International Trade with Endogenous Sunk Export Costs: The Case of Marketing

Paper to the NOITS Annual Workshop 2005, Helsinki

Arne Melchior, Norwegian Institute of International Affairs
Version 1 November 2004, comments are welcome.

Abstract

In the growing literature on sunk costs in international trade transactions, these have been modelled as exogenous fixed costs of entering foreign markets. This paper shows that when the sunk export costs are *endogenous*, their impact is partly different. We allow firms to invest in marketing in each market, including the home market. The total level of marketing depends on how responsive demand is to marketing, but its allocation across markets depends on variable trade costs. Firms therefore tend to undertake more marketing at home than in export markets. This trade-reducing impact is stronger; the higher is the overall level of marketing. Contrary to the case with exogenous sunk costs, sales to each market are positively related to the level of sunk costs in that market. In the presence of marketing, the measurement of demand elasticities is changed.

Key words: International trade, marketing, investments, monopolistic competition, sunk export costs, gravity.

JEL classification numbers: F10, F12, F13, M30, R12.

Correspondence: Arne Melchior, Norwegian Institute of International Affairs, P.O. Box 8159 Dep., N-0033 Oslo, Norway. Tel. 47+99791209, e-mail arne.melchior@nupi.no

1. Introduction[1]

In research on international trade, marketing and distribution costs represent a large "black box" which is currently being opened. According to Anderson and van Wincoop (2004), trade costs broadly defined, including distribution costs, on average amount to a 170% ad valorem tax equivalent for a representative rich country. This paper contributes to a recent and growing research literature that attempts to shed light on this large component of trade transactions. While the majority of contributions in this literature have been empirical, this paper adds to a yet small theoretical literature. Using an international trade model with monopolistic competition, we examine how international trade and industrial location are affected by investments in *marketing* or other sales efforts.

The renewed interest in the magnitude and characteristics of international trade costs has been stimulated by empirical research on international trade. The "border effects" literature compares trade between regions within and across national borders, and concludes that national borders still represent a significant obstacle to trade; even between countries that are well integrated in terms of trade policy (see e.g. McCallum 1995, Helliwell 1996, Chen 2004). The empirical "gravity" literature confirms that bilateral trade flows are strongly impeded by geographical distance, and this impact has not become weaker over time (see e.g. Brun et al. 2002, Coe et al. 2002, Carrere and Schiff 2004). In spite of some remaining methodological challenges (see Anderson and van Wincoop 2003, Buch et al. 2003), this literature suggests that dreams about the "global village" or "borderless world" are premature, and trade costs still bite.

The continued importance of trade costs is also confirmed by micro-econometric evidence on firm-level export performance: Only a fraction of the firms are exporters, and the exporters are generally larger and more productive (Roberts and Tybout 1997, Clerides et al. 1998, Bernard and Jensen 1999, 2004, Bernard and Wagner 2001). Previous entry also increases the probability of exporting in a given period, and in the literature, this is taken to confirm the presence of trade costs that are sunk and not proportional to export value. The theoretical foundation for this empirical work has therefore often been models that combine firm heterogeneity, e.g. with respect to productivity, with sunk export costs (see e.g. Jean 2002, Helpman et al. 2004). In this framework, sunk export costs represent a threshold that only the most productive firms can surpass. This approach builds on the earlier literature on exchange rate hysteresis; which showed that temporary exchange rate shocks could have a permanent impact on trade due to sunk export costs (see

[1] Financial support from the Norwegian Research Council, project no. 137515/510 under the programme SKIKT (Social and Cultural Preconditions for Information and Communication Technologies) is gratefully acknowledged. I thank Hege Medin for comments to an earlier draft. An earlier version was printed as NUPI Working Paper No. 634/2002.

e.g. Baldwin 1988).[2] As shown by Bernard et al. (2003), the fact that only some firms export can also be explained theoretically by models without sunk export costs. Hence the hysteresis model is but one possibility.

This new evidence about high and persistent trade costs is useful and convincing, but it has one limitation: Some of the evidence is based on *unobserved* trade costs. Border effects are shown by dummies, gravity models use distance as a proxy for unobserved trade costs, and in the literature on firm-level export performance, sunk export costs are implied but not directly observed. For this reason, we need more evidence about the nature of the hidden barriers as well as on the economic mechanisms involved. This represents an extensive agenda for theoretical and empirical work. Responding to this need, a growing literature now addresses topics such as transport costs and infrastructure (see e.g. Bougheas et al. 1999), networks and migration (see e.g. Rauch 2001), contract enforcement and risk (Dixit 2003), trade and the internet (Freund and Weinhold 2004), and other issues and references could be added.

The solution to the "border effect puzzle" or "missing globalisation puzzle" is not necessarily that trade costs are so large: Another possibility is that there exist mechanisms that limit trade even if measurable trade costs are modest. Uncertainty and risk (Dixit 2003) is a case in point. In this article, we will show that *marketing* is another mechanism of this kind, which amplifies the impact of trade barriers. Marketing is intended to promote trade so it is not a trade barrier as such, but it may interact with other trade costs in a way that limits trade.

Our analysis is motivated by two of the stylised facts emanating from the empirical literature: Some trade costs are sunk, and distribution costs are important. We therefore examine theoretically how sunk costs in distribution affect international trade. More specifically, we analyse sunk costs that have the purpose of influencing demand in a particular market. This may include advertising, adjusting products to local demand, investment in sales offices, travel to meet customers, or other sunk costs in distribution that affects demand. We use *marketing* as a shorthand term for all such costs.

Common to earlier theoretical contributions on sunk export costs is that these costs are *exogenous*; i.e. independent of the firms' decisions. This assumption is plausible for some forms of sunk costs; e.g. adapting products to a national standard. For marketing, however, it is likely that the magnitude of investments depends on the firm's assessment of their profitability. In the analysis undertaken here, sunk export costs are therefore *endogenous* (using the terminology of Sutton (1991)).

A second difference between this paper and earlier contributions is that firms undertake marketing in their home market, so there are sunk costs also at home. The model predicts that firms will invest more in marketing, the lower are trade costs in a particular market. Sunk costs

---

[2] Venables (1994) and Medin (2003) show that firm heterogeneity can occur in international trade models with monopolistic competition, if an assumption of national product differentiation is added. Medin also examines a case where sunk export costs are borne by an immobile specific factor.

are therefore larger, and sales correspondingly higher, in the home market. In models with exogenous sunk costs, high sunk costs imply low sales, and sunk costs only apply abroad. In the analysis presented here, this logic is reversed.

We deviate from earlier theoretical contributions by not modelling firm heterogeneity. While such heterogeneity is more realistic in the light of empirical evidence, we sacrifice this here in order to obtain a tractable model.

