

Three Essays in International Trade

by

Thomas CHANEY

M.A., Ecole Normale Supérieure (2000)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2005

© Thomas CHANEY. All rights reserved.

The author hereby grants to Massachusetts Institute of Technology permission to
reproduce and
to distribute copies of this thesis document in whole or in part.

Signature of Author
Department of Economics
16 August 2005

Certified by
Xavier GABAIX
Rudi Dornbusch Career Development Associate Professor of Economics
Thesis Supervisor

Certified by
Marc MELITZ
John and Ruth Hazel Associate Professor of the Social Sciences
Thesis Supervisor

Certified by
Daron ACEMOGLU
Charles P. Kindleberger Professor of Applied Economics
Thesis Supervisor

Accepted by
Peter TEMIN
Elisha Gray II Professor of Economics
Chairman, Departmental Committee on Graduate Studies

Three Essays in International Trade

by

Thomas CHANEY

Submitted to the Department of Economics
on 16 August 2005, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis is a collection of three essays in international trade.

Chapter 1 explains how firm heterogeneity and market structure can distort the geography of international trade. By considering only the intensive margin of trade, Krugman (1980) predicts that a higher elasticity of substitution between goods magnifies the impact of trade barriers on trade flows. In this chapter, I introduce firm heterogeneity in a simple model of international trade. I prove that the extensive margin, the number of exporters, and intensive margin, the exports per firm, are affected by the elasticity of substitution in exact opposite directions. In sectors with a low elasticity of substitution, the extensive margin is highly sensitive to trade barriers, compared to the intensive margin, and the reverse holds true in sectors with a high elasticity. The extensive margin always dominates, and the predictions of the Krugman model with representative firms are overturned: the impact of trade barriers on trade flows is dampened by the elasticity of substitution, and not magnified. To test the predictions of the model, I estimate gravity equations at the sectoral level. The estimated elasticities of trade flows with respect to trade barriers are systematically distorted by the degree of firm heterogeneity and by market structure. These distortions are consistent with the predictions of the model with heterogeneous firms, and reject those of the model with representative firms.

Chapter 2 demonstrates the importance of liquidity constraints in international trade. If firms must pay some entry cost in order to access foreign markets, and if they face liquidity constraints to finance these costs, only those firms that have sufficient liquidity are able to export. A set of firms could profitably export, but they are prevented from doing so because they lack sufficient liquidity. More productive firms that generate large liquidity from their domestic sales, and wealthier firms that inherit a large amount of liquidity, are more likely to export. This model predicts that the scarcer the available liquidity and the more unequal the distribution of liquidity among firms, the lower are total exports. I also offer a potential explanation for the apparent lack of sensitivity of exports to exchange rate fluctuations. When the exchange rate appreciates, existing exporters lose competitiveness abroad, and are forced to reduce their exports. At the same time, the value of domestic assets owned by potential exporters increases. Some liquidity constrained firms start exporting. This dampens the negative competitiveness impact of a currency appreciation. Under some circumstances, it may actually reverse it altogether and increase aggregate exports. This model provides some argument for competitive revaluations.

In chapter 3, I build a dynamic model of trade with heterogeneous firms which extends

the work of Melitz (2003). As countries open up to trade, they will experience a productivity overshooting. Aggregate productivity increases in the long run, but it increases even more so in the short run. When trade opens up, there are too many firms, inherited from the autarky era. The most productive foreign firms enter the domestic market. Competition is fierce. The least productive firms that are no more profitable are forced to stop production. Not only do the most productive firms increase their size because they export, but the least productive firms stop producing altogether. Aggregate productivity soars. As time goes by, firms start to exit because of age. Competition softens. Some less productive firms resume production. This pulls down aggregate productivity. The slower the exit of firms, the larger this overshooting phenomenon. This model also predicts that the price compression that accompanies trade opening may be dampened in the long run. It also predicts that inequalities should increase at the time when a country opens up to trade, and then gradually recede in the long run.

Thesis Supervisor: Xavier GABAIX

Title: Rudi Dornbusch Career Development Associate Professor of Economics

Thesis Supervisor: Marc MELITZ

Title: John and Ruth Hazel Associate Professor of the Social Sciences

Thesis Supervisor: Daron ACEMOGLU

Title: Charles P. Kindleberger Professor of Applied Economics

à Claude Arbelet

Acknowledgements

I owe an enormous debt of gratitude to my advisors, Xavier Gabaix, Marc Melitz, and Daron Acemoglu. I am proud to have worked under their supervision. Xavier has been a constant support during all these years. He is the most ambitious researcher I have ever met. His ambition has pulled me up, it has given me the hope of making science move forward. His contagious enthusiasm has helped me believe in what I was doing. Following in the footsteps of Marc Melitz has been extremely rewarding. He welcomed me warmly, and offered me tremendous support. His ability to get to the simplest expression of a complex idea has allowed me to go far beyond where I would ever have dreamt of going. Daron's Herculean intellectual strength has helped me and influenced me all those years. I have been lucky enough to benefit from his infinite knowledge. Daron en büyük! My only regret is not having spent more time bothering him.

I am grateful to my professors, Rudi Dornbusch, Jaume Ventura, the faculty of the economics department at MIT, and Daniel Cohen. I have had the immense honor of being introduced to international economics by Rudi Dornbusch. Rudi embodied the spirit of elegance, curiosity and honesty in economics, and was one of the pillars of the economics department at MIT. He is an ideal I thrive to get close to, and his memory has guided my research. I was taught international trade by Jaume Ventura. Jaume is a marvelous and inspiring teacher. Without Rudi and Jaume's classes, I would not have oriented my research towards international trade. I discovered economics in the hectic intellectual atmosphere of the economics department at MIT. I am grateful to the entire faculty there. I am proud to have been part of this amazing department. I discovered the beauty of economics in the lectures of Daniel Cohen, my professor at the Ecole Normale Supérieure. He awoke in me an intellectual enthusiasm for economics that never left. Without him, I would not have come to MIT. I am indebted to Daniel for his trust and his guidance.

I am grateful to the people at LSE, who have so kindly welcomed me for an entire year. I wrote part of this thesis while working there. For their hospitality, I want to thank all the people I met during those intense days of interviews and flyouts of the job market. Those meetings gave me both ideas and confidence.

I wish to thank all the friends I've met in Boston. They have allowed me to survive those years, and actually enjoy it despite the distance from my true love. Endless discussions have greatly

contributed to this thesis. A special thank to all the glorious and beautiful Ibbetson street's extended family. The founding fathers, Ivan Fernandez, Ruben Segura, Karna Basu and Geraint Jones. All the subsequent generations of Ibbetsonians, Pascaline Dupas and Kudzai Takavarasha, Karine Serfaty, Gero Jung, Isabel Neto, Sylvain Chassang, Antara Dutta, and all the others. Ibbetson has been a great house, not because of my spacious and luxurious room, but thanks to the great people who have been living there over the years. And I wish to thank all the extended Ibbetson family, Raphael Auer, Veronica Rappoport and Daniel Paravisini, Brice Tariel, Paula Bustos, Nancy Qian and Ashley Lester, and Matilde Bombardini. In the last two years, I have had to spend more time in my office than I wished for, and I am grateful to Francesco Franco, Emmanuel Farhi and Shawn Cole for alleviating the pain, helping me out with a beer during the ball games, or with a line of code when I got stuck in Stata. And thanks to David Sraer and Nicolas Coeurdacier for their help all along, whether at dusk on a steep freezing couloir in the deep snow with a dislocated shoulder, or any other time when I got dreadful doubts about my thesis.

For those marvelous nights I spent OnZi-r, and for all the rest, I am grateful to my friend Gab. I could trust him all along to be up in the middle of the east coast night, chatting with me. For never letting me down, I am grateful to Anaïs. Eric Oswald gave me back the taste and smell of Istanbul. François' tongue in cheek cynicism gave me the necessary distance from all this mess. I also want to thank my beloved family, my brothers Matthieu and Alain, the crazy Boulcquiens Billy, Prune and Claire, my parents, Eric and Claude. They have always been here for me, they have given me their love and trust. And as much as to a friend, I am grateful to Chauplane. Part of my heart is anchored in this house. The most substantial part of my thesis was written within its thick walls, in front of the mighty Veymont.

Finally, this thesis would never have existed without my love, Marianne. She has been on my side all along, through all the ups and downs, those manic-depressive states of a thesis writer. She encouraged me when I doubted, and she never (seemingly) got bored of listening to my ideas. I can never repay her those years spent away from one another.

New York

16 August, 2005

Contents

1	Distorted gravity: Heterogeneous Firms, Market Structure and the Geography of International Trade	13
1.1	Introduction	14
1.2	A simple model of trade with heterogeneous firms	20
1.2.1	Set-up	20
1.2.2	Trade with heterogeneous firms	23
1.2.3	Intensive versus extensive margin	27
1.3	Estimating distorted gravity equations	31
1.3.1	Data	32
1.3.2	Firm heterogeneity distorts gravity	37
1.3.3	Market structure distorts gravity	39
1.4	Conclusion	42
2	Liquidity Constrained Exporters	49
2.1	Introduction	50
2.2	A model of trade with liquidity constrained exporters	58
2.2.1	Demand	59
2.2.2	Production and trade	60
2.2.3	Liquidity constraints	62
2.2.4	Open economy equilibrium	64
2.2.5	Liquidity constraints and missing trade	69
2.3	The ambiguous impact of exchange rate shocks on trade flows	70

2.4	Conclusion	75
2.5	Appendix	81
2.5.1	Proof: monotonicity of $\bar{x}(A, C_d)$	81
2.5.2	Proof of proposition 5	82
2.5.3	Proof of proposition 7	82
2.5.4	Proof of proposition 9	83
3	The Dynamic Impact of Trade Opening: Productivity Overshooting with Heterogeneous Firms	85
3.1	Introduction	86
3.2	A simple model of trade with heterogeneous firms	88
3.2.1	Demand	88
3.2.2	Production and trade	89
3.2.3	Entry and exit of firms	91
3.2.4	Autarky steady state	93
3.2.5	Trade steady state	94
3.3	The dynamics of trade opening: productivity overshooting	95
3.3.1	Transitional dynamics	96
3.3.2	Productivity overshooting	101
3.3.3	Price compression and inequality	104
3.4	Conclusion	105
3.5	Appendix	108
3.5.1	Non stationary productivity distribution and dynamic adjustments	108
3.5.2	Criterion 13 and proofs of propositions 14, 15 and 16	109
3.5.3	Expected opening to trade	112

List of Figures

- 1-1 σ magnifies the impact of trade barriers on the intensive margin ($B' > B$), whereas it dampens the impact on the extensive margin ($C' < C$). 30
- 2-1 Liquidity constrained exporters. Note: No firm with a productivity below \bar{x}_f can profitably export. On top of this this, firms below the curve $\bar{x}(A)$ are liquidity constrained. All firms in the area Ω between \bar{x}_f and $\bar{x}(A)$ are liquidity constrained exporters. They could profitably export, but are prevented from doing so because of liquidity constraints. 67
- 2-2 The ambiguous impact of an exchange rate appreciation. Note: An appreciation of the domestic currency erodes the competitiveness of exporters, and forces the least productive exporters to exit (light shaded area). At the same time, it relaxes the liquidity constraint, and allows some liquidity constrained firms to enter the export market (dark shaded area). 73

List of Tables

1.1	High and low elasticities of substitution.	37
1.2	Firm heterogeneity distorts gravity.	39
1.3	Market structure distorts gravity.	41
1.4	35 manufacturing sectors.	47

Chapter 1

Distorted gravity: Heterogeneous Firms, Market Structure and the Geography of International Trade

Summary 1 *By considering only the intensive margin of trade, Krugman (1980) predicts that a higher elasticity of substitution between goods magnifies the impact of trade barriers on trade flows. In this chapter, I introduce firm heterogeneity in a simple model of international trade. I prove that the extensive margin, the number of exporters, and intensive margin, the exports per firm, are affected by the elasticity of substitution in exact opposite directions. In sectors with a low elasticity of substitution, the extensive margin is highly sensitive to trade barriers, compared to the intensive margin, and the reverse holds true in sectors with a high elasticity. The extensive margin always dominates, and the predictions of the Krugman model with representative firms are overturned: the impact of trade barriers on trade flows is dampened by the elasticity of substitution, and not magnified.*

To test the predictions of the model, I estimate gravity equations at the sectoral level. The estimated elasticities of trade flows with respect to trade barriers are systematically distorted by the degree of firm heterogeneity and by market structure. These distortions are consistent with the predictions of the model with heterogeneous firms, and reject those of the model with representative firms.

1.1 Introduction

The model most widely used for predicting bilateral trade flows was developed by Paul Krugman in 1980. In this model, identical countries trade differentiated goods despite the presence of trade barriers because consumers have a preference for variety. If goods are more differentiated, consumers are willing to buy foreign varieties even at a higher cost, and trade barriers have little impact on bilateral trade flows. Total exports from country A to country B are given by the following expression:

$$\text{Exports}_{AB} = \text{Constant} \times \frac{\text{GDP}_A \times \text{GDP}_B}{(\text{Trade barriers}_{AB})^\sigma}$$

where σ is the elasticity of substitution between varieties. Trade barriers have a strong impact on trade flows when the elasticity of substitution between goods is high, or when goods are highly substitutable. Competition is fierce when the elasticity of substitution is high, and any cost disadvantage translates into large losses of market share. A crucial assumption in this model is that all firms are identical, and that the only form of transportation cost is a variable cost. Under these assumptions, every firm exports to every country in the world. The amount exported to a given country depends on how competitive it is against other foreign exporters. Differences in competitiveness due to transportation costs have a greater or lesser impact on trade flows depending on whether goods are more or less substitutable.

In this paper, I add firm heterogeneity in productivity, as well as fixed costs associated with exporting. These simple amendments introduce a new margin of adjustment: the extensive margin. When transportation costs vary, not only does each exporter change the size of its exports (the intensive margin), but the set of exporters varies as well (the extensive margin). The main finding of this paper is that the elasticity of substitution has opposite effects on each margin. A higher elasticity makes the intensive margin more sensitive to changes in trade barriers, whereas it makes the extensive margin *less* sensitive. The reason is the following. When trade barriers decrease, new and less productive firms enter the export market, attracted by the potential for higher profits. When the elasticity of substitution is high, a low productivity is a severe disadvantage. These less productive firms can only capture a small market share. The impact of those new entrants on aggregate trade is small. On the other hand, when the

elasticity is low, each firm is sheltered from competition. The new entrants capture a large market share. The impact of those new entrants on aggregate trade is large. So the elasticity of substitution magnifies the sensitivity of the intensive margin to changes in trade barriers, whereas it dampens the sensitivity of the extensive margin.

Which effect dominates? Which margin is the most important? I prove that the effect on the extensive margin dominates. My augmented model predicts that total exports from country A to country B are given by the following expression:

$$\text{Exports}_{AB} = \text{Constant} \times \frac{\text{GDP}_A \times \text{GDP}_B}{(\text{Trade barriers}_{AB})^\zeta} \text{ with } \zeta'(\sigma) < 0$$

The elasticity of aggregate trade with respect to trade barriers, ζ , is negatively related to the elasticity of substitution, σ . I find strong support for this prediction in the data. The elasticity of substitution systematically dampens the impact of trade barriers on trade flows.

My model with heterogeneous firms also predicts that the same trade barriers will have a larger impact on trade flows than in the model with representative firms. In addition to the adjustment of the intensive margin of trade described in existing models, there are important adjustments of the extensive margin. When trade barriers decrease, each firm exports more. In addition, new firms start exporting. The entry of new firms is quantitatively important. Given the observed empirical distribution of firm size, I prove that this effect is large. Calibrating the model on the actual distribution of firm size in the US, the elasticity of trade with respect to trade barriers will be twice as large in a model with heterogeneous firms as in a model with representative firms. Anderson and van Wincoop (2004) argue that, using existing models of trade, observed trade flows are consistent with average trade barriers between the US and Canada equivalent to a 46% tariff (table 7 p. 717, Anderson and van Wincoop's results with $\sigma = 8$). This number is unrealistically large. 46% is the punitive tariff imposed by the US on exports from Laos¹. If my model with heterogeneous firms were the correct model underlying observed trade flows, I would infer from the same data, and assuming the same elasticity, that trade barriers between the US and Canada are equivalent to a 21% tariff. This calibration

¹Along with Cuba and North Korea, Laos is the only country that has not been offered normal trade relations with the US.

exercise puts trade barriers back into a more plausible range².

The prediction, that the effect of trade barriers on trade flows is magnified by the elasticity of substitution, is not specific to Krugman's model of trade. Obstfeld and Rogoff (2000) explain the six major puzzles in International Macroeconomics by the existence of trade barriers. The simple model they spell out to illustrate how plausible values for trade barriers can have a large impact on trade flows relies on the magnification by the elasticity of substitution. Anderson (1979) presents a theoretical foundation for the gravity equation based on trade in goods differentiated by country of origin, and consumers with CES preferences. Deardorff (1995) derives predictions equivalent to the gravity equations of trade from a Heckscher-Ohlin model. All these models find that a higher elasticity of substitution magnifies the effect of trade barriers on trade flows, without the need for increasing returns or monopolistic competition. All that is needed is some degree of specialization between countries, and CES preferences to ensure that the elasticity of substitution is a well defined concept. Since these models implicitly or explicitly assume that firms are identical, they can only describe how each firm, or a representative firm, adjusts its exports decision depending on trade barriers and the structure of demand. In such a framework, it is natural that the effect of trade barriers should be magnified by the degree of substitutability between goods. A notable exception is Eaton and Kortum (2002). Even though they have a CES structure for demand, they predict that the sensitivity of trade flows to trade barriers does not depend on the elasticity of substitution, but on parameters shaping the distribution of comparative advantages. Although this prediction is derived from different foundations, it is similar to ours.

