intercepts (at least one standard deviation above the mean) implying large amounts of intraindustry trade. Two things are remarkable. One, Ireland appears as one of the countries in five of the top seven pairs. These intercepts come from a regression that included variables for relative size ($MINGDP$ and $MAXGDP$). When we reestimate equation (7) without the size variables, Ireland is no longer among the country-pairs with large intercepts, and in fact, can be seen as a low-end outlier in some cases. This seems to indicate that size adjusts these estimates in important ways, and that Ireland, given its small size, has an especially large amount of intraindustry trade. This may be because of Ireland’s
tax policies with respect to multinational corporations. Also, of those country-pairs that do not include Ireland, nearly all share a border.

The right panels of Table VII contain country-pairs with very small intercepts (at least one standard deviation below the mean) and hence imply very little intraindustry trade. Fourteen of the fifteen country-pairs include either Canada, Japan, or the United States. Of the countries in our OECD sample, these are the only three outside of Europe, suggesting that perhaps oceans matter. The difficulty with interpreting these intercepts, though, is that they contain more than geographical information. Anything affecting intraindustry trade that is specific to country-pairs and does not change much over time will be captured in them. This might include geography, culture and language, trade barriers, or endowments of land, which mattered greatly in our cross-sectional results. For example, in the results reported above, one cannot ascertain whether Canada and the United States have low levels of intraindustry trade with their partners because they are geographically distant, because they are outside the European customs union, or because they have very different land endowments.
### Table VIII

**Country-Pair Outliers from Fixed-Effects Estimates of Equation (7) after Controlling for Land**

<table>
<thead>
<tr>
<th>Country-pair</th>
<th>Intercept</th>
<th>Country-pair</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ireland</td>
<td>2.88</td>
<td>Canada</td>
<td>Japan</td>
</tr>
<tr>
<td>Ireland</td>
<td>2.31</td>
<td>Japan</td>
<td>United States</td>
</tr>
<tr>
<td>Denmark</td>
<td>1.80</td>
<td>Canada</td>
<td>Italy</td>
</tr>
<tr>
<td>Germany</td>
<td>1.69</td>
<td>Japan</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Canada</td>
<td>1.44</td>
<td>France</td>
<td>Japan</td>
</tr>
<tr>
<td>Austria</td>
<td>1.37</td>
<td>Germany</td>
<td>Japan</td>
</tr>
<tr>
<td>Ireland</td>
<td>1.25</td>
<td>Italy</td>
<td>Japan</td>
</tr>
<tr>
<td>Canada</td>
<td></td>
<td>Germany</td>
<td>-2.12</td>
</tr>
<tr>
<td>Canada</td>
<td></td>
<td>France</td>
<td>-1.95</td>
</tr>
<tr>
<td>Italy</td>
<td></td>
<td>United States</td>
<td>-1.84</td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td>United States</td>
<td>-1.83</td>
</tr>
<tr>
<td>France</td>
<td></td>
<td>United States</td>
<td>-1.65</td>
</tr>
<tr>
<td>Italy</td>
<td></td>
<td>United Kingdom</td>
<td>-1.59</td>
</tr>
<tr>
<td>Canada</td>
<td></td>
<td>United Kingdom</td>
<td>-1.49</td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td>Italy</td>
<td>-1.47</td>
</tr>
<tr>
<td>United Kingdom</td>
<td></td>
<td>United States</td>
<td>-1.35</td>
</tr>
<tr>
<td>Germany</td>
<td></td>
<td>United Kingdom</td>
<td>-1.30</td>
</tr>
</tbody>
</table>

Large intercepts are defined as one standard deviation above the mean, while small intercepts are one standard deviation below the mean. Intercepts are normalized around zero for purposes of this table.

To address this confusion of effects, we regressed our IIT index on land variables\(^{30}\) in order to sweep out the effect of relative land endowments, and used the residuals from this regression as the dependent variable in a reestimation of equation (7). Very large and very small country-pair intercepts are reported in Table VIII. When comparing these with Table VII, it is remarkable that the outliers are very similar before and after sweeping out land effects. Put another way, controlling for land endowments does not change our notion of which countries have particularly large or small levels of intraindustry trade.

