Henry G. Overman, Stephen Redding and Anthony Venables

The economic geography of trade, production and income : a survey of empirics

Originally published as:
Choi, E. Kwan and Harrigan, James (Eds), *Handbook of international trade.* Blackwell handbooks in economics. Malden, USA : Blackwell Publishing, 2003, pp. 353-387 © 2003 Blackwell Publishing

Abstract

This paper surveys the empirical literature on the economic geography of trade flows, factor prices, and the location of production. The discussion is structured around the empirical predictions of a canonical theoretical model. We review empirical evidence on the determinants of trade costs and the effects of these costs on trade flows. Geography is a major determinant of factor prices, and access to foreign markets alone is shown to explain some 35% of the cross-country variation in per capita income. The paper documents empirical findings of home market (or magnification) effects, suggesting that imperfectly competitive industries are drawn more than proportionately to locations with good market access. Sub-national evidence establishes the presence of industrial clustering, and we examine the roles played by product market linkages to customer and supplier firms, knowledge spillovers, and labour market externalities.

This paper was produced as part of the Centre's
Globalisation Programme

The Economic Geography of Trade, Production, and Income: A Survey of Empirics

Henry G. Overman, Stephen Redding and Anthony J. Venables

September 2001

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

ISBN 0 7530 1496 3

Individual copy price: £5

The Economic Geography of Trade, Production, and Income: A Survey of Empirics

Henry G. Overman, Stephen Redding and Anthony J. Venables

The Centre for Economic Performance is financed by the Economic and Social Research Council

Acknowledgements

This work is supported by the UK ESRC funded Centre for Economic Performance at the London School of Economics. We are grateful to James Anderson, James Harrigan, James Markusen, and seminar participants at the NBER Spring Meetings, for helpful comments.

The authors contact details are:

H.G. Overman	S.J. Redding	A.J. Venables
Dept of Geography	Dept of Economics	Dept of Economics
London School of Economics	London School of Economics	London School of
Economics		
Houghton Street	Houghton Street	Houghton Street
London WC2A 2AE, UK	London WC2A 2AE, UK	London WC2A 2AE, UK
h.g.overman@lse.ac.uk	s.j.redding@lse.ac.uk	a.j.venables@lse.ac.uk

1. Introduction

Both first- and second-nature geography are major determinants of production structure, trade and income. First-nature is the physical geography of coasts, mountains, and endowments of natural resources, and second-nature is the geography of distance between economic agents. Elements of first-nature are the subject matter of factor endowment based trade theory, and our focus in this paper is largely -- although not entirely -- on second-nature. We shall ask: how does the spatial relationship between economic agents determine how they interact, what they do, and how well off they are?

How does geography shape interactions between economic agents? Distance directly increases transactions costs because of the transport costs of shipping goods, the time cost of shipping date sensitive products, the costs of contracting at a distance, and the costs of acquiring information about remote economies. The familiar gravity model indicates how rapidly distance reduces the volume of trade between countries.

Geography also shapes the activities undertaken in each country, as profits depend on proximity to linked activities. Thus, in addition to taking place where there are factor supplies, production will locate close to markets and to suppliers of intermediate goods. These obvious sounding statements immediately raise several questions. How are proximity to markets and to suppliers to be measured? To be operational we have to able to make statements that one country has better market-access or better supplier-access than another.[1] And having measured these geographical characteristics of countries, which industries are most influenced by them? All activities would, other things being equal, locate in countries with good market-access and supplier-access, but in equilibrium other things are not equal. Prices of immobile factors adjust so that some activities locate in central countries and others go to more remote locations, but which activities go where? It depends on industry characteristics including the cost of transporting final output and the share of intermediate goods and services in costs. Also important is the extent to which it is possible for firms to divide production and operate in many locations. If production in all activities is perfectly divisible then economic geography effects are likely to be small.[2] But if firms have to make 'either-or' choices and produce in only a subset of locations, then the effects will generally be larger. Thus, if there are industries with increasing returns at the plant level there will generally be 'home market effects', leading these industries to be disproportionately represented in countries with good market-access.

Much of the interest of economic geography derives from the fact that the location of demand (determining market-access) and input supply (determining supplier-access) is not exogenous. From the theory standpoint this generates the possibility of 'cumulative causation', agglomeration, and multiple equilibria; locations have one activity only because they have another, and vice-versa. From the empirical standpoint it raises several questions. Is there evidence that industries are more agglomerated than would be suggested by the location of factor endowments or by chance? What sorts of industries -- or what functional activities -- tend to agglomerate? What are the sources of agglomeration: linkages to customer and supplier firms, technological externalities, or effects arising in factor markets? More fundamentally, how are the endogeneity issues associated with co-location of industries to be handled? And how should econometrics proceed if theory suggests that there is not a unique mapping from exogenous variables to endogenous ones?

As well as influencing trade flows and production structure, geography is also one of the determinants of how well-off people are. How disadvantaged are remote countries, and how much of the cross-country income distribution can be explained by geography? Spatial variations in goods prices will lead to spatial variation in factor prices, as predicted by the Stolper-Samuelson effects of traditional trade theory. Real returns to all factors may be low in remote locations, as the value added that firms can pay to immobile factors is squeezed by transport costs reducing export receipts and raising the costs of imported inputs. Where value added is only a small fraction of total costs it is possible that quite modest transport costs translate into large reductions in value added attributable to immobile factors.

The impact of geography on income levels may come not just through the mechanism of goods prices and transport costs, but also through spatial differences in institutions and in technology. For example, productivity may depend on the spatial density of economic activity, and technology transfer may depend on distance from technology producers. Empirical work has found such effects, although we argue that true productivity differences are very difficult to disentangle from price effects.

These three sets of issues -- geography and trade flows, geography and income, and geography and the location of activity -- are the subject matter of sections 3 - 5 of this review. The next section provides some of the theoretical structure that will be used at various stages.

2. A Canonical Model

In this section we outline some key elements of a canonical model that we draw on at various stages in the paper. The oldest model in which the effects of economic geography on the structure of production and incomes is shown is that of von-Thunen (1826), and this can easily be set in an international context (see eg Venables and Limao 1999). The disadvantage of this model is that outlying regions trade with a single central location. To capture a full structure of bilateral trade flows in a tractable way we need a model that has product differentiation in at least some sectors, this possibly -- although not necessarily -- combined with monopolistic competition.

The model we use contains some number of countries (or more generally 'locations') and a number of industries. Country specific variables are sub-scripted and industries represented by super-scripts. Thus, x_{ij}^k is the quantity of an industry *k* good produced in country *i* and sold in country *j*. Underlying the demand side of the model is a price index (or expenditure function) for each industry that aggregates different varieties in the industry. This takes a CES form, is denoted G_j^k and defined by

$$G_j^k = \left[\sum_i n_i^k \left(p_i^k t_{ij}^k\right)^{1-\sigma^k}\right]^{1/(1-\sigma^k)} \qquad (1)$$

In this equation n_i^k is the number of varieties of industry *k* products produced in country *i*, p_i^k their fob prices, and t_{ij}^k the iceberg cost factor on trading industry *k* products from country *i* to country *j*. σ^k is the elasticity of substitution between varieties, and sectors in which $\sigma^k \rightarrow \infty$ produce homogeneous products.

If E_j^k is the total expenditure on industry *k* products in country *j*, then the sales of a single industry *k* product produced in country *i* and sold in *j* are given by

$$x_{ij}^k = \left(p_i^k\right)^{-\sigma^k}\left(t_{ij}^k\right)^{1-\sigma^k} E_j^k \left(G_j^k\right)^{\sigma^k - 1} \qquad (2)$$

The relationship is derived by using Shepard's lemma on the price index (see for example Dixit and Stiglitz 1977). It contains information about bilateral trade flows between each pair of countries, *i* and *j*, and we use it for assessing the impact of geography on these flows. Adding over all markets and over all n_i^k varieties of industry *k* products produced in country *i*, we derive the following expression for the total value of industry *k* output produced by country *i*, y_i^k

$$y_i^k \equiv n_i^k p_i^k x_i^k = n_i^k \left(p_i^k\right)^{1-\sigma^k} \sum_j \left(t_{ij}^k\right)^{1-\sigma^k} E_j^k \left(G_j^k\right)^{\sigma^k - 1} \qquad (3)$$

where $x_i^k \equiv \sum_j x_{ij}^k$.

On the production side, prices are set proportional to marginal costs, according to

$$p_i^k = \theta^k c^k(w_i, G_i) \qquad (4)$$

where θ^k equals unity in perfectly competitive industries, and is greater than unity if firms mark up price over marginal cost. $c^k(w_i, G_i)$ is marginal cost, and is a function of prices of primary factors in country *i*, w_i, and price indices of intermediates, G_i.[3] If there is more than one primary factor or intermediate input these are vectors, so intermediate prices are given by the vector of industry price indices, $G_i = G_i^1 \ldots G_i^k$.

Some sectors of the economy are perfectly competitive, and in these sectors the numbers of varieties produced in each country, n_i^k, are exogenously determined – an 'Armington' assumption. Other sectors are monopolistically competitive, and the numbers are determined by zero profit conditions. Given that prices are proportional to marginal costs and assuming further that cost functions in these sectors have increasing returns and are homothetic, zero profits are made if firms' sales reach a given level, $\bar{x}^k$. Firms in monopolistically competitive industry *k* therefore make zero profits if their sales satisfy (using (4) in (2) with the definition of x_i^k),

$$\left[\theta^k c^k(w_i, G_i^k)\right]^{-\sigma^k} \sum_j \left(t_{ij}^k\right)^{1-\sigma^k} E_j^k \left(G_j^k\right)^{\sigma^k - 1} = x_i^k = \bar{x}^k. \qquad (5)$$

The other main relationships in the model are factor market clearing and the determination of expenditure. Factor market clearing is

$$w_i L_i = \sum_k y_i^k \left(\frac{\partial c^k(w_i, G_i)}{\partial w_i} . \frac{w_i}{c^k(w_i, G_i)} \right) \qquad (6)$$

where L_i is the endowment (or vector of endowments), and the expression is written in value form. The term in large brackets is the share of the primary factor in marginal costs (equal, with a homothetic cost function, to the share in average costs). Expenditure on each industry in each country is

$$E_i^k = f_i^k + \sum_\ell y_i^\ell \left(\frac{\partial c^\ell(w_i, G_i)}{\partial G_i^k} . \frac{G_i^k}{c^\ell(w_i, G_i)} \right) \qquad (7)$$

where the first term, f_i^k, is final expenditure (itself depending on income and prices), and the second is derived demand, so the term in large brackets is the share of intermediates from industry k in industry ℓ costs.

The sets of equations (1) - (7) characterise the international general equilibrium, and can be solved for quantities (x_{ij}^k, n_i^k, y_i^k) and for prices and expenditures (p_i^k, w_i^k, G_i^k, E_i^k). What are the properties of the model, and what hypotheses does it generate? We outline the answer here in very general terms, and are more specific in the following sections of the paper.

The first broad property is that geography matters for factor prices and for the structure of production in each country. Geography enters the model through the trade costs, t_{ij}^k, which vary systematically with distance and other geographical forces, and also vary across industries. Trade costs are composed of a package of transport costs, time costs, and information costs, and section 3 presents evidence on the size of these costs. Trade costs prevent goods price equalisation from occurring, and hence also prevent factor price equalisation. Since they vary across both locations and industries they provide a basis for comparative advantage.

The geographical structure of trade costs mean that some locations will be attractive to industry because of good market-access, and also because of good intermediate supplier-access. How does this show up in equilibrium? One manifestation will be through spatial variations in the prices of immobile factors, which will be bid up in regions with good market- and supplier-access. This will be the subject of Section 4 of the paper. Another manifestation is in the structure of production. Some types of industry will be particularly drawn to these locations, and in Section 5 we show how this can be combined with factor endowment theory to give hypotheses about industrial structure.