The main contribution of this paper is to introduce the extensive margin of trade in a simple framework, and to prove that the elasticity of substitution dampens the effect of trade barriers on the extensive margin. The dampening effect on the extensive margin always dominates the magnifying effect on the intensive margin. I find strong support for this prediction in the data. This sheds a new light on many interpretations of the effect of trade barriers. The elasticity of trade flows with respect to trade barriers remains large in my model. Once the extensive

²In order to generate an extensive margin, I need to introduce fixed costs on top of those variable costs. Total trade barriers will be larger than the simple estimate of variable costs. However, since I impose that domestic firms also have to pay this fixed cost, I can safely describe these estimated variable costs separately from fixed costs. Anderson and van Wincoop (2004) estimate those fixed costs separately.

margin is considered, it is even larger than what traditional models would predict. However, it is not equal to the elasticity of substitution, and is even inversely related to the elasticity of substitution.

In the remaining of this introduction, I review previous work related to this issue. First, there is a growing body of research linking firm heterogeneity and international trade, both theoretically and empirically. Second, the interaction between market structure and the patterns of trade has long been a central piece in explaining the patterns of international trade. Finally, my empirical procedure is based on the vast literature on estimating gravity equations in international trade.

Melitz (2003) pioneered the study of firm heterogeneity in international trade in a general equilibrium framework. He describes the reallocation of firms within a sector between local and foreign markets triggered by trade opening and extends the classical model of trade with monopolistic competition developed by Krugman (1980) to allow for firm heterogeneity. I expand Melitz's model in the following way. I consider a world with many countries. I then study the strategic choice of firms to export or not, and if they export, which countries to target. I embed my model in a global equilibrium. Such a model generates predictions for the structure of bilateral trade flows. I can pin down exactly which firm from which country is able to enter a given market, and how it is affected by competition from local and other foreign firms. The presence of fixed costs associated with entering foreign markets provides a simple foundation for the extensive margin of trade³. Ruhl (2003) incorporates firm heterogeneity à la Melitz in a dynamic setting in order to explain the so called elasticity puzzle: he argues that in response to high frequency transitory shocks, only the intensive margin adjusts, whereas in response to permanent shocks such as trade liberalization, both the intensive and the extensive margin adjust. This provides an explanation for the difference between the low elasticity needed to explain the patterns of International Business Cycles, and the high elasticity needed to explain the growth of trade following a tariff reduction.

³In a parallel and independent work, Helpman, Melitz and Rubinstein (2004) develop a similar model with heterogeneous firms and fixed costs of accessing each foreign market. Using bounded support for the productivity shocks, they generate predictions for the extensive margin of trade. They can make use of the information contained in the zeros of the trade tables and improve on the traditional gravity regressions. Eaton, Kortum and Kramarz (2004, unpublished) also develop a similar model to calibrate firm level data on French exporters. None of those make any prediction on the impact of market structure on the geography of trade flows.

An alternative approach has been developed in Eaton and Kortum (2002). They propose a Ricardian model of international trade in the spirit of Dornbush, Fischer and Samuelson (1977), with many countries. Trade flows are determined by comparative advantages arising from productivity differences. Firm level productivity differences directly shape the patterns of international trade. Eaton, Kortum and Kramarz (2004a and 2004b) use this theoretical framework to analyze firm level trade data on French firms. They find that the number of exporters is a crucial variable of adjustment. Aggregate trade flows are mostly determined by the number of French firms, rather than by the amount exported by each individual French producer. My model is similar in spirit to the model they build. However, it presents the advantage of greater tractability and greater flexibility. I get simple closed form solutions for aggregate trade, and more importantly, for the intensive margin and the extensive margin separately. I get clear predictions for the impact of both variable and fixed costs on each margin, and for the interaction between these margins and measures of market structure and firm heterogeneity.

Bernard and Jensen (1999, 2001a, 2001b, 2002), Harrigan (1995), Tybout (2003), Bernard, Eaton, Jensen and Kortum (2003), Eaton, Kortum and Kramarz (2004a and 2004b) describe a series of stylized facts on firm level trade. Only a few firms export. Among exporters, only a few firms export to more than a few countries. Most exporters only sell a small fraction of their output abroad. Exporters are different from non exporters in many respects. They are much larger, and they tend to be more productive as well as more capital intensive. Are exporters more productive because they export, or do they export because they are more productive? This question is still a matter for debate. Bernard and Jensen (2001a) find the strongest support for the latter: the best predictor of whether a firm will export tomorrow is its productivity today. There is also significant evidence for the presence of sunk costs associated with exporting as well. A firm exporting today is 36% more likely to export in the future than a firm not exporting today. Anecdotal evidence collected from entrepreneurs also suggests that a large fraction of the costs associated with exporting take the form of fixed or sunk costs. My model matches most of the stylized facts on firm level trade described in this literature.

The importance of market structure in shaping trade flows has long been acknowledged. The path-breaking model of trade by Krugman (1980) explains the existence of intra-industry

trade by the mere presence of product differentiation and monopolistic competition. However, relatively little attention has been given to the difference in the patterns of trade across sectors. Davis (1998) points out that the home market hypothesis⁴ hangs on differentiated goods with scale economies having greater trade costs than homogeneous goods. Hummels (2001) performs a precise analysis of the impact of trade barriers on trade flows at the sectoral level. Anderson and van Wincoop (2004) extract results from Hummels and show that trade costs are more responsive to distance in sectors with more differentiated goods. The closest to my empirical findings is Rauch (1999). He finds that trade barriers have a lower impact on trade volumes when trade is done on organized markets or when reference prices exist. He argues that those goods are more homogeneous. Differentiated goods on the other hand are harder to compare, and it would be difficult for a trader to quote a single price for them. The explanation put forward by Rauch is that the cost of acquiring information about differentiated goods is high. Therefore differentiated goods are more costly to trade. Yet whether or not a good has a reference price is not a direct measure of the degree of differentiation of a good. Moreover, such a reasoning cannot directly explain why each extra mile has a larger impact in sectors with differentiated goods. I focus my empirical analysis on direct measures of product differentiation. I offer an alternative explanation for the interaction between product differentiation and trade barriers. I spell out a clear theoretical channel through which product differentiation affects trade barriers.

The interaction between market structure and the equilibrium distribution of firm size has been studied in a different context in the Industrial Organization literature. Syverson (2004) describes the effect of product substitutability on the selection of firms and the equilibrium dispersion of firm productivity. When products become more substitutable, production within an industry is reallocated. Less productive firms disappear, and output shifts towards the most productive firms. Syverson finds strong evidence that a higher degree of substitutability leads to a narrower productivity dispersion (less productive firms disappear), and a higher median productivity (output shifts towards the most productive firms).

I build my empirical analysis upon the large literature regarding estimating gravity equations in trade, founded by Tinbergen (1962). Harrigan (2002), Evenett and Keller (2002) and Feenstra (2003) offer recent surveys of the existing theories behind the gravity equations. Dear-

⁴The home market hypothesis is the fact that big countries produce more goods with scale economies.

dorff (1995) gives an elegant derivation of the gravity equations in a neoclassical framework. Anderson and van Wincoop (2003) provide generalizations of the theoretically founded gravity equations, and explain the so-called border effect with a well specified model. I augment traditional estimations of the gravity equations to include measures of the interaction between market structure and trade barriers. I find strong support for the predictions of my model with heterogeneous firms in the data.

The remainder of the paper is organized as follows. Section 1.2 introduces a simple model of trade with heterogeneous firms, and generates a series of testable predictions. Section 1.3 describes the empirical estimation and tests of the main predictions of the model. Section 1.4 concludes.

1.2 A simple model of trade with heterogeneous firms

In the next three sections, I develop a theoretical model of trade with heterogeneous firms. In section 1.2.1, I present partial equilibrium results. In section 1.2.2, I compute the equilibrium of the world economy. Finally, in section 1.2.3, I identify separately the adjustments of each margin of trade, in response to changes in both variable and fixed trade barriers.

1.2.1 Set-up

In this section, I introduce the basic ingredients of the model. I define preferences and technologies, and I characterize the optimal strategies of both firms and consumers in partial equilibrium.

There are N countries that produce goods using only labor. Country n has a population L_n . There are $H + 1$ sectors. Sector 0 provides a single homogeneous good that can be freely traded. It is produced under constant returns to scale with unit labor requirement. This homogeneous good is used as the numeraire. Its price is set equal to 1 so that if every country produces this good, then in every country wages are equalized to 1. I shall only consider equilibria where this assumption holds. The other H sectors supply a continuum of differentiated goods. Each firm is a monopolist for the variety it produces.

Preferences: The workers are the only consumers, each endowed with one unit of labor. They all share the same CES preferences over the $H + 1$ groups of goods. A consumer that receives q_o units of the homogeneous good, q_h^x units of each variety x of good h , and varieties of good h in the set X_h (to be determined in equilibrium) gets a utility U :

$$U \equiv q_o^{\mu_0} \prod_{h=1}^H \left(\int_{X_h} (q_h^x)^{\frac{\sigma_h-1}{\sigma_h}} dx \right)^{\frac{\sigma_h}{\sigma_h-1} \mu_h}$$

with $\mu_0 + \sum_{h=1}^H \mu_h = 1$ and $\sigma_h > 1$

where σ_h is the elasticity of substitution between two varieties in sector h .

Trade barriers: There are two types of trade barriers, a fixed cost and a variable cost. If a firm in country i in sector h exports to country j , it must pay a fixed cost C_{ij}^h . The variable cost takes the form of an "iceberg" transportation cost. If one unit of any differentiated good h is shipped from country i to country j , only a fraction $1/\tau_{ij}^h$ arrives. The rest melts on the way. The higher τ , the higher the variable trade cost⁵.

Strategies and equilibrium definition: Each firm in every country chooses a strategy, taking the strategies of other firms and all consumers as given. A strategy for a firm is both a subset of countries where it sells its output, and prices it sets for its good in each market. A strategy for a consumer is the quantity consumed of each variety of every good available domestically, given its price⁶. From the optimal strategies of each firm and each consumer in every country, I can compute the world equilibrium. It is the set of prices and quantities that correspond to a fixed point of the best response graph of each agent worldwide.

Production and pricing: All countries have access to the same technology. Due to the presence of fixed costs, firms in the differentiated sectors operate under increasing returns to scale technology. Each firm in sector h draws a random unit labor productivity x . The cost of producing q units of good and selling them in country j for a firm with productivity x is: $c(q) = q/x + C_{ij}^h$. Firms are price setters. Given that demand functions are isoelastic, the

⁵ $\tau_{ij}^h > 1$ for any $i \neq j$ and $\tau_{ii}^h = 1$. I also impose a triangular inequality to prevent transportation arbitrages: $\forall (i, j, k), \tau_{ik} \leq \tau_{ij} \times \tau_{jk}$.

⁶To prevent arbitrage by consumers, we implicitly assume that consumers in j who try and buy varieties in i would have to pay a fixed cost higher than potential exporters in i . Trade is done by firms, and not by consumers.

optimal price charged in country j by firm x from country i is a constant mark-up over the unit cost (including transportation costs): $p_{ij}^h(x) = \frac{\sigma_h}{\sigma_h - 1} \times \frac{\tau_{ij}^h}{x}$.⁷

Firm heterogeneity: For simplicity and as in Melitz (2001) and Helpman, Melitz and Yeaple (2004), I assume the following form for the productivity shocks. It is drawn from a Pareto distribution with scaling parameter γ_h .⁸ This means that productivity is distributed according to $P(\tilde{x} < x) = F_h(x) = 1 - x^{-\gamma_h}$, and $dF_h(x) = \gamma_h x^{-\gamma_h - 1} dx$, for $x \geq 1$. γ_h is an inverse measure of the heterogeneity in sector h , with $\gamma_h > 2$ and $\gamma_h > \sigma_h - 1$. Sectors with a high γ are more homogeneous, in the sense that more output is concentrated among the smallest and least productive firms⁹.

I also assume that the total mass of firms in country i in each differentiated sector is proportional to the size of country i , L_i .¹⁰

Demand for differentiated products: Given the optimal pricing of firms and the optimal decision of consumers, exports from country i to country j , by a firm with a labor productivity x , in sector h are:

$$t_{ij}^h(x) = p_{ij}^h(x) q_{ij}^h(x) = \mu_h L_j \left(\frac{p_{ij}^h(x)}{P_j^h} \right)^{1 - \sigma_h}$$

where P_j^h is the price index for good h in country j .

If only those firms above the productivity threshold \bar{x}_{kj}^h in country k and sector h export to country j , the ideal price index for good h in country j , P_j^h , is defined as:

⁷This price prevents any arbitrage either by domestic firms that might want to resell these goods at home or abroad, or by foreign firms.

⁸See Kortum (1997) and Gabaix (1999) for justifications of this distribution. An alternative justification is provided by Eaton and Kortum (2002): if the observed distribution of productivities among firms is the realization of the maxima of a generic distribution for which all moment are not defined, the distribution of the maxima converges in probability towards a Fréchet distribution. A Fréchet distribution with scaling parameter γ approaches a Pareto distribution with scaling parameter γ for x large.

⁹In x has a standard deviation equal to $1/\gamma$. The assumption $\gamma > \sigma - 1$ ensures that, in equilibrium, the size distribution of firms has a finite mean. If this assumption were violated, firms with an arbitrarily high productivity would represent an arbitrarily large fraction of all firms, and they would overshadow less productive firms. Results on selection into export markets would be degenerate. This assumption is satisfied in the data for almost all sectors.

¹⁰Implicitly, I assume that there is a group of entrepreneurs proportional to the size of the country. I could remove this assumption, and allow for the free entry of entrepreneurs, with an infinite set of potential entrepreneurs. Provided that trade barriers are not negligible, I would get qualitatively the same results.

$$P_j^h = \left(\sum_{k=1}^N L_k \int_{\bar{x}_{kj}^h}^{\infty} \left(\frac{\sigma_h - 1}{\sigma_h} \times \frac{x}{\tau_{kj}^h} \right)^{\sigma_h - 1} dF(x) \right)^{-1/(\sigma_h - 1)}$$

For now, I will consider only sector h . The other sectors are analogous. For notational clarity, I drop the h subscript and all sectoral variables will refer to sector h when there is no ambiguity.

1.2.2 Trade with heterogeneous firms

In this section, I compute the global equilibrium of this world economy. To do so, I define the selection of firms into the export market. I give predictions for aggregate bilateral trade flows.

If firms are heterogeneous and if there are fixed costs for entering foreign markets, there will be selection among exporters. Less productive firms are not able to generate enough profits abroad to cover the fixed cost of entering foreign markets. Exporters are therefore only a subset of domestic firms. The subset of exporters varies with the characteristics of the foreign market.

Productivity threshold: As long as net profits generated by exports in a given country are sufficient to cover the fixed entry cost, a firm will be willing to export there. The profits firm x in i earns from exporting to j are $\pi_{ij}(x) = \frac{\mu}{\sigma} \left(\frac{\sigma-1}{\sigma} \right)^{\sigma-1} \left(P_j^{\sigma-1} L_j \right) \left(\frac{x}{\tau_{ij}} \right)^{\sigma-1} - C_{ij}$. Call \bar{x}_{ij} the productivity threshold for the least productive firm in country i able to export to country j . \bar{x}_{ij} corresponds to the productivity of a firm in country i for which gross profits earned in country j are just enough to cover the fixed cost of entering market j :

$$\begin{aligned} \pi_{ij}(\bar{x}_{ij}) &= 0 \\ \Rightarrow \bar{x}_{ij} &= \lambda_1 C_{ij}^{\frac{1}{\sigma-1}} \left(P_j^{\sigma-1} L_j \right)^{\frac{-1}{\sigma-1}} \tau_{ij} \end{aligned} \quad (1.1)$$

with λ_1 a constant ¹¹. I assume that trade barriers are always high enough to ensure that $\forall k, l$, $\bar{x}_{kl} > 1$.

¹¹ $\lambda_1 = \left(\frac{\sigma}{\mu} \right)^{\frac{1}{\sigma-1}} \left(\frac{\sigma}{\sigma-1} \right)$.

Price indices: Until now, I have considered prices as given. However they do adjust depending on country characteristics. I now know exactly the set of firms that export to country j . This set only depends on country j 's characteristics. By definition, the price index in country j is given by $P_j^{1-\sigma} = \sum_{k=1}^N L_k \int_{\bar{x}_{kj}}^{\infty} \left(\frac{\sigma-1}{\sigma} \times \frac{\tau_{kj}}{x}\right)^{1-\sigma} dF(x)$. Plugging in the productivity thresholds from Eq. (1.1), I can solve for the equilibrium price index:

$$P_j = \lambda_2 \times \left(\frac{L_j}{L}\right)^{\frac{1}{\gamma}} \times L_j^{-\frac{1}{\sigma-1}} \times \theta_j \quad (1.2)$$

$$\text{with } \theta_j^{-\gamma} \equiv \sum_{k=1}^N s_k \times \tau_{kj}^{-\gamma} \times C_{kj}^{-\left(\frac{\gamma}{\sigma-1}-1\right)}, \quad s_k \equiv \frac{L_k}{L} \text{ and } L \equiv \sum_{k=1}^N L_k$$

λ_2 being a constant¹².

θ_j is a aggregate index of j 's remoteness from the rest of the world¹³. It is similar to the "multilateral resistance variable" introduced by Anderson and van Wincoop (2003). In addition to their measure, it takes into account the impact of fixed costs and of firm heterogeneity on prices.

