# [634] Paper
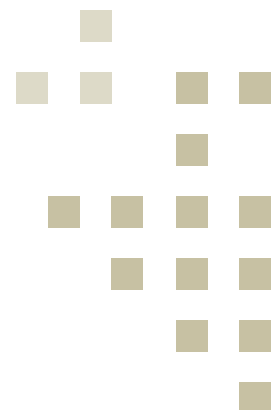
## Sunk costs in the exporting activity: Implications for international trade and specialisation

Arne Melchior

# Sunk costs in the exporting activity: Implications for international trade and specialisation

## Arne Melchior

[Abstract] International trade costs may be sunk and not proportional to sales. The paper explores this theoretically, by allowing firms to invest in sales channels or marketing in order to increase demand in each market. The returns to such investments will, ceteris paribus, be higher in markets with lower variable trade costs (e.g. transport costs). Firms will therefore invest and sell more at home than in foreign markets, and more in foreign markets with low variable trade costs. Sunk export costs will therefore amplify the trade-reducing impact of other trade barriers, and dampen the «home market effect» whereby large countries tend to be net exporters of differentiated goods.

Correspondence: Arne Melchior, Norwegian Institute of International Affairs, P.O. Box 8159 Dep., N-0033 Oslo, Norway. Tel. 47+22056520, e-mail arne.melchior@nupi.no

## Introduction[1]

Empirical research on international trade suggests that international trade is more limited than we should expect from current trade models and known trade barriers. For example, the recent literature on "border effects" in international trade (see, for example, McCallum 1995, Helliwell 1996, Brenton and Vancauteren 2001) shows that trade is much more intense between regions within a country than across borders. This also applies to trade between countries that are fairly well integrated in terms of trade policy. For the EU, Chan (2002) shows that non-tariff barriers explains some of the border effects; but after controlling for these EU countries still trade 4.3 times more "with themselves" than with their EU partners. A second example of the gap between predicted and real trade flows is found in CGE (computable general equilibrium) trade models: When such models are calibrated to fit observed trade patterns, it is necessary to introduce some kind of "trade friction" in the models if they are to fit the data. A frequently used option is the "Armington approach", whereby goods from different countries are by assumption considered as different, with some specific substitution elasticity that makes the models fit. A third example is found in research on the "gravity equation" for international trade: In spite of great expectations about globalisation and the reduced impact of borders and geography, the inverse relationship between geographical distance and the bilateral trade volumes has remained stubbornly stable over recent decades (Brun et al. 2002).

Instead of "missing trade", we therefore have "missing explanations". There are two main strategies to fill this gap: The first one is to undertake more empirical research in order to trace the "hidden barriers". A growing literature in this field now exists, and more needs to be done. There is now evidence on the importance of sunk costs in international trading activity (see, for example, Roberts and Tybout 1997, Medin and Melchior 2002), the importance of ethnic networks and migration for trade flows (for a survey, see in Rauch 2001), and also learning-by-doing in exporting firms (see, for example, Bernard and Wagner 2001).

A second possibility is that known trade barriers may have a stronger impact on trade flows than current trade models suggest. A contribution by this paper is to show that this is possible: When firms are allowed to invest in each market to increase demand for their products, they will invest more in markets with low variable trade costs (including their home market), and thereby sell more in these markets. Fixed trade costs of this type will therefore amplify the trade-reducing impact of trade barriers.

The paper is motivated by empirical research on internationalisation of the information technology industries (Melchior and Øi 2003). In this work, it was observed that some software and IT consultancy firms had huge fixed costs in foreign sales, a low export share, and exports concentrated in geographically close markets. Other IT firms, on the other hand, were selling globally and had lower fixed sales costs. The model presented here is an

attempt to provide a consistent explanation of these contrasting patterns of internationalisation. The mechanisms described are of a general nature and may potentially be relevant in a wide variety of contexts, for example with respect to the analysis of marketing in general, or with respect to the internationalisation of service industries. The model is formulated in the context of international trade, but we may also think of the "countries" in the model as cities or regions; the model describes a general microeconomic mechanism.

The paper adds to the theoretical analysis of fixed trade costs, on which some contributions already exist. Venables (1994) and Medin (2002) examine fixed costs in monopolistic competition models, with assumptions creating an outcome where exporting and non-exporting firms coexist within an industry. Export growth may then occur by new firms being exporters, and in the presence of fixed costs, access to large markets may increase the share of exporting firms in small countries (Medin 2002). Evans (2000) assume that individual firms face different fixed trade costs. Hence only a fraction of the firms will export in equilibrium, and trade will be smaller than in the case when all firms export. In the analysis of Jean (2002), firms face the same fixed trade costs but their marginal production costs differ, and this affects the share of exporting firms.

