Trade, Reallocations and Productivity: A Bridge between Theory and Data in Öresund*

Anders Akerman
Stockholm University
*** REVISED VERSION ***

July 24, 2009

Abstract

The paper estimates the causal effect of trade liberalisation on aggregate productivity through mechanisms related to firm selection. The construction of a bridge in 2000 across the Öresund Strait linking Copenhagen with Malmö, Sweden's third largest city, provided a natural experiment with which to analyse this effect. A difference-in-difference methodology is applied using both geographic and sectoral variation in how much the bridge affected export patterns and productivity. Firms based in Malmö raise exports to Denmark substantially, mostly by firms selecting into exporting, and the aggregate productivity in Malmö increases. I find that almost all of Malmö's productivity growth is due to the reallocation of production from less productive to more productive firms. When decomposing the productivity gain, I find that these efficiency gains come mostly from the exit of the least productive firms but also from firms with an above-average productivity that start to export and therefore expand their output share. The two largest sectors in Malmö are wholesale trade and manufacturing. Exports by the wholesale sector in Malmö are strongly affected by the bridge whereas those of manufacturing are not. The productivity effects are also the strongest in the wholesale sector.

*I am grateful for comments from seminar participants at the International Monetary Fund, the ETSG Meeting in Athens 2007, NOITS in Bergen 2008, EIIE in Ljubljana 2009 and for helpful comments by Richard Baldwin, Karolina Ekholm, Doireann Fitzgerald, Rikard Forslid, Nicola Fuchs Schündeln, Elhanan Helpman, Martin Bech Holte, Marc Melitz and Diego Puga. Financial support from the Jan Wallander's and Tom Hedelius' Foundation is gratefully acknowledged.

1 Introduction

Understanding the welfare effects of globalisation requires knowing how international trade affects aggregate productivity. This paper estimates the causal effect from a specific mechanism through which this can occur: the reallocation of production among firms that are heterogeneous in productivity levels. This link is important to understand for two reasons. First, aggregate productivity is crucial in determining an economy's welfare since it is a main determinant of real wages. Second, much recent empirical work has highlighted the substantial heterogeneity in firm characteristics in virtually all sectors and countries. Theoretical work has built on this evidence and has shown how this heterogeneity can play an important role in determining aggregate productivity in an open economy.

Melitz (2003)[1] argues that two processes raise aggregate productivity when trade costs decrease: a) the output of non-exporters decreases due to foreign competition, and b) existing exporters expand their output. Because exporters are more productive than non-exporters, a decrease in trade costs reallocates production from less productive to more productive firms. Although ample empirical evidence traces the static differences between exporters and non-exporters, it has been more difficult to test the causality of these links when trade costs decrease. This stems from the difficulty in finding exogenous reductions in trade costs. Trade agreements and trade liberalisation policy are rarely determined in isolation from unobserved industry variables; moreover, trade liberalisation is often associated with other types of economic liberalisation in time.

This paper provides an analysis of how trade affects aggregate productivity through firm selection, using a truly exogenous change in trade costs. It examines the natural experiment arising from the 2000 construction of the Öresund bridge linking the Swedish city Malmö with the Danish capital Copenhagen.

[1] Other important theoretical contributions showing similar dynamics include Bernard, Eaton, Jensen, and Kortum (2003) and Yeaple (2005).

Data at the firm level is available for all Swedish enterprises during the years before and after construction, making it possible to study how dynamics at the firm level drive aggregate variables. Before 2000, the Öresund Strait separating Sweden and Denmark could only be crossed by ferry or commuter passenger boats; now, however, the trip can be made in 20 minutes by car or lorry. The fall in trade costs caused by the bridge is exogenous for two main reasons. First, its construction was decided six years prior to its completion and can therefore not be related to temporary unobserved shocks in productivity. Second, except for the fact that sectors depending on road transport benefit the most, a bridge cannot intentionally be designed to favour trade in certain sectors more than others (trade policy, for example, is often designed to protect vulnerable and less competitive sectors). Moreover, macroeconomic policy and regulation related to productivity, such as labour market and other industry regulation, are relatively stable in Sweden and uniform across the country. This makes it possible to compare Malmö, which is situated right next to the bridgehead, with the two other large cities, Stockholm and Gothenburg (located about 610 and 270 kilometers away).

I apply a difference-in-difference methodology using geographic and sectoral variation. Exports to Denmark by firms in Malmö increase by more than by firms in other cities and the main sector that is affected is the wholesale industry; manufacturing seems not to be affected (wholesale and manufacturing account for about 70% of output in these cities). I find evidence of a causal effect of trade liberalisation on aggregate productivity. The main mechanism that drives aggregate productivity is the reallocation of production across existing firms; the efficiency of allocation improves. More specifically, the least productive firms exit and firms with high productivity enter into exporting and therefore capture a larger output share.

I begin by describing the previous literature and empirical issues involved in estimating causal effects of international trade in Section 2. Section 3 describes the natural experiment and how it can be used to overcome many of the empirical problems discussed in Section 2. Section 4 presents the model, which is based on Melitz (2003)[2], and derives the main predictions from falling trade costs according to the model. Section 5 describes my method for estimating firm productivity. I use the semiparametric method from Levinsohn and Petrin (2003) to find unbiased estimates of sector specific production functions that generate measures of total factor productivity for each firm and year. Section 6 describes the data set and finds that patterns of heterogeneity related to trade and productivity observed in previous literature exist also in the Swedish data. Section 7 describes my empirical method and the results. Section 8 concludes the paper.

2 Previous literature

The literature on how trade drives aggregate productivity through firm selection originates in an empirical literature that has focused on how firm characteristics correlate with trade patterns. These results are by now firmly established. Bernard and Jensen (2004), Aw, Chen, and Roberts (2001) and Clerides, Lach, and Tybout (1998) among others find that exporters are on average more productive and larger than non-exporters[3]. Also, the effect of turnover of firms (in the sense of firm entry and exit) on aggregate productivity growth is large as shown in Foster, Haltiwanger, and Krizan (1998). They show that the simplifying assumption that the economy consists of identical firms disregards dynamics related to firm heterogeneity that affect aggregate productivity levels. It is also shown in Bernard and Jensen (2005) that significant fixed costs are involved in exporting, and that these are larger than fixed costs involved in entering the firm's domestic market. Roberts and Tybout (1997) estimates these costs for Colombia and find them to be of substantial size. Finally, Helpman, Melitz, and

[2] My simplifying assumptions are from Helpman, Melitz, and Yeaple (2004) and consist of using a homogenous sector that equalises wages across countries and parameterising the probability distribution of productivity across firms to a Pareto distribution.

[3] See Tybout (2003) for a survey and also Schank, Schnabel, and Wagner (2007) for a literature overview on wage premia paid to workers in exporting firms in a large set of countries.

Rubinstein (2008) show that firm heterogeneity are important also for understanding aggregate trade flows since the productivity distributions of firms help explaining bilateral trade links without any trade at all. Several papers have built general equilibrium models incorporating these facts, such as Bernard, Eaton, Jensen, and Kortum (2003), Yeaple (2005) and Melitz (2003). The common feature used for this analysis is that they build on firm selection as the driving channel for how trade affects aggregate productivity. What this literature shows is that firm heterogeneity in productivity is of substantial size, that certain patterns relating to trade hold across countries and that it is important to incorporate firm heterogeneity to explain trade patterns and the evolution of aggregate productivity.

There have been previous studies on how trade liberalisation affects productivity through the reallocation of production. Pavcnik (2002) estimates the effect of trade liberalisation in Chile in the 1970s. During this period, Chile liberalised its trade regulation dramatically and she finds strong effects on reallocation of production across firms in import-competing sectors. Also, Loecker and Konings (2006) find similar results for Slovenia in the 1990s, where the main drivers are job destruction at state firms and reallocation of employment to private firms. Trefler (2004) estimates the effect of the Free Trade Agreement (FTA) between the United States and Canada. This paper has many advantages. Industry in all of Canada was affected by the FTA and the cut in tariff levels was substantial. However, I argue that my experiment builds on a fall in trade costs that was more exogenous in nature than in the papers mentioned here. I will outline my argument in the following passages.

Empirical research on the effect of a trade liberalisation is often susceptible to a problem of endogeneity. The decision to reduce tariffs is likely to coincide with other reforms that have effects on production choices and productivity. Several countries have, for example, liberalised trade at the same time as they have removed domestic regulatory and taxation barriers to productivity growth such as India in 1990s. Using these would overestimate the true effect due to the inclusion of productivity increases that come from other sources than lower trade costs alone.

Also, trade policy is frequently effected by industry lobbying (see for example Grossman and Helpman (1994) and Koujianou Goldberg and Maggi (1999)). The size of a tariff cut in a specific sector could in such a setting be correlated with other characteristics that affect production and productivity. This would bias the estimates for the effect of trade liberalisation on productivity. For example, a sector with, in an international context, relatively unproductive firms may lobby its government to retain tariff protection while a sector with some very productive firms may want the government to sign free trade treaties with countries to which these firms can successfully export. Both of these problems would bias the estimates upward. I claim, however, that the natural experiment in my paper provides a setting where the fall in trade costs is more exogenous than in the previously mentioned papers since policy plays a very limited role. Both the timing and the selection of industries to be exposed are very unlikely to be affected by unobserved shocks to productivity in the affected region around the time of the bridge being opened. This is discussed in more detail in the following section.

3 The Öresund bridge as a natural experiment

The natural experiment used in this study is the construction of the Öresund bridge. The Öresund bridge connects the Danish capital of Copenhagen with Sweden's third largest city of Malmö. These two cities are separated by the Öresund Strait which was previously only connected by large ferries and somewhat smaller but faster boats that only carried people. The region was, however, already before the bridge was constructed, well integrated due to free trade, similar cultures, similar and mutually comprehensible languages and the fact that passing the border did not require a passport. In 2000, the metropolitan population of Copenhagen was 495,699 and that of Malmö 259,579.[4]

There are several advantages of this experiment. First, the timing can be assumed to be exogenous for reasons that will be discussed shortly. Second, the countries involved have similar relative factor endowments and access to similar technology as shown by the similarities in the GDP per capita levels. Firm dynamics can therefore be assumed to be related to intraindustry trade which is the focus of Melitz (2003). Third, there is firm-level data available for Swedish firms for both the time before and after the bridge was constructed.

Regarding the exact timing of the bridge, it is important to note that whether to build the bridge was a controversial issue in both Danish and Swedish politics. It was discussed for a very long time. But it was only in 1991 that decisions in the Danish and Swedish parliaments were taken. However, due to public concern and tensions within the Swedish coalition government regarding possible environmental effects, the bridge was only accepted by the Swedish government in 1994 and finally completed in June 2000. The long construction time is useful for this paper because politicians at the time in 1994 and even less in 1991 could only have had a very vague idea of what the business cycle would be in 2000 and what other industrial policies would be implemented in 2000. Policy makers could not know how such variables would change from 1998 to 2002, which is the sample period that I will focus on. I therefore claim the timing of the construction of the bridge to be largely exogenous.

