DISCUSSION PAPER SERIES

No. 3455

MARKET POTENTIAL AND THE LOCATION OF JAPANESE INVESTMENT IN THE EUROPEAN UNION

Keith Head and Thierry Mayer

INTERNATIONAL TRADE

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP3455.asp**

ISSN 0265-8003

MARKET POTENTIAL AND THE LOCATION OF JAPANESE INVESTMENT IN THE EUROPEAN UNION

Keith Head, University of British Columbia
Thierry Mayer, Université de Paris I, CEPII, CERAS and CEPR

Discussion Paper No. 3455
July 2002

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3455

July 2002

ABSTRACT

Market Potential and the Location of Japanese Investment in the European Union*

We investigate the hypothesis that firms prefer to locate 'where the markets are'. We use a theoretical model of location choice under imperfect competition to formalise this concept. The model yields an equilibrium profit equation incorporating a term closely connected to the market potential index introduced by Harris in 1954. The location decision is a function of demand in all locations weighted by accessibility to consumers. We also show how the spatial distribution of competitors should be factored into the location choice. We then implement the model empirically, comparing our theoretically-derived measure of market potential with Harris' term and with *ad hoc* agglomeration variables. Our sample consists of firm-level location choices by Japanese firms between 1984-95 and we use both the information on the choice of country and the choice of region inside each country in our analysis. Our results show that market potential does matter for location choice but that traditional agglomeration variables retain an important role in the location decision.

JEL Classification: F12 and F15
Keywords: borders, European Union and location choice

Keith Head
University of British Columbia
Faculty of Commerce
2053 Main Hall
Vancouver BC, V6T 1Z2
CANADA
Tel: (1 604) 822 8492
Fax: (1 604) 822 8477
Email: keith.head@ubc.ca

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=154019

Thierry Mayer
TEAM
Université de Paris I
106-112 Boulevard de l'Hôpital
F-75647 Paris Cedex 13
FRANCE
Tel: (33 1) 44 07 82 67
Fax: (33 1) 44 07 82 67
Email: tmayer@univ-paris1.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=152707

*This Paper is produced as part of a CEPR Research Network on 'The Economic Geography of Europe: Measurement, Testing and Policy Simulations', funded by the European Commission under the Research Training Network Programme (Contract No: HPRN-CT-2000-00069).

We would like to thank participants at the CEPR Workshop 'Economic Geography of Europe: Measurement, Testing and Policy Simulations' Villars, 12–13 January 2001 and at the seventh 'Empirical Investigations in International Trade' Annual Conference, 3–5 November 2000, University of Colorado, Boulder. Participants at seminars in Catholic University of Leuven (KUL), 15 March 2001, University of Paris XII, 19 March 2001, University of Paris XIII 29 June 2000, and University of Geneva - HEI, 10 May 2001 are also to be thanked. Richard Baldwin, Bruce Blonigen, Pierre-Philippe Combes, Matthieu Crozet, Gianmarco Ottaviano and Jean Marc Siroën provided very helpful comments on early versions for which we are particularly grateful.

Submitted 02 June 2002

1 Introduction

"We want to build our plants where the markets are."

In December 1997, Hiroshi Okuda, chairman of Toyota, used the quote above to justify Toyota's decision to build a factory in northern France. At that time, analysts in the press largely attributed that decision to the low market share of the Japanese car manufacturer in France (1%) and in Europe in general (3%).[1] The Toyota example suggests that even in the zero-tariff internal market of Europe, firms still seek locations with superior market access. The managers of Toyota apparently thought that their existing production sites in Great Britain were not "close" enough to the French market. The *Wall Street Journal* reported that "Toyota ... hopes to capture 3% [of the French market] after opening its factory here in 2001."

This paper connects two previously disparate strands of the economic geography literature. The first strand demonstrates a statistical tendency of firms to make the same location decision as previous firms with similar attributes (such as industry and national origin).[2] While such "agglomeration effects" appear regularly in empirical work, they are consistent with a variety of explanations. The second strand comprises a large number of theoretical papers that focus on a particular mechanism of agglomeration: namely, that producers concentrate where demand is highest and serve smaller markets via exporting.[3]

We link the two strands by showing how to derive the firm's location choice probabilities as a function of production costs and a demand variable closely linked to the measure of "market potential" introduced by Harris (1954). We then take the model to the data, investigating whether location choices of Japanese affiliates in Europe are driven by market-access motivations *a la* Krugman or some other form of agglomeration effect. We find the demand-pull mechanism has some explanatory power but it does not appear to explain away the entire empirical agglomeration effect.

Three recent papers have examined the impact of market potential on economic geography. Hanson (1998) estimates the relationship between county-level wages in the United States and

[1] Douglas Lavin in the *Interactive Edition*, December 10, 1997 and *AP-Dow Jones News Service*, December 9, 1997.

[2] Devereux and Griffith (1998) and Head *et al.*, (1999) are recent examples from this literature.

[3] Krugman (1980 and 1991) wrote the seminal papers in this literature; the monograph of Fujita, Krugman, and Venables (1999) thoroughly analyzes the basic model and its extensions.

a market potential term based on Krugman (1992). Structural estimation of this equation reveals that wages in a county are increasing in demand emanating from all American counties weighted by the bilateral distance. Redding and Venables (2000) follow the same line of reasoning for international data using the gravity-like bilateral trade equation of the same theoretical model to obtain estimates of bilateral transport costs and of each country's market and supply accessibility. They find that international inequality is closely linked to market access. Crozet (2000) uses the same theoretical framework but focuses on the migration flows generated by real wage differentials in this model. He estimates the core parameters of the model in an equation relating migration to price index differentials for 5 European countries and finds (as did Hanson) values that are consistent both with the constraints of the theory and the previous findings in the literature concerning price elasticities and the impact of distance on trade flows (see Hummels, 1999 and Head and Ries, 2001 for instance).

The literature on firm location choice has not previously estimated models directly derived from the Krugman model. Prior work has, of course, considered demand but typically only local demand.[4] Knowing the size of demand in each of the districts a firm might choose is not sufficient since firms can export to nearby locations. Some studies such as Friedman et al. (1992), Henderson et al. (1995), and Head et al. (1999) consider non-local demand but not using measures derived from theory.[5] In particular, theory suggests that non-local demand must be discounted based on bilateral trade impediments. Furthermore, a given amount of market access contributes less to profits when a firm's competitors have access to the same markets. We follow Krugman (1992) in adjusting the market potential measure to take into account the location of competitors.

The paper proceeds as follows. In section 2 we derive a linear-in-logs equation that relates the profitability of a location to a prospective foreign investor to a measure of that location's access to demand. We then show how to estimate the distance and border effects that impede market access using bilateral trade data. In section 3 we report the results from the trade equation and show how we use them to calculate market potential. We then discuss our sample of Japanese investors and the set of possible location choices. Our location-choice results are

[4]See for instance, Coughlin et al. (1991).

[5]Friedman et al. (1992) use a distance weighted sum of per-capita incomes, Henderson et al. (1995) use distance to nearest major business district, and Head et al. (1999) sum the personal incomes of adjacent states.

detailed in section 4 and we conclude and propose directions for future work in section 5.

2 The Model

We present a model derived from the widely used Dixit and Stiglitz (1977) monopolistic competition framework, applied by Krugman (1980) to international trade. We demonstrate that the profitability of investing in a given country for a foreign affiliate is a function of its market potential. We then show how to estimate the parameters of that function using the trade equation implied by the same model.