These contributions assume that fixed trading costs are exogenous; i.e. a given investment of a certain size is required in order to enter a new market. Contrary to this, the model here follows Sutton (1991) by letting the magnitude of market investments be endogenously determined. Hence the paper attempts to derive new results by merging the "endogenous sunk costs" approach of Sutton with a standard modelling framework from international trade theory.

A second way in which this paper deviates from the above-mentioned theoretical contributions is that we do not model firm heterogeneity. While such heterogeneity is certainly more realistic in the light of empirical evidence, we sacrifice this here in order to obtain a tractable model.

In the model to be presented, we assume (realistically) that all firms incur a minimum fixed sales cost, equal to all firms, in each market. This minimum cost may be thought of as travel to meet customers, obtaining market information, adjusting products to national standards or preferences and so on. In addition, firms may invest more in each market (including the domestic market) and thereby increase demand. We may think of advertising to launch products, buying office facilities close to markets, investment in sales offices, investing more in market analysis, and the like. We shall use the term "market investments" for these sunk costs.

Our "endogenous sunk cost" approach implies that firms may choose how much to invest in each market beyond the minimum investment. How much firms invest, depends crucially on how effective these investments are in terms of generating increased demand. In the model, this efficiency is measured by a parameter that lets the impact on demand of a given investment vary from little to large. It is useful to observe that if the impact is close to zero, we approach a model that is qualitatively identical to a standard "home market model" along the lines of Krugman (1980) or Venables (1987). Hence the situation when market investments have no impact on demand provides a benchmark that illustrates how endogenous

sunk costs modify the conclusions compared to the standard literature. In the standard home market model, large countries will have a disproportionately large share of production for differentiated goods. As we shall see, this effect remains in our model but it tends to be eliminated if market investments have a strong impact on demand.

We present the model as a partial equilibrium story for one sector, hence assuming that there are no resource constraints or factor price effects in the model. Firms may expand production by drawing on a homogenous production factor (think of labour) at factor prices that are constant and the same in different countries. When undertaking marketing investments abroad, they may use domestic or foreign resources, but we do not specify this since it is not crucial for showing the basic mechanisms in this market. The model may be expanded by adding another sector and by specifying the factor constraints. We comment on this in section 4.

In Section 2, we set out the model and solve it for the case of two countries. Section 3 derives some implications for the international trade patterns, and generalises the model to many countries in order to examine its relevance for the gravity relationship in international trade. Section 4 discusses some implications and limitations of the model, and concludes.

## The model

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 for X goods

$$(1) \quad X_j = \left[ \Sigma_k \, a_{kj}^{1/e} \; x_{kj}^{\frac{e-1}{e}} \right]^{\frac{e}{e-1}}$$

where $\varepsilon$ 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.[2] With prices $p_{kj}$ for individual varieties, we obtain the demand functions for variety k in market j

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

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 $X_j P_j = M_j$). $P_j$ has the form

$$(3) \quad P_j = \left[ \Sigma_k \, a_{kj} \, p_{kj}^{1-e} \right]^{1/(1-e)}$$

---

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

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:

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

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 $\gamma$ is a measure of the efficiency of these investments; a large $\gamma$ implies that a given investment has a large impact on demand. $a_j$ 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 $a_j=1$ 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 $\gamma>0$.[3] When $\gamma$ 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.

Now all firms in country i (or alternatively 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

$$(5) \quad x_{ij} = F_{ij}^{g} \, p_{ij}^{-e} \, P_{j}^{e-1} M_{j}$$

In addition to their fixed market investments, firms face a fixed production cost f, and constant marginal costs c. They also have to spend resources on real trading 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 market there are no such costs, $t_{ii}=t_{jj}=1$. When selling to a market, the marginal costs are then equal to $ct_{ij}$.[4]

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

---

3    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 model solutions (e.g. equation (13)) apply only with $\gamma>0$, and the case with $\gamma=0$ would then have to be treated separately.

4    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.

$$(6) \quad \boldsymbol{p}_i = -f - \sum_j F_{ij} - \sum_j x_{ij} c \, t_{ij} + \sum_j x_{ij} \, p_{ij}$$

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

We assume that firms first choose how much to invest in each market ($F_{ij}$), and then how much to sell in each market. Maximising profits with respect to $x_{ij}$, given $F_{ij}$, gives the first-order conditions

$$(7) \quad p_{ij} = \frac{e}{e-1} c \, t_{ij}$$

so the price is a mark-up on marginal costs, with the mark-up determined by the elasticity of substitution. This can also be expressed as

$$(7a) \quad \frac{p_{ij} - c \, t_{ij}}{p_{ij}} = \frac{1}{e}$$

which will be useful in the following. This pricing condition is the same as in a standard trade model with monopolistic competition. Due to the presence of fixed trade costs, however, the market outcome will be different.