The endogeneity problem when the selection of which sectors that are exposed is correlated with their productivity levels is unlikely to be present in this experiment. It is difficult to design a bridge so that sectors are protected in a way that is related to variables affecting productivity. A lorry carrying goods from any industry is able to drive across the bridge. It is probably the case, however, that a bridge of this type favours certain sectors more than others. Some trade, for example, is simply transported through information technology communication and not subject to any change due to a bridge. Sectors with

[4] Statistics Denmark and Statistics Sweden.

goods that are very costly to transport by sea but not by road might, however, be greatly effected. But since which sectors benefit from a bridge is largely exogenous, the productivity levels in those sectors at the time of construction cannot be related to the bridge actually being constructed.

A possible criticism specifically against the experiment used here could be that rational firms change their behaviour in advance since they are expecting the bridge to open in 2000. This might mean that I do not fully capture the true effect of the bridge by simply looking at levels before and after the experiment happened. This would cause a downward bias in the results. In terms of the significance of an effect existing at all, however, it is not that problematic since this phenomenon would only make it more difficult to find significant results.

Another issue is that labour markets in the region change. There is strong evidence of this in the region. During the very first years, this mostly involved skilled Swedish labour starting to work in Copenhagen, somewhat draining firms in Malmö of skilled labour. Since I use Swedish data this factor would, again, make it more difficult for me to find significant results. If I find productivity to increase in Malmö due to an increased exposure to foreign competition and a larger market, this would happen despite any movement of skilled labour to Copenhagen.

The Malmö city region is therefore selected as the "treatment" area since it is the geographic area closest to the bridgehead. In proportional terms, the largest decrease in trade costs should be for areas closest to the bridge. The control group should be one that faces the same macroeconomic or national shocks as Malmö. I therefore pick the two largest cities in Sweden: Gothenburg and Stockholm (located about 270 and 670 kilometers from the bridge, respectively). Any national policy or national economic shock should theoretically affect Malmö, Gothenburg and Stockholm in the same direction. Other control groups could, of course, be considered. But these three cities are usually referred to as the only "large" cities of Sweden. Any smaller towns would probably be much more limited as regards what sectors are active there Also, using only firms located within the city boundary somewhat narrows down the sample since large cities are typically surrounded by smaller cities that participate in the economy of the big city. However, it would be difficult to expand the definition of Malmö, Stockholm and Gothenburg to include surrounding cities without introducing an arbitrary and questionable selection system. Where the boundary of the "larger city area" ends is difficult to define in a consistent way. Therefore only firms actually registered in the city area of the three cities are included.

Finally, it could be argued that this experiment is more focused on the effect of infrastructure improvements and cuts in transport costs rather than the effect of tariff cuts. It is true to the extent that the paper uses a decrease in transport costs but in terms of external validity this is not so problematic. Theoretically in the trade literature, the effect of transport costs versus tariffs of a variable type is almost never treated differently. This is because transport costs map into the cost of trade in exactly the same way as tariffs. As long as the cost is proportional to the value of goods being transported, it does not matter for the actions of firms whether the costs is accrued to the government or spent on the actual transportation of the good.

4 Model

4.1 Basic setup

The model outlined here is a simplified version of Melitz (2003) in the sense that it uses a homogenous sector (here labelled as agriculture) with constant returns to scale and parameterises the probability distribution of productivity to a Pareto distribution[5]. The model has two countries, Home and Foreign (marked with an asterix "*"). There is one factor of production, L and L^*. Production is composed of two different industries: agriculture, A, and manu-

[5] Helpman, Melitz, and Yeaple (2004) introduces these assumptions as well as foreign direct investment (which I do not).

Figure 1: Map of the Öresund region.

facturing, M. The agricultural sector is characterised by costless trade, a unit input requirement of labour and constant returns to scale. The prices of the homogenous agricultural good is therefore equalised across countries and chosen as the numeraire: $p^A = p^{A*} = 1$. Due to the unit labour requirement it also follows that $w = w^* = 1$. The manufacturing sector is characterised by increasing returns due to positive fixed costs and has a monopolistic competition setting as in Dixit and Stiglitz (1977). The marginal cost is heterogeneous across firms. International trade in manufacturing varieties is costly with an iceberg trade cost of $\tau > 1$.[6]

There are three types of fixed costs that firms have to pay: f_E is the innovation cost for starting a new firm, $\widehat{f}_X$ is a fixed cost a firm has to pay only if they export to the other country and $\widehat{f}_D$ is a fixed cost which the firm has to pay for entering the domestic market. The marginal cost is wa_i where a_i is the labour requirement for producing one unit of the good x_i and is specific to firm i. In equilibrium, there is a constant rate of entry and exit of firms due to an exogenous per period death probability $\delta < 1$. Potential entrants do not know their labour requirement a_i but instead face an *ex ante* probability distribution $G(a)$.

[6] A firm has to ship τy units for y unites to arrive at the destination.

Having sunk the entry cost, f_E, firms learn their labour requirement and decide whether they exit immediately, serve the domestic market (for which they have to sink an entry cost of $\widehat{f}_D$) or serve both the domestic and foreign markets (for which they have to invest both $\widehat{f}_D$ and $\widehat{f}_X$). The value of a firm, in the absence of intertemporal discounting, is bounded from above by the exogenous death probability δ.

Consumers in both countries derive utility from, first, an upper tier (Cobb-Douglas) combination of manufacturing and agricultural goods:

$$U = C_M^{\mu} C_A^{1-\mu} \tag{1}$$

where C_A is the consumption of the agricultural homogenous good. C_M is an index of consumption of manufacturing varieties as in Dixit and Stiglitz (1977):

$$C_M = \left[\int_0^n c_i^{(\sigma-1)/\sigma} di \right]^{\sigma/(\sigma-1)} \tag{2}$$

where n is the mass of varieties consumed, c_i the amount of variety i consumed and $\sigma > 1$ the elasticity of substitution. Consumers spend μ of their income on the manufacturing varieties and this generates the following demand functions for the domestic firm i:

$$x_i = \frac{p_i^{-\sigma}}{P^{1-\sigma}} \mu Y, \qquad x_i^* = \frac{(\tau p_i)^{-\sigma}}{P^{*1-\sigma}} \mu Y^*$$

where x_i and x_i^* are the demand levels of firm i's good in Home and Foreign, respectively. p_i is the producer price of variety i, Y is income in Home and Y^* income in Foreign. $P \equiv \left(\int_0^n p_i^{1-\sigma} di + \int_0^{n_X^*} (\tau p_i^*)^{1-\sigma} di \right)^{\frac{1}{1-\sigma}}$ is the price index of manufacturing goods in Home where n is the number of domestic firms serving the domestic market and n_X^* the number of foreign firms exporting to Home. This setup generates the following optimal pricing rules for firms:

$$p_{ijj} = \frac{\sigma}{\sigma - 1} a_i, \qquad p_{ijl} = \frac{\sigma}{\sigma - 1} \tau a_i \tag{3}$$

where p_{ijl} is the price for firm i based in country j but selling the good in country l.

It can be shown that an equilibrium is characterised by six equations. The first four consist of "cutoff" equations showing the labour requirement a_i for a firm that is just productive enough to serve the domestic or both markets so that the operating profit is equal to the per period equivalent of the fixed costs involved in serving that market. This has to hold for firms in both countries.

$$a_D^{1-\sigma} B = f_D \tag{4}$$

$$a_X^{1-\sigma} \phi B^* = f_X \tag{5}$$

$$a_D^{*1-\sigma} B^* = f_D \tag{6}$$

$$a_X^{*1-\sigma} \phi B = f_X \tag{7}$$

where $f_D \equiv \delta\sigma \widehat{f}_D$, $f_X \equiv \delta\sigma \widehat{f}_X$, $B = \frac{\mu L}{P^{1-\sigma}}$, $B^* = \frac{\mu L^*}{P^{*(1-\sigma)}}$, and $\phi \equiv \tau^{1-\sigma} \in [0, 1]$ represents trade freeness. a_D denotes the cutoff productivity to enter the domestic market and a_X is the equivalent for the export market. It is assumed that $f_X > \phi f_D$ always holds. The term B is best understood as a "per firm market potential". Free entry into both markets generate two more equations stating that for the potential entrant, the expected profit must equal the entry cost f_E.

$$\int_0^{a_D} \left(a_D^{1-\sigma} B - f_D\right) dG(a) + \int_0^{a_X} \left(\phi a_X^{1-\sigma} B^* - f_X\right) dG(a) = f_E, \tag{8}$$

$$\int_0^{a_D^*} \left(a_D^{*(1-\sigma)} B^* - f_D\right) dG(a) + \int_0^{a_X^*} \left(\phi a_X^{*(1-\sigma)} B - f_X\right) dG(a) = f_E. \tag{9}$$

4.2 Solution to the long run equilibrium

The probability distribution $G(a)$ determining the labour requirement is assumed to follow a Pareto distribution. This is in line with empirical evidence, as for example in Axtell (2001).

$$G(a) = a^k. \tag{10}$$

It is assumed that $\beta \equiv \frac{k}{\sigma-1} > 1$ to ensure that the integrals in the free entry conditions (8) and (9) are bounded. Using this in the free entry conditions and solving the equilibrium generates the following expressions for the endogenous variables:

$$a_D^k = a_D^{*k} = (\beta - 1) \frac{f_E}{f_D} \frac{1}{1+\phi^\beta} \tag{11}$$

$$a_X^k = a_X^{*k} = (\beta - 1) \frac{f_E}{f_X} \frac{\phi^\beta}{1+\phi^\beta} \tag{12}$$

$$B = B^* = \left((\beta - 1) \frac{f_E f_D^{\beta-1}}{1+\phi^\beta} \right)^{\frac{1}{\beta}}. \tag{13}$$

The variables do not differ across countries, even if these differ in size, because all cost parameters and determinants of the distribution are assumed to be the same across countries. It can also be seen that since $f_X > \phi f_D$:

$$a_D > a_X.$$

meaning that the marginal productivity for surviving in the home market is lower than the marginal productivity needed to be able to export.[7] It follows from this that exporters (which always have $a \leq a_X$) are always more productive than non-exporters (which have $a_X < a \leq a_D$).

