FOREIGN DIRECT INVESTMENT WITH HETEROGENEOUS ENTRY COSTS*

BY SHISHU PENG

Institute of Economics, Academia Sinica

February 22, 2010

ABSTRACT

In this paper, I argue that multinational activity is not just a privilege of large firms. Models with fixed entry costs and heterogeneous productivity such as Melitz (2003) and Helpman, Melitz and Yeaple (2004) predict that only firms above a certain size or productivity threshold should engage in international trade or foreign direct investment (FDI). I show that this prediction is not supported by the data: the share of Taiwanese multinationals with foreign investments in both developing and advanced economies is smoothly increasing in firm size. I then argue that this fact can be explained by introducing entry cost heterogeneity to the model, which allows smaller firms to engage in multinational activity if their associated fixed cost is low. This framework also predicts that, relative to the standard model, (1) the effect of expanding foreign markets on multinational activity is reduced, because new entrants in foreign markets are on average smaller; (2) policies which reduce entry costs have a smaller effect on total sales of the firms investing abroad and their extent of global investments.

Keywords: Foreign Direct Investment, firm heterogeneity, fixed entry costs.

JEL Classification: F12

---

* E-mail: ssp@econ.sinica.edu.tw. I would like to thank Pranab Bardhan, Maurice Obstfeld, Ann Harrison, Pierre-Olivier Gourinchas, Professor Yuriy Gorodnichenko and in particular Adam Szeidl for helpful discussions.

1. INTRODUCTION

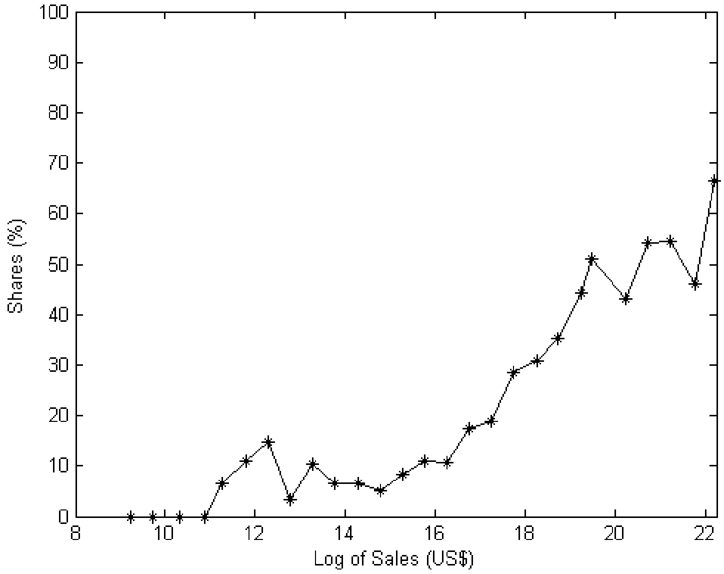
Firm size and productivity appear to exhibit considerable heterogeneity both in developing and in advanced countries.[1] This heterogeneity applies to multinational activity as well — while many firms only serve their domestic markets, some firms export their goods, and others set up subsidiaries in foreign markets. Recent theoretical work in international trade, including Melitz (2003) and Helpman, Melitz and Yeaple (2004), has modeled these patterns by combining heterogeneity in productivity with fixed costs of export or foreign direct investment (FDI). These models predict the sorting of multinational activity as a function of productivity: least productive firms produce domestically, more productive ones export their products, and the most productive companies perform FDI.

A key advantage of this class of models is that they allow for an analysis of firm entry and exit in international markets. In standard models with homogenous firms, the effect of a policy change such as a reduction in trade costs would be that all firms increase exports. In a Meltiz-type framework, this effect on the "intensive margin" is complemented by additional exports through an "extensive margin", where new firms who now find it profitable to export enter international markets.

The particular fixed entry cost formulation of these Melitz-type models has a sharp prediction about the relationship between firm productivity and multinational activity. Essentially, these models predict a complete separation of multinational activity as a function of firm productivity, or equivalently, given monopolistic competition, as a function of firm size. Companies below a given productivity and size are all predicted to stay domestic; firms with productivities in a higher range are all expected to export; and the most productive and largest firms tend to all engage in FDI. This prediction is a consequence of the assumption that all firms face the same fixed entry costs when engaging in a given international activity.

In this paper I first document that this sharp prediction of Melitz-type models fails to

[1] Bernard and Jensen (1995, 1999), Sofronis Clerides, Saul Lach, and James R. Tybout (1998), and Bee Yan Aw, Sukkyun Chung, and Mark J. Roberts (2000) document stylized facts about the behavior and performance of firms across a number of countries. They find that exporters are in the minority; they tend to be larger and more productive.



Source: Department of Statistics, the Ministry of Economic Affairs, Taiwan, 2003 and 2004.

Figure 1: Shares of Taiwanese Firms Investing in both Developing and Developed Countries

describe well the foreign investment behavior of Taiwanese multinational companies. Using data from 3,229 Taiwanese firms in 2003 and 2004, Figure 1 plots, as a function of firm size, what share of Taiwanese companies who invest abroad choose to invest in both developing economies (the South) and advanced economies (the North).[2] The share of firms engaging in FDI in both the South and the North is a slowly increasing function of firm size, rather than the piecewise constant function with a cutoff predicted by the theory.[3] Following Helpman *et al.* (2004), this figure uses total sales as a proxy for firm productivity. A similar smoothly increasing graph would be obtained when sales per worker is used as a proxy for productivity.

One possible explanation for this finding is that fixed entry costs may exhibit variation across firms and markets. Market-specific fixed costs could be the result of variation in distance to markets, differences in language, culture and institutions, among

[2] I explain the construction of the Figure in Section 2 below. Firm size is measured by the firm sales.

[3] All firms can be categorized into four main industries: 1) metal and machinery, 2) information and technology, 3) chemistry and 4) food, textile and other. The patterns are similar even if we look at the firms in each specific industry instead of all firms.

others. Firm-specific entry costs can arise as a consequence of differences in access to international trade networks, corporate culture, reputations, or the ability to raise funds, etc.[4]

To evaluate this possible explanation, I incorporate heterogeneity in entry costs into the Melitz framework. I assume that firms are heterogeneous both in their productivity and in the fixed entry costs they face. In this framework, firms with the same size might make different FDI choices: those facing high fixed costs only invest in developing countries; while others with low fixed costs can afford to conduct FDI in advanced economies as well.

I then calibrate this model to match the pattern documented in Figure 1. I assume that both firm productivity and fixed entry costs follow lognormal distributions. To estimate the means and variances of these distributions, I compute, for each possible firm size, the share of firms with that size performing FDI in the South and the share of firms performing FDI both in the South and in the North. I then match these predicted shares to the actual shares from the Taiwanese data, and estimate the parameters of the productivity and fixed cost distributions using a weighted least squares approach.

Figure 2 shows the predicted share of firms performing FDI in both the North and the South in this calibration. As the figure shows, this model allows me to accurately match the share of firms investing in both developed and developing regions, improving on the standard formulation. My setup also outperforms the standard approach in predicting the distribution of firm size for both types of firms. A central prediction of heterogeneous entry costs is that multinational activity is not just a privilege of large firms: Firms with a broad range of size and productivity can end up investing abroad if business conditions are favorable. For example, in these calibrations, 37% of all firms investing in both regions are in the lower half of the firm size distribution; in contrast, a calibrated version of the model with homogenous fixed costs predicts no multi-regional FDI for these firms.

The heterogeneous entry cost model also has quantitative and policy implications that are different from the standard model. To explore these differences, I consider the

[4] For instance, Rauch (2001) studies the role of transnational networks in international trade. The network might help to improve the resource allocation by creating trade, and generate benefits for its members.

