



ELSEVIER

Journal of International Economics 59 (2003) 1–23

Journal of
INTERNATIONAL
ECONOMICS

www.elsevier.com/locate/econbase

Market access, economic geography and comparative advantage: an empirical test

Donald R. Davis, David E. Weinstein*

Department of Economics, Columbia University, 420 W. 118th St. MC 3308, New York, NY 10027, USA

Received 25 October 1999; received in revised form 12 November 2001; accepted 10 December 2001

Abstract

Traditional neoclassical models of comparative advantage suggest that, all else equal, a country with idiosyncratically strong demand for a good will be an importer of that good. However, there is a contrary tradition that emphasizes the advantages of a large home market as a foundation for *exports* of a good. One recent formalization of this home market approach falls within what is termed the new economic geography. This paper integrates core models of Heckscher–Ohlin and Krugman [American Economic Review 70 (1980) 950] to investigate whether such home market effects matter empirically in manufacturing for a set of OECD countries. The evidence suggests that home market effects are important for a broad segment of OECD manufacturing.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Increasing returns; Economic geography; Comparative advantage

JEL classification: F1; D5; E1

1. Market access, economic geography and comparative advantage: an empirical assessment

The empirical trade literature has focused strongly in recent years on understanding determinants of the pattern of trade. In a seminal contribution, Leamer

*Corresponding author. Tel.: +1-212-854-6880; fax: +1-212-854-8059.

E-mail address: dew35@columbia.edu (D.E. Weinstein).

(1984) examines this question in a version of the Heckscher–Ohlin model featuring factor price equalization (FPE) and equal numbers of goods and factors. Leamer identifies a set of twelve productive resources that correlate well with countries' net trade vectors. One limitation of Leamer's approach is the focus on explaining the net trade vector. The reason this is a limitation is that it is well understood that the real intellectual capital of the theory is the predictions about the cross-country pattern of production, which is then coupled with a rudimentary and implausible theory positing identical structures of absorption.

Recognition of this limitation led researchers such as Harrigan (1995, 1997) and Bernstein and Weinstein (2002) to focus directly on the model's ability to predict cross country patterns of production. Within the framework of the frictionless FPE Heckscher–Ohlin model, this divorce between the patterns of production and absorption makes perfect sense. However this makes less sense in a world in which trade frictions matter. And other contributions to the recent research agenda, including inter alia McCallum (1995) and Davis and Weinstein (1999) have emphasized instead the importance of these frictions for understanding trade.

Trade frictions segment markets, giving rise to differences in price indices across locales. Moreover, these frictions imply that the geographical distribution of demand across markets will matter for local production patterns. In such a case, investigation of the pattern of production cannot proceed without attention to underlying demand conditions. In order to investigate this empirically, one needs to commit to specific models. A fundamental divide may be identified between two classes of models. In the first class, unusually strong demand for a good, *ceteris paribus*, makes a country an importer of a good. An example would be a conventional two-sector neoclassical model with strictly downward sloping import demands. However, there is an alternative tradition within the trade literature which emphasizes an important interaction between demand conditions and production opportunities in which the production response to local demand conditions is so powerful that strong local demand for a product leads a country to export that product. When such conditions exist, the literature terms it a home market effect.

It is important to recognize that there are a variety of models which potentially could give rise to such home market effects. For example, the so-called biological model of trade posits that nationally-differentiated products arise in response to peculiarities in local demand (cf. Bhagwati, 1982; Feenstra, 1982), in turn effectively giving rise to Ricardian advantages that lead a country to export that product. It is likewise important to recognize that the exact determinants of the home market effect may well differ across models so that it will be important in subsequent work to consider alternative specifications of the home market effect that correspond to the distinct underlying models.

This paper develops one approach to identifying empirically the existence of home market effects. The foundation for our approach is Krugman (1980) as extended by Weder (1995). The underlying model is Dixit–Stiglitz monopolistic

competition, with the novelty that the relative strength of demand for different classes of differentiated products leads production to rise more than one-for-one in response to idiosyncratic local demand. Building further on Krugman (1980) and Helpman (1981), we show how to place this model within a multi-sector Heckscher–Ohlin framework to ready it for empirical estimation.

The present paper builds on earlier work in this vein, including Davis and Weinstein (1996, 1998). The most important alteration to our earlier analytic framework is a more careful approach to characterizing idiosyncratic local demand which takes fuller account of the geographical structure of absorption and the resulting opportunities this provides to producers considering location in the various countries. The results prove favorable to the hypothesis of the existence and importance of home market effects.

In recent years, a number of papers have extended the approach developed here of assessing economic geography by examining the relationship between demand and output. In particular, Trionfetti (2001a,b) demonstrates that home market effects need not always arise in increasing returns models and develops alternative tests. Head and Ries (2001) test for home market effects using firm level data. Finally, in a very interesting paper, Feenstra et al. (2001) demonstrate that home market effects can arise even in models in which all output is homogeneous.

2. Theoretical framework for hypothesis testing

The broad outlines of our theoretical framework follow Davis and Weinstein (1996). The objective is to distinguish a world in which trade arises due to increasing returns as opposed to comparative advantage. This is very difficult if we focus on the class of zero transport cost increasing returns models deriving from Krugman (1979) and Lancaster (1980). However this is possible if we focus instead on the class of trade models that have come to be known as economic geography, which interact increasing returns and trade costs in general equilibrium.

We begin by sketching the model of Krugman (1980). The model is one of monopolistic competition. There are two classes of goods, each with many varieties. All varieties are symmetric in production and demand. Each variety is produced under increasing returns to scale with a fixed cost and constant marginal costs in units of labor. Preferences are the iso-elastic Dixit–Stiglitz form. The novelty in Krugman’s paper is the introduction to this framework of costs of trade in an iceberg form (for one unit of a good to arrive, $\tau > 1$ units must be shipped). He further assumes that there are two countries which are mirror images of each other. They have the same labor forces. The difference lies in their demand structure. For simplicity, he assumes that consumers come in two types, each specialized to consume all varieties of only one of the two classes of goods. In addition, he assumes that the sole difference between the countries is that one

country is predominantly populated by those who consume varieties of one of the classes of goods, and vice versa (in perfect mirror fashion) for the other. The symmetry insures factor price equalization in spite of the trade costs.

An important feature of the model is that the combination of constant mark-ups and free entry implies that in equilibrium output per firm is the same across markets in spite of the trade costs. This means that a full description of the equilibrium can be given by the number of varieties of each of the two types produced in each country. Let μ be the number of varieties of good g produced at Home relative to those produced abroad. Let a $\sigma < 1$ be the ratio of demand for a typical import relative to a domestically produced variety in the same class. Let λ represent the ratio of demanders for good g at Home relative to the number in Foreign. Krugman shows that in the range of incomplete specialization, the relative production levels μ can be described as:

$$\mu = \frac{\lambda - \sigma}{1 - \lambda\sigma}$$

When $\lambda = 1$, demand patterns are identical and the countries produce the same number of varieties in each industry, leaving a zero net balance. This will play an important role when we turn to our empirical implementation as it suggests that predictions of production structure, *ceteris paribus*, should be centered around an even distribution of the industries across the countries. Idiosyncratic demand components will then explain deviations from this neutral production structure.