In each country, consumers have Cobb-Douglas utility functions in terms of two sets of differentiated products, namely those produced by European firms and those produced by foreign-owned affiliates. Within each set there is a common constant elasticity of substitution, σ, between symmetric varieties. Each foreign affiliate chooses a location from a choice set of N European countries. Home firms produce n_j varieties and foreign affiliates produce n^*_j varieties in each country j. The Cobb Douglas functional form implies that consumers spend fixed shares of their income, Y_j, on each set of products. We denote the share spent on foreign affiliate products as α. Symmetry in the sub-utility functions implies equal expenditure on each variety from a given country j.

2.1 The Profit Equation for Foreign Affiliates

The resulting demand curve of consumers in country i for a representative foreign variety (affiliate) from country j is given by

$$q^*_{ij} = \frac{p^{*-\sigma}_{ij}}{\sum_k n^*_k p^{*(1-\sigma)}_{ik}} \alpha Y_i, \tag{1}$$

where p^*_{ij} is the delivered price faced by consumers in country i for products from j. It is the product of the mill price p^*_j and the iceberg trade cost, τ_{ij}. Trade costs include all transaction costs associated with moving goods across space and national borders. The marginal production cost in each country is denoted c_j. Increasing returns come from a plant-specific fixed cost F. Each foreign affiliate maximizes the following gross profit function for each market:

$\pi^*_{ij} = (p^*_j - c_j)\tau_{ij} q^*_{ij}$. The resulting mill prices are simple mark-ups over marginal costs:

$$p^*_j = \frac{c_j \sigma}{\sigma - 1}.$$

Substituting into (1), we obtain the equilibrium quantity the affiliate delivers to each market:

$$q^*_{ij} = \frac{(\sigma - 1)}{\sigma} \frac{(c_j \tau_{ij})^{-\sigma}}{C^*_s} \alpha Y_i, \tag{2}$$

where $C^*_s \equiv \sum_k n^*_k (c_k \tau_{sk})^{1-\sigma}$. The gross profit earned in each market i for an affiliate located in country j is thus:

$$\pi^*_{ij} = \frac{(c_j \tau_{ij})^{1-\sigma}}{\sigma C^*_s} \alpha Y_i \tag{3}$$

This gross profit is an increasing function of the expenditure of country i on the good we consider. The magnitude of the fraction multiplying Y_i depends on costs between this representative firm and all its competitors. At the numerator, we see that profits are decreasing in own production costs. The profits are also higher with a better access to country i. The denominator contains the characteristics of each affiliate's foreign competitors. Note also that the denominator term is multiplied by σ, which underlines the fact that with high values of the elasticity of substitution (which is also the price elasticity of demand here), competition is fiercer and profits are therefore lower.

Summing the profits earned in each market and subtracting the plant-specific fixed cost, we obtain the net profit to be earned in each potential location j:

$$\Pi^*_j = \frac{\alpha}{\sigma} (c_j)^{1-\sigma} \sum_{s=1}^{N} \frac{\tau_{sj}^{1-\sigma} Y_s}{C^*_s} - F \tag{4}$$

When an affiliate chooses its location, the only relevant information is the ordering of those profit functions. We are therefore able to make monotonic transformations in order to obtain an additive expression for the profitability of each location. Specifically, we add F, multiply by σ/α, and take logs, yielding

$$V^*_j \equiv \ln[(\sigma/\alpha)(\Pi^*_j + F)] = -(\sigma - 1) \ln c_j + \ln M^*_j, \tag{5}$$

where

$$M_j^* \equiv \sum_s \frac{\tau_{sj}^{1-\sigma} Y_s}{C_s^*}.$$

We will refer to M_j^* as the "Krugman market potential" since it first appears in Krugman (1992).

Equation (5) expresses the profitability for a firm of locating in country j as a decreasing function of production costs and an increasing function of market potential.[6]

The Krugman market potential aggregates the expenditures of all countries while adjusting for country j's access $\tau_{sj}^{1-\sigma}$, and for competition from affiliates located in other countries, C_s^*. Considering only the numerator of M_j^*, we can recover the original Harris (1954) formulation of market potential. Specifically, if we set $C_s^* = 1$ and $\tau_{sj}^{1-\sigma} = 1/d_{sj}$, M_j^* reduces to $\sum_s Y_s/d_{sj}$, i.e. the inverse-distance weighted sum of incomes. In the Krugman market potential function, M_j^*, the denominator, C_s^* takes into the account the competition that affiliates from country j face from rival affiliates in other countries for the demand in each country s. The competition adjustment increases with the number of rivals and decreases with their trade and production costs.

The Krugman market potential has the advantage of being derived rigorously from theory. However, unlike the Harris form, its calculation requires estimates of the unknown parameters τ_{ij} and σ. We use the trade equation derived from the same model to estimate those parameters.

2.2 The Trade Equation

The value of country i's imports of European varieties from country j, denoted m_{ij}, is given by the quantity exported by a representative variety firm from j multiplied by the price and

[6]The profitability equation is closely related to the wage equation analyzed by Fujita et al. (1999, p. 53). The assumption of free entry sets equation (4) equal to zero. Specifying production to require F units of labour as overhead and one unit of labour per unit of output, wages are given by

$$w_j = \left[\frac{\alpha}{\sigma F} M_j^*\right]^{1/\sigma}.$$

That is, for firms to be indifferent between locations, wages must be a power function of market potential. Hanson (1998) and Redding and Venables (2000) estimate variations of this relationship.

the number of European varieties from j:

$$m_{ij} = p_{ij} q_{ij} n_j = \frac{n_j p_{ij}^{1-\sigma}}{\sum_k n_k p_{ik}^{1-\sigma}} (1-\alpha) Y_i.$$

We will obtain a linear in logs equation by dividing each bilateral trade flow by the internal trade of the importing country and then taking logs. The resulting equation, also used by Head and Mayer (2000), is given by

$$\ln\left(\frac{m_{ij}}{m_{ii}}\right) = \ln\left(\frac{n_j}{n_i}\right) - (\sigma - 1)\ln\left(\frac{p_{ij}}{p_{ii}}\right). \tag{6}$$

As we do not have data on the number of European varieties in each country we rely on the theory to establish a proportional relationship between the number of firms and the value of production, $v_j = p_j q_j n_j$. If European firms operate in a zero profit equilibrium, then $(p_j - c_j)q_j = F$. The Dixit-Stiglitz markup rule implies $(p_j - c_j) = p_j/\sigma$. Hence, $v_j = F\sigma n_j$, and the ratio of the number of varieties can be replaced with the ratio of the value of production.