Maximising profits with respect to $F_{ij}$, given that prices are determined by the mark-up rule above, gives the first-order condition

$$(8) \quad \frac{\partial \boldsymbol{p}_i}{\partial F_{ij}} = -1 - c \, t_{ij} \frac{\partial x_{ij}}{\partial F_{ij}} + p_{ij} \frac{\partial x_{ij}}{\partial F_{ij}} = 0$$

or equivalently

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

From (5) we also obtain:

$$(9) \quad \frac{\partial x_{ij}}{\partial F_{ij}} = g \, F_{ij}^{g-1} \, p_{ij}^{-e} \, P_j^{e-1} M_j$$

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

$$(10) \quad v_{ij} = x_{ij} \, p_{ij} = F_{ij}^{g} \, p_{ij}^{1-e} \, P_{j}^{e-1} \, M_{j}$$

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

$$(11) \quad \mathbf{e} = \mathbf{g} \; F_{ij}^{-1} \, v_{ij} \quad \text{or equivalently} \quad F_{ij} = \frac{\mathbf{g}}{\mathbf{e}} \, v_{ij}$$

Hence at equilibrium, sales in each market are proportional to the market investments.

With monopolistic competition, total profits must be zero. Using (7a) and (11), we must therefore have:

$$(12) \quad \mathbf{p}_i = -f - \sum_j F_{ij} + \sum_j \frac{F_{ij}}{\mathbf{g}} = 0$$

So we obtain

$$(13) \quad \sum_j F_{ij} = \frac{\mathbf{g} \, f}{1 - \mathbf{g}}$$

Hence the more efficient market investments are (i.e. $\gamma$ approaches 1), the larger are the market investments compared to the fixed production cost. Observe that this expression applies to firms in either country, so total market investments are the same for firms in small and large countries. Observe also that total market investments are independent of the level of variable trading costs. Hence there is no "tariff-jumping" effect here, although it is true that trade costs affect the allocation of fixed trading costs across markets.

Using (11) and (13), we may derive the total sales of an individual firm in country i, which will be:

$$(14) \quad \sum_j v_{ij} = \frac{\mathbf{e}}{\mathbf{g}} \sum_j F_{ij} = \frac{\mathbf{e} \, f}{1 - \mathbf{g}}$$

Hence when $\gamma=0$ (market investments have no effect), firm size is $\varepsilon f$, as in a standard model with monopolistic competition. If market investments 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

$$(15) \quad F_{ij}^{1-g} = \frac{g}{e} \, p_{ij}^{1-e} \, P_j^{e-1} \, M_j$$

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

$$(16) \quad F_{ij} = F_{jj} \, t^{\frac{1-e}{1-g}}$$

With equal trade barriers $t_{12}=t_{21}=t$, this applies to either country. Using this as well as (13) and (16), we must then have:

$$(17) \quad F_{ii} + F_{ii} t^{\frac{1-e}{1-g}} = \frac{g \, f}{1-g}$$

and we find that

$$(18) \quad F_{ii} = \frac{g \, f}{1-g} \, \frac{1}{1+t^{\frac{1-e}{1-g}}}$$

and $F_{ij}$ then follows from (16), given that $F_{ii}=F_{jj}$. From (11) we may then derive how much the firms sell in each market.

From these results it is evident that firms invest more in marketing in their home market. Since market investments and sales in each market are proportional (from (11)), it is also clear that firms sell more in their home market. By examining the derivatives of $F_{ii}/F_{ij}$ (equation (16) applies since $F_{ii}=F_{jj}$) we find that firms will be more export oriented if
-    variable trade costs $t$ are low
-    products are not too close substitutes ($\varepsilon$ is low).
-    market investments are not very efficient ($\gamma$ is low)
It is unsurprising that firms are more export oriented if variable trade costs are low. Also the second result is similar to what is obtained in standard trade models with imperfect competition (see e.g. Melchior 1997a): Homogeneous goods sectors are more closed. The reason is mainly that trade costs "bite harder" if products are close substitutes.[5]

More novel is the third result: If market investments are efficient ($\gamma$ is high), these investments will be larger but at the same time more focused on the home market. Hence large market investments correspond to a smaller proportion of international trade. The intuition is that in the presence of variable trading costs, the payoff to market investments will be larger in the home market, and the more so if market investments are more effective.

