# Essays on the effects of trade liberalization

by

Sabine D'Costa

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2010

Doctoral Committee:

Professor Katherine Terrell (Deceased), Chair
Professor Alan V. Deardorff
Professor John E. DiNardo
Assistant Professor Andrei A. Levchenko

To Emmanuel

## ACKNOWLEDGEMENTS

I thank my parents for giving me the best education possible, and for their unwavering support, love, and encouragement throughout my studies. I also owe this achievement to the support and encouragement I have received from family, friends of many years, as well as colleagues at the University of Michigan and the London School of Economics.

I am very grateful to my committee for their precious help in the completion of this dissertation, and particularly to my chair the late Professor Terrell for her enthusiasm and constant encouragement. She will remain in my memory as an exceptionally caring and dedicated advisor, and I hope to continue to be inspired by her in my future career and in particular in my relationship with my students. I am also indebted to faculty at the LSE for their help and suggestions on my work, and I thank the Centre for Economic Performance at the LSE for its hospitality during the latter years of my doctoral studies.

# TABLE OF CONTENTS

# LIST OF FIGURES

**Figure**

# LIST OF TABLES

**Table**

# LIST OF APPENDICES

**Appendix**

# CHAPTER I

# Introduction

Most of the early literature on the impact of trade liberalization has been concerned with the impact on aggregate welfare indicators. More recently the literature has focused on the impact on the productivity of export oriented firms or import competing firms. In this dissertation I analyze three additional ways in which trade liberalization can affect the economy of the liberalizing country. Chapter 1 explores the impact on input linkages at the firm level, using firm-level data. Chapter 2 considers the effects on the geographical concentration of industries using regional industry-level data. Finally chapter 3 looks at the impact on the size distribution of plants, by exporter status, using plant level data.

In chapter 1, I investigate how the availability of imported intermediates affects firm productivity through the channel of input linkages. Input linkages imply that firms are more productive if the region where they are located offers them proximity to input suppliers.

I first set up a simple model of international trade inspired by Krugman (1980) to show that as trade costs go down, productivity of the domestic composite good increases less in the number of domestic varieties. To my knowledge, the notion that input linkages are weakened as trade is liberalized has only been supported by simulation results in theoretical models of new economic geography trying to assess the impact of trade on regional concentration (Krugman and Elizondo (1996), Fujita, Krugman, and Venables (1999), chapter 18 and Crozet and Koenig (2004)).

I then provide empirical evidence combining firm-level data with trade data and measures of input linkages. I show that the positive impact of input linkages with local suppliers on firm productivity is reduced as the imports of similar inputs increase. This paper is related to Amiti and Konings (2007), who focus on the impact of the liberalization of trade in intermediates on firm-level productivity during Indonesia's trade liberalization. Their dataset allows them to identify the effect of trade liberalization on firms that import their inputs compared to other firms in the same sector. In contrast, my paper focuses on local input linkages and their interaction with the liberalization of trade in intermediates, and it also offers a theoretical model that motivates the empirical work.

The main contribution of this paper is that it combines trade liberalization with the notion of input linkages and is able to identify their combined effect on firm-level productivity. It investigates a new channel through which trade liberalization can affect firm productivity. The results may also lend empirical support to those economic geography models that rely on the weakening of input (and output) linkages by trade. Moreover, the empirical implementation avoids the endogeneity issues related to the cumulative causation process linking firm productivity and firm location. Finally, this paper proposes a theoretical foundation for such empirical investigations.

In chapter 2, I consider the fact that industries may geographically concentrate or spread. This is because trade re-orientation or liberalization impacts the sub-national location of industries by altering the balance of dispersion and concentration forces and by modifying the patterns of market and supplier access. This paper aims to answer the question empirically given that the results of the relevant theoretical models hinge on a set of initial conditions. I estimate the relation between trade and the spatial concentration of industries and the relation between trade and the share of the economic center (Budapest) in an industry. Regressions of the Krugman index of industrial concentration on industry exports and imports of inputs reveal that exports increase the geographical concentration of industries, and more so for industries that were initially dispersed. Furthermore, higher exports reduce the share of the capital Budapest in an industry, and more so for industries that initially had a low share of Budapest. Combining these results reveals that increased exports in an industry have brought about higher geographical concentration away from the capital region. This means that trade has increased the geographical concentration of industries without necessarily reinforcing existing industrial networks.

Most of the literature on industry localization (also called geographical concentration of industries) has either described the evolution of industry concentration over time, or estimated the determinants of industry concentration. The literature investigating the determinants of the geographical concentration of industries began with Kim (1995). The findings from this literature are that there is little evidence that labor, capital or resource intensity increases the geographical concentration of industries, and mixed evidence on the effect of demand and cost linkages as well as trade liberalization. There is positive evidence that technology intensive and high returns to scale industries are more concentrated. In this chapter I take into account the results of the literature by accounting for the characteristics of industries that have been proved to explain industry localization.

In contrast there are very few empirical papers estimating the sub-national effects of trade liberalization on the location of industries. Overman and Winters (2003) analyze the impact of the accession of the United Kingdom into the EU on the location of industries in the United Kingdom. They find evidence of relocation towards the South East, caused by market access to mainland Europe. My results shed light on the impact of liberalization on the geographical concentration levels of industries as well as on relocation away from the capital region.

In chapter 3, I investigate how trade liberalization affects the size distributions of exporters and non exporters. I rely on predictions that can be extracted from heterogeneous-firms trade models and confront them with actual distributions derived from Colombian data before and after policy changes. This is interesting from a policy perspective because it can explain whether a policy that favors exporters has generated an export boom where the largest and most productive incumbent exporters grow, or one where the bulk of the increase in exports comes from massive entry of smaller plants into the export market.

Overall most of the results in this chapter support the predictions from Melitz (2003) as well as the long-run predictions of Melitz and Ottaviano (2008). I identify a distinct leftward shift in the size distribution of exporters during a period of currency depreciation and unilateral tariff increases, which remains later while the currency is stable and tariffs are cut drastically. In contrast, the size distribution of non exporters does not shift but becomes more peaked with shorter tails in the first period, then becomes less peaked again during the second period. I then show that the bulk of the leftward shift in the size distribution of exporters is due to massive entry of smaller plants into exporting, which is consistent with both models.

So far the literature on the distribution of firm size has focused on establishing stylized facts about the size distribution of firms, irrespective of their exporter status. And the existing empirical literature examining the impact of international trade liberalization on firm size has used mostly parametric estimation methods, and focused on average effects or on effects on specific quantiles of the distribution, without comparing the impact on exporting vs. that on non-exporting firms or plants. To summarize the empirical evidence, there is ample evidence that trade liberalization increases exit, however there is mixed evidence regarding the impact on plant size. Most of the evidence shows that greater import competition reduces average plant size, while greater export possibilities tend to increase average plant size. Finally, Nocke and Yeaple (2006) focus directly on the impact of trade on the overall distribution of firm size, however they summarize the distribution of firm size using a measure of its dispersion. This does not explain leftward or rightward shifts in the distribution. Moreover their framework does not allow them to distinguish between the impact on exporters and non-exporters. On the other hand, this chapter, by using the Melitz (2003) and the

Melitz and Ottaviano (2008) models, is able to derive theoretical predictions on how the distribution of firm size would be affected by trade policy changes, for the sample of exporters and non-exporters.

This chapter therefore contributes to the literature on the distribution of plant size by establishing that the size distribution of exporters is lognormal whereas that of non exporters is more right skewed. It also contributes to the trade literature that estimates the effects of trade liberalization, by testing predictions derived from two heterogeneous-firms models regarding effects on the size distribution of plants by exporter status. The results also add to the wide literature on how exporting plants differ from non exporters.

# CHAPTER II

# Input linkages, productivity and trade liberalization

## 2.1 Introduction

Many researchers and policy makers are concerned with the impact of trade liberalization on the productivity of export oriented firms, or of import competing firms. Empirical studies of trade liberalization focus on relationships between trade policy in certain sectors and the performance of domestic firms in those same sectors.[1] However, much has still to be said on how domestic firm performance can also be affected by the availability of imported *intermediates*. Indeed, imported intermediates can come to replace locally produced ones, as input linkages such as those that prevail in regional industrial clusters are weakened. Consequences of this would include the decline of regional industrial clusters (also called local industrial networks) and the relocation of industries.

In this paper, I investigate how the availability of imported intermediates affects firm productivity through the channel of input linkages. Input linkages imply that firms are more productive if the region where they are located offers them proximity to input suppliers. The relevance of input linkages to firm productivity is straightforward in a closed economy, but it may weaken when foreign intermediates become available. This paper therefore tries to answer the following question: Does the availability of inputs in a firm's region have a sustained positive impact on firm productivity as imports are liberalized?

On the theoretical side, Ethier (1982) pointed out that when trade liberalization increases the number of varieties of intermediates available, firm productivity may increase. Furthermore, a main feature of the Dixit-Stiglitz model is the love for variety effect. Ciccone and Hall (1996) use this notion in an empirical investigation of the role of economic density on regional productivity in the absence of trade. They use a simple closed economy model of monopolistic competition with a two-

[1] See Tybout (2003) for a survey on plant-level evidence.

tier production function to show that increasing the number of intermediate varieties available in a location increases the productivity in the production of the local composite good. To my knowledge, the notion that input linkages are weakened as trade is liberalized has only been supported by simulation results in theoretical models of new economic geography trying to assess the impact of trade on regional concentration (Krugman and Elizondo (1996), Fujita, Krugman, and Venables (1999), chapter 18 and Crozet and Koenig (2004)).

On the empirical side, the economic geography literature has linked agglomeration economies and productivity. However these papers do not analyze input linkages as sources of agglomeration economies, and they abstract from international trade. They focus on establishing the role of market access, knowledge spillovers, and labor pooling in productivity (Lall, Salizi, and Deichmann (2004)), or try to discriminate between the contributions of specialization versus diversification of a region in its productivity (Henderson, Lee, and Lee (2001)). Amiti and Cameron (2007) have demonstrated the role of supplier access on firm wages in Indonesia, while Javorcik (2004) has identified positive FDI spillovers going from foreign owned clients towards domestic suppliers. Encouraging evidence of the positive contribution of a wider set of inputs to productivity is found in Feenstra, Markusen, and Zeile (1992). Using the case of the vertically integrated Korean *chaebols*, they show that TFP growth of final goods producers within the *chaebols* increases in the availability of new inputs coming from within-*chaebol* firms.

This paper is related to Amiti and Konings (2007), who focus on the impact of the liberalization of trade in intermediates on firm-level productivity during Indonesia's trade liberalization. They find that a 10% fall in input tariffs has a higher impact on productivity than a 10% fall in output tariffs. Their dataset allows them to identify the effect of trade liberalization on firms that import their inputs compared to other firms in the same sector. In contrast, my paper focuses on local input linkages and their interaction with the liberalization of trade in intermediates, and it also offers a theoretical model that motivates the empirical work.

I offer empirical evidence combining firm-level data with trade data and input linkages data. I use a panel of Hungarian manufacturing firms extracted from the Amadeus database. The data span eight years (1995-2002) during which Hungary liberalized its trade with the European Union. The Europe Agreement with Hungary was signed in December 1991, and the bulk of the liberalization of manufacturing imports from the EU into Hungary occurred between 1997 and 2000. There was considerable variation in the import flows during the period of this study. I show that the positive impact of input linkages with local suppliers on firm productivity is reduced as the imports of similar inputs increase.

The main innovation of this paper is that it combines trade liberalization with the notion of local input linkages and is able to identify their combined effect on firm-level productivity. It investigates a new channel through which trade liberalization can affect firm productivity. The results may also lend empirical support to those economic geography models that rely on the weakening of input (and output) linkages by trade. Moreover, the empirical implementation avoids the endogeneity issues related to the cumulative causation process linking firm productivity and firm location. Finally, this paper proposes a theoretical foundation for such empirical investigations.

In the next section I lay out a simple model of international trade inspired by Krugman (1980) to show that as trade costs go down, productivity of the domestic composite good increases less in the number of domestic varieties. Section 3 then briefly describes the data and develops the empirical strategy. Section 4 presents the results.

## 2.2 Theoretical background

I begin by using a model inspired by Krugman (1980)'s open economy model of monopolistic competition. There is a final good and an intermediate composite good sector. The composite good is made of intermediate varieties which are tradable. We focus on derivations concerning the home country $H$. Variables denoting the foreign country are identified by an asterisk and derivations are analogous to those for the home country.

### Final good

The final good is produced according to the following Cobb-Douglas production function:

$$G = l^{\lambda} y^{\gamma} \tag{2.1}$$

where $\lambda, \gamma < 1$, $l$ is labor and $y$ is the composite good.
Let $L$ denote the total amount of labor in $H$, and $w$ the domestic wage. The price of the final good is normalized to 1. Since the share of the final good paid to labor is $\lambda$,

$$wl = \lambda G.$$

Also, since all final output is paid to labor working in the final and intermediate sectors,

$$wL = G.$$

Therefore $l = \lambda L$, $\lambda$ is the fraction of $L$ working in the final good sector, and $1-\lambda$ is the fraction working in the composite good sector.

**Composite good**

Good $y$ is produced following a CES production function:

$$y = [\int_0^n c(t)^{1/\mu} dt + \int_0^{n^\star} c^F(t)^{1/\mu} dt]^\mu \tag{2.2}$$

where $c$ is the absorption of each domestic variety in $H$, $c^F$ the absorption of each foreign variety in $H$, $n$ and $n^\star$ are the number of intermediate varieties at home and in the foreign country, $\mu > 1$ and $\sigma = \frac{1}{1-\frac{1}{\mu}}$ is the elasticity of substitution.

**Production decision of each variety**

All intermediates are produced with the same technology at home and abroad. Let $x$ and $x^\star$ be the amount of each domestic and foreign variety produced. The cost function for domestic varieties is:

$$C(w, x) = \alpha w + \beta w x, \tag{2.3}$$

and that for varieties produced abroad is:

$$C(w^\star, x) = \alpha w^\star + \beta w^\star x,$$

with $\alpha, \beta > 0$.

So the price of each domestic variety is:

$$p = w\beta\mu, \tag{2.4}$$

and that of each foreign variety is $p^\star = w^\star \beta\mu$.

The zero-profit condition in the production of each variety yields:

$$x = x^\star = \frac{\alpha}{\beta(\mu - 1)}. \tag{2.5}$$

**Full employment of labor**

The number of units of labor required to produce $x$ is

$$\alpha + \beta x = \alpha \frac{\mu}{\mu - 1},$$

and the number of varieties produced in $H$ is therefore

$$n = \frac{\mu - 1}{\mu} \frac{(1 - \lambda)L}{\alpha}. \tag{2.6}$$

Similarly, $n^\star = \frac{\mu-1}{\mu}\frac{(1-\lambda)L^\star}{\alpha}$, and we have in equilibrium:

$$\frac{n}{n^\star} = \frac{L}{L^\star}.$$

**Trade liberalization**

Let the two countries engage in trade, with iceberg trade cost g being the fraction of the goods that reaches. In the home country, $\hat{p^\star} = p^\star/g$ is the price of imported intermediates. In the foreign country the price of imported intermediates is $\hat{p} = p/g$.

The home producers of y will use a combination of domestic and imported intermediates in the following proportions:

$$\frac{c^F}{c} = (\frac{\hat{p^\star}}{p})^{-\frac{1}{1-1/\mu}}$$

therefore

$$\frac{c^F}{c} = (g\omega)^{\frac{1}{1-1/\mu}}, \tag{2.7}$$

where $\omega$ is the relative wage $p/p^\star$.

Let $d$ and $d^F$ be the domestic demand for domestic varieties and for foreign varieties, respectively.

The total demand from home will follow:

$$\frac{d^F}{d} = \frac{c^F}{gc}$$

such that

$$\frac{d^F}{d} = \omega^{\frac{\mu}{\mu-1}} g^{\frac{1}{\mu-1}}. \tag{2.8}$$

**Productivity in autarky**

First I determine the impact of a greater number of varieties on productivity in the closed economy. In autarky, $c^F = 0$, $c = x$ and the production function of $y$ reduces to

$$y = [\int_0^n x(t)^{1/\mu} dt]^\mu$$

By substituting equation (2.5) into this function, the output of y becomes

$$y = xn^\mu \tag{2.9}$$

and the productivity of the production process of good y in autarky is output divided by input: $\frac{xn^\mu}{nx}$, that is:

$$prod^a = n^{\mu-1}. \tag{2.10}$$

This is the closed economy result highlighted in Ciccone and Hall (1996), that productivity is increasing in the number of varieties available.

**Productivity in the open economy**

In the open economy, in order to determine the amount of output produced, I substitute for $c^F$ in the production function for y.

$$y = [nc^{1/\mu} + n^\star c^{1/\mu} (g\omega)^{\frac{1/\mu}{1-1/\mu}}]^\mu$$

i.e.,

$$y = c[n + n^\star(g\omega)^{\frac{1}{\mu-1}}]^\mu. \tag{2.11}$$

By dividing output by $cn$, the amount of local input used, I find that the productivity of the local input is:

$$prod = \frac{[n + n^\star(g\omega)^{\frac{1}{\mu-1}}]^\mu}{n} \tag{2.12}$$

which is analogous to the expression found in the closed economy case.

I first want to know the impact of a marginal increase in the number of local varieties on productivity.

I obtain:

$$\frac{\partial prod}{\partial n} = \frac{[n + n^\star(g\omega)^{\frac{1}{\mu-1}}]^{\mu-1}}{n^2} \cdot Z_1, \tag{2.13}$$

where

$$Z_1 = n(\mu - 1) + n^\star(g\omega)^{\frac{1}{\mu-1}}[\frac{\mu n}{\mu - 1}\frac{J}{\omega} - 1].$$

$J$ is the second order effect: $J = \frac{\partial\omega}{\partial n}$.

Intuitively we expect the sign of this derivative to be positive: Productivity of the composite good increases in the number of varieties available locally. I evaluate the sign of expression (2.13) graphically, as a function of $\omega$ and $\mu$. Since the first term on the right hand side is positive, I plot $Z_1$, in order to evaluate its sign.

I choose reasonable parameter values by which to plot $Z_1$:

- I choose $n < n^\star$, since we are in a context of trade liberalization where a small country opens up to trade with a larger group of countries. For the graph I choose $n = 100$ and $n^\star = 1000$.
- Since a standard result of Krugman (1980) is that 'the larger country, other things equal, will have the higher wage', we know that $J = \frac{\partial\omega}{\partial n} > 0$.
- From the same result, if $L < L^\star$ then $\omega < 1$. Therefore I plot $Z_1$ for $\omega$ ranging from 0 to 1.
- The 'markup' $\mu$ is chosen to range from 1 to 2.
- I choose any $g$ in (0,1). Let $g$=0.7.

We see on Figure 2.3 that this expression is positive, for most of the reasonable parameter values. The restriction for this to hold is that $J = \frac{\partial \omega}{\partial n}$ is not too small (say $J > 0.005$). There are no restrictions on $g$ or $\mu$.

This paper aims at identifying whether this positive impact of the number of local varieties on firm productivity is maintained as trade is liberalized. In order to evaluate this, I look at the sign of the second derivative of productivity with respect to $n$ and $g$:

$$\frac{\partial^2 prod}{\partial L \partial g} = \delta \frac{\mu}{L} A^{\mu-1} Z_2 \tag{2.14}$$

where

$$\begin{aligned} Z_2 \quad = \quad & L^\star (g\omega)^{\frac{1}{\mu-1}} \left(\frac{1}{g} + \frac{K}{\omega}\right)[A^{-1}(1 + \frac{L^\star}{\mu-1}(g\omega)^{\frac{1}{\mu-1}}\frac{J}{\omega}) \\ & - \frac{1}{L(\mu-1)}] + \frac{L^\star}{(\mu-1)^2}(g\omega)^{\frac{1}{\mu-1}} J\left(\frac{1}{g\omega} + (2-\mu)\frac{K}{\omega^2}\right) \end{aligned} \tag{2.15}$$

and $K = \frac{\partial \omega}{\partial g}$.

Since we are looking at the productivity of domestic inputs, we intuitively expect this expression to have a negative sign: as trade is liberalized ($g$ increases) we expect the contribution of the number of domestic varieties to the productivity of domestic inputs to decrease. Again, I identify the sign of this partial effect by graphical simulation. Since $\delta \frac{\mu}{L} A^{\mu-1} > 0$, I plot $Z_2$ as a function of $\omega$ and $\mu$. I keep the values of the parameters common to the previous simulation unchanged. Following the same reasoning employed for $J$, we know that $K = \frac{\partial \omega}{\partial g} < 0$.

The second derivative is thus negative, as appears in Figure 2.4. However this expression is highly nonlinear and the sign is very sensitive to parameter values: $Z_2$ ceases being negative for all $0 < \omega < 1$ and $1 < \mu < 2$ for lower values of $g$ or higher (less negative) values of $K$.

To sum up, we can draw from this stylized model that, under certain reasonable parameter restrictions, productivity of the composite good is increasing in the number of local varieties available, but less so as trade is liberalized.

## 2.3 Empirical analysis

### 2.3.1 Empirical model

The stylized model yields non-linear results, yet if we restrict parameters to reasonable values we can draw clear predictions. I therefore take a reduced-form approach and test empirically how trade liberalization affects the relationship between input linkages and productivity.

I estimate a Cobb-Douglas production function, where firm output is a function of inputs (capital K, labor L, and materials M) and external factors that influence productivity (the presence of input firms and the level of imports of input industries):

$$y_{ijrt} = A.f(K, L, M), \tag{2.16}$$

$$A = g(PIL_{jrt}, II_{jt}) \tag{2.17}$$

where $y_{ijrt}$ is output of firm $i$ in industry $j$, region $r$ and year $t$, $PIL_{jrt}$ is a measure of potential input linkages[2] at the industry-region-year level, and $II_{jt}$ is a measure of the time $t$ volume of the imports from the EU[3] which are relevant to industry $j$.

The estimable equation is:

$$\begin{aligned} \ln y_{ijrt} &= \alpha_0 + \alpha_1 \ln K_{ijrt} + \alpha_2 \ln L_{ijrt} + \alpha_3 \ln M_{ijrt} + \beta \ln PIL_{jrt} \\ &\quad + \gamma \ln II_{jt} + \delta \ln PIL_{jrt} \times \ln II_{jt} + \eta_i + \epsilon_{ijrt}. \end{aligned} \tag{2.18}$$

What I call input linkages is the presence in a firm's region of firms that provide its inputs. So a firm in industry $j$ benefits from high input linkages if it is located in a region that has relatively many firms that provide inputs for industry $j$. The variable tracking trade liberalization is the volume of imports of those industries providing inputs to industry $j$, and is therefore industry and year specific. According to the stylized model, the higher the volume of imports of input industries, the lower the contribution of input linkages (the number of varieties in our stylized model) to firm productivity.

I construct a measure of input linkages based on standard practise in the literature,[4] and weight them by interregional distances to obtain potential input linkages:

$$PIL_{jrt} = \Sigma_{r'} \frac{1}{d_{rr'}} IL_{jr't},$$

$$IL_{jr't} = \Sigma_{k \neq j} I_{kj} \times \frac{l_{kr't}}{l_{kt}}.$$

[2] As will become clearer later, I speak of potential input linkages because I take into account the input linkages in neighboring regions.

[3] Imports from the EU are used as a measure of trade liberalization, because these account for the bulk of the variation in import flows.

[4] See Dumais, Ellison, and Glaeser (1997) for the input linkages measure.

$I_{kj}$ is industry $j$'s share of inputs coming from industry $k$ ($k \neq j$),[5] and $\frac{l_{kr't}}{l_{kt}}$ is the share of region $r'$ in industry $k$ employment in year $t$,

$d_{rr'}$ is the road distance in kilometers between the main city in region $r$ and that in region $r'$, with $d_{rr} = 1$.

We see that a firm's level of input linkages, $IL_{jrt}$, depends on the relative abundance in region $r$ (compared to the rest of the country) of those industries that represent a high share of industry $j$'s inputs.[6] $II_{jt} = \Sigma_{k \neq j} \frac{Imports_{kt}}{GDP_{kt}} \times I_{kj}$, so this measure is a weighted average of imports of other industries (normalized by industry size) weighted by their share in $j$'s total inputs.

Finally, $\eta_i$ is a firm level fixed effect (which includes the part of firm-level productivity which is time invariant), and $\epsilon_{ijrt}$ is the idiosyncratic error term.

According to the predictions of the stylized model, we expect to find the coefficient on the interaction term, $\delta$, negative and significant. I estimate equation (2.18) in two different ways, first by OLS regression, then using the Levinsohn-Petrin productivity estimation method in order to address the problems caused by the endogeneity of firms' input choices.

### 2.3.2 Data description

The unbalanced firm-level panel dataset for 1995-2002 was constructed using the Amadeus database (Top 1.5 million). The firms included in the database fulfill at least one of the following three criteria: operating revenue of at least 1 million euros, total assets of at least 2 million euros, and at least 15 employees.[7] These data are extracted from financial statements and cover 17,600 firms. This represents about 9% of Hungarian firms (only 40% of Hungarian firms are required to file accounts). After removing non-manufacturing firms, and firms with missing observations, the final dataset contains 6,500 observations covering 1,300 firms.

The information available includes operating revenue, assets, cost of employees and of materials, the industrial classification (according to the NACE classification, at the two-digit level) and the location of the firms within the twenty administrative counties in Hungary, which we call regions in this paper. I was therefore able to merge this firm-level dataset with regional characteristics and with trade flow data. Tables 2.1 and 2.2 give the geographic and industrial breakdown of the dataset.

[5] Intermediate inputs coming from the same industry are excluded from the measure for input linkages, because I wish to focus on externalities from the agglomeration of diverse industries (Jacobs externalities) rather than from the geographical concentration of a single industry (MAR externalities).

[6] Unlike in Dumais, Ellison, and Glaeser (1997), this measure of input linkages does not take values between 0 and 1. This is because the $I_{kj}$ do not sum to 1: They are the share of industry $k$ in industry $j$'s *total* inputs, including other non-manufacturing industries.

[7] So there are firms with very few employees as long as they have large turnover or assets.

There is considerable variation in output across regions and across industries. I further aggregate industries into eight major categories (see table 2.3), to match the classification of the data coming from the Hungarian statistical institute (see Appendix 1 for a description of the data).

The fact that the time series begins in 1995 whereas trade liberalization began in 1992 should in principle be a problem. However, as shown in figures 2.1 and 2.2, most of the growth in manufacturing imports occurred after 1997 (although there was a first smaller increase in 1994), and import flows by industry vary considerably during our time period.

Table 2.4 presents summary statistics for potential input linkages and import of inputs. There is considerable variation in these measures, at least across groups.

### 2.3.3 Econometric difficulties

The most fundamental issue is that the potential input linkages are endogenous. They are part of a cumulative causation process in which industry concentration and regional agglomeration influence firm profits, which in turn influence firm location as well as survival in a region. This selection bias issue is very common in economic geography models, in which the locations of demand and production are determined jointly, and it makes these models difficult to test empirically.[8] We discuss below the various ways in which endogeneity issues may hamper our estimations, and offer some solutions.