Equilibrium exports and threshold: Exports by an individual firm depend on its productivity, the trade barriers it must overcome, and the prices set by its competitors. I have solved for the price indices in every country. By plugging the general equilibrium price index from Eq. (1.2) into the demand function, and into the productivity threshold from Eq. (1.1), I can solve for firm level exports and the productivity threshold.

In general equilibrium, exports $t_{ij}(x)$ from country i to country j by an individual firm with productivity x , and the productivity threshold \bar{x}_{ij} above which firms in i export to j , are given by:

$$t_{ij}(x|x \geq \bar{x}_{ij}) = \lambda_3 \times \left(\frac{L_j}{L}\right)^{\frac{\sigma-1}{\gamma}} \times \left(\frac{\theta_j}{\tau_{ij}}\right)^{\sigma-1} \times x^{\sigma-1} \quad (1.3)$$

$$\bar{x}_{ij} = \lambda_4 \times \left(\frac{L}{L_j}\right)^{\frac{1}{\gamma}} \times \left(\frac{\tau_{ij}}{\theta_j}\right) \times C_{ij}^{\frac{1}{\sigma-1}} \quad (1.4)$$

¹² $\lambda_2 = \left(\frac{\gamma-(\sigma-1)}{\gamma}\right)^{1/\gamma} \left(\frac{\sigma}{\mu}\right)^{1/(\sigma-1)-1/\gamma} \left(\frac{\sigma}{\sigma-1}\right)$.

¹³A simple way to interpret this aggregate index is to look at a symmetrical case: when $\tau_{kj} = \tau_j$ and $C_{ij} = C_j$ for all k 's, $\theta_j = C_j^{\frac{1}{\sigma-1}-\frac{1}{\gamma}} \times \tau_j$. In asymmetric cases, θ_j is a weighted average of bilateral trade barriers.

with λ_3 and λ_4 constants¹⁴. They are functions of the size L_j , the trade barriers C_{ij} and τ_{ij} , and the measure of j 's remoteness from the rest of the world, θ_j .

As expected from this simple monopolistic competition model, exports by individual firms depend on the transportation cost τ_{ij} with an elasticity $\sigma - 1$. Individual firm's exports depend on the size of the destination market L_j with an elasticity less than one, because of the impact of market size on the degree of price competition. This equation is very similar to what a traditional model of trade with representative firms would predict for aggregate trade flows. In contrast, in my model with firm heterogeneity, because of the selection into the export market, aggregate trade will look radically different.

Proposition 2 (aggregate trade) *Total exports (f.o.b.) T_{ij}^h in sector h from country i to country j are given by:*

$$T_{ij}^h = \mu_h \times \frac{L_i L_j}{L} \times \left(\frac{\tau_{ij}^h}{\theta_j^h} \right)^{-\gamma_h} \times \left(C_{ij}^h \right)^{-\left(\frac{\gamma_h}{\sigma_h - 1} - 1 \right)} \quad (1.5)$$

*Exports are a function of the sizes L_i and L_j , the bilateral variable cost τ_{ij}^h , the bilateral fixed cost C_{ij} , and the measure of j 's remoteness from the rest of the world, θ_j^h .*¹⁵

Proof. By definition, aggregate exports are the sum of exports by all exporting firms: $T_{ij}^h = \int_{\bar{x}_{ij}^h}^{\infty} t_{ij}^h(x) L_i dF(x)$. Using Eq. (1.4) for the productivity threshold, Eq. (1.3) for individual firms' exports, and the specific Pareto distribution for the productivity shocks, I solve for aggregate trade. ■

The gravity structure of trade has been dramatically distorted by the presence of firm heterogeneity.

First note that the elasticity of exports with respect to variable trade barriers, γ_h , is larger than in the absence of firm heterogeneity, and larger than the elasticity for each individual

¹⁴ $\lambda_3 = \mu \left(\frac{\gamma - (\sigma - 1)}{\gamma} \right)^{\frac{\sigma - 1}{\gamma}} \left(\frac{\sigma}{\mu} \right)^{1 - \frac{\sigma - 1}{\gamma}}$ and $\lambda_4 = \left(\frac{\sigma}{\mu} \times \frac{\gamma}{\gamma - (\sigma - 1)} \right)^{\frac{1}{\gamma}}$.

¹⁵ Interestingly, note that the ratio of i 's market share in k , and j 's market share in k , only depends on the ratio of i 's trade barriers and j 's trade barriers. If I define the composite measure of trade barriers $\varphi_{ik} = \tau_{ik}^{-\gamma} \times C_{ik}^{-\left(\frac{\gamma}{\sigma - 1} - 1 \right)}$, I get: $\frac{T_{ik}/L_i}{T_{jk}/L_j} = \frac{\varphi_{ik}}{\varphi_{jk}}$. Similarly, i 's market share in k only depends on trade barriers from i relative to trade barriers from other countries: $\frac{T_{ik}}{L_k} = \frac{s_i \varphi_{ik}}{\sum_j s_j \varphi_{jk}}$.

firm (both equal to $\sigma_h - 1$). An increase in variable costs not only causes a reduction in the size of exports of each exporter, but it also forces some exporters to pull out. The extensive margin comes on top of the intensive margin and amplifies the impact of variable costs. This amplification effect is quantitatively important. Anderson and van Wincoop (2004) argue that if one assumes that trade is governed by an underlying model of trade with identical firms, trade barriers between the US and Canada must be equivalent to a 46% tariff in order to explain the observed bilateral trade flows (table 7 p. 717, Anderson, van Wincoop's results with $\sigma = 8$). This number is an indirect measure. It depends crucially on what assumption is made about the underlying trade model. If my model were correct, and using an average heterogeneity parameter estimated from firm level data ($\frac{\gamma}{\sigma-1} \approx 2$), I would infer from the same trade volume data that trade barriers are equivalent to a 21% tariff ($1.21 = \sqrt{1.46}$). This is far below their 46% estimate.

Second, the elasticity of exports with respect to transportation costs depends on the degree of firm heterogeneity, γ_h . In more homogeneous sectors (γ_h high) large productive firms represent a smaller fraction of firms. The productivity threshold moves in a region where most of the mass of firms lies. In those sectors, aggregate exports are sensitive to changes in transportation costs because many firms exit and enter when variable costs fluctuate.

Third and most importantly, the elasticity of exports with respect to variable costs *does not* depend at all on the elasticity of substitution between goods, σ_h ¹⁶, and the elasticity of exports with respect to fixed costs is *negatively* related to the elasticity σ_h . This prediction is in stark contrast with models with representative firms. In such models, the elasticity of exports with respect to transportation costs would be equal to $\sigma_h - 1$. In the following section, I will examine how σ_h has exact opposite effects on the intensive and the extensive margins of trade.

¹⁶Eaton and Kortum (2002) derive a similar prediction from a different set-up. In a Ricardian model of trade, they find that bilateral trade flows do not depend on the elasticity of substitution between goods, but only on the scaling parameter of the underlying distribution of productivity shocks. They use Fréchet distributions, which approach Pareto distributions in their right tails: the distribution for shocks they consider is $1 - F(z) = 1 - e^{-Tz^{-\theta}} \underset{z \rightarrow +\infty}{=} Tz^{-\theta} + o(z^{-\theta})$. In equilibrium, they predict that the elasticity of trade flows with respect to trade barriers (variable only) is equal to θ .

1.2.3 Intensive versus extensive margin

In this section, I separately examine the intensive and the extensive margins of trade. I describe how the elasticity of substitution magnifies the sensitivity of the intensive margin to trade barriers, and dampens the sensitivity of the extensive margin. I prove that the dampening effect on the extensive margin dominates the magnifying effect on the intensive margin.

Thus far I have seen that after taking firm heterogeneity into consideration, the selection of firms into the export market becomes a key feature of the adjustment of trade flows. This is the extensive margin of trade. The main prediction of the model is that the extensive margin and the intensive margin are affected in opposite directions by the elasticity of substitution. If the elasticity of substitution is high, then the impact of trade barriers on the intensive margin is strong, and mild on the extensive margin. The reverse holds true when the elasticity of substitution is low. The dampening effect of the elasticity of substitution on the extensive margin always dominates the magnifying effect on the intensive margin.

Proposition 3 (intensive and extensive margins) *The elasticity of substitution (σ) has no effect on the elasticity of trade flows with respect to variable trade costs (ζ), and a negative effect on the elasticity of trade flows with respect to fixed costs (ξ):*

$$\text{if } \zeta \equiv -\frac{d \ln T_{ij}}{d \ln \tau_{ij}} \text{ and } \xi \equiv -\frac{d \ln T_{ij}}{d \ln C_{ij}}, \text{ then } \frac{\partial \zeta}{\partial \sigma} = 0 \text{ and } \frac{\partial \xi}{\partial \sigma} < 0$$

Proof. I go into much details to prove this proposition. In doing so, I introduce formally the intensive and the extensive margins of trade. I describe the adjustment of each margin, and the sensitivity of these adjustments to the elasticity of substitution.

The impact of trade barriers, both variable cost and fixed cost, on aggregate trade flows can be decomposed into two different margins. The intensive margin is defined by how much each existing exporter changes the size of its exports. The extensive margin is defined by how much new entrants export (in the case of a reduction in trade barriers).

Differentiating Eq. (1.5), I get the following expressions for each margin¹⁷:

$$d\Gamma_{ij} = \underbrace{\left(\int_{\bar{x}_{ij}}^{\infty} \frac{\partial t_{ij}(x)}{\partial \tau_{ij}} f(x) dx \right)}_{\text{Intensive margin}} d\tau_{ij} - \underbrace{\left(t(\bar{x}_{ij}) f(\bar{x}_{ij}) \times \frac{\partial \bar{x}_{ij}}{\partial \tau_{ij}} \right)}_{\text{Extensive margin}} d\tau_{ij} \\ + \underbrace{\left(\int_{\bar{x}_{ij}}^{\infty} \frac{\partial t_{ij}(x)}{\partial C_{ij}} f(x) dx \right)}_{\text{Intensive margin}} dC_{ij} - \underbrace{\left(t(\bar{x}_{ij}) f(\bar{x}_{ij}) \times \frac{\partial \bar{x}_{ij}}{\partial C_{ij}} \right)}_{\text{Extensive margin}} dC_{ij}$$

Following a reduction of trade barriers, each existing exporter (all $x > \bar{x}_{ij}$) exports more. This is the intensive margin. At the same time, higher potential profits attract new entrants (\bar{x}_{ij} goes down). This is the extensive margin.

In elasticity notations, I get the following expression for each margin for changes in the variable cost, τ_{ij} :

$$\zeta = \underbrace{(\sigma - 1)}_{\substack{\text{Intensive margin} \\ \text{Elasticity}}} + \underbrace{(\gamma - (\sigma - 1))}_{\substack{\text{Extensive margin} \\ \text{Elasticity}}} = \gamma$$

σ magnifies the intensive margin when variable costs move ($\sigma - 1$ increases with σ), whereas it dampens the extensive margin ($\gamma - (\sigma - 1)$ decreases with σ)¹⁸. The effect of σ on each margin cancels out, so that:

$$\frac{\partial \zeta}{\partial \sigma} = 0$$

In elasticity notation, I get the following expression for each margin for changes in the fixed costs, C_{ij} :

$$\xi = \underbrace{0}_{\substack{\text{Intensive margin} \\ \text{Elasticity}}} + \underbrace{\left(\frac{\gamma}{\sigma - 1} - 1 \right)}_{\substack{\text{Extensive margin} \\ \text{Elasticity}}} = \frac{\gamma}{\sigma - 1} - 1$$

σ has no impact on the intensive margin when fixed costs move, whereas it dampens the impact on the extensive margin ($\frac{\gamma}{\sigma - 1} - 1$ decreases with σ). The impact of σ on the elasticity of trade

¹⁷I use Leibniz rule to separate the intensive from the extensive margin. I apply Lebesgue's monotone convergence theorem to ensure the existence of and to compute the intensive margin.

¹⁸I have implicitly assumed that changes in both τ_{ij} and C_{ij} have no significant impact on the general equilibrium. That is, I have assumed that $\frac{\partial \theta_i}{\partial \tau_{ij}} = \frac{\partial \theta_i}{\partial C_{ij}} = 0$. This is a fair approximation as long as country i is not too large compared to the rest of the world (s_{ij} small). Relaxing this assumption would reinforce my results, but it would make calculations cumbersome.

flows with respect to fixed costs is always negative:

$$\frac{\partial \xi}{\partial \sigma} < 0$$

■

The intuition for these results is the following. When goods are substitutable, the demand for each individual variety is highly sensitive to changes in trade costs. In other words, when σ is high, the intensive margin of trade is strongly affected by trade barriers. This margin is the only one in the Krugman model of trade with representative firms.

The interaction between the elasticity of substitution and the extensive margin is more complex. When σ is low, the market share that each firm is able to capture is relatively insensitive to differences in productivity. Less productive firms are still able to get a relatively large market share, despite having to charge a higher price than other firms. In the limiting case of a Cobb-Douglas ($\sigma = 1$), differences in productivity have no effect on the market share of each firm. As trade barriers decrease, some firms with a low level of productivity are able to enter. When goods are highly differentiated, these new entrants are relatively large compared to the firms that are already exporting. Therefore the extensive margin is strongly affected by trade barriers when σ is low. The reverse holds when σ is high.

I can describe the impact of a decrease of trade barriers on both the intensive and the extensive margin of trade graphically. This is illustrated on Figure 1-1. On this graph, quantities exported by each firm are represented for two sectors, one where goods are differentiated (σ_{low}), and one where goods are easily substitutable (σ_{high}). Aggregate trade is the sum of exports of all firms with a productivity above the productivity threshold \bar{x} . It is represented graphically by area A for σ_{low} , and A' for σ_{high} .

When variable trade barriers go down, each firm is able to export a larger volume. The density of exports shifts up. This is the intensive margin of trade. With σ_{low} , each exporter only increases its exports moderately. With σ_{high} on the other hand, the cost advantage from lower trade barriers allows exporters to capture a large market share in the foreign market, and each exporter increases its exports substantially. Aggregate trade is increased by the area B for σ_{low} , and B' for σ_{high} , with $B < B'$. The higher the elasticity of substitution σ , the more

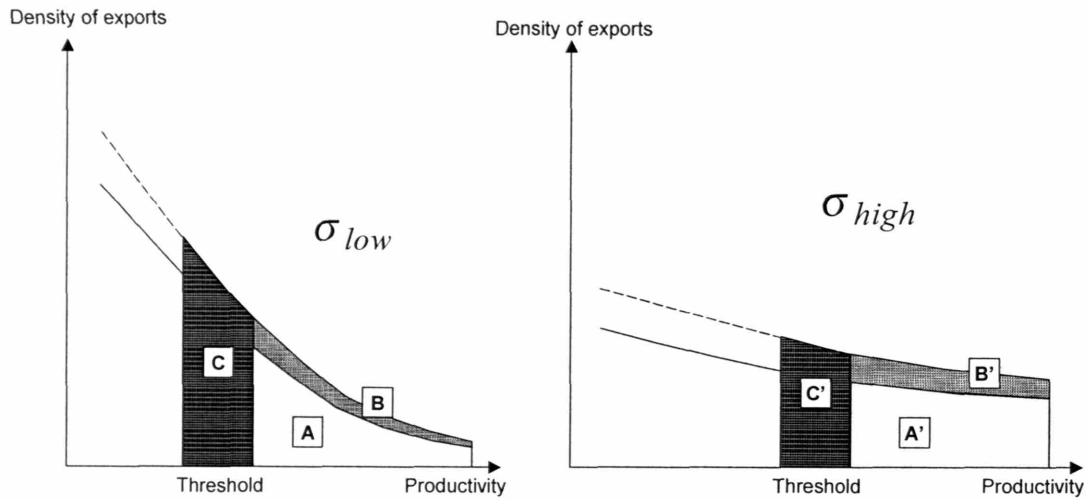


Figure 1-1: σ magnifies the impact of trade barriers on the intensive margin ($B' > B$), whereas it dampens the impact on the extensive margin ($C' < C$).

sensitive the intensive margin.

In addition, following a decrease in variable trade barriers, new exporters are able to enter the export market. These new entrants are firms with a productivity below the initial productivity threshold. The productivity threshold shifts to the left. This represents the extensive margin of trade. With σ_{low} , the new entrants, despite their lower productivity, capture a large market share in the foreign market. Total exports by new exporters are large. On the other hand, with σ_{high} , new entrants capture only a small market share in the foreign market. This is because their lower productivity is a severe handicap in this highly competitive environment. Total exports by new entrants are small. Aggregate trade is increased by the area C for σ_{low} , and C' for σ_{high} , with $C > C'$. The lower the elasticity of substitution σ , the more sensitive the extensive margin.

Adjustments to changes in fixed trade barriers are simpler. The intensive margin does not move in response to a reduction of fixed costs, $B = B' = 0$. However, the extensive margin is affected. In contrast with adjustments to changes in variable trade barriers, not only are new entrants larger with σ_{low} than with σ_{high} , but in addition, the productivity threshold moves

more when σ_{low} than when σ_{high} . $C > C'$, and even more so than in the case of a reduction of variable trade barriers.

I have proven that σ , the elasticity of substitution between goods, has opposite effects on the extensive and intensive margins. Which effect dominates? What is the net effect of σ on aggregate trade? Does a larger σ make aggregate trade flows more sensitive to trade barriers (if the intensive margin effect dominates), or less sensitive (if the extensive margin dominates)? In Figure 1-1, is $B + C$ larger or smaller than $B' + C'$? I have proven in proposition 2 that with Pareto distributed productivity shocks, the effect of σ on the extensive margin always dominates the effect on the intensive margin: $B + C > B' + C'$ if fixed costs are reduced, and $B + C = B' + C'$ if only variable costs are reduced.