We repeat this strategy, this time regressing our IIT index on land and distance variables\(^{31}\) and using the residuals as the dependent variable in a reestimation of equation (7). Outliers are

---

30. This has been done variously with factor difference measures that include land (i.e., TLDIF and TKDIF) and with differences in land alone. Results are robust.

31. We interact a distance variable with year dummies and use these interaction terms to control for distance. This allows distance effects to change over time and reflect changing transport costs.
reported in Table IX. The contrast between this table and the two previous is notable. The six largest country-pairs include the United States, and eleven of the largest thirteen include the United States or Japan. Prior to controlling for distance effects, these countries were seen as low-end outliers; after, they are seen as high-end outliers. These results are also robust to an estimation that controls for distance but not land effects.

Finally, to further address the competing effects of land endowments and distance, we include a distance variable in our estimation of equation (5). We report the results in Table X, and compare them with those in Table IV. We see that the coefficient on \( DISTANCE \) is negative and highly significant. \( TLDIF \) (differences in land-to-labor ratios) is negative as in Table IV, but the significance of the estimates is reduced, especially in the later years. The explanatory power of the estimates that include \( DISTANCE \) is much higher than that of the estimates that do not. As noted earlier, there are a large number of combinations of factor difference variables we might include in estimating equation (5), and that only \( TLDIF \) and \( TKDIF \) had previously been negative and significant. After including \( DISTANCE \) in the estimation, only
**Table X**

Equation (7) OLS Estimates with Distance, Capital-to-Labor, and Land-to-Labor Ratios (1962–1983)