The revenues from selling to the domestic market and the export market

[7] This is not a strong assumption since it is generally viewed as more expensive to enter a foreign and less well-known market than the home market.

are:

$$y_D = a^{1-\sigma} B = \gamma a^{1-\sigma} \left(\frac{1}{1+\phi^\beta} \right)^{\frac{1}{\beta}} \tag{14}$$

$$y_X = a^{1-\sigma} \phi B^* = \gamma a^{1-\sigma} \left(\frac{\phi^\beta}{1+\phi^\beta} \right)^{\frac{1}{\beta}} \tag{15}$$

where $\gamma \equiv \left[(\beta - 1) f_E f_D^{\beta-1} \right]^{\frac{1}{\beta}}$. The combined revenue of an exporter (since $a_X < a_D$ all firms that export also serve the domestic market) is therefore

$$y_D + y_X = \gamma a^{1-\sigma} \frac{1+\phi}{\left(1+\phi^\beta\right)^{\frac{1}{\beta}}}. \tag{16}$$

Finally, using the price index together with the free entry conditions (8) and (9) and the solutions to the productivity cutoffs and B and B^* I find the expressions for the number of firms:

$$n = \frac{\beta - 1}{\beta} \mu \frac{1}{f_D} \frac{(L - L^* \Omega)}{(1 - \Omega^2)} \tag{17}$$

$$n^* = \frac{\beta - 1}{\beta} \mu \frac{1}{f_D} \frac{(L^* - L \Omega)}{(1 - \Omega^2)} \tag{18}$$

where $\Omega \equiv \phi^\beta \left(\frac{f_X}{f_D} \right)^{1-\beta}$.

To derive the expression for total M-sector trade, I note that the export level of an existing exporter is

$$y_X(a) = a^{1-\sigma} \phi B^* = \left(\frac{a}{a_X} \right)^{1-\sigma} f_X$$

due to (5). Integrating from 0 to a_X and multiplying with the number of firms as in (17), I find that the total value of M-sector exports from home, V_T, is

$$V_T = n \int_0^{a_X} y_X(a) \, dG(a | a < a_D) = \mu \left(\frac{f_X}{f_D} \right)^{\beta - 1} \Omega \frac{(L - L^* \Omega)}{(1 - \Omega^2)} \tag{19}$$

As opposed to the variables analysed so far, total exports depend also on country size (L and L^*). This is since population size affects the total number of firms in each country's manufacturing industry due to the "home market effect".[8]

4.3 Comparative statics when trade costs decrease

It is now possible to analyse what happens when trade costs decrease (here characterised by an increase in trade freeness, $d\phi > 0$). First, the cutoffs for survival (a_D) and exporting (a_X) change in the following way:

$$\frac{da_D^k}{d\phi} = \frac{da_D^{*k}}{d\phi} = -\beta(\beta-1)\frac{f_E}{f_D}\frac{\phi^{\beta-1}}{(1+\phi^\beta)^2} < 0. \tag{20}$$

$$\frac{da_X^k}{d\phi} = \frac{da_X^{*k}}{d\phi} = \beta(\beta-1)\frac{f_E}{f_X}\frac{\phi^{\beta-1}}{(1+\phi^\beta)^2} > 0. \tag{21}$$

It is the case that a_D decreases in trade freeness, meaning that the least productive firms exit the economy when trade is liberalised. It is often argued that this takes place due to the increase in foreign competition but what is technically operating in the model is that there is an increase in the demand for factors of production from existing exporters which now expand production.

The cut-off for exporting increases which means that new firms enter the export market. These "switchers" have productivity levels which were marginally below the previous cut-off for exporting but are now productive enough to engage in exporting.

It can also be seen that the allocation of production changes from non-exporters to exporters. Equation (14) show that domestic sales, y_D, decreases in ϕ and (15) shows that foreign sales, y_X, increases in ϕ. Non-exporters therefore lose market share. Exporters, however, more than makes up for the loss of domestic sales by increasing exports. The derivative of $y_D + y_X$ as in (16) with regard to ϕ is shown in the appendix to be positive.

These are essentially the processes that raise aggregate productivity in Melitz (2003). Production is reallocated from non-exporters to exporters and the least

[8] See Krugman (1980).

productive non-exporters exit production.

Formally, I denote aggregate productivity in home by Ψ.

$$\Psi = \left(s_D \int_0^{a_D} a^{1-\sigma} dG(a \,|a < a_D) + s_X \int_0^{a_X} a^{1-\sigma} dG(a \,|a < a_D) \right)^{\frac{1}{\sigma-1}} \tag{22}$$

where s_D is the share of non-exporters and s_X is the share of exporters. The ratio of exporters to non-exporters is $\left(\frac{a_X}{a_D}\right)^k = \frac{f_D}{f_X}\phi^\beta$ so it follows that $s_D = \frac{1}{1+\frac{f_D}{f_X}\phi^\beta}$, and $s_X = \frac{\frac{f_D}{f_X}\phi^\beta}{1+\frac{f_D}{f_X}\phi^\beta}$. It is shown in the appendix that $\frac{d\Psi}{d\phi} > 0$ implying that aggregate productivity increases when trade costs decrease.

It can be seen in (19) that lower trade costs increase total manufacturing exports if the two countries are not too asymmetric in size. This is easiest shown if I assume that the two countries are of equal size:

$$\left.\frac{dV_T}{d\phi}\right|_{L=L^*} = \frac{d}{d\phi}\left[\mu L \left(\frac{f_X}{f_D}\right)^{\beta-1} \frac{\Omega}{(1+\Omega)}\right] > 0 \tag{23}$$

since $\frac{d\Omega}{d\phi} > 0$.

4.4 Predictions

The main analytical predictions for trade and aggregate productivity can now be summarised. As regards trade, exports always increase at the so-called intensive margin among existing exporters and among new exporters at the extensive margin. Moreover, if size asymmetries are not too large, exports increase also at the aggregate level. Aggregate productivity increases through the reallocation of production from exiting firms and non-exporters to existing exporters.

5 Estimating firm productivity

Productivity in this analysis will be estimated by using an estimated production function to find the total factor productivity of each firm and year. To find this, I follow a semiparametric method outlined in Levinsohn and Petrin (2003).

Having estimated production function coefficients for labour and capital, productivity, p_{ist}, is defined as

$$p_{ist} = \exp\left(va_{ist} - \beta_l l_{ist} - \beta_k k_{ist}\right).$$

where va_{ist} is the value added of firm i in sector s in year t, l_{ist} is labour used, k_{ist} is capital used. The estimation technique uses intermediate inputs (energy and raw materials) as a proxy to control for the correlation between input levels of factors of production and unobserved productivity shocks.

There are many different ways of estimating firm productivity.[9] To calculate the total factor productivity of a firm there is a need to estimate the firm's production function. This involves some well known problems. There is, first, a problem of simultaneity because the productivity level of a firm could have an effect on its choice of factor input levels. Olley and Pakes (1996) finds evidence of this using data from the US telecommunications industry in 1980s. For example, with time varying productivity, a firm that experiences a positive shock can be assumed to increase its production temporarily to reap the higher profits available. If it can do this immediately by increasing for example labour (as opposed to capital, which might be assumed to take longer time to accumulate), the coefficient on labour in the estimated production will have a positive bias.[10]

Also, productivity affects entry and exit decisions by firms. Much research has so far used "balanced panels" when estimating production functions and has thereby excluded all new entrants and exiting firms. This methodology is, however, prone to be subject to selection bias if the productivity of a firm affects its decision of staying or exiting. If, for each productivity shock in time, the exit decision is correlated with the current capital level (it can been argued that firms with more capital may be less likely to default given productivity) there will be a negative bias on the coefficient on capital in the production function if the balanced panel is used. These two biases are the reasons for

[9] See Arnold (2005) for an excellent and practical summary.

[10] This problem was first highlighted by Marschak and Andrews (1944).

using semiparametric methods as in Olley and Pakes (1996) or Levinsohn and Petrin (2003). The methods are fairly similar but Olley and Pakes (1996) uses investment instead of intermediate inputs as its instrument. However, as is often the case in many other datasets, there are more observations in my sample that report non-zero or non-missing values for intermediate inputs than for investment. So since the method by Levinsohn and Petrin (2003) gives me more observations for my dataset, I opt for this method. Production functions are estimated at the 2 digit industry level. A more disaggregated industry level would provide two few observations for many sectors.

6 Data

All data I use is from the Swedish database "Företagsdatabasen" except for trade data which is collected by the Swedish Customs. Both datasets have been provided by Statistics Sweden. It contains information on all active Swedish firms during the time period 1996 to 2002.[11] In the absence of industry-specific price indices, I deflate all variables measured in monetary values by the national Swedish price level. The location variable is in what city ("kommun") the firm is registered so unfortunately there is not information about the location of plants. However, for the purpose of the experiment I argue that, on average, more firms that have most of their plants close to the bridge should also be registered in Malmö than in Gothenburg or Stockholm.

Generally, the data show that similar relationships between trade and productivity as found in the literature for other countries (regarding the difference between exporters and non-exporters, export intensity and which industries that export) hold also in Sweden. There is clear evidence of heterogeneity among firms which is a key motivation for the model developed by Melitz (2003). In

[11] I exclude some sectors, either because they are not affected by changes in trade levels or that they are not relevant. These include the financial and public sectors as well as utilities and construction. There is also a small subset of observations for which I cannot estimate productivity due to the lack of data on either capital, labour or value added. These observations, which account for 6% of all observations with exports, are dropped from the analysis.

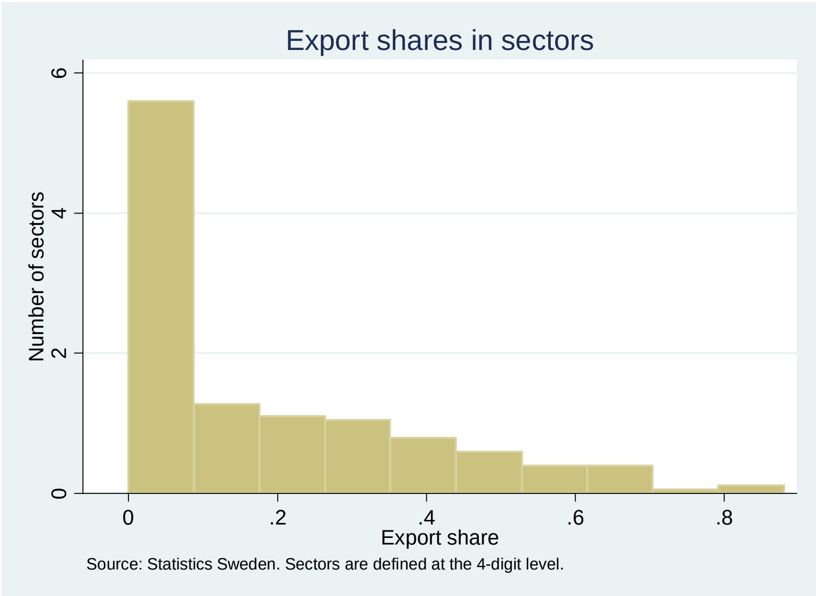

Figure 2: Export shares in sectors.

the year of 1999, the year just before the bridge was constructed, the standard deviation of the log of sales of firms in Sweden was 1.98. When only including variation within sectors, the variance is still as high as 1.84. In terms of productivity there is also substantial variance. The variance of the log of productivity is 0.65 and when only using variation within sectors it is 0.62.