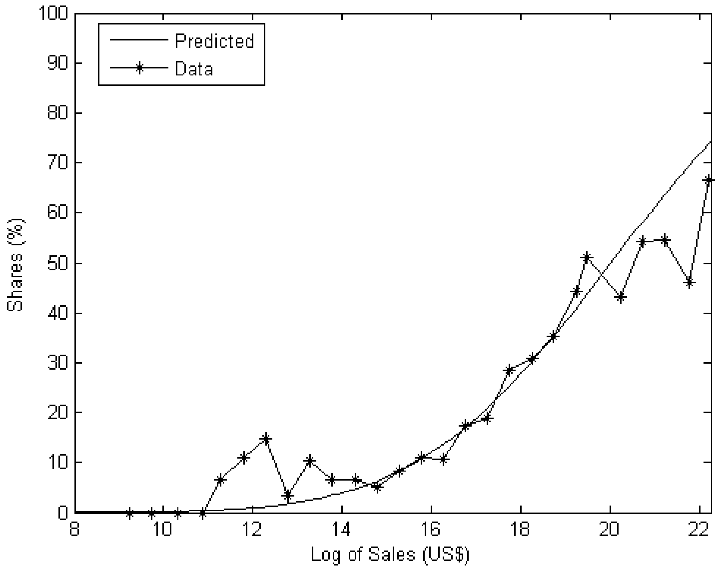


Figure 2: Predicted Shares of Taiwanese Firms Investing in both Developing and Developed Countries

following two scenarios to explore the effects of two environmental changes: (1) an expansion in the market size of the North; (2) a policy change which reduces the fixed entry costs of all firms. I analyze the impact of these changes in calibrated versions of the standard Melitz-type model and the heterogeneous entry cost framework, with parameters chosen to best match the distribution of sales shown in Figure 1.

In scenario 1, as the market size of the North expands, both the homogeneous and heterogeneous entry cost models predict increases in the share of the firms who invest in both regions, and in the proportion of these firms' sales in the total sales. However, these shares respond much faster in the homogeneous entry cost model. Intuitively, as the North expands, in the homogenous cost model the new firms investing on the extensive margin are all large, and hence generate a high volume of new foreign sales. In contrast, with heterogeneous entry costs, the firms entering on the extensive margin are drawn from the entire size distribution and hence are on average much smaller. As a result, their effect on total sales is smaller. Since the heterogeneous entry cost model fits the data better, this finding suggests that the quantitative importance of the extensive margin of adjustment might be smaller than previously thought.

In scenario 2, the effect of the policies that reduce entry costs depends on the initial level of entry costs. When the initial level of entry costs is high, a small reduction results in a greater increase in multinational activity in the model with homogenous entry costs, because the firms that enter through the extensive margin are large. In contrast, when the initial level of entry costs is small, a further reduction predicts larger increases in total sales with heterogeneous entry costs, because it allows for increased entry across the entire firm size distribution. For realistic values of entry costs, it appears that the first effect dominates, and hence the effect of globalization policies on multinational activity is smaller in the heterogeneous entry cost framework.

This paper builds on the literature about heterogeneous firms in international trade. In an influential paper, Melitz (2003) combines heterogeneity in productivity with a fixed entry cost to study intra-industry reallocation in international trade. He finds that only the most productive firms choose to export and that trade liberalization would induce more trade through the extensive margin.

Helpman *et al.* (2004) use the Melitz framework to explore the proximity-concentration tradeoff between horizontal FDI and exports. They also find such sorting by productivity: the most productive firms engage in FDI, less productive ones export, and the least productive firms stay in the home market. Grossman *et al.* (2006) incorporate both horizontal and vertical FDI into this framework, and find a complex pattern of integration strategies governed by market size, productivity and trade costs.

Aw and Lee (2008) study the pattern of FDI empirically using the data on Taiwanese multinationals that I do in this paper. They find that among firms investing abroad, most productive firms invest in both the U.S. and China; less productive ones invest only in the U.S.; while the least productive firms invest only in China. This pattern is consistent with the basic logic of the above models, but, as I show above, the data does not match the sharp cutoff prediction of the theories.

The rest of this paper is organized as follows. In Section 2, I document the basic empirical facts underlying Figure 1. Section 3 presents the model, and Section 4 develops the estimation approach. Section 5 explores the effect of expanding the market in the North and the implications of trade policies. Section 6 concludes.

Table 1: Descriptive Statistics for Taiwanese Multinationals Investing Abroad

	Developing	**Developed**	**Both**
Average Sales (US$)	69.87m	128.42m	239.15m
Average Fixed Assets Purchases (US$)	5.14m	19.65m	17.94m
Average R&D (US$)	1.66m	5.79m	4.77m
Average Employment	725	505	1,473
Average Investment Duration (Years)	5.87	5.21	6.15
Firm Number	2,586	362	643

Source: Department of Statistics, the Ministry of Economic Affairs, Taiwan (2003 and 2004).
Notes: Sample size is 3,591. The developed region includes: the U.S., Canada, Western Europe, Hong Kong, Japan, Singapore, Australia and New Zealand. The developing region contains: Mexico, Central and Southern America, East Europe, Mainland China, Malaysia, Thailand, Indonesia, Philippines, Vietnam, South Asia and Africa.

2. FIRM SIZE AND MULTINATIONAL ACTIVITY: EVIDENCE

2.1 Data

In this paper I use data from the survey of "Outward FDI Survey in Manufacturing" about Taiwanese manufacturing sector in the years of 2003 and 2004. This survey is conducted by the Department of Statistics of the Ministry of Economic Affairs of Taiwan. The sample contains 1,880 and 1,711 Taiwanese manufacturing firms investing abroad in these two years. The data contains the firms' industrial classification, sales levels, total employment, and the destination countries of their FDI. Table 1 contains summarized statistics of these data.

As the table shows, Taiwanese multinationals which invest in both developed and developing regions are larger than those which invest only in developing regions in every dimension: they have higher levels of sales, make higher fixed assets purchases and R&Ds, and employ more workers.[5] This comparison also holds on a per worker basis.

[5] This relationship does not hold for those firms investing only in developed regions. These firms have higher per worker sales, fixed asset purchases, and R&D than the firms investing in both regions. They also hire fewer workers and invest for a shorter duration. I find that two thirds of these firms are in the IT industry. Their behavior might relate have more to other concerns such as access to high technology and are worth researching in the future. Hereafter I drop these 362 firms to make the analysis less complicated and

Table 2: Firm Size and Multiple-Region FDI

	Small	Mid-Sized	Large
All firms	12.14%	26.45%	61.41%
Metal and Machinery	9.55%	21.21%	69.24%
Information and Technology	16.83%	32.93%	50.24%
Chemistry	9.55%	21.85%	68.60%
Food, Textile and Others	10.62%	22.73%	66.65%

Source: Department of Statistics, the Ministry of Economic Affairs, Taiwan (2003 and 2004).
Notes: Sample size is 3,229. Each row shows the proportions of firms investing in both developed and developing regions that belong to small, mid-sized or large sizes. The measurement of the size is each firm's sales.

This pattern confirms the basic prediction of Melitz-type models: the most productive firms would engage in international activities with higher entry costs.

However, if we look at the breakdown of the multinationals who invest in both regions, we will find that they are not all large firms. Table 2 shows the proportion of firms investing in both regions that fall in different firm size groups. The firms are divided, by sales, into three equal-sized groups: small, mid-sized and large firms. The first row shows that out of all firms investing in both regions, only 61% of them are large; while 39% are small or mid-sized. As the next rows show, within individual industries, small and mid-size firms constitute 30% to 50% of all firms which engage in multiple-region FDI as well. This table tells us that multi-region FDI might not be a privilege of large firms.

2.2 Construction of Figure 1

I divide the 3,229 firms in the sample into two groups. (1) The South firms: these are firms investing only in the developing countries; (2) The Global firms: firms investing in both developing and developed countries. I then sort these firms by their log sales levels and divide them into twenty seven bins, each of which has the same length in units of log size.[6] For each bin, I compute the share of firms investing in both regions relative to the total number of firms in that bin. The result has already been shown in Figure 1.

focus only on the firms investing only in developing region and in both developing and developed regions.

[6] This length is 0.5, in terms of log sales in U.S. dollars.

Melitz-type models which only consider firm heterogeneity in productivity predict a sharp relationship between the firm size and the FDI in multiple regions, since firms above a certain threshold can all overcome the unique fixed entry costs. In the next section I will modify this Melitz-type framework by introducing another dimension of firm heterogeneity in entry costs to try to interpret the pattern in Figure 1.