Moreover, we need to consider closely the way in which idiosyncratic demand components will translate into alterations in production structure. From above, and for the range of incomplete specialization for which these relations are valid,

$$\frac{\partial \mu}{\partial \lambda} = \frac{1 - \sigma^2}{(1 - \sigma\lambda)^2} > 1$$

Krugman emphasized that this will imply that countries with a large home market for a good will be net exporters of that good. For our purposes it is convenient to focus on an equivalent statement of this result that speaks directly to the implications for production. That is, idiosyncratic demand patterns (indexed by λ) have a *magnified* impact on production patterns. This will play a crucial role in our empirical implementation, helping to separate the influences of economic geography from that of comparative advantage.

Why does the home market effect arise? In the presence of trade costs, producers will have an incentive to locate near the larger source of demand. This is counterbalanced by the fact that as more and more producers leave the smaller market, those who remain experience the trade costs not only as an inhibition on their deliveries to the larger market, but also as protection against the many producers who have located in that larger market. *Ex ante* it may not seem obvious which of these influences will dominate. However, it is possible to show that if the share of varieties produced moved exactly one-for-one with the idiosyncratic

demand that those producers located in the large country would have higher demand for their products than those located in the smaller market for that good. Since equilibrium requires that the derived demand be the same for all producers, this implies residual incentives for producers to move to the large market—hence the home market effect.

It is likewise important to think about why the home market effect does not arise in the conventional constant returns to scale comparative advantage framework. The logic turns out to be very simple. Consider a positive shock to the home demand structure for a good. Will this call forth additional local supply, and if so will supply move more than one-for-one (as required for the home market effect)? If the production set is strictly convex, additional supply of the good will be forthcoming only if its relative price rises. But then, provided the foreign export supply curve has the conventional positive slope, this will also call forth additional net exports from abroad. In such a case, the idiosyncratic demand will be partly met by additional local supply and partly by higher imports. Local supply, then, moves less than one-for-one with the idiosyncratic demand. In this conventional comparative advantage world, there is no home market effect.

Of course, Krugman (1980) cannot be taken straight to data. Such models of economic geography contemplate highly abstract worlds in order to provide clear theoretical insights. Even in such stark models, the inherent complexity of the problems frequently defies analytic solution. While the robustness of the home market effect has been explored along a variety of dimensions (e.g. Weder, 1995), there is no single fully-solved model that has simultaneously incorporated the myriad elements essential for empirical implementation. Our approach is to hew as closely as possible to the theory, and so provide a highly-structured interpretation of the models. Where it is not possible to provide a full solution, we make what we consider the most sensible match between theory and specification.

3. Implementing the search for home market effects

3.1. Methodology

We begin with a sketch of the theoretical framework. The specification and data work consider three levels of product aggregation: Varieties, Goods, and Industries. *Varieties* play an important theoretical role within the model of economic geography. In the Dixit and Stiglitz (1977) formulation, they are the locus of increasing returns in production, as well as the elements across which consumers have a preference for variety. While they play an important theoretical role, we assume they exist at a greater level of disaggregation than exists in our data. *Goods*, in our formulation, can be thought of in two ways. Under the hypothesis of increasing returns, a good is a collection of a large number of varieties produced under monopolistic competition. It is at the goods level that differences in the

composition of demand give rise to home market effects. By contrast, under the hypothesis of comparative advantage, a good is a traditional homogeneous commodity. *Industries*, in both frameworks, consist of a collection of goods produced using a common technology. In the comparative advantage framework, we interpret these as simple Leontief input coefficients. In the increasing returns framework, we assume that both fixed and marginal costs of all varieties of all goods within an industry use inputs in a fixed proportion. In our data work, industries and goods are typically 3- and 4-digit ISIC data respectively.

The null hypothesis that we consider is that comparative advantage determines production and trade. The particular model of comparative advantage that we implement is the so-called square Heckscher–Ohlin (HO) model, i.e. with equal numbers of goods and factors (cf. Ethier, 1984). All countries share identical Leontief technologies of production, which are linearly independent, so that the technology matrix is invertible. Let n be an index of industries, g of goods, and c of countries. Let it stand for the whole world, and ROW stand for the rest of the world (excluding country c). Let X_{ngc} and $X_{ng\text{ROW}}$ be total output in industry n of good g for country c and the rest of the world respectively. Let V_c be the vector of endowments of country c . Let Ω be the inverse of the technology matrix, and Ω_{ng} be the row corresponding to the g th good in industry n . Then our Heckscher–Ohlin model of goods production is given by:

$$X_{ngc} = \Omega_{ng} V_c \quad (1)$$

The alternative that we consider is what we term the Helpman–Krugman specification. It is inspired by Helpman’s (1981) integration of Heckscher–Ohlin with a zero transport cost model of monopolistic competition. But in place of the latter we substitute the Krugman (1980) model of economic geography.

Accordingly, we assume output structure is determined in two stages. We assume the Heckscher–Ohlin model determines the broad industrial structure of a country. Let $\bar{\Omega}_n$ be the n th row of an inverse technology matrix for industry output, where the coefficients indicate average inputs at the equilibrium scale per variety (which is constant within an industry). Let G_n be the number of products in industry n . Then output in industry n in country c is given by:

$$X_{nc} = \sum_{g=1}^{G_n} X_{ngc} = \bar{\Omega}_n V_c \quad (2)$$

While we assume endowments map perfectly to industry-level output, we also assume they tell us nothing about the composition of production across the goods within an industry. Since all varieties of all goods within an industry are assumed to use the same mix of factors, these may be thought of as a composite factor—an analogue to the single factor ‘labor’ of Krugman (1980). Because of the Leontief technology assumption, resource constraints become industry-specific within a country.

We may think of the determination of the output of the various goods within an industry in two stages. Absent idiosyncratic elements of demand, each country allocates its resources across the goods within a particular industry in the same proportion as all other countries. This provides the country with a base level of production for each good in an industry that we denote SHARE. The second component arises when there are idiosyncratic elements of demand across the goods—what we term IDIODEM. These give rise to home market effects, here a more than one-for-one movement of production in response to idiosyncratic demand.