---

[5]    An implication in standard models is that the "home market advantage" of large countries is stronger for more differentiated goods, for which markets are more open.

Hence the more receptive buyers or consumers are to market investments, the more home market oriented will the firms be.

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, based on the following two equations telling that total sales must add up to total market size in each market:

(19)    $n_1 \; v_{11} + n_2 \; v_{21} = M_1$

(20)    $n_1 \; v_{12} + n_2 \; v_{22} = M_2$

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

(21)    $n_i = \dfrac{1-g}{e\,f} \dfrac{M_i - M_j \, t^d}{1 - t^d}$

where we have used $d=(1-\varepsilon)/(1-\gamma)$ in order to simplify notation. Observe that $d<0$ given our assumptions, and the solution for $n_i$ is identical to the standard home market model when $\gamma=0$.

Now consider the ratio

(22)    $\dfrac{n_i}{n_j} = \dfrac{M_i - M_j \, t^d}{M_j - M_i \, t^d}$

or eqivalently, using $M=M_i/M_j$ as the market size ratio

(22a)    $\dfrac{n_i}{n_j} = \dfrac{M - t^d}{1 - M \, t^d}$

We find that the first- as well as the second-order derivative of $n_i/n_j$ 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. When market investments are very effective ($\gamma$ approaches 1 so the absolute value of d gets large), however, the first-order derivative approaches one and the second-order derivative approaches zero. Hence the home market effect gradually vanishes as market investments become more effective. The sign of the relevant derivatives depend on whether $M=M_i/M_j$ is smaller or larger than 1. In the case when M>1, we have

$$(23) \quad \frac{\partial {}^{n_i}\!\!\diagup\!\!{}_{n_j}}{\partial \boldsymbol{g}} < 0 \quad \text{and} \quad \frac{\partial {}^{n_i}\!\!\diagup\!\!{}_{n_j}}{\partial \boldsymbol{e}} < 0 \quad \text{and} \quad \frac{\partial {}^{n_i}\!\!\diagup\!\!{}_{n_j}}{\partial t} < 0 \quad \text{if M>1}$$

With M<1, all the signs are reversed.

Since our results concerning variable trade costs and the elasticity of substitution are similar to those obtained in the standard model, the focus in the following will be on market investments. As noted before, it is the case that when market investments become more effective (large $\gamma$), investments are also more focused on the home market, so the export share of firms is lower. Hence also in this case, the home market effect is also more pronounced, the more "internationalised" markets are. Diagram 1 illustrates how the ratio $n_i/n_j$ changes as the market size ratio M increases, for three different values of $\gamma$ (0.01, 0.5, 0.99).[6]



Diagram 1: Market investments and the size advantage of large countries
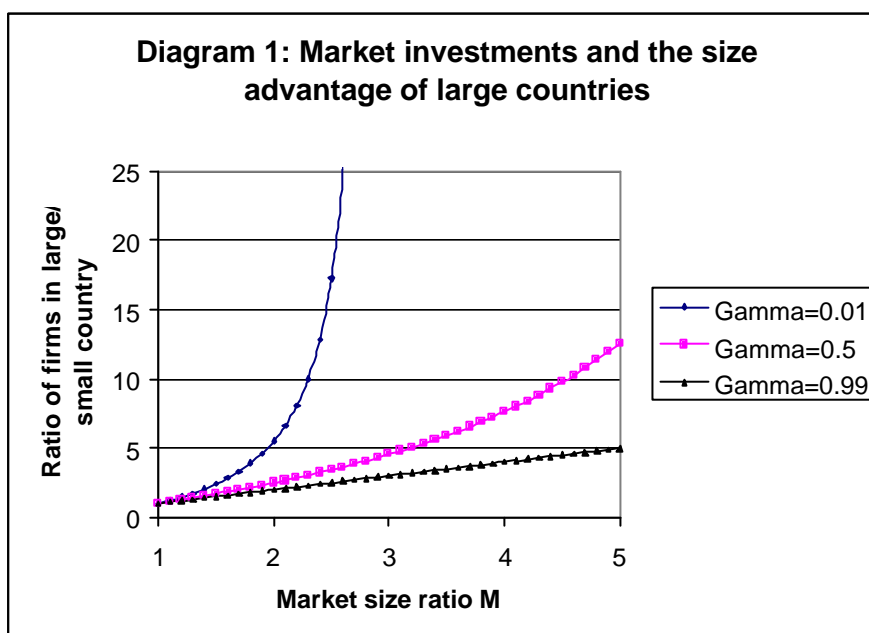
With $\gamma$=0.01 there is an extreme home market advantage for the large country; in fact the smaller country reaches zero production when the size ratio M approaches 3. For an intermediate level of $\gamma$, it is still the case that the large country has a more than proportional share of production, but the impact is more modest. And when $\gamma$ is large (0.99), the number of firms is approximately proportional to country size. Hence the more efficient market investments are in generating increased demand, the weaker will the home market effect be. Large market investments hence correspond to less international trade and a weaker home market effect.