The delivered price in country i for a product made in country j takes the form $p_{ij} = \tau_{ij} p_j$. Substituting these expressions into (6) yields

$$\ln\left(\frac{m_{ij}}{m_{ii}}\right) - \ln\left(\frac{v_j}{v_i}\right) = -(\sigma - 1)\ln\left(\frac{p_j}{p_i}\right) - (\sigma - 1)\ln\left(\frac{\tau_{ij}}{\tau_{ii}}\right). \tag{7}$$

Relative trade costs (τ_{ij}/τ_{ii}) will be estimated as a linear in logs function of relative distance (d_{ij}/d_{ii}), border effects (β) and common language ($L_{ij} = 1$ if i and j share a language and 0 otherwise) and an error term, ϵ_{ij}. The estimated equation will therefore be

$$\ln\left(\frac{m_{ij}}{m_{ii}}\right) - \ln\left(\frac{v_j}{v_i}\right) = -(\sigma - 1)\ln\left(\frac{p_j}{p_i}\right) - \beta - \delta\ln\left(\frac{d_{ij}}{d_{ii}}\right) + \lambda L_{ij} + \epsilon_{ij}. \tag{8}$$

The estimated parameters $(\hat{\sigma}, \hat{\beta}, \hat{\delta}, \hat{\lambda})$ are then used to construct the market potential variable that will be included in the location choice analysis of Japanese firms in Europe. The formulas are

$$\tau_{ij}^{1-\sigma} = e^{-\hat{\beta} + \hat{\lambda} L_{ij}} d_{ij}^{-\hat{\delta}} \text{ for } i \neq j \text{ and } \tau_{ii}^{1-\sigma} = d_{ii}^{-\hat{\delta}}.$$

3 Econometric model and data

We estimate a model of location choice of 452 Japanese-owned affiliates that were established in 57 regions belonging to 9 European countries (Belgium, France, Germany, Ireland, Italy, the Netherlands, Spain, Portugal and the United Kingdom) during the period 1984–1995. As can be seen in equation (5), we hypothesize that market potential is a central component of this decision. We construct the market potential for 18 industries using parameters estimated from the trade equation (8).

3.1 Estimation of the Trade Equation

We estimate (8) using Eurostat data on bilateral trade matched with production and price indexes at the NACE70 2-digit level. Our sample runs from 1980 to 1995 and includes all twelve pre-1995 members of the European Union. Internal trade flows, m_{ii}, subtract multilateral exports in an industry from national production. Both internal and external distances are weighted averages of point-to-point distances between sub-national regions. Head and Mayer (2000) provide the complete distance matrix and greater detail on its construction. A dummy variable identifies late entrants (Greece, Portugal, and Spain) prior to their respective years of accession. The common language variable, L_{ij} takes a value of one for the U.K. and Ireland, Belgium and France, and Belgium and the Netherlands. As Head and Mayer (2000) show a pattern of declining trade costs, we divide our sample into two periods, before (1980–87) and after (1988–95) the start of the Single Market Programme.

Table 1 provides the border and distance effect estimates for each 2-digit industry. We find distance effects that average -1.04 across the two periods. This number aligns closely with the results of other estimates of gravity equations and suggests Harris's assumption of an inverse distance rule is a good approximation. Where Harris' specification appears inappropriate for Europe is in its omission of the impact of national borders.[7] Border effects average 2.12 in the first period and 1.85 in the later period. Expressing their magnitude in the McCallum (1995) manner, within-border trade after 1987 remains more than six (exp(1.85)) times as large as cross-border trade after controlling for the effect of relative distance and economy size. While sizeable, these effects are considerably smaller than the value of 20 first reported by McCallum

[7] His pioneering study considered the market potential of counties *within* the United States.

Table 1: Estimates of Border and Distance Effects

Industry Name	NACE Code	No. of Obs.	1980–1987 Border ($\hat{\beta}$)	1980–1987 Dist. ($-\hat{\delta}$)	1988–1995 Border ($\hat{\beta}$)	1988–1995 Dist. ($-\hat{\delta}$)
Metal-Primary	22	513	1.68	-.8	1.24	-1.18
Non-metallic Mineral Products	24	720	2.84	-1.3	2.77	-1.25
Chemicals & Fibres	25/26	648	1.95	-.79	1.9	-.74
Metal-Fabricated	31	720	3.17	-1.03	2.96	-1.12
Machinery	32	621	1.84	-.61	1.72	-.63
Office Machines	33	342	.67	-.73	.55	-.45
Electronics	34	702	2.06	-.82	2.02	-.62
Motor Vehicles and Parts	35	576	1.45	-1.22	1.39	-1.05
Cycles	363	448	2.47	-.87	1.96	-.86
Precision Instruments	37	522	.7	-1.43	-.09$^{\bullet}$	-1.63
Food, Drink, & Tobacco	41/42	720	2.87	-1.19	2.81	-1.22
Textiles	43	720	2.95	-1.00	2.73	-.82
Leather	44	531	2.09	-1.02	1.20	-1.62
Clothing and Footwear	45	630	2.74	-.79	2.46	-.8
Wood and Wooden Furniture	46	720	3.15	-1.12	3.18	-1.10
Paper, Printing, & Publishing	47	720	2.82	-1.43	2.75	-1.34
Rubber and Plastics	48	675	1.75	-1.13	2.01	-1.00
Toys & Sports	494	423	1.01	-1.04	-.11$^{\bullet}$	-1.84

Note: All coefficients are significantly different from zero at the 10% level except those with designated with $^{\bullet}$.

for the Canada-US border. In contrast with distance effects, there appears to be a downward trend in the impact of national borders in the E.U.; all but two border effects declined. We also find large common language effects. Country pairs sharing a language trade over three times as much as pairs lacking a common language.

The other parameter we use to calculate market potential is the estimate of $\sigma - 1$ obtained from the coefficient on relative prices. Here we find values averaging 1.1, i.e. $\hat{\sigma} = 2.1$. This magnitude, while in line with the usual estimates of price elasticities in trade equations, falls short of the values used in the simulations of Fujita et al (1999) as well as estimates based on tariff changes found by Head and Ries (2001).

3.2 Specification of the Location Choice Model

We estimate the parameters of the profit equations (5) using a discrete choice model. As we do not observe the potential profitability of each location, we rely upon the assumption that firms choose the country yielding the highest profit. The location choice literature makes extensive use of the conditional logit model (CLM) proposed by Daniel McFadden. This model requires error terms that are independent across locations. As it seems likely that the unobserved component of profitability is correlated among regions in the same nation, we use a generalisation, also due to McFadden, that permits such a structure of the error term, the nested logit model (NLM).[8]

The firm's expected profitability of region r is given by $V_r^* + V_c^* + \varepsilon_r$ where V_r^* contains the determinants from equation (5) that vary across regions in a given country c, and V_c^* contains the determinants that have no intra-national variation. McFadden (1978) shows that, when the distribution of the error term, ε_r, is given by a multivariate extreme value, the conditional probability that firms choose region r out of a set $\mathcal{B}_c$ of alternatives in country c is $\exp(V_r^* - I_c)$, where $I_c \equiv \ln \sum_{r \in \mathcal{B}_c} \exp(V_r^*)$ is termed the "inclusive value" for country c.

The probability of choosing country c is $\exp(V_c^* + \rho I_c - \hat{I})$, where

$$\hat{I} \equiv \ln[\sum_c \exp(V_c^* + \rho I_c)]. \tag{9}$$

Parameter ρ measures the relative strength of correlations within and between subsets. A value of $0 < \rho < 1$ implies that cross elasticities are largest in magnitude for alternative regions belonging to the same country compared to regions belonging to different countries. For $\rho = 0$, regions are perfect substitutes inside a country whereas for $\rho = 1$ patterns of substitution are the same within and between countries (and the NLM collapses to the CLM).

We also consider the un-nested version of the model corresponding to $\rho = 1$. In that case the probability of choosing a region r is $\exp(V_c^* + V_r^* - \hat{I})$.