3. A MODEL OF FDI WITH HETEROGENEOUS ENTRY COSTS

The model is a modification of the framework of Helpman *et al.* (2004), so most of the structure is similar to their work. I make a change to their model by introducing the heterogeneous entry costs to replace their assumption of a unique homogeneous fixed entry cost. Also, due to the limitation of the data, I focus only on a firm's choice between the single-region and multi-region FDIs, instead of on trade and FDI.

There are three countries, Home, the South and the North. The South and North refer to the developing and developed regions, respectively. In each country, firms use labor to produce goods in H + 1 sectors. In one sector firms produce a homogeneous product, which is the numeraire, using one unit of labor per unit output, while in the other H sectors, firms produce differentiated products. In each country, an exogenous fraction β_h of income is spent on differentiated products of sector h, and the remaining fraction $1-\sum_h \beta_h$ on the homogeneous good. Country i is endowed with L^i units of labor and wage rates w^i, where i = Home, South or North.

Consider a particular sector h in the home country that produces differentiated products. From now on the index h is dropped for simplicity but be aware that all sectoral variables refer to h. A firm j draws a labor-per-unit-output coefficient "a_j" from a distribution $G(a)$ and produces and sells domestically.[7]

This firm may choose to serve foreign markets through FDI. In order to enter the South, a firm incurs a market-specific fixed entry cost f_S, measured in labor units, which

[7] I make a simplified assumption that there is no fixed entry cost for entering this industry, and also no fixed overhead labor costs in the Home country, since the decision of exit and entry is not the issue of interest. Given the former assumption, firms with different coefficients a_j will all serve the Home market.

is the same for all firms. Likewise, when entering the North, firm j bears an additional firm-specific fixed entry cost f_{Nj}, which is different across all firms. I assume f_{Nj} is drawn from a distribution $F(f_N)$. Both f_S and f_{Nj} include the distribution and servicing network costs, as well as the costs of forming a subsidiary in the South or North.

For the preferences across varieties of product h, the standard CES form with an elasticity of substitution $\varepsilon = 1/(1-\rho) > 1$ is assumed for all countries. Thus the demand function for each brand in country i is $A^i p^{-\varepsilon}$, where A^i is exogenous from the viewpoint of each individual firm.[8] As a result, each brand of the monopolistic firm with labor coefficient a_j is facing sales at the price $p = w^i a/\rho$, where $1/\rho$ is the markup factor.

I hereafter focus on the FDI strategies of a firm in the Home country. A firm in the Home country that stays in that sector will always produce domestically to serve its domestic market. It may further also serve foreign markets by investing and building a subsidiary there to produce and sell its differentiated products. To focus only on the two FDI choices for the firms: investing in the South, or in both the South and North, I assume the following constraint for the fixed costs of entering the South and North:

(1) $$f_{Nj} > \left(\frac{W^N}{W^S}\right)^{1-\varepsilon} \frac{B^N}{B^S} f_S;\ \forall j,$$

to exclude the FDI choice of investing only in the developed countries.[9] Following subsection 2.2, I call the firms who make the former FDI choice the “South firms”, the ones who make the latter FDI choice the “Global firms.”

For simplicity, assume by now $w^i = 1$ for all countries, which will be the case as long as the numeraire good is produced in every country and freely traded. The operating profits from serving the Home market will then be $\pi_j^H = a_j^{1-\varepsilon} B^H$ for firm j with a

[8] The utility function here implies $A^i = \beta E^i / [\int_0^{n^i} p^i(v)^{1-\varepsilon} dv]$ where E^i is the aggregate level of spending in country i; n^i is the number of varieties available in country i; and $p^i(v)$ is the consumer price of variety v.

[9] Helpman *et al*. (2004) discuss the decision between exports versus FDI.

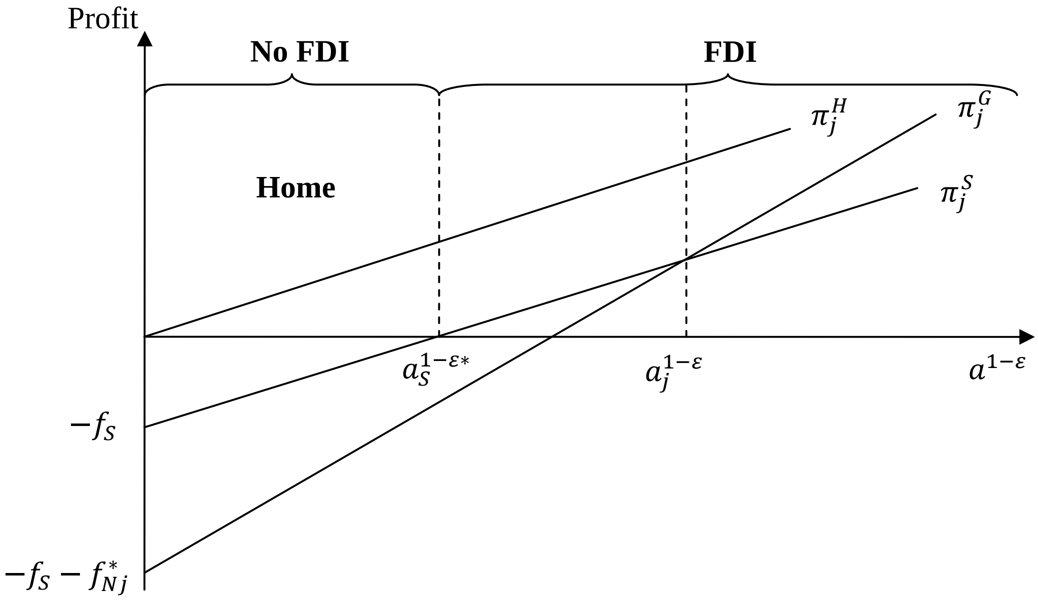


Figure 3: Profits from Serving Different Markets

labor-output coefficient a_j, where $B^H = (1-\rho)A^H/\rho^{1-\varepsilon}$.[10] If a firm selling domestically decides to invest in the South, its additional operating profits from serving the South market will be $\pi_j^S = a_j^{1-\varepsilon}B^S - f_S$ where $B^S = (1-\rho)A^S/\rho^{1-\varepsilon}$. The additional operating profits from serving both the South and North markets will be $\pi_j^G = a_j^{1-\varepsilon}(B^S + B^N) - f_S - f_{Nj}$. These profit functions are illustrated in Figure 3 for the case of equal demand levels $B^H = B^S = B^N$ and $w^i = 1$ for all countries.

Figure 3 shows the operating profits for a given firm *j*. The variable for the horizontal axis is $a^{1-\varepsilon}$, which can be used as a productivity proxy since it increases monotonically with labor productivity $1/a$, due to the assumption $\varepsilon > 1$. All profit functions are linear and increasing, implying that more productive firms are more profitable in all strategies.

The profit functions π_j^H and π_j^S are parallel due to the assumption $B^H = B^S$, but the profits from the South are lower because of the fixed entry cost of f_S. The profit

[10] The demand function $A^H p^{-\varepsilon}$ implies the output level $A^H(a/\rho)^{-\varepsilon}$ when the price is a/ρ. The costs are then $\rho A^H(a/\rho)^{1-\varepsilon}$, while the revenue is $A^H(a/\rho)^{1-\varepsilon}$. Thus the operating profits are $\pi^H = (1-\rho)A^H(a/\rho)^{1-\varepsilon}$.

function π_j^G will be different across firms due to heterogeneity in entry costs of entering the North. It is steeper than π_j^H and π_j^S because by investing in both the South and North, firms can access larger markets. These relationships in Figure 3, together with the first inequality in (1), guarantee all firms will first consider investing in the South before investing in the North.[11]

The sorting of the market access over the horizontal line of productivity is as follows. The least productive firms whose productivity level is below $a_S^{1-\varepsilon *}$ earn positive operating profits by serving domestically, but expect to lose by engaging in FDI. They will choose to just serve domestic but not foreign markets. The cutoff $a_S^{1-\varepsilon *}$ is the productivity level at which domestic firms earn zero profits from performing FDI.