In this section I have explained why the elasticity of substitution has exactly opposite effects on the intensive and the extensive margins of trade. A higher elasticity of substitution makes the intensive margin more sensitive to changes in trade barriers, whereas it makes the extensive margin less sensitive. What is the net impact of σ on the two margins? I prove in Proposition 3 that the extensive margin always dominates. Contrary to the predictions of the Krugman model with representative firms, the elasticity of substitution σ always dampens the impact of trade barriers on trade flows.

The next section is devoted to testing the predictions from the model. I find strong support for the heterogeneous firms model, thus rejecting the predictions of the Krugman model with representative firms. The interaction between the elasticity of substitution and the sensitivity of trade flows to trade barriers suggests that the extensive margin plays a crucial role in international trade.

1.3 Estimating distorted gravity equations

In the following sections, I test empirically the predictions of the model with heterogeneous firms against the predictions of the Krugman model with representative firms. If a higher elasticity of substitution dampens the sensitivity of trade flows to trade barriers, the Krugman model with representative firms will be rejected in favor of the model with heterogeneous firms. I find strong support for the model with heterogeneous firms in the data, and reject the Krugman

model with representative firms. This finding is consistent with parallel micro evidence on the importance of firm heterogeneity and the extensive margin in international trade.

1.3.1 Data

In order to test the main prediction of the model, I need data from several sources. I need data on bilateral trade flows, disaggregated at the sector level. I need measures of the degree of heterogeneity between firms within each sector. I need measures of the elasticities of substitution between goods within each sector. Finally, I need measures of trade barriers between trading partners at the sectoral level with either direct measures of trade barriers or proxies for trade barriers.

Bilateral trade flows data

I use bilateral exports data from the World Trade Database and from the World Trade Analyzer. All details from Statistics Canada and NBER preparations are given in Feenstra, Lipsey and Bowen (1997) and Feenstra (2000). Only data for the period 1980-1997 are presented. The results are similar on other time periods. A total of 169 countries are represented. Results are robust and hold when restricting the analysis to different subsets of countries.

Products are disaggregated according to different classification systems, to ensure that the results are robust to changing the definition of sectors. The most disaggregated classification I use corresponds to the 3-digit SITC revision 3. In this classification, I have data on 265 sectors. I also use a much coarser classification with 34 manufacturing sectors only. This classification is based on the US 1987 Standard Industrial Classification (SIC) and corresponds roughly to 3 digits SIC sectors¹⁹.

Sectoral heterogeneity data

The model predicts that the more heterogeneous a sector, the milder the impact of trade barriers on trade flows. In order to test this prediction, I need an estimate of firm heterogeneity.

¹⁹See Table ?? in the appendix for a detailed description of the classification.

Following Melitz, Helpman and Yeaple (2004), I can construct a measure of sectoral heterogeneity by looking directly at the distribution of firm size within sectors. The size distribution of national firms is shaped by the distribution of productivity shocks and the degree of competition. However, the link between the productivity distribution and the size distribution will be more or less distorted by the accessibility of foreign markets.

A firm receiving a random productivity shock \tilde{x} has total sales $S_i(\tilde{x}) = \sum_{j=1}^N p_{ij}(\tilde{x}) q_{ij}(\tilde{x})$. More productive firms are able to capture a larger demand. They are also able to reach more countries. They sell more than other firms not only because they can charge lower prices and capture a larger demand in each market, but also because they have access to more markets. In the case of a small and integrated economy, this selection process will magnify the impact of productivity differences between firms.

In large and rather closed economies however, most firms sell only at home, and exporters sell only a fraction of their output abroad. Access to foreign markets has only a mitigated impact on the size distribution of firms within a sector. This is typically the case of the US economy. Bernard et al. (2003) report that 21% of US manufacturing plants export. Even though those plants are large and account for 60% of the US manufacturing sector, the vast majority of exporters sell no more than 10% of their output abroad. Hence when looking at the entire distribution of firm sizes for an economy like the US, I can safely assume that the size of a firm is almost entirely determined by the size of the domestic market: $S(\tilde{x}) \approx p_{US}(\tilde{x}) q_{US}(\tilde{x}) = \lambda_{US} \times \tilde{x}^{\sigma-1}$ with λ_{US} a US specific term common to all US firms²⁰. The probability that a firm has a size (measured by sales) larger than S is:

$$P_{US}(\tilde{S} > S) \approx P\left(\tilde{x} > \left(\frac{S}{\lambda_{US}}\right)^{1/(\sigma-1)}\right)$$

$$P_{US}(\tilde{S} > S) \approx \lambda_{US}^{\gamma/(\sigma-1)} \times S^{-\gamma/(\sigma-1)}$$

If empirically I have N (large) draws from this distribution, I can estimate the coefficient $\frac{\gamma}{\sigma-1}$ by looking at the rank-size relationship. I order firms according to their size, the largest firm first. Since there are i out of N firms that are larger than the i -th firm, i/N is an estimator of

²⁰ $\lambda_{US} = \mu \left(\frac{\sigma-1}{\sigma}\right)^{\sigma-1} P_{US}^{\sigma-1} L_{US}$.

the probability that a firm has a size larger than Size_i . For a firm i :

$$\ln \left(\text{Rank}_i^h / N_h \right) \approx \ln \left(P \left(\tilde{S} > \text{Size}_i^h \right) \right) = a_h - \frac{\gamma_h}{\sigma_h - 1} \ln \left(\text{Size}_i^h \right) \quad (1.6)$$

Estimating this equation with OLS provides us with an estimator of $\frac{\gamma_h}{\sigma_h - 1}$, the scaling coefficient of the size distribution in sector h ²¹. This measure of sectoral heterogeneity should amplify the impact of transportation costs on bilateral trade flows. That is in sectors where the distribution of the log of firm size has a lower variance ($\frac{\gamma}{\sigma - 1}$ is larger), transportation costs should have a larger negative impact on bilateral trade flows.

In order to measure heterogeneity using the distribution of firm size within a sector, I use data from Compustat on the distribution of sales of all publicly traded companies listed in the US stock markets in the year 1996 (I obtain similar results for the years between 1970 and 1997). Following the guidance of the model, I restrict the sample to US firms only and exclude affiliates of foreign firms. I compute this measure for only the broad 34 manufacturing sectors. This allows us to get information on a large enough number of firms to compute heterogeneity measures for every sector. A finer definition of sectors such as the 3-digit SITC sectors would provide too few datapoints per sector, and would give us unreliable estimates of heterogeneity.

Sectoral elasticity of substitution data

The model also predicts that trade barriers have the largest impact on trade flows in sectors where goods are the most differentiated. It is therefore crucial that I get estimates of sectoral elasticities of substitution. However, the selection that takes place among exporters prevents me from using aggregate sectoral demand elasticities as a measure of the elasticities of substitution. The elasticity of substitution between any two varieties in a given sector, σ_h , is the demand elasticity that one firm in sector h faces. It is not the demand elasticity that the whole sector faces when it exports. I predict that interpreting the elasticity of aggregate exports with respect to trade barriers in gravity-type regressions as a measure of the elasticity of substitution is incorrect. The elasticity of exports with respect to trade barriers *is not* an estimate of the elasticity of substitution in that sector. It is a composite measure of both the sectoral elasticity

²¹See Gabaix and Ioannides (2004) for a discussion of the various procedures and pitfalls to estimate equations such as Eq. (1.6).

of substitution (σ_h) and the sectoral heterogeneity between firms (γ_h). If my model were correct, it should actually be inversely related to the elasticity of substitution. Since almost all empirical studies assume that the Krugman model is correct, and interpret the elasticity of exports with respect to trade barriers to be a measure of the sectoral elasticity of substitution, I cannot use their estimates. A proper estimation must be able to estimate the demand elasticity at the variety level, and not on aggregate trade flows.

For that reason, I use estimates of sectoral elasticities of substitution built by Broda and Weinstein (2004)²². They widely extend the seminal work of Feenstra (1994). Using the panel dimension of data, they use the second moments of demand and supply variations to infer demand and supply elasticities separately²³. Even though they do not account directly for heterogeneity between firms, they use data at a sufficiently fine level of disaggregation to capture most of the heterogeneity between firms. The information they use is about substitution between two extremely narrowly defined subsectors, rather than substitution between imports from one country versus imports from another country. They use price and volume data at the highest level of disaggregation available (10-digit Harmonized System), on consumption by US consumers of imported foreign goods. They estimate how much demand shifts between two (10-digit) varieties when relative prices vary, within each (3-digit) sector. I use their estimates of the elasticities of substitution over the period 1990-2001 for 3-digit SITC (revision 3) sectors. My results also hold over the earlier period they consider, 1972-1988.

It must also be noted that consistent with my model, they do find that a substantial part of the increase in trade flows comes from the extensive margin of trade: the US not only import more of each variety, but they import more and more varieties, from more and more countries.

Trade barriers data

I use both direct and indirect measures of trade barriers. I use several indirect measures of trade costs. The most widely used and those for which I present results here are the bilateral

²²I am immensely indebted to Christian Broda and David Weinstein for providing me with their estimates of elasticities of substitution.

²³In my model, the supply elasticity is infinite: I assume a constant returns to scale technology (for the variable costs), and a sufficiently large labor force so that wages do not respond to demand shocks. It is however important to distinguish empirically between demand and supply shocks, and identify separately the demand and the supply elasticities for each sector.

geographical distance between two trading partners, the fact of sharing a common border, and the fact of sharing a common language. I use great distance circles to measure distance between capital cities from L. Eden, Texas A&M University. I use the CIA world factbook for contiguity and language data. The common language variable is a dummy equal to 1 if both countries have the same official language. I do not use more continuous measures of language proximity, such as those used by Jacques Melitz (2003). Implicitly, I assume that those proxies for trade barriers are correlated with both fixed and variable costs.

I use data on freight and tariff from Hummels (2001) as a direct measure of trade costs²⁴. Data are disaggregated at the 3-digit SITC level. They correspond to the freight and tariff reported by exporters as a fraction of the value of their exports. Unfortunately, those data are incomplete: they only cover trade towards the US.

Descriptive statistics

In Table 1.1, I give summary statistics for a few sectors. I show sectors with distance elasticities of trade among the highest, and sectors with elasticities among the lowest. A model with representative firms would predict that, controlling for the transportation technology, sectors with low distance elasticities of trade should have low elasticities of substitution, and the reverse for sectors with high distance elasticities. I observe exactly the opposite. The sectors where distance has a large impact on trade flows are sectors with elasticities of substitution among the lowest, and sectors where distance has a mild impact on trade flows are sectors with elasticities of substitution among the highest. These patterns cannot be explained by differences in transportation technologies. High distance elasticity sectors are neither sectors where freight rates are large, nor sectors where freight rates are highly sensitive to distance.

This anecdotal evidence suggests that models with representative firms generate incorrect predictions. Introducing firm heterogeneity and the extensive margin of trade provides with an answer for this apparent puzzle.

²⁴I am immensely grateful to David Hummels for providing me with those data and helping me to organize them at the correct level of disaggregation.

Table 1.1: High and low elasticities of substitution.

Sector (3-digit SITC)	Distance elasticity of trade flows	Distance elasticity of trade costs	Average trade cost	Elasticity of substitution
LOW ELASTICITY OF SUBSTITUTION				
Printed Matter	1.5	0.08	9%	2.8
Non alcoholic beverages	1.4	0.15	17%	1.7
Equipment for distributing electricity	1.3	0.1	5%	1.9
HIGH ELASTICITY OF SUBSTITUTION				
Steam and vapor generating boilers	0.6	1.1	8%	12
Road motor vehicles	0.6	0.7	4.5%	19
Pulp and waste paper	0.5	0.08	15%	18
Min < Average < Max	-1.46 < .89 < 1.8	-.7 < .27 < 1.1	1.4% < 10% < 38%	1.1 < 3.9 < 58.5

Source: bilateral trade flows, Feenstra (2000); average freight rate towards the US, Hummels (1999); elasticities of substitution, Broda and Weinstein (2004); data are aggregated at the 3-digit SITC (rev 3) level; year 1997, all countries with GDP/capita above \$3000 (PPP) and population above 1 million.

1.3.2 Firm heterogeneity distorts gravity

In this section, I test whether or not the degree of firm heterogeneity affects the sensitivity of trade flows with respect to trade barriers. I find strong support for the predictions of the model with heterogeneous firms. In sectors where the output is concentrated among a few large firms, trade barriers have a mild impact on trade flows, and they have a strong impact in sectors where small firms account for a larger share of output.

My model predicts that the degree of heterogeneity between firms will affect the sensitivity of trade flows to trade barriers. In heterogeneous sectors, defined as sectors where the largest firms account for a large fraction of output, the selection among exporters takes place among small firms. It does not have much of an impact on aggregate trade flows. On the other hand, in homogeneous sectors, defined as sectors where small firms account for a large fraction of output, the entry and exit of less productive firms has a large impact on aggregate trade.

In order to test this prediction, I run the following equation using OLS:

$$\ln(\text{Exports}_{ij}^h) = B_{ij}^h + X_{ij}^h B_1 + \left(\frac{\widehat{\gamma}_h}{\sigma_h - 1} \times X_{ij}^h \right) B_2 + \varepsilon_{ij}^h \quad (1.7)$$

Exports from country i towards country j in sector h are a function of a constant and a set of dummies (country of origin dummies, country of destination dummies, and sector dummies, B_{ij}^h), a vector of trade barriers (X_{ij}^h includes the log of bilateral distance, common language and common border dummies), and the interaction between the sectoral heterogeneity and trade barriers ($\frac{\widehat{\gamma}_h}{\sigma_h - 1} \times X_{ij}^h$). ε_{ij}^h is assumed to be a normally distributed random shock orthogonal to the right hand side variables. $\frac{\widehat{\gamma}_h}{\sigma_h - 1}$ is estimated from the sectoral distribution of firm size in Compustat. A larger coefficient corresponds to a thinner tail for the distribution of firm size, and therefore a more homogeneous sector. The country fixed effects sum up the impact of size and the impact of relative prices on trade flows²⁵. I cluster observations by country pairs to allow for shocks affecting trade flows in all sectors to differ across country pair, and report robust standard errors.

I expect firm heterogeneity to dampen the effect of trade barriers on bilateral trade flows. More heterogeneous sectors are sectors where the largest firms account for a larger fraction of output; hence the selection among the less productive firms has a minor impact on aggregate trade. I expect B_1 and B_2 to have the same sign.

The regression results are reported in Table 1.2. Column (1) is the benchmark gravity regression with no interaction term between sector heterogeneity and trade barriers. The other specifications take into account the interaction between sector heterogeneity and trade barriers. When using distance as a proxy for trade barriers, the predictions from the model are confirmed. Whether I control for other measures of trade barriers or not, in columns (2) and (5) of Table 1.2, the negative effect of distance on trade is dampened by sectoral heterogeneity (magnified by $\frac{\widehat{\gamma}_h}{\sigma_h - 1}$), as predicted by the theory.

Results on the other proxy measures of trade barriers are more ambiguous. The size, sign and significance of the coefficients vary from one specification to the next.

These qualifications put aside, the predictions of the model with heterogeneous firms are

²⁵The impact of prices on trade flows corresponds in our model to the indices of remoteness, the θ_j 's. I also test the prediction of the model more directly by using measures of country size, and get similar results.

Table 1.2: Firm heterogeneity distorts gravity.

Variable	(1)	(2)	(3)	(4)	(5)
$\ln(\text{Distance}_{ij})$	-0.9 (.04)***	-0.8 (.04)***			-0.7 (.04)***
$\frac{\widehat{\gamma}_h}{\sigma_h-1} \times \ln(\text{Distance}_{ij})$		-0.09 (.002)***			-0.09 (.003)***
Language_{ij}	.3 (.1)***		1.6 (.2)***		-0.4 (.02)**
$\frac{\widehat{\gamma}_h}{\sigma_h-1} \times \text{Language}_{ij}$			-0.4 (.05)***		.4 (.05)***
Border_{ij}	.8 (.02)***			3.9 (.3)***	1.4 (.3)***
$\frac{\widehat{\gamma}_h}{\sigma_h-1} \times \text{Border}_{ij}$				-0.7 (.008)***	-0.3 (.08)*
R^2	30%	31%	23%	25%	32%
Number of obs.	65,687	65,687	65,687	65,687	65,687

Note: Dependent variable, log of exports from country i to country j in sector h in 1996. All regressions include sector dummies, origin country and destination country dummies. Observations are clustered within country pairs. Robust standard errors are given in parentheses. Significant at the 1% (***) , 5%(**), 10% level (*). Source: 1996 bilateral trade flows, Feenstra (2000); firm heterogeneity, Compustat, rank-size scaling coefficient of sales in 1996; data are aggregated over 35 BEA sectors; countries with a GDP/capita lower than \$3000 (in PPP) or a population smaller than 1 million have been ignored.

confirmed. In sectors where output is concentrated among the few largest firms, trade barriers have a mild impact on trade flows, whereas the reverse holds in sectors where output is more uniformly spread across firms.

1.3.3 Market structure distorts gravity

In this section, I evaluate the impact of the elasticity of substitution between goods on how sensitive trade flows are to trade barriers. I find strong support to my model with heterogeneous firms, and reject the model with representative firms. Sectors where the elasticity of substitution is high are sectors where trade barriers have little impact on trade flows. The opposite is true in sectors where the elasticity is low.