<table>
<thead>
<tr>
<th>Year</th>
<th>KLDIF</th>
<th>TLDIF</th>
<th>MINGDP</th>
<th>MAXGDP</th>
<th>DIST</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1962</td>
<td>-0.086</td>
<td>-0.056</td>
<td>0.464**</td>
<td>0.160*</td>
<td>-0.601**</td>
<td>.647</td>
</tr>
<tr>
<td>1963</td>
<td>-0.124</td>
<td>-0.073*</td>
<td>0.456**</td>
<td>0.124</td>
<td>-0.527**</td>
<td>.630</td>
</tr>
<tr>
<td>1964</td>
<td>-0.097</td>
<td>-0.103**</td>
<td>0.412**</td>
<td>0.123</td>
<td>-0.427**</td>
<td>.558</td>
</tr>
<tr>
<td>1965</td>
<td>-0.059</td>
<td>-0.095**</td>
<td>0.408**</td>
<td>0.165*</td>
<td>-0.461**</td>
<td>.593</td>
</tr>
<tr>
<td>1966</td>
<td>-0.086</td>
<td>-0.104**</td>
<td>0.372**</td>
<td>0.140*</td>
<td>-0.447**</td>
<td>.624</td>
</tr>
<tr>
<td>1967</td>
<td>-0.071</td>
<td>-0.100**</td>
<td>0.373**</td>
<td>0.178**</td>
<td>-0.443**</td>
<td>.631</td>
</tr>
<tr>
<td>1968</td>
<td>-0.072</td>
<td>-0.115**</td>
<td>0.337**</td>
<td>0.137*</td>
<td>-0.428**</td>
<td>.650</td>
</tr>
<tr>
<td>1969</td>
<td>-0.059</td>
<td>-0.110**</td>
<td>0.358**</td>
<td>0.141*</td>
<td>-0.418**</td>
<td>.633</td>
</tr>
<tr>
<td>1970</td>
<td>-0.040</td>
<td>-0.118**</td>
<td>0.354**</td>
<td>0.126*</td>
<td>-0.428**</td>
<td>.642</td>
</tr>
<tr>
<td>1971</td>
<td>-0.030</td>
<td>-0.111**</td>
<td>0.334**</td>
<td>0.118*</td>
<td>-0.458**</td>
<td>.668</td>
</tr>
<tr>
<td>1972</td>
<td>-0.028</td>
<td>-0.102**</td>
<td>0.341**</td>
<td>0.154*</td>
<td>-0.463**</td>
<td>.641</td>
</tr>
<tr>
<td>1973</td>
<td>0.011</td>
<td>-0.080**</td>
<td>0.279**</td>
<td>0.167**</td>
<td>-0.478**</td>
<td>.66</td>
</tr>
<tr>
<td>1974</td>
<td>0.037</td>
<td>-0.108**</td>
<td>0.232**</td>
<td>0.166**</td>
<td>-0.441**</td>
<td>.669</td>
</tr>
<tr>
<td>1975</td>
<td>0.029</td>
<td>-0.075*</td>
<td>0.245**</td>
<td>0.197**</td>
<td>-0.509**</td>
<td>.654</td>
</tr>
<tr>
<td>1976</td>
<td>0.071</td>
<td>-0.061*</td>
<td>0.211**</td>
<td>0.206**</td>
<td>-0.546**</td>
<td>.668</td>
</tr>
<tr>
<td>1977</td>
<td>0.057</td>
<td>-0.065*</td>
<td>0.200**</td>
<td>0.199**</td>
<td>-0.558**</td>
<td>.652</td>
</tr>
<tr>
<td>1978</td>
<td>0.064</td>
<td>-0.067*</td>
<td>0.151*</td>
<td>0.213**</td>
<td>-0.489**</td>
<td>.559</td>
</tr>
<tr>
<td>1979</td>
<td>0.078</td>
<td>-0.030</td>
<td>0.169**</td>
<td>0.200**</td>
<td>-0.544**</td>
<td>.642</td>
</tr>
<tr>
<td>1980</td>
<td>0.094*</td>
<td>-0.035</td>
<td>0.161**</td>
<td>0.190**</td>
<td>-0.546**</td>
<td>.644</td>
</tr>
<tr>
<td>1981</td>
<td>0.116**</td>
<td>-0.030</td>
<td>0.152**</td>
<td>0.190**</td>
<td>-0.570**</td>
<td>.697</td>
</tr>
<tr>
<td>1982</td>
<td>0.127**</td>
<td>-0.021</td>
<td>0.146**</td>
<td>0.175**</td>
<td>-0.541**</td>
<td>.709</td>
</tr>
<tr>
<td>1983</td>
<td>0.116**</td>
<td>-0.043</td>
<td>0.094</td>
<td>0.203**</td>
<td>-0.515**</td>
<td>.653</td>
</tr>
</tbody>
</table>

The estimated regression is

\[
HIT_{kt} = \beta_0 + \beta_1 \log \left( \frac{K_i^1}{L_i^1} - \frac{K_i^2}{L_i^2} \right) + \beta_2 \log \left( \frac{T_i^1}{L_i^1} - \frac{T_i^2}{L_i^2} \right) + \beta_3 \min (\log GDP_i^1, \log GDP_i^2) + \beta_4 \max (\log GDP_i^1, \log GDP_i^2) + \beta_5 DISTANCE + \epsilon_{kt}. \]

*indicates statistical significance at the 95 percent level.

**indicates statistical significance at the 99 percent level.

**TLDF** retains any significance, and even that is much reduced. These results all suggest that the distance between trading partners is quite important in understanding intraindustry trade both in cross-section and time series.

In this section we tested the relationship between the share of intraindustry trade and factor differences. Existing studies employ per capita income as a factor proxy, utilize cross-sectional analysis, and find a negative correlation between intraindustry trade and factor differences. We find initially that using actual factor data provides even stronger evidence of this negative correlation in the
cross section, although this holds only for certain factor measures that include land. We investigate the behavior of intraindustry trade over time, and find that is largely explained by country-pair-specific effects, not by time-varying factor measures. We attempt to decompose the country-pair effects into land and distance effects and find that the distance effect seems to be much stronger. We also investigate the competing effects of land and distance using cross-section tests, and again find distance to have a greater influence.