#### Reverse-causation from TFP to input linkages

The location of a firm $i$ in a particular region does not directly increase this firm's measure of $PIL$, because $PIL$ is a function of the presence of input firms in *other* industries than the one to which $i$ belongs. However, $i$'s location decision can affect its measure of $PIL$ indirectly. For instance, if $i$ is a large, productive firm, (such as IBM or a car manufacturer) its location in $r$ may attract many input firms, making $PIL_{jrt}$ increase in time. There would therefore be some positive correlation between size (or productivity) and $PIL$ and some causation from TFP to $PIL$. I first investigate this possible issue by checking if larger firms have higher input linkages.

Table 2.5 shows that, far from having higher input linkages, larger firms, on average, have lower input linkages. Still, in the next subsection, I run regressions separately for small and large firms as a robustness check, and compare the results.

[8]This issue also arises in region-industry level regressions in Dumais, Ellison, and Glaeser (1997), and in Ellison and Glaeser (May 1999). See Overman, Redding, and Venables (2003) for a brief discussion of the endogeneity issue in these papers.

### Location decision

Since firms choose where they initially locate, they can choose their initial level of potential input linkages ($PIL$). We can therefore not assert that this initial level is assigned randomly. However, because firms in our sample do not change their location throughout the period, all time invariant regional (or industry-region) effects are captured in the firm fixed effect. What we exploit is therefore the variation of $PIL$ in time. Moreover, using Probit regressions of location in different regions on $PIL$ and other characteristics, we notice that $PIL$ is a strong positive predictor of location in Budapest and Pest (the capital and the region around the capital), and a strong negative predictor of location in all the other regions. This indicates that firms do not always choose to locate in regions where they will have high input linkages. In the next subsection, I run the model separately for the subsample of firms outside Budapest and Pest.

### Omitted variable bias

The regressions already control for omitted time-invariant regional characteristics through the firm fixed effect. We may however suspect that there is an omitted variable that can be described as the 'attractiveness' of region $r$ in year $t$, which is difficult to measure and may be correlated with potential input linkages. For example, improvement in a region's infrastructure will not be captured in our model. It would influence a firm's output, and would definitely be correlated with changes in the industrial composition of the region - and therefore with changes in $PIL$.

I do not offer an instrumental variable solution to this problem. The main reason is the lack of an instrument for $PIL$ that satisfies the exclusion restriction, i.e., that affects output only through its impact on potential input linkages and not directly. For instance, a variable reporting improvements in infrastructure would violate the exclusion restriction. Moreover, it will be difficult to find an instrument that predicts potential input linkages well but is orthogonal to the attractiveness of the region.

We can nevertheless try to sign the bias. Since the regional effect we have in mind is likely to increase output, and it is likely to be positively correlated with $PIL$, we expect the bias to be downward.

### Double indicator solution

Instead of an instrumental variable solution, I use a double indicator solution, as suggested in Wooldridge (2002). Rewriting equation (2.18), explicitly including the omitted variable $q_{rt}$, the attractiveness of region $r$ in year $t$, yields:

$$
\begin{aligned}
\ln y_{ijrt} &= \alpha_0 + \alpha_1 \ln K_{ijrt} + \alpha_2 \ln L_{ijrt} + \alpha_3 \ln M_{ijrt} \\
&\quad + \beta \ln PIL_{jrt} + \gamma \ln II_{jt} \\
&\quad + \delta \ln PIL_{jrt} \times \ln II_{jt} + \nu \ln q_{rt} + \eta_i + \epsilon_{ijrt}
\end{aligned}
\tag{2.19}
$$

with $E[\epsilon|x, q] = 0$, where

$$x = [\ln K, \ln L, \ln M, \ln PIL, \ln II, \ln PIL \times \ln II].$$

Instead of putting $\ln q_{rt}$ in the error term and instrumenting for $PIL$, I choose two indicators for $q_{rt}$. The first indicator will substitute for $q_{rt}$ in (2.19), and the second one will be used to instrument for the first one. The two indicators must be redundant in (2.19) and correlated with $q_{rt}$.

The first indicator of $q_{rt}$ is regional GDP, and can be written:

$$\ln GDP_{rt} = \delta_0 + \delta_1 \ln q_{rt} + a_{1ijrt} \tag{2.20}$$

with $\delta_1 \neq 0$, $Cov(q, a_1) = 0$ and $Cov(x, a_1) = 0$ by assumption. Rearranging, we get

$$\ln q_{rt} = -\frac{\delta_0}{\delta_1} + \frac{1}{\delta_1} \ln GDP_{rt} - \frac{1}{\delta_1} a_{1ijrt}. \tag{2.21}$$

Now $GDP$ and $a_1$ are correlated, and substituting for the above expression of $\ln q_{rt}$ in (2.19) requires that we instrument for $GDP_{rt}$. We therefore introduce a second indicator of $q_{rt}$, regional population density:

$$\ln den_{rt} = \rho_0 + \rho_1 \ln q_{rt} + a_{2ijrt} \tag{2.22}$$

with $\rho_1 \neq 0$, $Cov(q, a_2) = 0$ and $Cov(x, a_2) = 0$. I further assume that $Cov(a_1, a_2) = 0$.

The structural equation (2.19) becomes:

$$\begin{aligned} \ln y_{ijrt} &= (\alpha_0 - \nu\frac{\delta_0}{\delta_1}) + \alpha_1 \ln K_{ijrt} + \alpha_2 \ln L_{ijrt} + \alpha_3 \ln M_{ijrt} \\ &+\beta \ln PIL_{jrt} + \gamma \ln II_{jt} \\ &+\delta \ln PIL_{jrt} \times \ln II_{jt} \\ &+\frac{\nu}{\delta_1} \ln GDP_{rt} + \eta_i - \frac{\nu}{\delta_1} a_1 + \epsilon_{ijrt}. \end{aligned} \tag{2.23}$$

Although it is still doubtful whether *den* satisfies the exclusion restriction, it is orthogonal to the error term: *den* is uncorrelated with $\epsilon$ since it is redundant in (2.19), and it is also uncorrelated with $a_1$ since both $q$ and $a_2$ are uncorrelated with $a_1$ by assumption. Therefore we use $\ln den$ as an instrumental variable for $\ln GDP$ in (2.23).

**Autocorrelated standard errors**

For three main reasons, we may suspect that the errors are serially correlated. The first reason is the length of our panel (8 years, with the average number of years per firm being 4.9), and the second, that we suspect serial correlation in our measure of $PIL$ within a region-industry group. The third reason is that the dependent variable, firm output, is also likely to be serially correlated. This can lead to understating the standard errors.

I evaluate the magnitude of this issue. The serial correlation in $PIL$ is 0.18 and that of output is 0.22. I also estimate the autocorrelation of errors, by regressing the standard errors on their lagged values. The correlation at a one-period lag is 0.14, which is different enough from -0.5 and therefore means a rejection of the assumption that standard errors are serially uncorrelated. At a two-period lag, the correlation is -0.08.

Following one of the solutions suggested by Bertrand, Duflo, and Mullainathan (2004), I aggregate the data into two periods and run the OLS regression on a two-period panel. Although Bertrand, Duflo, and Mullainathan (2004) suggest three methods, this one is the simplest. I choose 1997 as the cutoff year, it being the year when manufacturing imports from the EU obtained full and free access into Hungary. The two periods are therefore 1995-1997 and 1998-2002.

**Moulton correction**

The dependent variable is at the firm level, while some of the explanatory variables are at a more aggregated level. As shown by Moulton (1990), using standard OLS would underestimate the standard errors if the latter are correlated within region-industry-year groups. To avoid this, I cluster the standard errors by region-industry-year groups.

**Levinsohn-Petrin productivity estimation**

Finally, the input coefficients obtained when estimating a production function by the ordinary least squares method are likely to be biased, due to the endogeneity of the input use. The coefficient on capital is likely to be underestimated, and the coefficients on labor and materials overestimated. I propose to correct these biases using the Levinsohn-Petrin methodology. I choose this method over Olley and Pakes because of the poor quality of the capital variable (and therefore of the data on investment) and because data on intermediate materials was available.

## 2.4 Results

### 2.4.1 Basic specification

The results of the basic OLS model are presented in column (3) of Table 2.6. As predicted, the coefficient on the interaction between imports and input linkages is negative, and it is significant at the 5% level. In order to understand the economic significance of the estimates, I can compute the partial effect of potential input linkages on firm output:[9] when imports of inputs are at their mean value, a 1% increase in potential input linkages yields a 1.6% increase in output, and when imports of inputs are near their minimum value, the corresponding increase in output is about 6%.

### 2.4.2 Corrections for endogeneity and serially correlated standard errors

Column (4) reports the results of the double-indicator method. I also obtain a negative and significant coefficient on the interaction term. Again we can interpret the coefficients in the following way: when imports of inputs are at their mean value, a 1% increase in potential input linkages yields a 5.3% increase in output. As expected, this corrects upwards the estimates obtained from the standard procedure (column (3)).

Finally, the tentative correction to the correlated standard errors problem is shown in column (5). After collapsing the dataset into two time periods, we are left with 1874 observations. Although I again obtain a negative coefficient on the interaction term, it is no longer significant. These results could mean that the standard errors in the basic model were indeed underestimated, or they could be due to a poor collapsing of the data into two arbitrary periods.

[9] I compute $0.623 - 0.031 \times \ln(II)$.

Table 2.6 and the robustness checks are done on the pooled dataset. However it makes more sense to analyze industry-level production functions. Tables 2.7 and 2.8 respectively show results of specifications (1) and (3) (fixed effects model) for individual industries. The coefficient on the interaction term is negative for all industries but one, although it is significant only for industry 7 (Machinery and Equipment).

### 2.4.3 Robustness checks

Table 2.9 presents the basic OLS results for the first and the last quartiles of firms by size. We see that the results still hold for the sub-sample of small firms, which tends to refute the objection that the results were driven by large firms attracting many input firms.

Table 2.10 presents again the basic model, run on the two capital regions and on the rest of the country separately. We see that the result seems to be driven not by firms in the capital regions with high potential input linkages (see Table 2.11 for summary statistics) but by firms in the rest of the country, where firm location is driven by other factors.

Finally, I check the robustness of the estimates to the exclusion of outliers. To this effect, I remove firms which have observations in the top or bottom 1% of the distribution of TFP. We are left with 6060 observations, or 93% or our initial sample. Table 2.12 shows that this truncation of the sample has very little impact on mean values of input linkages and imports of inputs. The results in Table 2.13 correspond to the first three columns of Table 2.6. The production function estimates are robust. The estimates for elements of TFP (input linkages, imports and their interaction) are also similar in magnitude, sign and significance to those found on the whole sample, and the overall effect of a 1% increase in potential input linkages found at the mean value of imports of inputs is now slightly higher, at 2.7%.

### 2.4.4 Levinsohn-Petrin productivity estimation

I finally estimate one value-added production function by industry, using the Levinsohn-Petrin method. The production function coefficients for each industry are shown in table 2.14. As a check of the validity of the LP estimates, Table 2.15 shows the corresponding estimates obtained using OLS. These have the expected biases, with the capital coefficient having a negative bias for all industries except 5 and 8, and the labor coefficient having a positive bias for all industries.

I then regress ln productivity on ln potential input linkages, ln import of inputs and their interaction term (as well as a firm fixed effect). The results are shown in Table 2.16. Again I find a negative and significant coefficient on the interaction term, meaning that an increase in the import of inputs has lowered the contribution of input linkages to productivity.

## 2.5 Conclusion

This paper contributes to the trade and productivity literature by investigating the changing impact of input linkages on productivity. It departs from usual studies of trade liberalization by concentrating on the role of imported inputs on firm productivity, and their interaction with the availability of local inputs. It offers a theoretical explanation for why input linkages should contribute less to firm productivity as trade is liberalized. The empirical results are fairly robust and match the predictions.

Although input linkages are highly endogenous and firm location and performance are subject to cumulative causation, the empirical analysis manages to circumvent these issues by using a panel of firms with constant location. The paper also deals with omitted variable bias in which the underlying attractiveness of a firm's region may explain firm performance.

The results are in line with the intuition and with the predictions of the model: the positive impact of high input linkages on firm productivity declines as trade of intermediates used as inputs by this industry increases. This is robust to corrections for omitted variable bias and endogeneity. This result is maintained in both the OLS and Levinsohn-Petrin methods of production function estimation. A future improvement to the empirical methodology will consist in adapting the Levinsohn-Petrin method to this set-up where input linkages enter TFP.

From a broader perspective, these empirical results lend support to theoretical models of economic geography which predict that trade liberalization causes economic activity to spread in space in the country that opens to trade. We see that if a firm's productivity is no longer enhanced by proximity to local inputs, it may have an incentive to relocate away from an industrial cluster. The results therefore shed light on the mechanism through which trade liberalization can cause the decline of old industrial centers. This decline may be both welcome (because it means a more even repartition of economic activity and less congestion) and dreaded. Either way it must be taken into account by policy makers when evaluating the impact of trade liberalization.

## Tables and Figures

Table 2.1: Summary statistics by region

| **Region** | **Mean operating revenue** (thousand constant HUF) | **N** |
|---|---|---|
| Budapest | 7576 | 1653 |
| Pest | 6135 | 771 |
| Gyor | 38226 | 448 |
| Vas | 13199 | 176 |
| Komarom | 15423 | 245 |
| Veszprem | 27207 | 223 |
| Fejer | 22984 | 356 |
| Zala | 4087 | 193 |
| Somogy | 6812 | 125 |
| Tolna | 3974 | 151 |
| Baranya | 17358 | 258 |
| Nograd | 3275 | 79 |
| Heves | 9294 | 174 |
| Borsod | 6308 | 228 |
| Szabolcs | 6041 | 215 |
| Jasz | 13860 | 196 |
| Hajdu | 9330 | 259 |
| Bacs | 4574 | 378 |
| Csongrad | 8457 | 249 |
| Bekes | 3501 | 125 |
| Total | 11607 | 6502 |

Table 2.2: Summary statistics by industry

| **Industry** | **Mean operating revenue** (thousand constant HUF) | **N** |
|---|---|---|
| 15 Food products and beverages | 10152 | 1034 |
| 16 Tobacco products | 6375 | 6 |
| 17 Textiles | 3822 | 188 |
| 18 Wearing apparel, fur | 11665 | 213 |
| 19 Leather, luggage, footwear | 7156 | 55 |
| 20 Wood and cork products | 3244 | 232 |
| 21 Pulp, paper and paper products | 5370 | 165 |
| 22 Publishing, printing | 3802 | 392 |
| 23 Coke, refined petroleum products and nuclear fuel | 88569 | 11 |
| 24 Chemicals and chemical products | 23119 | 339 |
| 25 Rubber and plastic products | 4274 | 546 |
| 26 Other non-metallic mineral products | 7175 | 297 |
| 27 Basic metals | 27701 | 190 |
| 28 Fabricated metal products | 3268 | 1039 |
| 29 Machinery and equipment n.e.c. | 8051 | 618 |
| 30 Office machinery and computers | 8834 | 47 |
| 31 Electrical machinery and apparatus n.e.c. | 9888 | 282 |
| 32 Radio, television and communication equipment | 32366 | 203 |
| 33 Medical, precision and optical instruments | 4834 | 154 |
| 34 Motor vehicles, trailers and semi-trailers | 99062 | 197 |
| 35 Other transport equipment | 7851 | 44 |
| 36 Furniture; manufacturing n.e.c. | 3667 | 250 |
| Total | 11607 | 6502 |

Table 2.3: Manufacturing industries

| **Number** | **Name** | **NACE industries** |
|---|---|---|
| 1 | Manufacture of food products, beverages and tobacco products | 15,16 |
| 2 | Textiles, wearing apparel, leather products | 17-19 |
| 3 | Manufacture of wood and paper products, printing | 20-22 |
| 4 | Chemical industry | 23-25 |
| 5 | Other non-metallic mineral products | 26 |
| 6 | Manufacture of basic metals and fabricated metal products | 27,28 |
| 7 | Machinery and equipment | 29-35 |
| 8 | Manufacturing n.e.c. | 36-37 |

Table 2.4: Summary statistics

| **Variable** | **Mean** | **Std. Dev.** | **Min.** | **Max.** | **N** |
|---|---|---|---|---|---|
| Potential input linkages | 0.0156 | 0.0166 | 0.0001 | 0.0740 | 6,502 |
| Imports of inputs | 315,269,328 | 162,095,903 | 76,728,432 | 714,649,408 | 6,502 |

Table 2.5: Mean potential input linkages and imports by firm size category

| **Quartile** | **Potential input linkages** | **Imports of inputs** | **N** |
|---|---|---|---|
| 1 | 0.0178 | 3.39e+08 | 1623 |
| 2 | 0.0153 | 3.32e+08 | 1627 |
| 3 | 0.0157 | 3.11e+08 | 1623 |
| 4 | 0.0135 | 2.78e+08 | 1629 |
| Total | 0.0156 | 3.15e+08 | 6502 |

Table 2.6: Production function OLS regressions

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Ln capital | 0.039** | 0.045** | 0.044** | 0.026* | 0.064** |
| | (0.018) | (0.019) | (0.019) | (0.015) | (0.025) |
| Ln labor | 0.190*** | 0.189*** | 0.190*** | 0.193*** | 0.082*** |
| | (0.035) | (0.035) | (0.035) | (0.010) | (0.023) |
| Ln materials | 0.723*** | 0.724*** | 0.724*** | 0.717*** | 0.808*** |
| | (0.035) | (0.036) | (0.036) | (0.010) | (0.023) |
| Ln potential input linkages | | 0.038 | 0.623** | 0.934*** | 0.349 |
| (ln PIL) | | (0.036) | (0.293) | (0.337) | (0.490) |
| Ln import of inputs | | -0.037 | -0.177** | -0.417** | -0.116 |
| (ln II) | | (0.023) | (0.072) | (0.166) | (0.115) |
| Ln PIL X ln II | | | -0.031** | -0.045*** | -0.020 |
| | | | (0.015) | (0.017) | (0.024) |
| Ln GDP | | | | 0.424 | |
| | | | | (0.269) | |
| Firm fixed effect | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.807 | 0.807 | 0.807 | | 0.809 |
| N | 6502 | 6502 | 6502 | 6502 | 1874 |

* significant at 10% level, ** significant at 5% level, *** significant at 1% level. Clustered standard errors.

Table 2.7: Production functions by industry (OLS)

| Industry | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Ln capital | -0.015 | 0.024 | 0.148*** | -0.053 | 0.132* | 0.071 | 0.070** | 0.109*** |
| | (0.032) | (0.029) | (0.046) | (0.049) | (0.073) | (0.044) | (0.034) | (0.029) |
| Ln labor | 0.179*** | 0.399*** | 0.172*** | 0.282*** | 0.356* | 0.172*** | 0.136 | 0.030 |
| | (0.046) | (0.041) | (0.049) | (0.066) | (0.181) | (0.045) | (0.094) | (0.048) |
| Ln materials | 0.719*** | 0.533*** | 0.585*** | 0.613*** | 0.624*** | 0.710*** | 0.825*** | 0.786*** |
| | (0.057) | (0.046) | (0.102) | (0.068) | (0.211) | (0.069) | (0.078) | (0.040) |
| $R^2$ | 0.816 | 0.909 | 0.787 | 0.788 | 0.864 | 0.767 | 0.829 | 0.831 |
| N | 1040 | 456 | 789 | 896 | 297 | 1229 | 1545 | 250 |

* significant at 10% level, ** significant at 5% level, *** significant at 1% level. Clustered standard errors.

Table 2.8: Production functions with interaction, by industry (OLS)

| Industry | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Ln capital | -0.018 | 0.043 | 0.151*** | -0.022 | 0.118* | 0.062 | 0.058* | 0.166*** |
| | (0.032) | (0.033) | (0.051) | (0.051) | (0.064) | (0.044) | (0.033) | (0.039) |
| Ln labor | 0.179*** | 0.385*** | 0.167*** | 0.279*** | 0.373* | 0.174*** | 0.140 | 0.033 |
| | (0.046) | (0.043) | (0.048) | (0.063) | (0.195) | (0.045) | (0.095) | (0.044) |
| Ln materials | 0.719*** | 0.539*** | 0.586*** | 0.614*** | 0.615*** | 0.706*** | 0.818*** | 0.774*** |
| | (0.056) | (0.043) | (0.101) | (0.067) | (0.221) | (0.069) | (0.081) | (0.040) |
| Ln PIL | 0.641 | 1.908 | -0.341 | 0.159 | 1.158 | 1.764 | 2.040*** | 0.676 |
| | (0.952) | (1.217) | (1.050) | (0.587) | (1.288) | (1.233) | (0.690) | (1.514) |
| Ln II | -0.138 | -0.686* | 0.122 | -0.160 | -0.207 | -0.371 | -0.383** | -0.378 |
| | (0.267) | (0.393) | (0.282) | (0.130) | (0.354) | (0.247) | (0.188) | (0.364) |
| Ln PIL X ln II | -0.031 | -0.096 | 0.027 | -0.004 | -0.052 | -0.100 | -0.106*** | -0.036 |
| | (0.048) | (0.063) | (0.052) | (0.030) | (0.066) | (0.068) | (0.037) | (0.075) |
| Firm fixed effect | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.817 | 0.912 | 0.789 | 0.794 | 0.865 | 0.770 | 0.830 | 0.839 |
| N | 1040 | 456 | 789 | 896 | 297 | 1229 | 1545 | 250 |

* significant at 10% level, ** significant at 5% level, *** significant at 1% level. Clustered standard errors.

Table 2.9: Comparison of small and large firms

| | **Small** | **Large** |
|---|---|---|
| Ln capital | 0.071* | 0.065** |
| | (0.040) | (0.032) |
| Ln labor | 0.102*** | 0.185*** |
| | (0.034) | (0.068) |
| Ln materials | 0.779*** | 0.721*** |
| | (0.055) | (0.067) |
| Ln potential input linkages | 2.277* | 0.357 |
| | (1.171) | (0.436) |
| Ln import of inputs | -0.475* | -0.106 |
| | (0.245) | (0.109) |
| Ln PIL X ln II | -0.117* | -0.015 |
| | (0.063) | (0.022) |
| Firm fixed effect | Yes | Yes |
| $R^2$ | 0.770 | 0.849 |
| N | 1506 | 1690 |

* significant at 10% level, ** significant at 5% level, *** significant at 1% level. Clustered standard errors.

Table 2.10: Regional comparison

| | **Budapest and Pest** | **Rest of the country** |
|---|---|---|
| Ln capital | 0.048* | 0.026 |
| | (0.028) | (0.025) |
| Ln labor | 0.102* | 0.247*** |
| | (0.054) | (0.042) |
| Ln materials | 0.732*** | 0.723*** |
| | (0.046) | (0.051) |
| Ln potential input linkages | 0.126 | 1.548*** |
| | (0.885) | (0.581) |
| Ln import of inputs | -0.035 | -0.459*** |
| | (0.152) | (0.164) |
| Ln PIL X ln II | -0.002 | -0.082*** |
| | (0.046) | (0.030) |
| Firm fixed effect | Yes | Yes |
| $R^2$ | 0.814 | 0.809 |
| N | 2424 | 4078 |

* significant at 10% level, ** significant at 5% level, *** significant at 1% level. Clustered standard errors.

Table 2.11: Potential input linkages and imports in Budapest and Pest vs. all other regions

| **Region** | **Potential input linkages** | **Imports of inputs** | **N** |
|---|---|---|---|
| Budapest and Pest | 0.0328 | 3.17e+08 | 2424 |
| Rest of the country | 0.0054 | 3.14e+08 | 4078 |
| Total | 0.0156 | 3.15e+08 | 6502 |

Table 2.12: Potential input linkages and imports after exclusion of outliers

| **Potential input linkages** | **Imports of inputs** | **N** |
|---|---|---|
| 0.0157 | 3.15e+08 | 6060 |

Table 2.13: Exclusion of outliers

| | (1) | (2) | (3) |
|---|---|---|---|
| Ln capital | 0.018** | 0.026*** | 0.025** |
| | (0.009) | (0.010) | (0.010) |
| Ln labor | 0.170*** | 0.168*** | 0.168*** |
| | (0.013) | (0.012) | (0.013) |
| Ln materials | 0.746*** | 0.748*** | 0.747*** |
| | (0.012) | (0.012) | (0.012) |
| Ln potential input linkages | | 0.038 | 0.497** |
| | | (0.024) | (0.203) |
| Ln import of inputs | | -0.040*** | -0.149*** |
| | | (0.015) | (0.053) |
| Ln PIL X ln II | | | -0.024** |
| | | | (0.010) |
| Firm fixed effect | Yes | Yes | Yes |
| $R^2$ | 0.913 | 0.913 | 0.914 |
| N | 6060 | 6060 | 6060 |

* significant at 10% level, ** significant at 5% level, *** significant at 1% level. Clustered standard errors.

Table 2.14: Value-added production functions by industry (Levinsohn-Petrin)

| Industry | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Ln capital | 0.040 | 0.149 | 0.125** | 0.147** | 0.106 | 0.413*** | 0.328*** | 0.133 |
| | (0.081) | (0.090) | (0.061) | (0.069) | (0.099) | (0.079) | (0.063) | (0.138) |
| Ln labor | 0.442*** | 0.629*** | 0.390*** | 0.451*** | 0.452*** | 0.389*** | 0.348*** | 0.389*** |
| | (0.055) | (0.059) | (0.068) | (0.051) | (0.063) | (0.052) | (0.041) | (0.140) |
| N | 1029 | 450 | 784 | 883 | 296 | 1216 | 1525 | 248 |

* significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 2.15: Value added production functions by industry (OLS)

| Industry | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Ln capital | -0.072 | 0.091* | 0.115** | -0.084* | 0.174** | 0.177*** | 0.284*** | 0.273*** |
| | (0.050) | (0.052) | (0.044) | (0.045) | (0.086) | (0.051) | (0.051) | (0.065) |
| Bias v. LP | - | - | - | - | + | - | - | + |
| Ln labor | 0.721*** | 0.850*** | 0.580*** | 0.678*** | 0.867*** | 0.700*** | 0.645** | 0.418*** |
| | (0.075) | (0.070) | (0.076) | (0.086) | (0.070) | (0.095) | (0.105) | (0.116) |
| Bias v. LP | + | + | + | + | + | + | + | + |
| N | 1029 | 450 | 784 | 883 | 296 | 1216 | 1525 | 248 |

* significant at 10% level, ** significant at 5% level, *** significant at 1% level.

Table 2.16: OLS regression of LP productivity estimates

| | (1) |
|---|---|
| Ln potential input linkages | -0.396*** |
| (ln PIL) | (0.115) |
| Ln import of inputs | -0.764*** |
| (ln II) | (0.249) |
| Ln PIL X ln II | -0.190*** |
| | (0.054) |
| Firm fixed effect | Yes |
| $R^2$ | 0.0067 |
| N | 6431 |

* significant at 10% level, ** significant at 5% level, *** significant at 1% level.