Firms with productivity levels higher than $a_S^{1-\varepsilon *}$ will choose to invest abroad since they begin to earn positive operating profits from foreign markets, in addition to serving the domestic market only. There will not be a universal break-even productivity level for these more productive firms due to heterogeneity in their fixed entry costs investing in the North. Instead, each firm has its own cutoff fixed entry cost level of investing in the North f_{Nj}^*.

For example, a firm with productivity level $a_j^{1-\varepsilon}$ in the figure will compare its real f_{Nj} with this cutoff value: if $f_{Nj} \leq f_{Nj}^*$, it will invest in both regions; if $f_{Nj} > f_{Nj}^*$, this firm will invest only in the South.

The universal cutoff coefficients $a_S^{1-\varepsilon *}$ and the firm-specific f_{Nj}^* for a firm with productivity level of $a_j^{1-\varepsilon}$, are determined by

(2) $$B^S a_S^{1-\varepsilon *} = f_S,$$

(3) $$B^N a_j^{1-\varepsilon} = f_{Nj}^*.$$

Conditions (2) and (3) offer implicit solutions for the cutoff coefficients a_S^* and a_{Nj}^* and the demand levels B^i in each country. These solutions do not depend on the country-size variables L^i as long as wages w^i are still equalized.

[11] Now inequality in (1) becomes $f_{Nj} > f_S;\ \forall j$.

4. ESTIMATION

4.1 Linkages between Sales and Productivity

Melitz-type framework bases its analysis on the firms' productivity levels, but in the data only the sales level is observable. I now develop the theoretical connection between unobserved productivity and observed sales in the model.

The connection between productivity, measured by productivity proxy $a^{1-\varepsilon}$ and sales, denoted by $sales$, for firms investing only in the South is given by

$$a_S^{1-\varepsilon} = \frac{\rho^{1-\varepsilon}}{(w^H)^{1-\varepsilon}A^H + (w^S)^{1-\varepsilon}A^S} \cdot sales. \tag{4}$$

For firms that invest in both regions, the relationship changes because in addition to the sales in the Home country, we must now take into account sales in both the North and the South:

$$a_G^{1-\varepsilon} = \frac{\rho^{1-\varepsilon}}{(w^H)^{1-\varepsilon}A^H + (w^S)^{1-\varepsilon}A^S + (w^N)^{1-\varepsilon}A^N} \cdot sales. \tag{5}$$

Here $a_S^{1-\varepsilon}$, $a_G^{1-\varepsilon}$ are the productivity levels of the South and Global firms, respectively, that generate the same sales level of $sales$.

Recall from (3) that the cutoff level of the fixed entry costs at which firms with a given productivity $a_j^{1-\varepsilon}$ start to conduct FDI in the North is

$$f_{Nj}^{*} = (w^N)^{1-\varepsilon} B^N a_j^{1-\varepsilon}. \tag{6}$$

In these equations I relax the assumption that wages are the same in all countries, to take into account potential wage differentials in the estimation.

4.2 Proxies for the Parameters

To match the model with the data summarized by Figure 1, I need to calibrate the following parameters: the elasticity of substitution (ε), the markup factor coefficient of

Table 1.3: Parameters

Parameters	Notations	Values
Elasticity of Substitution	ε	5
Markup Factor Coefficient	ρ	0.8
	w^H	\$10.81
Wage Rates (Home, South and North)	w^S	\$3.33
	w^N	\$15.02
	A^H	\$16,106m
Demand Levels (Home, South and North)	A^S	\$100,059m
	A^N	\$178,658m

the firms (ρ), the wage rates in the Home, South and North (w^H, w^S and w^N), and the demand levels of the Home, South and North (A^H, A^S and A^N). I list the parameters in Table 3.

Gallaway *et al.* (2000), find an elasticity of substitution of ε = 1.55 for the United States.[12] This, however, will result in a high markup $1/\rho$ of 2.86. In the calibration I then assume ε = 5 so that the corresponding markup factor is a more reasonable number of 1.25, namely ρ = 0.8. The elasticity of substitution is assumed to be the same for all countries.

The proxies for the wage rates in the Home, South and North are computed based on the data in the "Statistics and Databases" of the International Labor Organization.[13] I take the annual wage data during 2003 and 2004 for the developing (South) and

[12] Gallaway *et al.* (2000) make disaggregated and comprehensive estimates covering 311 industries at the four-digit U.S. SIC level (monthly data; from Jan. 1989 to Dec. 1995). The average long-run industry-level estimate of Armington elasticity of substitution of 118 SIC manufacturing industries of the U.S. is 1.55 (ranging from 0.52 to 2.83). It should be noted that the elasticities they estimate are the ones between domestic goods and imported goods within the same sector in the Armington CES utility function, which views these two goods of the same kind as differentiated ones.

[13] The data is from the website of the International Labour Organization (ILO): http://www.ilo.org/. The Taiwanese wage rates are taken from The Economic Statistical Indicator, Department of Statistics, MOEA, Taiwan. The same method of computing the wage proxies of the South and North is applied to compute the wage proxy of Taiwan (the Home country).

developed (North) countries, adjust them by the year 2000 CPI and the corresponding PPP exchange rates.[14] Then I compute the annual average wages of the two years for each country and finally use the PPP-adjusted GDP of each country as weights to calculate the wage rates for the South and North. I obtain an hourly wage of $3.33 for the South, and $15.02 for the North. The same method is applied to compute the hourly wage rate for Taiwan and the result is $10.81.

In the model, A^H, A^S and A^N are the demand levels of sector h in the Home, South and North. The Taiwanese data is obtained from the website of the Department of Statistics in MOEA in Taiwan. For the latter two variables, I take the data of industry output levels for different countries from the Industrial Statistics Database (INDSTAT4 2006 ISIC Rev.3), published by United Nations Industrial Development Organization (UNIDO). The categorization of the South and North countries is a bit different from that used in computing wage proxies, due to the data limitation. [15]

I compute the average output level of 61 ISIC 3 digit manufacturing industries for each country from 2003 and 2004, adjusted by the corresponding CPI of year 2001 and PPP exchange rates; then add them up for the South and North respectively to obtain the output levels as proxies for A^S and A^N. I use the same method to compute A^H, except that I get the data of Taiwan from the Department of Statistics in MOEA in Taiwan. The result is: $16,106 million for Taiwan, $100,059 million for the South and $178,658 million for the North.

[14] The developed and developing countries are defined as follows. Developed countries: Australia, Austria, Belgium, Canada, France, Germany, Hong Kong, Japan, Luxemburg, Netherlands, New Zealand, Singapore, Switzerland, United Kingdom and United States; developing countries: Argentina, Brazil, Chile, China, Costa Rica, Czech Republic, El Salvador, Guatemala, Hungary, India, Mexico, Nicaragua, Philippines, Poland, South Africa, Sri Lanka, and Thailand. These countries are chosen because they are the countries where Taiwanese firms tend to invest. Due to the lack of the data, however, Honduras, Indonesia, Malaysia, and Vietnam are not chosen.

[15] The developed countries include: Australia, Austria, Belgium, Canada, France, Germany, Japan, Luxemburg, Netherlands, New Zealand, Singapore, Switzerland, United Kingdom, and the United States. The developing countries involve: Argentina, Brazil, China, Czech Republic, Hungary, India, Indonesia, Malaysia, Philippines, Poland, South Africa, and Vietnam. Note that the data of China is the value added of industries from China Statistical Yearbook, 2001 through 2006, National Bureau of Statistics of China.

4.3 Estimation Approach

This subsection outlines my strategy for matching the data summarized in Figure 1 for both the homogeneous and heterogeneous entry costs models. Notice that in the calibration I assume $f_S = 0$ for simplicity.

I begin by defining two variables: the ratios of the South and Global firms to all firms. They can be calculated directly from the data for each of the twenty seven bins defined in subsection 2.2. I denote the ratio of the South firms in bin *k* in all firms by z_{Sk}, and the ratio of the Global firms in bin *k* in all firms by z_{Gk}. The subscript of *k* will be omitted below since all variables refer to the *k*th bin.