---

[6]    The simulation is based on the following parameter values: $\varepsilon$=5, t=1.3. We let country 1 be the (weakly) larger, and study M$\geq$1.

## Implications for international trade and the gravity relationship

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: An implication is that the elasticity of trade with respect to changes in variable trade costs may be much larger than standard models predict. If we consider the value of international trade in our model, it will be equal to $(n_i+n_j)v_{ij}$. Using the solutions above (and still using the notation $d=(1-\varepsilon)/(1-\gamma)$), we find that total international trade in the sector is equal to

$$(24) \quad T = (M_i + M_j)\frac{g}{e}\frac{t^d}{1+t^d}$$

The elasticity of total trade value with respect to t is then

$$(25) \quad El_t T = \frac{d}{1+t^d}$$

When $\gamma$ increases (remember $0<\gamma<1$ and $d<0$) the denominator approaches one while the absolute value of the numerator becomes larger. So the absolute value of the elasticity is increasing in $\gamma$. Diagram 2 illustrates how $\gamma$ affects the elasticity, for given t (t=1.5 is used in the simulations) and two different values of $\varepsilon$ (3 and 5, respectively).[7]



Diagram 2: The elasticity of international trade with respect to variable trade costs

---
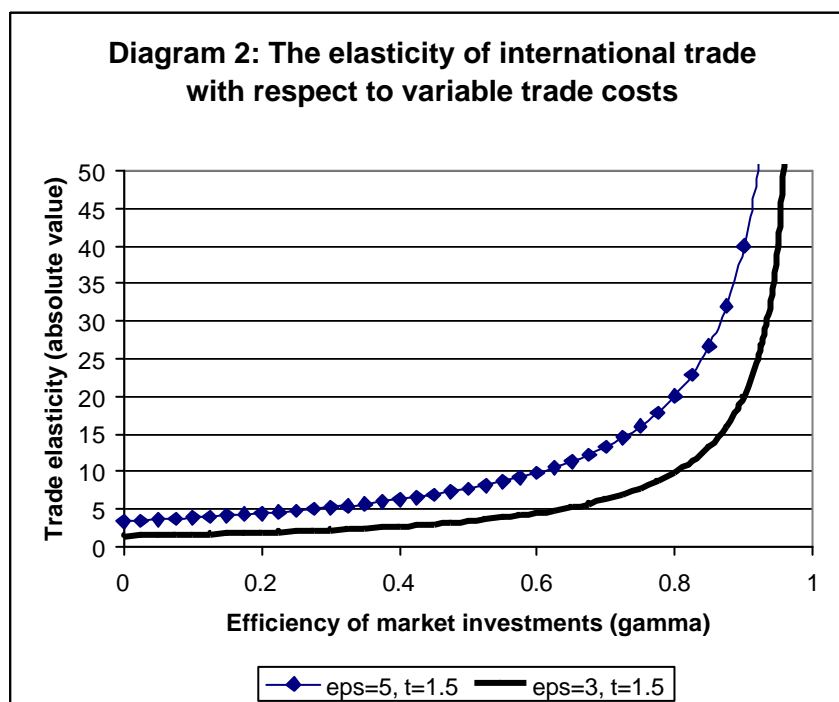
[7]    The elasticity is also increasing in t (the larger t, the closer to 1 is the denominator), but the numerical impact of t on the elasticity is not that large for "normal" ranges of t.

By definition, the trade impact of changed trade costs is stronger, the higher is $\varepsilon$ (as seen from comparing the two curves). The interesting thing here is that the impact also depends strongly on $\gamma$. The more efficient are market investments, the larger is the absolute value of the elasticity. To the extent that such a mechanism is empirically confirmed, it has important implications for the trade pattern.

It is the presence of variable trade costs that makes market investments in the home market more profitable for the firms. In the model above, there is only one foreign market. But if firms in a country has many foreign markets and variable trade costs differ between them, it seems likely that market investments would also vary across markets, depending on the level of trade barriers. This would have important implications for the trade pattern, by "disproportionally" increasing trade with trade partners with low trade costs. Empirical research strongly confirms that bilateral trade flows are inversely related to distance (see, for example, Brun et al. 2002). An interesting extension of the model would therefore be to generalise it to many countries and examine the impact of distance on market investments and the trade pattern. In the following, we shall undertake this in the simplest possible way.