While this paper is rooted the international trade literature, it relates to the economic literature on advertising (for an overview, see e.g. Martin 1993, Chapter 6). By analysing sunk costs that shift demand, we focus on *persuasive* rather than informative marketing. An early result in this literature was the Dorfman-Steiner (1954) condition or Theorem: For a monopolist, the advertising-to-sales ratio should equal the ratio between the "advertising elasticity" (the responsiveness of demand to advertising) and the demand elasticity. This Theorem also applies with monopolistic competition (see e.g. Valdes-Prieto 2002), and it may or may not apply under oligopoly (Lambin et al. 1975). While the oligopoly advertising literature addresses how specific forms of marketing is affected by strategic behaviour, the purpose here is to obtain knowledge about how marketing affects international trade and industrial location. For this purpose, we need a model where the number of firms is endogenous. In order to make the analysis tractable, we disregard strategic interaction and use an international trade model with monopolistic competition. Our behavioural assumptions about marketing are therefore stylised and simple.

In essence, this paper incorporates marketing into a standard model of international trade with monopolistic competition, along the lines of Krugman (1980) or Venables (1987). To our knowledge, marketing has not earlier been examined in a similar context. In the model, there are also traditional *ad valorem* trade costs, but these are unaffected by marketing. In the theoretical literature on foreign direct investment (FDI) (see e.g. Markusen and Venables 2000), trade costs provide a "tariff-jumping" motive for FDI – firms invest abroad in order to avoid trade costs. In the model here, we obtain the opposite effect: The returns to marketing are, *ceteris paribus*, lower in markets with high *ad valorem* trade costs. Firms will therefore invest more at home than abroad, and marketing thereby amplifies the impact of *ad valorem* trade costs, and reduces international trade.

It turns out that also in our model, the Dorfman-Steiner condition applies; on aggregate and in each market. Hence the volume of marketing depends on the "marketing elasticity" or the responsiveness of demand to marketing. If marketing is effective in generating increased demand, marketing will be larger, but at the same time a smaller share of marketing will be undertaken abroad. If marketing has little impact on demand, the model approaches standard international trade model with "home market effects": Large countries will have an advantage due to their larger home market. If marketing is effective, however, it will reduce international trade and thereby also dampen the "home market effect".

We deliberately present the model as partial equilibrium for one sector, hence implicitly assuming that there are no resource constraints or factor price effects in the model. When firms invest in marketing abroad, they may use domestic or foreign resources, but we do not specify this since it is not crucial for showing the basic mechanism. The purpose of the paper is to demonstrate a general microeconomic mechanism, and this may in later work be embedded in models with more sectors or a richer supply-side structure. While the results are derived in the context of international trade, the microeconomic effects are of a general nature and could apply to regions or cities as well.

In Section 2, we set out the model and derive the main implications of marketing. Section 3 shows that that in the presence of marketing, the measurement of elasticities is fundamentally changed. Section 4 examines the gravity relationship in international trade, extending the model to more countries. Section 5 concludes.

2. A model of international trade with marketing

There are two countries *i* and *j* (i,j=1,2). We consider the production and consumption of different varieties indexed k of goods in sector X. The representative consumer in country *j* has the CES utility function

(1) $$X_j = \left[\sum_k a_{kj}^{\frac{1}{\varepsilon}} x_{kj}^{\frac{\varepsilon-1}{\varepsilon}} \right]^{\frac{\varepsilon}{\varepsilon-1}}$$

where ε is the elasticity of substitution between varieties, x_{kj} is consumption of variety k in market j, and a_{kj} is a parameter allowing consumers to have special preferences for individual varieties.[3] With prices p_{kj} for individual varieties, we obtain the demand functions for variety k in market j

(2) $$x_{kj} = a_{kj}\, p_{kj}^{-\varepsilon} P_j^{\varepsilon-1} M_j$$

where M_j is the size of the market for X goods in j, and P_j is the price index for X goods in market j, dual to X_j (we have $X_jP_j=M_j$). P_j has the form

(3) $$P_j = \left[\sum_k a_{kj}\, p_{kj}^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}$$

Now we assume that firms in country i may invest an amount F_{ij} in market j (i,j=1,2) in order to increase the consumers' willingness to pay for their products. Moreover, we assume that the taste parameters a_{ij} depend on these investments as follows:

[3] See Armington (1969) or Venables (1987) for other examples using such taste parameters in CES utility functions.

(4) $$a_{ij} = F_{ij}^{\gamma} \qquad F_{ij} > 0,\ 0 < \gamma < 1$$

Hence we assume that there is a positive minimum investment in order to sell in each market, but firms are free to scale up their investments in order to increase the demand for their products. The parameter γ, the "marketing elasticity", is a measure of the efficiency of these investments; a large γ implies that a given investment has a large impact on demand. a_{ij} increases with F_{ij}, but the second-order derivative is negative so that the impact of increased investments is reduced as F_{ij} grows. For the model to be well behaved, we need $\gamma<1$.

With $\gamma=0$ we would have $a_{ij}=1$ in all cases, and the model would collapse to a standard "home market model" along the lines of Krugman (1980). In this case, if there is a positive minimum market investment, it will have no impact on demand, but enter the cost function as a standard exogenous fixed cost, together with fixed production costs. This will matter for the size of firms, but it will have no impact on market shares or the trade pattern. We shall therefore assume that $\gamma>0$.[4] When γ approaches zero, the model will approach the standard home market model, which may be thought of as a base case to which our results may be compared.

To what extent is the range assumed for γ empirically plausible or realistic? According to Shy (1995, 282), the advertising-to-sales ratio ranges empirically from close to zero up to 0.6. According to business sources, such ratios are below 0.1 for the majority of industries, and rarely above 0.2.[5] Since our marketing concept is wider than advertising, other sunk costs should be added, but data for these are limited. Nevertheless, our theoretical assumption seems reasonable in the light of evidence.

All firms in country *i* (or *j*) face exactly the same market conditions so that we can be sure that in the equilibrium situation, they all invest the same amounts in each market and hence face the same demand, and they will also charge the same prices. Using (2) and (4), we may therefore express demand for a variety from country *i* in market *j* as

(5) $$x_{ij} = F_{ij}^{\gamma} p_{ij}^{-\varepsilon} P_j^{\varepsilon-1} M_j$$

In addition to their marketing investments, firms face a fixed production cost *f*, and constant marginal costs *c*. They also have to spend resources on real *ad valorem* trade costs when selling in foreign markets; we express this as a mark-up on marginal costs t_{ij}, $t_{ij}>1$ for $i \neq j$. In the home

[4] The assumption of $\gamma>0$ (rather than $\gamma\geq0$) is not technically necessary, but is applied in order to simplify the exposition. Some of the solutions for F_{ij} apply only with $\gamma>0$, and the case with $\gamma=0$ would then have to be treated separately.