In my sample, 26% of manufacturing firms export and 20% when nonmanufacturing sectors are included. However, of firms with at least 50 employees, 90% of manufacturing firms and 77% of all firms export. There is no cutoff in firm size for exporting, except that total exports to an EU country has to exceed SEK 1.5 million to be recorded. Most exporters, as is typical in most countries, only export a small share of their output. Moreover, most sectors export only a smaller share of their output which confirms the absence of a clear division into exporting and nonexporting sectors; most sectors serve both the Swedish and the foreign market as can be seen in Figure 2. Among firms in 1999, I find that exporters are on average 26% more productive than non-exporters (33% in manufacturing). This is in line with research using U.S. data; Bernard and

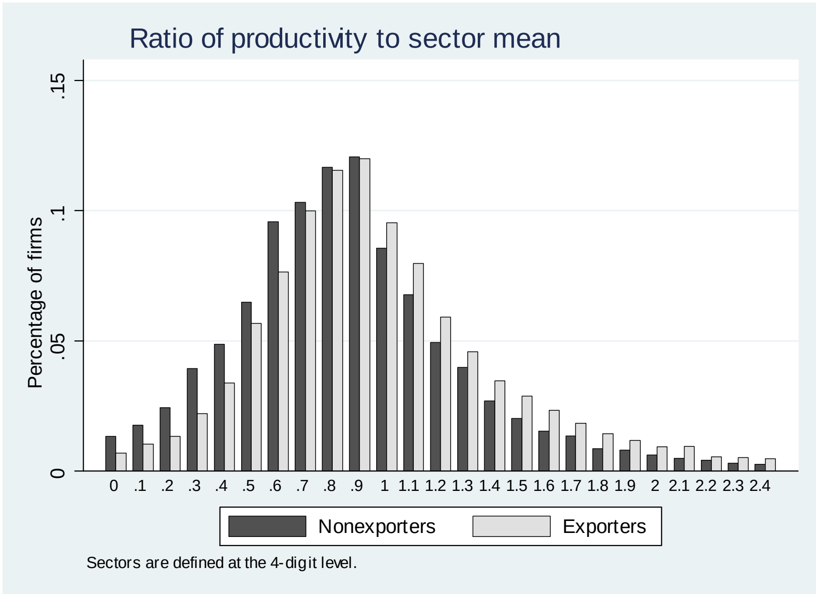

Figure 3: Ratio of productivity to sector mean.

Jensen (2004) found a 39% difference in US manufacturing data. This pattern can also be seen in more detail in Figure 3 where I divide the productivity of each firm by the mean in its sector and compare non-exporters with exporters. There are more exporters than non-exporters in all categories above 1.

Finally, a common finding is that exporters have a lower probability of being shut down. This can be seen as indicative of sunk costs being involved in exporting, which is assumed in Melitz (2003). Bernard and Jensen (2005) found in US data that, after introducing controls that explain plant shutdowns, there is a 5% reduction in the probability of the plant being shut down if the plant is exporting. Although the Swedish data is at the firm and not the plant level, I find a 4% reduction in shutdown probability for exporters as compared to non-exporters.

These findings tell us that there are indeed clear signs of heterogeneity among Swedish firms, that exporters and non-exporters operate within the same sectors and that there is an indication of fixed costs involved in exporting.

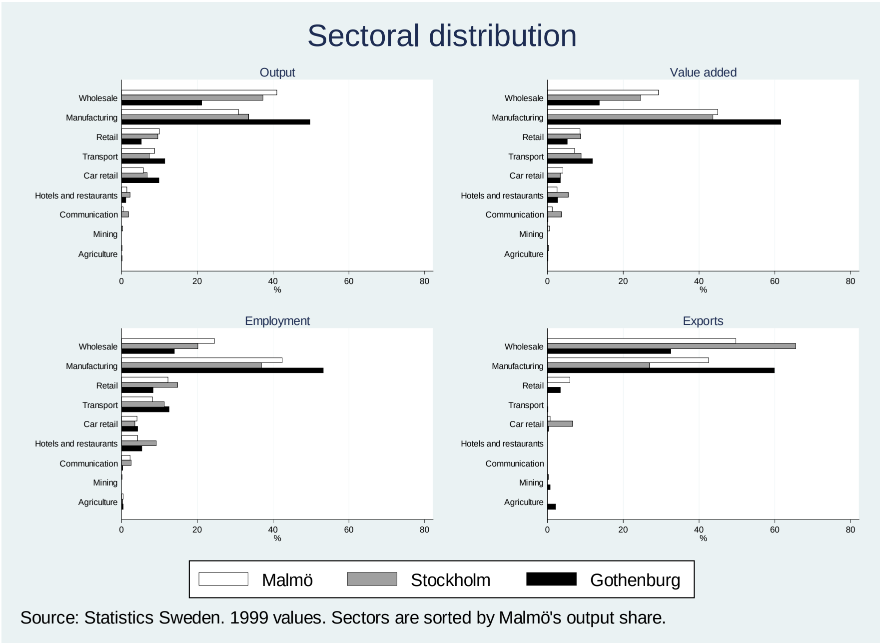

Figure 4: Sectoral distribution.

6.1 Comparability of cities

Figure 4 shows the sectoral composition of the three economies used for the study. It uses the year 1999 since it is the year immediately before the bridge is built. The industrial structure is relatively similar. Exports, value added and the number of firms seem to be located in about the same way in the three cities. The largest sectors in all three economies are wholesale trade and manufacturing; these two sectors combined account for around 70% of output in all cities. These sectors are most likely very different in many aspects and are probably affected to a different extent by the bridge. I will therefore examine these two sectors individually in addition to the aggregate economy. In terms of sectoral composition, Stockholm and Malmö are the most similar economies, while Gothenburg appears to be more manufacturing intensive. When possible in the analysis, however, I control for sectoral effects at the 4-digit level.

Table 1 provides descriptive statistics for four variables in each city: output, number of employees, productivity and exports. It also does so for the year of 1999. The table shows that, also from this aspect, it seems as if the cities are relatively comparable. It is seen that Stockholm and Gothenburg has more and, on average, larger firms than Malmö. The median is very similar for all variables except exports where Stockholm has a lower median. The fact that the mean is considerably larger than the median for all variables and in all cities is in line with the assumption that the distribution of firm productivity follows a Pareto distribution.[12] The frequency of firms is larger towards the lower levels of productivity but there is a thin tail with firms endowed with higher productivity. If the distribution of productivity of existing firms had been symmetric, such as in a non-truncated normal distribution, the mean would have been the same as the median which is clearly not the case here.

7 Results

The key assumption for examining whether any effect on productivity in Malmö builds on the reallocation of production, as in Melitz (2003), is that there has been a larger decrease in trade costs for the economy in Malmö, especially as regards trade with Denmark, than in the economies of Stockholm and Gothenburg. I therefore examine the changes in exports to Denmark first and then turn to the effects on aggregate productivity.

7.1 The effect on exports

The assumption in this experiment is that trade costs to a foreign country have been lowered more for firms based in Malmö (the treatment city) than for firms based in either of Gothenburg and Stockholm (the control cities). I examine this in some different ways according to the model described in Section 4. Less formally, however, I first use data and a survey from Öresundskonsortiet[13] that

[12] It would also be in line with a truncated normal distribution.

[13] The firm owning and managing the Öresund bridge. It is jointly owned by the Danish and Swedish governments.

Table 1

Descriptive statistics II

Variable	Stockholm	Gothenburg	Malmö
Observations	11 251	5 587	2 782
Output			
Mean	63 331	62 576	29 242
Standard deviation	848 660	1 177 315	128 591
Median	4 129	4 403	4 527
Employees			
Mean	22	19	14
Standard deviation	228	223	60
Median	3	3	3
Value added (Levinsohn Petrin)			
Mean	12 579	12 070	6 374
Standard deviation	201 633	207 363	28 160
Median	1 042	1 148	1 131
Export			
Mean	13 241	9 328	4 863
Standard deviation	90 016	36 334	10 463
Median	625	950	936

Note: The table describes properties of important variables in 1999, the last year before the construction of the bridge. Output, value added and exports are reported with a multiplier of 1000 Swedish kronor. Source: Statistics Sweden.

describes the changes in traffic across the Strait over time and of what characteristics their corporate customers are. Then, turning to the dataset described in Section 6, I, first, examine the changes in aggregate exports from the three cities. Second, I look at the per-firm increase in exporting and likelihood of being an exporter; while doing this I control for any sectoral changes that might confound the results. Finally, I look at how much of the changes in exporting levels come from the intensive versus extensive margin in the three cities.

Data from Öresundskonsortiet shows that there has been an increase in traffic across the Öresund Strait since the bridge was constructed. Figure 5 demonstrates the sharp increase since the construction of the bridge from a largely stable level. The light grey bars show the ferry traffic between Swedish Helsingborg and Danish Helsingör, two smaller towns slightly north of Malmö and Copenhagen. The black bars show ferry traffic between Malmö and Copenhagen before the bridge and dark grey shows the new traffic across the bridge. It is evident from the figure that traffic between Malmö and Copenhagen has increased rapidly. However, from this we do not know the purpose of the traffic. But Figure 6 shows that the proportion of traffic by car which travel in "business", including transportation of goods, is relatively stable. This indicates that the total number of vehicles crossing the bridge on business purposes has increased. Figure 7 shows the number of trucks crossing the bridge. Since I do not know the destination of the goods being transported, it can of course be argued that the goods are destined further south in Europe or to large ports. But it is at least an indication that the bridge has had an effect on trade patterns.