The predicted ratios of the South and Global firms to all firms in bin *k*, denoted as $\widehat{z_S}$ and $\widehat{z_G}$, are obtained in different ways in the homogeneous and heterogeneous entry cost models and will be explained below.

4.3.1 The Homogeneous Fixed Entry Cost Model

In Melitz-type models, firms are assumed to be heterogeneous only in productivity, but homogeneous in the fixed entry costs investing in the North. There is thus a clear cutoff productivity level, $a_G^{1-\varepsilon*}$, for all firms performing FDI. Firms which are more productive than this level would all invest in both regions since they are able to cover this homogeneous entry cost; those whose productivity levels are below this level would invest only in the South simply because they cannot bear this cost.

Let me call the bins with all firms investing in the South “South bins”, and the bins with all firms investing in both regions “Global bins”. For the South bins, the Melitz-type predicted probability of the South firms, i.e. $\widehat{z_S}$, is equal to the probability of the firms with the productivity level $a_S^{1-\varepsilon}$ (corresponding to the sales level of that bin) and $\widehat{z_G}$ is zero since it is predicted that no firms invest in both regions for the South bins. Similarly, $\widehat{z_S}$ and $\widehat{z_G}$ are zero and $g(a_G^{1-\varepsilon})$ respectively for the North bins. Assuming log normal distribution to the productivity $\log a^{1-\varepsilon} \sim N(\mu_a, \sigma_a^2)$, I use the weighted least square method to estimate the two parameters μ_a and σ_a^2:

$$\text{(7)} \qquad \min_{\mu_a, \sigma_a^2} \frac{1}{n_S} \sum_{k=1}^{27} (\widehat{z_{Sk}} - z_{Sk})^2 + \frac{1}{n_G} \sum_{k=1}^{27} (\widehat{z_{Gk}} - z_{Gk})^2,$$

where n_S and n_G represent the numbers of all the South and Global firms, respectively.

Since we do not know the exact level of $a_G^{1-\varepsilon*}$ (it is determined by the unknown homogeneous fixed entry cost investing in the North), it is unclear how many South bins there are. Different combinations of the South and Global bins will result in different values of residual sum of squares (RSS). By looking at the data, we see that there are 28 combinations correspond to different values of the homogeneous fixed entry costs of investing in the North. I choose the combination that gives the lowest RSS to estimate the parameters.

4.3.2 The Heterogeneous Entry Costs Model

In this paper, firms are assumed to be heterogeneous both in the productivity levels and fixed entry costs of entering the North. The predicted ratios of the South and Global firms to all firms in the kth bin, denoted as $\widehat{z_S}$ and $\widehat{z_G}$, are obtained as follows. The predicted probability of the South firms, $\widehat{z_S}$, is the probability of the firms with the productivity level $a_S^{1-\varepsilon}$ (corresponding to the sales of kth bin) which end up with high entry costs and decide not to invest in the North. Similarly, $\widehat{z_G}$ is the probability of the firms with the productivity level $a_G^{1-\varepsilon}$ who end up with low entry costs and decide to invest also in the North. Technically, they can be expressed as:

(8) $$\widehat{z_S} = g(a_S^{1-\varepsilon}) \cdot \left(1 - F(f_{NS}^*)\right) \text{ and } \widehat{z_G} = g(a_G^{1-\varepsilon}) \cdot F(f_{NG}^*)$$

where f_{NS}^* denotes the cutoff fixed entry cost entering the North for the South firms, f_{NG}^* the counterpart for the Global firms.[16]

Here I assume both the productivity and fixed entry cost investing in the North follow log normal distribution: $\log a^{1-\varepsilon} \sim N(\mu_a, \sigma_a^2)$ and $\log f_N \sim N(\mu_f, \sigma_f^2)$.[17] I also use the weighted least square method to estimate these four parameters:

(9) $$\min_{\mu_a, \sigma_a^2, \mu_f, \sigma_f^2} \frac{1}{n_S} \sum_{k=1}^{27} (\widehat{z_{Sk}} - z_{Sk})^2 + \frac{1}{n_G} \sum_{k=1}^{27} (\widehat{z_{Gk}} - z_{Gk})^2.$$

[16] The expression of (8) is based on the assumption of independence of the productivity level and the fixed entry costs of investing in the North. Also, the subscript j is omitted since all the South (or Global) firms are assumed to have the same productivity levels and cutoff fixed entry cost entering the North in the same bin.

[17] I also tried Pareto distribution, but it seems the fitting is better with log normal distribution.

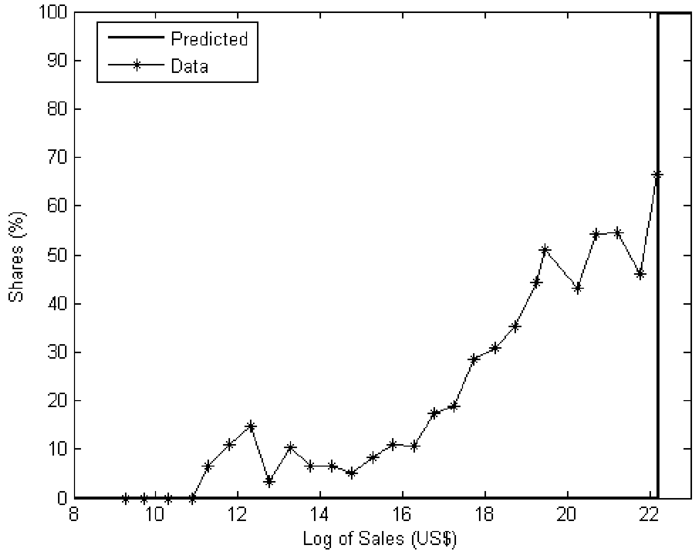


Figure 4: Predicted Shares of Global Firms (Homogeneous Entry Cost Model)
Note: This figure compares between the shares of Global firms derived from data (starred, smoothly increasing as in Figure 1) and those predicted by the homogeneous fixed entry cost model (solid). The latter is a step function: with log of sales below 22.2, the shares are zeros; while above 22.2, they are 100%.

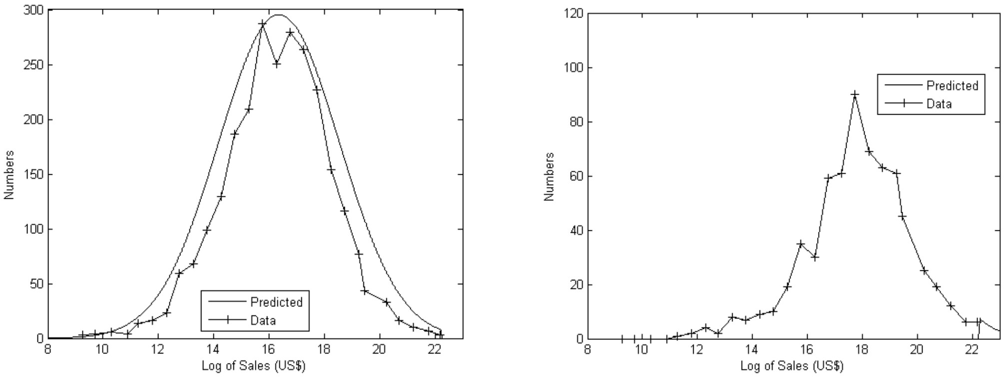
4.4 Results and Comparisons of the Two Approaches

4.4.1 The Homogeneous Fixed Entry Cost Model

The estimated values for μ_a and σ_a^2 are -3.26 and 2.18, respectively. This implies the average output per worker is 0.44 in terms of labor productivity. The predicted universal fixed entry cost of entering the North facing all firms is \$822,415 higher than that of entering the South. The R-squared is 0.52. With the parameters estimated, the share of the North firms in each bin can be obtained by $\widehat{z_G}/(\widehat{z_S}+\widehat{z_G})$ and depicted in Figure 4.

The predicted share curve in Figure 4 is a step function: almost all firms invest in the South (share = 0) and only a few very large firms (log sales greater than 22) invest in both regions (share = 1).