[5] See e.g. information from Schonfeld & Associates Inc. (http://www.saibooks.com/).

market there are no such costs, t_{ii}=t_{jj}=1. For sales from market i to market j, the marginal costs are then equal to ct_{ij}.[6]

With these costs, profits of a firm in country i are

(6) $$\pi_i = -f - \sum_j F_{ij} - \sum_j x_{ij} c\ t_{ij} + \sum_j x_{ij} p_{ij}$$

The market structure is Chamberlinian monopolistic competition; firms choose F_{ij} as well as x_{ij} while neglecting the impact of their actions on the industry average. Technically, this amounts to treating the price index P_j as a constant when maximising profits.

Maximising profits with respect to x_{ij} (also treating F_{ij} as given) gives the first-order conditions

(7) $$p_{ij} = \frac{\varepsilon}{\varepsilon - 1} c\ t_{ij}$$ or equivalently (7a) $$\frac{p_{ij} - c\ t_{ij}}{p_{ij}} = \frac{1}{\varepsilon}$$

i.e. the standard pricing condition in trade models with monopolistic competition.

Maximising profits with respect to F_{ij} gives the first-order condition

(8) $$\frac{\partial x_{ij}}{\partial F_{ij}} = \frac{1}{p_{ij} - c\ t_{ij}}$$

From (5) we also obtain:

(9) $$\frac{\partial x_{ij}}{\partial F_{ij}} = \gamma\ F_{ij}^{\gamma-1} p_{ij}^{-\varepsilon} P_j^{\varepsilon-1} M_j$$

Now observe from (5) that a firm's total sales in market j (using the notation $v_{ij}=x_{ij}p_{ij}$) is equal to

(10) $$v_{ij} = x_{ij} p_{ij} = F_{ij}^{\gamma} p_{ij}^{1-\varepsilon} P_j^{\varepsilon-1} M_j$$

Multiplying (8) as well as (9) by p_{ij}, and then using (7a) and (10), we obtain:

(11) $$\frac{F_{ij}}{v_{ij}} = \frac{\gamma}{\varepsilon}$$

[6] This gives results that are qualitatively similar to the case with "iceberg" trading costs, where some of the goods shipped "melt away" on the way to their destination. With this approach, firms charge the same price in all markets, but since consumers receive less of the goods, the "real price" is higher. The prices used here correspond to the "real prices" in the iceberg modelling approach.

which is the Dorfman-Steiner (1954) condition:

Result 1: The Dorfman-Steiner condtion applies in each market, and therefore also on aggregate: The marketing-to-sales ratio equals the ratio between the marketing elasticity and the demand elasticity perceived by firms.

As we would expect, marketing is higher if the marketing elasticity is high. Observe also that the proportion of marketing is lower for more homogeneous goods. The intuition is that for such goods, the price mark-up is lower, so the profit increase for an additional unit sold is smaller.

With monopolistic competition, total profits (6) must be zero. Also using (7a) and (11), we obtain (also using the notation $F_i=\sum_j F_{ij}$):

(12)
$$F_i = \sum_j F_{ij} = \frac{\gamma f}{1-\gamma}$$

The function $\gamma/(1-\gamma)$ is a non-linear increasing function of the marketing elasticity, with a positive second-order derivative. Hence the more marketing affects demand, the larger are the investments in marketing relative to the fixed production cost. Observe that the expression applies to firms in either country, so the total level of marketing is the same for firms in small and large countries. Furthermore, aggregate marketing is independent of the level of ad valorem trade costs.

Using (11) and (12), total sales of an individual firm in country i will be:

(13)
$$\sum_j v_{ij} = \frac{\varepsilon f}{1-\gamma}$$

The function $1/(1-\gamma)$ is also a strictly increasing function of the marketing elasticity. Hence we may conclude:

Result 2: Total investments in marketing and total sales are convex strictly increasing functions of the marketing elasticity, and independent of the level of ad valorem trade costs.

When $\gamma \rightarrow 0$ (investments in marketing have no effect), firm size approaches εf, as in a standard model with monopolistic competition. If investments in marketing are effective, however, firms will be larger and the market will be more concentrated.

Firms invest in marketing in their home market as well as in the foreign country. In order to derive their investments in the respective markets, we simplify by assuming that variable trade costs are the same in both directions; i.e. $t_{12}=t_{21}=t$. Using (10) and substituting for v_{ij} from (11) ($v_{ij}=\varepsilon F_{ij}/\gamma$) we obtain

(14) $$F_{ij}^{1-\gamma} = \frac{\gamma}{\varepsilon} p_{ij}^{1-\varepsilon} P_j^{\varepsilon-1} M_j$$

Using this to compare F_{ij} and the corresponding expression for F_{jj}, also substituting from (7) for prices, we obtain

(15) $$F_{ij} = F_{jj} t^{\frac{1-\varepsilon}{1-\gamma}}$$

In a standard model without marketing, the ratio between sales abroad and sales at home will be $t^{1-\varepsilon}$, so $1-\varepsilon$ is a core "trade elasticity" in that model. With marketing, the corresponding elasticity is $(1-\varepsilon)/(1-\gamma)$. Since this shows up frequently in the following analysis, we use the notation

$$\theta = \frac{1-\varepsilon}{1-\gamma}$$

for this composite "trade elasticity". This is negative since $\varepsilon > 1$, $0<\gamma<1$, and the absolute value increases with γ as well as ε. Observe also that $t^{\theta}<1$, so (15) implies that firms undertake more marketing at home than abroad.

With equal trade barriers $t_{12}=t_{21}=t$, the result (15) applies to either country. Using (12) and (15), we solve for marketing in the home market:

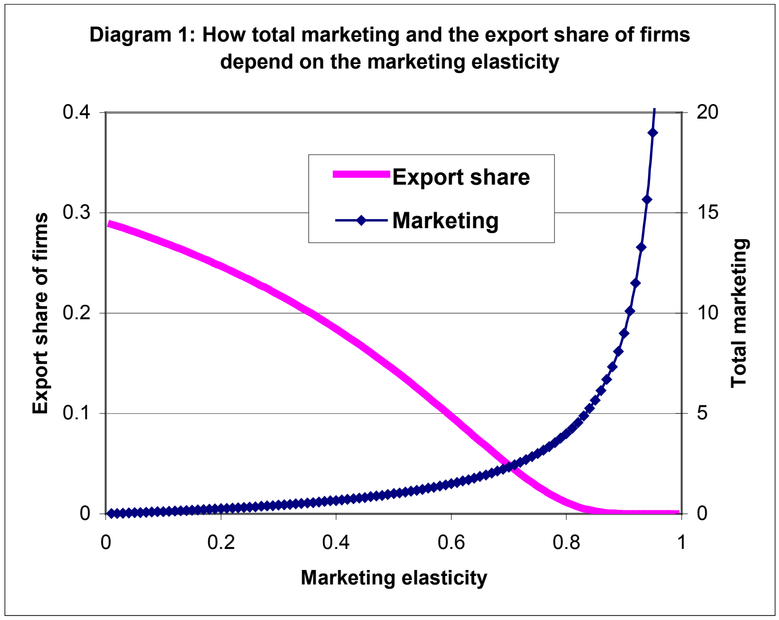
(16) $$F_{ii} = \frac{\gamma f}{1-\gamma} \frac{1}{1+t^{\theta}}$$

and F_{ij} then follows from (15), given that $F_{ii}=F_{jj}$. Sales in each market then follow from the Dorfman-Steiner condition (11).