In order to test the prediction of the model, I estimate the following equation with OLS:

$$\ln(\text{Exports}_{ij}^h) = B_{ij}^h + X_{ij}^h B_1 + (\hat{\sigma}_h \times X_{ij}^h) B_2 + \varepsilon_{ij}^h \quad (1.8)$$

Exports from country i towards country j in sector h are a function of a constant and a set of dummies (B_{ij}^h), a vector of trade barriers (X_{ij}^h includes the log of bilateral distance, common language and common border dummies), and the interaction between the sectoral elasticity of substitution and trade barriers ($\hat{\sigma}_h \times X_{ij}^h$). ε_{ij}^h is assumed to be a normally distributed shock orthogonal to the right hand side variables. $\hat{\sigma}_h$ is the estimated elasticity of substitution in sector h from Broda and Weinstein (2004). The country fixed effects sum up the impact of size and the impact of relative prices on trade flows. I cluster observations by country pairs to allow for shocks affecting trade flows in all sectors to differ across country pair, and report robust standard errors.

This specification enables us to separate the direct impact of trade costs on trade flows from the dampening or magnifying effect of the elasticity of substitution. If the extensive margin effect dominates, I expect the coefficients on trade barriers to be of the opposite sign to the interaction coefficients. B_1 and B_2 should be of opposite signs.

The regression results are given in Table 1.3. All predictions from the model are confirmed. All coefficients have the expected signs. Simply put, trade barriers reduce trade, but less so in more competitive sectors. This result directly invalidates the prediction of the model with representative firms. The distortion of the elasticity of trade with respect to trade barriers due to the elasticity of substitution between goods is quantitatively important. A one standard deviation increase in the elasticity of substitution (σ increases by 5) corresponds to a reduction of the distance elasticity of trade by a fifth of a standard deviation (from column (5) in Table 1.3, the distance elasticity of trade decreases by $.015 \times 5 = .075$, which represents 22% of the standard deviation of the distance elasticity of trade, .34).

I use direct measures of transportation costs (freight rates from Hummels (1999)) to verify that the sectoral distance elasticity of freight is not correlated with the sectoral elasticity of substitution between goods. The correlation is equal to -5.6% (-1.6% if one removes the 10% sectors with the largest elasticities). I conclude that my results are not a consequence of more

Table 1.3: Market structure distorts gravity.

Variable	(1)	(2)	(3)	(4)	(5)
$\ln(\text{Distance}_{ij})$	-0.8 (.02)***	-1 (.02)***			-0.9 (.02)***
$\hat{\sigma}_h \times \ln(\text{Distance}_{ij})$.02 (.001)***			.015 (.001)***
Language_{ij}	.4 (.04)***		1.2 (.09)***		.5 (.05)***
$\hat{\sigma}_h \times \text{Language}_{ij}$			-0.02 (.004)***		-0.02 (.004)***
Border_{ij}	.5 (.08)***			2.3 (.1)***	.6 (.09)***
$\hat{\sigma}_h \times \text{Border}_{ij}$				-0.04 (.006)***	-0.01 (.006)*
R^2	39%	40%	33%	35%	41%
Number of obs.	270,607	257,583	257,583	257,583	257,583

Note: Dependent variable, log of exports from country i to country j in sector h in 1997. All regressions include sector dummies, origin country and destination country dummies. Observations are clustered within country pairs. Robust standard errors are given in parentheses. Significant at the 1% (***), 5% (**), 10% level (*). Source: 1997 bilateral trade flows, Feenstra (2000); elasticities of substitution, Broda and Weinstein (2004), 1980-1997 estimates; data are aggregated at the 3-digit SITC level; countries with a GDP/capita lower than \$3000 (in PPP) or a population smaller than 1 million have been ignored.

competitive sectors being sectors where trade barriers are less responsive to distance.

This result seems to contradict findings by Hummels (2001, table 4) and presented by Anderson and van Wincoop (2004, figure 1). Based on estimates computed by David Hummels²⁶, James Anderson and Eric van Wincoop find a strong negative correlation between the distance elasticity of trade costs and the elasticity of substitution between goods. I believe that this correlation is an artefact. Oversimplifying unduly the empirical procedure adopted in Hummels (2001), the regression implicitly imposes that the product of the distance elasticity of trade costs and the elasticity of substitution between goods is equal to the distance elasticity of trade. In other words, the elasticity of substitution between goods is by construction equal to the ratio of the distance elasticity of trade and the distance elasticity of trade costs. But since

²⁶I used directly the data on freight from Hummels (2001). Hummels however has access to a larger set of data on direct measures of trade costs than we have.

the distance elasticity of trade varies very little from one sector to the next (far less than the distance elasticity of trade costs in any case), this amounts to imposing a negative relationship between the distance elasticity of trade costs and the elasticity of substitution between goods.

In addition to this, my theoretical model gives reasons to believe that because of the selection of firms into the export market, the elasticity of aggregate trade with respect to trade barriers is not a relevant measure of the elasticity of substitution between goods. It is even inversely related to the elasticity of substitution between goods. The data provides strong support to this interpretation.

In this section and the previous one, I have found strong support for the model with heterogeneous firms in the data, and I have rejected the model with representative firms. The patterns of international trade suggest that the extensive margin of trade plays a crucial role in the adjustments of trade flows to trade barriers, and that this margin tends to quantitatively dominate the intensive margin of trade. Specifically, trade barriers have relatively little impact on trade flows in sectors where there is a lot of heterogeneity between firms, and in sectors where the elasticity of substitution between goods is high.

1.4 Conclusion

I have shown that, contrary to the prediction of the Krugman (1980) model with representative firms, the impact of trade barriers is dampened by the elasticity of substitution, and not magnified by it. I introduce adjustments on the extensive margin in a simple model of international trade. I prove that the elasticity of substitution has opposite effects on the sensitivity of each margin to trade barriers. In sectors where the elasticity of substitution is high, the intensive margin of trade is highly sensitive to changes in trade barriers, whereas the extensive margin is not, and the reverse holds true in sectors where the elasticity of substitution is low. The dampening effect of the elasticity on the substitution on the extensive margin always dominates. High competition sectors are global, in the sense that differences in trade barriers have little impact on bilateral trade flows. Low competition sectors are local.

Bibliography

- [1] Anderson, James E., 1979, "A Theoretical Foundation for the Gravity Equation," *American Economic Review*, v. 69 no. 1, pp. 106-116.
- [2] Anderson, James E. and Eric Van Wincoop, 2004, "Trade Costs" *Journal of Economic Literature*, vol. 42 no. 3, pp. 691-751.
- [3] Axtell, Robert L., 2001, "Zipf Distribution of U.S. Firm Sizes," *Science*, 293 (September 2001), pp. 1818-1820.
- [4] Bernard, Andrew B.; Jonathan Eaton; J. Bradford Jensen and Sam Kortum, 2003, "Plants and Productivity in International Trade," *American Economic Review*, v. 93, pp. 1268-1290.
- [5] Bernard, Andrew B. and J. Bradford Jensen, 1999, "Exceptional exporter performance: cause, effect, or both?", *Journal of International Economics*, v. 47, pp. 1-25.
- [6] Bernard, Andrew B. and J. Bradford Jensen, 2001a, "Why Some Firms Export," *The Review of Economics and Statistics*, v. 86 no. 2.
- [7] Bernard, Andrew B. and J. Bradford Jensen, 2001b, "Who Dies? International Trade, Market Structure, and Industrial Restructuring," *NBER Working Paper # 8327*.
- [8] Bernard, Andrew B. and J. Bradford Jensen, 2002, "Entry, Expansion, and Intensity in the US Export Boom, 1987-1992," *Review of International Economics*, forthcoming.
- [9] Broda, Christian and David E. Weinstein, 2004, "Globalization and the Gains from Variety," *Federal Reserve Bank of New York, mimeo*.

- [10] Davis, Donald, 1998, "The Home Market Effect, Trade and Industrial Structure," *American Economic Review*, v.88 no. 5, pp. 1264-1276.
- [11] Deardorff, Alan V., "Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?" (Nov. 7, 1995). In Jeffrey A. Frankel, ed., *The Regionalization of the World Economy*, Chicago: University of Chicago Press, 1998, pp. 7-28
- [12] Dornbusch, Rudiger; S. Fischer, and Paul A. Samuelson, "Comparative Advantage, Trade, and Payments in a Ricardian Model with a continuum of Goods," *American Economic Review*, vol. 67, pp. 823-839.
- [13] Eaton, Jonathan; Sam Kortum, 2002, "Technology, Geography, and Trade," *Econometrica*, vol. 70 no. 5 (September 2002), pp. 1741-1779.
- [14] Eaton, Jonathan; Sam Kortum, and Francis Kramarz, 2004, "An Anatomy of International Trade: Evidence from French Firms," *American Economics Review, Papers and Proceedings*, forthcoming.
- [15] Eaton, Jonathan; Sam Kortum, and Francis Kramarz, 2004, "Dissecting Trade: Firms, Industries, and Export Destinations," *NBER Working Paper 10344*.
- [16] Eden, L., Texas A&M University: www.wcrl.ars.usda.gov/cec/java/capitals.htm
- [17] Evenett, Simon J. and Wolfgang Keller, 2002, "On Theories Explaining the Success of the Gravity Equation." *Journal of Political Economy*, v. 110 no. 2, pp. 281-316.
- [18] Feenstra, Robert C., 1994, "New Product Varieties and the Measurement of International Prices," *American Economic Review*, v.84 no. 1 (March 1994), pp. 157-177.
- [19] Feenstra, Robert C., 2000, "World Trade Flows 1980-1997," *Center for International Data, UC Davis, Mimeo (March 2000)*.
- [20] Feenstra, Robert C., 2003, *Advanced International Trade: Theory and Evidence*, Princeton, NJ: Princeton University Press.
- [21] Feenstra, Robert C.; Robert E. Lipsey and Harry P. Bowen, 1997, "World Trade Flows, 1970-1992, with Production and Tariff Data," *NBER Working Paper 5910 (January 1997)*.

- [22] Gabaix, Xavier, 1999, "Zipf's Law for Cities: An Explanation", *Quarterly Journal of Economics*, v. 114 n. 3 (August 1999), pp.739-67.
- [23] Gabaix, Xavier and Yannis Ioannides, 2004, "The Evolution of City Size Distributions", in *Handbook of Urban and Regional Economics*, volume 4, V. Henderson and J.-F. Thisse eds., North-Holland, Chapter 53, p. 2341-2378.
- [24] Harrigan, James, 1995, "The Volume of Trade in Differentiated Intermediate Goods: Theory and Evidence," *The Review of Economics and Statistics*, v. 77 no. 2 (May 1995), pp. 283-293.
- [25] Harrigan, James, 2002, "Specialization and the Volume of Trade: Do the Data Obey the Laws?" in *Handbook of International Trade*. E. Kwan Choi and James Harrigan, eds. London: Basil Blackwell.
- [26] Helpman, Elhanan; Marc Melitz and Yona Rubinstein, 2004, "Trading Partners and Trading Volumes," *Working Paper*.
- [27] Helpman, Elhanan; Marc Melitz and Stephen Yeaple, 2004, "Export Versus FDI with Heterogeneous Firms," *American Economic Review*, v. 94 no.1 (March 2004), pp. 300-316.
- [28] Hummels, David, 2001, "Toward a Geography of Trade Costs," *Purdue University mimeo* (September 2001).
- [29] Hummels, David and Peter J. Klenow, 2004, "The Variety and Quality of a Nation's Exports," *American Economic Review* (forthcoming).
- [30] Hummels, David and J. Levinsohn, 1993, "Product differentiation as a Source of Comparative Advantage," *American Economic Review, Papers and Proceedings*, v. 83 (May 1993), pp. 445-449.
- [31] Kortum, Sam, 1997, "Research, Patenting, and Technological Change," *Econometrica*, v. 65, no. 6. (Nov., 1997), pp. 1389-1419.
- [32] Krugman, Paul. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review*, vol. 70, pp. 950—959.

- [33] Melitz, Jacques, 2003, "Language and Foreign Trade," *University of Strathclyde Working Paper*.
- [34] Melitz, Marc, 2003, "The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations," *Econometrica*, v. 71 no. 6 (November 2003), pp. 1695-1725.
- [35] Obsfeld, Maurice and Kenneth Rogoff, 2000, "The Six Major Puzzles in International Macroeconomics. Is There a Common Cause?" in *NBER Macroeconomics Annual 2000*. B. S. Bernanke and K. Rogoff, eds. Cambridge, MA: The MIT Press, pp. 339-390.
- [36] Rauch, James E., 1999, "Networks versus Markets in International Trade," *Journal of International Economics*, vol. 48, pp. 7-35.
- [37] Ruhl, Kim, 2003, "Solving the Elasticity Puzzle in International Economics," *University of Minnesota, Mimeo (November 2003)*.
- [38] Syverson, Chad, 2004, "Product Substitutability and Productivity Dispersion," *Review of Economics and Statistics*, v. 86 no. 2 (May 2004), pp. 534-550.
- [39] Tinbergen, J., 1962, "An Analysis of World Trade Flows," in *Shaping the World Economy*, Tinbergen, ed. New York: Twentieth Century Fund.
- [40] Tybout, James R., 2003, "Plant and Firm-Level Evidence on 'New' Trade Theories," in *Handbook of International Trade*. E. Kwan Choi and James Harrigan, eds. Oxford: Basil Blackwell, 2003, pp. 388-415.
- [41] Yeaple, Stephen R., 2003, "Firm Heterogeneity, International Trade, and Wages," *University of Pennsylvania, Mimeo*.

Table 1.4: 35 manufacturing sectors.

<i>WTDB</i>	<i>Industry</i>	<i>ISIC, Rev.2</i>	<i>US 1987 SIC code</i>
1	Grain, mill and bakery products	311	204, 205
2	Beverages	313	208
3	Tobacco products	314	21
4	Other food and kindred products	311	201, 202, 203, 206, 207, 209
5	Apparel and other textile products	321+322	22, 23
6	Leather and leather products	323+324	31
7	Pulp, paper and board mills	341	261, 262, 263
8	Other paper and allied products	341	265, 267
9	Printing and publishing	342	27
10	Drugs	352	283
11	Soaps, cleaners and toilet goods	352	284
12	Agricultural chemicals	351	287
13	Industrial chemicals and synthetics	351	281, 282, 286
14	Other chemicals	352	285, 289
15	Rubber products	355	301, 302, 305, 306
16	Miscellaneous plastic products	356	308
17	Primary metal industries: Ferrous	371	331, 332, 339
18	Primary metal industries: Non-ferrous	372	333, 334, 335, 336
19	Fabricated metal products	381	34
20	Farm and garden machinery	382	352
21	Construction, mining, etc machinery	382	353
22	Computer and office equipment	382	357
23	Other non-electric machinery	382	351, 354, 355, 356, 358, 359
24	Household appliances	383	363
25	Household audio and video, etc	383	365, 366
26	Electronic components	383	367
27	Other electrical machinery	383	361, 362, 364, 369
28	Motor vehicles and equipment	384	371
29	Other transportation equipment	384	372, 373, 374, 375, 376, 379
30	Lumber, wood, furniture, etc	331+332	24, 25
31	Glass products	362	321, 322, 323
32	Stone, clay, concrete, gypsum, etc	361+369	324, 325, 326, 327, 328, 329
33	Instruments and apparatus	385	38
34	Other manufacturing	390	39

Chapter 2

Liquidity Constrained Exporters

Summary 4 I build a model of international trade with liquidity constraints. If firms must pay some entry cost in order to access foreign markets, and if they face liquidity constraints to finance these costs, only those firms that have sufficient liquidity are able to export. A set of firms could profitably export, but they are prevented from doing so because they lack sufficient liquidity. More productive firms that generate large liquidity from their domestic sales, and wealthier firms that inherit a large amount of liquidity, are more likely to export. This model predicts that the scarcer the available liquidity and the more unequal the distribution of liquidity among firms, the lower are total exports. I also offer a potential explanation for the apparent lack of sensitivity of exports to exchange rate fluctuations. When the exchange rate appreciates, existing exporters lose competitiveness abroad, and are forced to reduce their exports. At the same time, the value of domestic assets owned by potential exporters increases. Some liquidity constrained exporters start exporting. This dampens the negative competitiveness impact of a currency appreciation. Under some circumstances, it may actually reverse it altogether and increase aggregate exports. This model provides some argument for competitive revaluations.

2.1 Introduction

Trade economists view the patterns of imports and exports mainly as the outcome of a competition game between producers in different countries. If goods from a given country become cheaper, provided that demand are not too inelastic, producers from this country will export larger quantities. Eventually, general equilibrium adjustments will have to take place to restore trade balance, but in the short and medium run, we should observe such competitiveness effects of devaluations. The only departure from this competition mechanism may happen in the very short run, and has been described as the J-curve: if demand is inelastic in the short run, a devaluation may have a negative impact on the current account in the short run, before the competitiveness effect comes into play and exports catch up. However, we observe relatively little response of trade flows to exchange rate fluctuations. The euro-dollar exchange rate has experienced wide fluctuations since the inception of the euro, without any significant and systematic effect on the patterns of trade between Europe and the US. The US dollar has been steadily depreciating vis à vis most foreign currencies, without any evidence of a reduction in the US trade deficit. On the contrary, there is evidence that foreign companies take advantage of the relatively cheap US domestic prices to enter the US market for the first time. In a different context, there are many examples of middle income countries undergoing massive devaluations of their currency without much gain in terms of current account imbalances. The only way to reconcile the observed impact of exchange rate fluctuations with the existing theory of competitive trade is to assume extremely inelastic demands for foreign goods. Such elasticities are at odds with other evidence of the impact of trade barriers on trade flows, as well as micro economic evidence.