[8]Precise descriptions of the nested logit model methodology can be found in Maddala (1983) or Train (1993). Hansen (1987) provides an early application of the nested logit model to location choice. Devereux and Griffith (1998) also use this model nesting location choices inside an export/FDI choice of American firms in Europe and Mayer and Mucchielli (2002) use the same sample used here to test for the validity of a structure nesting a location choice among regions beneath a location choice among nations.

3.3 Implementation of the location choice model

The sample of Japanese firms is extracted from the 1996 *Survey of Current Manufacturing Operations of Japanese Firms in Europe* issued by the JETRO. More than 700 manufacturing Japanese investments are listed in this survey with corresponding date when operation started, country of location, employment and other details including a detailed description of the product manufactured. In order to assign investments to sub-national regions, the *Directory of Japanese-Affiliated Companies in the EU: 1996–1997* also issued by the JETRO was used to determine the precise city where the plant was located. Almost all explanatory variables come from industrial statistics issued by Eurostat either at the national or at the regional level. The selection of Japanese investments was essentially driven by the availability of regional and national data. Further details concerning the data can be found in the data appendix as well as a descriptive statistics table of different variables used.

Figure 1 plots the Japanese affiliates in the NUTS1 region where they invested. Several important features of Japanese investment patterns are immediately apparent: The strong attractiveness of the United Kingdom as a whole, the agglomeration in the Northern part of Europe, as well as a tendency of investors to locate in the economic core of each country (Japanese investors cluster in London in the United Kingdom, Paris in France, Milan in Italy, and Barcelona in Spain).

Recall that the profit equation is given by

$$V_j^* = -(\sigma - 1)\ln c_j + \ln M_j^*.$$

We consider three different specifications of production costs, c_j. The first follows Krugman (1992) in assuming that costs are proportional to wages. The second adds typical labour market and fiscal determinants of production costs. The "full" specification incorporates two proxies for agglomeration effects into the cost function. Here we intend to capture the possibility that clustering leads to direct economic benefits such as access to workers with specialized skills or knowledge sharing between competitors. Those three specifications are described in table 2.

The second type of determinant consists of the market potential term: $M_j^* = \sum_s (\tau_{sj}^{1-\sigma} Y_s)/(C_s^*)$. The Y_s term is calculated by combining data at the national and regional level. We first com-

Figure 1: Japanese investors in Europe at the end of 1995

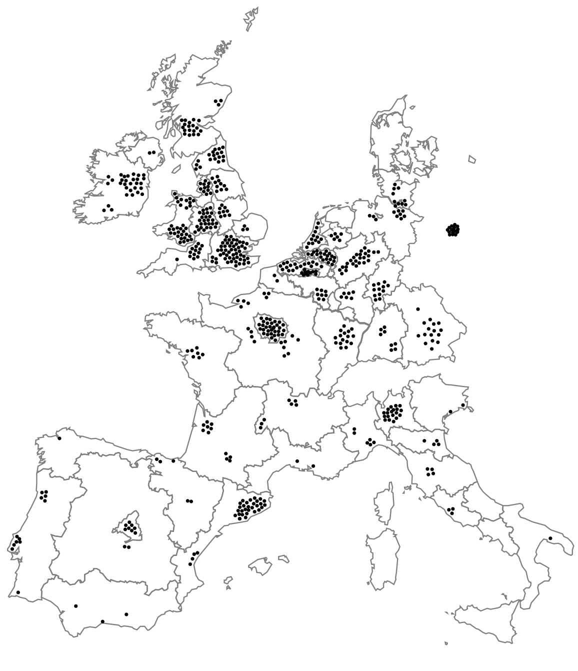


Table 2: Production costs specifications

Specification	Variable	Definition
Basic	Wages	Total wage bill divided by number of employees in the 2-digit industry-region
Extended	Unemployment	Percent of the labour force defined as unemployed at the regional level
	Subsidy eligibility	A dummy variable set equal to one when the per capita income of the region is less than 75% of the EU average, the critical value for qualifying for European Union structural funds
	Corporate tax	Corporate tax receipts divided by gross operating income at the national level
	Social charges	Non-wage labour costs such as payroll taxes and pension contributions divided by the number of employees in the 2-digit industry-country
Full	Affiliate agglomeration	1+ count of Japanese affiliates in the 3-digit industry-region
	Domestic agglomeration	Share of European firms in the 2-digit industry-region[9]

pute industry-level national expenditure as apparent consumption (value of production − exports + imports). We allocate this expenditure across regions in proportion to their shares of national GDP. Accessibility to this expenditure ($\tau_{sj}^{1-\sigma}$) is calculated from the trade equation as described at the end of section 2.2. In calculating the competition component of the market potential, C_s^*, we employ our basic specification in which c is proportional to wages. This variable must be raised to the $(1-\sigma)$ according to the theory. We obtain our estimates of σ from the coefficient on the relative price variable in equation (8). In line with the assumption that Japanese goods enter a separable sub-utility function, we measure n_k using the count of Japan-owned affiliates in country k in the same industry. Figure 2 displays the cross-regional variation of the market potential (M_j^*) calculated in 1990 for four industries that are representative of our sample of investors.