The fact that large markets are profitable locations and tend to have high factor prices creates a potential positive feedback, as large markets attract firms and mobile factors, so becoming still larger. As a consequence the model may have multiple equilibria, some unstable, and others manifesting agglomeration. For example, in Krugman (1991a) and (1991b) there are two sectors, one monopolistically competitive and the other perfectly competitive and freely traded. Production uses sector specific factors (and no intermediate goods), the factor used in the monopolistically competitive industry being perfectly mobile between locations. Krugman shows how an increase in the amount of manufacturing in one location increases income and the size of the market and reduces the price index. If trade costs are low enough then this location attracts the mobile factor and leads to the agglomeration of all of manufacturing in one location. Krugman and Venables (1995) have the same two sectors and a single immobile factor. However, the presence of intermediate goods (manufacturing uses manufacturing as an input) creates agglomeration, as firms gain from being close to customer and supplier firms (see also Venables 1996, 1999). Theoretical analysis of these models is synthesised in Fujita, Krugman and Venables (1999), but there is as yet little empirical investigation of clustering at the international level.[4] This survey reviews some sub-national studies (Section 5.3) that have attempted to identify agglomeration effects.

While the full model outlined above endogenises all the main variables, the empirical studies we review below typically focus on a few key relationships, holding other variables exogenous. Thus, in Section 3 we look at trade costs and examine trade flows based on equation (2), while holding all other variables exogenous. In Section 4 we look at factor prices and incomes, and much of this is based on equation (5) with equation (1), which give values of factor prices and price indices, conditional on values of expenditure and numbers of firms in each location. Section 5 turns to equation (3), giving the structure of production of each location. We discuss measurement issues, descriptive studies, and attempts to econometrically estimate (3). A number of studies look at the relationship between expenditure and production, searching for home market effects, and we review one study that endogenises input prices and derived demands, estimating (3) with (1), (6) and (7).

3. Trade Costs and Trade Volumes

The dependence of trade volumes on geography is well known through the widespread use of gravity models. In this section we start by investigating the trade costs -- the t_{ij}^{k} -- that underly the gravity relationship, and then turn to the relationship between trade and geography.

Trade Costs

Trade costs have many different elements, some observable (such as transport costs), while others, such as costs of acquiring information, are much more difficult to observe directly although inferences can be made from trade flows.

There are three main sources of data for transport costs between countries. The most readily available are the bilateral cif/fob ratios produced by the IMF by matching export data (reported by countries fob) and import data (reported cif). However, problems with this data include the fact that it is an aggregate over all commodities so depends on the composition of trade, and that a high proportion of observations are imputed (see Hummels 1999b for discussion). The second source is national customs data, made available by a few countries in a form that allows extraction of very detailed information. For example, the US Census Bureau make available data on US imports at the 10 digit level by exporter country, mode of transport, district of entry, and valued both inclusive and exclusive of freight and insurance changes (see Hummels 1999a).

The third source is direct industry or shipping company information. These include indices of ocean shipping prices and air freight rates from trade journals (Hummels 1999b), or direct quotes from shipping companies (eg Limao and Venables 2001 who obtain quotes for shipping a standard container from Baltimore to various destinations).

We learn a number of things from studies of these data. First, there is a very wide dispersion of transport costs across commodities and across countries. Thus, for the US in 1994, freight expenditure was only 3.8% of the value of imports, but equivalent numbers for Brazil and Paraguay are 7.3% and 13.3% (Hummels 1999a, from customs data). These values incorporate the fact that most trade is with countries that are close, and in goods with low transport costs. Looking at transport costs unweighted by trade volumes gives much higher numbers; thus, the median cif/fob ratio, across all country pairs for which data is available, is 1.28 (implying 28% transport and insurance costs). Looking across commodities, an unweighted average of freight rates is typically 2 to 3 times higher than the trade weighted average rate.

Estimates of the determinants of transport costs are given in Hummels (1999b) and

Limao and Venables (2001). These studies find elasticities of transport costs with respect to distance of between 0.2 and 0.3. Limao and Venables find that sharing a common border substantially reduces transport costs, and overland distance is around 7 times more expensive than sea distance. Being landlocked increases transport costs by approximately 50%. Infrastructure quality (as measured by a composite index of transport and communications networks) is important; for example, while the median cif/fob ratio is 1.28, the predicted value of this ratio for a pair of countries with infrastructure quality at the 75th percentile rises to 1.40.

Trade Volumes

Equation (2) provides the basis for a gravity trade relationship.[5] It is usually estimated on aggregate data, so

$$n_i p_i x_{ij} = n_i p_i^{1-\sigma} \left(t_{ij}\right)^{1-\sigma} E_j \left(G_j\right)^{\sigma-1} \tag{2'}$$

(derived by dropping the industry specific superscript, and multiplying by the number of varieties produced in each country and their price). The left hand side is simply the value of trade between country *i* and *j*; the main data source for this is the UN COMTRADE data base, made available by the NBER (Feenstra et al., 1997). The right hand side contains exporter country information (numbers of varieties and their prices), importer country information (expenditure and the price index), and trade cost information, t_{ij}. The exporter and importer country information is typically proxied by income in each country. However, if the focus is on the geography of trade, then these terms can simply take the form of fixed effects for exporter and importer countries.[6]

Trade costs, t_{ij}, are typically assumed to be a function of a number of geographical variables, and perhaps also cultural or political variables. We look first at the geographical ones. Distance is the most important, and the elasticity of trade volumes with respect to distance is usually estimated to be in the interval -0.9 to -1.5.[7] Sharing a common border increases trade volumes, analogous to its effect on transport costs. Country characteristics that bear on trade costs include (see Limao and Venables) being an island, which increases trade volumes somewhat, and being landlocked, which reduces trade volumes by a massive 60%. Infrastructure also matters, with predicted trade volumes between two countries with infrastructure quality at the 75th percentile 28% lower than at the median.

Gravity estimates tell us that geography matters greatly for trade volumes, although it does not reveal whether this is through the impact of geography on trade costs, or the impact of trade costs on trade volumes. Several attempts have been made to make this separation, by combining information from estimates of trade costs and trade volumes. This can be done either by taking the ratio of the distance elasticity of trade to the distance elasticity of transport costs (Hummels 1999a, Limao and Venables 2001) or by using predicted values of t_{ij} derived from the estimated transport cost equation as an independent variable in a gravity model. The latter approach gives an elasticity of trade with respect to transport costs of approximately -3, and the former a range of around -2 to -5.[8]

We have so far concentrated on transport costs, and the role of geography in determining these. However, trade costs include a wider package of transactions costs, as well as policy measures. Hummels (2000) estimates the cost of time in transit. He uses data on some 25 million observations of shipments into the US, (imports classified at the 10-digit commodity level by exporter country and district of entry to the US for 25 years), some by air and some by sea. Given data on the costs of each mode and the shipping times from different countries he is able to estimate the implicit value of time saved in shipping. The numbers are quite large. The cost of an extra day's travel is (from estimates on imports as a whole) around 0.3% of the value shipped. For manufacturing sectors, the number goes up to 0.5%. These costs are around 30 times larger than the interest charge on the value of the goods. They also carry the implication that transport costs have fallen much more through time than suggested by looking at shipping charges. The share of US imports going by air freight rose from zero to 30% between 1950 to 1998, and containerization approximately doubled the speed of ocean shipping; this gives a reduction in average shipping times of 26 days over 50 years, equivalent to a shipping cost reduction worth 12-13% of the value of goods traded.

Many studies have used a variety of further 'between country' measures in the gravity estimation in order to try and capture the role of culture, history, and politics in influencing trade flows (see Frankel 1997 for a synthesis of some of this material). A recent example is the work of Rauch and Trindade (1999), who seek to explore the role of ethnic Chinese networks in promoting trade. Their gravity estimation includes dummies for sharing a common language, having shared colonial ties, and a variable which is the product of the share of ethnic

Chinese in the populations of the importing and exporting countries. They find that colonial ties and Chinese networks have large significant effects in promoting trade (although the effects of language are mixed). Studies of this type remind us that while trade costs are important determinants of trade volumes, they are not just functions of physical geography.[9]

The Research Agenda

The studies above give some indication of the role of geography in determining transport costs and in choking off trade. There are several areas where much more work is needed.

Borders create very large trade barriers (Helliwell 1996; McCallum 1995; Wei 1996), and work is needed to understand the difference between international and inter-regional trade. One aspect of this is to recognise that there are fixed (and perhaps sunk) costs, as well as marginal costs, to firms entering new markets. Understanding the nature of these costs is important as countries seek to promote 'deep integration' to overcome international market segmentation, and also as these costs may pose important barriers to developing country export growth (Roberts and Tybout 1997).

Research is also needed to better understand the geography of information flows. Much trade involves a process of searching and matching between firms. Once a match has been made there may be monitoring and control issues (as downstream agents are concerned with the quality and delivery of supplies). These are areas where new technologies might possibly transform the geography of trade and production, but where very little is so far known.[10]

4. Factor Prices and Income

The fundamental determinants of the spatial variation of per capita income can be grouped into three broad headings. First nature geography: the second nature geography of access to markets, suppliers, and ideas: and third, the effects of social infrastructure, "the institutions and government policies that determine the economic environment within which individuals accumulate skills and firms accumulate capital and produce output" (Hall and Jones 1999, p84). Each of these determinants affects income directly, as well as by changing the incentives to make investments and accumulate factors of production.

This is not the place to review the literature on social infrastructure and we simply note that attempts to quantify its role have drawn heavily on geographical variables as proxies.[11] For example, in Hall and Jones (1999), social infrastructure is modelled as a function of distance from the equator, a measure of openness to international trade, the fraction of the population speaking English, and the fraction of the population speaking a European language. They find that these four variables account for 41% of the cross-country variation in social infrastructure and 60% of the cross-country variation in income per capita

The work of Sachs and his coauthors has focussed largely on first nature geography (Gallup et al. 1999; Gallup and Sachs 2000; McArthur and Sachs 2001; Sachs and Warner 1999; and Sachs 2000). Thus, Gallup et al. (1999) find that countries with a large percentage of their population close to the coast, low levels of malaria, large hydro-carbon endowments, and low levels of transport costs (as measured by IMF data on the cif/fob import price ratio) have higher levels of income per capita. These four variables alone explain nearly 70% of the variation in per capita income for a sample of 83 developed and developing countries.

Second-nature geography, or the location of economic agents relative to one another, affects per capita income through several different mechanisms. One is technology spillovers, which may diminish with the geographical distance between economic agents, as will be discussed further below. Another is countries' distance from the markets in which they sell output and from sources of supply of manufactured goods, intermediate inputs, and capital equipment. Trade costs reduce export receipts and increase prices of these inputs, squeezing the value added attributable to domestic factors of production.[12]

The idea that access to markets is important for factor incomes dates back at least to Harris (1954), who argued that the potential demand for goods and services produced in a location *i* depends upon the distance-weighted GDP (or, more generally, distance weighted economic activity) in all locations

$$MP_i = \sum_{j=1}^{R} GDP_j d_{ij}^{\gamma}, \tag{8}$$

where MP_i is the 'market potential' of country i, d_{ij} is the bilateral distance between locations i and j and γ is a distance weighting parameter, traditionally set at -1.