While the solutions for market investments, sales to each market and the number of firms above was undertaken for two countries only, equations (1)-(15) of the model also apply to the general case with n countries. The following analysis thus departs from equation (15). With differing trade costs, equation (16) becomes

$$(26) \quad F_{ij} = F_{jj}\, t_{ij}^{\ d}$$

As the simplest possible "gravity" model, now consider 4 countries that are evenly spaced around the circumference of a circle. Between adjacent countries, we assume that variable trade costs are equal to t (still assuming t>1), and between countries two steps away from each other, trade costs are $t^2$. Each country therefore has two neighbours and one "remote" trade partner. From (13) and (26) we know that the market investments are independent of market size. Since all countries face similar trade barriers, we then know that the firms' market investments in their home markets must be identical in all countries. Using that $F_{ii}=F_{jj}$ for all i, j and (26), equation (13) may then be expressed as:

$$(27) \quad F_{ii}\,(1 + t^d + t^{2d} + t^d) = \gamma f\,/(1-\gamma)$$

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

$$(28) \quad F_{ii} = \frac{\gamma\, f}{(1-\gamma)\left(1 + t^{\ d}\right)^2}$$

This defines home market investments, which will be 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 market investments are therefore lower than in the two-country case, since firms now spread their investments across more markets.

Investments in the foreign markets then follow from (26), 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, d<0). From (11) it is therefore also clear that firms will sell more to the neighbours than to the remote country. This applies to any model where trade costs increase with distance; the point made here is that the distance effect is amplified by the market investments. As shown in Diagram 2 for the two-country case, it also applies here that the elasticity of trade with respect to trade costs is magnified by the presence of fixed market investments. The elasticity of trade with respect to distance will therefore be larger, the more homogenous are the goods, and the more effective are the marketing investments.[8] While the home bias of firms still applies to the multi-country case, we have shown that market investments will be larger in geographically close countries, and hence trade with these countries would be disproportionally larger in sectors where market investments are effective. The model could therefore explain why trade is so much more intense between adjacent countries. The mechanism should not be expected to apply generally, but to specific sectors.

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 strongly focused on the neighbour countries if $\gamma$ is large. If size differences are introduced, bilateral trade in X goods will become imbalanced – at least in some cases. Analysis of market size effects is in this case complicated by the fact that the impact depends on the location of the large countries in space. If one country grows in size, the impact of this 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

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

where $v_{ij}$ – as before – is the value of sales from a firm in country i to market j. Similar equations apply to countries 2-4. Now using (11) and (16), we may express this equation system in matrix form as

---

[8]   For the sake of brevity, we omit a detailed examination of this. The model may also easily be expanded to include any number of countries, but this does not produce more insight here and is therefore dropped.

$$(30) \quad \frac{e}{g} F_{ii} \begin{bmatrix} 1 & t^d & t^{2d} & t^d \\ t^d & 1 & t^d & t^{2d} \\ t^{2d} & t^d & 1 & t^d \\ t^d & t^{2d} & t^d & 1 \end{bmatrix} \times \begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{bmatrix} = \begin{bmatrix} M_1 \\ M_2 \\ M_3 \\ M_4 \end{bmatrix}$$

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

(30a)   $(\varepsilon/\gamma) F_{ii}$ **TN** = **M**

so the solution for the number of firms is then

(31)    **N** $= \gamma/ (\varepsilon F_{ii})$ **T$^{-1}$ M**

The inverse T-1 is equal to

$$(32) \quad T^{-1} = \frac{1}{\left(t^{-d} - t^d\right)^2} \begin{bmatrix} t^{-2d} & -t^{-d} & 1 & -t^{-d} \\ -t^{-d} & t^{-2d} & -t^{-d} & 1 \\ 1 & -t^{-d} & t^{-2d} & -t^{-d} \\ -t^{-d} & 1 & -t^{-d} & t^{-2d} \end{bmatrix}$$

By using (31), (32) and substituting for $F_{ii}$ from (28), the solutions for the number of firms in each country may then easily be found. This exercise is left to the reader.