In order to make this precise, we must distinguish between a country's demand for a good produced in many locations, which we denote D_{ngc} , from the derived demand facing producers in a particular locale which forms the basis for the construction of IDIODEM, the latter of which we denote \tilde{D}_{ngc} (and which is discussed in more detail below). We may denote the correlate for the rest of the world as \tilde{D}_{ngROW} . Because output and demand shares figure prominently in our discussion, it is convenient to define some additional variables. Let $\gamma_{ngROW} \equiv X_{ngROW}/X_{nROW}$ be the share of good g industry n in the rest of the world (and which, of course, varies with c). Let $\tilde{\delta}_{ngc} \equiv \tilde{D}_{ngc}/\tilde{D}_{nc}$ be the share of good g in industry n 's derived demand in country c . With these definitions in hand, the specification may be written in a general form as:

$$X_{ngc} = \alpha_{ng} + \beta_1 \text{SHARE}_{ngc} + \beta_2 \text{IDIODEM}_{ngc} + \epsilon_{ngc} \quad (3)$$

where

$$\text{SHARE}_{ngc} \equiv \gamma_{ngROW} X_{nc}, \quad \text{IDIODEM}_{ngc} \equiv (\tilde{\delta}_{ngc} - \tilde{\delta}_{ngROW}) X_{nc}.$$

IDIODEM is our measure of the extent of idiosyncratic derived demand. The term in parentheses measures the extent to which the relative demand for a good within an industry differs from that in the rest of the world. If all countries demand goods in the same proportion, then IDIODEM is identically zero. When relative demand for producers of a good in one country is higher (lower) than that in the rest of the world, IDIODEM is positive (negative). Multiplying this term by X_{nc} gives IDIODEM the correct scale and units to include in the regression.

If instead we believe that endowments may matter for the structure of 4-digit production, then Davis and Weinstein (1996) show that an appropriate way of nesting the models is as follows:

$$X_{ngc} = \alpha_{ng} + \beta_1 \gamma_{ngROW} X_{nc} + \beta_2 (\tilde{\delta}_{ngc} - \tilde{\delta}_{ngROW}) X_{nc} + \Omega_{ng} V_c + \epsilon_{ngc}, \quad \text{or} \quad (4)$$

$$X_{ngc} = \alpha_{ng} + \beta_1 \text{SHARE}_{ngc} + \beta_2 \text{IDIODEM}_{ngc} + \Omega_{ng} V_c + \epsilon_{ngc} \quad (4')$$

Within the set of models contemplated in this paper, this approach allows us to use the estimate of β_2 to distinguish three hypotheses. In a frictionless world

(comparative advantage or increasing returns), the location of demand does not matter for the pattern of production, so we would predict $\beta_2 = 0$. When there are frictions to trade, demand and production are correlated even in a world of comparative advantage, reaching exactly one-for-one when the frictions force autarky. However production does not rise in a more than one-for-one manner. Accordingly, if we find $\beta_2 \in (0, 1]$, we conclude that we are in a world of comparative advantage with transport costs. Finally, in the world of economic geography, we do expect the more than one-for-one response, hence $\beta_2 > 1$. Summarizing, the estimate of β_2 allows us to distinguish three hypotheses:

- $\beta_2 = 0$ Frictionless world (comparative advantage or IRS)
- $\beta_2 \in (0, 1]$ Comparative advantage with frictions
- $\beta_2 > 1$ Economic geography

These form the basis for our hypothesis tests.

Direct estimation of Eq. (4) is not possible because of the simultaneity problem arising from having industry output on the right-hand side and the output of a good within that industry on the left. We can eliminate this simultaneity by remembering that, in our framework, endowments determine industry output. Using endowments as instruments for X_{nc} eliminates the simultaneity problem.

There are a number of ways in which we can estimate Eq. (4) in addition to estimating the full system. If one believes that endowments do not matter at the goods level, then one can force Ω to equal zero for every factor and industry. If one excludes factor endowments, one should expect the coefficient β_1 to equal unity. This is due to the fact that ceteris paribus one expects the share of goods production within an industry to be the same across countries. While Davis and Weinstein (1996) confirm this, the parameter often has much larger standard errors and deviates far from unity in specifications including endowments. This owes to the high degree of multicollinearity between SHARE (which is formed in part using endowment instruments) and the endowments. Since we found that the crucial coefficient on β_2 in specifications with endowments is largely invariant to the inclusion of SHARE, we dropped the latter from our specifications with endowments.¹

¹Davis and Weinstein (1996, 1999) found that in specifications with endowments and SHARE, β_1 is negative and significant. This likely results from an identification problem that arises when we include SHARE and endowments. Since \hat{X}_{nc} is a linear function of endowments, if there were no movement in γ_{ngROW} across countries, SHARE would be perfectly collinear with endowments and we could not estimate a coefficient. This is what would have occurred if we had calculated SHARE using γ_{ngW} (where W indicates world values) instead of γ_{ngROW} . The linear relationship between endowments and \hat{X}_{nc} means that we would obtain an identical coefficient if we replaced SHARE with $(\gamma_{ngROW} - \gamma_{ngW})\hat{X}_{nc}$. Identification here is achieved by examining the difference between γ_{ngROW} and γ_{ngW} . This is likely to produce a negative coefficient because the share of four-digit output in the rest of the world is likely to be below the world average precisely when output in a country is above average.

Our construction of IDIODEM contrasts with the approach in the earlier Davis and Weinstein (1996) by focusing not only on locally idiosyncratic demand, but rather a derived demand that takes account of idiosyncracies in demand in neighboring countries as well. Let this derived demand be denoted by \tilde{D}_{nge} . This is derived in a few steps. First we run industry-level gravity regressions to obtain parameters reflecting the impact of distance on demand. We then use these elasticities to form weighted derived demands for each market subject to a scaling so that the sum of derived demands equals world demand for the entire market. We follow Leamer (1997) in our specification of the distance of a country from itself. These derived demands then serve as the input for calculations of IDIODEM. Our estimates of the gravity equation are reasonable and the measures of IDIODEM depart in sensible ways from the corresponding measures in Davis and Weinstein (1996).²

3.2. Data

The theories examined in this paper relate the structure of output to the structure of a country's available factor endowments and idiosyncratic components of demand facing a country's producers. This idiosyncratic demand, in turn, is a weighted average of absorption across a country's trading partners. The weights themselves depend on estimates from a gravity equation of trade, hence on economic size, bilateral distance, and the characteristics of the particular industry which determine how demand dissipates with distance. These define the data required for our study.³

The OECD's Compatible Trade and Production (COMTAP) data set is the foundation for our study. It provides comparable trade and production data for 13 members of the OECD disaggregated through the four-digit ISIC level [Australia, Belgium/Luxembourg, Canada, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Sweden, UK, USA] and for 22 members of the OECD through the three-digit ISIC level [Austria, Denmark, Greece, Ireland, New Zealand, Portugal, Spain, Turkey, Yugoslavia]. National absorption is measured as the residual between output and net trade. World outputs and absorption levels are calculated by summing across all available countries. Country capital stocks are from the

²A great deal more detail on our procedures is available in the working paper version of this paper, Davis and Weinstein (1999).