Much of the traditional geography focussed on the implications of market potential for the location of economic activity (see, for example, Clark et al. 1969; Dicken and Lloyd 1977;

and Keeble et al. 1982) with relatively little structural econometric estimation. Early econometric investigations of the role of market access in determining the cross-country distribution of income include Hummels (1995) and Leamer (1997). Hummels (1995) explores the role of three alternative measures of geographical location within the Solow and augmented Solow models. One is a measure of distance-weighted GDP in all other countries, constructed according to (8). The second two measures relate to a country's distance from the three main centres of world economic activity (the US, Japan, and Germany) and are respectively the sum and minimum of these three distances. In an equation for steady-state levels of per capita income, the geography measures are highly statistically significant, reduce the estimated magnitude of the coefficients on the Solow model variables, and improve the fit of the regression. Leamer (1997) examines the importance of access to Western Europe markets for post-reform income per capita in Eastern Europe. He uses a measure of market-access based on equation (8), with the distance weighting parameter γ derived from estimating a gravity equation. (Data on internal area is used to evaluate 'own distance', d_{ii}). The variation in access to Western markets within Eastern Europe suggests that these countries differ markedly in terms of their potential to achieve higher standards of living.

Although the focus is not on access to markets *per se*, Frankel and Romer (1999) explore the relationship between a measure of international openness (the ratio of trade to GDP) and levels of per capita income. One of the central problems in the literature concerned with openness and growth is the potential endogeneity of international openness. Therefore, Frankel and Romer (1999) use geography measures, including bilateral distance, area, land-locked status, and population, as instruments for bilateral trade flows. The predicted values for bilateral trade flows from this first-stage regression are then used to construct the ratio of total trade to GDP.[13] Evidence is found of a positive and statistically significant relationship between levels of per capita income and exogenous variation in the trade ratio due to the geography measures.

Redding and Venables (2000) use the structure of the Krugman and Venables (1995) model to obtain theory-consistent measures of both market-access and supplier-access.[14] From the theoretical discussion in Section 2, a firm in a monopolistically competitive industry will make zero profits if it achieves a volume of sales equal to $\bar{x}$ in equation (5). The volume of sales achieved depends on prices, which are a constant mark-up over marginal cost.

Equation (5) thus implicitly defines the maximum wage that a manufacturing firm in location *i* can afford to pay consistent with zero equilibrium profits. Dropping the superscript, and assuming that the marginal cost function is Cobb-Douglas in labour (with share β) and intermediate inputs (with share α) equation (5) is,

$$w_i^{\beta\sigma} = \left(\frac{\theta}{\bar{x}}\right) G_i^{-\alpha\sigma} \sum_j \left(t_{ij}\right)^{1-\sigma} E_j \left(G_j\right)^{\sigma-1} \qquad (5')$$

The term in the summation is the *market-access* of country *i*,

$$MA_i \equiv \sum_j \left(t_{ij}\right)^{1-\sigma} E_j \left(G_j\right)^{\sigma-1}, \qquad (9)$$

and is the theoretically founded analogue of market potential. It is comprised of expenditure in each market *j*, together with the price index (this measuring the amount of competition in the market, between which expenditure has to be shared), and adjusted according to transport costs from *j* to *i*. Terms in this expression are not directly observable, but can be derived from gravity estimation. We saw earlier how gravity models generate estimates of the between country trade frictions, $\left(t_{ij}\right)^{1-\sigma}$. The 'own distance', t_{ii}, can be constructed using a number of alternative approaches, some of which exploit information on internal area. Country dummies are used to capture importer effects, $E_j\left(G_j\right)^{\sigma-1}$. Combining these yields an estimate of the market-access of country *i*.

If intermediate goods are used in production, $\alpha > 0$, then transport costs also reduce the wage payable via an increase in the price of intermediates. This is captured in the term $G_i^{-\alpha\sigma}$ in equation (5'). Using the definition of the price index (equation (1)) we define country *j*'s *supplier access* analogously to market access as,

$$SA_j \equiv \sum_i n_i \left(p_i t_{ij}\right)^{1-\sigma} = G_j^{1-\sigma} \qquad (10)$$

Once again, estimates of country dummies (now the exporter rather than the importer dummy) from the gravity model provide the information needed to construct the series.

Having used estimates from a gravity model to construct the market-access and supplier-access variables, these are then combined with cross country data on per capita income to estimate (5') which, in logs, is

$$\ln w_i = \xi + \varphi_1 \ln SA_i + \varphi_2 \ln MA_i + u_i \qquad (11)$$

where the parameters φ_1 and φ_2 are functions of underlying structural parameters of the model, $\varphi_1 \equiv \alpha/\beta(\sigma-1)$, $\varphi_2 \equiv 1/\beta\sigma$. The stochastic error u_i includes cross-country variation in the price of other factors of production that enter manufacturing unit costs, technical differences, and other stochastic determinants of manufacturing wages.

Table 1 reports the results of estimating (11) using a cross-section of data on 101 developed and developing countries with GDP per capita as a proxy for manufacturing wages.[15] Because of the potential endogeneity of domestic market and supply capacity, only measures of *foreign* market and supplier access are considered (i.e. own effects are ignored, so summations in (9) and (10) are over $j \neq i$).[16] Column (1) presents the results using foreign market access alone. The estimated coefficient is positive and explains about 35% of the cross-country variation in income per capita. Column (2) includes information on supplier access as well. Separately identifying the coefficients on these two variables is difficult given their high degree of correlation. However, choosing values for α and σ implies a linear restriction on the estimated coefficients, $\varphi_1 = \varphi_2\, \alpha\sigma/(\sigma\text{-}1)$, and column (2) reports the results of estimating for values of $\alpha = 0.5$ and $\sigma = 10$, both of which are broadly consistent with independent empirical estimates. Including foreign supplier access reduces the magnitude of the estimated coefficient on foreign market access, but it remains highly statistically significant.

There are a number of concerns that one might have about these results. Is one really identifying an effect of economic geography, or just picking up that rich countries tend to be located next to rich countries, particularly within the OECD? Could the results not be explained by some third variable (eg unobserved technology differences), that is correlated with both income per capita and foreign market / supplier access? Redding and Venables undertake a number of robustness tests to address such concerns. These include augmenting the specification with a large number of control variables for factor endowments, physical geography, and social, political, and institutional considerations. For example, column (3) reports the results for non-OECD countries only, and column (4) presents results for non-OECD countries only, with their foreign market access calculated on the basis of their distance from OECD markets only: it asks, to what extent can variation in income per capita across developing countries be explained by access to OECD markets? In both cases, the effect of foreign market access is robust and highly statistically significant.

Table 1: World Market Access, Supplier Access, and GDP per capita

ln*(GDP per capita)*	(1)	(2)	(3)	(4)
Obs	101	101	79	79
Year	1996	1996	1996	1996
α		0.5		
σ		10		
$\ln(FMA_i)$	0.476	0.320	0.425	0.307
	[0.076]		*[0.081]*	*[0.074]*
$\ln(FSA_i)$	-	0.178	-	-
		[0.039]		
R^2	0.346	0.360	0.248	0.152

Notes: The results reported are from Redding and Venables (2000). Dependent variable is ln(GDP per capita). Independent variables are ln(Foreign Market Access), $\ln(FMA_i)$, and ln(Foreign Supplier Access), $\ln(FSA_i)$. $\ln(FMA_i)$ and $\ln(FSA_i)$ are generated from estimating the trade equation (equation (5)). Since these variables are generated from a prior regression bootstrapped standard errors are reported in square parentheses (200 replications). The wage equation estimation sample in Columns (1) and (2) is 101 countries. Column (3) estimates the model for the sample of 79 developing countries only. Column (4) estimates the model for 79 developing countries with a measure of $\ln(FMA_i)$ constructed only using data on OECD market capacities.

Wage gradients can be estimated on sub-national as well as international data, and Hanson (1998a, 2000a) performs such an estimation using a panel of US counties. Ignoring intermediate goods, his specification is equation (5') with $\alpha = 0$. In his basic specification this is estimated using county data on average earnings, and taking as independent variable the aggregate income of counties in a set of concentric circles at increasing distance around each observation, each distance weighted according to a factor $\exp(\beta_2 d_{ij})$, (where d_{ij} is distance, and this weighting factor corresponds to $(t_{ij})^{1-\sigma}$). The equation is estimated in first differences so that any time-invariant features of counties are swept out. Hanson finds a powerful wage gradient effect, with his measure of market access having a positive effect on earnings, and within this measure, distance (coefficient β_2) having a highly significant effect.

In the augmented version of his model Hanson addresses the endogeneity of the price index, G_j, by assuming that labour is perfectly mobile across counties (as in Krugman 1991a), so that real wages are equalized. Hypothesizing that housing is the only immobile factor (as in

Helpman 1998), and that it takes a fixed share 1 - μ of income, real wages are $w_j / G_j^{\mu} P_j^{1-\mu}$ where P_j is the price of housing (so the denominator is the cost of living index in the *j*th county). The value of housing expenditure satisfies $P_j H_j = (1-\mu) Y_j$ where H_j is the (exogenous) housing stock, so the equilibrium value of the price index is

$$G_j^{\mu} = \frac{1}{w_j} \left(\frac{(1-\mu) Y_j}{H_j} \right)^{\mu - 1} \qquad (12)$$

Using this in (5), together with manufacturing expenditure $E_j = \mu Y_j$ gives estimating equation,

$$w_i = \xi + \sigma^{-1} \ln \left(\sum_j^R Y_j^{\frac{\sigma(\mu-1)+1}{\mu}} H_j^{\frac{(1-\mu)(\sigma-1)}{\mu}} w_j^{\frac{\sigma-1}{\mu}} e^{-\tau(\sigma-1) d_{ij}} \right) + u_i \qquad (13)$$

where transport costs are modelled as an exponential function of distance: $T_{ij} = e^{-\tau d_{ij}}$.

Columns (1) and (2) of Table 2 present the results of estimating this specification using non-linear least squares for the periods 1970-80 and 1980-90. All variables are signed according to economic priors and are highly statistically significant. The inclusion of controls for the manufacturing price index, G_j, is found to improve the fit of the regression. The estimated values of the elasticity of substitution, σ, are broadly consistent with independent econometric estimates of this parameter, and are found to have fallen between the two sample periods. As implied by theory, the estimated expenditure share on tradable goods, μ, lies between 0 and 1, although a value above 0.9 is somewhat high. The estimated value of transport costs, τ, rises over time, and this may reflect a shift in production away from low-transport-cost manufactures to high-transport-cost services during the sample period. The estimated values of σ imply a markup factor of price over marginal cost that ranges between 1.15 and 1.25.[17][18]

The time-differenced specification controls for unobserved heterogeneity across counties in the level of manufacturing wages. However, it could be that wages have risen faster in counties with favourable exogenous amenities (eg weather or natural geography) or that have accumulated human capital (both through the private rate of return to human capital acquisition and through any externalities) and that these omitted variables are correlated with changes in market access. Since human capital accumulation may, in part, be determined by economic geography, it is not clear that one wants to exclude this component of the change in wages from the analysis. However, Hanson (2000a) shows that his results are robust to including a whole range of controls for levels of human capital, demographic composition of the working age population, and exogenous amenities.[19] Results including these controls are shown for the main estimation sample for the period 1980-90 in column (3).

Table 2: Market Potential and Wages Across US Countries

	(1)	(2)	(3)
Obs	3705	3705	3705
Time Period	1970-80	1980-90	1980-90
σ	7.597	6.562	4.935
	(1.250)	*(0.838)*	*(1.372)*
μ	0.916	0.956	0.982
	(0.015)	*(0.013)*	*(0.035)*
τ	1.970	3.219	1.634
	(0.328)	*(0.416)*	*(0.523)*
Wage Controls	no	no	yes
Adj. R^2	0.256	0.347	0.376
Log Likelihood	-16698.1	-16576.9	-16479.9
Schwarz Criterion	-16714.0	-16592.9	-16575.5

Notes: Reported results are from Hanson (2000a). Estimation is by non-linear least squares. Sample is all US counties in the continental United States, and the equation estimated is the time-difference of equation (13). All variables are scaled relative to weighted averages for the continental United States. The dependent variable is the log change in average annual earnings from Regional Economic Information System (REIS), US BEA. Regional income is total personal income from REIS. The housing stock is measured by total housing units from the US Census of Population and Housing. The specification in column (3) includes controls for human capital, demographic characteristics, and exogenous amenities. Heteroscedasticity-consistent standard errors are in parentheses. The Schwartz Criterion is written as ln(L) - k*ln(N)/2, where k is the number of parameters.