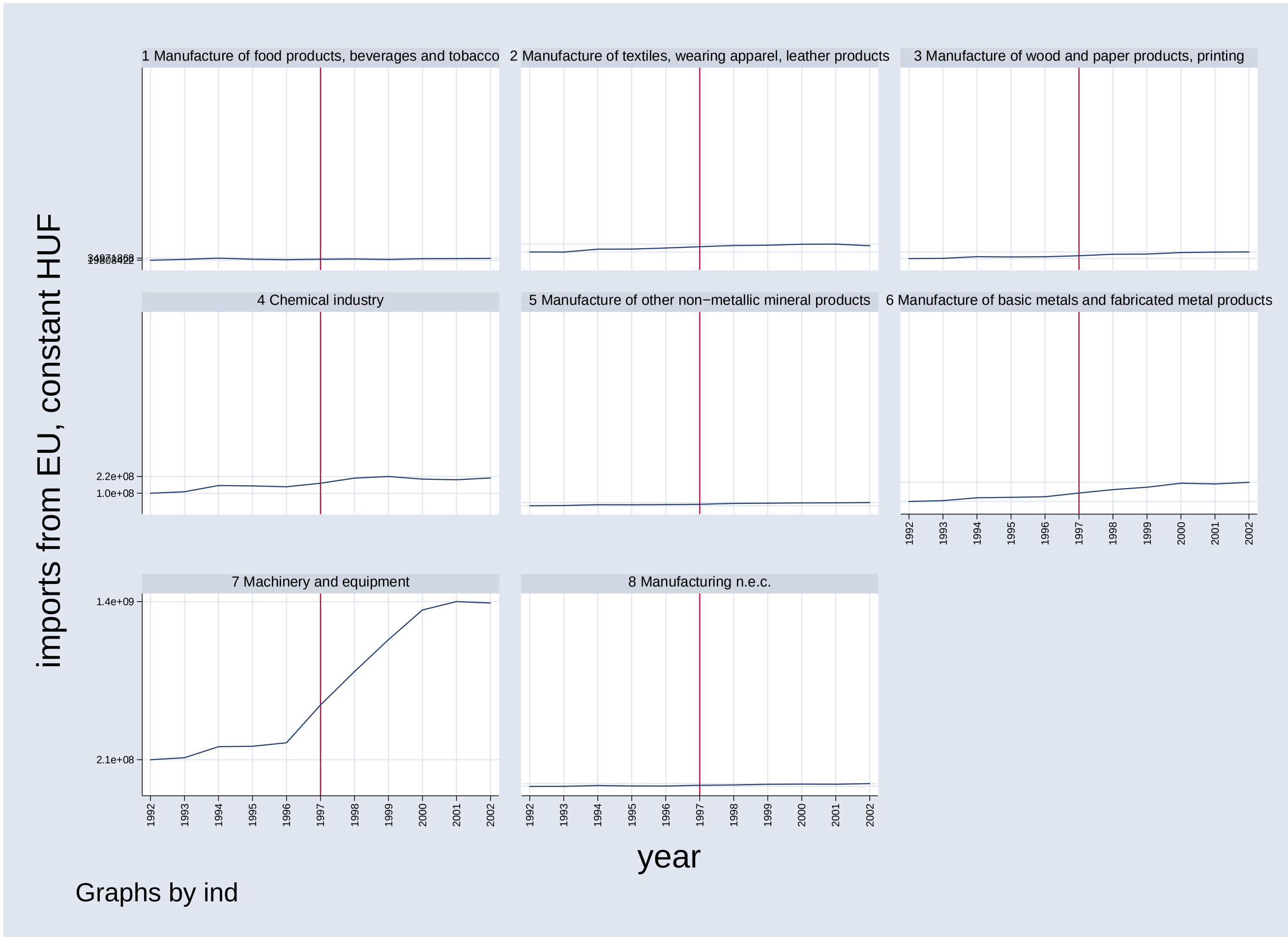


Figure 2.1: Imports by industry, constant HUF, 1992-2002

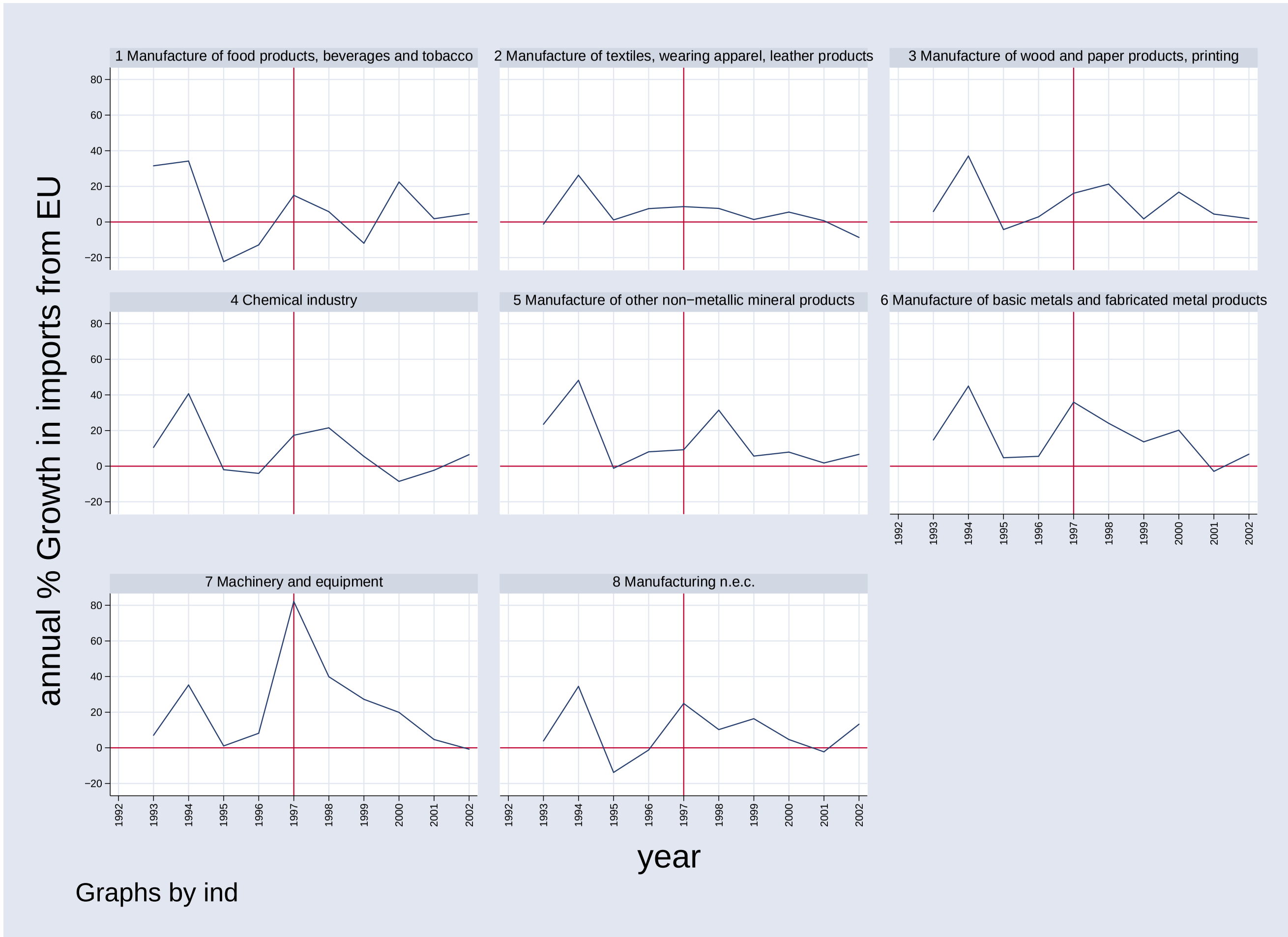


Figure 2.2: Annual growth in imports, 1992-2002

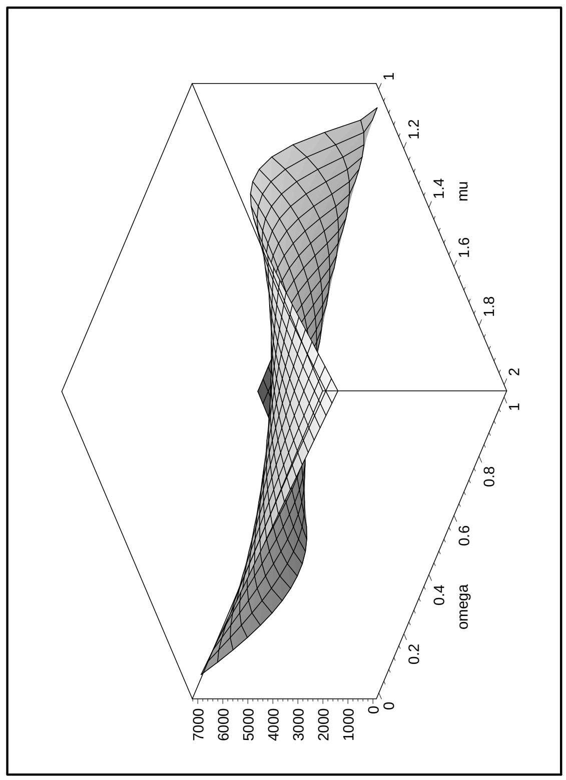


Figure 2.3: First derivative

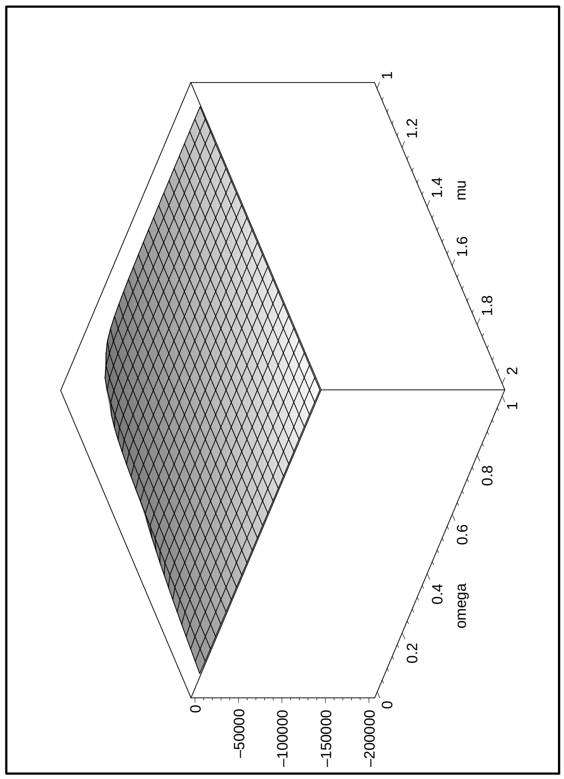


Figure 2.4: Second derivative

# CHAPTER III

# Trade liberalization and the geographical concentration of industries

## 3.1 Introduction

Trade liberalization is usually evaluated in terms of national performance indicators: economic growth, industry or firm-level productivity, wages. However its sub-national consequences are also important. More specifically, industries may geographically concentrate or spread because trade re-orientation or liberalization impacts the sub-national location of industries by altering the balance of dispersion and concentration forces and by modifying the patterns of market and supplier access.

Whether it is beneficial for industries to concentrate or spread depends on the initial conditions in the country. If, before trade liberalization, the country initially had an economic center that contained a large share of the economy, the spread of industries might be desirable as it would relieve congestion. Geographically concentrated industries also make regions more vulnerable to asymmetric shocks. As regional specialization is often the counterpart of industry localization, a shock to a given industry would have a disproportionate effect on the welfare of the region where this industry is localized. On the other hand, the existence of industrial networks is considered desirable as policy makers tend to give incentives for firms in a same industry to agglomerate in order to benefit from localization economies such as labor pooling, intra-industry linkages, or technological spillovers. In the presence of powerful industrial networks, further geographical concentration (also called clustering) of industries is considered desirable.

In this paper I determine the impact of external trade on the sub-national concentration of industries using the case of Hungary in the period following the Europe Agreement. Hungary is a particularly interesting case because its economy was highly geographically concentrated in the capital Budapest before the period of trade liberalization. Hungary already engaged in international trade with the West but it underwent a further trade liberalization with the EU-15 after the Europe Agreement in 1991: Between 1992 and 1997 all Hungarian industrial products obtained duty-free access to EU markets, and all manufacturing goods from the EU had full access into Hungary by the end of 2000. This has translated into a surge in trade flows with the EU, particularly after 1997. During the period of observation (1995-2002), overall manufacturing has spread, but the situation is more complex at the industry level, with some industries becoming more concentrated and others spreading. The share of Budapest county has also decreased for most industries. It is therefore relevant to ask to what extent trade liberalization has impacted the geographical concentration of industries in Hungary, and has caused spreading away from the capital.

The impact on the geographical concentration of the economy as a whole or of individual industries are separate questions. In this paper I focus on estimating the impact of trade liberalization on the concentration of individual industries, using sub-national data for Hungary, although I will also be able to check the prediction for the overall economy. In order to understand the theories analyzing the impact of trade on the spatial concentration of industries, it is necessary to first consider some of the mechanisms that cause overall economic concentration. Krugman and Elizondo (1996) develop a two-country, three-region model of new economic geography inspired by Krugman (1991), that predicts the impact of trade liberalization on the overall concentration of the economy. In autarky, a concentrated economy is maintained by strong input and output (or forward and backward) linkages. When the closed economy opens to trade, these linkages are weakened by the availability of foreign markets and the economy tends to spread geographically. Krugman and Elizondo (1996) thus explain the existence of huge metropolitan centers in countries that were closed to (or had limited) international trade. The predictions are based on simulations, the results of which depend on country-specific initial conditions. The limited empirical evidence on the subject seems to support this result: in a cross-country study, Ades and Glaeser (1997) find that the size of the largest city in a country is negatively correlated with its trade openness.

I now turn to predictions at the industry level. Two theories investigate the two-industry case. The first one is a theory of regional production networks used in Hanson (1996) for the apparel industry in Mexico during the country's trade liberalization. In autarky, high skill increasing return activities concentrate in the urban center (Mexico City), and low skill constant return activities locate in the periphery. With trade liberalization, the country specializes in low skill subcontracting, and industry therefore relocates towards the periphery (the US border). Location is determined by access to inputs and to foreign markets (both being in the US). In autarky, the low skill activities were present in both regions, but in the open economy they leave the urban center: the low skill industry therefore becomes more concentrated. So trade increases the geographical concentration of the dispersed industry and reduces the share of the economic center in this industry's employment, due to changes in market access.

In our context it is important to note that the results in Hanson (1996) could also have been brought about by the overall trade liberalization in effect in Mexico, independently of the relative location of the economic center and the trade partner: Krugman and Elizondo (1996) show that trade liberalization causes overall de-concentration away from the initial economic center, regardless of the geographical location of regions and of relative market access, because international trade reduces the importance of having local suppliers and customers nearby. In other words, trade weakens forward and backward linkages.

The second theory is based on this idea. It is a three-region (two domestic and one foreign), two-industry stylized theoretical model presented in chapter 18 of Fujita, Krugman, and Venables (1999), in which the authors focus directly on evaluating the impact of external trade on the internal geographical concentration of industries. This model is inspired by Krugman and Elizondo (1996). The main result is that although the economy as a whole tends to spread as trade is liberalized, individual industries become in fact more concentrated. The location of workers and production is the result of an equilibrium between two forces: Forward and backward linkages act as a centripetal force as firms locate in the economic center near their suppliers and customers, and final consumer demand as well as congestion costs act as a centrifugal force. As a country opens to trade, forward and backward linkages and final consumer demand are both weakened.[1] The initial conditions assumed in the model are that the economy is highly concentrated in the largest region (the center), the largest region is diversified and contains most of both industries, and the smallest region (called the periphery) is specialized in one industry (say industry 2). As trade is liberalized, international trade provides new sources of inputs and new customers, which weakens local forward and backward linkages. The location of firms is then driven by pre-existing intra-industry advantages such as spillovers and labor pooling. Firms in industry 2 that were located in the core region therefore relocate to the periphery, avoiding congestion costs while also being near other firms in industry 2. So industry 1 remains concentrated in the core while industry 2 concentrates in the periphery. With trade, the economy as a whole therefore becomes more spread out (as the smallest region receives more of industry 2 and the largest region decreases in size) and each industry more geographically concentrated. The mechanism at play here is the weakening of industrial linkages.

The prediction we can carry from the theory (in a two-industry, two-region set up) is therefore

[1]In Chapter 1 of this dissertation, I show that trade opening weakens input-output linkages at the firm level.

that trade increases the geographical concentration of industries. Industries that were initially spread out leave the center and concentrate in the periphery, while industries that were initially concentrated are expected to be less impacted by trade. Table 3.1 summarizes the predictions of the three papers mentioned above. Only Hanson (1996) and Fujita, Krugman, and Venables (1999) provide predictions at the industry level, that individual industries that were initially dispersed should become more concentrated, which I will test using Hungarian data. All three models make predictions on the concentration of the economy as a whole, predicting that the economy should spread[2] and the share of the economic center should decrease.

Although Hanson (1996) and Fujita, Krugman, and Venables (1999) make a clear prediction in terms of the geographical concentration of industries during trade liberalization, this is not directly testable empirically. Firstly, the theoretical models say nothing when there are more than two industries. Secondly, the relation between trade and geographical concentration identified in Fujita, Krugman, and Venables (1999), like most relationships in new economic geography models, is highly nonlinear: the pattern of production will at some point jump discontinuously between concentration in one region and a symmetric equilibrium with equal repartition across regions. As Head and Mayer (2004) pointed out, if we are to take this prediction literally, this means that a linear regression would be misspecified.

For these reasons, instead of trying to estimate these theoretical predictions directly, this paper aims to answer the question empirically. I will be estimating the relation between trade and the spatial concentration of industries (column 2 of Table 3.1) and the relation between trade and the share of the economic center (Budapest) in an industry. I control for other determinants of industry localization that have been highlighted in the prior empirical literature, as well as for industry fixed effects. Regressions of the Krugman index of industrial concentration on industry exports and imports of inputs reveal that exports increase the geographical concentration of industries, and more so for industries that were initially dispersed.

The empirical literature on economic integration and industry concentration has given mixed results. Most of the literature has focused on concentration patterns over time, or on estimating the determinants of industrial concentration. Conclusions may depend on whether the papers look at industry concentration patterns *across* countries or across regions *within* countries. Brulhart (2001) finds a slight increase in industrial concentration across countries in Europe in time, while others find decreasing trends (Aiginger and Davies (2004), Midelfart-Knarvik, Overman, Redding,

[2]Hanson (1996) deals with production networks, we should therefore see this as spreading of a production network rather than of the whole economy.

and Venables (2002)). Midelfart-Knarvik, Overman, Redding, and Venables (2002) use Gini coefficients of industrial concentration for 36 manufacturing industries across EU countries, for four periods between 1970 and 1997, and find that patterns of geographical concentration vary greatly across industries and over time. Overall, concentration is decreasing for 24 of the 26 industries. Brulhart (2001)'s contrasting results may be due to the fact that he excludes certain countries due to data limitations and compares concentration patterns in 1972 and 1996 while Midelfart-Knarvik, Overman, Redding, and Venables (2002) look at several time averages between 1970 and 1997.

The literature investigating the determinants of the geographical concentration of industries began with Kim (1995). He computes Gini indices for US industries and regresses these on industry characteristics, notably the average firm size in the industry as a proxy for returns to scale and the share of raw materials used in the industry as a proxy for the role of input-output linkages. He also includes industry and year fixed effects in order to remedy the fact that many variables are omitted, notably the spatial distribution of factors and trade variables such as capital and labor intensities. He finds that both returns to scale and linkages positively affect the degree of industrial concentration in the US. A similar study by Amiti (1999) using Gini coefficients for 65 manufacturing industries in Europe also finds positive effects of returns to scale and linkages. The empirical evidence on the determinants of the geographical concentration of industries is summarized in Combes and Overman (2004): there is little evidence that labor, capital or resources intensity increases concentration, and mixed evidence on the effect of demand and cost linkages as well as trade liberalization. There is positive evidence that technology intensive and high returns to scale industries are more concentrated.

Finally, there are very few papers looking for subnational effects of external trade liberalization. A notable example is Overman and Winters (2003), who analyze the impact of the accession of the United Kingdom into the EU on the location of industries in the United Kingdom. They find evidence of relocation towards the South East, caused by market access to mainland Europe.

The next section describes the international trade patterns observed during the period, and section 3 how the geographical concentration of industries has evolved. Section 4 estimates the effect of trade on industrial concentration. Section 5 concludes.

## 3.2 Trade patterns

Hungary already engaged in international trade with the West but it underwent a further trade liberalization after the Europe Agreement in 1991, which was accelerated after the Copenhagen

Council of 1993. A ten year schedule for the lowering of tariffs and quotas was set up between Hungary and the EU. As early as 1992, some industrial products could be exported to the EU duty-free, and by early 1995 all industrial products except textiles and steel benefited from full duty-free access. Steel obtained free access at the beginning of 1996, and textile products at the beginning of 1997. This trade liberalization was asymmetric in the sense that imports into Hungary from the EU obtained free access with a time lag: all manufacturing goods from the EU had full access into Hungary by the end of 2000.

This trade policy change began when Hungary was coming out of a deep recession, which lasted from 1990 until 1993, as illustrated in Appendix 2. This seems to have caused a delay in the impact of the policy on trade flows with the EU, which is why I focus the analysis on the period 1995-2002. Between 1995 and 2002 both the volume of trade with the EU and the share of the EU in Hungarian exports greatly increased. Exports to the EU increased by 220% between 1995 and 2002 in real terms (while GDP increased by 31%), and imports from the EU increased by 135%. The share of exports to the EU in total Hungarian exports increased from 63% in 1995 to 76% in 1999, while the share of imports from the EU remained rather stable, increasing from 60% in 1995 to 64% in 1999, then decreasing to 56% in 2002. Looking at manufactured goods only, 90% of exports went to the EU in 1999. Figure 3.1 shows the total import and export flows with the EU-15, in 1992 Hungarian Forints deflated by import and export PPI (base 100 in 1992). They have roughly quadrupled between 1992 and 2002, with a sharp rise after 1997. The process of trade integration was therefore both rapid and deep.

I now consider the trade patterns at the industry level. I use import and export flows with the EU, from the OECD's STAN Bilateral Trade database, for 18 manufacturing industries. Most of these industries correspond to a single two-digit NACE industry (equivalent to ISIC Rev. 3.3), however some combine two or three two-digit industries. Table 3.2 indicates the breakdown of manufacturing into these 18 industries which will be used for the analysis throughout the paper, and which two-digit NACE industries they correspond to.

Figures 3.2 and 3.3 present respectively the industry-level import and export flows adjusted by industry value added.[3] We can see that there is considerable variation in trade patterns across industries. Exports of textiles, leather and footwear have more than tripled between 1992 and 2002, and exports of office, accounting and computing machinery have increased by a multiple of seven, while the exports of several industries have in fact decreased (food products, beverages and tobacco,

[3] I deflate trade flows by industry value added rather than industry-level GDP because of the high proportion of re-exports in the Hungarian economy. The source of the value added data is the OECD's STAN Industry database.

wood and products of wood and cork, and basic metals for instance). There is less variation in import patterns, as import flows deflated by value added have decreased for almost all industries, although they have increased in value terms (in constant Hungarian Forints).

## 3.3 Industry concentration patterns

### 3.3.1 Measuring geographical concentration

Having identified variation in the industry-level trade patterns, I now check if this was accompanied by changes in the geographical concentration of industries. I first consider various measures of concentration, before focusing on the Krugman concentration index for the empirical analysis, and I finally present the patterns.

In order to measure the geographical concentration of industries, I use region-industry employment data for the period 1995-2002 from the Hungarian Labor Force Survey. The regions are the 20 Hungarian counties (NUTS3 level of geographical aggregation). These data can be used to compute different measures of geographical concentration.

A simple measure is given by the geographical Herfindahl index, the sum of the squared shares of regions $r$ in the employment of industry $j$. However the main drawback of this measure is that it assumes that regions have equal sizes, which is far from being the case in Hungary. In contrast three other measures, the locational Gini coefficient, the Krugman concentration index and the Ellison and Glaeser index, which I discuss below, are more appropriate as they have the advantage of taking into account the underlying regional repartition of the economy.

A second possible measure, the Theil index of inequality, from the class of generalized entropy indices, is particularly useful compared to other measures because it is additively decomposable.[4] The general form of Theil's index is:

$$T^j = \frac{1}{N} \sum_i \frac{l_i^j}{\bar{l}^j} log(\frac{l_i^j}{\bar{l}^j})$$

where $i \in \{1...N\}$ denotes basic spatial units of equal size (regions differ in their number of basic spatial units), and $\bar{l}^j = \frac{l^j}{N}$.

The decomposability of this index is a useful feature if one wishes to compare the within-country and the between-country components of economic concentration or to decompose overall economic concentration in a country into industry-level components. However this is not the focus of this

[4]See Brulhart and Traeger (2005) who use this index to account for geographical concentration patterns in Europe.

paper. An important drawback of the additive decomposability feature is that there is no benchmark of equality attached to the entropy index. In our context it is crucial to use a measure of geographical concentration that takes into account at least the underlying distribution of manufacturing,[5] so this class of indices does not match the trade models as well as the locational Gini, the Krugman index or the Ellison and Glaeser index.

I now turn to the locational Gini coefficient. Computing the locational Gini coefficient for each industry requires that we first compute the location index (or Hoover-Balassa index) for each industry-region combination:

$$r_r^j = \frac{l_r^j/l^j}{l_r/L}$$

where $l_r$ is employment in region $r$ and $L$ is total manufacturing employment. Therefore $r_r^j$ is the share of region $r$ in industry $j$ normalized by the share of region $r$ in manufacturing.

For a given industry $j$, the distribution of $r_r^j$ across regions $r$ is an indication of the localization (or geographical concentration) of this industry. The locational Gini coefficient summarizes this distribution: it is twice the area between the Lorenz curve associated with the location indices and the 45° line. The Lorenz curve ranks regions in ascending order of their location index and plots cumulated values of $l_r^j/l^j$ against cumulated values of $l_r/L$. The Gini coefficient therefore takes value zero if for each region $r$, the share of $r$ in industry $j$ is equal to the share of $r$ in manufacturing. Therefore the benchmark of no concentration is that the repartition of industry $j$ across regions is the same as that of overall manufacturing. High values of the Gini coefficient mean that an industry is geographically concentrated compared to the overall economy.

In the Gini coefficient, the impact of transfers across regions is determined by the rankings of regions in the distribution of region-industry employment. If two regions keep the same employment of a given industry but more regions appear in between these two employment levels, this would change the Gini coefficient. Moreover, the associated Lorenz curves cross if over time industry shares $l_r^j/l^j$ decline both in low and high share regions, with an ambiguous effect on overall concentration.

A main drawback of the Gini coefficient - as well as of all the other measures - is that it does not take into account the underlying industrial structure. Some industries may be geographically concentrated just because they have a small number of establishments, as pointed out by Ellison and Glaeser (1997). The Ellison and Glaeser index has the advantage of also taking into account the underlying concentration level of the industry. However implementation of the Ellison and Glaeser index requires that we know the Herfindahl index of industrial concentration, which is not available

[5] For a more detailed discussion of the main criteria for a desirable index of geographical concentration, see Combes and Overman (2004). For a more technical discussion of inequality measures, see Sen (1997) and Atkinson (1998).

and may not make much sense at this rather broad level of disaggregation (Ellison and Glaeser (1997) use 4-digit level data). I am therefore not able to use this index in this paper.