There has also been more detailed studies done by Öresundskonsortiet into the characteristics of their corporate customers. The report Öresundsbrokonsortiet (2006) describes how the patterns of trade change after the construction of the bridge after interviewing 2000 randomly selected corporate customers (1000 customers each from Sweden and Denmark, although the results are sometimes pooled) in 2006. First, they find that, among Swedish firms, only 29% of firms using the bridge in 2006 had any activity (of what kind is not specified) on the

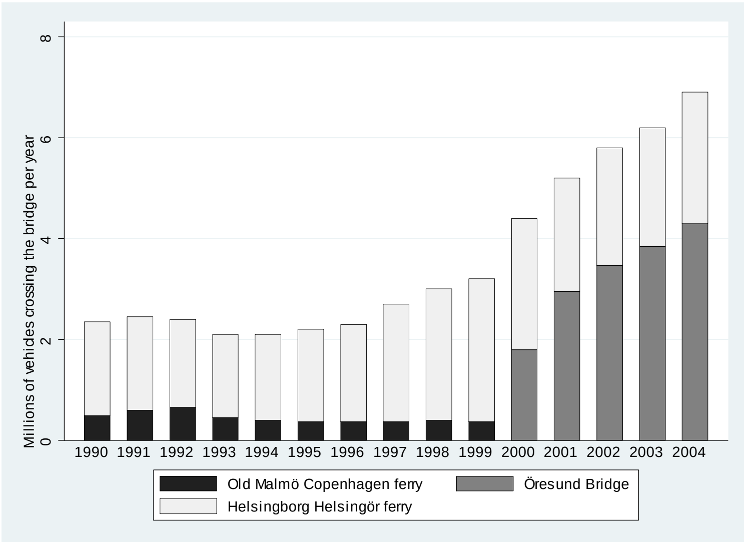

Figure 5: Millions of vehicles crossing the Öresund Strait per year. Source: Öresundsbrokonsortiet.

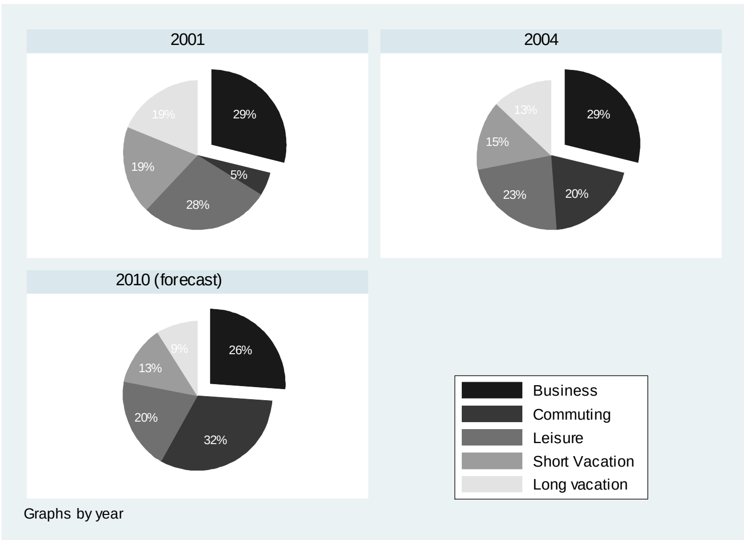

Figure 6: Car traffic across the bridge broken down by purpose. Source: Bridge Authorities.

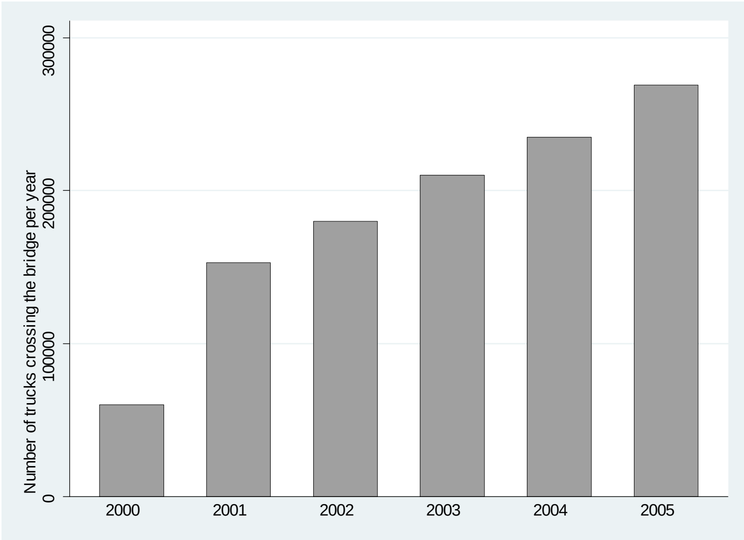

Figure 7: Number of trucks crossing the bridge per year. Source: Öresunds-brokonsortiet.

other side of the Strait before the decision to build the bridge was taken. Of the remaining 71%, 26% initiated their activities after the decision was taken to build the bridge but before it was completed (1991 - 1999) and 45% started after the bridge was constructed (2000 and afterwards).[14] In the pooled sample of firms that started their activities on the other side after 2000, smaller firms are more common. 53% of firms employing 0-5 people initiated their activities on the other side of the Strait after 2000, 44% of firms employing 6-10 people and only 19% of the largest firms employing more than 1000 people.[15] This means that, first, most Swedish firms that today use the bridge did not have any activity on the other side of the Strait prior to the bridge's construction. Second, the report tells us that smaller firms were affected more by the bridge than larger firms. The group of large firms seemed to have initiated their activities abroad already before the bridge was constructed whereas small firms to a greater extent chose to do so only after the bridge was constructed. This is what would happen, on average, across sectors according to the model. If

[14] Öresundsbrokonsortiet (2006), p. 15.
[15] Ibid, p. 13.

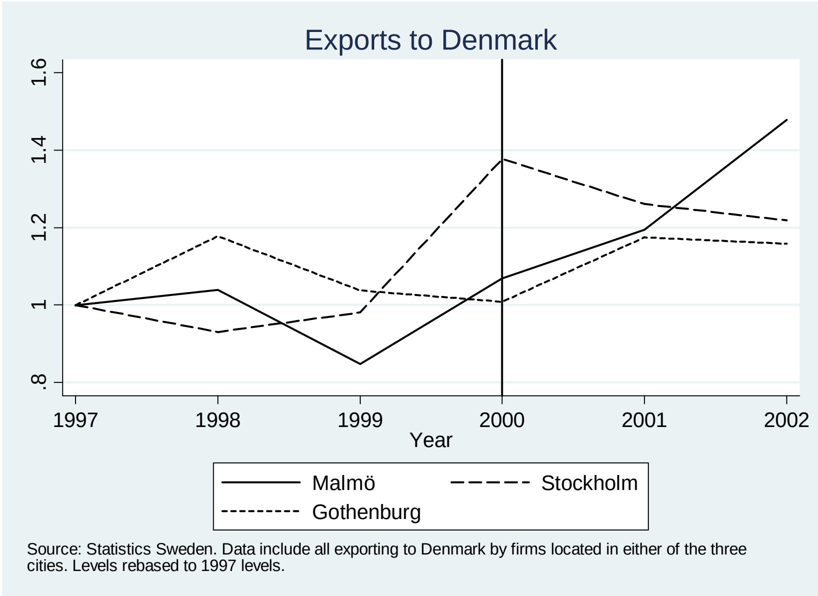

Figure 8: Exports to Denmark.

there was some sector-specific productivity cutoff prior to the bridge and that this was lowered by the bridge, then the larger firms would be less affected than smaller or medium-sized firms.

Having discussed these less formal indications of an effect from the bridge, I now turn to the formal analysis using the dataset from Statistics Sweden.

Here, I can examine the effect on aggregate exports by summing the export of each firm to get a measure of V_T. The prediction from the model, as shown in (23), is that total exports from Malmö will increase more than from Stockholm and Gothenburg. First, I plot graphically the sum of total exports to Denmark per city during the sample period in Figures 8 and 9. It can be seen that the aggregate series is, first, more volatile than one might expect and, second, that both Stockholm and Malmö have large increases starting in 1999. After 2000, however, Malmö's growth continues and stays at a high level while Stockholm's level drops a little bit. Proportionally, however, if I compare the years 1998-1999 with 2001-2002, i.e. the four years around the construction of the bridge, exports from Malmö increase by around 42% which is more than

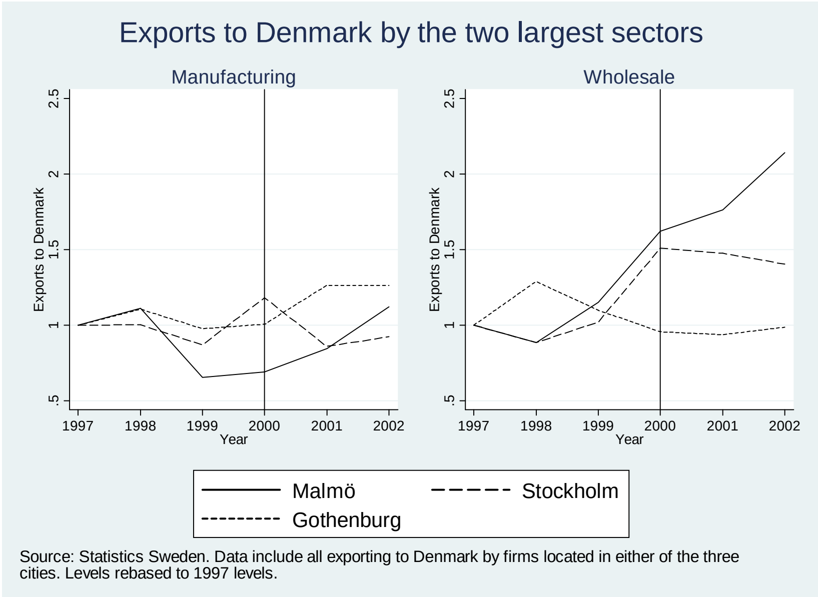

Figure 9: Exports to Denmark (two largest sectors).

Stockholm's 30% and Gothenburg's 5% (see Table 2). A comparison of the two largest sectors, manufacturing and wholesale trade, reveal large differences in how these two sectors are affected. While it is difficult to detect any effect in manufacturing, exports to Denmark by Malmö-based firms in wholesale trade almost doubles (increases by 92%). Stockholm and Malmö move closely together until 2000 when the bridge is constructed, but then Malmö continues the rapid increase while Stockholm levels stay at the level in 2000. Summing up, it seems as if aggregate exports to Denmark in Malmö respond to the construction of the bridge, and most clearly so in the wholesale sector. It will be seen in analysis of productivity whether this difference between the two sectors translate in different productivity effects as well.

In order to make a clean comparison, I also do a difference-in-difference-estimation on how existing exporters change their export levels where I can control for sectoral effects (that are the result of, for example, changes in international demand or technology and are uniform for all firms in Sweden). It was seen in (15) in Section 4 that foreign sales of an existing exporter, y_X, always

Table 2
Contributions, by margin, to the change in exports (1998/9 versus 2001/2)

	(a) All firms		
	Malmö	Stockholm	Gothenburg
Extensive margin	*35%*	*16%*	*17%*
Intensive margin	*6%*	*13%*	*-12%*
Total change	42%	30%	5%

	(b) Manufacturing			*(c) Wholesale*		
	Malmö	Stockholm	Gothenburg	Malmö	Stockholm	Gothenburg
Extensive margin	*22%*	*30%*	*20%*	*56%*	*13%*	*15%*
Intensive margin	*-11%*	*-35%*	*1%*	*36%*	*38%*	*-34%*
Total change	11%	-5%	21%	92%	51%	-19%

Note: The table compares how the different categories contribute to the changes in exporting levels between the two years following the construction of the bridge (2001 and 2002) with the two years preceding the bridge (1998 and 1999).
Source: Statistics Sweden.

increases when trade freeness, ϕ, increases. If all exporters at a given point in time are included, it would not necessary hold that exports in the exposed region increases since new and less productive firms ("switchers") would enter the export market with lower export levels than existing exporters.