The empirical results surveyed so far provide econometric evidence of wage gradients across geographical space (both across and within countries) consistent with the predictions of economic geography models.[20] *Ceteris paribus*, locations that are remote from markets and sources of supply of intermediate inputs are characterised by lower nominal wages. As always, there remain potential concerns relating to identification and simultaneity that could be resolved by observing a controlled or natural experiment that generates exogenous variation in market and supplier access. In the remainder of this subsection, we discuss a group of papers that have exploited trade liberalisation in Mexico as precisely such an experiment.

In 1985 Mexico opened its economy to international trade, bringing to an end four decades of import-substitution industrialization. Hanson (1996), (1998b), (1998c) finds that trade reform has contributed towards the breakup of the traditional manufacturing belt centred on Mexico City and the formation of new industry centres in Northern Mexico.[21] For example, in the apparel industry Hanson (1996) finds that prior to trade liberalization, production was concentrated around Mexico City and largely orientated towards the Mexican market, with design and marketing concentrated in Mexico City and assembly in the neighbouring states. With trade liberalization, there was a substantial relocation of manufacturing activity towards the US border, and the nature of manufacturing activity was also reorientated - away from domestic production towards offshore assembly for foreign (largely US) firms. There is evidence of a negative relationship between relative wages and distance from Mexico City prior to 1988, and of a statistically significant decline in the size of the estimated coefficient on distance from Mexico City between 1985 and 1988.[22] This provides support for the existence of a regional wage gradient centred on Mexico City prior to trade liberalization and of the partial breakdown of this regional wage gradient as production re-orientated towards the United States.

Hanson (1997) analyses the determinants of state relative to national manufacturing wages for a panel of two-digit Mexican manufacturing industries over the period 1965-88. Nominal wages are found to be negatively correlated with both distance from Mexico City and distance from the Mexico-USA border. A 10% increase in distance from Mexico City is associated with a 1.9% reduction in the relative state wage, while the same increase in distance from the Mexico-USA border is associated with a 1.3% reduction.

Geography and Technology

Much of the discussion in this section has been concerned with distance from markets and sources of supply as an explanation for spatial variation in factor prices. Distance is important because of the transportation costs incurred on deliveries to markets and shipments of intermediate goods and capital equipment. An alternative explanation for variation in factor prices across space is the existence of technology differences, which may arise, for example, because knowledge spillovers diminish with geographical space between economic agents.[23] A number of papers have presented empirical evidence that knowledge spillovers are much greater within than between countries: see, for example, Branstetter (2001), Coe and Helpman

(1995), Eaton and Kortum (1999a) (1999b), Jaffe and Trajtenberg (1998), and Keller (2000), (2001). Much of the literature has been concerned with the extent to which international knowledge spillovers are trade-related (see, in particular Coe and Helpman 1995). Since, as discussed in Section 3, distance plays a substantial role in explaining international trade flows, this suggests a potential role for geography in the diffusion of ideas. The role of international trade flows *per se* has been questioned by Keller (1998), and Keller (2001) examines the direct relationship between distance and international knowledge spillovers. A 10% higher distance from a major technology-producing country such as the U.S. is found to be associated with a 0.15% reduction in productivity.

5. The Location of Activity

We now turn to the question of how geography determines the structure of production across locations. We organise the material into three sub-sections. The first deals with some measurement issues and descriptive studies. Is it possible to make statements along the lines of 'the US is more regionally specialized than the EU', and what are the stylised facts concerning specialization and localization? The second and third sections seek to go behind the descriptives and ask what determines location. In 5.2 we look at studies on international data, and in 5.3 sub-national studies. This sub-section is also where we deal with the issues of clustering and agglomeration. Are industries more localised than would be suggested by alternative hypotheses, and if so, why?

5.1. Measurement Issues and Descriptive Studies

Localisation and Specialisation

The researcher may wish to ask two distinct, but related, questions. One is how localised is a particular economic activity, and the other is how specialised is a particular geographical unit? This question can be addressed using different measures of activity, typically employment or production, and in the following discussion we refer to production. Denoting the production of industry *k* in location *i* as y_i^k, the localisation of industry *k* can be addressed by looking at y_i^k relative to total production of that industry: $\ell_i^k = y_i^k / \sum_i y_i^k$. This measures the share of location *i* in the total production of industry *k*. Conversely, the specialisation of a location can be studied by looking at y_i^k relative to the total production of that location, $s_i^k = y_i^k / \sum_k y_i^k$. This measures the share of industry *k* in region *i*'s total production of all industries.

Recognising that regions and industries differ in size we might want to normalise these two measures. If we normalise the first by the share of the location in overall activity and the second by the share of the industry in overall activity we end up with a measure which we call the *location quotient,*[24]

$$r_i^k = \frac{y_i^k / \sum_i y_i^k}{\sum_k y_i^k / \sum_i \sum_k y_i^k} = \frac{y_i^k / \sum_k y_i^k}{\sum_i y_i^k / \sum_k \sum_i y_i^k} \qquad (14)$$

These are two equivalent expressions or interpretations of the location quotient. The first is as a measure of the localisation of industry *k* in *i*, relative to the localisation of activity as a whole in *i*. The second is as a measure of location *i*'s specialisation in industry *k* relative to the share of the industry in total world output. It is important to be clear that economic geography models make statements about both localisation and specialisation. We shall refer to statements about the distribution of r_i^k across locations *i* for given industry *k* as statements about the localisation of industry *k*, noting that *k* could be an aggregate of many or all sectors. And we shall refer to statements about the distribution of r_i^k across industries for a given location as describing the specialisation of location *i.*

Summary Statistics of Localisation and Specialisation

The matrix r_i^k contains the distributions of localisation and specialisation, and we typically want to be able to summarise these distributions in order to make statements like 'industry *k* has become more localised' or 'location *i* is more specialised than location *j*'. To do so requires calculation of a summary statistic of the distribution, and the problems of collapsing distributions down to a meaningful scalar representation is fraught with conceptual issues that have long occupied the literature on the distribution of personal income and wealth (see for example Cowell 1995). In addition to these concerns, Duranton and Overman (2001) suggest five properties that we would like such measures to satisfy, from a standpoint of economic geography: (1) they should be comparable across industries or locations; (2) they should take in to account the overall distribution of activity across different sectors (for specialisation) and across different locations (for localisation); (3) they should be able to distinguish between ‘lumpiness’ in the unit of observation and geographical concentration; (4) they should be defined over the correct spatial units; (5) they should allow the statistical significance of the measured specialisation or localisation to be assessed. The measures that we discuss in the remainder of this section only satisfy properties (1) and (2). We consider indices that satisfy conditions (3)-(5) in section 5.3, as they have so far only been applied to sub-national data.

Looking first at specialization, we seek a summary measure of the specialization of location *i*. Various measures of dispersion can be used, defined either on absolute production shares, s_i^k, or shares relative to industry size, r_i^k For example, the Herfindahl index of absolute specialisation, takes the form $h_i = \sum_k (s_i^k)^2$. For the bilateral comparison of the specialisation of two different locations Krugman (1991) computes the absolute value of the difference in production shares, $K_{ij} = \sum_k \left| s_i^k - s_j^k \right|$.

Analogous measures are used for localisation. Various authors compute locational gini coefficients, $g^k = gini^k(r_i^k)$, referred to by Kim (1995) as ‘Hoover’s coefficient of localisation’.[25] Haaland et al. (1999) argue that conditioning on the distribution of the location of activity as a whole is not consistent with the intuitive concept of agglomeration, and so they use an ‘absolute’ gini coefficient, $ga^k = gini^k(\ell_i^k)$, calculated analogously.[26]

Findings on Localisation and Specialisation

In this section we review some of the main stylised facts that emerge from descriptive studies of the location of economic activity, before moving on in sections 5.2 and 5.3 to more formal work that seeks to disentangle the various forces determining location.

It is taken as self evident by geographers that activity clusters. ‘In fact, the geographical concentration of economic activities, at a local or subnational level, is the norm not the exception.....’ P. Dicken (1998), p11. The most systematic evidence on overall agglomeration comes from the work of urban economists and historians on cities. In his classic book, Bairoch (1988) provides a wide range of data on the size of cities and the extent of urbanisation. In 1300, Bairoch’s estimates put the urban population at 41 million out of a total of approximately 460 million (an urbanisation rate of roughly 9%). By 1900 this had risen to 260 million (an urbanisation rate of 16%), while by 1990 the urbanisation rate had risen to 37.6%, with roughly 1670 million people living in urban areas. By 2025, he predicts a world level of urbanisation of 57%. Not only does a high proportion of the world’s population live in these cities, but there are also a large number of such agglomerations. In 1980, there were roughly 2,290,000 cities with more than 100 thousand inhabitants, and by 1995 15 cities had a population greater than 10 million. We also know how these patterns change during economic development. Studies of localisation in developing countries confirm the hypothesis of Williamson (1965) that in growing from low-income levels countries go first through a period of regional divergence and concentration of development and industrialization in a restricted region of the country, followed by industrial deconcentration, growth of hinterland regions and a move towards regional convergence (see for example Henderson 1988 and Henderson, Lee and Lee 1999).

Particular types of economic activity are also massively localised. In 1995 the OECD countries produced 78% of the world's manufacturing output, despite containing only 15% of the world's population. A key feature of the process of economic development is the reallocation of resources from agriculture to manufacturing. Analysis of the spatial deconcentration of manufacturing production is therefore central to our understanding of economic development.

Turning to a finer sectoral level, there are a number of studies of the evolution of specialisation and localisation within countries or groups of countries. The US provides the longest available time series, and using this data Kim (1995) calculates specialization measures, K_{ij}, and the 'coefficient of localization', g^k, and finds that changes from 1860-1987 have been non-monotonic. Industries became increasingly localised and states increasingly specialised up to 1930. From then state specialisation fell substantially and is lower today than it was in 1860. On average, industries became less localised during this later period, but individual industries show large variations around this average trend.

In the EU, data is available over a much shorter period. However, there is evidence that, in contrast to the US, EU countries are becoming increasingly specialised as European integration proceeds. Amiti (1999) uses data on both employment and production for her study of a selection of EU countries; Midelfart-Knarvik, Overman, Redding and Venables (2000) use data on gross output for the entire EU 12; while WIFO (1999) use value-added. The pattern of increasing specialisation with respect to the EU average seems to be consistent from the mid-1980 onwards, although the changes are not particularly large. Midelfart-Knarvik et al. use K_{ij} to show that countries are also becoming less similar to one another, with 71 out of 91 bilateral comparisons revealing increasing dissimilarity. Industrial localisation experiences (measured by g^k and ga^k) are diverse with some industries localising and others dispersing. This is consistent with earlier work by Brulhart and Torstensson (1996) and Brulhart (1998).

A number of these papers push such descriptive exercises further by constructing measures of industry characteristics - for example the extent of increasing returns to scale and the resource intensity - and running regressions of localisation coefficients on these industry characteristics. See, for example, Kim (1995, 1997), Brulhart and Torstensson (1996), Amiti (1999), Brulhart (1998), Haaland et al. (1999). Results are mixed, reflecting both the small number of observations, and the lack of any real theoretical foundation for the estimation. Some of the results are suggestive. For example, Kim (1995) argues that his findings "support explanations based on production scale economies and the Heckscher-Ohlin framework but are inconsistent with explanations based on external economies" (p. 881). However the lack of theoretical foundations is an important limitation of these studies.