From these results it is evident that firms invest more in marketing in their home market than abroad, and sell more at home. Since $F_{ii}/F_{ij}=t^{\theta}$, it follows that firms will be more export oriented if variable trade costs t are low, or if products are not too close substitutes (ε is low). These are standard features in international trade models. The derivative of F_{ij}/F_{jj} with respect to γ is more interesting: The first-order derivative is negative, and the second-order derivative is ambiguous (it is normally negative, but can become positive if t and/or ε are large). Hence the larger is the marketing elasticity γ, the lower is the share of marketing abroad, and the lower is the export share of total sales (given Result 1, marketing and sales in each market are proportional). We may therefore conclude:

Result 3: Firms undertake more marketing at home than abroad. The share of marketing abroad in total marketing, equal to the share of exports in total sales, is a strictly decreasing function of the marketing elasticity.

Recalling Result 2, a higher marketing elasticity raises total marketing expenditure but at the same time, the share of marketing abroad declines. Diagram 1 illustrates this, for given *ad valorem* trade costs and substitution elasticity.[7]



Hence the more firms spend on marketing, the more they will focus on the home market. The more important is marketing, the more closed will the economy be. The intuition is that in the presence of ad valorem trade costs, the returns to investments in marketing will – other things equal – be larger in the home market, and the more so if marketing is effective. Hence the more receptive buyers or consumers are to marketing, the smaller will international trade be. Marketing will thereby amplify the impact of ad valorem trade costs. If *ad valorem* trade costs approach zero ($t \rightarrow 1$), the impact of marketing on international trade will vanish.

We turn next to industrial location. As noted in the introduction, standard international trade models with monopolistic competition imply that large countries have a "home market advantage" for differentiated goods. In order to examine whether the "home market effect" also applies here, we solve for the number of firms n_i and n_j in each country. Total sales must add up to total market size in each market:

(17a) $n_1 v_{11} + n_2 v_{21} = M_1$ (17b) $n_1 v_{12} + n_2 v_{22} = M_2$

Using the former solutions for $v_{ii}=v_{11}=v_{22}$ and $v_{ij}=v_{12}=v_{21}$, we obtain:

(18) $$n_i = \frac{1-\gamma}{\varepsilon f} \frac{M_i - M_j t^{\theta}}{1 - t^{\theta}}$$

[7] Parameter values: t=1.25, ε=5, f=1.

For country j, the subscripts i and j are interchanged. When $\gamma \to 0$, the solution for n_i approaches the outcome in standard home market model. The number of firms can become zero; this occurs if $M_i/M_j \leq t^{\theta}$. The following results are only valid if there are a positive number of firms in both countries.

Writing $M=M_i/M_j$ for the market size ratio, we obtain (also using the notation $H=n_i/n_j$ for the ratio of firms)

$$(19) \qquad H = \frac{n_i}{n_j} = \frac{M - t^{\theta}}{1 - M t^{\theta}}$$

The first as well as the second order derivative of H with respect to M are positive, which shows that we have a "home market effect": Large countries have a *disproportionately* high share of world production of differentiated goods. The higher is the marketing elasticity, however, the less international trade there is, and this reduces the large-country advantage. When γ approaches 1, the first-order derivative of H with respect to M approaches one and the second-order derivative approaches zero. Hence the home market effect gradually vanishes as marketing becomes more effective. The sign of the relevant derivatives depend on whether M is smaller or larger than 1. In the case when $M>1$, we have

$$(20) \qquad \frac{\partial H}{\partial \gamma} < 0 \qquad \text{if } M>1$$

[8] With $M<1$, the sign is reversed. We can therefore conclude:

Result 4: By reducing international trade, marketing dampens the home market advantage for large countries in the production of differentiated goods.

3. Redefining elasticities

As noted in the introduction, "border effects" in international trade may either be explained by hidden trade barriers, or by known trade barriers having a stronger impact. The latter is precisely what follows from our analysis: The elasticity of trade with respect to changes in *ad valorem* is larger than standard models predict. In the standard model, the ratio between exports and home market sales value is $t^{1-\varepsilon}$; in the marketing model it is equal to t^{θ} (equation (15)). Since $\theta<(1-\varepsilon)<0$, the absolute value of the elasticity is larger.

For the individual firm, the perceived elasticity of demand is equal to the elasticity of substitution ε. In the market equilibrium, however, the true elasticity is different due to changes in the price indexes P_i and P_j. Substituting our solutions into the demand equation (2), exports from a firm (equal for firms in both countries) will be

---

[8] If $M>1$, the derivatives of H with respect to t and ε are also negative.

(21) $x_{ij} = \frac{(\varepsilon-1)f}{(1-\gamma)c}\frac{t^{\theta-1}}{(1+t^{\theta})}$ or in value (21a) $v_{ij} = \frac{\varepsilon f}{(1-\gamma)}\frac{t^{\theta}}{(1+t^{\theta})}$

The elasticity of firm-level exports with respect to *ad valorem* trade costs is then equal to

(22) $El_t x_{ij} = \frac{\theta}{1+t^{\theta}} - 1$ or (22a) $El_t v_{ij} = \frac{\theta}{1+t^{\theta}}$

(22a) is also the elasticity of total world trade value with respect to *ad valorem* trade costs. Note that since the export shares of firms in both countries are the same, the production shares of the two countries do not as such affect total world trade.

The absolute values of these trade elasticities are increasing with the marketing elasticity. Using (22a), we obtain:

(23) $$\frac{\partial El_t v_{ij}}{\partial \gamma} = \frac{1-\varepsilon}{\left(1-t^{\theta}\right)\left(1-\gamma\right)^2}\left\{1+\frac{t^{\theta}\ln t\,(\varepsilon-1)}{(1-\gamma)\left(1-t^{\theta}\right)}\right\} < 0$$

Given that the elasticity is normally negative, the absolute value of the trade elasticity will be larger if the marketing elasticity is high.