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 also increases because country 1 grows in size! Diagram 3 simulates this outcome:

Diagram 3: Gravity: The impact on production when one country grows in size
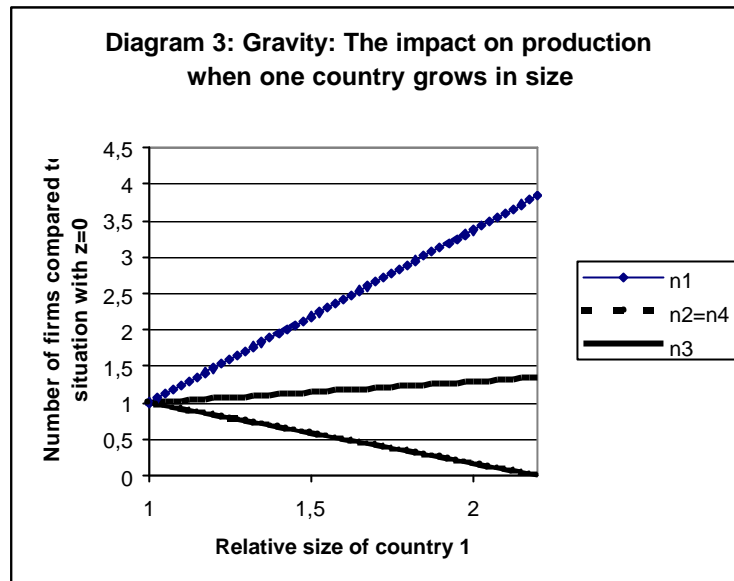
Here the horizontal axis measures the relative size of country 1, starting from 1 where all the four countries are of equal size. The vertical axis measures the relative number of firms in a country, compared to the situation with $z=0$ (all countries of equal size $M_0$). Although the increase in production is smaller for country 3 than for country 1, the remote country gains from the size increase in country 1. Due to geographical distance, the countries in-between act as "buffers" and their industrial decline also benefits country 3. Hence in spatial models, transport costs may have wave-like repercussions on the industrial location pattern, as also shown by Fujita et al. (1999, Chapter 17).[9]

Also in this case, we find that with large $\varepsilon$ or $\gamma$ or t, the home market effect is dampened and the number of firms in each country approaches proportionality with respect to country size. For small values of $\varepsilon$ or $\gamma$ or t, the impact of country size will be larger, but with an important distinction with respect to the number of firms in the "remote" country 3: Small t or $\varepsilon$ 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 $\gamma$, on the other hand, this is not the case – then country 3 will remain in the intermediate range, as in Diagram 3.

This analysis shows that in spatial models, simple conclusions about the relationship between country size, production and the trade pattern are hard to obtain. In the spatial case, the influence of "home market effects" would vary across bilateral trade flows and depend on the spatial location of the larger countries. Consider, for example, the proportion of intra-industry trade (two-way trade in the X sector) between the four countries. In the

---

[9]  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. See also Melchior (1997b, Chapter 3) for partly similar results in a model where countries are located in a one-dimesional space, a "Hotelling" line. In that case, only the neighbours are affected by a size increase for one country. See Fujita et al. (1999, Chapter 17) for more on industrial location in a circular space.

example shown in Diagram 3, this will be largest between countries 2 and 4 (far away from each other), at an intermediate level in all the bilateral trade flows of country 3 (with varying distance), and lowest in the trade between country 1 and its two neighbours. The exact pattern would, however, depend strongly on the spatial distribution of countries of varying size, and we could easily have generated other examples.

In spite of this ambiguity about the impact of country size differences in a gravity model, it still applies that if we consider the share of exports going from a country to different destinations, we have a "clean" inverse relationship between trade flows and distance. This inverse relationship will be stronger, the more effective are the fixed investment in generating increased demand. For the gravity relationship in international trade, an implication is that the elasticity of export shares to different markets trade with respect to distance could be magnified for sectors in which fixed market investments are important. Hence empirical work on gravity should be undertaken at a disaggregated level, in order to trace which models that contribute to explaining the observed trade pattern, and it is an idea to use export shares rather than gross bilateral trade flows at the sector level in order to trace the impact of distance on trade flows.

## Implications and empirical relevance

As demonstrated in the introduction, the model is relevant for the empirical analysis of "border effects" as well as the gravity relationship in international trade. While the empirical evidence on sunk costs in international trade is still limited, it confirms the importance of such costs. Theoretical work along the lines followed here may add to the theoretical foundation for future empirical work.

It should also be observed that a significant share of foreign direct investment is not related to production, but to sales organisation. Hence models that shed light on the motives for such trade-related investments provide hypotheses that may be used for empirical work on such investments.

Empirically, an important issue is whether market investments are exogenous or endogenous. According to the model, exogenous sunk costs related to market entry would increase firm size but have no impact on market shares and the international trade pattern, while endogenous sunk costs will have such an impact. A study on information technology exports (Melchior and Øi 2003) suggests that sunk costs on sales channels as well as product adjustment have an exogenous as well as an exogenous component, and that the endogenous part is more important for the firms' sunk costs related to sales channels. In accordance with the gravity model shown above, firms with large endogenous sunk costs tend to sell more at home and in neighbour countries, while firms with low entry costs have a more "global" export pattern (ibid.). An econometric study based on the same data set also provides support to many of the theoretical predictions derived here (Melchior 2003): The larger are the sunk costs on sales channels, the lower is the export share of firms, and the higher is the national production share in

Norway. Hence small countries such as Norway have a relatively higher share of production in sectors with high entry costs.