³In order to allow comparability with David and Weinstein (1996) we use the same data set, with one small amendment. For three sectors (other food products, rubber products, and professional and scientific equipment) Belgium–Luxembourg and Finland only report one four-digit sector within a three-digit sector. The values Belgium–Luxembourg and Finland report seem exceptionally large in these sectors and lead us to suspect that data from other four digit sectors is included in these sectors. We therefore delete these industries from the data set. However, before doing so we re-ran our equations with and without these sectors and found that the results in David and Weinstein (1996) are robust to the inclusion of these three outliers.

Penn World Tables v. 5.6. World endowments of labor force by educational level are from the UNESCO *Statistical Yearbook*. Fuel production is equal to the sum of the production of solid fuels, liquid fuels, and natural gas in coal-equivalent units as recorded in United Nations' *Energy Statistics Yearbook*. OECD's COMTAP bilateral import and export numbers as prepared by Harrigan (1993) and made available by Feenstra (1996) also underpin our gravity estimates. Country distance is measured as the distance between the major economic centers in the respective countries and comes from Wei (1996). Measurements of how far countries are from themselves are taken from Leamer (1997).

Implementation of the economic geography framework, as embodied in Eq. (4), requires data at two levels of aggregation. At the higher level of aggregation, endowments determine the structure of output, while at the more disaggregated level, economic geography is expected to exert its force. Unfortunately, theory does not indicate how to find a level of disaggregation where factor endowments cease determining production structure and specialization is driven by increasing returns and demand patterns. Our strategy is to use the most detailed cross-national data we can find, and then assume that goods at the most disaggregated levels represent a collection of monopolistically competitive varieties.

One concern about use of these data is whether the actual criterion for industrial classification is congruent with the underlying theoretical categories. It is not. Actual classification is by product usage rather than simply by factor input composition, as would be strictly required by the theory. Maskus (1991) examined this issue for ISIC three- and four-digit industries and found that while there is greater similarity of factor intensities within three-digit sectors than across them, there still is substantial variation within three-digit sectors. Thus, although it is true, for example, that the skilled to unskilled ratio in precision instruments exceeds that in textiles, there is no guarantee that this is true in comparing every good produced within the respective industries. This could pose problems for our tests. Within the economic geography framework, the assumption that all four-digit goods in a three-digit industry use common input proportions served to replicate the one-factor world of Krugman (1980). Heuristically this implied that our production possibility surface had Ricardian flats, so a constant marginal opportunity cost of shifting production from one good to another. Assuming instead that the goods use different input proportions could then imply a rising marginal opportunity cost of expanding one good in terms of the other. This might tend to diminish the responsiveness of production to idiosyncratic demand, implying that the IDIODEM coefficients might be less than unity even if the world is one of economic geography. We acknowledge this possibility. Yet we remain skeptical that this view is correct. Quite apart from the empirical issues of how the goods are classified into industries, we know that if the number of goods is large relative to the number of primary factors, the production surface (here in units of varieties) will again have flats [see Chipman (1992)]. Demand again could play the crucial role in making the production and export patterns determinate—the key being that

production expansion for a single good again need not imply rising marginal opportunity cost in terms of other goods.

In principle, working at the four-digit level enabled us to break manufacturing up into 82 four-digit sectors, but because in 13 cases there is only one four-digit sector within a three digit sector, our sample is reduced to 69 four-digit and 27 three-digit sectors. In addition, we dropped another 14 four-digit sectors due to missing observations for some countries. In two sectors (fur dressing and dyeing and manufacturing goods, not elsewhere classified), we obtained large negative numbers for domestic absorption for a number of countries so we dropped those industries. For a few out of the remaining 702 observations, imputed domestic absorption is negative but very small (1–2 per cent of production), and we attributed these negatives to measurement error and reclassified these amounts as zeros. There are 53 four-digit industries that we eventually used in the analysis. Many of the industries at the four-digit level, such as carpets and rugs, and motor vehicles, have been used as examples of monopolistically competitive industries. Indeed this level of disaggregation is basically the same as the one used by Krugman (1991) to support his hypothesis that geography matters for trade.

Because of data limitations, we are forced to measure domestic absorption as a residual. Measuring domestic absorption by using a residual potentially introduces a bias into our sample through the mis-measurement of production. If production is recorded at too high a level for a particular year, that will also tend to cause measured absorption to rise. This creates a simultaneity problem if we use contemporaneous demand. Furthermore, since the spirit of economic geography models is to explore how long-run historical demand deviations affect production, we thought it inappropriate to regress current production on current demand. In order to deal with both of these issues, we decided to use average demand over the period 1970–1975 to identify idiosyncratic components of demand, while other variables in our regressions are values for 1985. We also ran all specifications with demand calculated over the time period 1976–1985 and just 1985 and obtained results qualitatively the same.

4. Estimation

4.1. Pooling and aggregation

Our discussion makes a clear analytic distinction between various levels of aggregation—varieties, goods, and industries. No such neat division exists in the data. Thus the level of aggregation at which to implement our methodology is a matter of judgment and subject to data availability. If data were not a constraint, our inclination would be to think of goods as being at a level of disaggregation greater than exists in the currently available data. Accordingly, in considering only this aspect of the problem, our preference is to work with the most disaggregated

data available. We do, though, consider a case at a higher level of aggregation since this provides us with more observations and allows comparability with previous work.

A second important consideration is the extent to which we should pool observations across goods and industries. There is a clear advantage to pooling—it increases the number of observations. This is potentially important, since in our most disaggregated runs we will have only thirteen observations per good. However there is correlatively an important disadvantage of pooling—it forces us to impose more structure on the estimates, and so leads us further from the underlying analytic model. These include assumptions of common input proportions, demand symmetry, and equilibrium scale economies for all varieties of all goods within an industry. *Ex ante* it is difficult to know whether we should be happier with estimates in which the theoretical model is more appropriate but there are very few observations or the contrary case.

Our approach is to implement the estimation at a variety of levels of both pooling and aggregation. If home market effects exist, we would at least like to see some indication of their presence in the various exercises. However we should likewise be cognizant that since these place quite distinct constraints on the data, it will be asking too much to expect a perfect mapping among results from the varied runs.

We pursue four estimation exercises. In three of these, the dependent variable is four-digit production, with the runs distinguished by the extent of pooling, while the fourth treats three-digit output as the dependent variable for individual industry runs. Consider first what we term the ‘pooled’ run. This exercise pools all four-digit observations for the estimation of a single coefficient on *IDIODEM*. The great advantage of this exercise is that there are 650 observations. The disadvantages lie in that implicitly we must assume that either all industries are comparative advantage or all are economic geography, and that we must assume there is a common structure determining the coefficient on *IDIODEM* for all goods in all industries. We next move to the opposite extreme, that of individual ‘four-digit’ good runs. The advantages of this exercise are that it is closest to the analytic structure we posit and that it allows the most detailed comparison across sectors of the presence or absence of home market effects. The disadvantage is that data availability implies there are only thirteen observations per four-digit sector. We next report an intermediate approach which pools all observations for four-digit goods within a three digit industry. We may term these ‘industry-pooled’ runs. This approach trades off the advantages of the previous two exercises. It imposes less structure than the fully pooled runs, but typically has four times as many observations as the individual four-digit industry runs. This also suggests the downside, namely the fact that it forces us to impose some common structure within industries that may not be fully suggested by the results in the four-digit runs themselves.