Result 5: The higher is the marketing elasticity, the more elastic is trade with respect to changes in ad valorem trade costs.

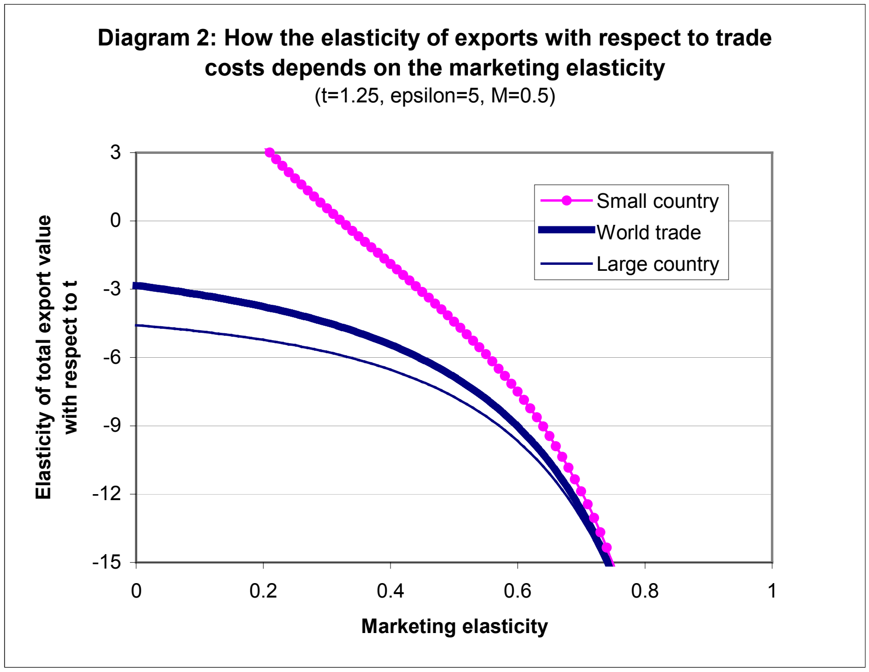
Changes in trade costs affect world trade via the export shares of firms. In addition, such changes will relocate firms between countries; trade liberalisation will amplify the home market effect and relocate firms from the small to the large country. For this reason, the elasticity of total exports (i.e. $n_i v_{ij}$) will differ between the two countries (see also Feenstra et al. 2001). For the large country, trade liberalization will increase the number of firms as well as their export shares, but for the small country, firm exit will limit export growth. The elasticity of total exports for country i with respect to *ad valorem* trade costs is

(24) $$El_t(n_i v_{ij}) = \frac{\theta\left(M_i + M_i t^{2\theta} - 2M_j t^{\theta}\right)}{\left(M_i - M_j t^{\theta}\right)\left(1-t^{2\theta}\right)}$$

If countries are of equal size, this collapses to (22a), but if country size differs, the elasticities also differ. The derivative of the elasticity (24) with respect to country size M_i is negative. Hence the total exports of large countries, or equivalently the imports of small countries, will be more elastic. We state our result as:

Result 5: The exports of large countries are more elastic with respect to ad valorem trade costs than the exports of small countries.

In Diagram 2, we illustrate these results:



All the three curves are falling, so the sensitivity of trade with respect to trade costs is larger, the higher is the marketing elasticity γ. The curves continue outside the range shown in the Diagram, with extremely high absolute values when γ approaches one. The Diagram also shows the difference between small and large countries: For the small country, the elasticity can even become positive! The divergence between large and small countries is greatest when the marketing elasticity is low, and we approach the standard home market effect model. When the marketing elasticity increases, trade is reduced, the home market effect is dampened, and the elasticities in small and large countries converge. Hence the impact of country size on trade elasticities should be lower for marketing-intensive industries.

These results have important implications for empirical work. In recent years, e.g. tariff data have become more easily available, and these may be used in order to estimate demand elasticities. What the analysis here shows, is that empirically measured elasticities using such data are composite functions where the demand elasticity is only one of the arguments. If marketing or country size differences are not taken into account, empirical estimates could become biased.[9]

[9] It can also be shown that the trade elasticity (22) is a decreasing function of the level of ad valorem trade costs t; hence trade is more elastic at high levels of trade costs. The intuition is that if trade costs are high, imports have a smaller share of the market, and cross-product substitution effects are then stronger.

In the model presented here, the elasticity of substitution is exogenous and not related to marketing. If differentiated goods (with a low substitution elasticity) are more marketing-intensive (higher γ), the marketing and substitution elasticities affect the trade elasticity in opposite directions. In that case, empirical estimates for different sectors could be similar, even if the underlying parameters are quite different. In empirical research, it is confirmed that differentiated goods have lower trade elasticities, but the difference is sometimes modest (see e.g. Kee et al. 2004, Table 2).

4. A simple gravity model

If firms can sell to many foreign markets, the former results suggest that they would undertake less marketing and sell less in markets with high *ad valorem* trade costs. If trade costs increase with geographical distance, marketing could amplify the impact of distance and strengthen the "gravity effect". In the following, we show this, by extending the model to many countries.

The earlier equations (1)-(14) also apply to the general case with n countries. The following analysis thus departs from equation (14). With differing trade costs, equation (15) becomes

(25) $$F_{ij} = F_{jj} t_{ij}^{\theta}$$

where the subscript on t is the only difference from (15).

As the simplest possible "gravity" model, consider 4 countries that are evenly spaced around the circumference of a circle, with distance equal to one unit between adjacent countries, and two units between "remote" countries. Each country therefore has two neighbours and one "remote" trade partner. For analytical simplicity, we assume that trade costs increase exponentially with distance; these are equal to t between adjacent countries (still assuming $t>1$), and t^2 between "remote" countries. From (12) and (25), investments in marketing are independent of market size. Since all countries face similar trade barriers, the firms' investments in marketing in their home markets must be identical in all countries. Using this, together with (11) and (25), equation (13) may be expressed as:

(26) $$F_{ii}\,(1 + t^{\theta} + t^{2\theta} + t^{\theta}) = \gamma f/(1-\gamma)$$

The bracketed sum on the left hand side is equal to $(1 + t^{\theta})^2$, so we have

(27) $$F_{ii} = \frac{\gamma f}{(1-\gamma)\left(1+t^{\theta}\right)^{2}}$$

This defines home investments in marketing, which is equal in all the four countries. Compared to the solution in the two-country case, the only difference is the power 2 in the denominator. Home investments in marketing are therefore lower than in the two-country case, since firms now spread their investments across more markets. But still, firms undertake more marketing at home than in any individual foreign market.