IV. INCONCLUSIONS

From the outset, our goal has been to test some hypotheses generated from a formal model of monopolistic competition and international trade. Previous tests had been encouraging. Studies that were not especially informed by the theory of monopolistic competition and international trade still found reasonable correlates of indexes of intraindustry trade. A study that was directly guided by the theory also found encouraging support for the theory. After reconsidering the evidence, we are not so sure.

The first test presented in this paper seems based on very unrealistic assumptions, but the theory passes with flying colors. We found that the volume of trade is well explained by a theoretically well-motivated index of the size similarity of trading partners. When confronted with data for which the theory is probably quite inappropriate, it still passes with high marks. Adding country-pair fixed or random effects, allowing for linear trends in the data, and accounting for issues of econometric endogeneity of regressors do not alter this conclusion. The relative unanimity of our results suggests that something other than monopolistic competition may be responsible for the empirical success of the gravity model. The second test we conducted allows a more reasonable underlying theoretical structure, but we find, at best, very mixed empirical support for the theory. Instead of factor differences explaining the share of intraindustry trade, much intraindustry trade appears to be specific to country-pairs. Further investigation suggests that distance is especially important to this relationship.

The results of the first test leave us genuinely puzzled. The results of the second test leave us, on the one hand, pessimistic. If much intraindustry trade is specific to country-pairs, we can only be skeptical about the prospects for developing any general theory to explain it. On the other hand, these results suggests ways in
which one might refine the theory to better fit the data. Distance, and possibly tax policy toward multinationals, are empirically important variables that are not well accounted for in the simple models we tested.\textsuperscript{32}

The theory of monopolistic competition and international trade is elegant and seems to address important aspects of reality. We hope our results also motivate others to investigate the empirical relevance of the theory, for, as promised in the introduction, we provide few answers.

**APPENDIX 1**

*Note.* Most of this Appendix is taken from work by Elhanan Helpman. It is reported here for reference purposes only.

*Theoretical Background for Section II*

Consider an economy with two countries, two factors ($K$ and $L$) and two sectors ($X$ and $Y$). Suppose that $X$ and $Y$ are differentiated products with an increasing returns to scale technology. Monopolistic competition prevails so that with free entry, equilibrium is characterized by a large number of firms, each producing a unique variety of $X$ and making zero profits.

Let $X$ and $X^*$ denote total production of good $X$ in the home and foreign country, respectively. The number of firms is given by

$$n = \frac{X}{x},$$

where $x$ is the number of home varieties and similarly for the foreign country.

Assume identical homothetic preferences and a utility function that rewards variety. Then, with costless transport, every variety of every good will be demanded in both countries. Further, each country will consume an amount of each variety proportional to its share in world GDP, $\overline{GDP}$.

Let $s$ be the home country's share in world GDP. That is,

$$s = \frac{GDP}{\overline{GDP}} \quad \text{and} \quad s^* = (1 - s),$$

where $GDP + GDP^* = \overline{GDP}$. Assuming balanced trade, then the home country consumes $spn^*x^* = (spX^*)$ of the foreign $X$ good, and the foreign country consumes $s^*pnx = s^*pX$ of the home $X$

\textsuperscript{32} An admirable move in this direction is Brainard [1995].
good, where \( p \) is the price of good \( X \) and the price of good \( Y \) is normalized to one. Since \( y \) is also differentiated, the volume of trade is given by

\[
VT = s(pX^* + Y^*) + s^*(pX + Y).
\]

The bracketed terms are just foreign and home GDP so

\[
VT = sGDP^* + s^*GDP.
\]

It follows that

\[
(A1) \quad VT = 2sGDP^* = \frac{2 \cdot GDP \cdot GDP^*}{GDP} \cdot \frac{GDP}{GDP} = 2ss^*GDP.
\]

The bilateral volume of trade achieves a maximum when \( s = s^* \).