The Krugman concentration index, introduced in Krugman (1991a), takes the form:

$$KC_j = \sum_r |s_r^j - \overline{s^j}|$$

where $s_r^j = \frac{l_r^j}{l_r}$ is the employment share of industry $j$ in region $r$ and $\overline{s^j} = \frac{l^j}{L}$ is the employment share of industry $j$ in manufacturing. This index takes values between 0 and 2, and is equal to zero if for each location $r$ the share of industry $j$ is equal to the share of $j$ in manufacturing. The benchmark of no concentration is therefore the same as for the Gini coefficient, that industry $j$ 's repartition across the regions maps that of manufacturing as a whole. High values of the Krugman index also indicate that the repartition of industry $j$ across locations departs from that of manufacturing as a whole.

I rule out the Gini coefficient in favor of the Krugman index because the latter is more adapted to the trade models. The Gini coefficient is rather a measure of inequality and it is sensitive to changes in the rankings of regions through transfers between the two ends of the distribution.[6]

### 3.3.2 The patterns

The level of geographical aggregation used in this paper is NUTS3, corresponding to the 20 Hungarian counties. The situation in 1995 and historically was one of strong economic concentration in the capital Budapest, which makes Hungary a good subject for this analysis. A political map of Hungary is given in Appendix 3. Figure 3.4 plots the regional shares of manufacturing against the rank of each region in manufacturing in 1995, for 1995 (dots) and 2002 (crosses). Focusing on the dots we can see that in 1995 the distribution of regional shares was very skewed towards Budapest (rank 1) and Pest (the region around Budapest, ranked 2), and that the economy was therefore highly concentrated in the capital. Budapest county alone (the smallest of 20 counties, accounting for only 0.5% of the total area) accounted for 17% of manufacturing employment in 1995.

In line with the predictions in Krugman and Elizondo (1996), overall manufacturing has spread during the period: Comparing the distributions of regional manufacturing shares in 1995 and 2002 in Figure 3.4, we can also see that by 2002 (shown by the crosses) there had been some degree of convergence in the regional shares of manufacturing, with Budapest's share in manufacturing falling

[6]The Gini coefficient is often used to measure income inequality, as in Garner and Terrell (1997) and Gottschalk and Smeeding (1997).

to 15% in 2002 and the shares of the lower ranked regions increasing. Moreover the Herfindahl index of geographical concentration for overall manufacturing decreased slightly, from 0.0694 to 0.0655 between 1995 and 2002.

The situation is more complex at the industry level. Figure 3.5 plots the Krugman concentration index for each industry in 2002 against that in 1995. Out of 18 industries, 4 exhibited the same level of concentration in 2002 as in 1995, 6 had higher concentration in 2002, and 8 had lower concentration. I now take a closer look at the evolution of the employment share of the capital region (Budapest county) in each industry. Figure 3.6 shows that the share of Budapest has decreased for almost all industries.

To conclude on these concentration patterns, the reduction in the concentration of overall manufacturing seems to give support to Krugman and Elizondo (1996), Fujita, Krugman, and Venables (1999) and Hanson (1996). However we see that the patterns of industry concentration are more complex, and I now turn to the econometric analysis of the relationship between trade and the geographical concentration of industries.

## 3.4 Testing for the relationship between trade and geographical concentration

The intuition we carry from the stylized models presented in Section 1 is that increased trade makes industries that were initially spread out leave the center and concentrate in the periphery, and industries that were initially concentrated are expected to be less impacted. However, as stated earlier there are factors other than international trade that affect the geographical concentration of industries. Since there is empirical evidence that geographical concentration is positively correlated with the degree of returns to scale in an industry, I control for returns to scale using average plant size in the industry. Moreover, the starting up of large FDI ventures is also likely to explain part of the changes in industry concentration in Hungary. Bearing this in mind, I will also control for FDI stock in the industry.

I therefore test the impact of international trade on the spatial concentration of industries by regressing the log of the Krugman coefficient of industry $j$ in year $t$ on log industry exports and log of imported inputs for that industry, controlling for log returns to scale and log FDI as well as industry fixed effects. I estimate the following equation:

$$\ln KC_t^j = \alpha_1 \ln Exports_t^j + \alpha_2 \ln II_t^j + \alpha_3 \ln Scale_t^j + \alpha_4 \ln FDI_t^j + \delta^j + \varepsilon_t^j \tag{3.1}$$

Since the insight from theory is that international trade affects industry concentration through the possibility of exporting and the access to foreign inputs, I include separately $ln\ Exports_t^j$, the log of industry-level value added adjusted exports to the EU, and $ln\ II_t^j$, the log of the measure of imported inputs from the EU. The source of trade flow and value added data is the OECD, as in Section 2.

Imported inputs for industry $j$ in year $t$ are defined as:

$$II_t^j = \Sigma_k \frac{Imports_t^k}{VA_t^k} \times I_{kj}$$

where $VA_t^k$ is value added of industry $k$ in year $t$ and $I_{kj}$ is industry $j$'s share of inputs coming from industry $k$ (from the input-output matrix published by the Hungarian Statistical Institute). So imported inputs are the sum of import flows of all industries weighted by their share in industry $j$'s inputs.

$FDI_t^j$ is deflated FDI stock for industry $j$ in year $t$, in million Euros deflated by investment deflators. These data are from the Hungarian National Bank and cover only the years 1998 to 2002.

$ln\ Scale_t^j$ is the log of average plant size in industry $j$ in year $t$ in terms of employment, a proxy for industry-level returns to scale. The source of the plant size data is the Hungarian Labor Force Survey. Industry fixed effects, $\delta^j$, are included to control for all time-invariant industry characteristics and omitted variables such as industry structure, the location and concentration of relevant endowments, and the time-invariant component of intra-industry linkages.

Table 3.3 gives summary statistics of the variables. The Krugman index varies greatly from 0.07 for industry 12 (Office, accounting and computing machinery) in 1995 to 1.56 for industry 2 (Textile, leather and footwear) in 1998. $FDI$ and $Exports$ are positively correlated with the Krugman index, while imports of inputs are practically uncorrelated with the Krugman index and negatively correlated with exports. Average plant scale is negatively correlated with the degree of concentration.[7]

Table 3.4 reports the results from the baseline regression of $ln\ KC$ on the explanatory variables.

[7]This feature of the data goes against previous empirical findings, however it is robust to the use of two alternative measures of returns to scale: the median and the average plant size from the Amadeus dataset.

The results in column 1, without taking returns to scale or FDI stock into account, indicate that a 1% increase in exports increases the Krugman coefficient by 32%, whereas imports of inputs seem to reduce concentration. The coefficients on $Scale$ and $FDI$ are not statistically significant from zero (columns 2-4). However the regressions including $FDI$ contain fewer observations as the data on FDI are available for only five years. In order to check whether using a reduced sample makes a difference, I show in columns 5 and 6 the same regressions as in columns 1 and 2 respectively, with the reduced sample. The results on columns 3-6 indicate that the inclusion of $Scale$ and $FDI$ does not change the coefficients on exports and imports of inputs much. So a 1% increase in exports leads to an increase in the Krugman index of 23-26% based on the reduced sample and 32% based on the full sample, whereas imports of inputs have a negative effect, which does not appear significant in the reduced sample.

In Table 3.5, I estimate the same model with a one-year lag on the independent variables. We can indeed expect trade flows and FDI to have a lagged effect on the geographical concentration of industries. The columns in Table 3.5 report the same regressions as the first four columns in Table 3.4. The results are very similar to those of the model without lags, with a positive effect of exports that however does not remain significant after I include $Scale$ and $FDI$.

I now check whether the effect of trade is stronger for industries that were less geographically concentrated in 1995, as suggested by the theory. In Table 3.6, the 'high concentration' industries are the five industries with a Krugman index in the top 25 percent in 1995 (industries 1, 2, 6, 8 and 9), and the 'low concentration' industries are the five industries with a 1995 Krugman index in the bottom 25 percent (industries 5, 12, 14, 15 and 17). Comparing columns 1 and 3, we see that the coefficient on exports is higher for high concentration industries, however once I introduce the full set of variables (columns 2 and 4) we have an effect of exports and imports of inputs not statistically different from zero for high concentration industries and a coefficient on exports that is large, positive and significant for low concentration industries. These results reinforce the previous results that trade positively impacts the geographical concentration of industries through export flows, and that most of the effect is seen among industries that were initially dispersed.

The theory tells us that with an increase in trade, industries that were not initially concentrated in the economic center will concentrate away from the economic center, and industries that were concentrated in the economic center will maintain that concentration. I test this by using the share of Budapest in industry employment as the dependent variable, and by isolating two groups of industries. The group of industries with a high employment share of Budapest are the five industries that were in the top 25 percent in terms of their share of Budapest in 1995 (industries 4, 6, 10, 14 and

15), and the group of low share of Budapest industries are the five industries that were in the bottom 25 percent in 1995 (industries 1, 2, 5, 9 and 16). I expect the coefficients on the trade variables to be positive or not significantly different from zero for the high Budapest share industries, and negative for the low Budapest share industries.

Table 3.7 reports the regression results for the full sample of industries. As in Table 3.4, the last two columns report the same specifications as columns 1 and 2 based on the reduced sample used in columns 3 and 4. Exports have an overall negative effect on the share of the industry located in Budapest, although this effect loses its significance after I include $Scale$ and $FDI$, while imports of inputs have a positive effect. This seems to indicate that overall, industries that experienced an increase in their exports to the EU have also relocated away from Budapest.

The fact that exports to the EU increase the Krugman index and at the same time decrease the share of Budapest suggests that in the case of Hungary market access has played a crucial role, as in Hanson (1996). There is trade-induced concentration of industries accompanied by relocation of these industries away from the economic center. This is hardly surprising as Budapest is located away from the border with the EU. The results in Table 3.7 also indicate that exports and imports of inputs have opposite effects on the share of Budapest: industries that experienced an increase in their imported inputs from the EU have increasingly located in Budapest, while industries that experienced at increase in their exports have located away from Budapest. Proximity to the client market (near the Western part of the country, away from Budapest) therefore appears more relevant for exports than proximity to the supplier market is for imports.

Finally, Table 3.8 presents the same specifications as the first four columns of Table 3.7 with lagged independent variables. The results are very similar to those of the model without time lags. Finally, Table 3.9 presents the results for the estimation on two groups of industries according to their initial share of employment located in Budapest county. First, the coefficients on exports and imports of inputs change with the inclusion of controls for scale and FDI. If we focus on the full specification (columns 2 and 4), we see that as predicted the effect of the trade variables is insignificant for the industries that had a high share of employment in Budapest, while it is significant for the industries with a low share in Budapest: for these industries, exports have a negative and significant effect (a 1% increase in exports decreases the share of Budapest by 0.12 points) and imports of inputs have a positive and significant effect.

## 3.5 Conclusion

This paper has estimated the impact of international trade on the geographical concentration of industries in Hungary following trade liberalization with the EU-15. The first set of results are obtained by regressing the Krugman index of industrial concentration on exports, imports of inputs, a proxy for industry returns to scale and FDI stock. Although the effects of the trade variables are significant, the number of explanatory variables is fairly limited, which the use of industry fixed effects partly compensates for. Moreover, the proxy for industry-level returns to scale, although commonly used in the literature, does not have the positive and significant effect on concentration that would be expected.

In the first set of results, exports to the EU increase the geographical concentration of industries, particularly for industries that were initially dispersed, while imports of inputs from the EU do not have a robust effect. Trade therefore seems to be affecting the geographical concentration of industries mainly through exports. This effect is stronger for industries that had a low initial level of concentration, as predicted by theory.

The second set of results is obtained by using the share of Budapest in industry employment as the dependent variable. Increasing exports reduces the share of the capital Budapest in an industry, and more so for industries that initially had a low share of Budapest. On the other hand, imports of inputs from the EU increase the share of Budapest in an industry's employment. Through the lens of Hanson (1996), this could mean that supplier access is not relevant, as Budapest is away from the EU border, or that inputs from the EU arrive into Hungary via the capital. Finally, combining the two sets of results about industrial concentration and the share of Budapest tells us that increased exports in an industry have brought about higher geographical concentration away from the capital region. This means that trade has increased the geographical concentration of industries without necessarily reinforcing existing industrial networks.

## Tables and Figures

Table 3.1: Predictions on the effect of an increase in international trade

| | Concentration of the economy | Concentration of individual industries | Share of economic center |
|---|---|---|---|
| Krugman and Livas Elizondo (1996) | - | | - |
| Fujita, Krugman and Venables (1999) | - | + | - |
| Hanson (1996) | - | + | - |

Table 3.2: Manufacturing industries

| Number | Name | NACE / ISIC |
|---|---|---|
| 1 | Food products, beverages and tobacco | 15-16 |
| 2 | Textile, leather and footwear | 17-19 |
| 3 | Wood and products of wood and cork | 20 |
| 4 | Pulp, paper, printing, publishing | 21-22 |
| 5 | Coke, refined petroleum products and nuclear fuel | 23 |
| 6 | Chemicals and chemical products | 24 |
| 7 | Rubber and plastic products | 25 |
| 8 | Other non-metallic mineral products | 26 |
| 9 | Basic metals | 27 |
| 10 | Fabricated metal products | 28 |
| 11 | Machinery and equipment, n.e.c. | 29 |
| 12 | Office, accounting and computing machinery | 30 |
| 13 | Electrical machinery and apparatus, n.e.c. | 31 |
| 14 | Radio, television and communication equipment | 32 |
| 15 | Medical, precision and optical instruments | 33 |
| 16 | Motor vehicles, trailers and semi-trailers | 34 |
| 17 | Other transport equipment | 35 |
| 18 | Manufacturing n.e.c.; recycling | 36-37 |

Table 3.3: Summary statistics of the variables

| | **Krugman index** | **Exports** | **Imported inputs** | **Scale** | **FDI** |
|---|---|---|---|---|---|
| Mean | 0.57 | 8.41 | 1.01 | 94.34 | 2.10e+11 |
| Standard deviation | 0.33 | 9.04 | 0.67 | 19.70 | 1.83e+11 |
| Min | 0.07 | 1.00 | 0.07 | 61.88 | 4.57e+08 |
| Max | 1.56 | 67.25 | 2.65 | 197.50 | 6.21e+11 |
| N | 144 | 144 | 144 | 144 | 90 |

Table 3.4: Trade flows and geographical concentration

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ln exports | 0.32*** | 0.32*** | 0.26*** | 0.25*** | 0.23*** | 0.23*** |
| | (6.64) | (6.75) | (3.18) | (3.17) | (3.19) | (3.06) |
| Ln imports of inputs | -0.21** | -0.26*** | -0.14 | -0.07 | -0.10 | -0.04 |
| | (2.39) | (2.70) | (0.90) | (0.42) | (0.71) | (0.25) |
| Ln scale | | 0.17 | | 0.17 | | 0.11 |
| | | (1.27) | | (0.83) | | (0.60) |
| Ln FDI | | | -0.02 | -0.02 | | |
| | | | (0.66) | (0.88) | | |
| Industry fixed effect | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.27 | 0.28 | 0.13 | 0.13 | 0.13 | 0.13 |
| N | 144 | 144 | 90 | 90 | 90 | 90 |

Dependent variable: ln KC. * significant at 10% level, ** significant at 5% level, *** significant at 1%level. Absolute value of t statistics in parentheses.

Table 3.5: Lagged trade flows and geographical concentration

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Ln exports$_{t-1}$ | 0.27*** | 0.28*** | 0.17 | 0.17 |
| | (5.77) | (5.85) | (1.52) | (1.53) |
| Ln imports of inputs$_{t-1}$ | -0.25*** | -0.29*** | -0.12 | -0.01 |
| | (3.09) | (3.25) | (0.67) | (0.06) |
| Ln scale$_{t-1}$ | | 0.13 | | 0.24 |
| | | (1.04) | | (0.80) |
| Ln FDI$_{t-1}$ | | | -0.02 | -0.03 |
| | | | (0.42) | (0.63) |
| Industry fixed effect | Yes | Yes | Yes | Yes |
| $R^2$ | 0.26 | 0.27 | 0.05 | 0.06 |
| N | 126 | 126 | 72 | 72 |

Dependent variable: ln KC. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. Absolute value of t statistics in parentheses.

Table 3.6: High vs. low concentration industries

| | **High concentration** | | **Low concentration** | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Ln exports | 0.85*** | -0.15 | 0.36*** | 0.37** |
| | (4.95) | (0.37) | (4.02) | (2.69) |
| Ln imports of inputs | -0.71** | 0.40 | -0.35** | -0.64 |
| | (2.52) | (0.79) | (2.15) | (1.25) |
| Ln scale | | 0.48* | | -0.65 |
| | | (0.99) | | (0.87) |
| Ln FDI | | -0.48** | | -0.01 |
| | | (2.31) | | (0.28) |
| Industry fixed effect | Yes | Yes | Yes | Yes |
| $R^2$ | 0.56 | 0.50 | 0.39 | 0.35 |
| N | 40 | 25 | 40 | 25 |

Dependent variable: ln KC. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. Absolute value of t statistics in parentheses.

Table 3.7: Trade flows and the employment share of Budapest

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Ln exports | -0.05*** | -0.05*** | -0.03 | -0.03 | -0.05** | -0.05** |
| | (4.54) | (4.34) | (1.35) | (1.36) | (2.14) | (2.17) |
| Ln imports of inputs | 0.12*** | 0.10*** | 0.10** | 0.12** | 0.13*** | 0.14*** |
| | (5.85) | (6.17) | (4.34) | (2.54) | (3.21) | (2.91) |
| Ln scale | | 0.09*** | | -0.06 | 0.03 | |
| | | (2.86) | | (1.03) | (0.46) | |
| Ln FDI | | | -0.01* | -0.02* | | |
| | | | (1.74) | (1.97) | | |
| Industry fixed effect | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.28 | 0.32 | 0.18 | 0.19 | 0.15 | 0.15 |
| N | 144 | 144 | 90 | 90 | 90 | 90 |

Dependent variable: share of Budapest. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. Absolute value of t statistics in parentheses.

Table 3.8: Lagged trade flows and the employment share of Budapest

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Ln exports$_{t-1}$ | -0.04*** | -0.04*** | -0.03 | -0.03 |
| | (3.92) | (3.76) | (1.05) | (1.08) |
| Ln imports of inputs$_{t-1}$ | 0.15*** | 0.14*** | 0.10** | 0.07 |
| | (8.29) | (6.74) | (2.42) | (1.33) |
| Ln scale$_{t-1}$ | | 0.06** | | -0.07 |
| | | (2.18) | | (1.04) |
| Ln FDI$_{t-1}$ | | | -0.00 | 0.00 |
| | | | (0.14) | (0.17) |
| Industry fixed effect | Yes | Yes | Yes | Yes |
| $R^2$ | 0.41 | 0.44 | 0.14 | 0.16 |
| N | 126 | 126 | 72 | 72 |

Dependent variable: share of Budapest. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. Absolute value of t statistics in parentheses.

Table 3.9: High vs. low Budapest share industries

| | **High Budapest share** | | **Low Budapest share** | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Ln exports | -0.07** | 0.02 | -0.01 | -0.12** |
| | (2.18) | (0.55) | (0.56) | (2.59) |
| Ln imports of inputs | 0.25*** | -0.02 | 0.02 | 0.23** |
| | (4.77) | (0.18) | (0.51) | (2.86) |
| Ln scale | | -0.19 | | 0.01 |
| | | (1.38) | | (0.15) |
| Ln FDI | | -0.05 | | 0.01 |
| | | (1.06) | | (0.79) |
| Industry fixed effect | Yes | Yes | Yes | Yes |
| $R^2$ | 0.47 | 0.44 | 0.01 | 0.47 |
| N | 40 | 25 | 40 | 25 |

Dependent variable: share of Budapest. * significant at 10% level, ** significant at 5% level, *** significant at 1% level. Absolute value of t statistics in parentheses.

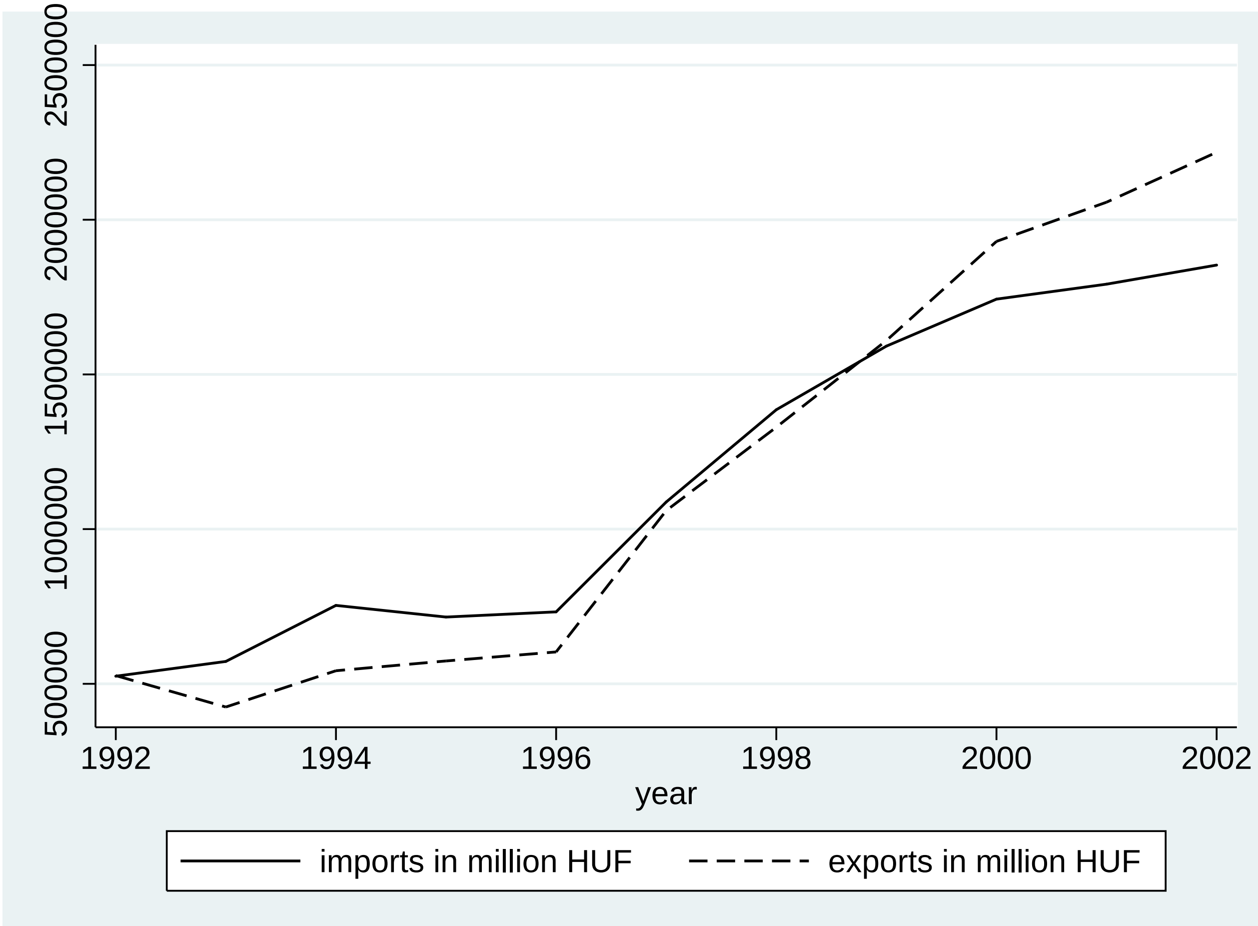


Figure 3.1: Total trade flows with the EU-15 in million HUF (base 1992), 1992-2002