Figure 5 shows the predicted distribution of the sales for the South and Global firms. The homogeneous model overestimates the number of South firms and underestimates that of Global firms. We can also see the clear cutoffs in the two figures in the right tails.



South Firms Global Firms

Figure 5: Firm Sales Distributions of South and Global Firms

(Homogeneous Entry Costs Model)

Note: For the left figure (South firms), the homogeneous entry cost model overestimates the number of South firms (solid curve is above the starred curve) and fails to capture the right tail of the distribution of South firms. For the right figure (North firms), however, the homogeneous entry cost model only reveals the right tail of the distribution of South firms without showing most Global firms of small and mid-sized.

4.4.2 The Heterogeneous Entry Costs Model

The estimated values for μ_a and σ_a^2 are -2.93 and 1.92, respectively. This implies the average output per worker is 0.48 in terms of labor productivity. The estimated values for μ_f and σ_f^2 are 12.97 and 3.44, respectively. This tells us that the predicted entry cost of entering the North is, on average, \$431,111 higher than that of entering the South.

The R-squared is 0.98, much higher than 0.52 in the previous case. This implies the heterogeneous entry costs approach fits the data better than the homogeneous approach. The share of the Global firms in each bin can also be obtained by $\widehat{z_G}/(\widehat{z_S}+\widehat{z_G})$ and is depicted in Figure 6.

The predicted share curve in Figure 6 basically exhibits a similar trend to the data — as the sales get larger, the share of the Global firms increases. We also observe that the matching is better for middle-size firms, probably because there are more firm samples in those bins.

Figure 7 shows the predicted distributions of the firm sales for the South and Global firms, respectively. Compared with Figure 5, the heterogeneous entry cost model predicts the distributions of both kinds of firms better than the homogeneous entry cost model. Also, there is no clear cutoff in the right tails.

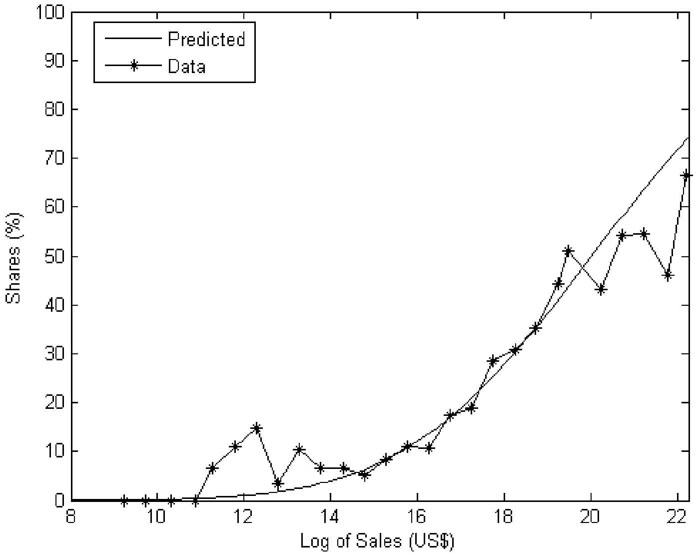
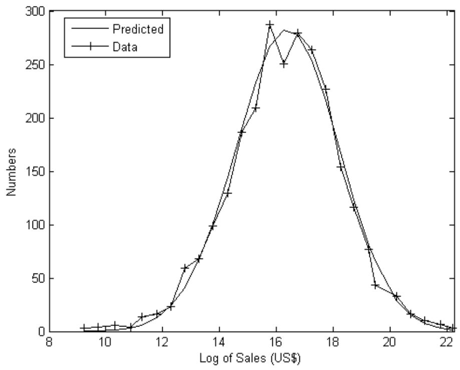
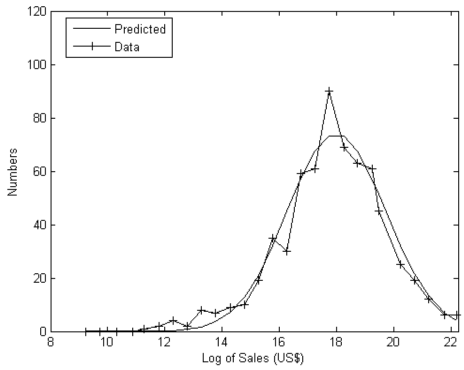


Figure 6: Predicted Shares of Global Firms (Heterogeneous Entry Costs Model)

Note: This figure is shown before as Figure 2. By introducing heterogeneous fixed entry costs, the share curve of Global firms becomes a gradually increasing function instead of a step function in Figure 4.





South Firms Global Firms

Figure 7: Firm Sales Distributions of South and Global Firms

(Heterogeneous Entry Costs Model)

Note: The two figures, like their counterparts in Figure 1.5, show, for South and Global firms, the firm sales distributions derived from the data (starred) and predicted by the heterogeneous fixed entry cost model (solid). In both figures, we see that by introducing the heterogeneous fixed entry costs, we can fit the data-derived distributions of both kinds of firms better, compared with Figure 5, because the solid curves match the starred curves more closely.

5. APPLICATIONS

5.1 Comparative Statics

In this subsection, I will analyze the comparative statics utilizing the calibrated versions of the two models, with the parameters chosen in section 4. In the model of section 3, changes of several parameters will cause the firms to re-consider their FDI decisions. These parameters include: the markup factor coefficient of the firms (ρ), wage rates in the South and North (w^S and w^N), and the demand levels of the South and North (A^S and A^N). Here I only consider the scenario in which the market size of the North (A^N) expands, from its original size (the benchmark case), by 10% to 100%.[18]

When the North market size expands, it is more profitable to invest in the North since serving the North will bring higher revenues. The difference between the homogeneous and heterogeneous entry cost models is the following: As the North market expands, in the homogenous cost model the new firms investing on the extensive margin are all large, and hence generate a high volume of new foreign sales. In contrast, with heterogeneous entry costs, the firms entering on the extensive margin are drawn from the entire size distribution and hence are on average much smaller. As a result, their average productivity and effect on foreign sales are smaller.

Figure 8 shows, for the two models, the trends of total sales and the sales of Global firms as the North market expands. The sales level of South firms is then the difference between these two sales. I hereafter call the sales made by Global firms "Global sales", and the sales of the South firms "South sales". Note that the share of the Global sales in the total sales can be viewed as a measure for the degree of globalization.[19]

Firstly, both models predict increases in total and Global sales, and decreases in the South sales with the expansion of the North market.[20] The homogeneous entry cost

---

[18] The shrinking of the market size of the South, the decrease of the wage rage in the North, the rise of the wage rate in the South, and the fall in markup have similar qualitative effects with the expansion of the North market size.

[19] Here I view "globalization" as "investing globally, namely, both in the South and North". In other words, if a firm invests only in the South, it is not a globalized firm in this view.

[20] The total sales in both models increase slightly in Figure 1.8. They are more obvious in Figure 1.9 and 1.10.

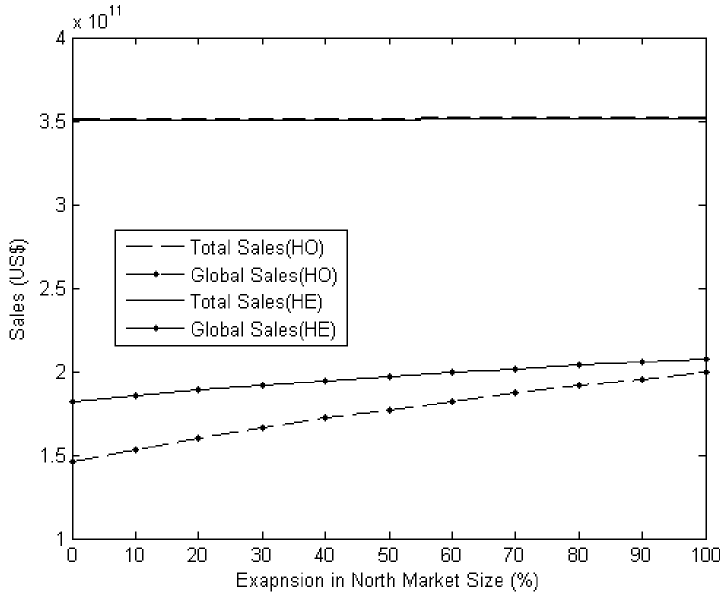


Figure 8: Sales by All Firms and by Firms Investing in Both Regions as North Market Size Expands

Note: (1) HO stands for the homogeneous entry cost model; HE denotes the heterogeneous model. (2) This figure shows the trends of sales created by all firms (total sales) and sales only by firms investing in both regions (Global firms). Notice that the upper two curves are overlapping because with the scale of the y-coordinates, meaning that the total sales of both models are similar in this figure. We can see that Global sales rise faster in the homogeneous model than in the heterogeneous model in this figure.

model predicts a faster increase in Global sales. Global sales rise because more new firms decide to also invest in the North. They contribute to Global sales through two channels — the original sales in the South, and the new sales in the North, and both are higher in the homogeneous entry cost model.