Our final exercise returns to individual sectoral runs. The departure is that

industries are now defined as two-digit output, and goods are three-digit output, and so are now the dependent variables. This has two important advantages. The first is that we do gain some observations relative to the four-digit runs, since twenty-two countries report the three-digit data. The second is that this structure and level of aggregation can be directly compared with results of Davis and Weinstein (1999) on Japanese regional data. There are three disadvantages to this exercise. The first is the loss of observations relative even to the industry-pooled runs. The second is that the additional observations relative to the four-digit runs come through the addition of countries that likely have lower quality data. Third, for related reasons, moving from the initial thirteen to twenty-two countries likely leads to a greater violation in our assumption of a common economic structure for all countries.

These four exercises provide different windows on the home market effect. As we have seen, each exercise has advantages and drawbacks. Hence to judge the results, we should not rely too heavily on any single exercise, but rather on the conjunction.

4.2. A first view of the data

Before running regressions, we feel it is informative to present a picture of what our data looks like. Eq. (3) is specified as a multivariate regression, so is impossible to plot. However, if we constrain the coefficient on SHARE to equal unity, then simple algebraic manipulation enables us to rewrite Eq. (3) as

$$\gamma_{ngc} - \gamma_{ngROW} = \frac{\alpha_{ng}}{X_{nc}} + \beta_2(\tilde{\delta}_{ngc} - \tilde{\delta}_{ngROW}) + \tilde{\epsilon}_{ngc}$$

If we plot the left-hand side of this equation against the term in parentheses, we can obtain an approximate idea of how production distortions move with demand distortions. What should we expect to see? In a frictionless comparative advantage world, one would expect the two variables to be uncorrelated. Frictions in a comparative advantage world would produce a positive correlation, but the slope of the line would be less than unity. Only in a world of home market effects should one see a positive correlation with a slope greater than unity.

We plot these in Fig. 1 for the four-digit sectors. The data clearly seems to be arrayed along a line with a slope that is greater than unity. Indeed, the fitted line has a slope of 1.8, indicating that demand deviations typically produce more than proportional production deviations.

4.3. Pooled tests for the home market effect

A more precise view of this relation comes from estimating Eq. (4) under a variety of specifications. The results from these pooled regressions appear in Table 1. The most striking fact is that the coefficient on IDIODEM exceeds unity in all

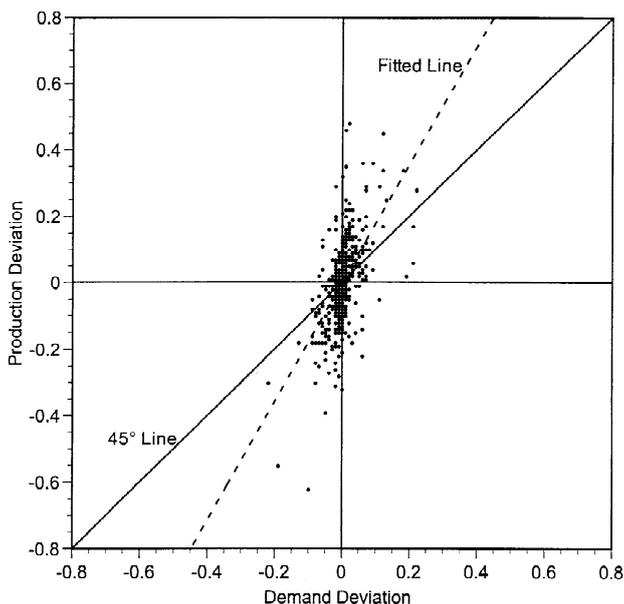


Fig. 1. The Idiosyncratic Demand Deviation is the share of 4-digit absorption in 3-digit absorption less that level for the rest of the world, i.e. $\delta_{ngc} - \delta_{ngROW}$. The Production Deviation is the share of 4-digit output in a 3-digit industry less that level for the rest of the world, i.e. $\gamma_{ngc} - \gamma_{ngROW}$. These variables indicate how different absorption and production are from the rest of the world.

specifications. This indicates that on average there is a strong home market effect. In the typical OECD industry, if the derived demand deviation rises by 1 per cent, then output rises by 1.6 per cent. What is quite striking is that we obtain this result

Table 1
Pooled runs (Dependent variable is 4-digit output; standard errors below estimates)

	1	2	3	4
IDIODEM	1.67	1.67	1.57	1.57
	0.05	0.05	0.10	0.10
SHARE	0.96	0.92		
	0.01	0.02		
EXPORTD		0.07		0.01
		0.02		0.04
FACTORS	No	No	Yes	Yes
Observations	650	650	650	650

IDIODEM is idiosyncratic demand, SHARE is the share of 4-digit output in 3-digit output in the rest of the world, EXPORTD is a dummy variable that is one if the country is a net exporter of the good, and FACTORS indicates whether the coefficients on factor endowments were allowed to differ across 4-digit sectors. No indicates that the coefficients on factor endowments were constrained to be the same for every 3-digit sector; Yes means they varied by 4-digit sector.

on the same data set used by Davis and Weinstein (1996). The crucial difference is that the relevant idiosyncratic demand now accounts for the real geography of the OECD economies.

A final econometric issue that we must address is simultaneity. Is idiosyncratic demand, as we posit, leading to a strong production response? Alternatively, is a level of production beyond that our model explains drawing in its wake idiosyncratic demand, creating only the appearance of home market effects? The ideal solution to this problem would be to find good instruments correlated with idiosyncratic demand, but not with output. Unfortunately we know of no such good instruments. Hence we cannot formally rule out the possibility that simultaneity influences our results. We can, though, take some steps to minimize its potential influence. Moreover we can give some reasons, based on the conjunction of our studies, to suggest that this is very likely not an appealing interpretation of our results.

First, we construct the demand variable based on an average of demands in the countries *ten to fifteen years prior* to the estimation period. This removes simultaneity arising from contemporaneous correlations. Second, while we cannot instrument for IDIODEM, we can control for some of the potential price effects in the regression. In columns 2 and 4 we include a variable EXPORTD in our specification. EXPORTD is a dummy variable that equals one if the country is a net exporter of that commodity times the (instrumented) three digit output in that sector. EXPORTD controls for the fact that countries that are net exporters tend to have lower prices than countries that are net importers. As one can see the coefficient on EXPORTD is positive as one should expect, but it hardly affects the overall magnitude or significance of the coefficient on IDIODEM. The absence of a strong impact of controlling for whether the country is a net exporter or not makes it less likely that price movements associated with being a net exporter or importer of a commodity are driving our results.