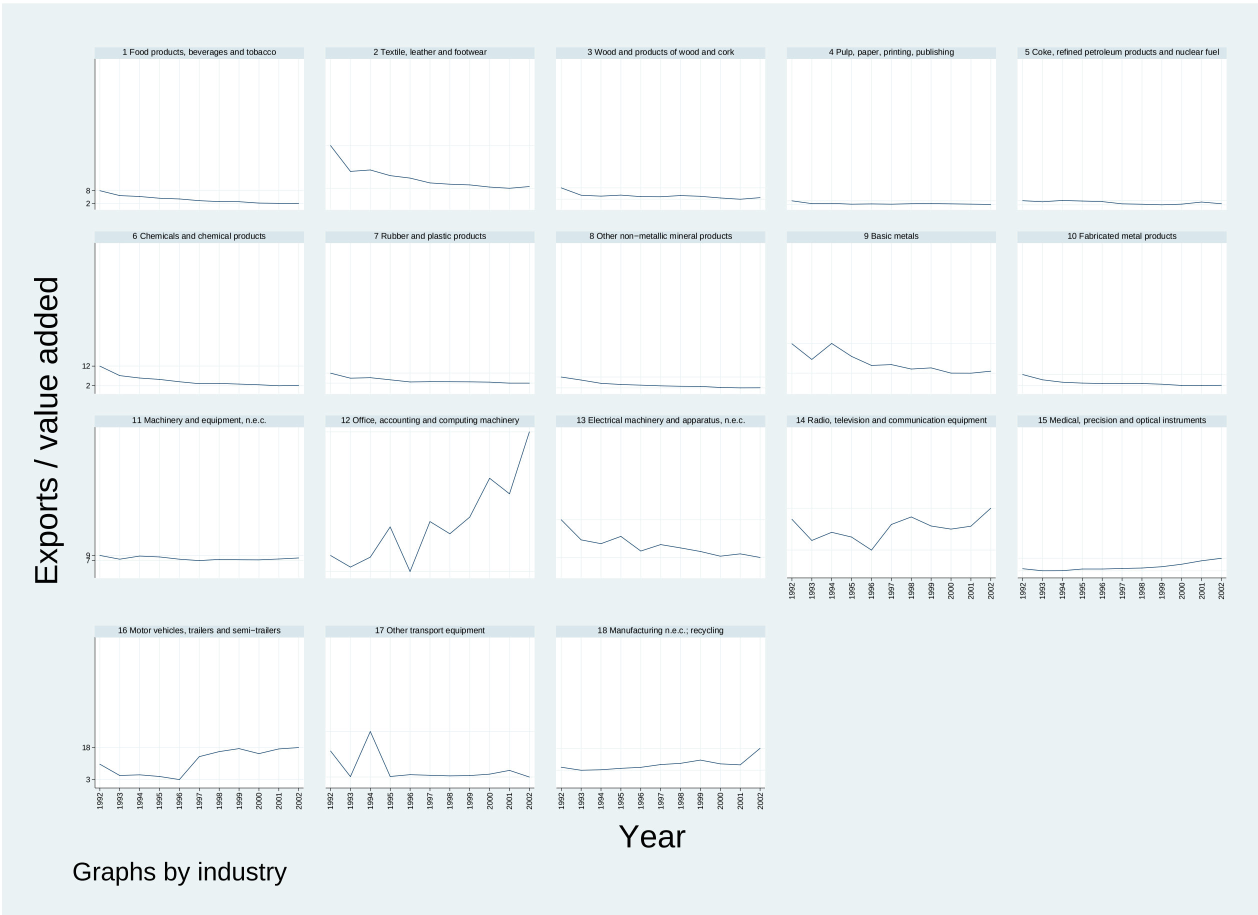


Figure 3.2: Exports by industry deflated by value added, 1992-2002

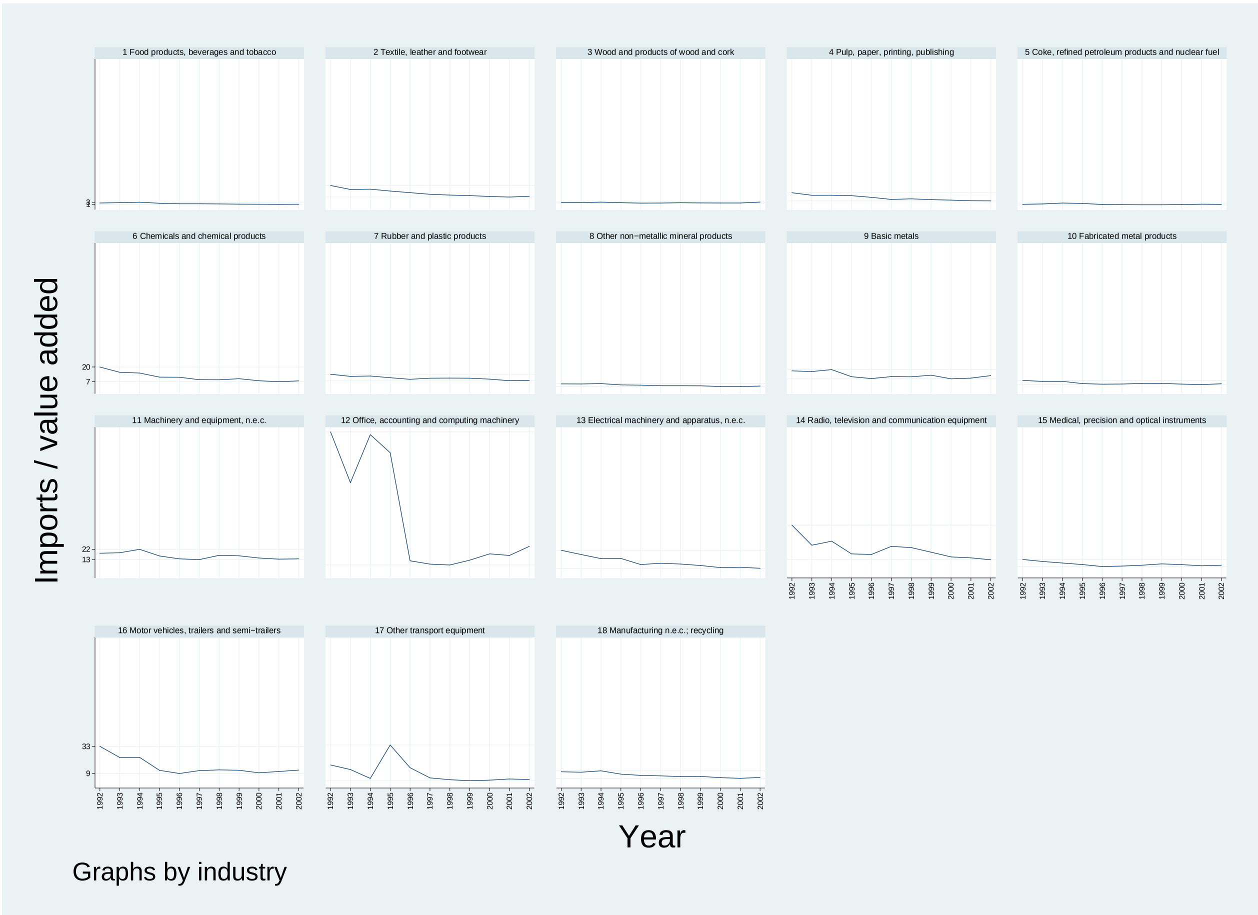


Figure 3.3: Imports by industry deflated by value added, 1992-2002

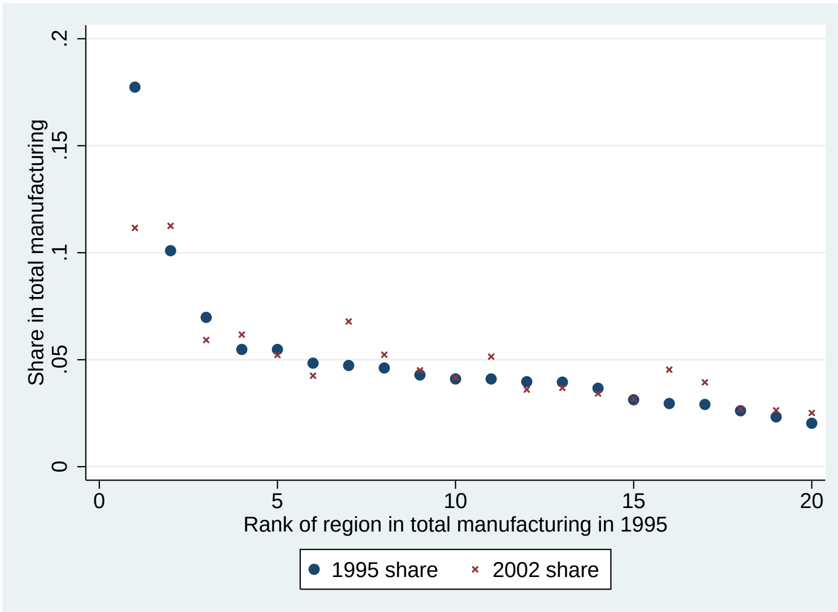


Figure 3.4: Share of each region in total manufacturing, 2002 v. 1995

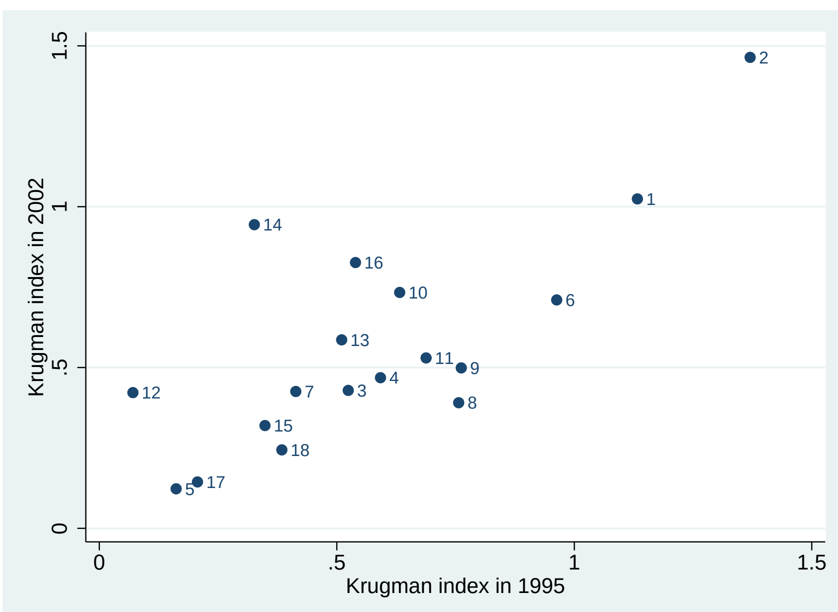


Figure 3.5: Krugman concentration index, 2002 vs. 1995

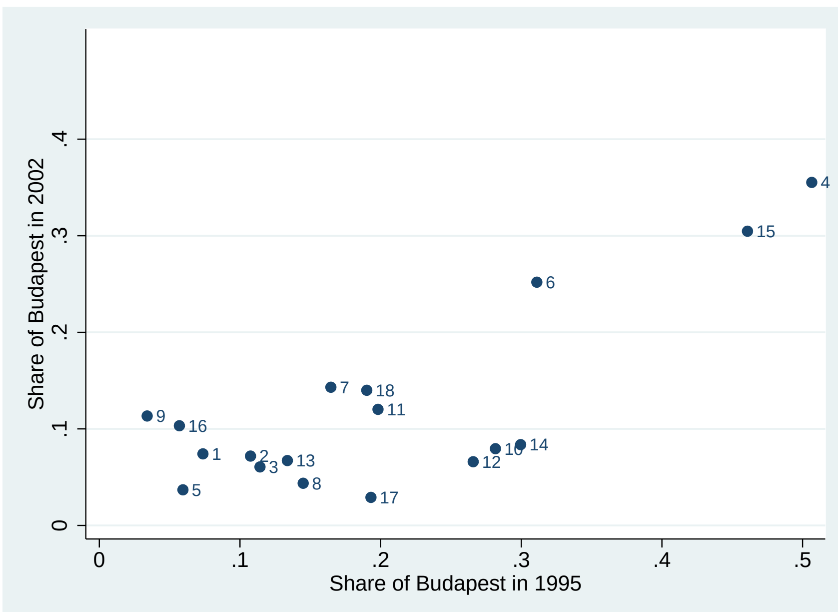


Figure 3.6: Share of Budapest county in each industry, 2002 vs. 1995

## CHAPTER IV

# Trade liberalization and the size distribution of plants in Colombia

### 4.1 Introduction

In this paper I ask how trade liberalization affects the size distributions of exporters and non-exporters. I rely on various predictions that can be extracted from heterogeneous-firms trade models and confront them with actual distributions derived from Colombian data before and after policy changes. This can explain whether a policy that favors exporters has generated an export boom where the largest and most productive incumbent exporters grow, or one where the bulk of the increase in exports comes from massive entry of smaller plants into the export market. Or given that most plants do not export, we may want to know how that population has evolved: in particular what happens to the smallest and the largest non-exporters?

So far the literature on the distribution of firm size has focused on establishing stylized facts about the size distribution of firms, irrespective of their exporter status. Although an implication of Gibrat's Law is that the distribution of firm size is lognormal,[1] later models, including Simon's model,[2] and more recent empirical evidence (such as Machado and Mata (2000), using a comprehensive dataset of all Portuguese firms) show that this distribution is in fact more right skewed than the lognormal and that there does not exist a single class of distribution that fits the data across countries and industries. Although the empirical literature so far is blind to the export status of firms, we might however expect that the subsample of firms (or plants) that engage in exports differs in its size distribution from the subsample of non-exporters. This is because exporters differ from

[1] See Gibrat (1931).

[2] See Ijiri and Simon (1977), and for a review of the literature on the size distribution of firms, Sutton (1997). Papers modeling the size distribution of firms and firm dynamics in the absence of international trade include Luttmer (2004), Luttmer (2007) and Rossi-Hansberg and Wright (2007).

non-exporters in their characteristics, notably in terms of their age, size, and the industries they belong to: it has been established that exporters tend to be larger, older and more productive than non-exporters.[3]

I focus on analyzing the evolution of the size distribution of Colombian manufacturing plants between 1981 and 1990, in light of two heterogeneous firms models. Between 1982 and 1986, exporting was facilitated by a depreciation of the domestic currency. This channel is similar to that operating in the Melitz (2003) heterogeneous-firms model: the new export possibilities increase production needs and the cost of labor, with implications for the size of surviving firms and of those who choose to export and not to export. Simultaneously there were tariff increases, which reduced the level of competition in the domestic market. Secondly, starting in 1985, there were also large unilateral tariff cuts which considerably reduced the initial differences in tariffs across sectors. This increased import competition in the sectors affected by tariff reductions, a channel explored in the Melitz and Ottaviano (2008) model.

The existing empirical literature examining the impact of international trade liberalization on firm size has used mostly parametric estimation methods, and focused on average effects or on effects on specific quantiles of the distribution.[4] Moreover, none of the previous work seems to have compared the impact on exporting vs. that on non-exporting firms or plants. To summarize the empirical evidence, there is ample evidence that trade liberalization increases exit (Head and Ries (1999), Eslava, Haltiwanger, Kugler, and Kugler (2008)), however there is mixed evidence regarding the impact on plant size. The impact on plant size might depend on the kind of liberalization observed (unilateral or bilateral tariff cuts, devaluation of the currency), and on the proportion of plants engaged in or affected by international trade. Most of the evidence shows that greater import competition reduces average plant size (Tybout, de Melo, and Corbo (1991), Roberts and Tybout (1991), using panel data from the manufacturing censuses for Colombia and Chile, and Head and Ries (1999), studying Canadian firms following CUSFTA), although the evidence on small plants is mixed (Head and Ries (1999)), while greater export possibilities tend to increase average plant size (Roberts and Tybout (1991)). The former finding is contrary to the predictions in the Melitz and

[3]See Bernard and Jensen (1999). Roberts and Tybout (1997) and Bernard and Jensen (2004) also analyze the determinants of exporter status. Eaton, Kortum, and Kramarz (2008), using French data, and Eaton, Eslava, Kugler, and Tybout (2008), using Colombian data study the behavior of exporters in terms of their participation in individual export markets.

[4]See Tybout and Westbrook (1995), Pavcnik (2002), Eslava, Haltiwanger, Kugler, and Kugler (2004), Bernard, Jensen, and Schott (2006) and Fernandes (2007) on the impact of trade liberalization on firm productivity. Taking a different approach, Di Giovanni and Levchenko (2009a) consider the implications of the size distribution of firms for the impact of lower entry costs and trade opening on welfare, and Di Giovanni and Levchenko (2009b) model the impact of international trade on macroeconomic volatility through idiosyncratic shocks to large firms in economies where the distribution of firm size is dominated by a few large firms.

Ottaviano (2008) model of heterogeneous-firms that average firm size should increase in the short run, but in accordance with the long-run prediction of a decrease in average firm size. The latter finding is consistent with the predictions from the Melitz (2003) model.

Directly focusing on the impact of trade on the overall distribution of firm size, Nocke and Yeaple (2006) show that in U.S. data, a reduction in shipping costs has flattened the distribution of firm size within an industry. They summarize the distribution of firm size by the gradient of the logarithm of firm size with respect to the logarithm of its sales rank: this gives information on the dispersion of firm size but not on potential leftward or rightward shifts. They do not study the impact on firms by exporter status.

To the best of my knowledge, the only paper that examines the impact of international trade on the distribution of firm size using non-parametric methods is Machado and Mata (2000). Using quantile regression, they find that high industry-level import intensity tends to shift the conditional firm size distribution leftward (i.e. to make firms smaller in all centiles of the distribution), whereas export intensity tends to shift it to the right (i.e. to make firms larger). This is consistent with the findings in Roberts and Tybout (1991).

Finally, using the same data for Colombia that are used in this paper and in Roberts and Tybout (1991), and analyzing the export boom between 1984 and 1991, Roberts, Sullivan, and Tybout (1997) find that more than half of the export growth in volume came from net entry into exporting. Incumbent exporters did not increase their export volumes much in response to devaluation. These results, combined with those of Roberts and Tybout (1991), indicate that there would have been a shift in the size distribution of exporters but also a change in its shape. These distributional effects can be expected to be long lasting due to the phenomenon of "export hysteresis" that has been documented in Colombia.[5]

The main results of this paper are as follows. I first find that the size distribution of exporters is lognormal, and to the right of that of non-exporters, whereas that of non-exporters is more right skewed than the lognormal. I then identify a distinct leftward shift in the size distribution of exporters during a period of currency depreciation and unilateral tariff increases (period 1), which remains later while the currency is stable and tariffs are cut drastically (period 2). In contrast, the size distribution of non-exporters does not shift but becomes more peaked with shorter tails in period 1, then becomes less peaked again during period 2. This is evidence in favor of predictions derived from both heterogeneous firms models: Melitz (2003) if there is massive entry of smaller

[5] See Das, Roberts, and Tybout (2007), who estimate entry costs into exporting to be substantial, meaning that firms may continue to export even though their current profits may be negative.

plants into exporting, and the long run predictions of Melitz and Ottaviano (2008). I then show that the bulk of the leftward shift in the size distribution of exporters is due to massive entry of smaller plants into exporting, which is consistent with both models. Overall most of the results support the predictions from Melitz (2003) as well as the long-run predictions of Melitz and Ottaviano (2008).

The next section summarizes the predictions that can be derived from two models of international trade with heterogeneous-firms, and section 3 describes the data and the policy changes that occurred during the period of observation. Section 4 tests the predictions of the models in terms of plant dynamics and of changes in the size distributions, and section 5 refines the results obtained in section 4 by conducting the analysis for separate groups of industries depending on industry-level characteristics. Section 6 concludes.

## 4.2 Theoretical predictions

In order to obtain a non-degenerate size distribution, it is necessary to depart from the representative firm framework and consider heterogeneous-firms. In the typical heterogeneous-firms trade models, firms are ordered according to their productivity level, there is a monotonically increasing relationship between firm size and productivity, and cut-off levels of productivity for firms to be able to operate and to export. A decrease in the variable trade cost causes the more productive firms (and the largest) to become or remain exporters. This increases aggregate productivity (and average firm size), firm exit, the number of exporters, and export sales of existing exporters. The result is that the smallest firms are the most hit by trade liberalization and the first ones to exit, while the largest firms are able to compensate their losses on the domestic market through export sales.

In Melitz (2003), the channel through which trade affects firm dynamics, and firm size in particular, is competition on the factor market. Due to facilitated exports, brought about by lower variable trade costs or in the particular case studied in this paper the depreciation of the Peso, the new production needs bid up the cost of labor, and the lowest productivity firms are forced to exit. This model does not take into account the import competition channel brought about by import tariff reductions. The Peso depreciation increases the productivity cut-off for operating and decreases the productivity cut-off for exporting. This forces the smallest firms to shut down, and enables more (and smaller) firms to become exporters. Domestic revenue shrinks for all firms, therefore non-exporters become smaller. However it is also shown that the combined domestic and export sales of exporters increase with trade liberalization, so exporters (new and incumbent) become larger at all points of the productivity distribution.

What are the implications of Melitz (2003) in terms of the overall distribution of firms? The total number of firms decreases. The smallest firms exit, the surviving small (domestic) firms shrink, and the largest (exporting) firms become larger. If the surviving domestic firms represent the bulk of the distribution (as only a small fraction of firms export) we can expect a leftward shift in the distribution (although the right tail becomes longer). If however there is high exit and massive entry into the export market, the number of non-exporters compared to that of exporters is greatly reduced, and there will be a rightward shift in the distribution except for the left tail. In either case, both tails of the distribution become thicker, the right tail becoming longer, and the distribution should be flatter.

The distribution of non-exporters should shift left. Since both tails drop out of the distribution (the left tail exits and the right tail become exporters), there should be more mass in the center.

Now let us consider the population of exporters. Since the incumbent exporters grow, we expect the right tail to become longer. However the new exporters are smaller than incumbent exporters, which should also make the left tail longer. The distribution of exporters should therefore have greater variance and become flatter. Again, the extent of entry of smaller exporters will determine the extent of flattening, as well as potential shifts (massive entry leading to more flattening and a leftward shift).

In Melitz and Ottaviano (2008), there is also firm heterogeneity in the form of productivity differences, although these are modeled as production cost differences, where production costs are an inverse function of productivity. The demand system is different: consumers have linear preferences as in Ottaviano, Tabuchi, and Thisse (2002) instead of constant elasticity of substitution preferences as in Melitz (2003). This framework allows for horizontal product differentiation and endogenous markups. This results in an endogenous distribution of markups which responds to trade liberalization, because trade liberalization changes the level of competition through the number of firms operating and their productivity levels. In this model trade liberalization therefore affects firm size through a different channel, that of increased competition on the product market. In contrast, in Melitz (2003) the firm dynamics respond to the access to a larger market and the increase in domestic factor costs. The Melitz and Ottaviano (2008) model leads to a similarly tractable pattern of production whereby only firms with cost draws below the domestic cost cut-off will operate.

Although Melitz and Ottaviano (2008) investigates several forms of trade liberalization, I only use the predictions from the case of a unilateral liberalization, where the domestic country unilaterally reduces its import tariffs. The model yields distinct predictions in the short run and the long run. In the short run, the number of firms is fixed in the home and the foreign country, while this is no longer the case in the long run. In the short run, increased import competition in the liberalizing country increases exit of the smallest domestic producers, so the number of firms decreases and average firm size increases. The variance of firm size decreases. Since competition in the non-liberalizing country is unchanged, the cut-off production cost for exporting in the liberalizing country is unchanged, and the number of exporters and their export sales are also unchanged (so entry into exporting is not a consideration of this model in the short run). The domestic sales of firms decrease, and therefore all non exiting firms (exporters and non-exporters) decrease in size. This effect is dominated by the exit of small firms and overall the average firm size increases.

In terms of short-run changes in the size distribution of firms, this implies that the distribution of all firms loses its left tail, has lower variance and shifts left; the size distribution of non-exporters exhibits the same patterns; and the distribution of exporters shifts left but has an unchanged variance.

The long-run predictions are different. Since now firms have an incentive to locate in the non-liberalizing country (in order to escape high competition in the domestic country and to benefit from low import tariffs into the domestic country), there is less competition in the liberalizing country and the cost cut-off for producing increases. There is greater entry of small firms, the number of firms operating increases and average size decreases. The variance of firm size increases. On the other hand competition in the non-liberalizing country increases (due to relocation of firms from the liberalizing country) so it becomes more difficult to be an exporter in the liberalizing country. The cost cut-off for exporting decreases, the number of exporters thus decreases in the long-run and the smaller exporters are the ones that exit the export market. Export sales of exporters decrease (at all levels of the cost distribution), and domestic sales of all firms increase. The increase in domestic sales outweighs the decrease in export sales of exporters.

So in the long run, a decrease in trade costs causes the size distribution of all firms to have a longer left tail and greater variance. The distribution of non-exporters has longer tails and greater variance, and shifts right. That of exporters loses its left tail and shifts right.

Table 4.1 summarizes the predicted effects of trade liberalization on firm dynamics according to Melitz (2003) and Melitz and Ottaviano (2008). Melitz (2003) predicts a shift in production from small to large firms and from non-exporters to exporters. In Melitz and Ottaviano (2008), the effects are pro-competitive in the short run, with greater exit and all firms decreasing in size. In the long run, competition decreases and all firms become larger.

Table 4.2 summarizes the effects on the size distribution of exporters and non-exporters. Melitz

(2003) predicts that the distribution of exporters becomes flatter and shifts left if there is massive entry into exporting, while that of non-exporters also shifts left but becomes more peaked. Melitz and Ottaviano (2008) predicts a leftward shift for both exporters and non-exporters in the short run, and a rightward shift in the long run.[6]

## 4.3 Data description and the environment

There were two main trade and other policy changes in Colombia that could have affected plant size between 1981 and 1990. Firstly, after having appreciated between 1975 and 1981, the Peso underwent a real depreciation between 1982 and 1986, then remained stable. The PPP conversion factor to exchange rate ratio published by the World Bank, which gives the number of US dollars required to buy a bundle of goods in Colombia as compared to the US, was 0.48 in 1981 and 1982, then began to decline sharply until 1986 (0.33), then slowly declined to 0.28 in 1990. As a result, the terms of trade became more favorable to exporters from 1982 onwards.[7] Second, there were major changes in the import tariff structure. The early 1980s were a period of tariff increases, in response to the appreciation of the Peso. Protection levels varied considerably across industries. When large unilateral tariff cuts started in 1985, the differences in tariffs across sectors were also lowered: There was therefore considerable sectoral variation in tariff cuts. To sum up, the first period (1981-1985) is characterized by a real depreciation of the Peso and an increase in tariffs, while the second period (1986-1990) is characterized by a reduction in tariffs.

I use a dataset of plants from the Colombian manufacturing survey, which includes all manufacturing plants with at least 10 employees. Ten years are covered, between 1981 and 1990. I end the dataset in 1990 because a vast labor market reform that significantly reduced the cost of firing was introduced in 1990, and might be confounded with the impact of trade liberalization on plant employment. Eslava, Haltiwanger, Kugler, and Kugler (2009) show that this reform lead to strong plant-level labor adjustments.

I follow the distributions in 1981, 1986 and 1990, because the real depreciation of the Peso began in 1982 and ended in 1986, and tariff cuts began in 1985. Following the densities every year shows that 1986 is a turning point in their evolution. After cleaning the data there are 6,589 plants in

[6] There exist other heterogeneous-firms models that yield predictions on the effects of trade liberalization on the size distribution of all firms, such as Bernard, Redding, and Schott (2007) and Nocke and Yeaple (2006). I do not test the predictions of Bernard, Redding, and Schott (2007) as this two-industry model is difficult to transpose empirically into a multi-industry setting. Nocke and Yeaple (2006) focuses on the impact of trade costs on the skewness of the overall distribution of firms. Changes in trade costs affect the skewness of the size distribution through the sale of product lines between small and large firms. Since in this model all firms export, this framework is not appropriate to derive predictions on the size distribution by exporter status.

[7] In addition, export subsidies increased between 1983 and 1984, before stopping as the terms of trade were more favorable to exporters.

1981, 6,532 in 1986 and 7,256 in 1990. There are still 12% of observations in 1981, 1.2% in 1986 and 9.5% in 1990 with employment less than 10. Their presence is likely to be due to sampling anomalies and changes in the sampling frame in certain years, when plants with less than 10 employees were sampled. For the purpose of this paper, since I focus on the years 1981, 1986 and 1990, for which there are a fair number of plants with less than 10 employees, I choose to keep all the data available, bearing in mind that the plants with less than 10 employees might be slightly under represented. The main variables of interest are plant employment, export sales, industry and age.

The data span 27 industries at the 3-digit SIC level. Table 4.3 gives the breakdown of the share of plants and their average size by industry. The first columns for each year indicate that the industrial composition of plants has remained stable over time. Labor shares (column 2 for each year) also indicate that there has been hardly any labor reallocation across industries between 1981 and 1990.[8] The third columns indicate great heterogeneity in average scale across industries.

Now let us consider the sample of exporters. Unlike the evidence in Table 4.3 against changes in the industrial composition of plants overall, it is clear from Table 4.4 that there have been changes in the industrial composition of exporters: The second columns for each year indicate that the shares of each industry in the total number of exporters have changed over time. Column 1 shows the heterogeneity in the proportion of exporters: For example 39% of the plants in industry 351 (engines and turbines) were exporters in 1981, while 313 (boot and shoe cut stock and findings) had only 3% of their plants exporting. There was also a lot of heterogeneity in the trajectories of individual industries, some increasing in their proportion of exporters and others decreasing. Finally, the labor shares indicate that there has been little labor reallocation across industries, as in the whole sample.

## 4.4 The distribution of plant size

In this section I first check whether the basic predictions in terms of plant dynamics laid out in Table 4.1 are confirmed in the data. I then turn to testing the predictions in terms of size distributions. The data reveal a distinct leftward shift in the size distribution of exporters. Finally, I analyze which of the roles of incumbents or net entry can best explain this distribution shift.

From the information in section 3, it is clear that several sometimes simultaneous policy changes occurred during the period of observation. In period 1 (1981-1985), the real depreciation that began in 1982 is a form of trade liberalization through facilitated exports according to the Melitz

[8] Attanasio, Goldberg, and Pavcnik (2004) come to the same conclusion.