The measurement exercises reported above are clearly important in describing the data and trends in its evolution, but we need to go further to investigate the economic forces driving these variables. For example, evidence of increasing specialisation in the EU does not, by itself, discriminate between comparative advantage and agglomeration as drivers of specialisation. In the next two sections we look at studies that attempt to identify the mechanisms at work. In 5.2 we report the studies based on international data, and in 5.3 look at some of the (much larger number of) studies using sub-national data.

5.2. International Studies

Home Market Effects

We saw in section 2 that the number of firms (or varieties) in each location, n_i^k, might either be determined exogenously or, for monopolistically competitive industries, endogenously through a zero profit condition (equation (5)).[27] These two cases give rise to different predictions about the effect of demand (or market access) on production. More specifically, the presence of increasing returns to scale and transport costs in the monopolistically competitive case implies a 'home market' or 'magnification' effect, whereby an increase in demand for a good results in a more than proportionate increase in production of the good. Intuitively, increasing returns to scale imply that firms would like to concentrate production in a single location, while the existence of transport costs implies that, other things equal, this concentration will occur close to large markets.

The argument can be seen more formally by referring back to section 2. Suppose that all economies are identical, except that we now give country 1 a small increase in E_1^k. Suppose also that there are no intermediate goods and all factor prices stay constant. If n_i^k is fixed and industry k not monopolistically competitive, then we see from equation (3) that this increase in E_1^k will raise outputs in all countries, while increasing country 1's output less than proportionately. But if there is monopolistic competition, equation (5) must hold for all countries. It will do so if G_1^k falls so that $E_1^k\left(G_1^k\right)^{\sigma-1}$ remains constant, while all other price indices, G_j^k, $j \neq 1$, stay constant. From inspection of the price index (1), given t_{ii},= 1 and t_{ij}, > 1 ($i \neq 1$) this requires an increase in n_1^k and *fall* in all other n_i^k, $i \neq 1$. The falls in n_i^k, $i \neq 1$, must mean that country 1 output increases more than proportionately, if supply is to equal the new value of demand.[28]

Davis and Weinstein (1998, 1999) use this home market effect as a basis for testing between models of imperfect competition/increasing returns to scale and perfect competition/constant returns to scale. It requires estimating the relationship between variations in expenditure and variations in output across countries and industries, and seeing whether there is a proportional response of more or less than unity. Davis and Weinstein (1998) consider a nested specification, where factor endowments are assumed to determine production at the more aggregate level (3 digit), while economic geography effects operate in disaggregated industries. Using data for 13 OECD countries, they first construct measures of ‘idiosyncratic demand’ for each 4-digit industry based on demand in the country and its trading partners, distance weighted.[29] Estimating the effects of this demand variable on production in a pooled sample across countries and all 4 digit industries they find an elasticity of production with respect to demand of 1.6, indicating a strong home market effect. Estimating a single coefficient across all industries is unsatisfactory, as we expect that industries have different market structures. Disaggregating and running separate regressions for each 3 digit industry (with the sample of countries and 4 digit sub-industries), they find evidence of a home market effect (coefficient greater than unity) in a majority of industries, the estimated coefficient being significantly greater than unity in four industries, and significantly less than unity in two.

These results are broadly similar to those obtained using a related specification on data for 29 sectors and 47 Japanese prefectures in 1985 (Davis and Weinstein 1999). Statistically significant home market effects are found in 8 out of 19 manufacturing sectors, including transportation equipment, iron and steel, electrical machinery, and chemicals. These effects are found to be quantitatively important: for the 8 sectors with statistically significant home market effects, a one standard deviation movement in idiosyncratic demand is found to move production, on average, by half a standard deviation.

Home market effects have also been found by several other authors. Head and Ries (2001) look at US-Canada trade at the 3-digit level, 1990-95. Since they only have a single pair of countries they have to rely on cross industry or time series variation in the data to identify the home market effect, and only estimate a single effect for all industries (like Davis and Weinstein's pooled regression). They find a weak home market effect in their industry cross section (an elasticity of production with respect to local demand of 1.12), which becomes less than unity once the time series variation is employed. This is probably explained by the short time series – the home market effect is essentially a long-run relationship driven by entry and exit of firms or varieties.

Feenstra, Markusen and Rose (2001) identify a home market effect from estimating a gravity model separately for differentiated products, reference priced exports, and homogeneous goods. The coefficient on income of the exporting country rises as they go from homogeneous to more differentiated products. For differentiated products the coefficient is slightly greater than unity and significantly greater than the coefficient on importer country income, indicating the presence of a home market effect in these goods.

Geography and Comparative Advantage

Whereas Davis and Weinstein separate out factor endowment effects (operating at an aggregate level) and geography effects (operating at a disaggregate level), Midelfart-Knarvik, Overman and Venables (2000) show how the effects can be combined. The basis of their approach is to estimate a linearised version of equation (3) on a panel of European countries and industries.

To implement this they assume first that all industries are perfectly competitive and that the numbers of varieties of each industry produced in each country are exogenously determined and proportional to the size of the industry and size of the country, thus $n_i^k = \sum_k y_i^k \sum_i y_i^k$. Using this together with (4) and (14) in (3) gives[30]

$$r_i^{\,k} \;\equiv\; \left(\theta c^{\,k}(w_i, G_i)\right)^{1-\sigma^k} \sum_j \left(t_{ij}^{\,k}\right)^{1-\sigma^k} E_j^{\,k} \left(G_j^{\,k}\right)^{\sigma^k - 1} \qquad (15)$$

Although the numbers of varieties are set exogenously, (15) indicates how both cost and demand factors determine the matrix of location quotients.

Linearisation of the model gives, on the right hand side, a sum of interactions between country characteristics and industry characteristics. Denoting the country characteristics x_i[j] and industry characteristics y^k [j], where j is an index running across the set of interactions, gives estimating equation of the form,

$$\ln(r_i^{\,k}) \;=\; \zeta \;+\; \sum_j \beta[j]\left(x_i[j] \;-\; \bar{x}[j]\right)\left(y^{\,k}[j] \;-\; \bar{y}[j]\right) \;+\; \varepsilon_i^k. \qquad (16)$$

The interpretation of this is seen by thinking of the interaction between, say, skilled labour abundance and skilled labour intensity. Countries which have skilled labour abundance greater than some reference level $x_i[j] \;>\; \bar{x}[j]$, will have high production in industries with skill labour intensity above a reference level $y^{\,k}[j] \;>\; \bar{y}[j]$, and vice- versa -- a Rybzcynski effect. This multiplicative form of interaction holds for other pairs of country and industry characteristics. Expanding the products in (16) yields an equation in which the parameters to be estimated are β[j], $\beta[j]\bar{x}[j]$, and $\beta[j]\bar{y}[j]$, and the estimates of β[j] are given in table 3.

The first three are interactions of factor endowments with factor intensities. We see that all are significant by the end of the period, with the absolute magnitude of the scientist abundance/ R&D intensity interaction having nearly trebled in size. The fourth interaction captures forward linkages, so interacts a measure of supplier access (proximity to other manufacturing sectors, as defined above) with the share of intermediates in production; the coefficient has the correct sign, although is barely significant. The fifth term measures backwards linkages. This is the relative importance of derived demand (measured as the difference between market access computed for final products and market access computed for intermediates) interacted with the share of each industry’s output that is sold to industry. Backwards linkages are significant, although become less important over the period. Finally, to capture in a rigorous manner the possibility that high transport cost industries are drawn to central locations, the transport intensity of products is interacted with the elasticity of market access with respect to transport intensity. The estimated coefficient is insignificant and has the wrong sign.

Although this paper abstracts from monopolistic competition, it does show how geography can be combined with comparative advantage, and indicates the relative contributions of the two sets of forces.

Table 3: Comparative advantage and geography: Dependent variable $\ln(r_i^k)$

	Variable	1980-83	1985-88	1990-93	1994-97
	Interactions: β[*j*]				
β[1]	Agric. endowment	0.078	0.140	0.166**	0.158**
	* agricultural input intensity	(0.114)	(0.097)	(0.085)	(0.079)
β[2]	Skill endowment	1.503***	1.484***	1.479***	1.663***
	* skill intensity	(0.439)	(0.420)	(0.463)	(0.582)
β[3]	Researchers+scientists endowment	0.584*	0.741**	1.108**	1.624***
	* R&D intensity	(0.325)	(0.389)	(0.536)	(0.581)
β[4]	Intermediate prices (supplier access)	0.570	0.754	0.799	1.096*
	* intermediate goods intensity	(0.811)	(0.771)	(0.667)	(0.689)
β[5]	MA final demand - MA intermediate demand	0.182***	0.171***	0.130***	0.083**
	* share of output to industry	(0.059)	(0.052)	(0.043)	(0.041)
β[6]	Elasticity of MA wrt transport intensity	-0.395	-0.270	-0.319	-0.382
	* transport intensity	(0.315)	(0.299)	(0.290)	(0.275)
	Adjusted R^2	0.105	0.116	0.143	0.137
	Number of obs	456	456	456	456

Note: Standard errors reported in brackets; * * *= significant at 1% level; * * = significant at 5% level; * = significant at 10%. All regressions are overall significant according to standard F-tests.

5.3. Evidence From Sub-national Studies

In this section we consider the lessons that we can learn from sub-national empirical work. The literature here is much larger than the corresponding international literature, and addresses the issues of the existence and determinants of clustering. The reasons for the greater amount of empirical evidence would appear to be twofold. First, urban and regional economists have been interested in agglomeration economies for a longer period; second comparable production data is more readily available for sub-national units, particularly in the US. We organise our review in two sections. The first considers a number of papers that take a step back and test for the existence of localisation against the alternative hypothesis of the random location of ‘lumpy’ activity. The second section considers the determinants of specialisation and localisation at the sub-national level.

Location, Lumpiness and Randomness

In Section 5.1 we identified five properties that we think a good measure of localisation or specialisation should possess. The measures that we discussed earlier generally satisfied the first two of these properties (comparability and controlling for the distribution of aggregate activity). In this section we discuss the literature that proposes measures of localisation that satisfy some of the remaining properties. That is, measures that control for industrial lumpiness (property 3); that consider the problem of spatial unit boundaries (property 4) and that assess the extent of localisation against the null hypothesis of randomness (property 5).

Industries characterised by higher increasing returns to scale (larger plants) will *ceteris paribus* appear more spatially concentrated than industries characterised by low increasing returns (small plants), simply because they have relatively few plants. This observation is important if we want to compare the location of (say) aircraft manufacturing with that of textiles. After controlling for this, we might want to assess whether the patterns that we see are systematic or whether they could have occurred by chance. This emphasis on randomness is particularly appropriate in a situation where departures from randomness (localisation) is driven mainly by cumulative causation, because the multiple-equilibrium properties of these models tell us that localisation will occur, but not where it will occur.

Ellison and Glaeser (1997) were the first to address these issues directly. They specify a stylised location model where industries may be localised because (i) overall activity is agglomerated; (ii) activity within the industry is concentrated in a few randomly located plants; (iii) activity within the industry is concentrated in non-randomly located plants. To separate out the third cause of localisation from the first two, they proceed as follows. First, they define a measure of sectoral relative to overall localisation, $G_{EG}^{k} = \sum_i (\ell_i^k - \ell_i)^2 / \left(1 - \sum_i \ell_i^2\right)$ where using our earlier notation $\ell_i^k = y_i^k / \sum_i y_i^k$ is location *i*'s share of industry *k*, and $\ell_i = \sum_k y_i^k / \sum_i \sum_k y_i^k$ is location *i*'s share in overall activity. This measures the extent of localisation for industry *k* over and above localisation of activity as a whole. To allow for plant size, they first construct the standard Herfindahl index of *industrial* concentration for industry *k*, $H^k = \sum_j \left(z_j^k\right)^2$ where z_j^k is the share of *plant j* in total industry *k* output. They then use this to construct a measure, γ_{EG}^k, that controls for the localisation that would arise just as a consequence of plant size and industrial concentration. The measure takes the form, $\gamma_{EG}^k = (G_{EG}^k - H^k) \,/\, (1 - H^k)$. They show that the expected value of this measure is zero if plants are randomly located, so positive values of γ_{EG}^{k} indicate 'excess localisation', relative to activity as a whole and relative to random location of the industry's plants.