Investments in the foreign markets then follow from (25), and it is evident that firms will invest most in their home market, but less in the remote country than in the neighbour countries (since $t>1$, $\theta<0$). From (11) it follows that firms will sell more to the neighbours than to the remote country. Hence the impact of distance is amplified by marketing. As in the two-country case, marketing magnifies the elasticity of trade with respect to trade costs. The elasticity of trade with respect to distance will therefore be larger, the more homogenous are the goods, and the more effective is marketing. We may conclude:

Result 6: Marketing increases the sensitivity of trade with respect to trade costs that are correlated with geographical distance, and thereby strengthens the gravity relationship in international trade flows.

If all the four countries are of equal size, they will have an equal number of firms (equal to $M_i (1-\gamma)/\varepsilon f$), and trade will be balanced in each bilateral trade flow and all trade will be intra-industry trade. As shown above, international trade will be smaller and more focused on neighbour countries if γ is large. If size differences are introduced, the impact on trade and the distribution of industry depends on the location of large countries in space. If one country grows in size, the impact will be different in its neighbour countries and the remote country. In order to show this, we write the market clearing equations for each market, which for country 1 will be

(28) $$n_1 v_{11} + n_2 v_{21} + n_3 v_{31} + n_4 v_{41} = M_1$$

Similar equations apply to countries 2-4. Now using (11) and (25), we may express this equation system in matrix form as

(29) $$\frac{\varepsilon}{\gamma} F_{ii} \begin{bmatrix} 1 & t^{\theta} & t^{2\theta} & t^{\theta} \\ t^{\theta} & 1 & t^{\theta} & t^{2\theta} \\ t^{2\theta} & t^{\theta} & 1 & t^{\theta} \\ t^{\theta} & t^{2\theta} & t^{\theta} & 1 \end{bmatrix} \times \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{bmatrix} = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix}$$

Denoting the three matrixes and vectors as **T**, **N** and **M**, respectively, we have

(29a) $$(\varepsilon/\gamma)\, F_{ii}\, \boldsymbol{TN} = \boldsymbol{M}$$

and the solution for the number of firms then follows:

(30) $$\boldsymbol{N} = \gamma/\,(\varepsilon F_{ii})\; \boldsymbol{T}^{-1} \boldsymbol{M}$$

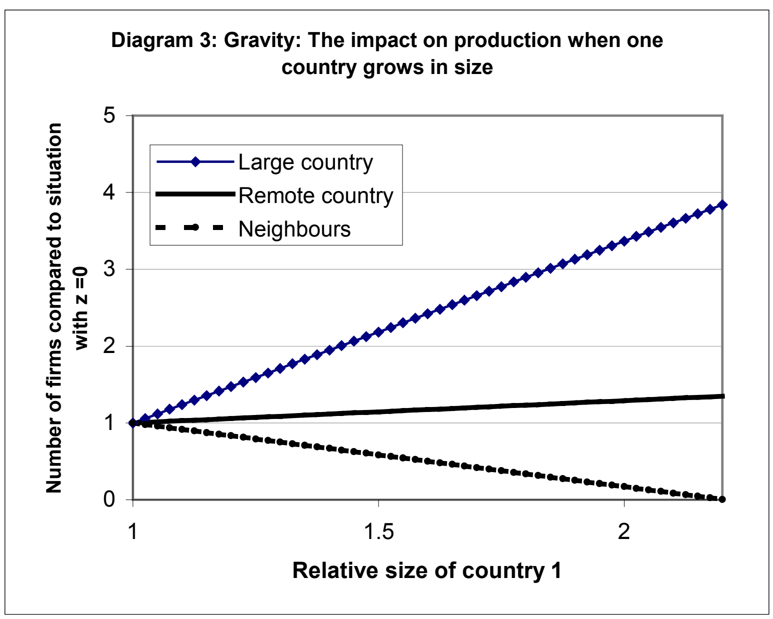
The inverse $\boldsymbol{T}^{1}$ is equal to

$$(31) \qquad \boldsymbol{T}^{-1} = \frac{1}{\left(t^{-\theta} - t^{\theta}\right)^2} \begin{bmatrix} t^{-2\theta} & -t^{-\theta} & 1 & -t^{-\theta} \\ -t^{-\theta} & t^{-2\theta} & -t^{-\theta} & 1 \\ 1 & -t^{-\theta} & t^{-2\theta} & -t^{-\theta} \\ -t^{-\theta} & 1 & -t^{-\theta} & t^{-2\theta} \end{bmatrix}$$

Using (30-31), and substituting for F_{ii} from (27), the number of firms in each country follows. Further analytical results may then also be derived, but we drop this in order to avoid excessive detail, and use numerical simulations to show some properties of the model.

As an example, consider that the four countries are originally of equal size M_0, and that country 1 grows in size to M_0+z, where z is a positive amount. It is then easily shown that the number of firms in country 1 increases, and the number of firms in the neighbour countries 2 and 4 is reduced, and there is a "home market effect" so that country 1 gets a disproportionately large share of production, while countries 2 and 4 obtain a disproportionately low share.

This is what we should expect, given the results from the two-country model. The qualitative difference between the spatial model and the two-country model shows up when we consider the remote country 3: In fact, the number of firms in country 3 increases because country 1 grows in size. Diagram 3 simulates this outcome:[10]



The horizontal axis measures the relative size of country 1, starting from 1 where all the four countries are of equal size. The vertical axis

[10] Other parameter values: t=1.3, ε=3, γ=0.5, f=1.

measures the relative number of firms in a country, compared to the situation with z=0 (all countries of equal size M_0). As country 1 grows in size, there is a "home market effect", but this only hurts the neighbour countries, and the remote country gains. Due to geographical distance, the countries in-between act as "buffers" and their industrial decline benefits the remote country. [11]

Hence in spatial models, transport costs may create wave-like repercussions on the industrial location pattern, as also shown by Fujita et al. (1999, Chapter 17).[12] In the model used here, there are no true cross-border externalities (such as knowledge spillovers, economies of scope for differentiated inputs, etc.). This is the reason why the *neighbours* of the large country are subject to industrial decline. In the presence of cross-border positive externalities, neighbours to large manufacturing producers benefit from these externalities and this modifies the outcome. In models with such externalities, it may not be the neighbour countries that are worse off, but countries further away. Hence the spatial impact depends on the exact model assumptions, but with agglomeration, it will generally be the case that growth in some areas has its counterpart in industrial decline elsewhere.