(2003) model. However, using Melitz and Ottaviano (2008), this real depreciation is equivalent to a unilateral increase in protection, as it reduces competition from imported varieties. The simultaneous increase in tariffs is also a unilateral increase in protection according to Melitz and Ottaviano (2008), yielding the same predictions as the simultaneous real depreciation of the currency. Domestic tariff changes without foreign tariff changes are not interpretable using Melitz (2003).[9] In period 2 (1986-1990), the Melitz and Ottaviano (2008) model can predict the impact of trade liberalization in the form of a unilateral reduction in tariffs.

### 4.4.1 The impact of trade liberalization on plant dynamics

It can be inferred from Table 4.1 that, according to the Melitz (2003) model, we can expect to see the share of exporters and their size increase in period 1, while the size of non-exporters decreases. According to Melitz and Ottaviano (2008), in period 1 all plants should increase in size in the short run. In the long run all plants will shrink and the share of exporters will increase. In period 2, all plants should decrease in size and the share of exporters should increase in the short run. In the long run all plants should become larger and the share of exporters should decrease.

The final lines of Tables 4.3 and 4.4 confirm the long run predictions of Melitz and Ottaviano (2008) for period 1. In period 1, average plant size decreases overall, but more sharply for exporters during period 1.[10] The share of exporters increases. In period 2 however, average plant size is unchanged and exporter share continues to increase. This is consistent with the short run effects predicted by Melitz and Ottaviano (2008), or with a combination of the long run effects of the policy changes that occurred in period 1and the long run effects of those that occurred in period 2. However these are average effects and I now turn to a more detailed analysis in terms of the distribution of plant size.

### 4.4.2 The impact of trade liberalization on the size distributions

In this subsection, I first examine the shape of the distribution of plant size, overall and by exporter status. In particular I will check for lognormality. I then lay out the changes expected during each of the two periods of policy changes, as predicted by Melitz (2003) and Melitz and Ottaviano (2008). Finally, I check whether the expected changes are observed in the data by comparing the distributions before and after each period.

[9]Melitz (2003) only considers the case of a symmetric reduction in the variable trade cost, most probably for reasons of tractability.

[10]Tables 4.3 and 4.4 also show that on average exporters are larger and older than non-exporters.

#### 4.4.2.1 Testing for lognormality

In order to examine the size distribution of plants, and to be able to compare exporters and non-exporters in terms of class of distribution and also in terms of relative size, I use kernel density estimation. As discussed previously, earlier work in the literature on the distribution of firm size finds that lognormality, which follows from Gibrat's Law, does not hold for comprehensive datasets but that it holds for the subset of more mature firms (Cabral and Mata (2003)). As exporters differ from non-exporters in many ways, we can also expect the subset of exporters to not only be larger but also to follow a different distribution from that of non-exporters.

Figure 4.1 displays kernel density estimates of plant size distributions, in terms of employment, for the whole sample and that of exporters and non-exporters, in 1981.[11] The size distribution of the whole sample of plants is right skewed (dotted line), which corroborates the results obtained by Cabral and Mata (2003) using a comprehensive survey of all Portuguese firms. However, comparing the size density of exporters (dashed line) to that of non-exporters (solid line), we can see that the density of exporters is more symmetric (less skewed), to the right, with a greater mode. So exporters are larger and their size distribution appears lognormal.

In order to check formally if Gibrat's prediction of lognormality holds in this data, I use the skewness-kurtosis test of normality as well as plots comparing quantiles of the distribution of log employment with quantiles of the normal distribution. Table 4.5 reports the results of the skewness-kurtosis tests performed on ln plant employment, the null hypothesis being that ln employment is normally distributed.[12] N is the number of observations used in the test, and the joint p-value is a combined p-value from separate skewness and kurtosis tests. If the joint p-value is greater than 0.01, then the null hypothesis that the data is normally distributed cannot be rejected. The hypothesis that the distribution of ln employment is normal is rejected at the 0.01 level of significance for the whole sample and for the sample of non-exporters in 1981, 1986 and 1990. But for the sample of exporters, it is not rejected in 1981 and 1990.

The quantile plots compare ordered values of ln employment with quantiles of the normal distribution. If the two distributions match, the dots on figures 4.2 and 4.3 should replicate the 45-degree line. So Figure 4.2 shows that the distribution of the log of the size of exporters is very close to the normal distribution in all three years, while the quantile plots in Figure 4.3 indicate that the distribution of the log of the size of non-exporters differs significantly from the normal distribution.

[11] I report here the data for 1981. The relative positions and shapes of the distributions of plant size for exporters and non-exporters are similar in all years.

[12] I use the skewness-kurtosis test rather than other available tests for normality because of the large number of observations.

This paper therefore presents an interesting and novel result: the size distribution of exporters is lognormal,[13] while that of non-exporters is more right skewed than the lognormal. This means that exporters are neither a random sample of the overall population nor exactly the right tail of that distribution. In the heterogeneous firms models, the most productive and therefore the largest firms export (ie at a given point in time all the firms above a certain size export), while these data indicate that the larger the plant the more likely it is of being an exporter. This difference in distributions is not driven by the fact that exporters are older, as Appendix 4 indicates that the same patterns are observed when one scales plant size by plant age.

#### 4.4.2.2 Testing for the effects of trade liberalization

I now consider the size distribution of plants, by exporter status, before and after the policy changes, in each of the two periods. The predictions from theory follow from Table 4.2 and have been reorganized in Table 4.6 to lay out more clearly, period by period, the predictions given by each model, given the combination of policy changes occurring in each period. The two models yield different predictions in terms of shifts and changes in the tails of the distribution, and in terms of changes in variance, which depend on the extent of entry into exporting and on whether we consider short run or long run effects. For each period, I will first outline the expected changes in the distribution of plant size, as laid out in Table 4.6. Then, in order to identify which model best predicts the impact of policy changes on the size distributions, I use kernel density estimates of the size distributions of exporters and non-exporters, at the beginning and at the end of each period, as well as statistics reporting the evolution of plant size for several centiles of the distribution of plant size.

For period 1, between 1981 and 1986, the Melitz (2003) model predicts an increase in the variance of the size distribution of exporters and a leftward shift if there has been massive entry into exporting. The size distribution of non-exporters should also shift left but become more peaked. According to the Melitz and Ottaviano (2008) framework however, period 1 is a period of trade protection. This model therefore predicts a rightward shift for exporters in the short run and a leftward shift with a longer left tail in the long run. non-exporters also shift right in the short run and left in the long run, losing the left tail. Figure 4.4 shows kernel density estimates of the distributions of plant size for exporters and non-exporters. Focusing on the distributions of plant size in 1981 (solid line) and 1986 (dashed line), we see a distinct leftward shift in the size distribution of exporters, and the size distribution of non-exporters becomes more peaked. Now turning to the evidence in Table 4.7,

[13] To be more precise: the hypothesis that the distribution is lognormal cannot be rejected.

which reports plant sizes for several centiles of the size distribution in 1981, 1986 and 1990, we can see that the whole distribution of exporters has shifted left between 1981 and 1986, and that the variance has decreased. Table 4.7 also confirms that both tails of the distribution of non-exporters become shorter and the distribution is more peaked (as the standard deviation decreases). This is evidence for the long run predictions of Melitz and Ottaviano (2008), as well as for Melitz (2003) if there is massive entry into exporting and exit of the smallest non-exporters. These results do not suggest any evidence for the short run predictions of Melitz and Ottaviano (2008), indicating that Colombian plants may have had a high level of flexibility.

Now moving on to period 2, the short run effects predicted by Melitz and Ottaviano (2008) are a leftward shift for both exporters and non-exporters, with the size distribution of non-exporters losing its left tail and reducing in variance. In the long run, both distributions shift right, with that of exporters losing its left tail and that of non-exporters increasing in variance, with longer tails. However Figure 4.4 shows that between 1986 (dashed line) and 1990 (dotted line) the left tail of the size distribution of exporters becomes longer while the rest of the distribution is largely unchanged. This is confirmed by Table 4.7 where we can see that although the size of plants at every centile has become smaller between 1986 and 1990, plants in the first centile of the size distribution of exporters are 50% smaller in 1990 than in 1986, while plants in the $95^{th}$ to the $99^{th}$ centiles are less than 2% smaller. Now turning to non-exporters, Figure 4.4 shows that the left tail becomes thicker and longer, consistent with the fact that Table 4.7 reports a reduction in the size of plants at all centiles of the size distribution that is more pronounced at the lower centiles. Table 4.7 also reports a decrease in variance.

The observed patterns are compared to the corresponding predictions in Table 4.8. *Prima facie,* the patterns for period 2 partly corroborate the short run predictions of Melitz and Ottaviano (2008). The thickening of the left tail of the distribution of non-exporters between 1986 and 1990 is inconsistent with higher exit of the smallest plants.[14]

Also of high relevance is the fact that the patterns observed in period 2 must be analyzed taking into account the lingering effects of the policy changes that occurred in period 1. What is striking in Figure 4.4 is that although the size distribution of non-exporters seems to revert toward its initial (1981) position by the end of period 2, that of exporters does not. So the distribution of exporters

[14] One could argue that the reappearance of the left tail of the distribution after 1986 is merely due to a sample anomaly, as fewer small plants were sampled in 1986 than in 1981 or 1990. Although there have been temporary changes in the sampling frame that allowed more plants with less than 10 employees into the dataset in 1981 and 1990 than in 1986, this does not drive the results: The bottom panel of Table 4.7 shows that the size of plants in the first and fifth centiles of the size distribution (with plant sizes less than 10) follows the same evolution in time as the size of plants in the $10^{th}$ and $25^{th}$ centiles (with plant sizes greater than 10).

observed after period 2 could be the result of a combination of the long run effects of the policy changes that occurred in period 2 and of the lingering long run effects of the changes that occurred in period 1. In period 1, in both the Melitz (2003) and the Melitz and Ottaviano (2008) models, the leftward shift of exporters is caused by massive entry into exporting of smaller plants as the cutoffs for exporting become more favorable. I will look for evidence of the role of massive entry of smaller exporters in the next subsection. The long run effect predicted for period 2 by Melitz and Ottaviano (2008) is a rightward shift caused by the exit of exporters as the cost cutoff for exporting decreases. However, as mentioned earlier, there is evidence in the trade literature of export hysteresis, in particular in the case of Colombian plants. The fact that the size distribution of exporters does not shift back rightward in period 2 appears to be evidence of this export hysteresis, where once plants have started exporting, they are unlikely to exit the export market even if their profits become negative. Since the patterns observed in period 2 do not correspond clearly to a single prediction, it is also important to check if the results for period 2 become clearer when we consider separately industries with high and low tariff cuts. I will address this issue in the next section.

In terms of robustness, Appendix 4 deals with the fact that exporters are also older. Cabral and Mata (2003) had found that older firms have a size distribution that is more symmetric, to the right, and with a higher mode than younger firms. So it is not clear whether exporters are larger (and their size lognormally distributed) because they are older or whether they are older (i.e. they have survived longer) because they are larger. I therefore examine the distribution of plant employment scaled by plant age: Appendix 4 shows that the patterns are very similar to those identified using plant employment: exporters are still larger and the shape and shifts of the size distributions of exporters and non-exporters are the same as in the un-scaled case. Similarly, in Appendix 5, in order to account for the changes in the industrial composition of exporters observed in Table 4.4, I scale plant employment by mean industry employment and I find similar patterns.

### 4.4.3 The role of incumbents vs. that of net entry

I now focus on explaining the observed leftward shift in the size distribution of exporters observed during period 1. As discussed above, this observation is consistent with the predictions in Melitz and Ottaviano (2008) and in Melitz (2003) when there is massive entry of small exporters into exporting. To test whether it is indeed massive entry into exporting of smaller plants that has caused this leftward shift, I separate exporters into three categories, incumbents, entrants and exiters, and I assess their respective roles in the leftward shift of exporters using kernel density estimates.

The incumbents are plants that exported both in 1981 and 1986 (which I will call stayers), the exiters are plants that exported in 1981 but did not export in 1986 either because they stopped exporting or because they exited ( I call this category stoppers), and the entrants are plants that exported in 1986 but were not exporting in 1981 either because they were operating but non-exporters in 1981 or because they entered after 1981 (starters).

Using this categorization there can be three explanations for the leftward shift in the density of exporters. First, the stayers could have become smaller. Figure 4.5, shows that the distribution of stayers in 1986 is very similar to that in 1981, with a slight leftward shift. So this hypothesis can partly explain the leftward shift of exporters. Figure 4.5 also shows that plants that were exporting both in 1981 and in 1986 (dotted line) were initially larger than exporters overall (solid line). A second explanation would be that the stoppers were among the larger exporters in 1981. I can rule this out as Figure 4.6 clearly shows that the stoppers (dotted line) were smaller than exporters overall (solid line) in 1981.

The third possible explanation is that in 1981 the starters (i.e. future exporters) were smaller than the existing exporters. I examine this possibility using Figure 4.7. The 1981 size density of starters (dotted line) reports the data only for the plants that were operating but not exporting in 1981 (The proportion of starters that began operating after 1981 is 19.5%). The distribution is considerably more to the left than that of existing exporters (solid line), which means that in 1981 future exporters were smaller than current exporters. Moreover, the density of starters in 1981 (dotted line) is closer and more similar to that of exporters (solid line) than to that of all non-exporters (dashed line). This is evidence for the selection of future exporters among larger non-exporters, a common result of the heterogeneous-firms models. I now turn to the 19.5% of starters who began operating after 1981. Figure 4.8 reports the kernel density estimate of their size distribution in 1986, comparing it to the 1986 size distribution of all exporters. Quite logically, these new entrants are smaller than the whole group of exporters which includes older plants, and their entry into exporting therefore also explains the leftward shift in the size distribution of exporters. Finally, the extent of entry of small plants into exporting can be seen in Table 4.7: The number of exporters increased by 6% between 1981 and 1986, while the number of non-exporters decreased by 1.7%, and the size of plants in the lower centiles of the size distribution of exporters shrank.

The above results demonstrate that the leftward shift in the size distribution of exporters is due to some extent to the shrinking of stayers, and to a greater extent to the entry of smaller plants into exporting. The latter is evidence towards both heterogeneous-firms models. However the shrinking of stayers is evidence against the Melitz (2003) model, which predicts that incumbent exporters should grow during trade liberalization. It is evidence in favor of the Melitz and Ottaviano (2008) model, since in this framework period 1 is a period of trade protection during which exporters (as well as non-exporters) should become smaller in the long run.

## 4.5 Industry variation

The magnitude of the effects predicted by the two heterogeneous-firms models tested in this paper can vary across industries, along two dimensions. The first dimension is the export extensiveness of industries: industries that have a smaller proportion of exporting plants have a greater margin of entry into exporting, and we can expect that trade liberalization would induce a greater leftward shift of the size distribution of exporters in these industries. The second dimension is the degree of tariff changes: the effects predicted by the Melitz and Ottaviano (2008) model should be more pronounced for industries that had greater tariff cuts in period 2.

### 4.5.1 Variation according to export extensiveness

In order to test if the size distribution of plants in non export extensive industries shifts left to a greater extent than that of plants in export extensive industries, I compare kernel density estimates for the size distributions of exporters in each group of industries, at the beginning and at the end of period 1. The group of non export extensive industries is defined by industries that have a proportion of exporters in 1981 of 6% or less, as indicated in column 1 of Table 4.4. These are industries 311, 312, 313, 314, 324, 332, and 369, and this group has a plant count of 2047 in 1981. I define the group of export extensive industries by industries with a share of exporters of at least 24% in 1981, i.e. industries 323, 351, 361, 362, and 385, with a plant count of 682 in 1981.[15]

Figure 4.9 shows that although there is hardly any shift between 1981 and 1986 in the size distribution of exporters in export extensive industries, there is a large leftward shift in non export extensive industries. This strengthens the evidence in subsection 4.2 in favor of both heterogeneous-firms models.

### 4.5.2 Variation according to the degree of tariff changes

The second type of variation we can expect is across industries with differing tariff cuts. I define high tariff cut industries as having an overall tariff decrease of 46% or more between 1986 and 1990, and low tariff cut industries as having either no tariff decrease or a decrease of 9% or less. This

[15] The median industry has a share of exporters of 13%. The group of non export extensive industries represent the bottom 26% and the group of export extensive industries the top 19% of the distribution of exporter shares (as seen in column 1 of Table 4.4). Because of the skew in the distribution of exporter shares I choose not to include industries 382 and 383, with respective exporter shares of 20% and 18%, into the export extensive sample.

amounts to 649 high tariff cut plants and 1474 low tariff cut plants in 1986.[16] The list of industries belonging to each group and their associated tariff changes are reported in Appendix 6. I focus on the long run predictions of Melitz and Ottaviano (2008), since section 4 has shown little evidence that any short run effects can be observed, and on period 2, the period of tariff cuts. As described in Table 4.6, we expect to see the size distributions of exporters and non-exporters shift to the right to a greater extent for industries with high tariff cuts.

Figure 4.10 shows that the evolution of the size distributions of exporters and non-exporters are very different for the two industry groups. Although the size distribution of exporters has remained stable for low tariff cut industries, it has shifted right for high tariff cut industries, as predicted by Melitz and Ottaviano (2008). Also in accordance with the predictions, although the size distribution of non-exporters has shifted left for low tariff cut industries, it has somewhat shifted right for high tariff cut industries.

## 4.6 Conclusion

In this paper I have investigated the effects of trade liberalization on the size distributions of exporting and non exporting plants, using three cross sections of the annual survey of Colombian manufacturing plants. The data span two periods of policy changes. The first period, from 1981 to 1985 saw a real depreciation of the domestic currency and an increase in import tariffs, and the second period, from 1986 to 1990 saw a stabilization of the currency and decreases in import tariffs. I use two heterogeneous-firms trade models, Melitz (2003) and Melitz and Ottaviano (2008), to derive period-by-period predictions on the impact of these policy changes on the size distribution of exporters and non-exporters. I then use kernel density estimates as well as statistics of the plant size distribution to assess the differences in plant size distributions for exporters and for non-exporters before and after each period of policy changes.

The first result is a contemporaneous comparison of the size distributions of exporters and non-exporters. I find that the two distributions differ in their shape and in their relative positions. The size distribution of exporters is lognormal, and to the right of that of non-exporters. That of non-exporters is more right skewed than the lognormal. The second set of results concerns the plant dynamics that occur during trade liberalization and that underly the changes in the distribution of plant size. The patterns observed support the long run predictions of Melitz and Ottaviano (2008) in both periods. The third set of results reveal period by period the changes in the distribution of

[16]The median industry is industry 381 with a tariff cut of 29.5%. The groups of high and low tariff cut industries represent respectively the top and the bottom 19% of the distribution of tariff cuts.

exporters and non-exporters. There is a distinct leftward shift in the size distribution of exporters during period 1 (meaning that plants at all centiles of the distribution are smaller), followed by no shift but a longer left tail during period 2 (i.e. there are more small plants). The size distribution of non-exporters becomes more peaked with shorter tails in period 1 (i.e. there are fewer small and large plants), then becomes less peaked again during period 2. This is evidence in favor of the predictions of Melitz (2003) with massive entry of smaller exporters, and for the long run predictions of Melitz and Ottaviano (2008). In period 2 there is some indication that export hysteresis prevents the new entrants into exporting from period 1 from exiting the export market in period 2 (as should be the case according the the long run predictions of Melitz and Ottaviano (2008)). The fourth result is that the leftward shift in the size distribution of exporters is due to massive entry of smaller plants into exporting (which further confirms the results of both heterogeneous-firms models), and to a lesser extent to the downsizing of continuous exporters (which is consistent with the long run predictions of Melitz and Ottaviano (2008) and inconsistent with Melitz (2003)). Finally, in the fifth set of results I find variation in the strength of the results according to characteristics of industries. I first find stronger results in non export extensive industries, which strengthens the evidence on the importance of entry of smaller plants into exporting. I also find that comparing the results obtained from the subsamples of high and low tariff cut industries yields clearer results for period 2: the long run effects predicted by Melitz and Ottaviano (2008) can now clearly be seen for the subsample of high tariff cut industries, whereas there is no evidence of these effects in the subsample of low tariff cut industries.

This paper therefore contributes to the literature on the distribution of plant size by establishing that the size distribution of exporters is lognormal whereas that of non-exporters is more right skewed. It also contributes to the trade literature that estimates the effects of trade liberalization, by testing predictions derived from two heterogeneous-firms models regarding effects on the size distribution of plants by exporter status. The results also add to the wide literature on how exporting plants differ from non-exporters. Possible extensions of this work include exploring the plant dynamics that make the size distribution of exporters different from that of non-exporters, and discriminating between the effects of different policies on the size distributions.

**Tables and Figures**

Table 4.1: Summary of predictions on firm dynamics

| Effects of trade liberalization | Melitz (2003) | Melitz and Ottaviano (2008): unilateral liberalization case. |
|---|---|---|
| **Firm exit** | Increases | Short run (SR): increases. Long run (LR): decreases. |
| **Smallest firms** | Exit or shrink | SR: exit or shrink. LR: Get larger. |
| **Largest firms** | Get larger | SR: Shrink. LR: Get larger. |
| **Nbr of exporters** | Increases | SR: unchanged. LR: decreases. |
| **Export sales of incumbent exporters** | Increase | SR: unchanged. LR: decrease. |
| **Total sales of incumbent exporters** | Increase | SR: decrease. LR: increase. |

Table 4.2: Summary of predictions on the size distribution of exporters and non-exporters

| Effects of trade liberalization | Melitz (2003) | Melitz and Ottaviano (2008): unilateral liberalization case. |
|---|---|---|
| **Size distribution, exporters** | Flatter. Both tails become longer. If no massive entry into exporting: rightward shift. If massive entry: leftward shift. | SR: unchanged variance, shifts left. LR: left tail drops out, shifts right. |
| **Size distribution, non-exporters** | Leftward shift, more kurtosis, shorter tails. | SR: lower variance, shifts left, left tail drops out. LR: greater variance, longer tails, shifts right. |

Table 4.3: Summary statistics by industry

| | 1981 | | | | 1986 | | | | 1990 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Industry | Share of plants | Labor share | Mean employ-ment | Mean age | Share of plants | Labor share | Mean employ-ment | Mean age | Share of plants | Labor share | Mean employ-ment | Mean age |
| 311 | 0.15 | 0.13 | 64.77 | 15.98 | 0.15 | 0.14 | 63.19 | 17.11 | 0.15 | 0.14 | 63.09 | 18.38 |
| 312 | 0.03 | 0.02 | 51.10 | 15.61 | 0.03 | 0.02 | 55.64 | 17.30 | 0.03 | 0.02 | 47.48 | 19.03 |
| 313 | 0.02 | 0.06 | 210.87 | 25.94 | 0.02 | 0.05 | 189.28 | 28.98 | 0.02 | 0.05 | 176.05 | 30.52 |
| 314 | 0.00 | 0.00 | 102.60 | 23.40 | 0.00 | 0.00 | 122.67 | 33.56 | 0.00 | 0.00 | 109.13 | 23.63 |
| 321 | 0.07 | 0.13 | 134.55 | 12.99 | 0.07 | 0.10 | 101.45 | 14.29 | 0.07 | 0.11 | 112.15 | 16.39 |
| 322 | 0.13 | 0.10 | 58.68 | 9.05 | 0.16 | 0.11 | 45.71 | 9.16 | 0.14 | 0.10 | 48.03 | 11.05 |
| 323 | 0.02 | 0.02 | 73.09 | 13.45 | 0.01 | 0.01 | 66.72 | 13.78 | 0.02 | 0.02 | 72.62 | 14.81 |
| 324 | 0.03 | 0.02 | 47.86 | 10.28 | 0.04 | 0.03 | 45.93 | 9.70 | 0.04 | 0.03 | 57.52 | 11.60 |
| 331 | 0.03 | 0.01 | 32.34 | 14.19 | 0.02 | 0.01 | 36.29 | 15.41 | 0.02 | 0.01 | 36.64 | 16.48 |
| 332 | 0.03 | 0.01 | 35.60 | 10.59 | 0.03 | 0.02 | 39.32 | 10.62 | 0.03 | 0.02 | 38.09 | 12.13 |
| 341 | 0.02 | 0.02 | 81.37 | 16.50 | 0.02 | 0.02 | 77.26 | 17.55 | 0.02 | 0.02 | 80.27 | 19.67 |
| 342 | 0.05 | 0.04 | 57.00 | 16.79 | 0.05 | 0.05 | 57.77 | 17.11 | 0.05 | 0.05 | 60.66 | 19.49 |
| 351 | 0.02 | 0.03 | 143.46 | 16.52 | 0.02 | 0.03 | 138.65 | 17.04 | 0.02 | 0.03 | 121.44 | 20.55 |
| 352 | 0.04 | 0.05 | 84.21 | 20.90 | 0.05 | 0.05 | 79.90 | 21.24 | 0.04 | 0.05 | 81.09 | 22.47 |
| 355 | 0.01 | 0.02 | 108.35 | 15.16 | 0.01 | 0.02 | 92.10 | 14.76 | 0.01 | 0.01 | 88.69 | 19.03 |
| 356 | 0.04 | 0.03 | 62.92 | 10.17 | 0.04 | 0.04 | 61.38 | 11.18 | 0.05 | 0.04 | 51.42 | 12.72 |
| 361 | 0.01 | 0.01 | 128.52 | 13.43 | 0.00 | 0.01 | 158.43 | 18.93 | 0.00 | 0.01 | 204.23 | 23.58 |
| 362 | 0.01 | 0.01 | 134.92 | 12.55 | 0.01 | 0.01 | 98.05 | 13.20 | 0.01 | 0.01 | 95.30 | 15.61 |
| 369 | 0.05 | 0.04 | 65.09 | 15.59 | 0.04 | 0.05 | 73.40 | 17.84 | 0.04 | 0.04 | 67.36 | 19.64 |
| 371 | 0.01 | 0.03 | 265.91 | 16.22 | 0.01 | 0.02 | 154.97 | 13.49 | 0.01 | 0.02 | 140.04 | 15.64 |
| 372 | 0.01 | 0.00 | 71.97 | 16.97 | 0.00 | 0.00 | 68.26 | 17.19 | 0.00 | 0.00 | 64.91 | 19.32 |
| 381 | 0.09 | 0.06 | 52.09 | 13.42 | 0.08 | 0.06 | 52.08 | 14.22 | 0.08 | 0.06 | 48.61 | 16.37 |
| 382 | 0.05 | 0.03 | 44.60 | 15.15 | 0.05 | 0.03 | 47.05 | 16.65 | 0.05 | 0.03 | 47.72 | 17.98 |
| 383 | 0.03 | 0.04 | 88.08 | 13.94 | 0.03 | 0.04 | 80.79 | 16.24 | 0.03 | 0.04 | 86.94 | 18.28 |
| 384 | 0.03 | 0.04 | 99.61 | 13.63 | 0.03 | 0.04 | 83.47 | 15.57 | 0.03 | 0.04 | 75.04 | 16.75 |
| 385 | 0.01 | 0.01 | 51.82 | 15.56 | 0.01 | 0.01 | 49.23 | 16.40 | 0.01 | 0.01 | 54.77 | 19.53 |
| 390 | 0.02 | 0.02 | 47.45 | 14.70 | 0.02 | 0.02 | 54.11 | 18.22 | 0.02 | 0.02 | 50.64 | 18.11 |
| Total | 1.00 | 1.00 | 73.63 | 14.19 | 1.00 | 1.00 | 66.94 | 14.94 | 1.00 | 1.00 | 66.68 | 16.76 |