Ellison and Glaeser calculate this measure for the location of employment in 459 industries across 50 states in 1987. They find that 446 of the 459 industries are more localised than we would expect to arise randomly. However, although localisation is ubiquitous, many industries are only slightly localized, suggesting that previous literature may have over-emphasized the extent of localization. No clear classification of industries by extent of excess localisation is possible - the characteristics of least and most localised industries are quite variable. Ellison and Glaeser calculate the index at a number of different spatial scales. The results suggest that departures from randomness are strongest at the county level, substantial between counties in the same state, but fairly weak at the regional level. They also calculate a related index of co-agglomeration. For both three and two-digit industries, there is some evidence that sub-industries within these categories co-agglomerate, although the extent of co-agglomeration varies substantially across industries. There is also some evidence of co-agglomeration among industries with strong upstream and downstream linkages.

Similar indices have been calculated by Maurel and Sedillot (1999) for France, and Devereux, Griffith and Simpson (1999) for the UK. There are cross-country similarities in the most and least localised industries. For example, Hosiery, Jewellery, Other Carpets, and Spirit Distilling appear in the top 20 most concentrated industries in both the UK and US, while Cutlery, Woollen, and Periodicals are among the top 20 most localised industries in France and the UK. One could argue that this is indicative of second-nature effects given the very different first-nature geographies of the three countries. Finally, Duranton and Overman (2001) suggest a further development of Ellison and Glaeser which allows for the fact the location decisions are made over a continuous rather than discrete space (property 4); and that also allows them to assess whether departures from randomness are statistically significant or not (property 5).

Determinants of Localisation and Specialisation

The results outlined above suggest that internal returns to scale play an important role in understanding the distribution of activity at the sub-national level, and that there is localisation in a large number of industries that cannot purely be explained by industrial concentration (plant size). In attempting to explain this excess localisation, regional and urban economists seem far readier to admit that *both* comparative advantage and economic geography factors could matter for determining location. Possibly, this reflects the fact that the assumption of exogenous endowments of some factors (capital, skilled labour) is a much stronger one at the sub-national level where mobility is substantially higher. We start by considering a small number of papers that attempt to assess the proportion of excess localisation that can be explained by internal returns and the distribution of endowments. We then consider attempts to explain the residual excess localisation.

Ellison and Glaeser (1999) consider how much localisation can be explained by natural advantage by studying the shares of US states in different industries as a function of the interaction between industry and state characteristics. As discussed in Section 5.2, Midelfart-Knarvik, Overman, and Venables (2000) show that such an interaction formulation can be derived as the solution to a fully specified trade model. Ellison and Glaeser (1999) use data on four digit manufacturing for 1987. They have information on a range of state characteristics, including: electricity price; natural gas price; coal price; farmland; average manufacturing wage; percentage labour force with high skill; population density. The corresponding industry characteristics are; electricity use; natural gas use; coal use; agricultural input share; wages/value added; skilled labour intensity; percentage of output sold to consumers.

Their estimation takes the form of regressing state-industry employment shares (ℓ_i^k in our notation) on a non-linear function of state characteristics interacted with industry characteristics. They find that about 20% of the variation in these shares is explained by the characteristics, and suggest that this could increase to 50% with inclusion of more characteristics. There are problems with their approach, particularly in so far as it is not clear that some of the location characteristics are first nature - this is particularly true of the wage, skill composition and population density measures. If they are not first-nature, they are surely endogenous and this problem is not corrected for. This makes both interpretation and evaluation of the results difficult. Even ignoring these problems, their results suggests that between 50% and 80% (at least) of localisation at the state level is unexplained by natural advantages.

So, what explains the residual excess localisation? Assuming that we have correctly conditioned for all other factors, the simple, but uninformative answer is “some sort of agglomeration economy”. There are several related but separate strands to the empirical research on agglomeration economies. The first strand attempts to assess the importance of localisation versus urbanisation economies in explaining the location of activity. Localisation economies occur when there is a positive externality on firm productivity from other firms in the same sector; urbanisation when there is a positive externality on firm productivity from other firms in different sectors. Either type of externality could arise from Marshall’s three agglomeration forces - knowledge spillovers, labour market externalities or input-output linkages. Research also considers whether or not these returns are static or dynamic and what characteristics of the local environment matter for determining the extent of these externalities. See Henderson (1999) for a recent discussion of the issues. Henderson (1988) is most closely associated with the finding that localisation increases firm productivity, while Henderson et al. (1995) find that localisation also increases growth. This contrasts with the results of Glaeser et al. (1992) who find that diversity raises growth. The issue remains unresolved and Combes (2001) identifies problems with the empirical approach in this literature.

A related literature has examined the effect of the scale or density of economic activity on productivity levels. Early studies of the effect of city population size on productivity include Sveikauskas (1975), Segal (1976), and Moomaw (1981). Ciccone and Hall (1996) use information on employment densities for US counties to construct an index of the density of economic activity at the state-level. According to their preferred instrumental variables estimates, doubling the employment density in a county increases state labour productivity by 6% and total factor productivity (TFP) by 4%. Caballero and Lyons (1990), (1992) find substantially larger returns to scale for aggregate manufacturing in both the US and in Europe than for individual two-digit industries. One explanation is the existence of external economies of scale, although this interpretation has recently been questioned. Basu and Fernald (1997) find that estimates of returns to scale vary substantially according to whether one uses data on value-added or gross output, and argue that aggregation bias provides a more plausible explanation for these empirical findings.

Of more interest for us here, are the few papers that attempt to distinguish between the micro-economic mechanisms that might cause agglomeration economies. Dumais, Ellison and Glaeser (1997) use plant births and deaths to attempt to distinguish between the three possible sources. For selected years they have data on the population of manufacturing establishment in the US. Plants can be classified into 134 sectors located in one of 307 metropolitan areas and by US state. They find that, despite large plant turnover (73% of plants that existed in 1972 had closed by 1992) the extent of localisation remains constant. To examine the determinants of localisation, they construct three different measures. They use input-output tables to construct measures of supplier presence and product customer presence. To capture labour market agglomerations they construct a measure based on the risk of closure and a comparison of the plant's labour market mix to the average labour mix in the area. Finally, they construct a proxy for information flows using weights derived on co-ownership across multiple industries. With some caveats, their broad findings suggest that inputs help explain where existing firms locate new plants, while output matters more for plants created by new firms. Neither effect was very strong compared to the importance of labour mix (which was particularly important for new firms). Technology spillovers, although poorly proxied by their measure also seem to be important. Finally, input/output linkages seemed to be more important at the state level than the metropolitan level, consistent with assertions in some of the theoretical literature that these are generally useful for explaining large scale agglomerations (Krugman 1991b). Devereux, Griffith, and Simpson (1999), (2001) present some results along the same lines for the UK.

As we suggested, the technological proxy used by Dumais et al. is not a particularly good one. We do however have some additional evidence on the importance of *local* technological spillovers between firms. A series of papers starting with Jaffe et al. (1993) have compared the location of patent citations with the location of cited patents. For the US they find that the citation to domestic patents is more likely to be from domestic patents. They find a similar pattern at the state and particularly at the standard metropolitan statistical area level, even after conditioning for the localisation of particular industries. This is the strongest evidence we have to date on the importance of local knowledge spillovers in determining location.

We have only been able to provide a very brief overview of a larger literature on the determinants of sub-national localisation. What lessons can students of international location take from this collection of subnational studies? First, and perhaps most importantly, nearly all the evidence that we have at a sub-national level suggests that both endowments *and* geography matter in determining location. Second, the most informative descriptions of localisation try to address all of the properties that we outline above. In addition, there is a clear feeling that these descriptive measures will be more informative if they can be closely related to theory. Finally, we may need to concentrate on developing micro (firm) level data rather than aggregate data if we wish to separate out the forces driving both subnational and international location decisions.

6. Concluding Comments

The evidence surveyed here strongly suggests the importance of geography in determining international economic interactions, in influencing cross-country income distribution, and in shaping the structure of production across space.

While the current state of knowledge establishes that geography matters, we know much less about exactly why it matters. Distance clearly chokes off economic interactions, but is it because of transport costs, time costs, fixed costs of entering new markets, informational barriers, or difficulties encountered in managing remote supply chains or production operations? Similarly, activity benefits from being agglomerated, but are the benefits from demand and supply linkages, from pools of labour market skills, or from technical spillovers; and if the last of these, exactly how are they transmitted?

Answering these questions is crucially important as new technologies and further trade liberalizations continue to drive the process of globalization. What activities will relocate to developing countries, and what stay in established centres? What should be the trade and investment priorities of geographically disadvantaged regions? What are the implications of globalization for international inequality? Fortunately, both the analytical frameworks and the rich micro-economic data sets needed to address these questions are now becoming available, although much work remains to be done.

Endnotes:

1. We shall use the term market-access to measure how well placed a location is with respect to markets, and supplier-access to measure how well placed it is with respect to suppliers.

2. The 'folk theorem' of location theory says that, in the absence of increasing returns there will be 'backyard capitalism', with production potentially locating wherever there is demand.

3. We make the usual assumption that the same price indices, G_i^k, hold for both consumers and intermediate users of good *k* in country *i*.

4. For recent reviews of the theoretical economic geography literature, see Neary (2001) and Ottaviano and Puga (1997). Hanson (2000b) surveys empirical work, concentrating largely on regional and urban research. For a review of the empirical trade literature that concentrates on the predictions of neoclassical theory, see Leamer and Levinsohn (1995).

5. See also Deardorff (1998).

6. These have the advantage of controlling for the extent of a country's barriers to trade with all of its partners – what Anderson and van Wincoop (2000) call 'multilateral resistance'.

7. See, for example, Feenstra et al. (2001), Frankel (1997), and Soloaga and Winters (1999). The difference in estimated coefficients arises, at least in part, because of the treatment of zeros. Tobit estimation typically yields larger coefficients.

8. This elasticity is on the transport cost factor, thus doubling transport costs from 20% to 40% reduces trade volumes to $(1.4/1.2)^{-3} = 0.63$ of their initial level.

9. See Anderson and Marcouiller (2001) for an analysis of the role of insecurity and risk of appropriation in determining trade patterns.

10. For further discussion of the impact of new technologies see Leamer and Storper (2000) and Venables (2001).

11. For a recent study of the effect of institutions on economic performance, which uses variation in settler mortality as an instrumental variable for the type of institutions adopted by European colonists, see Acemoglu et al. (2000). McArthur and Sachs (2001) emphasize the role of both institutions and physical geography.

12. For example, suppose that intermediates account for 50% of costs and transport costs are borne by the producing country. *Ad valorem* transport costs of 10% on both final output and intermediate goods have the effect of reducing domestic value-added by 30% (compared to a country facing zero transport costs); the reduction in value-added rises to 60% for transport costs of 20% and to 90% for transport costs of 30%. See Radelet and Sachs (1998) for further discussion of this point.

13. See Leamer (1988) for a related measure of international openness.

14. See also Fujita, Krugman, and Venables (1999), Chapter 14.

15. A similar pattern of results is observed using data on manufacturing wages per worker for a subset of countries.

16. The full results including own country effects are given in Redding and Venables. The market access measure including both foreign and domestic effects explains up to 75% of the cross country income distribution.

17. These estimates imply a value of $\sigma/(\sigma-1)$ of greater than 1, and are thus consistent with increasing returns to scale. See Antweiler and Trefler (2000) for evidence of increasing returns to scale in a number of manufacturing industries from data on the net factor content of trade.