I run a panel data regression to test the hypothesis in (15) at the level of the firm with sector-specific controls, only including firms that were exporting before the bridge was constructed. This is in essence a difference-in-difference estimator.

$$\text{export}_{it} = \beta_0 + \beta_M M_{it} + \beta_T T_{it} + \beta_{MT} M_{it} T_{it} + \Gamma_{it} + \varepsilon_{it} \tag{24}$$

where export_{it} denotes exports in logarithms of firm i in year t, M_{it} is a dummy that takes the value 1 if the firm is based in Malmö and 0 otherwise, T_{it} is a time treatment dummy that takes the value 1 if $t \geq 2001$, that is if the bridge has been constructed, t_{it} is a linear time trend and Γ_{it} is a vector of industry fixed effects at the 4 digit level. More specifically, Γ_{it} takes into account sectoral *changes*. It is constructed in the following way:

$$\Gamma = [1 - T \qquad T]\, D$$

where T is a vector which is one for each observation where $t \geq 2001$ and zero otherwise. D is a matrix with fixed sector effects. Adding Γ to the regression controls for time specific sectoral means. That is, it adjusts for the fact that these sectors may change export patterns in all cities across the period. A fixed effect for the same sector with the same value for all years would not capture changes in a sector that occur after the bridge is built. To control for such common changes in time, I use instead one fixed effect per sector before the bridge was constructed and another fixed effect for the same sector after the bridge was built. My set of dummies is therefore twice as large as the set of sectors. This allows the intercept (or average levels) to change in each sector in order to control for common shocks due to, for example, international demand. The error term ε_{it} is robust to heteroskedasticity. The coefficient β_{MT} is therefore a measure of how much a firm in Malmö, on average, increases its exports from 1996-2000 to 2001-2002 in comparison to firms in Stockholm and Gothenburg.

Table 3 reports the result of the regression in (24), first for the aggregate economy and then separately for manufacturing and wholesale trade. The results indicate that Malmö-based exporters had a large increase in exports compared to firms in Stockholm and Gothenburg both with and without adjusting for sector-specific changes at the 4 digit level. This holds for the aggregate economy as well as manufacturing and wholesale. The effect is strongest in manufacturing which could seem like a paradox given how little manufacturing changed overall. This could be explained, however, by the fact that the increase in wholesale trade mainly originated in new firms starting to the exporting, the extensive margin instead of the intensive margin. I turn to this question now.

A decomposition of the changes in exporting into the intensive and extensive

Table 3
Changes in exporting

	1a	1b	2a	2b	3a	3b
	All sectors		*Manufacturing*		*Wholesale*	
Malmö	0.3007***	0.2705**	0.2301	0.1154	0.4644***	0.4225***
	[0.0963]	[0.1055]	[0.1474]	[0.1653]	[0.1378]	[0.1438]
Treat	0.0171		-0.0986		0.0972	
	[0.0780]		[0.1404]		[0.1014]	
Malmö * Treat	0.2879*	0.4495***	0.3945*	0.7240***	0.3494	0.4363**
	[0.1568]	[0.1679]	[0.2394]	[0.2733]	[0.2226]	[0.2221]
Fixed industry effects (changes)	NO	YES	NO	YES	NO	YES
Observations	5637	5637	1950	1950	3109	3109

Note: The export level of firms is regressed on a dummy for whether the firm is based in Malmö, a treatment dummy if the year is after the bridge was built and an interaction term of these two. The interaction term shows the difference in difference estimate of how much export changed per firm in Malmö after compared to before the bridge and compared to the two other cities. Robust standard errors in brackets. Industry fixed effects at the four-digit level control for changes in export levels in sectors.
* significant at 10%; ** significant at 5%; *** significant at 1%
Source: Statistics Sweden.

margins is reported in Table 2. Section 4 describes that exports increase due to (i) an increase in the exporting level of existing exporters ($\frac{dy_X}{d\phi}$ as discussed) and (ii) a selection of new firms, "switchers", into the export market. This is due to the change in the productivity cut-off for exporting, a_X, which increases as shown in (21) in Section 4. Table 2 shows that Malmö's increase mostly originates from changes at the extensive margin. It has a much larger increase due to switchers (new exporters). When looking at the two largest sectors, manufacturing does not show any of this pattern whereas it is strong in wholesale (where two thirds of the increase is due to changes at the extensive margin).

The importance of the extensive margin is also tested by running a linear probability regression similar to equation (24) above, but changing the dependant variable to a dummy for whether the firm i is exporting in year t. The results are reported in Table 4.

There is evidence at the aggregate level that the likelihood of exporting has increased in Malmö more than in the control cities, even if the coefficient is rather small. In wholesale, the coefficients are, as expected from Table 2, larger than in both the aggregate economy and in manufacturing, although the

Table 4
Changes in the share of firms that export

	1a	*1b*	*2a*	*2b*	*3a*	*3b*
	All sectors		*Manufacturing*		*Wholesale*	
Malmö	0.0122***	0.0088***	0.0281***	0.0180***	0.0116***	0.0103***
	[0.0013]	[0.0013]	[0.0040]	[0.0038]	[0.0030]	[0.0030]
Treat	0.0080***		0.0141***		0.0159***	
	[0.0009]		[0.0026]		[0.0022]	
Malmö * Treat	0.0061**	0.0046*	0.0064	0.0075	0.0106*	0.0092
	[0.0028]	[0.0026]	[0.0079]	[0.0075]	[0.0064]	[0.0063]
Fixed industry effects (changes)	NO	YES	NO	YES	NO	YES
Observations	221196	221196	43865	43865	67873	67873

Note: The export level of firms is regressed on a dummy for whether the firm is based in Malmö, a treatment dummy if the year is after the bridge was built and an interaction term of these two. The interaction term shows the difference in difference estimate of how much export changed per firm in Malmö after compared to before the bridge and compared to the two other cities. Robust standard errors in brackets. Industry fixed effects at the four-digit level control for changes in export levels in sectors.
* significant at 10%; ** significant at 5%; *** significant at 1%
Source: Statistics Sweden.

standard errors become larger due to smaller sample size and the coefficient is therefore insignificant when sector-specific effects are accounted for.

Taken together, the results show strong evidence that there has been a reduction in trade costs for exporters based in Malmö which is larger than that for exporters in Stockholm and Gothenburg. The effect seems to be stronger for the wholesale sector than manufacturing and is mainly driven by the extensive margin.

7.2 Changes in aggregate productivity

Having established that there has been a reduction in trade costs for firms based in Malmö but not to the same extent for firms in Stockholm and Göteborg, I turn to the see how this has affected the main variable of interest: aggregate productivity. First, however, I want to describe the method used to aggregate productivity across firms to the level of the city. I follow the methodology described in Olley and Pakes (1996).

The aggregate productivity in each economy is calculated as the average productivity weighted by output in the following way:

$$p_t = \Sigma_i s_{it} p_{it} \tag{25}$$

where $s_{it} \equiv \frac{y_{it}}{y_t}$ and p_{it} are the market share and productivity, respectively, of firm i in year t. I want to see how much that is due to average productivity levels across firms and how much is due to the efficiency in the allocation of production, namely how much more of production is performed by the more productive firms. This can be seen by decomposing the measure in the following way:

$$\begin{aligned} p_t &= \Sigma_{i=1}^{N_t} s_{it} p_{it} \\ &= \Sigma_{i=1}^{N_t} \left(\Delta s_{it} + \overline{s}_t\right) \left(\Delta p_{it} + \overline{p}_t\right) \\ &= \Sigma_{i=1}^{N_t} \left(\Delta s_{it} \Delta p_{it} + \Delta s_{it} \overline{p}_t + \overline{s}_t \Delta p_{it} + \overline{s}_t \overline{p}_t\right) \\ &= \Sigma_{i=1}^{N_t} \overline{s}_t \overline{p}_t + \Sigma_{i=1}^{N_t} \Delta s_{it} \Delta p_{it} \\ &= \overline{p}_t + \Sigma_{i=1}^{N_t} \Delta s_{it} \Delta p_{it} \end{aligned} \tag{26}$$

where

$$\Delta s_{it} = s_{it} - \overline{s}_t$$

$$\Delta p_{it} = p_{it} - \overline{p}_t$$

where $\overline{s}_t$ and $\overline{p}_t$ denote unweighted means of market share and productivity, respectively. The first term in the expression above, $\overline{p}_t$, represents how aggregate productivity is explained by the average productivity across firms. The second term, $\Sigma_{i=1}^{N_t} \Delta s_{it} \Delta p_{it}$, is the "allocation efficiency" of production.

The model predicts that if trade freeness, ϕ, increases, then aggregate productivity, Ψ, increases.[16] First, I look at the aggregate level in the three cities. Table 5 contains information about how aggregate productivity changes between

[16] This is formally shown in the Appendix.

the two years before 2000 and the two years after. The first column confirms that Malmö experiences the sharpest rise in aggregate productivity (10% versus small or no changes). The wholesale sector in Malmö seems to be strongly affected and its aggregate productivity increases by as much as 47%. This should be compared with the effects on trade, where wholesale was found to have a much stronger effect than manufacturing and the economy as a whole.

To see whether this is due to an average increase (a change in $\overline{p}_t$) or a change in the efficiency of allocation (a change in $\Sigma_{i=1}^{N_t} \Delta s_{it} \Delta p_{it}$), the rate of increase in both of these terms is reported in the second and third columns. It is clear that the change in average unweighted productivity is small and there does not seem to be large differences across the cities. However, allocation efficiency increases more in Malmö than in the control cities (where it actually decreases slightly). That production is shifted towards more productive firms is the main cause for the rapid change in productivity growth in Malmö. Again, this mechanism appears most strongly in wholesale which had the largest trade effect.

The change in aggregate productivity is in line with the prediction by Melitz (2003) if Malmö has been exposed to a decrease in trade cost: that although productivity is relatively constant within firms, the reallocation of production raises aggregate productivity. And as the model predicts, almost all of the increase comes from a reallocation of production across firms (shown in the third column in Table 5).

I have now analysed changes at the aggregate level but I also want to look at a decomposition of aggregate productivity which lies closer to the model. While the result above is a general outcome of the model, the specific mechanism relates to the partitioning of firms into exporters and non-exporters and the fact that some production is reallocated towards more productive exporters. This can be examined in more detail by decomposing the change in aggregate productivity into specific groups based on exporting status.