Table 4.4: Summary statistics for exporters by industry

| | 1981 | | | | | 1986 | | | | | 1990 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Industry** | **Exporter share** | **Share of plants** | **Labor share** | **Mean employment** | **Mean age** | **Exporter share** | **Share of plants** | **Labor share** | **Mean employment** | **Mean age** | **Exporter share** | **Share of plants** | **Labor share** | **Mean employment** | **Mean age** |
| 311 | 0.06 | 0.08 | 0.12 | 327.18 | 22.19 | 0.08 | 0.10 | 0.11 | 216.76 | 24.96 | 0.08 | 0.08 | 0.09 | 206.86 | 26.34 |
| 312 | 0.06 | 0.02 | 0.01 | 189.92 | 19.50 | 0.05 | 0.01 | 0.01 | 209.89 | 24.56 | 0.07 | 0.02 | 0.02 | 190.18 | 26.94 |
| 313 | 0.03 | 0.01 | 0.01 | 371.75 | 35.00 | 0.03 | 0.01 | 0.01 | 287.00 | 37.00 | 0.05 | 0.01 | 0.01 | 277.83 | 39.67 |
| 314 | 0.05 | 0.00 | 0.00 | 39.00 | 29.00 | 0.22 | 0.00 | 0.00 | 314.00 | 39.50 | 0.13 | 0.00 | 0.00 | 203.00 | 16.00 |
| 321 | 0.10 | 0.06 | 0.15 | 568.77 | 23.23 | 0.14 | 0.08 | 0.15 | 371.15 | 24.55 | 0.18 | 0.08 | 0.14 | 311.53 | 22.56 |
| 322 | 0.10 | 0.11 | 0.09 | 182.73 | 12.28 | 0.06 | 0.07 | 0.06 | 164.60 | 16.93 | 0.13 | 0.12 | 0.09 | 130.66 | 12.94 |
| 323 | 0.31 | 0.04 | 0.03 | 152.63 | 13.50 | 0.36 | 0.04 | 0.03 | 131.69 | 17.84 | 0.45 | 0.05 | 0.04 | 136.29 | 16.82 |
| 324 | 0.06 | 0.02 | 0.02 | 290.92 | 17.50 | 0.07 | 0.02 | 0.03 | 218.61 | 16.11 | 0.27 | 0.07 | 0.04 | 119.74 | 13.68 |
| 331 | 0.07 | 0.02 | 0.01 | 141.38 | 14.15 | 0.07 | 0.01 | 0.01 | 148.45 | 17.55 | 0.07 | 0.01 | 0.01 | 155.50 | 16.58 |
| 332 | 0.04 | 0.01 | 0.01 | 119.88 | 14.63 | 0.07 | 0.02 | 0.01 | 76.23 | 11.31 | 0.05 | 0.01 | 0.01 | 97.09 | 14.64 |
| 341 | 0.15 | 0.03 | 0.03 | 257.90 | 15.67 | 0.17 | 0.03 | 0.03 | 209.86 | 19.36 | 0.16 | 0.02 | 0.03 | 241.63 | 21.92 |
| 342 | 0.07 | 0.03 | 0.03 | 244.00 | 19.21 | 0.09 | 0.04 | 0.03 | 161.56 | 16.06 | 0.10 | 0.03 | 0.04 | 210.46 | 19.91 |
| 351 | 0.39 | 0.05 | 0.07 | 286.35 | 20.02 | 0.33 | 0.04 | 0.05 | 235.38 | 21.85 | 0.32 | 0.04 | 0.06 | 289.73 | 25.48 |
| 352 | 0.17 | 0.07 | 0.06 | 215.80 | 24.20 | 0.23 | 0.09 | 0.08 | 168.61 | 25.13 | 0.22 | 0.07 | 0.06 | 164.87 | 27.92 |
| 355 | 0.13 | 0.01 | 0.03 | 506.36 | 18.55 | 0.14 | 0.01 | 0.03 | 369.73 | 21.64 | 0.22 | 0.02 | 0.02 | 293.06 | 23.00 |
| 356 | 0.13 | 0.04 | 0.03 | 177.76 | 16.03 | 0.17 | 0.06 | 0.05 | 141.42 | 17.30 | 0.17 | 0.06 | 0.04 | 114.73 | 17.08 |
| 361 | 0.24 | 0.01 | 0.03 | 462.40 | 21.00 | 0.33 | 0.01 | 0.03 | 419.20 | 24.80 | 0.35 | 0.01 | 0.03 | 547.89 | 32.00 |
| 362 | 0.25 | 0.02 | 0.02 | 229.77 | 16.85 | 0.19 | 0.02 | 0.02 | 269.75 | 24.00 | 0.22 | 0.02 | 0.02 | 265.69 | 25.75 |
| 369 | 0.06 | 0.03 | 0.03 | 275.74 | 23.00 | 0.05 | 0.02 | 0.02 | 252.29 | 27.00 | 0.07 | 0.02 | 0.03 | 259.64 | 31.41 |
| 371 | 0.11 | 0.01 | 0.01 | 267.67 | 24.67 | 0.09 | 0.01 | 0.02 | 494.00 | 16.50 | 0.15 | 0.01 | 0.04 | 724.70 | 23.40 |
| 372 | 0.15 | 0.01 | 0.01 | 222.40 | 20.60 | 0.15 | 0.01 | 0.00 | 111.00 | 24.50 | 0.12 | 0.00 | 0.00 | 145.25 | 25.00 |
| 381 | 0.11 | 0.09 | 0.07 | 187.90 | 19.06 | 0.12 | 0.08 | 0.06 | 160.30 | 20.60 | 0.14 | 0.08 | 0.06 | 153.09 | 22.29 |
| 382 | 0.20 | 0.09 | 0.03 | 90.44 | 16.42 | 0.20 | 0.08 | 0.04 | 88.10 | 21.36 | 0.18 | 0.06 | 0.03 | 91.79 | 21.08 |
| 383 | 0.18 | 0.05 | 0.04 | 217.14 | 15.29 | 0.17 | 0.04 | 0.04 | 178.91 | 18.72 | 0.20 | 0.04 | 0.05 | 222.20 | 22.51 |
| 384 | 0.14 | 0.04 | 0.04 | 239.00 | 12.80 | 0.14 | 0.04 | 0.04 | 212.39 | 19.36 | 0.13 | 0.03 | 0.03 | 177.74 | 20.39 |
| 385 | 0.24 | 0.02 | 0.01 | 131.00 | 14.67 | 0.27 | 0.02 | 0.01 | 95.29 | 16.18 | 0.23 | 0.02 | 0.01 | 144.29 | 20.24 |
| 390 | 0.16 | 0.03 | 0.02 | 139.00 | 20.92 | 0.25 | 0.05 | 0.03 | 104.22 | 21.00 | 0.24 | 0.04 | 0.02 | 103.32 | 22.83 |
| Total | 0.11 | 1.00 | 1.00 | 237.99 | 18.17 | 0.12 | 1.00 | 1.00 | 191.56 | 20.93 | 0.15 | 1.00 | 1.00 | 185.93 | 20.99 |

Table 4.5: Skewness-kurtosis tests of normality on ln plant employment

| **Year** | **N** | **Joint Prob > $\chi^2$** |
|---|---|---|
| **Whole sample** | | |
| 1981 | 6589 | 0.0000 |
| 1986 | 6532 | 0.0000 |
| 1990 | 7256 | 0.0000 |
| **Exporters** | | |
| 1981 | 736 | 0.2923 |
| 1986 | 780 | 0.0000 |
| 1990 | 1058 | 0.0204 |
| **non-exporters** | | |
| 1981 | 5853 | 0.0000 |
| 1986 | 5752 | 0.0000 |
| 1990 | 6198 | 0.0000 |

Table 4.6: Summary of predictions, by period

| | **Melitz (2003)** | **Melitz and Ottaviano (2008): unilateral liberalization case.** |
|---|---|---|
| **Period 1: 1981-1986. Real depreciation and tariff increases.** | Exporters:<br><br>Greater variance, longer tails. Leftward shift if massive entry of small exporters, otherwise rightward shift.<br><br>non-exporters:<br>More peaked, shorter tails, leftward shift. | Exporters:<br><br>SR: rightward shift. LR: Longer left tail, leftward shift.<br><br>non-exporters:<br>SR: Longer left tail, rightward shift. LR: Left tail drops out, leftward shift. |
| **Period 2: 1986-1990. Tariff cuts.** | | Exporters:<br><br>SR: Unchanged variance, leftward shift. LR: Left tail drops out, rightward shift.<br><br>non-exporters:<br>SR: Lower variance, left tail drops out, leftward shift. LR: Greater variance, longer tails, rightward shift. |

Table 4.7: Distribution of plant employment - exporters vs. non-exporters

| **Year** | **N** | **Mean** | **Std. dev.** | **Skewness** | **Kurtosis** | **p1** | **p5** | **p10** | **p25** | **p50** | **p75** | **p90** | **p95** | **p99** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Exporters** | | | | | | | | | | | | | | |
| 1981 | 736 | 238 | 385 | 7.9 | 115 | 11 | 19 | 27 | 55 | 119 | 275 | 567 | 867 | 1513 |
| 1986 | 780 | 191 | 292 | 5.7 | 66 | 10 | 16 | 20 | 39 | 88.5 | 222 | 512.5 | 731 | 1283 |
| 1990 | 1058 | 186 | 300 | 5.8 | 61 | 5 | 11 | 17 | 37 | 87 | 208 | 428 | 715 | 1278 |
| **non-exporters** | | | | | | | | | | | | | | |
| 1981 | 5853 | 53 | 146 | 23.2 | 956 | 4 | 7 | 8 | 13 | 22 | 47 | 104 | 178 | 537 |
| 1986 | 5752 | 50 | 105 | 13 | 339 | 7 | 10 | 11 | 15 | 23 | 46 | 100 | 168 | 446 |
| 1990 | 6198 | 46 | 92 | 8.9 | 123 | 4 | 7 | 10 | 13 | 22 | 44 | 94 | 148 | 388 |

Table 4.8: Predictions vs. outcomes, by period

| | **Melitz (2003)** | **Melitz and Ottaviano (2008): unilateral liberalization case.** | **Observed outcome** |
|---|---|---|---|
| **Period 1: 1981-1986. Real depreciation and tariff increases.** | Exporters:<br><br>Greater variance, longer tails. Leftward shift. | Exporters:<br><br>SR: rightward shift. LR: Longer left tail, leftward shift. | Exporters:<br><br>Smaller variance, leftward shift. |
| | non-exporters:<br>More peaked, shorter tails, leftward shift. | non-exporters:<br>SR: Longer left tail, rightward shift. LR: Left tail drops out, leftward shift. | non-exporters:<br>More peaked, shorter tails, no shift. |
| **Period 2: 1986-1990. Tariff cuts.** | | Exporters:<br><br>SR: Unchanged variance, leftward shift. LR: Left tail drops out, rightward shift. | Exporters:<br><br>Unchanged variance, leftward shift, longer, thicker left tail. |
| | | non-exporters:<br>SR: Lower variance, left tail drops out, leftward shift. LR: Greater variance, longer tails, rightward shift. | non-exporters:<br>Lower variance, longer, thicker left tail. |

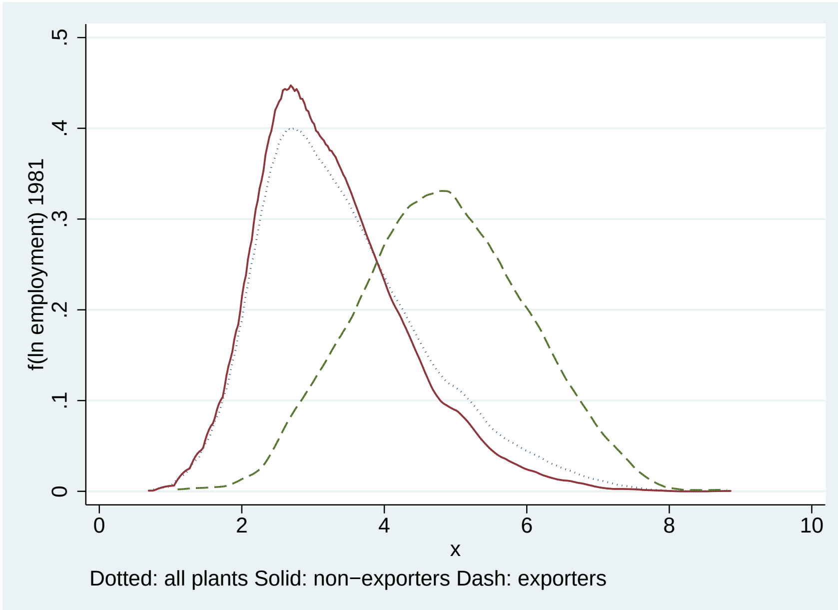


Figure 4.1: Distributions of plant employment by exporter status in 1981

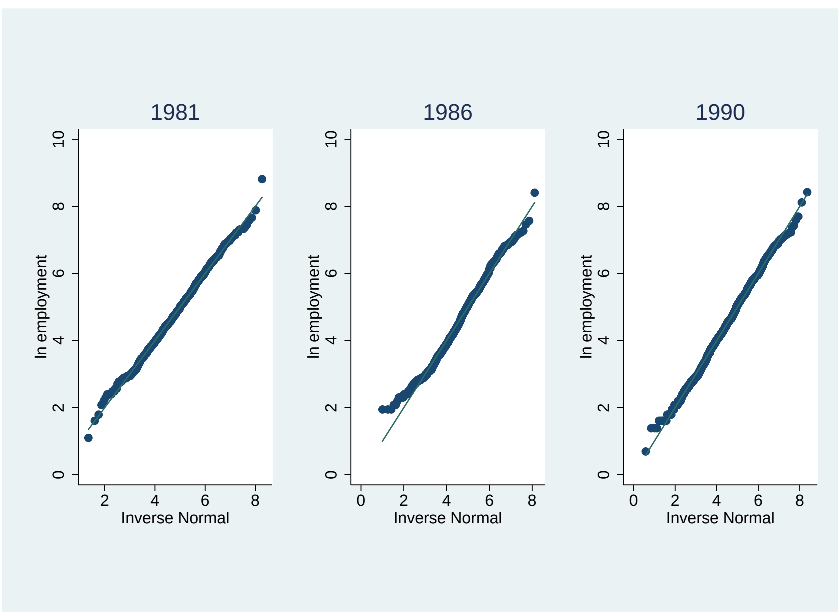


Figure 4.2: The distribution of ln employment vs.the normal distribution: exporters

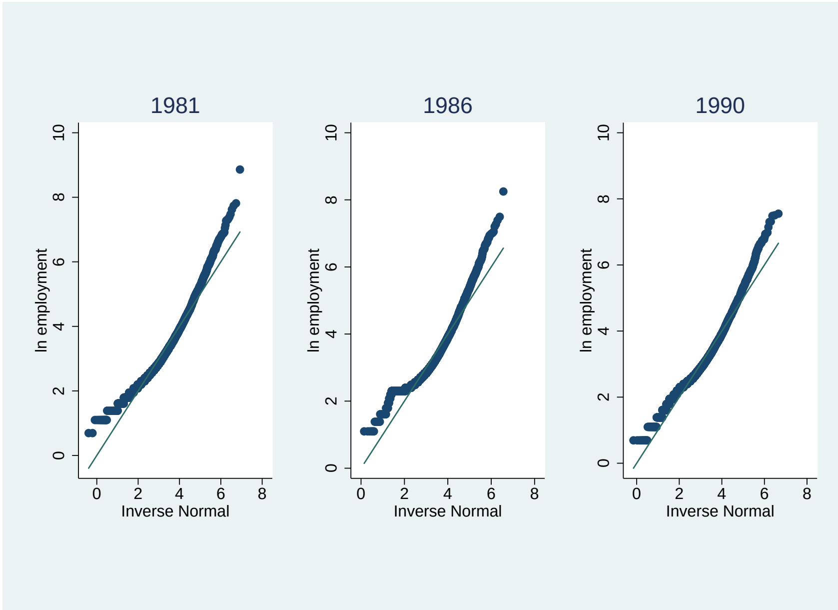


Figure 4.3: The distribution of ln employment vs.the normal distribution: non-exporters

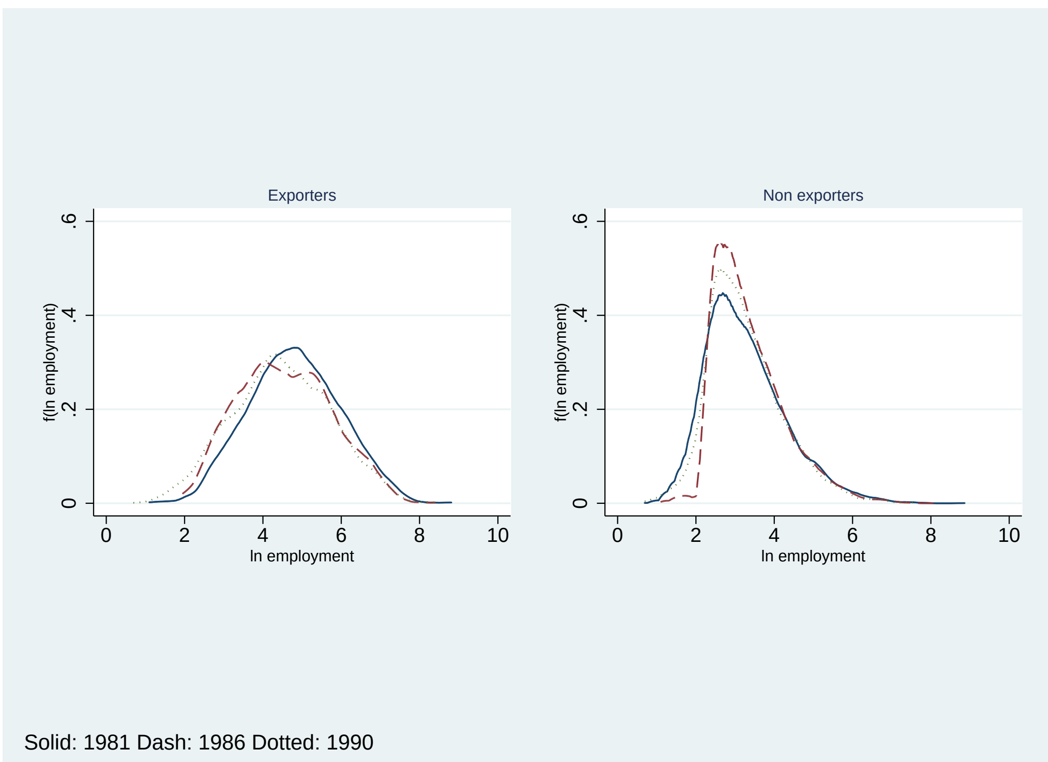


Figure 4.4: Distributions of plant employment over time, by exporter status

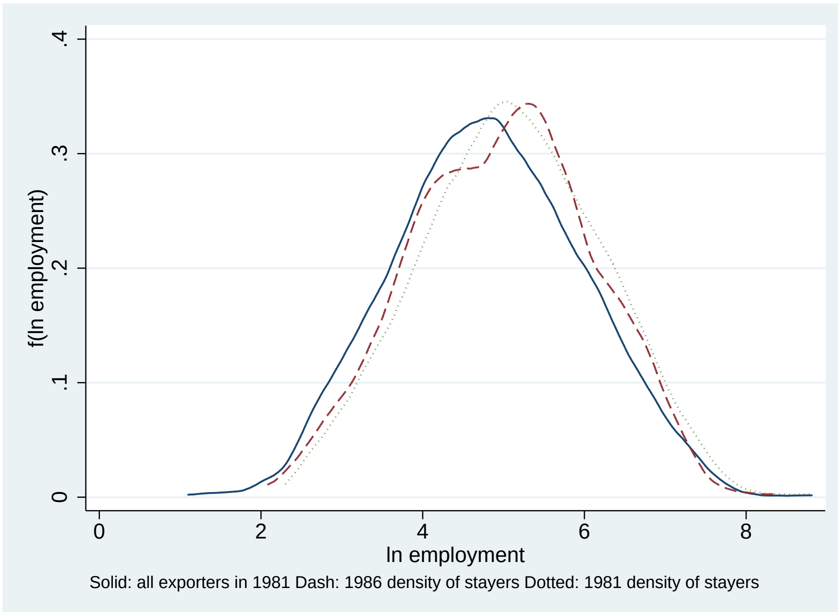


Figure 4.5: Stayers: plants exporting in 1981 and 1986

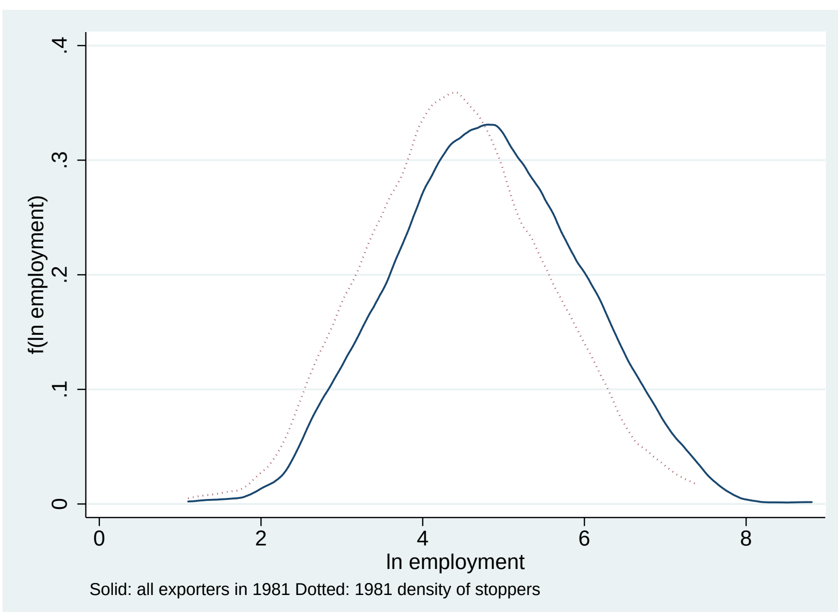


Figure 4.6: Stoppers: plants exporting in 1981 but not in 1986

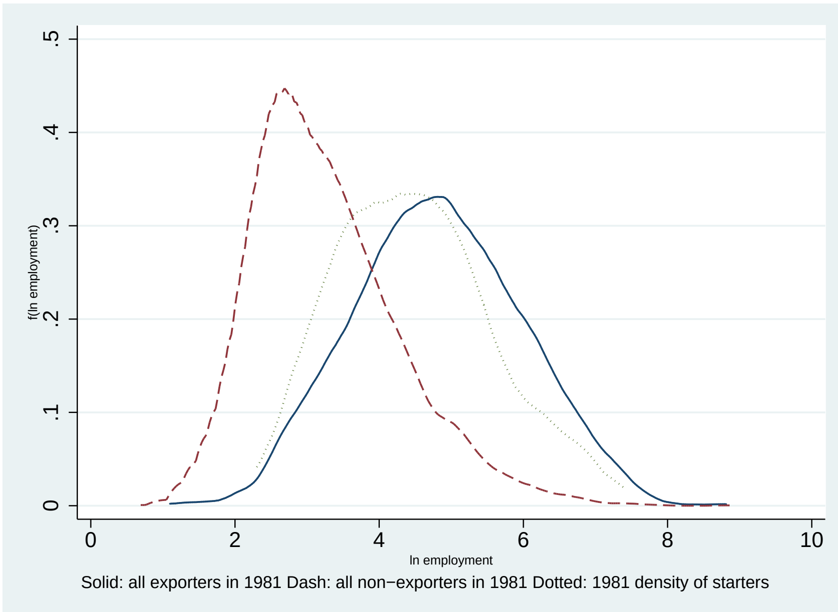


Figure 4.7: Starters: plants operating and not exporting in 1981, but exporting in 1986