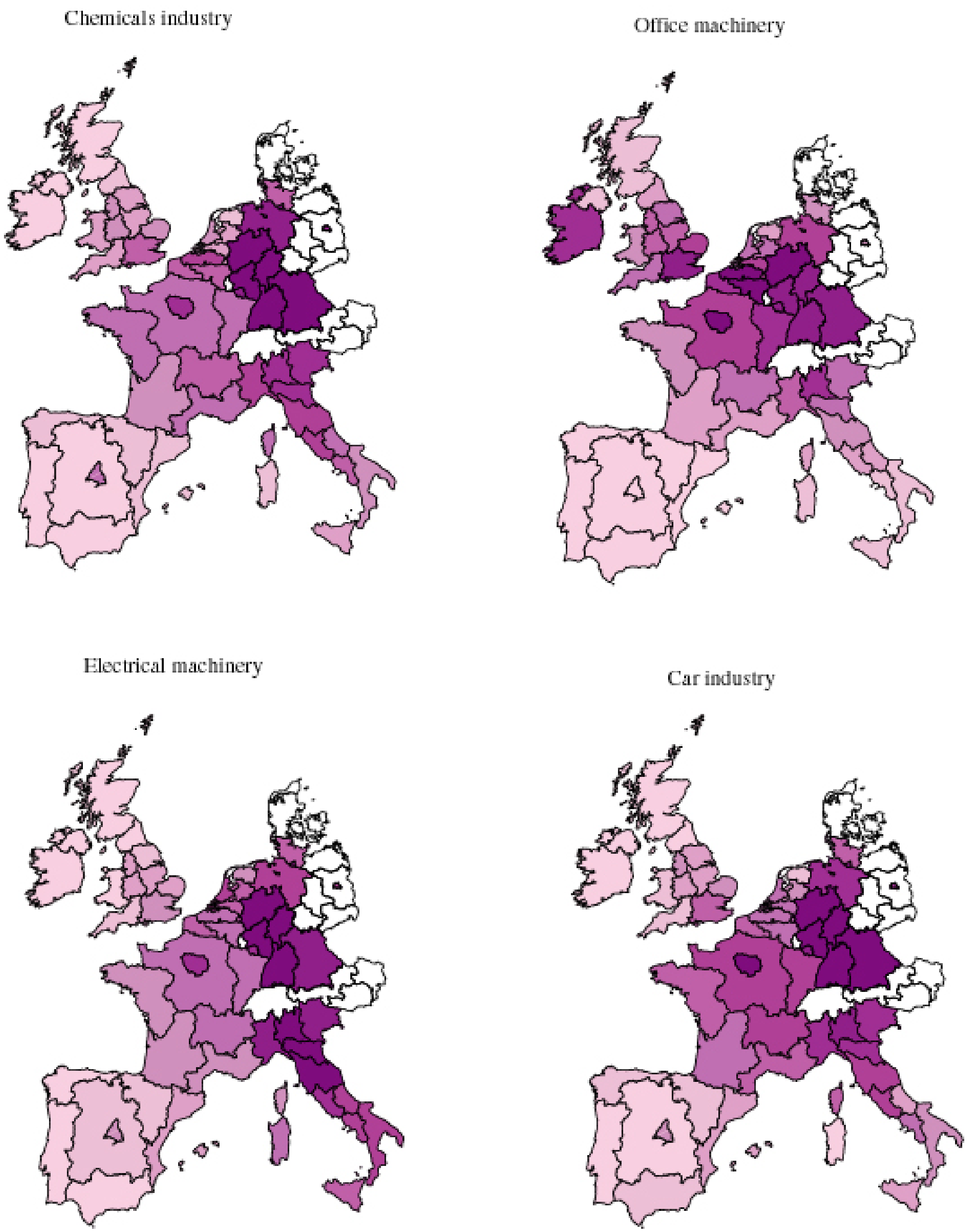
4 Location choice results

We begin with an assessment of the performance of our market potential variable in a conventional specification used in the literature. We therefore start with conditional logit estimation of the region (NUTS1) choices of Japanese firms in Europe. We then turn to a nested logit specification in which we first estimate choice of region within a given nation and then estimate the choice of nation taking into account the attractiveness of its constituent regions.

Table 3 presents the results of the conditional logit regional level choice. The first three columns move from our basic (column 1) to extended (column 2) and then full (column 3) specifications. All specifications in table 3 include—but do not report—nation-level dummy variables to control for a variety of omitted factors that differ across countries (language, legal system, employment regulations, etc.). In all three specifications, market potential has a statistically and economically significant effect on location choice. The average probability elasticity with respect to market potential is just under 0.53 (since there are 57 choices). This means that a 10% increase in market potential raises the probability a region will be selected by over 5%.

The insignificance of regional wages as a determinant of location choice is disappointing in light of the model's predictions. Nonetheless, the result is not out of line with other

Figure 2: Market Potentials ($\mathcal{P}_s^M$) in Europe in 1990



studies.[10] Since the standard model of wage determination (the Mincer equation) explains differences in wages with differences in human capital (education, experience, and ability) that are presumably valuable to the firm, ambiguous results should perhaps be expected. Unemployment enters negatively suggesting it is a better measure of labour market rigidities than labour availability. Policy variables generally lack statistical significance (although the national fixed effects are often significantly different from each other). Regions eligible for Objective One subsidies are also low-income and these effects seem to mainly cancel each other out. The social charge rate contains variation in labour costs that is unrelated to human capital and enters with a consistently negative sign, albeit insignificantly.

Column (3) introduces region-level agglomeration effects. While past work on location choice would lead to a strong prior that they would enter positively, the current paper is the first to control for market potential using the exact functional form dictated by theory. Thus, if prior findings of agglomeration effects merely reflected the absence of adequate controls for variation in demand, the agglomeration terms would *not* enter significantly in our specification. The data decisively reject the hypothesis of zero agglomeration effects, despite the presence of the Krugman market potential term. We interpret this result—which is robust to changes in econometric technique and formulation of market potential—as indicating that agglomeration arises for reasons independent of the forward linkages emphasized in most of the New Economic Geography literature.

In column (4) we investigate the extent that our results are driven by use of the Krugman market potential and its associated adjustments for differential distance and border and competition effects. This specification uses the Harris formulation of market potential. This has two consequences that are relevant for interpreting the new estimates. First, these results do not depend on our trade equation estimates. Second, there is no adjustment in the denominator to take into account where the competitors are. The change in specification leads to a slight reduction in estimated agglomeration effects, perhaps because they now include offsetting competition effects that previously entered the market potential term. The finding of a higher coefficient on the Harris term should not be overly interpreted since the confidence intervals overlap and the nested logit specification obtains a *lower* coefficient on the Harris form.

[10]See, for instance, Head et al (1999).

In columns (5) and (6) we follow Redding and Venables (2001) in decomposing the Krugman market potential into terms from the *chosen* region and market potentials of *other* regions. The latter variable, which we refer to as non-local market potential, is given by

$$\sum_{i \neq s} \tau_{ij}^{1-\sigma} \frac{Y_i}{C_i^*}.$$

One key advantage of such an approach is that it does not require an ad hoc specification for the distance of a region from itself. Instead the coefficient on $\ln Y_r$ incorporates an estimate of the average internal distance. The unappealing aspect of this approach is that is not founded in theory. The two-part specification of market potential improves the likelihood ratio index and suggests that local demand has a surprisingly strong influence on location choice.

While column (5) implicitly assumes that all intra-regional distances are equal, column (6) specifies only that internal distance be a power function of the square root of the region's area. To the extent that area is a proxy for internal distance, it should enter negatively since it implies larger transport costs for serving a given demand level. On the contrary, we find a positive effect. One obvious interpretation is that higher area conditional on region GDP indicates relative abundance of land for the construction of new factories. That is, area belongs in the cost function as well as in the internal distance function and the former effect outweighs the latter.

The use of the (unreported) country dummies in Table 3 helps to mitigate the problem associated with non-independent errors across regions belonging to the same nation.[11] However, the country fixed effects do not resolve problems associated with cross-industry and inter-temporal differences in the attractiveness of nations. By considering the choice of region for a given choice of nation, we condition on all aspects of the nation that do not vary across its constituent regions from the perspective of a given investor. The drawback is that we must omit the national tax and social charges variables. However, we reintroduce them when estimating the upper level of the decision tree (nation choice).

Table 4 reestimates the same specifications as table 3 in a nested structure. The results are broadly similar and we comment here only on the noteworthy differences. First, the Krugman market potential has twice as high an impact on choice of region (even after taking into account

[11]See Train (1993) for a detailed explanation.