I propose a theory of international trade with liquidity constraints that can account for these facts, along with micro evidence on the characteristics of exporters. The main predictions of the model are, first, that liquidity constraints are a key determinant of the export behavior of firms, and second, that exchange rate fluctuations (or more generally fluctuations of relative prices) may have the opposite effect as predicted by traditional theories. If there are fixed costs associated with exporting, then liquidity constraints at the firm level will come into play. In such a context, few firms will be able to export. Furthermore, an appreciation of the domestic currency, despite the negative effect on the competitiveness of exporters, will not have a large

impact on aggregate exports, even if demand is elastic. It may actually have a positive impact on exports in the medium run if goods are not too substitutable.

The reason is the following. In the presence of fixed costs associated with exporting and liquidity constraints, some firms could profitably export, but they are prevented from doing so because they cannot gather sufficient liquidity. Only those firms that are productive enough and generate sufficient cash flows from their domestic sales are able to export. If the exchange rate appreciates, potential exporters lose competitiveness abroad and therefore lose some market shares abroad. Existing exporters reduce their exports. This is the traditional competitiveness effect. But an appreciation of the exchange rate also means that the value of domestic assets abroad increases. Some firms that could not enter foreign markets because of liquidity constraints, enter now that the value of their assets has appreciated. Total trade does not change much: existing exporters export less, but new firms start exporting. If the competitiveness effect is mild enough, that is if goods are sufficiently differentiated, aggregate exports may increase following an appreciation of the exchange rate.

In other terms, it is the extensive margin of trade that responds differently to exchange rate fluctuations in the presence of firm heterogeneity and liquidity constraints. Following an exchange rate appreciation, some firms, favored by the increased value of their domestic assets, enter the export market. This entry of exporters, the extensive margin of trade, may offset the reduction of the volumes exported by existing exporters, the intensive margin of trade.

This theory also accounts for the fact that few firms export, and that exporters will typically be firms that are not liquidity constrained. There is a growing set of evidence from micro data that only a small fraction of firms export. Exporters are different from non exporters in many respects. Exporters are more productive than non exporters, they are larger, more capital intensive, and they tend to belong to large groups more frequently than non exporters. The same hierarchy applies between firms that export to many markets versus firms that export towards a few markets only. Although it is still a matter for debate, it seems that the direction of the causality goes from the characteristics of the firms towards the export status. It is because a firm is more productive that it is more likely to become an exporter, rather than because they export that they become more productive. I develop a model where the selection into the export market is similar. Only those firms that are not liquidity constrained are able

to export. The capacity to overcome liquidity constraints is endogenously determined in this model. Firms may have sufficient liquidity, inherited from their past activities, but they can also generate sufficient cash flow from their domestic activities in order to gain access to foreign markets. In equilibrium, only a subset of firms are able to gather enough liquidity and export. The export status is the outcome of the characteristics of the firm, even though the partition between exporters and non exporters are endogenously determined as the outcome of a competitive game.

Finally, this model has important implications for the link between financial development, macroeconomic stability and openness to trade. The model predicts that a deepening or a widening of the financial markets will increase total exports. When firms get easier access to external finance (what I call a deepening of financial markets), or when more firms get access to cheap external finance (a widening of the financial markets), they become able to overcome barriers associated with international trade. More firms export, and total exports increase. However, the model does not predict that better financial markets will stabilize or destabilize the current account. The predictions of the model about the volatility of exports and the degree of financial development are ambiguous. Only in the extreme case of perfect financial markets can we say that exports will be more volatile than if financial markets were not perfectly developed. Exchange rate fluctuations, for instance, will cause larger movements of the volume of exports if financial markets are perfectly developed than if they are not. The reason is the following. If financial markets are not perfectly developed, there exists a fringe of liquidity constrained exporters. When the exchange rate appreciates, some of those liquidity constrained firms start exporting. This entry of new exporters dampens the negative impact of the exchange rate appreciation on existing exporters. If financial markets are perfectly developed, that is if no firm is liquidity constrained, this dampening channel does not exist anymore. Exports will be more responsive to exchange rate fluctuations. This is the only case where the model can make a clear prediction about the link between financial development and current account volatility. For intermediate levels of development on the other hand, an improvement of financial markets will always increase total exports, but it may or may not increase export volatility. The primary purpose of this model is not to describe the link between financial development and macroeconomic instability. It offers however an interesting angle on a potential link between

the volatility of some aggregates (fluctuations in the volume of exports here), and the degree of financial development. It also gives specific predictions for the impact of financial development on the volume of exports.

In the remaining part of this introduction, I review the literature related to this model. First, recent research has widely documented the important of firm heterogeneity and the role of fixed costs in international trade, both empirically and theoretically. Second, there is a large body of literature on the importance of liquidity constraints for firms. Finally, there is a (scarce) literature on the interaction between firm level liquidity constraints and international trade.

Firm heterogeneity has recently been acknowledged to be a major feature of the export behavior of firms. Exporters are different from non exporters in many respects. They tend to be more productive, larger, more capital intensive. The link between productivity and trade has been analyzed in many different countries. Bernard and Jensen (1999, 2001a, 2001b, 2002) for the US, Aw and Huang (1995) for Taiwanese and Korean firms, Clerides, Lach and Tybout (1998) for Colombian, Mexican and Moroccan firms, and Delgado, Farinão and Ruano (1999) for Spanish firms. The same hierarchy that exists between exporters and non exporters also exists between firms that export to a few foreign markets, and firms that export to many foreign markets. Using firm level data on French exporters, Eaton, Kortum and Kramarz (2001a) uncover systematic regularities for the characteristics of exporters, and for the popularity of foreign market. Not only are exporters more productive, larger, and more capital intensive than non exporters, but firms that export to many markets are also more productive, larger, and more capital intensive than firms that export only to a few markets. In this paper, I propose to extend the study of the heterogeneity between exporters and non exporters to the severity of liquidity constraints. My model predicts that one dimension of heterogeneity along which exporters may differ from non exporters is their ability to access financial intermediaries. Less financially constrained firms are more likely to export.

This dichotomy between exporters and non exporters allows the extensive margin of trade, the entry and exit of firms into the export market, to play a crucial role in determining the volume of trade flows. When trade barriers change, or when the degree of competition in foreign markets evolves, some firms will go in and out of the export markets. The importance of the extensive margin of trade has been pointed out since the seminal work of Paul Krugman (1980).

From the microeconomic point of view, trade is as much about how much each exporting firm exports, as it is about how many firms export. Empirically, it seems that, at least in the medium run, most of the adjustment for aggregate trade flows comes from entry and exit of firms into the export market. Eaton, Kortum and Kramarz (2001b), using data on French exporters, disaggregated by trading partner, show that most of the variation in the aggregate French exports comes from variation in the number of exporters, rather than differences of exports per firm. Helpman, Melitz and Rubinstein (2004) use the zeros of trade matrices (which happen when no firm trades between two countries) to infer information about the extensive margin of trade, and derive an estimate of trade barriers between countries. Broda and Weinstein (2004) point out that in the last 30 years, the number of varieties of foreign goods available to US consumers has increased fourfold. They calibrate an extended Krugman model of trade to show that the extensive margin of trade may explain an annual increase in welfare worth 3% of GDP. Their measures of the number of varieties imported by the US, given the extremely high level of disaggregation they use, is a good proxy for the number of firms exporting to the US. Ruhl (2003) uses a theoretical framework with heterogeneous firms to show that the extensive margin of trade may explain the discrepancy between the short run and long run elasticities of trade with respect to trade barriers. Ruhl argues that high frequency variations in exchange rates, because exchange rates are mean reverting, will trigger only negligible adjustments of the extensive margin, which explains why exchange rate fluctuations seem to have so little impact on trade flows, whereas variations in tariffs or quotas have such a large impact. In this paper, I build on this literature and study the extensive margin of trade in the presence of liquidity constraints. If financial markets are underdeveloped, the extensive margin of trade reacts both because there are productivity differences between firms, but also because different firms face different degree of liquidity constraints. I show how predictions of traditional models of trade may be modified, or even overturned. Among others, I offer an alternative explanation for the relatively mild impact of exchange rate fluctuations from that of Ruhl (2003). Even if the extensive margin of trade does respond to exchange rate movements, since there will be simultaneous entry and exit of firms, the impact on aggregate exports will be mild. When the exchange rate appreciates, some existing exporters lose competitiveness in the foreign market and stop exporting. But at the same time, the value of domestic assets denominated in foreign

currency increases, so that liquidity constrained firms start exporting. The net effect on the extensive margin is mild. Under some circumstances, despite the loss in competitiveness, a real exchange rate appreciation may actually lead to an increase in aggregate exports.

Alongside the empirical relevance of firm heterogeneity for international trade, theoretical models of heterogeneous firms have recently been developed. The two main models have been separately developed by Eaton and Kortum (2002) and by Melitz (2003). In both models, firms differ in terms of productivity. Eaton and Kortum build a new Ricardian model of trade in the spirit of Dornbusch, Fischer and Samuelson (1977). Heterogeneous firms compete internationally for foreign markets. Competition is perfect. In each sector, only the most competitive firm in the world will service the market in a given country. Firms in a subset of sectors will be exporters. Melitz on the other hand uses a monopolistic competition framework, and fixed costs associated with exporting. Only the most productive firms are able to overcome fixed costs associated with exporting. I build on the Melitz model and add liquidity constraints to it. Those liquidity constraints interact with productivity heterogeneity. The most productive firms generate enough liquidity from domestic sales to overcome any liquidity constraints. However, some less productive firms would be profitable enough to export, but are prevented from doing so because they are liquidity constrained. On top of interacting liquidity constraints with productivity heterogeneity, one important contribution of this model is to break up the symmetry imposed by construction in Melitz (2003). Exchange rate fluctuations in my model are equivalent to fluctuations in relative real wages in different countries. This formalization is similar to the one used by Atkeson and Burstein (2005). Atkeson and Burstein, using a model with endogenous mark-ups that depend on a firm's market share, are able to generate endogenously pricing to market. Such pricing to market behavior explain first why there is incomplete pass-through of exchange rate shocks on domestic prices, and second why exports may not be as sensitive to exchange rate fluctuations as expected. Exchange rate fluctuations have important implications for the selection of firms into the export market. An important other model that breaks the symmetry imposed in the Melitz model is done by Ghironi and Melitz (2005). They derive an endogenous micro founded explanation for the Harrod-Balassa-Samuelson effect. In their model, as in this one, entry and exit of a specific subset of firms have important implications for the behavior of aggregate variables, such as real exchange rates or

total exports.

This paper contributes to the literature on hysteresis in trade, and provides a new angle of explanation for these phenomena. Empirical studies do find substantial hysteresis in aggregate trade flows, in import and export prices, as well as in the export status of firms. Baldwin (1988) documents the hysteresis in import prices. Roberts and Tybout (1994) derive a theoretical model of export decision with sunk costs in order to analyze the sluggish movements in the export status of firms. Campa (1998) calibrates the importance of hysteresis in international trade. Closely related is the fact that pioneer firms, that is firms that are the first to enter foreign markets, tend to differ from followers in many dimensions. In the same way as there is a strong hierarchy between firms that explains their export status, there is also a strong hierarchy in the order in which firms access foreign markets. Trade has been growing continuously since WWII. However, the apparent smooth increase in international trade flows over the last 60 years hides a vast heterogeneity of trade links. As trade grows, new trade links are created. New countries start trading, new product lines start being traded. The firms that pioneer the entry into a new country, or the firms that pioneer the export of a new product line, are substantially different from the firms that follow them up. Bernard, Jensen and Schott (2005), using a rich dataset on US exporters and multinational firms, find that multinational firms tend to be pioneers. Firms that already export in many other foreign markets are more likely to enter new and relatively isolated markets. Liquidity constraints may play an important role in these hysteresis phenomena. I do not develop a dynamic version of this model in this paper. Hence I can only describe qualitatively how the tools developed in this model may explain these patterns. In a dynamic setting, firms may gradually accumulate enough liquidity from their exporting activity in other countries to eventually enter new markets. The most productive and least liquidity constrained firms are likely to be the first to enter remote markets. Firms that have been trading in many markets (and therefore have been generating liquidity), are the first candidate to enter new markets. If liquidity constraints matter for accessing foreign markets, then the history of previous liquidity shocks matters as well. Even if it could profitably export, a firm has to gather sufficient liquidity to enter a new market. An ordering between pioneers and followers will endogenously emerge. Hysteresis is a direct consequence. Once a firm has covered the fixed costs of entering a set of foreign markets, it is somehow sheltered

from higher frequency shocks. By the same token, exporting to other countries provides a form of insurance to multi-country exporters, which explains hysteresis in exported volumes. If firms are heterogeneous, different firms will charge different prices. Since the set of exporters is history dependent (through the history of liquidity shocks), aggregate export prices will tend to display the same hysteresis as aggregate trade flows.

There is a vast literature on the importance of liquidity constraints for firms, which follows the pioneering work of Stiglitz and Weiss (1981). Fazzari, Hubbard and Peterson (1988) study the importance of financing constraints for investment. The importance of the lending channel has been stressed in Holmstrom and Tirole (1997), as well as in Stein (1998). Empirically, there is a strong evidence of the presence liquidity constraints given by the correlation between a firm's financial position and its investments. This is true for firms as well as banks, which would explain the transmission of monetary shocks to the economy. Bernanke and Gertler (1995) and Kashyap and Stein (2000) study the importance of credit constraints for banks. Gertler and Gilchrist (1994) point that small firms' production contracts when money is tight, which is further evidence of the importance of liquidity constraints. Hoshi, Kashyap and Scharfstein (1992) offer a model where net worth determines whether to use direct or indirect finance. I introduce this concept of liquidity constraint in a model of international trade with heterogeneous firms. Liquidity constraints in this model are modelled in the simplest way possible. Without going into any detail of moral hazard, or endogenous bargaining with financial intermediaries under incomplete contracts, I simply assume that firms cannot borrow externally in order to enter foreign markets. This is the reduced form expression of an unmodelled game between potential financial intermediaries, and potential exporters. However, in the presence of productivity heterogeneity, liquidity constraints will interact with trade barriers, exchange rates, firm level productivity in a complex way. Domestic sales may endogenously relax the liquidity constraints faced by a potential exporter.

To the best of my knowledge, only one paper looks at the relationship between international trade and liquidity constraints. Campa and Shaver (2001) use a panel of Spanish manufacturing firms in the 1990's to test whether there exists any link between the liquidity constraints a firm faces and its exporting status. They do find that liquidity constraints are less binding for exporters than for non exporters. They also find that cash flows are more stable for exporters

than for non exporters. They argue that it is the stability provided by foreign sales that relaxes the liquidity constraints of exporters, and not the reverse. Exporters earning profits in different markets the business cycles of which are imperfectly correlated can pledge more stable future earnings, which softens agency problems in their relationship with financial intermediaries, and relaxes their liquidity constraints. The model I build predicts the same raw correlation between liquidity constraints and export status. But I would claim that the causality runs in the opposite direction: it is because they are less liquidity constrained that some firms are able to export, and not the reverse. I believe that Campa and Shaver actually find some suggestive evidence that a relaxation of liquidity constraints causes firms to export, and not the reverse. In page 21, they report that exporting firms are less liquidity constrained than non exporters, but that the fraction of sales exported does not matter for liquidity constraints. The only thing that matters is whether a firm exports or not (a dummy variable for the export status), not how much it exports. This is consistent with my model: firms below a given level of liquidity constraints export. How much they export depends on their productivity, not on how constrained they are. Only the dummy for positive exports should matter, not how much is exported. If the insurance mechanism put forward by Campa and Shaver were at play, the more a firm exports, the more insurance from demand shocks it gets, and the less liquidity constrained it should be. In a signalling model, how much a firm exports carries information about how productive a firm is. It would be surprising that financiers would not use such easily accessible information.

The remainder of the paper is organized as follows. Section 2.2 introduces a simple model of trade with liquidity constraints and heterogeneous firms. Section 2.3 describes the impact of exchange rate fluctuation in the presence of liquidity constraints. Section 2.4 concludes.

2.2 A model of trade with liquidity constrained exporters

In this section, I develop a model of international trade with liquidity constrained firms. I introduce those liquidity constraints in the context of a model of trade with heterogeneous firms à la Melitz (2002).

There are 2 countries, home and foreign, that produce goods using only labor. All foreign

variables are denoted by an asterisk. The home country has a population L (L^* for the foreign country). There are 2 sectors. One sector provides a single homogeneous good that can be freely traded. This good is used as the numeraire, and its price is set equal to 1. It is produced under constant returns to scale. The unit labor requirement for producing the homogeneous good at home is $1/w$ ($1/w^*$ abroad). Provided that each country produces the homogeneous good, the wages will be w and w^* . I shall only consider equilibria where this assumption holds. The other sector supplies a continuum of differentiated goods. Each firm is a monopolist for the variety it produces.

2.2.1 Demand

The workers are the only consumers, each endowed with one unit of labor. They all share the same CES preferences over the differentiated good. A consumer that receives q_o units of the homogeneous good, $q(x)$ units of each variety x of the differentiated good, for all varieties x in the set X (to be determined in equilibrium) gets a utility U :

$$U \equiv q_o^{1-\mu} \left(\int_{x \in X} q(x)^{\frac{\sigma-1}{\sigma}} dx \right)^{\frac{\sigma}{\sigma-1}\mu}$$

with $\sigma > 1$

where σ is the elasticity of substitution between two varieties of the differentiated good.

If all varieties in the set X are available domestically, at a price $p(x)$ each, I can define the following ideal price index for differentiated goods domestically:

$$P = \left(\int_{x \in X} p(x)^{1-\sigma} dx \right)^{\frac{1}{1-\sigma}} \quad (2.1)$$

The representative consumer has an isoelastic demand function for each differentiated variety. She spends $r(x)$ on each variety x :

$$r(x) = \mu w L \left(\frac{p(x)}{P} \right)^{1-\sigma} \quad (2.2)$$

where μwL is the total expenditure spent on differentiated goods.