Table 5

Changes in aggregate productivity

	Total change %	Mean productivity %	Allocation efficiency %
All firms			
Malmö	*10%*	*1%*	*9%*
Stockholm	*0%*	*1%*	*-2%*
Gothenburg	*-2%*	*2%*	*-4%*
Manufacturing			
Malmö	*6%*	*0%*	*5%*
Stockholm	*-30%*	*1%*	*-31%*
Gothenburg	*-8%*	*1%*	*-9%*
Wholesale			
Malmö	*47%*	*5%*	*42%*
Stockholm	*27%*	*5%*	*23%*
Gothenburg	*20%*	*4%*	*16%*

Note: The table shows the difference in productivity between 1998 and 1999 versus 2001 and 2002 (before and after the construction of the bridge). The change in aggregate productivity is calculated as described in the text and seperated into how much of the change that is attributed to a change in the unweighted average productivity and the change in the "allocational efficiency" of production.

Source: Statistics Sweden.

Grouping of firms

Name	*Explanation*
Switcher	Firms that did not export before 2000 but do so after.
Export_before	Firms that exported before 2000.
Entrants	Firms that did not exist before 2000.
Exiters	Firms that cease to exist during some time during or after 2000.
Neither	Firms that belong to neither of the above categories.

Each of these categories of groups can affect aggregate productivity in two ways. First, a group's market share can change and its effect on aggregate productivity will then be determined by the change of the market share multiplied by the relative productivity of the group versus the aggregate productivity of the region as a whole. This is denoted as "between-group changes".[17] Second, the aggregate productivity within the group can change and its effect on aggregate productivity in the region will then be determined by the group's market share multiplied by the change in its productivity. This I call a "within-group" change.[18]

As shown in Section 4, the model predicts that aggregate productivity increases due to three "between-group effects". First, exporters expand their share of production and since these firms are relatively more productive than non-exporters, this increases aggregate productivity. Second, some non-exporters exit and since these are the least productive firms this has a positive effect on aggregate productivity. Third, surviving non-exporters who do not select into exporting reduce their share of production and since they are less productive

[17] This is related to "allocational efficiency" but describes instead whether the change in a group's market share generates an increase or decrease in aggregate productivity. This is determined by how high the group's productivity is compared to aggregate productivity.

[18] This decomposition is different from the previous one which is due to Olley and Pakes (1996). The previous one described how much aggregate productivity changes due to a change in average productivity versus how productivity and production is correlated, respectively. The current one describes how different groups of firms contribute to the change in aggregate productivity. The productivity of every group, however, is still calculated as in Olley and Pakes (1996) above.

than exporters this also raises aggregate productivity. The model assumes constant productivity within firms and that all reallocation of production across firms is due to exporting status. Therefore, the prediction is that I should not see any "within-group" effects since the decomposition separates firms based on exporting status. Regarding the "between-group" effect of switchers, the model does not give clear predictions. If this group's initial productivity is above (below) average, then its expansion would generate an increase (decrease) in aggregate productivity. The uncertainty is because switchers are at an intermediate range in the productivity distribution of firms.

Formally, I decompose the change in productivity as follows, using i to denote a group of firms and p_{it} the aggregate productivity in that group in year t:

$$
\begin{aligned}
\frac{p_t - p_{t-1}}{p_{t-1}} &= \frac{1}{p_{t-1}} \left(\Sigma_{i=1}^{N_t} s_{it} p_{it} - \Sigma_{i=1}^{N_t} s_{it-1} p_{it-1} \right) \\
&= \frac{1}{p_{t-1}} \left(\Sigma_{i=1}^{N_t} s_{it} p_{it} - \Sigma_{i=1}^{N_t} s_{it-1} p_{it-1} + \Sigma_{i=1}^{N_t} s_{it-1} p_{it} - \Sigma_{i=1}^{N_t} s_{it-1} p_{it} \right) \\
&= \frac{1}{p_{t-1}} \left(\Sigma_{i=1}^{N_t} \left(s_{it} - s_{it-1} \right) p_{it} + \Sigma_{i=1}^{N_t} s_{it-1} \left(p_{it} - p_{it-1} \right) \right) \\
&= \frac{1}{p_{t-1}} \left(\Sigma_{i=1}^{N_t} \left(s_{it} - s_{it-1} \right) \left(p_{it} - p_t \right) + \Sigma_{i=1}^{N_t} s_{it-1} \left(p_{it} - p_{it-1} \right) \right) \\
&= \frac{1}{p_{t-1}} \left(\Sigma_{i=1}^{N_t} \left(s_{it} - s_{it-1} \right) \left(\widetilde{p}_{it} \right) + \Sigma_{i=1}^{N_t} s_{it-1} \left(p_{it} - p_{it-1} \right) \right).
\end{aligned}
\tag{27}
$$

In the fourth equality I use the fact that $p_t \Sigma_{i=1}^{N_t} \left(s_{it} - s_{it-1} \right) = 0$ and I denote the difference in group i's productivity from that of the economy as a whole by $\widetilde{p}_{it} \equiv p_{it} - p_t$. Then the contribution of a specific group to the total change in aggregate productivity can be denoted as:

$$
\overbrace{\frac{\left(s_{it} - s_{it-1} \right) \left(\widetilde{p}_{it} \right)}{p_{t-1}}}^{\text{Between-group change}} + \overbrace{\frac{s_{it-1} \left(p_{it} - p_{it-1} \right)}{p_{t-1}}}^{\text{Within-group change}}. \tag{28}
$$

Each group listed in the table above therefore affects aggregate productivity in two ways. First, the change in allocation of production that is due to

Table 6a

Contributions, by firm category, to changes in aggregate productivity (1998/9 versus 2001/2)

	Malmö		Stockholm		Gothenburg	
	Between	Within	Between	Within	Between	Within
All firms						
Switchers	*2%*	*1%*	*-3%*	*1%*	*1%*	*0%*
Exporter_before	*0%*	*2%*	*-2%*	*1%*	*0%*	*0%*
Entrants	*-1%*		*0%*		*0%*	
Exiters	*5%*		*4%*		*0%*	
Neither	*0%*	*1%*	*0%*	*1%*	*0%*	*-2%*
Total	10%		0%		-2%	

Note: The table compares how the different categories of firms contribute to the change in aggregate productivity between the two years following the construction of the bridge (2001 and 2002) and the two years preceding the bridge (1998 and 1999).
Source: Statistics Sweden.

the change in market shares between groups with different productivity levels. Second, the productivity also moves within groups.[19]

The contribution of the groups at the aggregate level is reported in Table 6a. The decomposition follows equation (28) above. Both the contribution through between-group and within-group changes are reported for each city and category of firms.

Table 6a shows that what drives the results in Malmö is mainly what is described in Melitz (2003): the exit of the least efficient firms (the between-effect of exiters) and the expansion of firms with above average productivity as they enter the export market (the between-effect of switchers). There is, however, no sign of a productivity increase originating in the expansion of existing exporters, instead there is an increase in the allocation efficiency among these but inside the group. The strongest effect, however, is from firms with low productivity

[19] I do not have estimates for p_t for exiters after the 2000 and not for new entrants before 2000 (these firms did not exist at these times). I therefore assume that the productivity of the groups in the period for which there is no estimate would have been the same as in the period for which I have estimates, i.e. I assume constant productivity within these groups.

Table 6b

Contributions, by firm category,
to changes in aggregate productivity (1998/9 versus 2001/2)

	Malmö		Stockholm		Gothenburg	
	Between	Within	Between	Within	Between	Within
Manufacturing						
Switchers	*-1%*	*1%*	*-2%*	*1%*	*6%*	*0%*
Exporter_before	*-1%*	*7%*	*-2%*	*-25%*	*1%*	*-8%*
Entrants	*0%*		*0%*		*0%*	
Exiters	*1%*		*1%*		*0%*	
Neither	*0%*	*0%*	*0%*	*0%*	*0%*	*-6%*
Total	6%		-30%		-8%	
Wholesale						
Switchers	*11%*	*4%*	*-7%*	*-1%*	*-2%*	*-3%*
Exporter_before	*1%*	*8%*	*0%*	*33%*	*0%*	*23%*
Entrants	*0%*		*0%*		*1%*	
Exiters	*15%*		*5%*		*2%*	
Neither	*1%*	*6%*	*0%*	*1%*	*0%*	*2%*
Total	47%		27%		20%	

Note: The table compares how the different categories of firms contribute to the change in aggregate productivity between the two years following the construction of the bridge (2001 and 2002) and the two years preceding the bridge (1998 and 1999).
Source: Statistics Sweden.

closing down.

Given the differences between sectors noted above, it is instructive to see if the results differ according to the extent that the export pattern was affected for these sectors. These results are reported in Table 6b for the manufacturing and wholesale sectors. While the manufacturing sector does not show any of the model's mechanism, the wholesale sector does so in the same way as the aggregate economy but with larger numbers. A large share of the increase in aggregate productivity in wholesale stems from the exit of the least productive firms and expansion by firms that enter into exporting. The two strongest effects are exactly these two mechanisms described by Melitz (2003). Existing exporters, however, do not expand as a group, but they experience an increase in productivity due to reallocation within the group as was observed in the aggregate economy. There is to some extent also an effect from the reallocation of production among nonexporters.

This analysis on aggregate productivity has so far not controlled for sectoral effects. In the analysis in Tables 6a and 6b, I cannot do this. However, the process of firms closing down or exiting, which was seen to affect productivity, can be analysed when controlling for sectors. The logarithm of (11) in Section 4 gives

$$a_D^{-k} = (\beta - 1)^{-1} \frac{f_D}{f_E} \left(1 + \phi^{\beta}\right) \tag{29}$$

$$k \log\left(\frac{1}{a_D}\right) = c + \log\left(1 + \phi^{\beta}\right) \tag{30}$$

where $c \equiv \log\left((\beta - 1)^{-1} \frac{f_D}{f_E}\right)$ and $\frac{1}{a_D}$ is the productivity (the inverse of marginal cost) cut-off for survival. It is now seen that if $\phi^{''} > \phi^{'} > \phi$, then also $\frac{1}{a_D^{''}} > \frac{1}{a_D^{'}} > \frac{1}{a_D}$, i.e. the productivity cut-off moves more in an economy in which the trade cost moves more.