Due to this complex relationship between country size and industrial location in spatial models, the relationship between country size, distance and the trade pattern is also more complex. As an illustration, suppose that we generate trade data from the model, and estimate by ordinary least squares a standard gravity equation of the form

$$ln(n_i v_{ij} + n_j v_{ji}) = ln(\alpha_0) + \alpha_1\, ln(M_i) + \alpha_2\, ln(M_j) + \alpha_3\, ln(Distance_{ij}) + e_{ij}$$

This gravity model is clearly mis-specified; there is an omitted-variable bias because the number of firms is affected by country size as well as distance, and this affects trade flows. If we generate trade data from the model with different country size patterns, and run gravity regressions, we will observe that all the parameters, including the estimate α_3 on distance, change. In the model, there is a perfect gravity relationship since trade costs increase exponentially with distance, but this is only captured if all countries are of equal size: then the gravity regression fits 100%. If country size varies, R^2 falls and the parameter estimates vary. A better fit is obtained if we run regressions for exports and imports separately, since the home market effects will then be reflected in the parameters α_1 and α_2: large countries export more, and small countries

[11] If the number of countries is increased in the model used here, it can be shown that a size increase for an individual country only affects its two neighbours, and the three countries diametrically opposed to it on the circle.

[12] Also in this case, we find that with large ε or γ or t, the home market effect is dampened and the number of firms in each country eventually approaches proportionality with respect to country size. For small values of ε or γ or t, the impact of country size differences will be larger, but with an important distinction with respect to the number of firms in the "remote" country 3: Small t or ε tends to amplify the positive effect for country 3 – at the limit the number of firms in country 3 grows as strongly as in country 1. For small γ, on the other hand, this is not the case – then country 3 remains in the intermediate range, as in Diagram 3.

import more ($\alpha_1 > \alpha_2$, see also Feenstra et al. 2001). But also in that case, the model does not fit perfectly.

Result 7: The distance coefficient in gravity equations is affected by changes in industrial location that are caused by geography and country size differences, even if the impact of geographical distance on trade is unchanged.

The implication is that growth differences, with corresponding changes in industrial location, may affect the distance parameter in gravity equations even if the underlying relationship between trade and distance remains the same. If distance parameter estimates have increased over time (see references in the introduction), we cannot conclude from this that distance-related trade costs have become higher.[13]

In spite of this ambiguity concerning the impact of country size in gravity models, we still have a "clean" inverse relationship between trade flows and distance in the model if we consider the *shares* of exports from a country to different destinations. Hence in empirical work on gravity, an idea is to use export shares rather than gross bilateral trade flows in order to trace the impact of distance on trade flows. Our analysis also suggests that empirical work should be undertaken at a disaggregated level, since the strength of the gravity relationship should vary across sectors, due to different *ad valorem* trade costs as well as different marketing intensities. For sectors in which marketing is important, we expect that the inverse relationship between trade and distance will be stronger. In our model, there is only one sector, but if trade has to be balanced, trade surpluses in one sector must be matched by deficits in others. An extended model is necessary to shed light on this.

4. Concluding comments

In recent research on sunk costs in international trade, the "hysteresis model" with firm heterogeneity and exogenous sunk export costs has been frequently used as a theoretical framework. In such models, sunk trade cost only exist in export markets, and represent a threshold that only the most efficient firms can surpass. In this paper, we have shown that sunk export costs may affect trade in a very different way if they are *endogenous*, with a magnitude that depends on the firms' investment decisions. In the case of marketing, sales in each market are proportional to the sunk costs (following the Dorfman-Steiner condition), firms also undertake marketing at home, and the larger are total investments in marketing, the larger share of these investments will be in the home market.

For empirical research, the analysis has several implications, e.g.:

- The economic implications differ sharply with endogenous instead of exogenous sunk costs.

[13] Buch et al. (2004) have a similar point when they show that if trade costs increase linearly with distance, proportional changes in all trade costs will show up in the constant term in gravity equations, and not in the parameter estimate on distance.

- We should not only be looking for sunk trade costs abroad, but also in the home market.
- Marketing amplifies the trade impact of *ad valorem* trade costs, and this effect depends on the marketing-to-sales ratio in each sector.
- Marketing reinforces the inverse relationship between trade and geographical distance, if trade costs depend on distance.
- Empirical estimation of demand elasticities can be biased unless marketing is taken into account.
- Marketing dampens the "home market effect".
- The elasticity of exports with respect to trade costs differs between large and small countries.
- Empirical estimates on the impact of distance in gravity equations are affected by changes in industrial location due to agglomeration.

The last two implications also apply in similar models with imperfect competition, but the other ones are due to the presence of marketing. In future work, these hypotheses should be tested empirically. The mere size of distribution costs warrants further work in this area. The analysis shows that large "border effects" or "distance effects" in gravity models do not necessarily imply that trade costs are very large, but they can be due to the interaction between trade costs and microeconomic mechanisms such as the one demonstrated here.

On the other hand, it should be recalled that the mechanism presented here is but one of several possibilities. For example, firms could invest in sales offices in order to reduce *ad valorem* trade costs, rather than to influence demand. In that case, it is likely that conclusions could be different, and we might have a "tariff-jumping" motive for investments in sales organization, as shown in the FDI literature. There may also be economies of scale in the exporting activity for other reasons; e.g. learning-by-doing by exporters (or even customers)[14]; externalities between exporters (e.g. knowledge spillovers); externalities related to export infrastructure (e.g. transport supply); or related to risk (e.g. when larger trade creates trust, as shown by Dixit (2003)).

Firms frequently do not sell directly to final customers, and an issue is whether the analysis also applies if this is not the case. As defined here, however, marketing could also be related to sales to wholesalers or producers, so it is an empirical issue whether the analysis is relevant.

We have focused on the empirical implications for trade and industrial location, but it should be recalled that by affecting price determination, sunk export costs could have important macroeconomic implications. Bergin et al. (2004) and Ghironi and Melitz (2004) have recently demonstrated such implications. Market entry barriers could therefore play an important role in the explanation of price differences

[14] On learning-by-doing, empirical work shows that causality mainly runs from productivity to exporting (i.e. productive firms self-select as exporters) and not the other way round. This evidence has been interpreted to say that learning-by-doing is not so important (see e.g. Bernard and Wagner 2001, Clerides et al. 1998, Bernard and Jensen 1999). While this conclusion is plausible, there is also the possibility that firms learn to become better exporters rather than better producers. In order to test this, one should not only check whether firms export or not, but how they perform in selling to different markets.

across markets, and marketing is an obvious form of such barriers that should be explored. In their analysis of price convergence in the European internal market, DRI (1997) concluded that price convergence was more limited in sectors with high advertising intensity combined with a high level of R&D. A possible extension of the analysis here is to check how marketing interacts with R&D or supply-side differences in general. For this purpose, the mechanism could also be examined in a general equilibrium framework with more sectors and a more realistic supply side.

Firm heterogeneity is a stylized fact empirically, but we have sacrificed this here in order to make the model as simple as possible. An interesting extension is to extend the model by adding firm heterogeneity, either in the form of production cost differences, or by making the impact of marketing stochastic rather than deterministic.