Note that the same relationship between trade volume and relative size holds any time there is complete specialization in production. For example, let \( X \) and \( Y \) be homogeneous goods, and assume that the home country produces only \( X \) and the foreign country produces only \( Y \). Then, \( \bar{X} = X + X^* = X \) and \( \bar{Y} = Y + Y^* = Y^* \). Identical homothetic preferences imply that

\[
VT = sY^* + s^*X = sGDP^* + s^*GDP = 2ss^*GDP.
\]

It is possible to generalize \((A1)\) so that it holds for groups of countries of any size. For a group of countries, \( A \), we have

\[
GDP^A = \sum_{j \in A} GDP^j,
\]

where \( GDP^A \) is the GDP of group \( A \). The share of country \( j \) in group \( A \) is given by

\[
e^j_A = GDP^j/GDP^A.
\]

Similarly, the share of group \( A \) in world GDP is

\[
e^A_j = GDP^A/GDP.
\]

The within-group volume of trade is given by

\[
(A2) \quad V^A = \sum_{j \in A} \sum_{k \in A, j \neq k} s^jGDP^k = \sum_e \sum_1 s^j e^k_A GDP^A = GDP^A \sum_1 s^j(1 - e^j_A).
\]
With balanced trade, one obtains
\[ s^j = \left( e^j_A \cdot GDP^A \right) / GDP = e^j_A \cdot e_A, \]
and substitution yields (1) from Section II of the text:
\[ \frac{V^A}{GDP^A} = e_A \sum_j e^j_A (1 - e^j_A) \]
\[ = e_A \left[ 1 - \sum_j (e^j_A)^2 \right]. \]  

This is the equation Helpman graphs to study the relationship between trade volume and relative country size in the OECD.

To amend the estimating equation for trade imbalances, we employed the following correction (following Helpman exactly). With a trade imbalance,
\[ s^j = \frac{e^j_A GDP^A - T^j}{GDP}, \]
where \( T^j = X^j - M^j \), and one just substitutes for \( s^j \) into (A2). The order of the correction is the ratio of the trade imbalance to group GDP, and this is empirically negligible.

**Theoretical Background for Section III**

Now allow \( X \) to be differentiated (as before) and \( Y \) to be a homogeneous good produced with constant returns to scale. Assume that \( X \) is capital intensive and that the home country is relatively capital abundant. Then there will be two-way trade in the \( X \) good. Also, the home country will be a net exporter of \( X \) and an importer of \( Y \). In Figure II we see the direction of trade for this example. The total volume of trade is given by
\[ VT = s^pX + spX^* + sY - Y. \]

The volume of trade that is intraindustry is \( 2 \min(spX^*, s^pX) \), and the share of intraindustry trade is
\[ \text{II}T_{jk} = \frac{2 \min (s^pX, spX^*)}{s^pX + spX^* + (sY - Y)}. \]  

Helpman and Krugman [1985] show that constant intraindustry trade share-curves for endowments in the factor price equalization sets are given by Figure III. Along the \( OO^* \) diagonal, the
intraindustry trade share equals one. Factor reallocations that widen capital-to-labor differences without changing relative size decrease the share of intraindustry trade.

To see this, consider a factor reallocation from endowment point $E_1$ to $E_2$ in Figure III. We are above the diagonal at $E_1$, so the home country is relatively capital abundant. The move to $E_2$ further widens the gap between the home country's and the foreign country's capital-to-labor ratios. Also, since the move takes place along the wage-rental line, relative size is unchanged. We now ask, what happens to our intraindustry trade index?

Since incomes and preferences are unchanged, each country consumes exactly what it did before (the value of which is given by point $C$). The only thing that has changed is the location of production. The home country produces more $X$, and the foreign country produces more $Y$. Since total endowments in the world economy have not changed, $dX = dX + dX^* = 0$. Hence, $dX = -dX^*$, and similarly $dY = -dY^*$. Since we remain in the factor price equalization set, prices are unchanged; $dp = 0$. Finally, by construction, relative size has not changed; $ds = ds^* = 0$.