Finally, we need to think more closely about whether it is attractive to interpret our results as arising from simultaneity. The story would need to go something as follows: While our model does a good job of predicting the pattern of production, it is surely less than perfect. Indeed, there could be some systematic influences on the pattern of production left out, as for example Ricardian technical differences across countries. Hence a country or region may have a high level of production of a good for reasons outside the model. In turn, this unusually high production may suggest lower prices for the associated good, so lead idiosyncratic demand to respond to the production in a less than one-for-one manner. Thus the argument would be that by reversing the direction of true causality, we find home market effects of production responding more than one-for-one with idiosyncratic demand.

The issue is whether this interpretation is attractive in light of the various investigations we have pursued of the home market effect. It is straightforward to show that under the hypothesis that production patterns are driven by comparative

advantage, plausible assumptions lead one to conclude that the potential upward simultaneity bias in β_2 would diminish in the present paper relative to Davis and Weinstein (1996) because output is likely to have a much smaller effect on derived demand than on local demand.⁴ Since the estimated coefficient in that paper was 0.3, this alone would suggest that simultaneity is not the likely cause of our finding of home market effects.

4.4. *The home market effect in industry runs*

Having examined this by pooling all four-digit observations, we now move to the opposite extreme, considering each four-digit sector on its own. The results appear in Table 2. Because there are very few degrees of freedom, it is quite difficult to obtain statistical significance in these equations. Even so, we find that half of the sectors have coefficients on IDIODEM that are larger than unity and of these eleven are significantly greater than unity. By comparison, Davis and Weinstein (1996) only found half as many coefficients larger than unity and hardly any that are significant. This suggests that in our data, some industries are constant and others increasing returns to scale.⁵ Home market effects are very much in evidence.

One way to increase the number of degrees of freedom relative to the four-digit runs is to conduct industry-pooled estimation. This pools the four-digit observations within each three-digit industry, but allows the coefficient on IDIODEM to vary across three-digit industries. Relative to the fully-pooled runs, this allows us to relax the assumption that three-digit industries must either all be comparative advantage or all exhibit increasing returns. The results are presented in Table 3. In similar runs, Davis and Weinstein (1996) found that less than one-fifth of all sectors had point estimates above unity. Here, using our new measure of market access, we now find that over half of the industries exhibit home market effects. Furthermore, while the earlier study found that none of the point estimates were significantly larger than unity, we now find that four of our coefficients have this property. Moreover, while Davis and Weinstein (1996) rejected home market

⁴The story would specify an additional relation between idiosyncratic demand and production as follows: $\text{IDIODEM}_{ngc} = \omega X_{ngc} + \eta_{ngc}$. For $\omega \in (0, 1)$, it is straightforward to show that a sufficient condition for the degree of bias to be increasing in ω is that $\beta_2 < 1$, i.e. that we are in a world of comparative advantage. The final step would be to note that the relevant ω is likely to be lower in the present work than in Davis and Weinstein (1996), since local demand is plausibly more strongly related to local production than is a weighted average of local and rest-of-world demand.

⁵More subtle problems arise if individual industries themselves are composed of both IRS and CRS goods. In alternative frameworks, Krugman (1980), Krugman and Venables (1995) and Davis (1998) address this problem. The various contributions stress the potential role of absolute market size and the cross-good structure of trade costs in determining industrial structure. This remains an important direction for further empirical study.

Table 2

Four-digit runs (Dependent variable is 4-digit output; all regressions include factor endowments; standard errors below estimates; number of observations = 13)

Industry	Adj. R^2	IDIODEM	Industry	Adj. R^2	IDIODEM
Slaughtering, preparing and preserving meat	0.68	3.45 1.77	Cordage, rope and twine industries	0.95	1.68 0.22
Dairy products	0.17	0.20 2.48	Textiles nec	0.84	2.71 0.49
Canning and preserving of fruits and vegetables	0.91	12.08 3.46	Tanneries and leather finishing	0.92	5.87 0.63
Canning, preserving and processing of fish, crustacea and similar foods	0.96	2.63 0.30	Products of leather and leather substitutes, except footwear and wearing apparel	0.81	-0.15 0.50
Vegetable and animal oils and fats	0.68	-0.67 2.44	Sawmills, planing and other wood mills	0.52	-4.74 8.08
Grain mill products	0.50	-5.66 5.92	Wooden and cane containers and small cane ware	0.70	-0.10 0.66
Bakery products	0.69	2.78 1.63	Wood and cork products nec	0.97	2.48 0.25
Cocoa, chocolate and sugar confectionery	0.84	-1.62 1.67	Pulp, paper and paperboard	0.36	15.62 10.17
Distilling, rectifying and blending spirits	0.68	1.84 0.72	Containers and boxes of paper and paperboard	0.89	2.88 0.74
Malt liquors and malt	0.64	-0.88 1.18	Pulp, paper and paperboard articles nec	0.84	-2.11 0.89
Spinning, weaving and finishing textiles	0.95	-10.92 2.03	Basic industrial chemicals except fertilizer	0.89	5.02 2.30
Made-up textile goods except wearing apparel	0.90	3.36 1.06	Fertilizers and pesticides	0.61	0.74 1.03
Knitting mills	0.71	-1.08 11.42	Synthetic resins, plastic materials and man-made fibres except glass	0.78 2.20	-1.35
Carpets and rugs	0.40	0.34 4.66	Paints, varnishes and lacquers	0.91	-0.96 1.02
Drugs and medicines	0.80	-0.43 1.71	Electrical industrial machinery and apparatus	0.96	-1.88 0.57
Soap and cleaning preparations, perfumes, cosmetics and other toilet preps	0.90	12.95 4.35	Radio, television and communication equipment and apparatus	0.97	13.45 1.54
Chemical products nec	0.77	2.52 1.71	Electrical appliances and housewares	0.88	2.04 0.87
Structural clay products	0.71	-0.27 0.45	Electrical apparatus and supplies nec	0.70	-0.83 1.53
Cement, lime and plaster	0.52	-1.06 0.90	Shipbuilding and repairing	0.71	0.44 0.95
Non-metallic mineral products nec	0.96	2.45 0.43	Motor vehicles	0.9	4.68 3.74
Cutlery, hand tools and general hardware	0.83	0.25 0.81			

Table 2. Continued

Industry	Adj. R^2	IDIODEM	Industry	Adj. R^2	IDIODEM
Furniture and fixtures primarily of metal	0.90	1.50			0.82
Structural metal products	0.71	1.15			1.48
Fabricated metal products except machinery and equipment nec	0.92	4.03			1.12
Engines and turbines	0.94	-0.63			0.44
Agriculture machinery and equipment	0.69	1.82			0.96
Metal and wood working machinery	0.93	1.23			1.25
Special industrial machinery and equipment except metal and wood working machinery	0.88	0.13			2.05
Office, computing and accounting machinery	0.77	-7.49			4.36
Machinery and equipment, except electrical nec	0.85	3.81			1.48

effects in two-thirds of the three-digit sectors, we now reject economic geography only in two sectors, other chemicals and non-electrical machinery.