While these examples suggest that the model above may have some relevance for empirical work, it is highly stylised and for that reason unrealistic in some respects:

- While the assumption of monopolistic competition enhances the analytical tractability of the model, it creates a stylised outcome with equal-sized firms selling in all markets. If firms were allowed to have non-zero profits, the outcome could be different.

- A specific feature of the model is that market investments do not affect variable trading costs. Hence the model does not fit well if market investments have the effect of lowering variable costs. So production investments motivated by avoiding transport costs would not be well accounted for by the model. Further theoretical work should be undertaken to shed light on this. If a "tariff-jumping" motive for investments abroad is present, the results are likely to be different.

The model has deliberately been presented as an equilibrium analysis of one sector, with exogenous and given production costs.[10] The purpose has been to describe a microeconomic mechanism, which may however be modified if other forces are allowed to work. For example, if production costs or factor prices were allowed to respond to changes in specialisation, this could modify the results. With a two-factor supply side and two sectors, the country specialised in X production could then face a higher price for the factor used intensively in X production, and this would tend to dampen the "home market effect". It is well known that with such a framework, a lowering of variable trade costs will first (from high levels) lead to a more pronounced "home market effect", but when trade costs have become low enough, further reductions will lead to a relocation of production to the smaller country (as cost differences become relatively more important than market access differences) (Krugman and Venables 1990). In Melchior (2003), it is shown that a similar effect applies in the model shown here, if cost differences are introduced.

While the model draws on Sutton (1991), the results show that the total number of firms increases with market size, contrary to the predictions of Sutton on endogenous sunk costs: Sutton showed that with oligopoly and endogenous sunk costs, industrial concentration could increase with market size, due to "investment races". With the assumption of monopolistic competition, this result no longer obtains: In our case, the larger market, the less concentration. This demonstrates that the evolution of concentration also depends on the type of competition, and that endogenous sunk costs *per se* do not define the outcome.

---

10    We could have added a "numeraire sector" with constant returns to scale, a given endowment of labour for each country, and an upper-tier Cobb-Douglas demand function for the consumer choice between the numeraire good and the X aggregate. This would, however, not have added much to the results.

# References

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

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

Brenton, P. and M. Vancauteren, 2001, The Extent of Economic Integration in Europe: Border Effects, Technical Barriers to Trade and Home Bias in Consumption, Brussels: CEPS Working Document No. 171.

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.

Chen, N., 2002, Intra- National Versus International Trade in the European Union: Why Do National Borders Matter? London: Centre for Economic Policy Research, CEPR Discussion Paper No. 3407.

Evans, C.L., 2000, National Border Effects and Heterogeneous Fixed Costs of International Trade, Federal Reserve Bank of New York, Staff Report No. 127.

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.

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

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

Krugman, P. and A.J. Venables, 1990, Integration and the competitiveness of the peripheral industry, pp. 55-77 in Bliss, C. and J. Braga de Macedo (eds.), Unity with Diversity in the European Economy, Cambridge/ London: Cambridge University Press/CEPR.

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

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

Medin, H., 2001, Firms' export decisions. Fixed trade costs and the size of the export market., Oslo: NUPI Working Paper no 618. Forthcoming in Journal of International Economics.

Medin, H. and A. Melchior, 2002, *Learning, Networks and Sunk Costs in International Trade: Evidence from Norwegian Seafood Exports*, Oslo: NUPI Report No. 272.

Melchior, A. and V. Øi, 2003, Born global or local? Technology, market structure and export performance in the IT industries. Evidence from Norwegian Data, Oslo: NUPI, forthcoming.

Melchior, A., 1997a, New theory and evidence on the standard good hypothesis, pp. 75-90, Chapter 5 in Fagerberg, J., L. Lundberg, P. Hansson and A. Melchior (eds.), Technology and international trade, Edward Elgar 1997 (ISBN 1-85898-528-5).

Melchior, A., 1997b, On the economics of market access and international economic integration. Ph.D. dissertation, Oslo: Dept. of Economics, University of Oslo, Økonomiske Doktoravhandlinger No. 36/1997.

Melchior, A., 2003, Export performance and specialisation in the "new economy": Theory, and evidence from the information technology industries. NUPI Working Paper, forthcoming.

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

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.

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

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.