18. In Helpman (1998), the value of $\sigma(1-\mu)$ is crucial for the determinants of agglomeration. All the parameter estimates in Table 2 imply a value of $\sigma(1-\mu)<1$, so that an increase in transport costs increases the likelihood of agglomeration.

19. Roback (1982) and Kahn (1995) emphasize the relationship between local amenities and wages and land rents within cities. Rauch (1993) and Moretti (1998) provide empirical evidence of city-level human capital externalities, although Ciccone and Peri (2000) argue that these disappear when one controls for the potential complementarity between workers with different levels of human capital. Glaeser and Mare (1994) stress the role of human capital accumulation in explaining the urban wage premium.

20. See Dekle and Eaton (1999) for an analysis of wage and land rent gradients across Japanese prefectures. The wage and land rent data are used to estimate the effect of the agglomeration of economic activity on measured productivity. Relocating value-added 100km away is found to reduce its impact on productivity by 9% in Finance and 1% in Manufacturing.

21. There is a literature concerned with the more specific question of the effects of export manufacturing in *maquiladoras* on employment and the relative wages of skilled and unskilled workers. See, for example, Feenstra and Hanson (1997).

22. To isolate regional wage differentials that are specific to the Apparel industry, the data on wages in Apparel sector in each state relative to Mexico City are normalised by average manufacturing wages in each state relative to Mexico City. Similar estimation results are found using un-normalized wages. See Hanson (1996) for further discussion.

23. Even in the absence of underlying technology differences, measured aggregate Total Factor Productivity (TFP) may vary substantially across locations due to differences in the transport cost inclusive price of manufacturing inputs and output. Cross-country differences in measured productivity may partly reflect true underlying technology differences and partly reflect the considerations of access to markets and sources of supply emphasized above.

24. We are following Kim (1995), who followed Hoover (1936). Amiti (1999), points out the similarity to Balassa's (1965) measure of revealed comparative advantage.

25. The calculation of $gini^k(r_i^k)$ follows a method directly analogous to that used in the personal income distribution literature. First, rank regions by their location quotients in

descending order. Second, evaluate the cumulative percentage of employment in industry *k* over the regions ($cemp^k$). Third, evaluate the cumulative percentage of employment in all industries over the regions ($cemp_t^k$). Graphing $cemp^k$ (y-axis) against $cemp_t^k$ (x-axis), we obtain a 'localisation curve.' The area between the 45 degree line and the localisation curve divided by the area under the 45 degree line is the locational gini.

26. See Amiti (1998) for discussion and comparison of some of these measures.

27. The exogenous case corresponding to an Armington model; the difference is immaterial as $\sigma \rightarrow \infty$ and all varieties are perfect substitutes.

28. For a complete derivation of the home market effect, see Fujita, Krugman, and Venables (1999), Chapter 4. See also the discussion in Krugman (1980), Krugman and Venables (1990), Davis (1999), and Krugman and Venables (2001). External economies of scale provide an alternative candidate explanation for home market effects and are discussed in further detail below (see Markusen (1990) for an analysis of the micro-foundations of external economies).

29. The exponent on distance, γ, is found by estimating a gravity model like equation (8).

30. r_i^k is the location quotient of equation (14) up to the normalisation $\sum_k \sum_i y_i^k = 1$

References

Acemoglu D, S. Johnson, and J. Robinson (2000) 'The Colonial Origins of Comparative Development: An Empirical Investigation', *NBER Working Paper*, 7771.

Amiti, M (1998) 'New Trade Theories and Industrial Location in the EU: A Survey of Evidence' *Oxford Review of Economic Policy,* Vol 14 (2) 45-53.

Amiti, M (1999) 'Specialisation Patterns in Europe', *Weltwirtschaftliches Archiv*, 134(4), 573-93.

Anderson, J and S. Marcouiller, (2001) 'Insecurity and the Pattern of Trade: An Empirical Investigation', *Review of Economics and Statistics*, forthcoming.

Anderson J.E. and E. Van Wincoop,. (2000) 'Gravity with gravitas: a solution to the border puzzle', mimeo, Boston University.

Antweiler, W. and D. Trefler (2000), 'Increasing returns and all that; a view from trade', *NBER Working Paper*, 7941.

Balassa, B. (1965) 'Trade Liberalisation and 'Revealed Comparative Advantage', *The Manchester School of Economic and Social Sciences'* 33, 99-123.

Bairoch, P (1988) *Cities and Economic Development: From the Dawn of History to the Present,* University of Chicago Press.

Basu, S and J. Fernald (1997) 'Returns to Scale in US Production: Estimates and Implications', *Journal of Political Economy*, 105(2), 249-83

Branstetter, L (2001) 'Are Knowledge Spillovers International or Intra-national in Scope? Micro-econometric Evidence from the US and Japan', *Journal of International Economics*, 53(1), 53-79.

Brulhart, M (1998) 'Economic Geography, Industry Location, and Trade: The Evidence', *World Economy*, 21(6), 775-801.

Brulhart, M. and J.Torstensson, (1996) 'Regional Integration, Scale Economies and Industry Location', *CEPR Discussion Paper,* 1435.

Caballero, R and R. Lyons (1990) 'Internal Versus External Economies in European Industry', *European Economic Review,* 34(4), 805-30.

Caballero, R and R. Lyons (1992) 'External Effects in US Pro-cyclical Productivity', *Journal of Monetary Economics*, 29(2), 209-25.

Ciccone, A and R. Hall (1996) 'Productivity and the Density of Economic Activity', *American Economic Review*, 86(1), pages 54-70.

Ciccone, A and G. Peri (2000) 'Human Capital and Externalities in Cities', *CEPR Discussion Paper*, 2599.

Clark, C, F. Wilson, and J. Bradley (1969) 'Industrial Location and Economic Potential in Western Europe', *Regional Studies*, 3, 197-212.

Coe, D and E. Helpman (1995) 'International R&D Spillovers', *European Economic Review*, 39(5), 859-87.

Combes, P.P. (2001) 'Marshall-Arrow-Romer Externalities and City Growth', CNRS - CERAS, Paris, mimeograph.

Cowell, F (1995) *Measuring Inequality,* London School of Economics Handbooks in Economics Series, 2nd Edition, Prentice-Hall.

Davis, D.R. (1999) 'The Home Market, Trade, and Industrial Structure', *American Economic Review*, 88(5), 1264-76.

Davis, D and D. Weinstein (1998) 'Market Access, Economic Geography, and Comparative Advantage: An Empirical Assessment', *NBER Working Paper*, 6787.

Davis, D and D. Weinstein (1999) 'Economic Geography and Regional Production Structure: An Empirical Investigation', *European Economic Review*, 43, 379-407.

Deardorff (1998) 'Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?' Chapter 1 in (ed) Frankel, J, *The Regionalization of the World Economy*, NBER & Chicago University Press.

Dekle, R and J. Eaton (1999) 'Agglomeration and Land Rents: Evidence from the Prefectures', *Journal of Urban Economics*, 46, 200-14.

Devereux, M, R. Griffith, and H. Simpson (1999) 'The Geographic Distribution of Productive Activity in the UK', *Institute for Fiscal Studies Working Paper,* W99/26.

Devereux, M, R.Griffith, and H.Simpson (2001) 'The Geography of Firm Formation', Institute for Fiscal Studies, London, mimeo.

Dicken, P (1998), *Global Shift: Transforming the World Economy*, Chapman, London.

Dicken, P and P. Lloyd (1977) *Location in Space,* New York: Harper and Row.

Dixit,A.K. and J.E. Stiglitz, (1977), 'Monopolistic competition and optimum product diversity', *American Economic Review*, 67, 297-308.

Dumais, G., G. Ellison, and E. Glaeser (1997) ‘Geographic Concentration as a Dynamic Process’ *NBER Working Paper* 6270.

Duranton, G. and H.G. Overman (2001) ‘Localisation in UK Manufacturing Industries: Assessing Non-Randomness using Micro-Geographic Data’ processed, London School of Economics.

Eaton, J and S. Kortum,(1999a) ‘International Technology Diffusion: Theory and Measurement’, *International Economic Review,* 40(3), 537-70.

Eaton, J and S. Kortum, (1999b) ‘Technology, Geography, and Trade’, processed, Boston University.

Ellison, G and E. Glaeser (1997) ‘Geographic Concentration in US Manufacturing Industries: A Dartboard Approach’, *Journal of Political Economy*, 105(5), 889-927.

Ellison, G and E. Glaeser, (1999) ‘The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?’, *American Economic Review*, 89(2), Papers and Proceedings, 311-6.

Feenstra, R and G. Hanson (1997) ‘Foreign Direct Investment and Relative Wages: Evidence from Mexico’s Maquiladoras’, *Journal of International Economics*, 42, 371-94.

Feenstra, R, R. Lipsey, and H. Bowen, (1997), ‘World Trade Flows, 1970-1992, with Production and Tariff Data’ *NBER Working Paper* 5910.

Feenstra, R.C., J.A. Markusen and A.K. Rose (2001) ‘Understanding the Home Market Effect and the Gravity Equation: The Role of Differentiated Goods’, *Canadian Journal of Economics*, forthcoming.

Frankel, J.A. (1997), *Regional trading blocs in the world economic system*, Institute for International Economics, Washington DC

Frankel, J and D. Romer (1999) ‘Does Trade Cause Growth?’, *American Economic Review*, 89(3), 379-99.

Fujita, M, P. Krugman and A.J. Venables (1999) *The spatial economy: cities, regions and international trade*, MIT Press, Cambridge MA.

Gallup, J and J. Sachs (2000) ‘The Economic Burden of Malaria’, *CID Working Paper*, No. 52, Harvard University.

Gallup, J.L. and J. Sachs (with A.D. Mellinger), (1999) ‘Geography and economic development’, in B. Pleskovic and J.E. Stiglitz (eds), *Annual World Bank Conference* *on Development Economics, 1998*, World Bank, Washington DC.

Glaeser, E and D. Mare (1994) 'Cities and Skills', *NBER Working Paper*, 4728.

Glaeser, E, H. Kallal, J. Scheinkman, and A.. Shleifer (1992) 'Growth in Cities', *Journal of Political Economy*, 100, 1126-52.

Haaland, J, H. Kind, K. Midelfart-Knarvik, and J. Torstensson, (1999) 'What Determines the Economic Geography of Europe?', *CEPR Discussion Paper*, 2072.

Hall, R and C. Jones (1999) 'Why Do Some Countries Produce So Much More Output Per Worker Than Others?', *Quarterly Journal of Economics*, 114(1) 83-116.

Hanson, G (1996) 'Localization Economies, Vertical Organization, and Trade', *American Economic Review*, 86(5), 1266-78.

Hanson G (1997) 'Increasing Returns, Trade, and the Regional Structure of Wages', *Economic Journal*, 107, 113-133.

Hanson G (1998a) 'Market Potential, Increasing Returns, and Geographic Concentration', *NBER Working Paper*, 6429.

Hanson G (1998b) 'Regional Adjustment to Trade Liberalisation', *Regional Science and Urban Economics*, 28(4), 419-44.

Hanson, G (1998c) 'North American Economic Integration and Industry Location', *Oxford Review of Economic Policy*, 14(2), 30-44.

Hanson, G (2000a) 'Market Potential, Increasing Returns, and Geographic Concentration', University of Michigan, mimeograph, November. Revised version of *NBER Working Paper*, 6249, February 1998.

Hanson, G (2000b) 'Scale economies and the geographic concentration of industry', *NBER Working Paper*, 8013.

Harris, C (1954) 'The Market as a Factor in the Localization of Industry in the United States', *Annals of the Association of American Geographers*, 44, 315-48.

Head, K. and J. Ries, (2001) 'Increasing returns versus national product differentiation as an explanation for the pattern of US-Canada trade', Forthcoming, *American Economic Review*.

Helliwell, J (1996) 'Do National Borders Matters for Quebec's Trade?', *Canadian Journal of Economics*, 29(3), 507-22.