Two effects play roles here. On the one hand, in the homogeneous entry cost model the firms newly entering the North are on average more productive than those in the heterogeneous model. This is because in order to match the data, the weighted least square method predicts very high level of sales as the lower threshold for the marginal Global firms. On the other hand, there are fewer new firms entering the North in the homogeneous than in the heterogeneous model due to nature of log normal distribution.

In the calibration, when we compute Global sales, the former dominates the latter, namely, in the homogeneous model, although there are fewer new Global firms as the North expands, but since their (marginal) sales level are so high, their contribution to overall Global sales will be larger than those new Global firms in heterogeneous model. In other words, the faster increase in Global sales in the homogeneous model is caused by

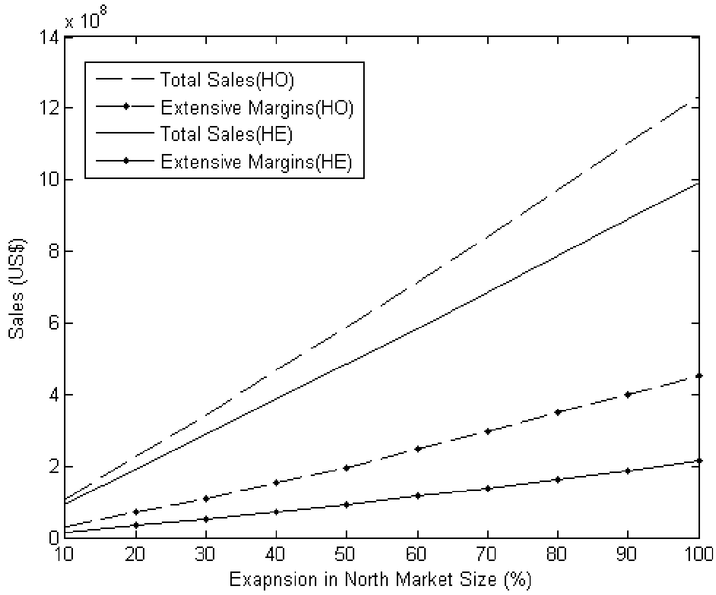


Figure 9: Extensive Margins as North Market Size Expands

Note: This figure illustrates, for the two models, the increases of sales created by all firms (total sales) and the increases in sales by all new Global firms (extensive margins). The lower two dotted curves show the trends of extensive margins for both models. We see that, compared with the homogeneous model, by introducing heterogeneous entry costs, the increases in total sales and extensive margin are both smaller.

fewer but much larger new firms investing in the North.

Secondly, the composition of total sales, or the degree of globalization, is different in the two approaches. The share of Global sales rises from 42% to 57% in the homogeneous approach while it rises from 52% to 59% in the heterogeneous approach. This higher share of Global sales (i.e. higher degree of globalization) is driven by the more global FDI activity of small and mid-sized companies.

Lastly, the increase in total sales has two sources — the intensive and extensive margins. As the North market expands, total sales rise because of increases in the sales per incumbent Global firm (intensive margin) and because of net increases in sales per new Global firm (extensive margin). By decomposing the total sales into these two margins, I can compare the relative importance of the extensive margin in the two models.

Figure 9 and 10 show the relationships between the increases in the total sales and the extensive margins as well as the intensive margins induced by the expansion of the North market. In Figure 9, the extensive margin in the homogeneous cost model (the dashed dotted curve) is above that in the heterogeneous approach (solid dotted curve).

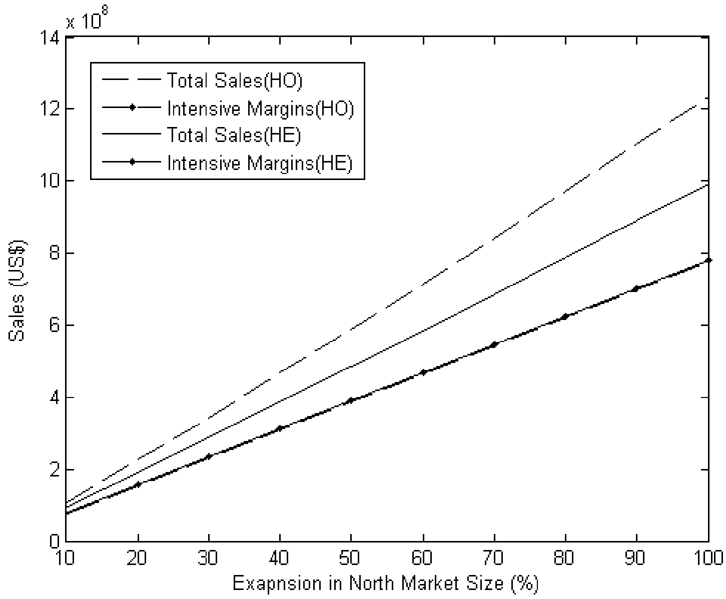


Figure 10: Intensive Margins as North Market Size Expands

Note: This figure shows, for the two models, the increases of sales created by all firms (total sales) and the increases in sales by all incumbent Global firms (intensive margins). This figure tells us that the intensive margins are the same in both models (two dotted curves overlap).

Also, the extensive margin always plays a less important role in the rise of the total sales in the heterogeneous approach (around 22%) than in the homogeneous cost model (around 34%).

In Figure 10, the intensive margins of the two models are at the same level since the incumbent Global firms' sales in the North grow at the same rate in the two models as the North market expands. In the figure, it is shown as the overlapping of the dashed and solid dotted curves. The intensive margin then plays a relatively more important role, in a relativity sense, in the rise of the total sales in the heterogeneous cost model, though the magnitudes of the intensive margins are the same in both models. Since the heterogeneous entry cost model fits the data better, the findings in Figure 9 and 10 suggest that the quantitative importance of the extensive margin of adjustment might be smaller than suggested by earlier literature.

5.2 Policy Implications

This subsection examines the effects of trade policies that reduce firms' entry costs. I assume that the government implements a policy which reduces each firm's fixed entry

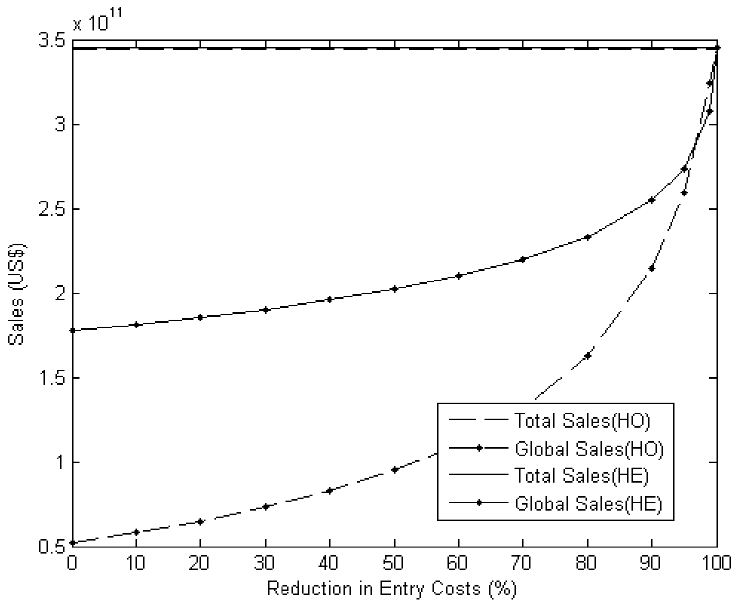


Figure 11: Sales by All Firms and by Firms Investing in Both Regions as the Entry Costs Fall

Note: This figure shows the trends of sales created by all firms (total sales) and sales only by firms investing in both regions (Global firms). Note the total sales of both models are similar in this figure (upper two curves overlap). In this figure, when the entry costs don't fall by that much (less than 95%), Global sales rise faster in the homogeneous fixed entry costs model.

cost by 10% to 100%. Examples of such policies include helping firms to promote their reputations by government quality assurances, enhancing their abilities to raise financing through the public banking system, improving their networks with foreign official institutions, and so on.