Table 3: Conditional Logit Model of Region Choice

	Dependent Variable: Region r choice					
Model :	(1)	(2)	(3)	(4)	(5)	(6)
ln Krugman market potential	0.77^a	0.57^a	0.53^a			
$\sum_s(\tau_{rs}^{1-\sigma}Y_s)/(C_s^*)$	(0.16)	(0.18)	(0.18)			
ln Harris market potential				0.91^a		
$\sum_s Y_s/d_{rs}$				(0.22)		
ln non-local market potential					0.35^c	0.45^b
$\sum_{s\neq r}(\tau_{rs}^{1-\sigma}Y_s)/(C_s^*)$					(0.20)	(0.21)
ln regional GDP					0.75^a	0.69^a
Y_r					(0.09)	(0.09)
ln $\sqrt{\text{area}_r}$						0.33^b
						(0.14)
ln Wages	0.31	0.45	0.12	0.01	-0.31	-0.06
	(0.29)	(0.34)	(0.34)	(0.34)	(0.34)	(0.36)
Unemployment rate		-9.82^a	-8.58^a	-6.83^a	-2.71	-2.58
		(2.20)	(2.26)	(2.33)	(2.37)	(2.38)
Subsidy eligibility (Obj 1)		-0.20	-0.26	-0.25	-0.21	-0.21
		(0.23)	(0.23)	(0.23)	(0.24)	(0.25)
ln 1+ social charges		-1.67	-1.70	-1.62	-1.64	-1.70
		(2.44)	(2.51)	(2.50)	(2.50)	(2.50)
ln 1+ corporate taxes		-1.12	0.67	1.87	1.43	1.06
		(3.30)	(3.33)	(3.30)	(3.32)	(3.33)
ln 1+ Jpn affiliates			1.11^a	1.03^a	0.85^a	0.86^a
			(0.13)	(0.13)	(0.14)	(0.14)
ln domestic firms share			0.49^a	0.42^a	0.48^a	0.50^a
			(0.10)	(0.10)	(0.10)	(0.10)
N	25764	24808	24762	24762	24762	24762
Likelihood Ratio Index	0.047	0.06	0.089	0.091	0.108	0.11

Note: Standard errors in parentheses with ^{a}, ^{b} and ^{c} respectively denoting significance at the 1%, 5% and 10% levels. Country dummies are included in the estimation but not reported here.

that there is an average of about 6 regions per nation).[12] Using Harris market potential still slightly reduces the agglomeration coefficients but in nested logit it also leads to a lower market potential effect. The lower likelihood ratio index in column (4) suggests a small preference for the Krugman formulation. However, it is possible to improve substantially upon the fit of the Krugman market potential by decomposing it into the regional and non-local components, as we do in columns (5) and (6). Comparing nested and conditional logit results for (6), the preferred specification, we find higher market potential effects and lower agglomeration effects in the nested model. However, non-local market potential is only marginally significant. Moreover, it is possible that the strong effect of regional GDP in column (6) reflects the more diverse economic base of larger local economies and is thus an indicator of "urbanization economies" in the cost function rather than market access.

The lower frame of table 4 reports estimates for the choice of nation. This specification directly considers national variables such as corporate tax rates but indirectly considers all the regional determinants of attractiveness as they enter the inclusive value for each nation. We do not include country dummies because there would not be be sufficient remaining variation in the data. However, we do include a dummy variable for English-speaking countries (the United Kingdom and Ireland) since anecdotal accounts of Japanese investment patterns often claim this feature is important.

We find consistently negative impacts of social charges but, as was also the case in the conditional logit region choice, the effect is statistically insignificant. Taxes are insignificant in the preferred specifications. The English language dummy is positive and statistically significant in all specifications. The magnitude of this effect indicates that use of English raises the odds of being chosen by a Japanese investor by a factor of 2.7 (= exp(1.0)) in the preferred specification.

The inclusive value, an index of the maximum expected profitability from locating in a given country considering the underlying characteristics of its regions, obtains reasonable values in all specifications, differing quite significantly from the value of one that would make conditional logit appropriate and the value of zero in which investors are indifferent between

[12] The probability elasticity is given by $\hat{b}(1-P_i)$ where $\hat{b}$ is the coefficient and P_i is the probability of choosing location i. On average, P_i is the inverse of the number of choices. Thus, using column (3) results we have $1.26(1-1/6) > 2(0.53)(1-1/57)$.

Table 4: Nested Logit Model of Region Choice

Model :	(1)	(2)	(3)	(4)	(5)	(6)
	Dep. Var.: Region r given country c choice					
ln Krugman market potential	1.76[a]	1.31[a]	1.26[a]			
$\sum_s(\tau_{rs}^{1-\sigma}Y_s)/(C_s^*)$	(0.25)	(0.27)	(0.28)			
ln Harris market potential				0.83[a]		
$\sum_s Y_s/d_{rs}$				(0.25)		
ln non-local market potential					0.40	0.64[c]
$\sum_{s\neq r}(\tau_{rs}^{1-\sigma}Y_s)/(C_s^*)$					(0.31)	(0.33)
ln regional GDP					0.82[a]	0.76[a]
Y_r					(0.10)	(0.10)
ln $\sqrt{\text{area}_r}$						0.33[b]
						(0.15)
ln Wages	0.34	-0.02	-0.13	0.05	-0.73	-0.31
	(0.46)	(0.48)	(0.48)	(0.49)	(0.50)	(0.53)
Unemployment rate		-7.07[a]	-5.58[b]	-6.75[b]	-1.75	-0.70
		(2.55)	(2.63)	(2.62)	(2.80)	(2.85)
Subsidy eligibility (Obj 1)		-0.68[b]	-0.63[c]	-0.65[c]	-0.63[c]	-0.65[c]
		(0.33)	(0.34)	(0.34)	(0.35)	(0.35)
ln 1+ Jpn affiliates			0.98[a]	0.92[a]	0.65[a]	0.66[a]
			(0.15)	(0.15)	(0.15)	(0.15)
ln Domestic firms share			0.43[a]	0.39[a]	0.42[a]	0.44[a]
			(0.14)	(0.14)	(0.14)	(0.14)
N	3882	3882	3882	3882	3882	3882
Likelihood Ratio Index	0.041	0.051	0.083	0.078	0.116	0.119
	Dependent Variable: Country c choice					
ln 1+ social charges	-1.24[c]	-1.13	-0.85	-0.46	-0.66	-0.71
	(0.70)	(0.70)	(0.71)	(0.71)	(0.74)	(0.74)
ln 1+ corporate taxes	5.21[a]	4.20[b]	3.02[c]	0.30	-0.05	0.55
	(1.66)	(1.75)	(1.72)	(1.89)	(1.93)	(1.88)
English-speaking	0.72[a]	0.76[a]	0.84[a]	0.94[a]	0.97[a]	1.00[a]
	(0.20)	(0.21)	(0.20)	(0.21)	(0.22)	(0.22)
Inclusive Value	0.31[a]	0.30[a]	0.38[a]	0.52[a]	0.56[a]	0.50[a]
	(0.05)	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)
N	3544	3544	3536	3536	3536	3536
Likelihood Ratio Index	0.102	0.101	0.114	0.12	0.132	0.131

Note: Standard errors in parentheses with [a], [b] and [c] respectively denoting significance at the 1%, 5% and 10% levels. The inclusive value is defined in equation 9.

regions inside a given country. In other words, the coefficient on the inclusive value supports the validity of our country-region nesting structure.

5 Conclusion

We analyze the determinants of location choices by Japanese firms in Europe during the 1984–1995 period. Our work is particularly concerned with the appropriate manner to take into account the spatial distribution of demand and competition in the location choice. We rigourously link the optimal location choice of Japanese investors to a theoretical model of imperfect competition in a multi-location setting. The underlying profit equation incorporates a term that is closely related to the market potential index originally introduced and used by geographers (Harris, 1954). This term aggregates the spatial distribution of demand weighted by transaction costs and the location of potential competitors. We estimate the border and distance effects that determine market accessibility using a bilateral trade equation implied by the same model that generates the profit equation.