The analysis has been based on a very stylised model, and an issue is to what extent our general results depend on these stylised assumptions. Sutton (1991) e.g. showed that with oligopoly, endogenous sunk costs in advertising could lead to an escalating "marketing race". As a result, industrial concentration could increase with market size. In our model, this is not the case: The total number of firms increase with world market size. This indicates that the form of competition also matters, and endogenous sunk costs *per se* do not define the outcome.

References

Anderson, J.E. and E. van Wincoop, 2004, Trade costs, *Journal of Economic Literature* XLII (September): 691-751.

Anderson, J.E. and E. van Wincoop, 2003, Gravity with Gravitas: A Solution to the Border Puzzle, *American Economic Review* 93(1): 170-192.

Armington, P., 1969, A Theory of Demand for Products Distinguished by Place of Production, *IMF Staff Papers* 16: 159-176.

Baldwin, R., 1988, Hysteresis in Import Prices: The Beachhead Effect, *American Economic Review* 78(4): 773-785.

Bergin, P., R. Glick and A.M. Taylor, 2004, Productivity, Tradability and the Long-Run Price Puzzle, NBER Working Paper No. W10569.

Bernard, A., J. Eaton, J. Bradford Jensen and S. Kortum, 2003, Plants and Productivity in International Trade, *American Economic Review* 93(4): 1268-1290.

Bernard, A.B. and J.B. Jensen, 1999, Exceptional export performance: cause, effect or both? *Journal of International Economics* 47: 1-25.

Bernard, A.B. and J.B. Jensen, 2004, Why Some Firms Export, *The Review of Economics and Statistics* 86(2): 561-569.

Bernard, A. and Wagner, J., 2001, Export entry and exit by German firms, *Weltwirtschaftliches Archiv* 137 (1): 105-123.

Bougheas, S., P.O. Demetriades and E.L.W. Morgenroth, 1999, Infrastructure, transport costs and trade, *Journal of International Economics* 47: 169-189.

Brun, J.-F., C. Carrere, P. Guillaumont and J. de Melo, 2002, Has distance died? Evidence from a panel gravity model, London: CEPR (Centre for Economic Policy Research), Discussion Paper No. 3500.

Buch., C.M., J. Kleinert and F. Toubal, 2004, The distance puzzle: on the interpretation of the distance coefficient in gravity equations, *Economics Letters* 83: 293-298.

Carrere, C. and M. Schiff, 2004, On the geography of trade: Distance is alive and well, Washington DC: World Bank Policy Research Working Paper No. 3206.

Chen, N., 2004, Intra- National Versus International Trade in the European Union: Why Do National Borders Matter? *Journal of International Economics* 63(1): 93-118.

Clerides, S. K., S. Lach and J.R. Tybout, 1998, Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico and Morocco. *Quarterly Journal of Economics* 113: 903-948.

Coe, D.T., A. Subramanian and N.T. Tamirisa, with R. Bhavnani, 2002, The Missing Globalization Puzzle, Washington DC: IMF Working Paper No. 02/171.

Dixit, A., 2003, Trade Expansion and Contract Enforcement, *Journal of Political Economy* 111(6): 1293-1317.

Dorfman, R. and P. Steiner, 1954, Optimal Advertising and Optimal Quality, *American Economic Review* 44: 826-36.

DRI, 1997, *Price competition and price convergence*, The Single Market Review, Subseries V: Impact on Competition and Scale Effects, Volume 1, Kogan Page/Earthscan/ Office for Official Publications of the European Communities.

Feenstra, R.C., J.R. Markusen and A.K. Rose, 2001, Using the gravity equation to differentiate among alternative theories of trade, Canadian Journal of Economics 34(2): 430-447.

Freund, C. and D. Weinhold, 2004, The effect of the Internet on international trade, *Journal of International Economics* 62(1): 171-189.

Fujita, M., P.R. Krugman and A.J. Venables A.J., 1999, *The spatial economy. Cities, Regions and International Trade,* Cambridge MA/London: MIT Press.

Ghironi, F. and M.J. Melitz, 2004, International trade and macroeconomic dynamics with heterogeneous firms, NBER Working Paper No. 10540.

Helliwell, J., 1996, Do national borders matter for Quebec's trade?, *Canadian Journal of Economics*, 29 (3), pp 507-522.

Helpman, E., M.J. Melitz and S.R. Yeaple, 2004, Exports Versus FDI with Heterogeneous Firms, *American Economic Review* 94(1): 300-316.

Jean, S., 2002, International trade and firms' heterogeneity under monopolistic competition, *Open Economies Review*, 13 (3), pp. 291-311.

Kee, H.L., A. Nicita and M. Olarreaga, 2004, Import demand elasticities and trade distortions, London: Centre for Economic Policy Research, CEPR Discussion Paper No. 4669.

Krugman, P., 1980, Scale Economies, Product Differentiation and the Pattern of Trade, *American Economic Review* 70(5): 950-59.

Lambin, J.-J., P. Naert and A. Bultez, 1975, Optimal Marketing Behaviour in Oligopoly, *European Economic Review* 6, 105-128.

Markusen, J. and A.J. Venables, 2000, The theory of endowment, intra-industry and multi-national trade*, Journal of International Economics* 52: 209-234.

Martin, S., 1993, *Advanced industrial economics*, Oxford UK and Cambridge USA: Blackwell.

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

Medin, H., 2003, Firms' export decisions. Fixed trade costs and the size of the export market, *Journal of International Economics* 61(1): 225-41.

Rauch, J.E., 2001, Business and Social Networks in International Trade, *Journal of Economic Literature*, XXXIX: 1177-1203.

Rauch, J.E. and J. Watson, 2004, Network Intermediaries in International Trade, *Journal of Economics and Management Strategies* 13: 69-93.

Roberts, M.J. and J.R. Tybout, 1997, The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs, *American Economic Review* 87(4): 545-564.

Shy, O., 1995, *Industrial Organization*, MIT Press.

Sutton, J., 1991, *Sunk Costs and Market Structure, Price Competition, Advertising and Evolution of Concentration*, Cambridge MA/London: MIT Press.

Valdes-Prieto, S., 2002, Pricing and marketing rules with brand loyalty, Pontificia Universidad Catolica de Chile, Instituto de Economia, Documento de Trabajo 210, February 2002.

Venables, A.J., 1987, Trade and trade policy with differentiated products: A Chamberlinian-Ricardian model, *The Economic Journal* 97 (No. 387): 700-717.

Venables, A.J., 1994, Integration and the Export Behaviour of Firms: Trade Costs, Trade Volumes and Welfare, *Weltwirtschaftliches Archiv* 130, 118-132.