2.2.2 Production and trade

There are two types of trade barriers, a fixed cost and a variable cost. If a firm exports, it must pay a fixed cost C_f in terms of foreign labor, or w^*C_f in terms of the numeraire. The assumption that the entry cost into the foreign market is denominated in foreign labor is important. An exporter must cover costs both in domestic and in foreign labor. I only need to assume that the part of the fixed entry cost denominated in foreign labor is positive¹. There is evidence that a large share of the cost of entering foreign markets consists of the cost of acquiring local information, setting up a local distribution network, and customizing goods to fit the local market. Arguably, those costs depend on the conditions in the local market. The variable cost takes the form of an "iceberg" transportation cost. If one unit of any differentiated good is shipped abroad, only a fraction $1/\tau$ arrives. The rest melts on the way. The higher τ , the higher the variable trade cost.

Each country has access to the same technology. The marginal product of labor is constant. In order to start production, a firm must pay a fixed entry cost C_d in terms of domestic labor, at a price wC_d in terms of the numeraire. The presence of fixed entry cost means that firms operate under increasing returns to scale. Each firm in the differentiated sector draws a random unit labor productivity $x \geq 0$. For a firm with productivity x , the cost of producing q_d units of good for the home market is $c_d(q_d)$, and the cost of producing q_f units for the foreign market is $c_f(q_f)$:

$$\begin{aligned} c_d(q_d) &= q_d \frac{w}{x} + wC_d \\ c_f(q_f) &= q_f \frac{\tau w}{x} + w^*C_f \end{aligned}$$

Firms are price setters. Given that demand functions are isoelastic, the optimal price is a

¹Adding another part to the cost of entering foreign markets, that would be denominated in domestic labor, would reduce the set of exporters and the total amount of exports, but it would not modify any qualitative result of this model.

constant mark-up over the unit cost (including transportation costs)²,

$$p_d(x) = \frac{\sigma}{\sigma-1} \times \frac{w}{x} \text{ at home, } p_f(x) = \frac{\sigma}{\sigma-1} \times \frac{\tau w}{x} \text{ abroad}$$

Given these pricing strategies, more productive firms are able to charge lower prices, capture a larger market share, and generate larger profits, both at home and abroad. A firm with productivity x potentially generates profits $\pi_d(x)$ in the domestic market, and $\pi_f(x)$ in the foreign market:

$$\begin{aligned} \pi_d(x) &= \frac{r_d(x)}{\sigma} - wC_d = \frac{\mu}{\sigma} wL \left(\frac{\sigma}{\sigma-1} \frac{w}{xP} \right)^{1-\sigma} - wC_d \\ \pi_f(x) &= \frac{r_f(x)}{\sigma} - w^*C_f = \frac{\mu}{\sigma} w^*L^* \left(\frac{\sigma}{\sigma-1} \frac{\tau w}{xP^*} \right)^{1-\sigma} - w^*C_f \end{aligned}$$

Only those firms that can profitably produce domestically will survive, and only those firms that can profitably produce for the export market could export. I can implicitly define two productivity thresholds, \bar{x}_d for survival on the domestic market, and \bar{x}_f for profitable entry into the foreign market, absent any additional constraint. Only those firms that generate non negative profits from domestic sales survive, and only those firms that generate non negative profits from selling in the foreign market could export. The productivity thresholds are defined by³,

$$\pi_d(\bar{x}_d) = 0 \text{ and } \pi_f(\bar{x}_f) = 0 \tag{2.3}$$

Firm heterogeneity and monopolistic competition gives the following partition among firms. More productive firms (higher x firms) are able to capture larger market shares, and generate larger profits. The least productive firms cannot cover the overhead costs and are not able to survive. However, despite the differences in productivity between firms, some low productivity firms can still survive because of the imperfect nature of competition. As long as the elasticity of

²This price prevents any arbitrage either by domestic firms that might want to resell these goods at home or abroad, or by foreign firms.

³Note that $(\bar{x}_f/\bar{x}_d)^{\sigma-1} = (\tau^{\sigma-1}C_d/C_f) \times (L/L^*) \times (P/P^*)$. I assume that trade barriers are always sufficiently high ($\tau^{\sigma-1}C_d/C_f$ sufficiently high) so that $\bar{x}_f > \bar{x}_d$ always holds. Only a subset of firms are able to export, and no firm is able to sell abroad but not domestically.

substitution between varieties, σ , is finite, low productivity firms are sheltered from competition and may survive. The same selection takes place among firms for the entry into the export market. The highest productivity firms generate enough profits to justify the entry cost into the foreign market. Less productive firms do not export.

Absent any other friction, all firms with a productivity above \bar{x}_f would export. But for the potential asymmetry between countries, this model is almost identical to the Melitz (2003) model of international trade. Among other things, the only reason why the export status of a firm is correlated to the size of its domestic sales, is that more productive firms sell more *and* are more likely to export. There is no direct link between what a firm exports, how many countries it exports to, and what it does at home.

In the next section, I introduce liquidity constraints. We shall see how the presence of financial imperfections creates a link between different markets, and modifies the adjustment to changes in exogenous variables.

2.2.3 Liquidity constraints

One crucial assumption of the above model is that there are some fixed costs associated with international trade. There is a growing set of evidence that a part of trade barriers take the form of fixed costs. Most of these costs must be paid up-front. These costs are substantial. All previous models assume that there exist perfect financial markets so that any firm that could profitably export will find some investors to finance the entry cost into the foreign markets.

However, there are reasons to believe that such investments may not be easy to finance. The nature of the contracting and informational environment is different from a similar entry cost investment made domestically. This is for two reasons mainly. First, export activities are essentially riskier than domestic ones. Part of it is due to the objective added risks, such as foreign exchange risk. Existing financial hedging products such as swaps and options may not be available, or available at a prohibitive cost for most potential exporters. Information about foreign markets is harder and more expensive to get. It may also be less verifiable. Part of the fixed cost associated with international trade actually corresponds to the cost of acquiring information on a foreign market. Potential investors may not be willing to pay this cost themselves. But since such information is harder to verify than similar information on

domestic markets, a potential investor may not be willing to trust a would-be exporter. Second, the contracting environment for international transactions is relatively weak, if existing at all. Sales are done in another country, and it is hard for investors to collect the proceeds of such sales in case of disagreement. In other words, a potential exporter cannot pledge much collateral for its foreign activities, and this translates into ex ante under investment.

The same problems apply for foreign investors: informational asymmetries and contract incompleteness plague such relations. A foreign investor has little information on foreign firms. If she does enter in a relationship with a foreign firm, and if the terms of their contract are violated, she will find it difficult to seize any asset the firm owns. Arguably, trade credits will alleviate many of these issues. However, trade credits are typically offered to existing exporters, that is firms with a known and verified history of exports into a given market. Such a firm arguably has already covered most of the entry cost into foreign markets.

In the remaining part of this paper, I will take an extreme view on the limitations of financial markets that potential exporters face. I will assume an extreme dichotomy between domestic and foreign markets. A firm may find investors for any investment regarding domestic activities, but none whatsoever for exporting activities. Therefore, a firm must rely on its own existing liquidity to cover entry costs into foreign markets. Moreover, I will assume that firms inherit an exogenous amount of liquidity (it may be thought of as a trustworthiness capital that gives access to financial markets). This is an extreme and oversimplified view of liquidity constraints. These assumptions are designed to carry two properties. First, liquidity constraints are more severe for international trade than for domestic trade. Second, firms are more or less severely hampered by liquidity constraints, and how much constraints they face is not perfectly correlated with their current productivity.

Liquidity constraints are formalized in the following way. Firms only face liquidity constraints for accessing foreign markets. I make the extreme assumption that domestic investors do not have any information on the conditions in foreign markets. Hence they are not willing to lend to firms for the purpose of exporting. Similarly, I assume that the incompleteness of international contracts is such that foreign investors are not willing to finance domestic exporters. Therefore, firms that want to export need to have enough liquidity on their own to cover the fixed cost of entering markets.

I further assume that each firm is endowed with a random liquidity shock A . Since A is a domestic liquidity shock, it is denominated in units of domestic labor, and has a value wA in terms of the numeraire. The profits generated from domestic sales, $\pi_d(x)$ are also pledgeable. (A, x) are drawn from a joint distribution with c.d.f. $F(A, x)$ over $\mathbb{R}^+ \times \mathbb{R}^+$, and $F_x(x) \equiv \lim_{A \rightarrow \infty} F(A, x)$ over \mathbb{R}^+ . A firm's productivity and its degree of liquidity constraint may or may not be correlated, depending on the specific shape of the distribution F . I also assume that the total mass of firms entering the lottery is proportional to the size of the country, L .⁴

In order to export, a firm must have enough liquidity to cover the fixed entry cost, w^*C_f . It generates some liquidity from domestic sales, $\pi_d(x)$, and it has access to some additional exogenous liquidity wA . So an exporter is subject to the following liquidity constraint,

$$\pi_d(x) + wA \geq w^*C_f \quad (2.4)$$

More productive firms generate larger profits at home, and therefore are less dependent on external finance. I define $\bar{x}(A)$ as the lowest productivity below which firms with liquidity A cannot gather enough liquidity to enter the foreign market. $\bar{x}(A)$ is defined by,

$$\pi_d(\bar{x}(A)) + wA = w^*C_f$$

All firms with a productivity below $\bar{x}(A)$ are prevented from exporting because of liquidity constraints, even if they could profitably export.

2.2.4 Open economy equilibrium

Since I am mainly interested in what happens in the home country, I assume that foreign firms face no liquidity constraint. I make one additional simplifying assumption: price indices only depend on prices set by local firms. In other words, prices set by foreign exporters have a negligible impact on the general price index domestically. This is a fair approximation for a

⁴Implicitly, we assume that there is a group of entrepreneurs proportional to the size of the country. We could remove this assumption, and allow for the free entry of entrepreneurs, with an infinite set of potential entrepreneurs. Provided that trade barriers are large enough, we would get qualitatively the same results.

relatively closed economy. Formally, I replace the price index Equation (2.1) by the following approximation,

$$P \approx \left(\int_{x \geq \bar{x}_d} p_d(x)^{1-\sigma} L dF_x(x) \right)^{\frac{1}{1-\sigma}} \quad (2.5)$$

It will be convenient to define the function $g(\cdot)$ in the following way:

$$g(\cdot) : \bar{x}^{\sigma-1} = \left(\frac{\sigma}{\mu} \int_{x \geq \bar{x}} x^{\sigma-1} dF_x(x) \right) \times C \Leftrightarrow \bar{x} = g(C) \quad (2.6)$$

It is straightforward to prove that $g' > 0$. Rearranging the conditions for the productivity thresholds in Eq. (2.3) and the liquidity constraints condition in Eq. (2.4), I have,

$$\bar{x}_d = g(C_d) \quad (2.7)$$

$$\bar{x}_f = \left(\frac{\tau w}{w^*} \right) \left(\frac{C_f}{C_d^*} \right)^{\frac{1}{\sigma-1}} g(C_d^*) \quad (2.8)$$

$$\bar{x}(A) = \left(\frac{C_d + \frac{w^*}{w} C_f - A}{C_d} \right)^{\frac{1}{\sigma-1}} g(C_d) \quad (2.9)$$

All the firms with a productivity above \bar{x}_d produce and sell their output domestically. Only those firms with a productivity above $\max\{\bar{x}_f, \bar{x}(A)\}$ are able to export.

What are the determinants of the liquidity constraint, $\bar{x}(A)$ in Eq. (2.9)? $\bar{x}(A)$ is a downward sloping schedule. Firms that only own a small amount of exogenous liquidity, A small, must have a very high level of productivity in order to generate sufficient liquidity on their own and enter foreign markets. Firms with a large amount of exogenous liquidity on the other hand, A large, do not require much additional liquidity, and do not need a high productivity in order to be able to export. The higher the entry cost into the foreign country, C_f , the higher the curve $\bar{x}(A)$. The fixed overhead production cost, C_d , has an ambiguous impact on the curve $\bar{x}(A)$ ⁵. An increase of the domestic fixed cost C_d eats up part of a firm's liquidity, and reduces its ability to enter foreign markets. However, an increase in C_d also makes it harder for firms to survive, and forces some domestic firms out of business. This softens

⁵See appendix 2.5.1 for a formal proof of this statement.

competition, increases the market share of each surviving firm, increases profits, and hence increases available liquidity. Which force dominates depends on the underlying distribution of firm productivity. The higher the exchange rate, that is the lower w^*/w , the lower the curve $\bar{x}(A)$. The reason is straightforward: if the exchange rate appreciates (w^*/w falls), the value of domestic assets in terms of foreign prices increases, and less domestic liquidity is required to enter the foreign market. The more substitutable the goods, that is the higher σ , the flatter $\bar{x}(A)$. This is because when goods are highly substitutable, any small difference in productivity implies large differences in profits, and therefore in the liquidity generated by domestic sales.

Interestingly, only a subset of firms are potentially subject to liquidity constraints. Firms with a very high productivity, no matter how little exogenous liquidity they own, are able to export. Formally, $\bar{x}(0)$ is bounded, so that firms with a productivity above $\bar{x}(0)$ do not need any exogenous liquidity. At the other extreme, firms with a sufficiently large amount of exogenous liquidity do not need to generate any additional liquidity from domestic sales. If their exogenous liquidity endowment A is sufficiently large to cover the fixed entry cost into both the domestic and the foreign market, even without any domestic sales, those firms could export⁶. In formal terms, $\bar{x}(C_d + \frac{w^*}{w}C_f) = 0$. From a social efficiency point of view, there is no need to allocate exogenous liquidity beyond $C_d + \frac{w^*}{w}C_f$. We will see in the next section that profitability conditions imply that the maximum liquidity a firm could ever need to enter foreign markets is actually below $C_d + \frac{w^*}{w}C_f$. We can already see that the distribution of liquidity among firms (and how that correlates with a firm's productivity) will have important implications for the ability of firms to export.

We have seen that a firm's productivity may allow it to overcome liquidity constraints without the need to accessing financial markets. Are liquidity constraints going to be binding for any firm? In other words, is any firm profitable enough to be a viable exporter, but prevented from accessing foreign markets because of liquidity constraints? The following proposition gives conditions under which there will be a set of liquidity constrained exporters.

Proposition 5 *If $\left(\frac{C_d^*}{C_f} + \frac{w^*C_d^*}{wC_d}\right)^{\frac{1}{\sigma-1}} \frac{g(C_d)}{g(C_d^*)} > \frac{\tau w}{w^*}$, then there is a non empty set of liquidity*

⁶Note however that such wealthy firms may not be able to survive in the domestic market, or profitably export. This is precisely described in the next section.

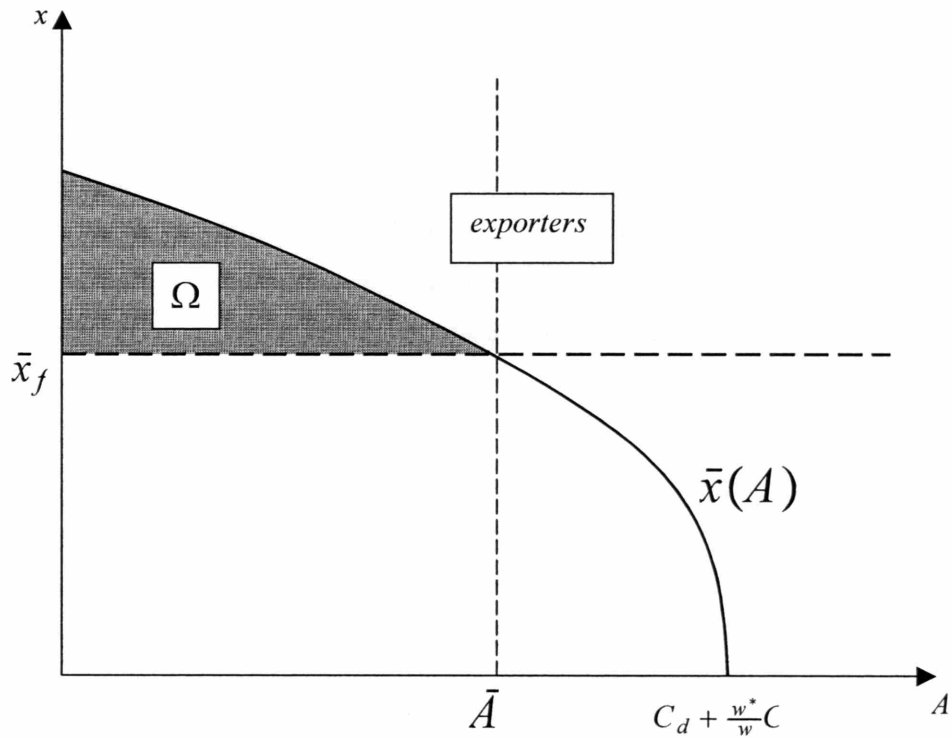


Figure 2-1: Liquidity constrained exporters. Note: No firm with a productivity below \bar{x}_f can profitably export. On top of this this, firms below the curve $\bar{x}(A)$ are liquidity constrained. All firms in the area Ω between \bar{x}_f and $\bar{x}(A)$ are liquidity constrained exporters. They could profitably export, but are prevented from doing so because of liquidity constraints.

constrained firms (denoted Ω). These firms could profitably export, but are prevented from doing so because they lack sufficient liquidity.

Proof. See appendix 2.5.2. ■

From now on, I assume that the condition in Proposition 5 holds, so that the set of liquidity constrained firms, Ω , is non empty.