We find that demand does matter for location choice: A ten percent increase in our market potential term raises the chance of a region being chosen by five to ten percent, depending on the specification. Despite the fact that we bring theory to empirical implementation in a structural way, traditional agglomeration variables retain an important influence. This suggests that the downstream linkages emphasized in Krugman (1991) are not the only or even the main cause of agglomeration. Future research should probably consider other reasons why firms cluster. It does not seem possible to falsify the hypothesis that observed agglomeration effects merely reflect omitted exogenous location attributes. However, a natural follow-up to this paper would be to consider "second-generation" economic geography models that implement the Venables (1996) setup with upstream and downstream linkages based on an input-output matrix.

References

Coughlin, C.C., J.V. Terza and V. Arromdee, 1991, "State Characteristics and the

Location of Foreign Direct Investment in the United States " *Review of Economics and Statistics* 68 : 675-683.

Crozet, M., 2000, "Do migrants believe in market potential? " Cahier de la MSE Série Blanche Université de Paris I # 2000-30.

Devereux, M. and R. Griffith, 1998, "Taxes and the Location of Production : Evidence from a Panel of US Multinationals" *Journal of Public Economics* 68(3) : 335-367.

Dixit, A. and J. Stiglitz, 1977, "Monopolistic Competition and Optimum Product Diversity" *American Economic Review.* 67(3) : 297–308

Friedman, J., D. Gerlowski and J. Silberman, 1992, "What Attracts Foreign Multinational Corporations" *Journal of Regional Science* 32(4) : 403-418.

Fujita , M., P. Krugman and A. Venables, 1999, *The Spatial Economy: Cities, Regions, and International Trade* Cambridge : MIT Press.

Hansen, E.R., 1987, "Industrial Location Choice in Sao Paulo, Brazil, a Nested Logit Model" *Regional Science and Urban Economics* 17(1) : 89-108.

Hanson, G., 1998, "Market Potential, Increasing Returns, and Geographic Concentration" NBER Working Paper #6429.

Harris, C. , 1954, "The Market as a Factor in the Localization of Industry in the United States" *Annals of the Association of American Geographers* 64 : 315-348

Head, K. and T. Mayer, 2000, "Non-Europe : The Magnitude and Causes of Market Fragmentation in Europe" *Weltwirtschaftliches Archiv* 136(2): 285-314.

Head, K. and J. Ries, 2001, "Increasing Returns Versus National Product Differentiation as an Explanation for the Pattern of US-Canada Trade," *American Economic Review*, 91(4) : 858-876.

Head, K., J. Ries and D. Swenson , 1999, "Attracting Foreign Manufacturing: Investment Promotion and Agglomeration" *Regional Science and Urban Economics* 29(2) : 197-218.

Henderson, J.V., A. Kuncoro and M. Turner, 1995, "Industrial Development in Cities" *Journal of Political Economy* 103(5) : 1067-1090.

Hummels, D., 1999, "Toward a Geography of Trade Costs" University of Chicago manuscript.

Krugman, P., 1980, "Scale Economies, Product Differentiation, and the Pattern of Trade" *American Economic Review* 70:950–959.

Krugman, P.R., 1991, "Increasing Returns and Economic Geography" *Journal of Political Economy* 99(3) : 483-499.

Krugman, P.R. , 1992, "A Dynamic Spatial Model" National Bureau of Economic Research Working Paper #4219.

Maddala, G.S., 1983, *Limited-Dependent and Qualitative Variables in Econometrics* Cambridge : Cambridge University Press

Mayer T. and J-L. Mucchielli, 2002, "Hierarchical Location Choice and Multinational Firms' Strategy" in Dunning J. and J-L. Mucchielli (eds.), *Multinational Firms, The Global and Local Dilemma*, pp.133-158. London: Routledge.

McCallum, J., 1995, "National Borders Matter: Canada-US Regional Trade Patterns," *American Economic Review* 85:615–23.

McFadden, D., 1978, "Modelling the Choice of Residential Location", in A. Karlquist et al. (eds), *Spatial Interaction Theory and Residential Location.*, pp. 75-96. Amsterdam: North-Holland.

Redding, Stephen and Anthony J. Venables, 2000, "Economic geography and International Inequality", Center for Economic Policy Research Discussion Paper # 2568.

Train, K., 1993, *Qualitative Choice Analysis : Theory, Econometrics, and an Application to Automobile Demand* Cambridge :MIT Press 2 ed.

Venables, A.J., 1996, "Equilibrium Locations of Vertically Linked Industries" *International Economic Review* 37(2) : 341-359.

Data Appendix

Trade equation estimation

The regressions and data used to estimate the negative impact of distance and borders on trade flows follow closely the ones in Head and Mayer (2000) where more details can be found on the data used. All data used in estimating equation (8) come from Eurostat databases: The COMEXT database for bilateral trade flows and the VISA database for production and prices. Production values are adjusted in order to account for the fact that most countries reported data only for firms larger than twenty employees in the years we are considering. The main difference in this paper is that the sectoral level of aggregation is NACE 2-digit instead of NACE 3-digit in order to match with the level of detail of the subsequent location choice estimation. All regressions take into account the fact that there are some combinations of country pairs at the industry level where flows are zero by using the Heckman two-step procedure.

Distances calculations are crucial in this paper for both the trade equation and profit equation estimations. We calculate the distance of one nation to another—or itself—as weighted averages of sub-national distances. Considering two countries O and D (the origin and destination countries of a given flow), respectively consisting of regions indexed $i \in O$ and $j \in D$, the following formula provides both external and internal distances.

$$d_{OD} = \sum_{i \in O} (\sum_{j \in D} w_j d_{ij}) w_i$$

We define d_{ij} as the distance between the centers of regions i and j and w_i as the weight of region i, calculated as the share of two-digit industry-level employment for origin weights and GDP for destination weights.

The distance of a region to itself is obtained using a simple geographical approximation. Each region is approximated as a disk in which all production concentrates in the center and consumers are uniformly distributed throughout the rest of the area. Then the average distance between a producer and a consumer is given by

$$d_{ii} = \int_0^R r \cdot 2r/R^2 dr,$$

where R denotes the radius of the disk, and $2r/R^2$ is the density of consumers at any given distance r to the center. We obtain R as the square root of the area, A, divided by π. Integrating, we obtain $d_{ii} = (2/3)R = (2/3)\sqrt{A/\pi} = .376\sqrt{A}$.

Location choice estimation

Regions and years used

The regional level choice sets incorporate 57 regions in Europe using NUTS 1 level of detail for Germany (11 regions), France (8 regions), Italy (11 regions), the UK (11 regions), Spain (6 regions), The Netherlands (4 regions), and Belgium (4 regions). Ireland and Portugal are considered as single region countries. Out of those 57 regions, 50 where chosen at least once by Japanese investors.

Industry-level regional data availability limits the sample to the years 1984–1995. Although there are some Japanese investments in the late seventies and early eighties, the vast majority of the investments took place in the late eighties.

Affiliates

The location choices of Japanese affiliates are mainly extracted from JETRO's *Survey of Current Manufacturing Operations of Japanese Firms in Europe*, 1996. This source provides in particular the country chosen and the date at which operations started for the 727 manufacturing affiliates which had established operations in Europe by the end of 1995.

We then identified for each firm the city in which the production unit was located. This information appears in a larger document also issued by JETRO: The Directory of Japanese-Affiliated Companies in the EU: 1996-1997. This directory lists a total of 2988 companies in the EU for all industries.