We wish to sign the change in (A4) that occurs as a result of this reallocation. In the numerator, $s^*pX$ is larger, but $spX^*$ is smaller, so the numerator decreases. For the denominator, take
the total derivative to yield \( s^*pdX + spdX^* - dY \). Since \( s^* = (1 - s) \) and \( dX = -dX^* \), we have \((1 - s)pdX - spdX - dY\) or \((1 - 2s)pdX - dY\). The factor reallocation causes the home country to produce more \( X \) and less \( Y \), so \( dX > 0 \), and \( dY < 0 \). Since \( s \) lies between 0 and \( \frac{1}{2} \), the term in brackets is always nonnegative. The denominator increases, so our IIIT index decreases as a result of a factor reallocation that widens factor differences without changing relative size.

APPENDIX 2

Trade Data

Trade data used in the first and second tests come from the United Nations Trade Database, years 1962–1983. The data are reported in four-digit SITC (revision 1). The volume of trade variable used in the first test is

\[
VT_{jk} = \sum_i (X_{ijk} + X_{ikj}).
\]

It comprises exports from country \( j \) to country \( k \), plus exports from country \( k \) to country \( j \), summed over industries \( i \).
The share of intraindustry trade was calculated using the Grubel-Lloyd [1975] index as described earlier. An industry is defined as a four-digit SITC group. All SITC categories were included in the calculation of both $VT_{j,k}$ and $IIT_{j,k}$.

The U. N. Trade Database contains both country $j$'s report of its exports to country $k$, and country $k$'s report of its imports from $j$. On the assumption that the importing country keeps better track of trade flows crossing its borders, we use the importing country's reported data. However, we have repeated tests in Sections II and III using importer and exporter data without much change in the reported results.

$RGDPW$ is per worker GDP in constant 1985 international prices (chain index). It is used to construct $YDIF$.

For the second test we use GDP, measured in 1985 international prices. This is used to construct $MINGDP$ and $MAXGDP$.

Factor Data

Factor data are used in the first test in the instrumental variables specification. Population data from the World Bank World Tables are used to proxy labor force. Our capital stock series has been constructed using the third method described in Appendix B of Leamer [1984]. Gross Domestic Investment, exchange rates (yearly average), and the GDP deflator, are taken from World Tables. Land endowments data are taken from Leamer [1984] and consist of land2 + land3 + land4. Distance is computed based on the latitude and longitude of the capital cities. Investment flows are converted year by year into dollars, deflated using the U. S. GDP deflator, then summed over years and depreciated appropriately.

This gives a capital stock for each year from 1962 to 1983, with accumulated investment flows denominated in the relevant year. That is, the 1970 capital stock is an accumulation of investment flows valued at 1970 prices. The World Tables Gross Domestic Investment series begins in 1960, so we assumed an initial capital stock for each country equal to 250 percent of its GDP in 1960. We assume a constant depreciation rate of 13.3 percent. This gives an asset life of fifteen years. We have constructed different series using different initial assumptions, and the first test results reported here are insensitive to these assumptions.

For the second test we require capital stock data valued in constant dollars. Leamer [1984] notes in his data appendix that the Penn World Tables provide a useful data set for constructing a capital stock series because GDP and investment flows are compa-
rable over time and across countries. We use the Penn World Tables, Mark V, series RGDPCH, POPULATION, RGDPW, and C. Using RGDPCH, WGDPC, and POPULATION, we get labor force. That is, RGDPCH/WGDPC = labor force participation rate. C is the year-by-year fraction of GDP that goes to investment.

Since the initial variables are already in 1985 international prices, we need only sum over investment flows and depreciate at 13.3 percent. That is,

\[ K_t = K_{t-1} (1 - \text{depreciation}) + \text{investment}. \]

Using Penn World Tables data, we can construct an investment series going back to 1950. We assume a 1950 capital stock equal to 250 percent of GDP.

UNIVERSITY OF CHICAGO
UNIVERSITY OF MICHIGAN

REFERENCES


Harrigan, James, "Openness to Trade in Manufactures in the OECD," mimeo, University of Pittsburgh, 1992.


---


---