One word of caution is in order. Looking at the sectors, it is somewhat disappointing that sectors like electrical machinery and transportation equipment do not have point estimates that exceed unity. A likely explanation is imprecision of the estimates. In both of these sectors, the standard errors are so large that we cannot reject home market effects. Indeed the four-digit runs presented in Table 2 indicate that in half the sectors within these industries (radio, television and communication equipment, electrical appliances and housewares, and motor vehicles), we do obtain point estimates for IDIODEM that exceed one.

Hence we conclude that, these problems notwithstanding, relative to previous work these results do represent a striking degree of support for the economic geography paradigm. Most sectors exhibit home market effects. Those that do not exhibit such effects typically have point estimates that are measured imprecisely.

It is useful to compare these results to those in our companion study on Japanese regional data. There we also found significant home market effects. Unfortunately, it is difficult to match our new results with those of Davis and Weinstein (1999) because that paper used Japanese data at a different level of aggregation. However, if we aggregate the data so that we assume industries are defined at the two-digit ISIC and goods at the three-digit ISIC, then we have a roughly comparable level of aggregation.

There are several issues to bear in mind about increasing the level of

Table 3

Industry-pooled estimation (Dependent variable is 4-digit output; coefficients on factor endowments vary at the 4-digit level; standard errors below estimates)

Industry	IDIODEM	Obs.
Food products	2.51 0.28	104
Beverage industries	1.11 0.61	26
Textiles	1.79 0.20	78
Leather	2.17 0.39	26
Wood products	2.16 0.23	39
Paper and pulp	0.89 0.57	39
Industrial chems.	1.02 0.86	39
Other chemicals	0.28 0.77	52
Other non-metallic mineral product	0.91 0.30	39
Fabricated metals	1.49 0.48	52
Machinery, except electric	0.11 0.36	78
Electrical mach.	0.42 0.44	52
Transportation equip.	0.69 0.92	26

aggregation. Because more countries report three-digit production data than four-digit, we have more degrees of freedom than on the four-digit runs. But the higher level of aggregation means that we increase the chance that we are pooling sectors that differ in many respects, including factor intensity. This may interfere with the operation of home market effects. For example, while it is plausible that high demand for motor vehicles might cause specialization in motor vehicles as opposed to motorcycles, it is less plausible that countries with high demand for transport equipment are less likely to produce precision instruments. On Japanese data, where we had compatible technology matrices, we could circumvent this problem by aggregating according to technological similarity, but on international data, this is not possible. Furthermore, we are faced with the problem that the variance in demand deviations shrinks at higher levels of aggregation. When we move from four- to three-digit data, the variance in our demand deviation variable falls by a factor of two for the countries for which we have comparable numbers. By comparison, Japanese regions had a demand deviation variance that was

comparable to international four-digit data. Finally the inclusion of countries like Turkey and Yugoslavia in the three-digit sample probably exacerbates problems such as measurement error.

These reasons may help explain why Davis and Weinstein (1996) found a smaller impact of demand deviations on production deviations on more aggregated data. Nevertheless, since we did find evidence of home market effects at a higher level of aggregation on Japanese data (albeit with more than twice the number of degrees of freedom), it may be useful to compare those results with our international results at a higher level of aggregation.

We present the results from goods-level estimation at the three-digit level in Table 4. Although only one sector, textiles, exhibits a coefficient on IDIODEM that is significantly larger than unity, 9 out of our 26 sectors have point estimates in excess of unity. By comparison, in Davis and Weinstein (1999), 9 out of 19 sectors had point estimates larger than unity and 8 out of these 9 were significant. No doubt many of the reasons that we have highlighted above explain the relative imprecision of our international results. Even so, there is a fair amount of overlap between the two sets of results. If we restrict attention to the 14 sectors that appear in both the international and regional data sets, we find that seven have coefficients on IDIODEM that are significantly larger than unity in the Japanese data and five have point estimates larger than unity in the international data. Interestingly, four of the five international sectors that come up as having home market effects—textiles, iron and steel, transportation equipment, and precision instruments—are among the seven sectors that also have measurable home market effects in the Japanese data. Although the large standard errors in these industry runs make it difficult to make strong statements, there is a striking degree of overlap. Furthermore, the fact that these sectors have often been presented as canonical examples of economic geography by Krugman (1991) and others bolsters the plausibility of our point estimates.

Returning to our individual four-digit sector runs, we next examine the issue of economic significance. Here we consider β -coefficients, which indicate how much a one standard deviation movement in the independent variable moves the dependent variable. Over all, our estimates for the pooled specification indicate that a one-standard-deviation movement in idiosyncratic demand moves production by about 0.15 standard deviations. While quite modest, it is still three times larger than the estimate in Davis and Weinstein (1996). However, since we are probably dealing with a mix of sectors, only some of which are monopolistically competitive, it makes sense to calculate these coefficients on a sector-by-sector basis.

If one inspects the β -coefficients (reported in Davis and Weinstein, 1998) in many sectors the home market effect is extremely important. For example in electrical machinery sectors, we obtain β -coefficients that are typically in the 0.9 range—indicating that the absorption linkage to production is very important. Overall, in the sectors where we detect coefficients on idiosyncratic demand that are larger than unity, β -coefficients are typically around 0.5.

Table 4

Three-digit runs (Dependent variable is 3-digit output; coefficients on factor endowments vary at the 3-digit level; standard errors below estimates; number of observations is 22)

Industry	Adj. R^2	IDIODEM	Industry	Adj. R^2	IDIODEM
Food products	0.72	18.28 17.41	Rubber products	0.82	-1.03 1.01
Beverage industries	0.70	0.15 0.45	Plastic products nec	0.91	1.32 0.20
Tobacco manufactures	0.69	0.81 0.29	Pottery, china and earthenware	0.64	3.05 0.86
Textiles	0.83	62.64 20.35	Glass and glass products	0.84	0.71 1.11
Wearing apparel, except	0.85	-0.53 0.18	Other non-metallic mineral products	0.74	1.61 0.72
Leather and products of leather, leather substitutes and fur, except footwear and wearing apparel	0.20	-0.32 0.80	Iron and steel basic industries	0.81	3.43 1.98
Footwear, except vulcanized or molded rubber or plastic footwear	-0.03	-0.12 0.44	Non-ferrous metal basic industries	0.86	-0.09 1.35
Wood and wood and cork products, except furniture	0.69	0.70 0.40	Fabricated metal products, except machinery and equipment	0.84	-0.33 0.78
Furniture and fixtures, except primarily of metal	0.65	0.56 0.90	Machinery except electrical	0.92	-5.40 2.48
Paper and paper products	0.59	12.94 10.35	Transport equipment	0.91	1.42 1.25
Industrial chemicals	0.91	-0.61 0.45	Professional and scientific and measuring and controlling equipment nec, and of photographic and optical goods	0.80	2.95 2.15
Other chemical products	0.88	0.71 1.14			
Petroleum refineries	0.82	-1.28 0.40			