A crucial matter for our study is the quality of this information: It had to be checked that, in the directory, the affiliate's location reported was not the headquarters but the actual production(s) unit(s). Fortunately, the directory almost always specifies both the location of the headquarters and the location of the plant. However, the information was double checked using three alternative sources: The database used in Yamawaki et al. (1996), mostly using data from Toyo Keisai and kindly made available by Hideki Yamawaki was of great help. The table 5.4 in Strange (1992) also confirmed the locations of Japanese subsidiaries in the U.K. Finally a document from the DATAR helped to check locations for France.

Market potential calculation

The main data needed for calculating market potential consists of apparent consumption (production value + imports − exports) in a given industry for each European region. Unfortunately, trade flows are not available at the regional level in Europe, which makes it impossible to directly calculate apparent consumption in a given industry-region combination.

To get a regional industry-level apparent consumption, we allocate the national apparent consumption of each industry between its regions according to the share of national GDP.

The main data issue for national apparent consumption computation is the availability of production value data. The member states might want, essentially for confidentiality reasons, to aggregate some NACE 2-digit industries figures. There are also some industries for which certain member countries give the figure for some years and not for the others. In the later case, we replace the industry-country figure by an estimation of the most proximate year for which data was available. If the data is missing for all years then we take the relevant aggregate. Often, countries give several aggregates for an industry, we systematically take the one for which the figure is lower and non missing.

After that procedure we still have countries which do not have any data in important industries: There are no production figures for Belgium and the Netherlands in the chemicals industry (Nace 25). However, all countries report production for chemicals aggregated with artificial fibers (Nace 2601 = 25 + 26). We thus take this aggregate for both Nace 25 and Nace 26. Another problem is related to Netherlands protecting data on its 3-digit aerospace industry which leaves us with no data for the 2-digit industry "Manufacture of other means of transport" in Netherlands. However, there is only a single subsector in this industry which is of some interest for us: Nace 363 (cycles). As the data is quite complete for this 3-digit industry, we replace Nace 36 with the data of Nace 363 for all countries. We choose to do the same for NACE 49, miscellaneous industries, which in our sample reduces to NACE 494: Toys and sports for which the data is rather complete.

Note finally that even though Denmark is not used in the final estimation of location choice, it is included in the market potential calculation in order to account for the fact that some regions in Germany have their market potential enhanced by Danish demand more than others.

Industry-level regional data

The main source of industry-level regional data comes from Eurostat publication *Structure and activity of industry annual inquiry, principal results, regional data*. It consists of 2-digit Nace data, essentially available for NUTS 1 regions. This database contains number of establishments, employment and the wage bill for each industry-region combination. For single-region countries, national level data is used.

An electronic version exists with regional data for the years 1989 to 1992 but in fact 1992 has many missing values. We additionally used the printed version for years 1984 and 1987. Observations for years 1984 to 1986 are matched with the 1984 data. Observations for years 1987 and 1988 are matched with the 1987 data. 1989, 1990 and 1991 observations are matched with same-year data. Observations from 1992 to 1995 are matched with 1991 data.

NACE 26 (Man-made fibres industry) was excluded of the sample because too few data was available.

When the data was missing for a particular NUTS 1 region, the following procedure was adopted: The missing values are often due to missing values in NUTS 2 small sub-regions (for instance Corsica in Mditérranée or Val d'Aoste in Nord Ovest have many missing employment and wages values because there are only one or two firms. In this case we just sum up the remaining NUTS 2 regions to get what appears to be a very precise approximation of the true data. In other cases too many data from sub-regions was missing for a particular year, we then replaced the figure with its value for the nearest year available. As a general pattern, the main problems in data availability concerned Netherlands and even more Belgium.

Table 5: Descriptive Statistics of European regions. Average values 1988–1991

Region	Code	# of Jpn. Inv.	Wage (000 ECUs)	Unemp. rate	GDP (Bn ECUs)
France					
Ile De France	FR1	21	24.2	7.6	254.98
Bassin Parisien	FR2	16	18.3	9.4	148.77
Nord-Pas-De-Calais	FR3	2	17.7	12.3	50.29
Est	FR4	16	17.8	7.3	73.66
Ouest	FR5	6	16.9	9.4	99.15
Sud-Ouest	FR6	8	18.7	9.5	81.66
Centre-Est	FR7	6	19.1	8.2	101.15
Mediterranee	FR8	2	20.1	11.9	90.8
Belgium					
Brabant	BE0	4	17.6	6.5	24.94
Brussels	BE1	2	18.2	9.3	22.07
Vlaams Gewest	BE2	11	17.2	5.7	86.87
Region Wallonne	BE3	5	17.1	10.2	38.48
Netherlands					
Noord-Nederland	NL1	0	19.8	10.1	22.79
Oost-Nederland	NL2	4	20.4	8	38
Zuid-Nederland	NL4	14	20.9	7.3	45.26
West-Nederland	NL3	9	21.9	7.3	109.3
Germany					
Schleswig-Holstein	DEF	3	21.8	6.1	39.67
Hamburg	DE6	2	25.5	8.1	45.91
Niedersachsen	DE9	9	22.4	6.7	112.41
Bremen	DE5	0	23	10	15.42
Nordrhein-Westfalen	DEA	18	24.6	7	291.91
Hessen	DE7	7	24.5	4.1	121.73
Rheinland-Pfalz	DEB	6	24.9	4.5	58.02
Baden-Wuerttemberg	DE1	6	25.7	3	190.22
Bayern	DE2	11	24	3.5	207.88
Saarland	DEC	0	22.5	7.7	17.26
Berlin	DE3	1	23.6	7.3	40.72

Table 6: Descriptive Statistics of European regions (Cont'd)

Region	**Code**	**# of Jpn. Inv.**	**Wage (000 ECUs)**	**Unemp. rate**	**GDP (Bn ECUs)**
Italy					
Nord Ovest	IT1	6	27.7	6.4	102.47
Lombardia	IT2	15	28.1	3.5	167
Nord Est	IT3	2	25.3	4.5	106.93
Emilia-Romagna	IT4	3	26.4	4.5	70.22
Centro (I)	IT5	4	26.2	7.1	87.02
Lazio	IT6	2	27	9.8	82.58
Campania	IT8	0	16	19.8	55.82
Abruzzi-Molise	IT7	0	16.1	9.6	19.5
Sud	IT9	1	23.3	16.7	64.63
Sicilia	ITA	0	17.3	19.2	48.47
Sardegna	ITB	0	17.3	17.1	17.43
United Kingdom					
North	UK1	17	17.6	11.3	35.85
Yorkshire & Humberside	UK2	8	16.7	9.1	59.69
East Midlands	UK3	7	17	6.9	51.09
East Anglia	UK4	1	18.6	5.1	27.47
South East	UK5	43	19.9	5.8	277.76
South West	UK6	11	17.6	6	59.01
West Midlands	UK7	26	17.1	8.3	64.16
North West	UK8	10	18.9	10	76.91
Wales	UK9	25	16.4	9	32.35
Scotland	UKA	17	18.3	10.9	64.2
Northern Ireland	UKB	2	16.4	17	16.26
Spain					
Noroeste	ES1	1	13.2	14.3	35.07
Noreste	ES2	3	15.4	15.8	45.46
Madrid	ES3	5	16	13.7	56.89
Centro	ES4	1	13.4	17.7	42.31
Este	ES5	22	15.3	14.8	112.84
Sur	ES6	3	12.9	25.5	59.07
Ireland	IE	22	16.9	14.7	33.91
Portugal	PT	9	6.4	4.8	50.70