However, to test if this happens for Malmö compared to Stockholm and Gothenburg, I have to assume that the initial trade cost is the same for all of the three economies. This is due to the fact that ϕ enters nonlinearly in (30) and the change therefore not only depends on the change in ϕ but also on its initial level. If this is assumed I can test the hypothesis that the reduction in trade costs has caused the least productive firms to exit also controlling for sectoral effects. I do this by comparing the means in productivity of firms exiting after the year 2000 with the surviving firms in the years before the bridge, 1998 and 1999. From (30), it is seen that if ϕ moves to $\phi^{''}$ in Malmö but in Stockholm and Gothenburg it only moves to $\phi^{'}$, then mean productivity of exiting firms in Malmö should be higher than those in Stockholm and Gothenburg since $\frac{1}{a_D^{''}} > \frac{1}{a_D^{'}}$ and productivity follows the same distribution in all cities.

I test this by running the following regression:

$$p_{it} = \beta_0 + \beta_X EXIT_{it} + \beta_M M_{it} + \beta_{XM} EXIT_{it} M_{it} + \Gamma_{it} + \varepsilon_{it} \qquad (31)$$

for the years 1998 and 1999. The variable $EXIT_{it}$ takes the value 1 if the firm is exiting after the bridge is constructed. M_{it} takes the value 1 if a firm is located in Malmö and Γ_{it} is a sector fixed effect. The interesting coefficient is therefore β_{XM} which is multiplied with the interaction term $EXIT_{it}M_{it}$. This compares the difference in productivity among exiters relative to surviving firms in Malmö with that in Stockholm and Gothenburg. This is a test of (30) saying that the reduction in trade costs has increased the productivity cut-off in Malmö more than in Stockholm and Gothenburg.

Table 7 shows that it seems that even with sectoral controls at a very disaggregated level, there is evidence that exiters in Malmö are more productive on average than in Stockholm and Gothenburg. The results hold in the aggregate economy both with and without sector fixed effects. For manufacturing the coefficients are close to zero but there are positive coefficients in the wholesale sector. These are, possibly due to the smaller sample size, not significant, however. The aggregate results support the notion that the effect of a decrease in trade costs forces out the least productive firms.

Table 7

Productivity of exiters

	Productivity					
	All		Manufacturing		Wholesale	
	(1)	(2)	(3)	(4)	(5)	(6)
Malmö	-0.037***	-0.035***	-0.007	-0.034**	-0.045***	-0.045***
	[0.008]	[0.008]	[0.016]	[0.016]	[0.014]	[0.015]
Exiter	-0.039**	-0.042***	-0.034	-0.047	0.009	0.018
	[0.015]	[0.015]	[0.034]	[0.032]	[0.032]	[0.031]
Malmö * Exiter	0.068*	0.091**	0.004	-0.010	0.028	0.042
	[0.037]	[0.037]	[0.079]	[0.080]	[0.070]	[0.070]
Fixed industry effects	NO	YES	NO	YES	NO	YES
Observations	77006	77006	16266	16266	23783	23783

Note: The regressions examine if the difference in productivity between firms that exit and firms that maintain production is different in Malmö than in the control cities. Firms that exit after 2000 are labelled as exiters. The regressions therefore only include the years before the bridge (1998 and 1999). Robust standard errors in brackets. Industry fixed effects at the four-digit level.

* significant at 10%; ** significant at 5%; *** significant at 1%

Source: Statistics Sweden.

8 Conclusion

The paper examines how aggregate productivity changes in a developed economy after a decrease in trade costs exogenous to industry-specific productivity and growth levels. Due to the stability and geographic symmetry of Sweden's macroeconomic and industrial policies, the risk of confounding effects is small since I compare the evolution of Malmö, which faced the trade cost reduction, with the two other large cities in Sweden. Using firm-level data, I calculate unbiased estimates of total factor productivity growth and estimate how much of productivity growth is due to the reallocation of production across firms and how much is due to average productivity growth within firms.

I find substantial evidence supporting models of firm selection such as Melitz (2003). Exports by firms based in Malmö rise substantially, mostly due to firms entering into the export market that did not export before. Importantly, of the two main sectors manufacturing and wholesale, it is clearly the wholesale sector that expands exports the strongest. Aggregate productivity in Malmö also increases. I find that almost all of Malmö's rapid growth in aggregate productivity occurs because of a reallocation of production from less productive to more productive firms. When decomposing the productivity gain, I find that the main drivers of aggregate productivity is a) exit of the least productive firms and b) expansion in output of firms with high productivity that enter the export market. When comparing the two largest sectors, the productivity changes are the strongest in the wholesale sector which is also the sector for which trade increased the most.

The findings provide insight into the causal link from market integration and increased trade to aggregate productivity through firm dynamics. It largely confirms the firm-selection patterns described in models such as Melitz (2003).

References

Arnold, J. M. (2005): "Productivity Estimation at the Plant Level: A Practical Guide," *note*.

Aw, B. Y., X. Chen, and M. J. Roberts (2001): "Firm Level Evidence on Productivity Differentials and Turnover in Taiwanese Manufacturing," *Journal of Development Economics*, 66(1), 51–86.

Axtell, R. L. (2001): "Zipf Distribution of U.S. Firm Sizes," *Science*, 293.

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

——— (2005): "Firm Structure, Multinationals and Manufacturing Plant Deaths," *Review of Economics and Statistics*, 89(2), 193–204.

Clerides, S. K., S. Lach, and J. R. Tybout (1998): "Is Learning by Exporting Important? Microdynamic Evidence from Colombia, Mexico and Morocco," *Quarterly Journal of Economics*, 113(3), 903–947.

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

Foster, L., J. Haltiwanger, and C. Krizan (1998): "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," *NBER Working Papers*, 6803.

Grossman, G., and E. Helpman (1994): "Protection for Sale," *American Economic Review*, 84(4), 833–850.

Helpman, E., M. J. Melitz, and Y. Rubinstein (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 123.

Helpman, E., M. J. Melitz, and S. R. Yeaple (2004): "Export versus FDI with Heterogeneous Firms," *American Economic Review*, 94(1), 300–316.

Koujianou Goldberg, P., and G. Maggi (1999): "Protection for Sale: An Empirical Investigation," *American Economic Review*, 89(5), 1135–1155.

Krugman, P. R. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70(5), 950–959.

Levinsohn, J. A., and A. K. Petrin (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), 317–341.

Loecker, J. D., and J. Konings (2006): "Job reallocation and productivity growth in a post-socialist economy: Evidence from Slovenian manufacturing," *European Journal of Political Economy*, 22.

Marschak, J., and W. Andrews (1944): "Random Simultaneous Equations and the Theory of Production," *Econometrica*, 12(3), 143–205.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

Olley, G. S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263–1297.

Pavcnik, N. (2002): "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *Review of Economic Studies*, 69(1), 245–276.

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.

Schank, T., C. Schnabel, and J. Wagner (2007): "Do exporters really pay higher wages? First evidence from German linked employer-employee data," *Journal of International Economics*, 72, 52–74.

Trefler, D. (2004): "The Long and Short of the Canada-U.S. Free Trade Agreement," *American Economic Review*, 94, 870–895.

Tybout, J. R. (2003): "Plant- and Firm-level Evidence on 'New' Trade Theories," in *Handbook of International Trade*, ed. by E. K. Choi, and J. Harrigan. Basil-Blackwell, Oxford.

Yeaple, S. (2005): "A Simple Model of Firm Heterogeneity, International Trade, and Wages," *Journal of International Economics*, 65(1), 1–20.

Öresundsbrokonsortiet (2006): "Näringslivets Erfarenheter över Öresund," *Report*.

9 Appendix

9.1 Proofs

9.1.1 Proof that $\frac{d\Psi}{d\phi} > 0$.

$$\Psi = \left(\frac{1}{1 + \frac{f_D}{f_X}\phi^\beta} \int_0^{a_D} a^{1-\sigma} \frac{g(a)}{G(a_D)} da + \frac{\frac{f_D}{f_X}\phi^\beta}{1 + \frac{f_D}{f_X}\phi^\beta} \int_0^{a_X} a^{1-\sigma} \frac{g(a)}{G(a_D)} da \right)^{\frac{1}{\sigma-1}} \tag{32}$$

$$= k^{\frac{1}{\sigma-1}} \left(\frac{1}{a_D^k} \frac{1}{1 + \frac{f_D}{f_X}\phi^\beta} \right)^{\frac{1}{\sigma-1}} \left(\int_0^{a_D} a^{k-\sigma} da + \frac{f_D}{f_X}\phi^\beta \int_0^{a_X} a^{k-\sigma} da \right)^{\frac{1}{\sigma-1}} \tag{33}$$

$$= \left(\frac{k}{k-(\sigma-1)} \right)^{\frac{1}{\sigma-1}} \left(\frac{1}{a_D^k} \frac{1}{1 + \frac{f_D}{f_X}\phi^\beta} \right)^{\frac{1}{\sigma-1}} \left(a_D^{k-(\sigma-1)} + \frac{f_D}{f_X}\phi^\beta a_X^{k-(\sigma-1)} \right)^{\frac{1}{\sigma-1}} \tag{34}$$

$$= \left(\frac{k}{k-(\sigma-1)} \right)^{\frac{1}{\sigma-1}} \left((\beta-1) \frac{f_E}{f_D} \right)^{1-\frac{1}{\beta}-\beta} \left(1+\phi^\beta\right)^{\beta+\frac{1}{\beta}-1} \tag{35}$$

$$* \left(\frac{1 + \left(\frac{f_D}{f_X}\phi^\beta \right)^{1+\frac{k-(\sigma-1)}{k}}}{1 + \frac{f_D}{f_X}\phi^\beta} \right)^{\frac{1}{\sigma-1}} . \tag{36}$$

It must be that Ψ increases in ϕ since $\beta > 1$ in the third bracket and that $\frac{k-(\sigma-1)}{k} > 0$ since $k-(\sigma-1) > 0$. ■

9.1.2 Proof that $\frac{d(y_D+y_X)}{d\phi} > 0$.

$$y_D + y_X = a^{1-\sigma} \left((\beta-1) \frac{f_E f_D^{\beta-1}}{1+\phi^\beta} \right)^{\frac{1}{\beta}} (1+\phi) \tag{37}$$

$$= a^{1-\sigma} \gamma \frac{1+\phi}{\left(1+\phi^\beta\right)^{\frac{1}{\beta}}} \tag{38}$$

where $\gamma \equiv \left((\beta-1) f_E f_D^{\beta-1} \right)^{\frac{1}{\beta}}$. The derivative of $\frac{1+\phi}{(1+\phi^\beta)^{\frac{1}{\beta}}}$ with regard to ϕ is:

$$\frac{d}{d\phi} \left(\frac{1+\phi}{\left(1+\phi^\beta\right)^{\frac{1}{\beta}}} \right) = \left(\left(1+\phi^\beta\right)^{\frac{1-\beta}{\beta}} \right) \left(1-\phi^{\beta-1}\right) > 0 \tag{39}$$

which is always true since $\phi < 1$ at trade costs strictly larger than zero and $\beta > 1$. ■