A second way to obtain a sense of how important economic geography is to OECD production is to examine the relative sizes of the sectors for which β_2 is larger than unity. At the four-digit level, of the 50 sectors for which we have data, the sectors with coefficients on IDIODEM exceeding unity account for 64 per cent of the total output. Repeating this exercise for the three-digit sectors, where we have 22 countries and all manufacturing output for each country, reveals that 50 per cent of all manufacturing production is governed by economic geography. This indicates that the sectors that appear to have home market effects account for a majority of manufacturing output.

5. Conclusion

This paper has examined data for a set of OECD countries to investigate the existence of home market effects from idiosyncratic demand on the pattern of production. We developed a framework that nests a conventional Heckscher–Ohlin framework (based on comparative advantage) with a model of economic geography. Within the context of these models, the simple comparative advantage model would not predict home market effects, while that of economic geography would predict such effects.

The results provide support for the economic geography hypothesis of the existence of home market effects. Within the context of the three models considered in this study, they also provide important evidence on the role and importance of increasing returns in determining production structure for the OECD. A parallel investigation of 40 Japanese regions in Davis and Weinstein (1999) also finds such home market effects.

The broad picture that emerges draws on insights from Helpman (1981) and Krugman (1980). Within the context of the models considered, comparative advantage matters both in affecting the broad and fine industrial structure. Even at the four-digit level, from one-third to one-half of OECD manufacturing output seem to be governed by simple comparative advantage. However increasing returns also play a vital role, in the particular form known as economic geography. These have measurable effects on production structure for as much as one-half to two-thirds of OECD manufacturing output. Finally, we saw that the key to identifying these effects is to introduce more geographical realism into our models of production and trade.

Acknowledgements

We are grateful to Tony Venables for very helpful suggestions.

References

- Bernstein, J., Weinstein, D.E., 2002. Do endowments predict the location of production? *Journal of International Economics* 56 (1), 55–76.
- Bhagwati, J.N., 1982. Shifting comparative advantage, protectionist demands, and policy response. In: Bhagwati, J.N. (Ed.), *Import Competition and Response*. University of Chicago.
- Chipman, J.S., 1992. Intra-industry trade, factor proportions and aggregation. In: Rader, T.J. (Ed.), *Economic Theory and International Trade: Essays in Memoriam*. Springer-Verlag, New York.
- Davis, D.R., 1998. The home market, trade, and industrial structure. *American Economic Review* December.
- Davis, D.R., Weinstein D.E., 1996. Does economic geography matter for international specialization? Mimeo, National Bureau of Economic Research Working Paper # 5706, August.
- Davis, D.R., Weinstein, D.E., 1998. Market access, economic geography and comparative advantage: an empirical assessment. National Bureau of Economic Research Working Paper # 6787, November.

- Davis, D.R., Weinstein, D.E., 1999. Economic geography and regional production structure: an empirical investigation. *European Economic Review* February.
- Dixit, A.K., Stiglitz, J.E., 1977. Monopolistic competition and optimum product diversity. *American Economic Review* 67 (3), 297–308.
- Ethier, W.J., 1984. Higher dimensional issues in trade theory. In: Jones, R.W., Kenen, P.B. (Eds.). *Handbook of International Economics*, Vol. 1. North-Holland, New York.
- Feenstra, R.C., 1982. Appendix. Product creation and trade patterns: a theoretical note on the 'biological' model of trade in similar products. In: Bhagwati, J.N. (Ed.), *Import Competition and Response*. University of Chicago.
- Feenstra, R.C., 1996. U.S. imports, 1972–1994: data and concordances. National Bureau of Economic Research #5515, and National Bureau of Economic Research Trade Database, Disk 1.
- Feenstra, R.C., Markusen, J.R., Rose, A.K., 2001. Using the gravity equation to differentiate among alternative theories of trade. *Canadian Journal of Economics* 34 (2), 430–447.
- Harrigan, J., 1993. OECD imports and trade barriers in 1983. *Journal of International Economics*.
- Harrigan, J., 1995. Factor endowments and the international location of production: econometric evidence from the OECD, 1970–1985. *Journal of International Economics*.
- Harrigan, J., 1997. Technology, factor supplies and international specialization: testing the neoclassical model. *American Economic Review* 87, 475–494.
- Head, K., Ries, J., 2001. Increasing returns versus national product differentiation as an explanation for the pattern of US–Canada trade. *American Economic Review* 91 (4), 858–879.
- Helpman, E., 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).
- Krugman, P.R., 1979. Increasing returns, monopolistic competition, and international trade. *Journal of International Economics* 9 (4), 469–479.
- Krugman, P.R., 1980. Scale economies, product differentiation, and the pattern of trade. *American Economic Review* 70, 950–959.
- Krugman, P.R., 1991. *Geography and Trade*. Massachusetts Institute of Technology, Cambridge.
- Krugman, P.R., Venables, A.J., 1995. Globalization and the inequality of nations. *Quarterly Journal of Economics* 110 (4), 857–880.
- Lancaster, K., 1980. Intra-industry trade under perfect monopolistic competition. *Journal of International Economics* 10 (2), 151–175.
- Leamer, E., 1984. *Sources of International Comparative Advantage*. Massachusetts Institute of Technology, Cambridge.
- Leamer, E., 1997. Access to western markets and eastern effort. In: Zecchini, S. (Ed.), *Lessons from the Economic Transition, Central and Eastern Europe in the 1990s*. Kluwer Academic Publishers, Dordrecht, pp. 503–526.
- Maskus, K.E., 1991. Comparing international trade data and product and national characteristics data for the analysis of trade models. *International Economic Transactions: Issues in Measurement and Empirical Research*, 17–56.
- McCallum, J., 1995. National borders matter: Canada–US regional trade patterns. *American Economic Review* 85 (3).
- Trionfetti, F., 2001a. Public procurement, economic integration, and income inequality. *The Review of International Economics* 9 (1), 29–41.
- Trionfetti, F., 2001b. Using home-biased demand to test for trade theories. *Weltwirtschaftliches Archiv* 137 (3), 404–426.
- Weder, R., 1995. Linking absolute and comparative advantage to intra-industry trade theory. *Review of International Economics* 3 (3).
- Wei, S.J., 1996. Intra-national versus international trade: how stubborn are nations in global integration? National Bureau of Economic Research Working Paper #5531.