Helpman, E (1998) 'The Size of Regions', in (eds) D. Pines, E. Sadka, and I. Zilcha, *Topics in*

Public Economics, Cambridge: Cambridge University Press.

Henderson, J. V. (1988), *Urban Development: Theory, Fact and Illusion*, Oxford University Press.

Henderson, J. V. (1999) 'Marshall's Scale Economies' *NBER Working Paper* no 7358.

Henderson, J. V. , A. Kuncoro, and M. Turner,(1995) 'Industrial Development and Cities', *Journal of Political Economy*, 103, 1067-81.

Henderson, J. V, T. Lee and J-Y Lee (1999), 'Externalities and Industrial deconcentration under Rapid Growth,' mimeo, Brown University.

Hoover E. (1936) 'Spatial Price Discrimination' in Greenhut, M and Norman, G eds *The Economics of Location. Volume 2: Space and Value'* Edward Elgar 1995.

Hummels, D (1995) 'Global Income Patterns: Does Geography Play a Role?', Chapter 2 of Phd Thesis, University of Michigan.

Hummels, D. (1999a), 'Towards a geography of trade costs', mimeo, Purdue University.

Hummels, D. (1999b), 'Have international transport costs declined', mimeo, Purdue University.

Hummels, D. (2000), 'Time as a trade barrier', mimeo, Purdue University.

Jaffe, A and M. Trajtenberg, (1998) 'International Knowledge Flows: Evidence from Patent Citations', *NBER Working Paper*, 6507.

Jaffe, A, M. Trajtenberg, and R. Henderson, (1993) 'Geographic Localisation of Knowledge Spillovers as Evidenced by Patent Citations', *Quarterly Journal of Economics*, 63(3), 577-98.

Kahn, M (1995) 'A Revealed Preference Approach to Ranking City Quality of Life', *Journal of Urban Economics*, 38, 221-35.

Keeble, D, P. Owens, and C. Thompson, (1982) 'Regional Accessibility and Economic Potential in the European Community', *Regional Studies*, 16, 419-32.

Keller, W (1998) 'Are International R&D Spillovers Trade-related? Analyzing Spillovers among Randomly Matched Trade Partners,' *European Economic Review*, 1469-1481.

Keller, W (2000) 'Do Trade Patterns and Technology Flows Affect Productivity Growth,' *World Bank Economic Review*, 17-47.

Keller, W (2001) 'Geographic Localization of International Technology Diffusion', *American Economic Review*, forthcoming.

Kim, S (1995) 'Expansion of Markets and the Geographic Distribution of Economic Activities: the Trends in US Regional Manufacturing Structure, 1860-1987', *Quarterly Journal of Economics*, 110, 881-908.

Kim, S (1997) 'Regions, Resources, and Economic Geography: Sources of US Regional Comparative Advantage, 1880-1987', *NBER Working Paper*, 6322.

Knack, S and P. Keefer, (1997) 'Does Social Capital Have an Economic Payoff?', *Quarterly Journal of Economics*, 112, 1251-88.

Krugman, P (1980) 'Scale Economies, Product Differentiation, and the Pattern of Trade', *American Economic Review*, 70, 950-9.

Krugman, P (1991a) 'Increasing Returns and Economic Geography', *Journal of Political Economy*, 99(3), 483-99.

Krugman, P (1991b) *Geography and Trade*, MIT Press.

Krugman, P and A.J.Venables (1990) 'Integration and the Competitiveness of the Peripheral Industry', in C. Bliss and J Braga De Macedo,(eds), *Unity With Diversity in the European Economy: The Community's Southern Frontier*, Cambridge University Press: Cambridge.

Krugman, P and A.J.Venables,(1995) 'Globalization and the Inequality of Nations', *Quarterly Journal of Economics*, 857-880.

Krugman, P and A.J.Venables (2001) 'How robust is the home market effect?' processed, LSE.

Leamer, E (1988) 'Measures of Openness', in (ed) Baldwin, R, Trade Policy Issues and Empirical Analysis, Chicago: University of Chicago Press, 147-200.

Leamer, E (1997) 'Access to Western Markets and Eastern Effort', in (ed) Zecchini, S, *Lessons from the Economic Transition, Central and Eastern Europe in the 1990s,* Dordrecht: Kluwer Academic Publishers, 503-26.

Leamer, E and J. Levinsohn, (1995) 'International Trade Theory: The Evidence', in (eds) G. Grossman and K. Rogoff, (eds) *Handbook of International Economics,* vol 3

Leamer, E. and M. Storper, (2000) 'The economic geography of the internet age', processed UCLA

Limao N. and A.J. Venables (2001), 'Infrastructure, geographical disadvantage, transport costs and trade', *World Bank Economic Review*, forthcoming.

McCallum, J (1995) 'National Borders Matter: Canada-US Regional Trade Patterns', *American*

Economic Review, 85(3), 615-23.

Markusen, J (1990) 'Micro-foundations of External Economies', *Canadian Journal of Economics*, August, 495-508.

Maurel, F and B. Sedillot, (1999) 'A Measure of the Geographic Concentration in French Manufacturing Industries' *Regional Science and Urban Economics*, 29 (5) 575-604.

McArthur, J and J. Sachs, (2001) 'Institutions and Geography: Comment on Acemoglu, Johnson, and Robinson (2000)', *NBER Working Paper*, 8114.

Midelfart-Knarvik, K, H.G. Overman, S.J. Redding and A.J. Venables, (2000) 'The location of European industry', *Economic Papers no. 142*, European Commission, D-G for Economic and Financial Affairs, Brussels.

Midelfart-Knarvik, K, H.G. Overman, and A.J. Venables, (2000) 'Comparative Advantage and Economic Geography', *CEPR Discussion Paper*, 2618.

Moomaw, R (1981) 'Productivity and City Size: A Critique of the Evidence', *Quarterly Journal of Economics*, 96(4), 675-88.

Moretti, E (1998) 'Social Returns to Education and Human Capital Externalities: Evidence from Cities', *Centre for Labour Economics, Working Paper*, no 9, University of California, Berkeley.

Neary, P (2001) 'Of Hype and Hyperbolas: Introducing the New Economic Geography', *Journal of Economic Literature*, 39, June, 536-61.

Ottaviano, G and Puga, D (1997) 'Agglomeration in the Global economy: A Survey of the New Economic Geography', *World Economy*, 21(6), 707-31.

Radelet, S and J. Sachs, (1998) 'Shipping Costs, Manufactured Exports, and Economic Growth', paper presented at the American Economic Association Meetings, Harvard University, mimeo.

Rauch, J (1993) 'Productivity Gains from Geographical Concentration of Human Capital: Evidence from Cities', *Journal of Urban Economics*, 34, 380-400.

Rauch, J and V. Trindade (1999) 'Ethnic Chinese Networks in International Trade', *NBER Working Paper*, 7189.

Redding, S. and A.J. Venables, (2000) 'Economic Geography and International Inequality', *CEPR Discussion Paper*, 2568.

Roback, J (1982) 'Wages, Rents, and the Quality of Life', *Journal of Political Economy*, 90, 1257-88.

Roberts, M and J.Tybout (1997) 'The Decision to Export in Columbia: An Empirical Model of Entry with Sunk Costs', *American Economic Review*, September, 545-64.

Sachs, J and A. Warner (1999) 'The Big Push, Natural Resource Booms and Growth', *Journal of Development Economics*, 59(1), 43-76.

Sachs, J (2000) "Tropical Underdevelopment", Center for International Development, Harvard University, September, mimeograph.

Segal, D (1976) 'Are There Returns to Scale in City Size?', *Review of Economics and Statistics*, 58(3), 339-50.

Soloaga, I and A. Winters (1999) 'Regionalism in the Nineties: What Effect on Trade?', *CEPR Discussion Paper*, 2183.

Sveikauskas, L (1975) 'The Productivity of Cities', *Quarterly Journal of Economics*, 89(3), 393-413.

Venables, A.J. (1996) 'Equilibrium Locations of Vertically Linked Industries', *International Economic Review*, 37, 341-59.

Venables, A.J. (1999) 'The International Division of Industries: Clustering and Comparative Advantage in a Multi-industry Model', *Scandanavian Journal of Economics*, 495-513.

Venables, A.J. (2001) 'Geography and international inequalities; the impact of new technologies', forthcoming discussion paper, Center for Economic Performance, LSE.

Venables, A and N. Limao, (1999) 'Geographical Disadvantage: A Heckscher-Ohlin-Von Thunen Model of International Specialisation, *CEPR Discussion Paper*, 2305.

Von Thunen, J.H. (1826) *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie*, Hamburg, (English translation by C.M. Wartenberg. *Von Thunen's Isolated State*, Oxford, Pergamon press, 1966).

Wei, S (1996) 'Intra-national versus International Trade: How Stubborn are Nations in Global Integration?', *NBER Working Paper*, 5531.

WIFO (1999), 'Specialisation and (geographic) concentration of European manufacturing', background paper for 'The competitiveness of European Industry; the 1999 report', EC Enterprise Directorate-General, working paper no 1, Brussels.

Williamson, J.G. (1965), 'Regional Inequality and the Process of National Development,' *Economic Development and Cultural Change*, 13, 3-45.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

507	A. J. Venables	Geography and International Inequalities: the Impact of New Technologies
506	R. Dickens D. T. Ellwood	Whither Poverty in Great Britain and the United States? The Determinants of Changing Poverty and Whether Work Will Work
505	M. Ghell	Fixed-Term Contracts and the Duration Distribution of Unemployment
504	A. Charlwood	Influences on Trade Union Organising Effectiveness in Great Britain
503	D. Marsden S. French K. Kubo	Does Performance Pay De-Motivate, and Does It Matter?
502	S. Nickell L. Nunziata W. Ochel G. Quintini	The Beveridge Curve, Unemployment and Wages in the OECD from the 1960s to the 1990s
501	S. Redding M. Vera-Martin	Factor Endowments and Production in European Regions
500	Edited by D. Marsden and H. Stephenson	Labour Law and Social Insurance in the New Economy: A Debate on the Supiot Report
499	A. Manning	A Generalised Model of Monopsony
498	A. Charlwood	Why Do Non-Union Employees Want to Unionise? Evidence from Britain
497	M. Keil D. Robertson J. Symons	Minimum Wages and Employment
496	A. Di Liberto J. Symons	Education and Italian Regional Development
495	S. Redding A. J. Venables	Economic Geography and International Inequality

494	A. Bryson	Union Effects on Managerial and Employee Perceptions of Employee Relations in Britain
493	D. Metcalf	British Unions: Dissolution or Resurgence Revisited
492	R. Gomez S. M. Lipset N. Meltz	Frustrated Demand for Unionisation: the Case of the United States and Canada Revisited
491	S. Burgess J. Lane D. Stevens	Jobs, Workers and Changes in Earnings Dispersion
490	S. Burgess S. Profit	Externalities in the Matching of Workers and Firms in Britain
489	S. Nickell G. Quintini	Nominal Wage Rigidity and the Rate of Inflation
488	S. Nickell J. Van Reenen	Technological Innovation and Performance in the United Kingdom
487	M. M. Tudela	Explaining Currency Crises: A Duration Model Approach
486	D. Sturm	Product Standards, Trade Disputes and Protectionism
485	G. Duranton V. Monastiriotis	Mind the Gaps: The Evolution of Regional Inequalities in the UK 1982-1997
484	H. G. Overman Y. Ioannides	Zipfs Law for Cities: An Empirical Examination
483	H. G. Overman Y. Ioannides	Cross Sectional Evolution of the US City Size Distribution
482	Y. Ioannides H. G. Overman	Spatial Evolution of the US Urban System
481	H. G. Overman	Neighbourhood Effects in Small Neighbourhoods
480	S. Gomulka	Pension Problems and Reforms in the Czech Republic, Hungary, Poland and Romania