Figure 11 shows the changes in the total and Global sales for the two models. The gap between these two sales is then South sales. The low levels of percentage to the left of the horizontal axis imply a slight decrease in the entry costs, thus meaning high entry cost levels. Both models predict increases in the total and Global sales, and decreases in South sales as the entry costs fall. [21] Notice that all changes in the total sales come from the new firms who enter the North through the extensive margin.

We can see that when the entry costs are high, Global sales (dashed dotted curve) in the homogeneous cost model increase faster than those (solid dotted curve) in the

[21] Again, the total sales of the two models increase slightly in the scale of Figure 11. The increases can be barely observed in the figure and the two total sales curves overlap with each other. The total sales in the homogeneous entry cost model increase faster when the entry costs are still high, as shown in Figure 12, due to the more important role of extensive margin.

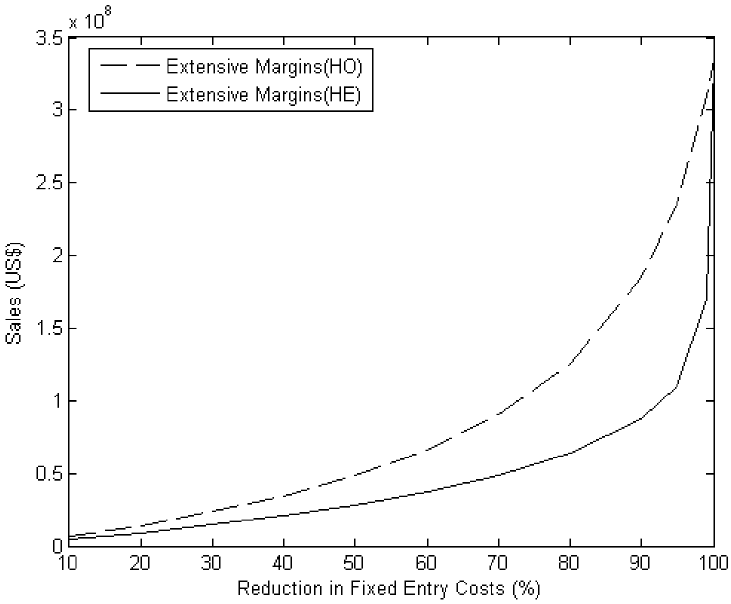


Figure 1.12: Extensive Margins as the Entry Costs Fall

Note: This figure shows the trends of the extensive margins created by new firms investing in the North when the fixed entry costs of all firms decrease by the same percentage. This figure shows that the extensive margin should be smaller when considering the heterogeneous fixed entry costs (the solid curve is always below the dashed one).

heterogeneous model. This means that with high initial levels of entry costs, a small reduction results in a greater increase in multinational activity in the homogenous approach, because large firms enter through the extensive margin.

In contrast, when the entry costs are extremely low, Global sales (solid dotted curve) in the heterogeneous approach increase faster than that (dashed dotted curve) in the homogeneous entry cost model. This implies that when the initial level of entry costs issmall, a further reduction predicts higher increases in foreign sales in the heterogeneous approach, because it allows for increased entry across the entire firm size distribution. If we consider a more realistic case where entry costs are not extremely low, both the levels and shares of Global sales change more slowly in the heterogeneous approach (as does the South sales).

The effectiveness of such policies on the total sales is shown in Figure 12, which presents the change in total sales that comes entirely from the extensive margin. The total sales increase slower in the heterogeneous cost model than in the homogeneous model when the entry cost is high due to the less important role of extensive margin in the former model.

By considering of firm heterogeneity in entry costs, the analysis above tells us that policies would be less effective in changing the levels of South, Global, and total sales. This implies the effectiveness of the policies which reduce all firms' entry costs on total sales and globalization might be smaller than previously thought.

In addition, the costs and policy objectives of such policies could be quite different in the two approaches. In the homogeneous approach, the government has to reduce the universal entry cost of all firms by the same percentage. In the heterogeneous approach, however, if the government can identify those firms with high fixed entry costs, it can implement policies that reduce only their entry costs. This means that the costs of the latter can be smaller since the government does not need to focus on all firms.

In summary, the analysis above tells us that the effects of economic environmental changes and trade policies on the total sales and the degree of globalization may be relatively smaller. This implication should also apply further to international trade with firm heterogeneity in trade costs.

6. CONCLUSION

This paper has introduced firm heterogeneity in entry costs of serving foreign markets to explain the observation that the share of multinationals with foreign investments in both developing and advanced economies is gradually increasing in firm size. This model matches data on the foreign investment behaviors of Taiwanese firms better and implies that investing in multiple regions is not just a privilege of large firms: small companies also engage in multinational activity if their associated entry costs are low. One implication I also show is that the extensive margin might be quantitatively smaller than suggested by earlier papers, thus mitigating the effectiveness of the policies which reduce entry costs of firms on foreign sales and also the degree of globalization.

The substantive new assumption in this paper is firm heterogeneity in entry costs. While such differences across firms are intuitive, it will be helpful to identify proxies for such entry costs in firm-level data.

This paper focuses on the analysis of foreign direct investment. In future research, I intend to explore whether heterogeneous entry costs also help explain the exporting behavior of firms, and trade-off between export and FDI.

References

Aw, Bee Yan and Lee, Yi. "Firm Heterogeneity and Location Choice of Taiwanese Multinationals." *Journal of International Economics*, 2008, 75(1), pp. 167-79.

Aw, Bee Yan; Chung, Sukkyun and Roberts, Mark J. "Productivity and Turnover in the Export Market: Micro Evidence from Taiwan and South Korea." *The World Bank Economic Review*, January 2000, 14(1), pp. 65-90.

Bernard, Andrew B. and Jensen, J. Bradford. "Exporters, Jobs, and Wages in U.S. manufacturing: 1967-87." *Brookings Papers on Economic Activity: Microeconomics*, Vol. 1995, pp. 67-112.

_______ . "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics*, February 1999, 47(1), pp. 1-25.

Clerides, Sofronis; Lach, Saul and Tybout, James R. "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco." *Quarterly Journal of Economics*, August 1998, 113(3), pp. 903-947.

Gallaway, Michael P.; McDaniel, Christine A. and Rivera, Sandra A. "Long-Run Industry-Level Estimates of U.S. Armington Elasticities." USITC Working Paper No. 2000-09-a, U.S. International Trade Commission, October 2000.

Grossman, Gene, Helpman, Elhanan and Szedl, Adam. "Optimal Integration Strategies for the Multinational Firm." *Journal of International Economics*, 2006, 70(1), pp. 216-38.

Helpman, Elhanan; Melitz, Marc J. and Yeaple, Stephen R. "Export versus FDI with Heterogeneous Firms." *American Economic Review*, March 2004, 94(1), pp. 300-316.

International Labour Organization. "Statistics and Databases." International Labour Office , Geneva.

Melitz, Marc J. "The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations." *Econometrica*, November 2003, 71(6), pp. 1695-725.

MOEA (Ministry of Economic Affairs). "Outward FDI Survey in Manufacturing." Executive Yuan, Taiwan, 2003-4.

Rauch, James E. "Business and Social Networks in International Trade." *Journal of Economic Literature*, December 2001, 39(4), pp. 1177-1203.

United Nations Industrial Development Organization (UNIDO). "Industrial Statistics Database." Statistics and Information Networks Branch, Vienna, Austria, 2006.