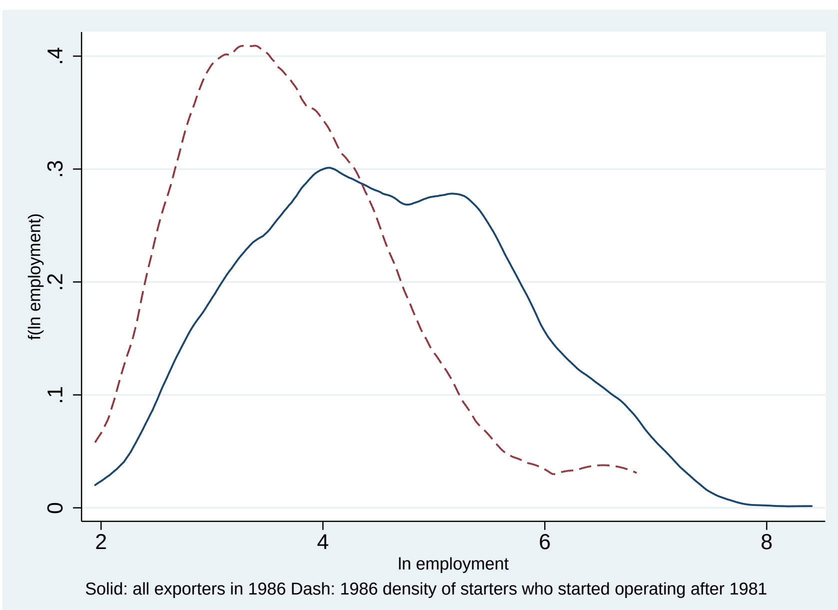


Figure 4.8: Starters: plants not operating in 1981, but exporting in 1986

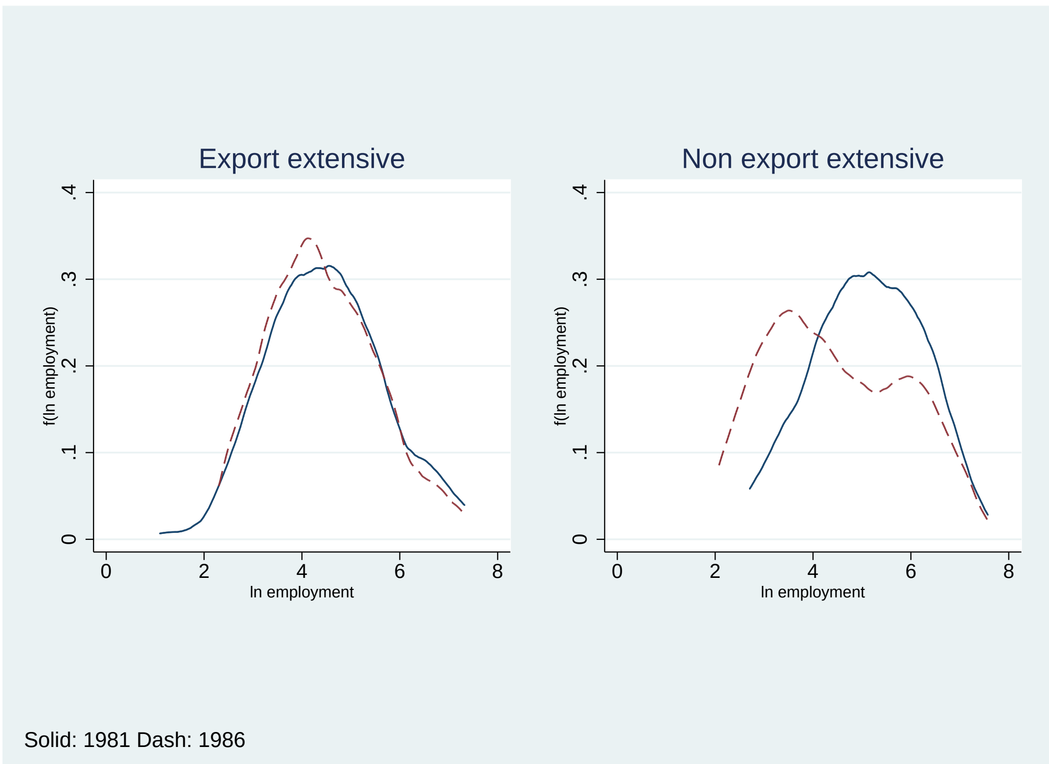


Figure 4.9: Exporters in export extensive vs. non extensive industries

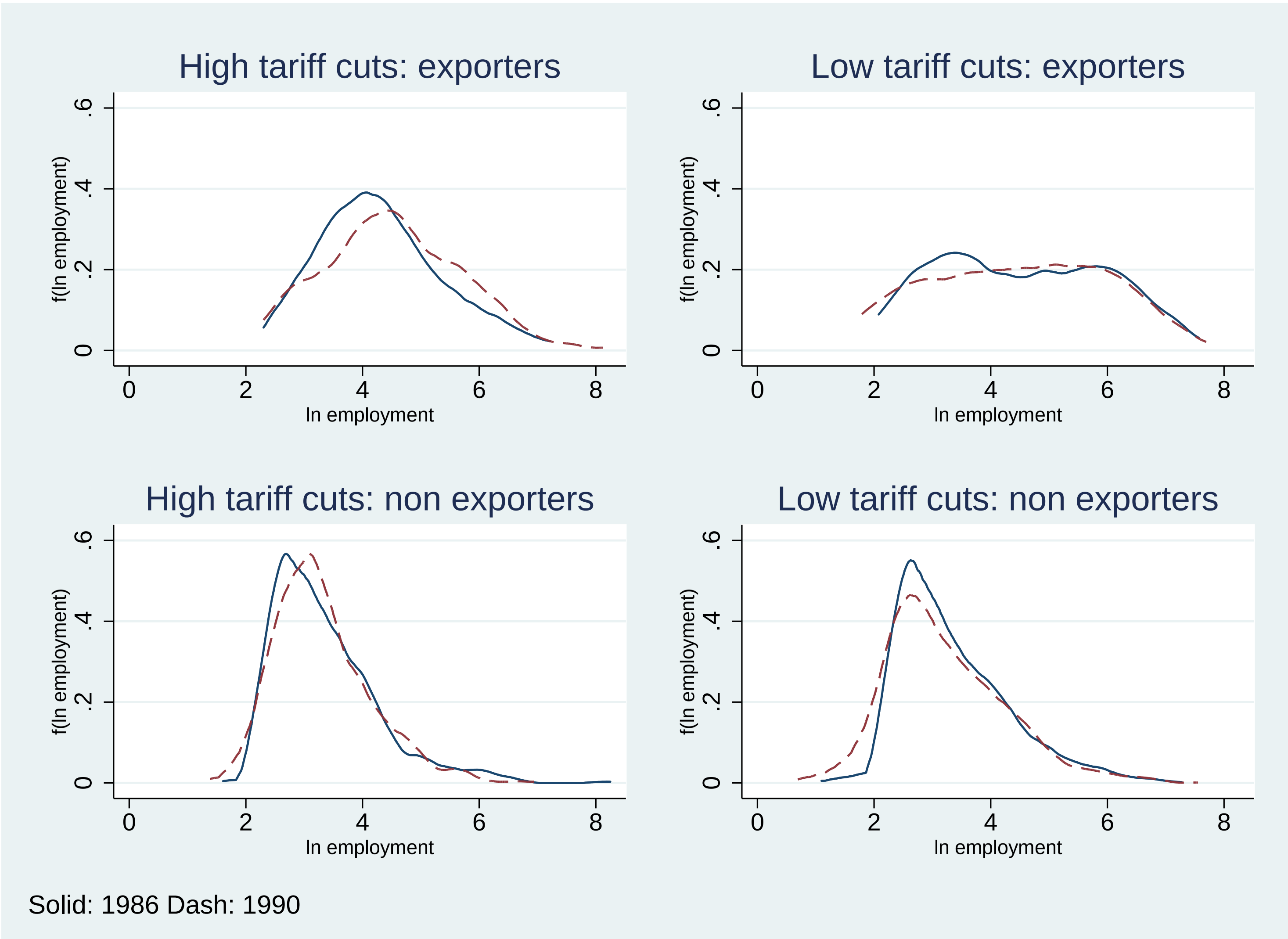


Figure 4.10: High vs. low tariff cut industries

# APPENDICES

# APPENDIX A

# Data used in chapter 1

**Amadeus data:** The firm-level panel was built using the online version of Amadeus downloaded in 2004 as well as three previous versions of the database on disk. These previous versions were merged with the online one because a given version of the database only contains those firms that filed statements in the last four years and drops the others.

**Regional data:** The measures of input linkages are constructed using the region-industry-level employment data and the input-output tables for Hungary in 2000, both obtained from the Hungarian Statistical Institute.

**Deflators:** Operating revenue is deflated by industry producer price index (PPI), obtained from the Hungarian Statistical Institute, labor is the wage bill deflated by national average wages, capital is deflated using the PPI for industry 7 (Machinery and Equipment), and the deflator for materials is an average of the PPI of several manufacturing industries. Industry-level imports (used in the import of inputs measure) are scaled down by industry value added, obtained from the OECD STAN Industry database.

## APPENDIX B

# World Bank estimates of GDP growth rates for the period 1986-2002

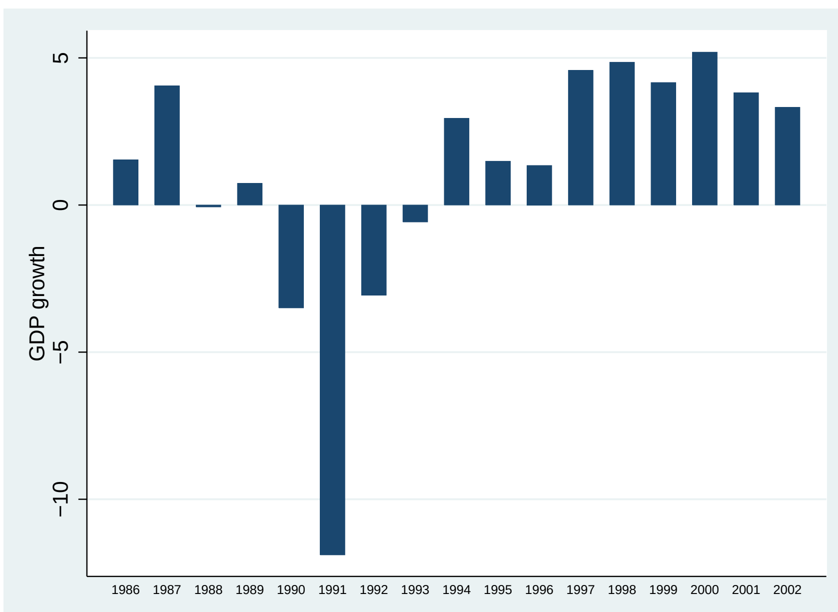


Figure B.1: GDP growth in Hungary, 1986-2002

## APPENDIX C

# Political map of Hungary


Figure C.1: Political map of Hungary

## APPENDIX D

# Distribution of employment scaled by plant age

Figure D.1 for 1981 shows that the scaled distributions are very similar to the unscaled ones.[1] Figures D.2 and D.3 show that scaling plant employment by plant age gives rise to similar evolutions in the distributions.

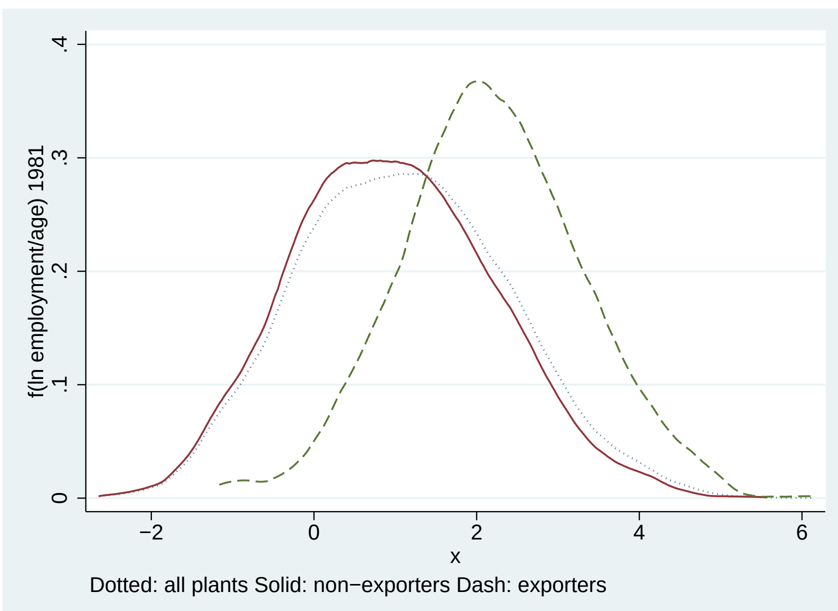


Figure D.1: Distributions of plant size scaled by plant age, by exporter status (1981)

[1] This is true for all years observed.

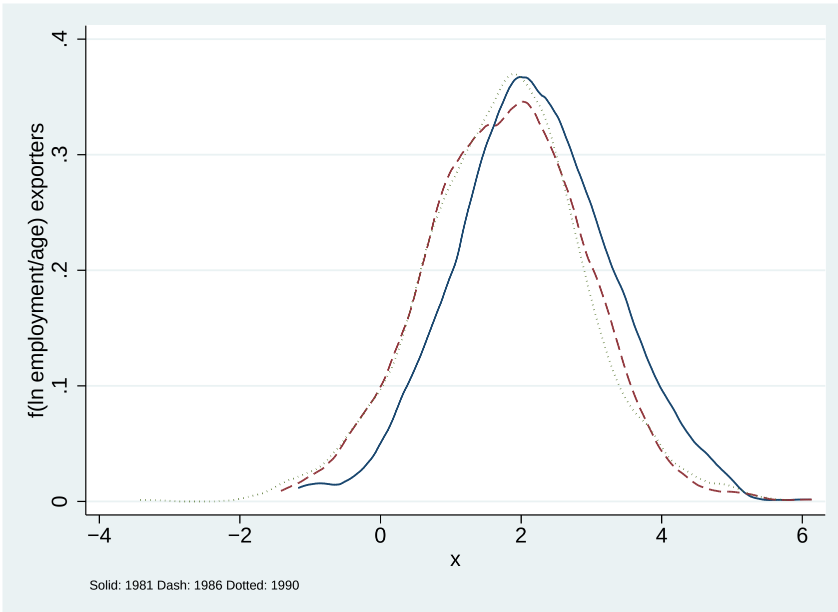


Figure D.2: Distributions of plant size scaled by age: exporters (1981, 1986 and 1990)

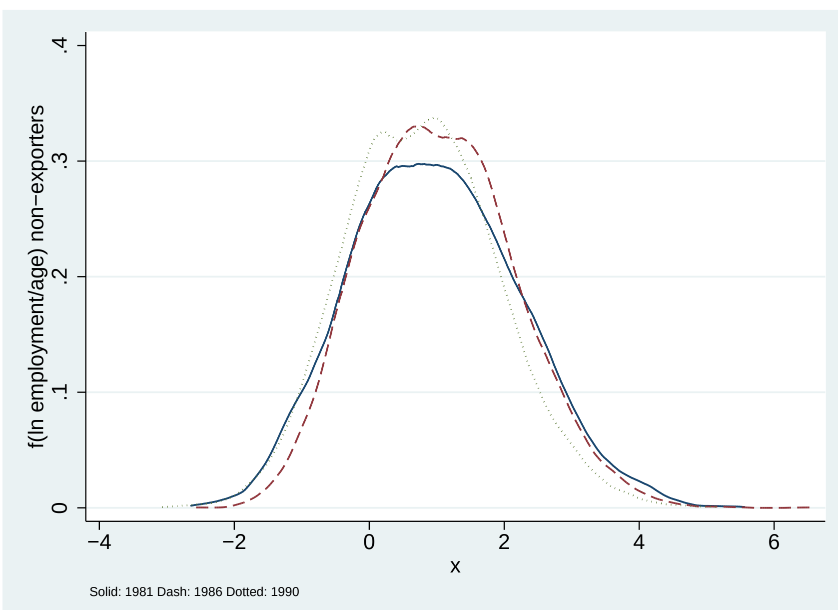


Figure D.3: Distributions of plant size scaled by age: non-exporters (1981, 1986 and 1990)

# APPENDIX E

## Check for a composition effect

Although Table 4.4 shows that there has been a redistribution of the number of exporters towards certain industries, this does not seem to affect the broad patterns in terms of the size distribution of plants: scaling plant employment by industry average employment in the same year yields the same patterns (Figures E.1 and E.2).

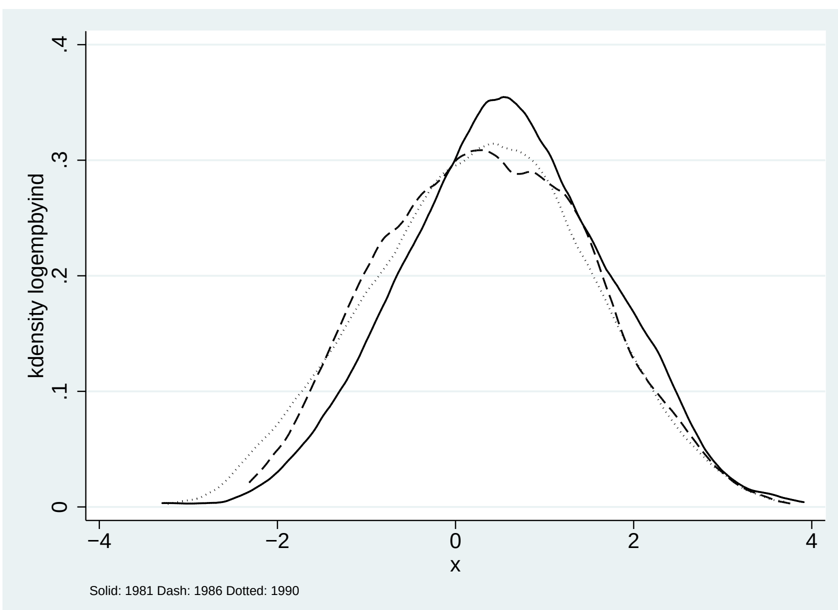


Figure E.1: Ln plant employment scaled by mean industry employment - exporters

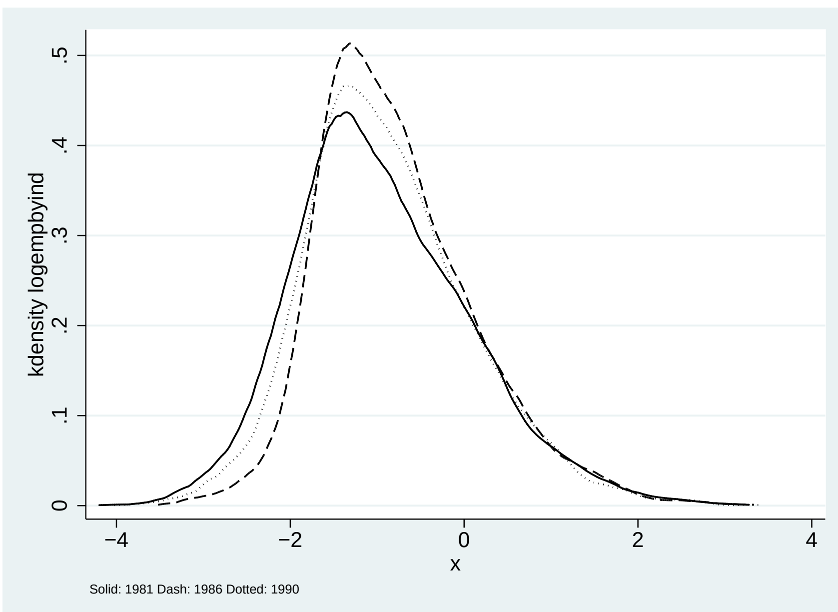


Figure E.2: Ln plant employment scaled by mean industry employment - non-exporters

# APPENDIX F

## Industry-level tariff changes

Table F.1: Tariff cuts between 1986 and 1990, by industry.

| **Industry** | **Tariff change** |
|---|---|
| **High tariff cut industries** | |
| 385 | -56% |
| 382 | -55% |
| 371 | -49% |
| 383 | -47% |
| 372 | -46% |
| **Low tariff cut industries** | |
| 362 | +3% |
| 314 | +1% |
| 311 | -7% |
| 312 | -8% |
| 369 | -9% |

# BIBLIOGRAPHY

# BIBLIOGRAPHY


Ades, A., and E. Glaeser (1997): "Trade and circuses: explaining urban giants," *Quarterly Journal of Economics*, 110(1), 195–227.

Aiginger, K., and S. W. Davies (2004): "Industrial specialisation and geographic concentration: Two sides of the same coin? Not for the European Union," *Journal of Applied Economics*, 7(2), 231–248.

Amiti, M. (1999): "Specialization Patterns in Europe," *Weltwirtschaftliches Archiv*, 135, 1–25.

Amiti, M., and L. A. Cameron (2007): "Economic Geography and Wages," *Review of Economics and Statistics*, 89(1), 15–29.

Amiti, M., and J. Konings (2007): "Trade Liberalization, Intermediate Inputs and Productivity: Evidence from Indonesia," *The American Economic Review*, 97(5), 1611–1638.

Atkinson, A. B. (1998): *Poverty in Europe*. Blackwell Publishing.

Attanasio, O., P. Goldberg, and N. Pavcnik (2004): "Trade reforms and wage inequality in Colombia," *Journal of Development Economics*, 74, 331–366.

Bernard, A., and B. Jensen (1999): "Exceptional exporter performance: cause, effect, or both?," *Journal of International Economics*, 47, 1–25.

——— (2004): "Why some firms export," *The Review of Economics and Statistics*, 86(2), 561–569.

Bernard, A., B. Jensen, and P. Schott (2006): "Trade Costs, Firms and Productivity," *Journal of Monetary Economics*, 53 (5), 917–937.

Bernard, A., S. Redding, and P. Schott (2007): "Comparative advantage and heterogeneous firms," *Review of Economic Studies*, 74(1), 31–66.

Bertrand, M., E. Duflo, and S. Mullainathan (2004): "How much should we trust differences-in-differences estimates?," *The Quarterly Journal of Economics*, 119(1), 249–276.

Brulhart, M. (2001): "Evolving geographical specialisation of European manufacturing industries," *Weltwirtschaftliches Archiv*, 137(2), 215–243.

Brulhart, M., and R. Traeger (2005): "An account of geographic concentration patterns in Europe," *Regional Science and Urban Economics*, 35, 597–624.

Cabral, L., and J. Mata (2003): "On the evolution of the firm size distribution: Facts and theory," *The American Economic Review*, 93 (4), 1075–1090.

Ciccone, A., and R. Hall (1996): "Productivity and the density of economic activity," *The American Economic Review*, 86(1), 54–70.

Combes, P.-P., and H. G. Overman (2004): *The spatial distribution of economic activities in the European Union, in Handbook of Regional and Urban Economics*vol. 4, pp. 2845–2909. Elsevier.

Crozet, M., and P. Koenig (2004): "EU enlargement and the internal geography of countries," *Journal of Comparative Economics*, 32, 265–279.

Das, S., M. Roberts, and J. Tybout (2007): "Market entry costs, producer heterogeneity, and export dynamics," *Econometrica*, 75(3), 837–873.

Di Giovanni, J., and A. Levchenko (2009a): "Firm entry, trade, and welfare in Zipfs world," mimeo, University of Michigan.

——— (2009b): "International trade and aggregate fluctuations in granular economies," mimeo, University of Michigan.

Dumais, G., G. Ellison, and E. Glaeser (1997): "Geographic concentration as a dynamic process," *NBER Working Paper*, 6270.

Eaton, J., M. Eslava, M. Kugler, and J. Tybout (2008): *The margins of entry into export markets: Evidence from Colombia*Harvard University Press.

Eaton, J., S. Kortum, and F. Kramarz (2008): "An anatomy of international trade: Evidence from French firms," working paper.

Ellison, G., and E. Glaeser (1997): "Geographic concentration in US manufacturing industries: A dartboard approach," *Journal of Political Economy*, 105, 889–927.

——— (May 1999): "The Geographic concentration of industry: Does natural advantage explain agglomeration?," *The American Economic Review, Papers and Proceedings*, 89(2), 311–316.

Eslava, M., J. Haltiwanger, A. Kugler, and M. Kugler (2004): “The effects of structural reforms on productivity and profitability enhancing reallocation: evidence from Colombia,” *Journal of Development Economics*, 75(2), 333–371.

——— (2008): “Trade reforms and market selection: Evidence from manufacturing plants in Colombia,” Mimeo.

——— (2009): “Factor adjustments after deregulation: Panel evidence from Colombian plants,” *Review of Economics and Statistics*, forthcoming.

Ethier, W. (1982): “National and international returns to scale in the modern theory of international trade,” *The American Economic Review*, 72(3), 389–405.

Feenstra, R., J. Markusen, and W. Zeile (1992): “Accounting for Growth with New Inputs: Theory and Evidence,” *The American Economic Review, Papers and Proceedings*, 82(2), 415–421.

Fernandes, A. (2007): “Trade policy, trade volumes and plant-level productivity in Colombian manufacturing industries,” *Journal of International Economics*, 71 (1), 52–71.

Fujita, M., P. Krugman, and A. J. Venables (1999): *The Spatial Economy*. MIT Press.

Garner, T., and K. Terrell (1997): “Changes in distribution and welfare in transition economies: Market vs. policy in the Czech Republic and Slovakia,” The William Davidson Institute Working Paper Number 77.

Gibrat, R. (1931): *Les inegalites economiques. Applications: Aux inegalites des richesses, a la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc., d’une loi nouvelle: La loi de l’effet proportionnel.* Sirey.

Gottschalk, P., and T. Smeeding (1997): “Cross-national comparisons of earnings and income inequality,” *Journal of Economic Literature*, 35, 633–687.

Hanson, G. (1996): “Localization economies, vertical organization, and trade,” *The American Economic Review*, 86(5), 1266–1278.

Head, K., and T. Mayer (2004): *The Empirics of Agglomeration and Trade*vol. 4, pp. 2609–2669. Elsevier.

Head, K., and J. Ries (1999): “Rationalization effects of tariff reductions,” *Journal of International Economics*, 47, 295–320.

Henderson, J. V., T. Lee, and Y.-J. Lee (2001): "Scale Externalities in a Developing Country," *Journal of Urban Economics*, 49, 479–504.

Ijiri, Y., and H. Simon (1977): *Skew distributions and the sizes of business firms.* North-Holland Pub. Co.

Javorcik, B. S. (2004): "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers Through Backward Linkages," *The American Economic Review*, 94(3), 605–627.

Kim, S. (1995): "Expansion of markets and the geographic distribution of economic activities: the trends in US regional manufacturing structure, 1860-1987," *Quarterly Journal of Economics*, 110, 881–908.

Krugman, P. (1980): "Scale economies, product differentiation, and the pattern of trade," *The American Economic Review*, 70(5), 950–959.

——— (1991): "Increasing returns and economic geography," *Journal of Political Economy*, 99(3), 483–499.

——— (1991a): *Geography and Trade*. MIT Press.

Krugman, P., and R. L. Elizondo (1996): "Trade policy and the Third World metropolis," *Journal of Development Economics*, 49, 137–150.

Lall, S., Z. Salizi, and U. Deichmann (2004): "Agglomeration economies and productivity in Indian industry," *Journal of Development Economics*, 73(2), 643–673.

Luttmer, E. G. J. (2004): "The size distribution of firms in an economy with fixed and entry costs," Federal Reserve Bank of Minneapolis, Working Paper 633.

——— (2007): "Selection, growth, and the size distribution of firms," *Quarterly Journal of Economics*, 122 (3), 1103–1144.

Machado, J., A., and J. Mata (2000): "Box-Cox quantile regression and the distribution of firm sizes," *Journal of applied econometrics*, 15, 253–274.

Melitz, M. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity," *Econometrica*, 71 (6), 1695–1725.

Melitz, M., and G. Ottaviano (2008): "Market size, trade, and productivity," *Review of Economic Studies*, 75, 295–316.

Midelfart-Knarvik, K. H., H. G. Overman, S. J. Redding, and A. J. Venables (2002): "The location of European industry," *European Economy*, 2, 216–273.

Moulton, B. (1990): "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *The Review of Economics and Statistics*, 72(2), 334–338.

Nocke, V., and S. Yeaple (2006): "Globalization and endogenous firm scope," PIER working paper 06-015.

Ottaviano, G., T. Tabuchi, and J.-F. Thisse (2002): "Agglomeration and trade revisited," *International Economic Review*, 43, 409–436.

Overman, H., S. Redding, and A. Venables (2003): *The Economic geography of trade, production and income: A Survey of empirics*pp. 353–387. Blackwell Publishing.

Overman, H., and A. Winters (2003): "Trade Shocks and Industrial Location: the Impact of EEC Accession on the UK," *mimeo LSE*.

Pavcnik, N. (2002): "Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants," *Review of Economic Studies*, 69, 245–76.

Roberts, M., T. Sullivan, and J. Tybout (1997): "Micro foundations of export booms," World Bank mimeo.

Roberts, M., and J. Tybout (1991): "Size rationalization and trade exposure in developing countries," World Bank WPS 594.

——— (1997): "The decision to export in Colombia: An empirical model of entry with sunk costs," *The American Economic Review*, 87 (4), 545–564.

Rossi-Hansberg, E., and M. Wright (2007): "Firm Size Dynamics in the Aggregate Economy," *American Economic Review*, 97 (5), 1639–1666.

Sen, A. K. (1997): *On economic inequality*. Oxford University Press.

Sutton, J. (1997): "Gibrat's legacy," *Journal of Economic Literature*, 35 (1), 40–59.

Tybout, J. (2003): *Plant and firm-level evidence on 'new' trade theories*Basil-Blackwell.

Tybout, J., J. de Melo, and V. Corbo (1991): "The effects of trade reforms on scale and technical efficiency. New evidence from Chile," *Journal of International Economics*, 31, 231–250.

Tybout, J., and D. Westbrook (1995): "Trade liberalization and the dimensions of efficiency change in Mexican manufacturing industries," *Journal of International Economics*, 39, 53–78.

Wooldridge, J. M. (2002): *Econometric analysis of cross section and panel data*. MIT Press.