No firm with a productivity below \bar{x}_f could profitably export. These firms have a productivity too low to allow them to generate enough profits in the foreign market to recover the fixed entry cost. Firms with a productivity above $\bar{x}(0)$ export no matter how little exogenous liquidity they hold: they are competitive enough in the foreign market to generate positive

profits, and they generate sufficient liquidity from their domestic activities to cover the entry cost into the foreign market, without the need for any additional liquidity. Firms with an intermediate productivity, $\bar{x}_f \leq x < \bar{x}(0)$, could profitably export, but do not generate sufficient liquidity from their sales on the domestic market. They need extra liquidity. Without this extra liquidity, despite being profitable, they would be prevented from exporting. This is shown on Figure 2-1. The dark shaded area Ω corresponds to liquidity constrained firms: these firms are willing to export ($x \geq \bar{x}_f$), but cannot do so because they lack sufficient liquidity ($x \leq \bar{x}(A)$).

It is interesting to note that the distribution of liquidity among firms does matter. As can be seen graphically, firms with more than \bar{A} (with $\bar{x}(\bar{A}) = \bar{x}_f$) exogenous liquidity have "too much" liquidity. \bar{A} corresponds to the amount of liquidity the least productive exporter would need to enter foreign markets. Any firm with a productivity below \bar{x}_f could not export, and therefore would have no use for exogenous liquidity. Following a similar reasoning, no firm with a productivity above $\bar{x}(0)$ has the need for any exogenous liquidity. Such high productivity firms are already able to generate sufficient liquidity from their domestic sales, and do not need additional source of funding. Only firms with an intermediate level of productivity (between \bar{x}_f and $\bar{x}(0)$) must have access to some exogenous source of liquidity in order to overcome financial constraints. However none of these firms would need more than a maximum \bar{A} of liquidity. I describe in the next section how the distribution of wealth matters for aggregate exports in the next section.

Proposition 5 is testable. It states that financially constrained firms cannot export. Financially constrained firms are firms that both lack sufficient exogenous liquidity, and that are not productive enough to generate sufficient liquidity on their own. Campa and Shaver (2001) find that more liquidity constrained firms are less likely to export. They define financially constrained firms as firms for which investment is correlated with cash flows. I expect financially constrained firms in my model, that is firms that both lack existing liquidity and generate little liquidity from their sales, to enter Campa and Shaver's category of financially constrained firms. Therefore there is some evidence supporting proposition 5: there exists a set of financially constrained firms that are prevented from exporting.

2.2.5 Liquidity constraints and missing trade

If it does export, the total value of exports (f.o.b.) by a firm with productivity x is $r_f(x)$. Using the expressions for the productivity thresholds in Eqs. (2.7), (2.8) and (2.9), plugging those and the price index equation (2.5) back into the revenue equation (2.2), I get,

$$r_f(x) = \sigma w^* C_d^* \left(\frac{w^*}{\tau w} \times \frac{x}{\bar{x}_d^*} \right)^{\sigma-1} \quad (2.10)$$

All firms with a productivity above $\max\{\bar{x}_f, \bar{x}(A)\}$ export. All firms in the set Ω are prevented from exporting. The total volume of missing trade (f.o.b.) from these constrained exporters, T_{missing} , and the total volume of exports (f.o.b.), T_{total} , are given by,

$$\begin{aligned} T_{\text{missing}} &= L \iint_{(A,x) \in \Omega} r_f(x) dF(A, x) \\ T_{\text{total}} &= L \int_{x \geq \bar{x}_f} r_f(x) dF_x(x) - T_{\text{missing}} \end{aligned}$$

The total volume of missing trade depends on several parameters of the distribution of productivity and liquidity shocks. It depends both on the average liquidity available economy wide, and on the distribution of this liquidity. If only highly productive firms are liquidity constrained, there will not be any missing trade. If only those low productivity firms that would not export anyway are liquidity constrained, there will not be any missing trade. To get a better understanding of the magnitude of this missing trade, I consider a special case for the distribution of liquidity and productivity shocks $F(A, x)$.

In order to get simple predictions for the export behavior of liquidity constrained firms, I will now use a simplified form for the joint distribution of productivity and liquidity shocks, $F(A, x)$. Assume that the liquidity shocks and the productivity shocks are uncorrelated. Further assume that a fraction θ of firms are liquidity constrained ($A < \bar{A}$), and the remaining $(1 - \theta)$ is not ($A \gg \bar{A}$). \bar{A} is defined as the minimum liquidity above which financial constraints are not binding, and the only constraint is the profitability constraint: $\bar{x}(\bar{A}) = \bar{x}_f$. Firms with a high enough positive liquidity shock will always be able to generate enough liquidity on the home market. They will export only if exporting is profitable. I can rewrite the equations for T_{missing}

and T_{total} in this special case,

$$T_{\text{missing}} = \theta L \int_{\bar{x}(A)}^{\bar{x}_f} r_f(x) dF_x(x) \quad (2.11)$$

$$T_{\text{total}} = L \int_{x \geq \bar{x}_f} r_f(x) dF_x(x) - T_{\text{missing}} \quad (2.12)$$

Definition 6 *A deepening of financial markets corresponds to an increase of A , the amount of liquidity available to financially constrained firms. A widening of financial markets corresponds to a reduction in θ , the number of financially constrained firms.*

Proposition 7 *Both a deepening and a widening of financial markets has a positive impact on total trade flows.*

Proof. See appendix 2.5.3. ■

This model predicts that financial constraints faced by potential exporters have a negative impact on trade flows. Both the absolute amount of liquidity and the distribution of liquidity among firms matter for the total volume of trade. These predictions are testable. The model predicts that more financially constrained industries should have lower trade flows. Sectors where the distribution of liquidity is more unequal should have lower trade flows. Moreover, financial constraints matter more when entry costs to foreign markets are larger (C_f large). If those entry costs increase with distance, more financially constrained sectors should have a larger distance elasticity of trade (in absolute value).

In the next section, I turn to the impact of exchange rate shocks on trade in the presence of liquidity constraints.

2.3 The ambiguous impact of exchange rate shocks on trade flows

In this section, I describe the impact of exchange rate shocks on trade. If the domestic currency appreciates vis à vis the foreign currency, domestic producers lose competitiveness in the foreign

market. This is a classic terms of trade effect. However, an appreciation of the domestic currency relaxes the liquidity constraint faced by potential exporters. The value of domestic assets in terms of the foreign currency increases. Liquidity constrained firms are now more likely to be able to pay foreign denominated entry costs and start exporting. Simultaneously, existing exporters lose competitiveness and export less, but new firms start exporting. The intensive margin of trade is negatively affected by an appreciation of the exchange rate, whereas the extensive margin is positively affected.

I model exchange rate shocks in this model as a shock on relative wages (in terms of the numeraire). I will define an appreciation of the domestic currency as an increase in the productivity in the homogeneous sector at home, which leads to an increase in the domestic wages w , all else equal. This definition is similar to the one used by Atkeson and Burstein (2005). If the domestic wage increases, the value of domestic assets (wA and $\pi_d(x)$) increases, whereas potential exporters lose competitiveness in the foreign market ($p_f(x)/P^*$ increases). These effects are exactly equivalent to an appreciation of the domestic currency vis à vis the foreign currency.

Proposition 8 *An appreciation of the exchange rate has 3 effects:*

- (i) *Existing exporters lose market shares abroad and reduce their exports: $\frac{\partial r_f(x)}{\partial w} < 0$.*
- (ii) *The least productive non constrained exporters are forced out of the export market: $\frac{\partial \bar{x}_f}{\partial w} > 0$.*
- (iii) *The most productive constrained firms start exporting: $\frac{\partial \bar{x}(A)}{\partial w} < 0$.*

Proof. (i) The first effect is the classic impact of a loss of competitiveness for exporters. As the value of domestic inputs increases, domestic exporters have to charge higher prices in order to maintain mark-ups, and therefore they lose market shares in the foreign market. This loss of market shares implies a reduction in exports. Formally, differentiating Eq. (2.10) with respect to w , I get,

$$\frac{\partial r_f(x)}{\partial w} = -(\sigma - 1) \frac{r_f(x)}{w} < 0$$

- (ii) The second effect is the natural corollary of the first effect. As exporters lose competi-

tiveness, they lose market shares, and therefore earn reduced profits:

$$\frac{\partial \pi_f(x)}{\partial w} = -\frac{\sigma-1}{\sigma} \frac{r_f(x)}{w} < 0$$

At the same time, the cost of entering the foreign market, w^*C_f , denominated in foreign currency, is unchanged. So the least productive firms, earning smaller profits, cannot cover the entry cost into the foreign market anymore. The productivity threshold \bar{x}_f goes up. Formally, differentiating Eq. (2.8) with respect to w , I get,

$$\frac{\partial \bar{x}_f}{\partial w} = \frac{\bar{x}_f}{w} > 0$$

(iii) The last effect comes from the relaxation of the liquidity constraint. As the domestic currency appreciates, the value of domestic assets (both exogenous liquidity and endogenous domestic profits) in terms of foreign currency increases. Since the entry cost into the foreign market is paid in foreign currency, this means a relaxation of the liquidity constraint for constrained exporters. Formally, differentiating Eq. (2.9) with respect to w , I get,

$$\frac{\partial \bar{x}(A)}{\partial w} = -\left(\frac{1}{\sigma-1}\right) \frac{w^*C_f}{w} \left(\frac{C_d}{C_d + \frac{w^*}{w}C_f - A}\right)^{\frac{1}{\sigma-1}} \frac{\bar{x}(A)}{w} < 0$$

■

An appreciation of the exchange rate causes both entry and exit. Non liquidity constrained firms with a low productivity are forced out of the export market because they lose competitiveness in the foreign market. Liquidity constrained firms with a high productivity (close to $\bar{x}(A)$) face a relaxed liquidity constraint and enter the export market. This can be seen on Fig. 2-2. Low productivity non constrained firms, in the light shaded area, exit the export market. This is due to the fact that they lose competitiveness in the foreign market, and therefore earn less profits. They cannot cover the fixed trade barrier any more, and exit the export market. At the same time, high productivity constrained firms, in the dark shaded area, enter the export market. These firms are sufficiently productive to export (even after the currency appreciation), but they were prevented from doing so because of liquidity constraints. The appreciation of their currency increases the value of their domestic assets and allows them to start exporting.

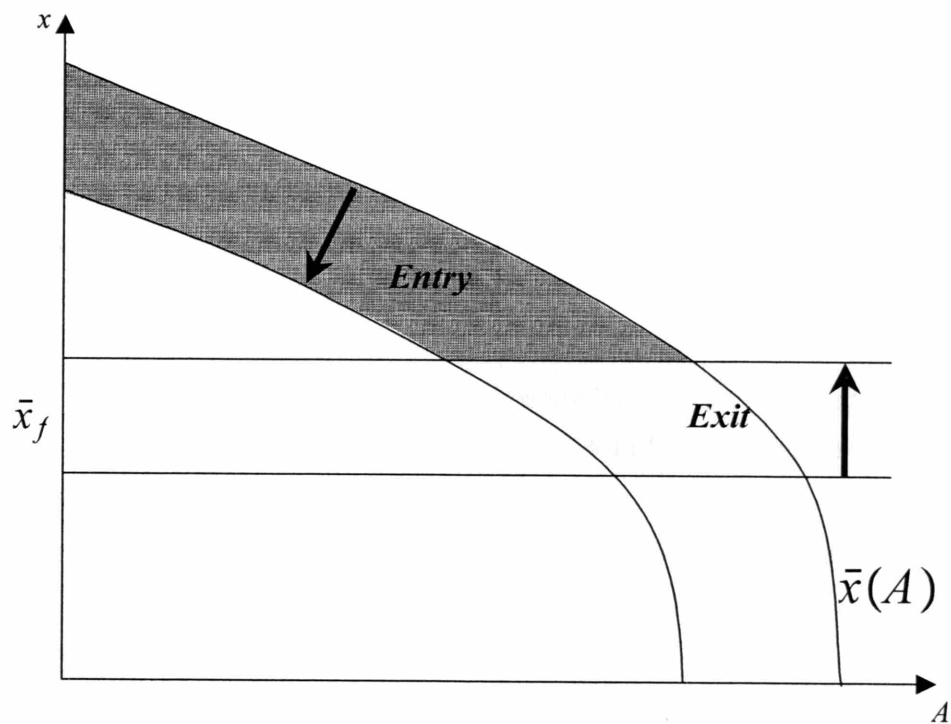


Figure 2-2: The ambiguous impact of an exchange rate appreciation. Note: An appreciation of the domestic currency erodes the competitiveness of exporters, and forces the least productive exporters to exit (light shaded area). At the same time, it relaxes the liquidity constraint, and allows some liquidity constrained firms to enter the export market (dark shaded area).

Depending both on the strength of the liquidity constraints (the overall scarcity of liquidity in the economy) and on the number of liquidity constrained firms, either effect can dominate. If there are relatively many liquidity constrained firms, there will be a net entry of firms following an appreciation of the exchange rate.

The presence of liquidity constraints introduces investment motive in international trade in goods. Exports do not depend only on the competitiveness of exporters, it also depends on the value of domestic assets relative to the "cost" of exporting. In the same way as an exchange rate appreciation will make investment abroad more accessible, it makes exporting more likely for a group of firms.

Proposition 9 *If competition is soft (σ close to 1), an appreciation of the exchange rate will have a positive impact on exports.*

Proof. See appendix 2.5.4 ■

This model of international trade with liquidity constrained exporters predicts that, under some conditions, an appreciation of the exchange rate, despite negative impact on the competitiveness of exporters, may have a strictly positive impact on exports. If competitiveness does not have too large an impact on the size of market shares, that is if goods are very differentiated (σ low), then the entry of liquidity constrained exporters following an appreciation of the exchange rate will dominate. Total exports increase after an appreciation of the exchange rate.

More generally, even if an appreciation of the exchange rate has an negative impact on exports, the negative impact of an exchange rate appreciation will be milder (or even become positive) the more unequal the distribution of liquidity within the sector (θ low). If liquidity is unequally shared among firms (θ low), many healthy and productive firms are liquidity constrained. This means that many firms could profitably export, but they lack sufficient access to financial markets to cover the entry cost into the foreign market. If the exchange rate appreciates, the liquidity constraint faced by all those firms is relaxed. A fraction of these firms will then start exporting, despite the loss of competitiveness. The more unequal the distribution of wealth among firms, the more firms will start exporting, and the more positive the impact of a exchange rate appreciation.

Liquidity constraints for the access to foreign markets allows effective policy interventions. There is room for temporary competitive revaluations. A temporary revaluation of the domestic currency, by increasing the value of domestic assets, may allow liquidity constrained firms to start exporting. The sunk cost nature of a fraction of fixed costs associated with exporting implies some asymmetry for the response of trade to exchange rate fluctuations. Once liquidity constrained firms have started exporting, insofar as they do not have to pay this fixed cost again, they will continue exporting, even after a devaluation of their currency.

It may also be possible to extend this model to describe phenomena of amplification and contagion in international trade. If a firm exports to a given foreign market, it generates some liquidity, in addition to domestic profits. Such extra liquidity will give this firm an edge for

entering other foreign markets. This corresponds to an amplification mechanism: an increase in exports by a firm may trigger additional exports, since it relaxes its liquidity constraint. In a multi country dynamic setting, firms may be able to accumulate liquidity from both their domestic sales, and from past sales to other foreign markets. Firms that have already entered many foreign markets are more likely to have sufficient liquidity to enter new and less accessible markets. Shocks that affects exports with one trading partner may influence the volume of exports with other trading partners, even absent any direct link between those countries. If trade links with a given trading partner are severed, the liquidity streams generated from exporting to this country cease, and some exporters may be forced to pull out of other markets as well. Moreover, if markets in the same region are characterized by similar trade barriers, they will attract similar exporters. Modifying the access to one of these markets will affect trade with all other countries in the same region. Hence, liquidity constraints may artificially generate contagion phenomena in international trade.

2.4 Conclusion

I have shown in this chapter that liquidity constraints may modify fundamentally the behavior of exporters, and the patterns of aggregate exports. If firms face liquidity constraints when accessing foreign markets, some firms are prevented from exporting. They could profitably enter foreign markets, but are prevented from doing so because they lack the ability to access financial markets and cover entry costs into foreign markets. The main prediction of the model is that financial underdevelopment hinders exports. Both the total amount of liquidity available, and the distribution of this liquidity matters for trade. The model also predicts that the presence of liquidity constraints will reduce the sensitivity of trade barriers to exchange rate fluctuations. When the exchange rate appreciates, exporters lose competitiveness, and they reduce their exports. However, since the value of domestic assets in terms of foreign prices increases, liquidity constraints for accessing foreign markets are relaxed. Some firms start exporting. This entry of liquidity constrained exporters dampens the negative competitiveness effect of an exchange rate appreciation. Under some circumstances, an exchange rate appreciation may have a positive impact on exports. Liquidity constraints also create artificial links between different markets,

and thus generate amplification and contagion phenomena. If a firm starts exporting to a new foreign market, it generates some liquidity from its exports. This additional liquidity may allow it to enter more foreign markets in the future.

Bibliography

- [1] Aw, Bee Yan and A. Hwang, 1995, "Productivity and the export market: A firm level analysis." *Journal of Development Economics*, v. 47 no.2, pp. 313-332.
- [2] Atkeson, Andrew and Ariel Burstein, 2005, "Trade Costs, Pricing to Market, and International Relative Prices." *Working Paper*.
- [3] Baldwin, R., 1988, "Hysteresis in Import Prices: The Beachhead Effect." *American Economic Review*, v. 78 no.4, pp. 773-785.
- [4] Bernanke, Ben S. and Mark L. Gertler, 1995, "Inside the Black Box: The Credit Channel of Monetary Policy Transmission." *The Journal of Economic Perspectives*, v. 9 no. 4, pp. 27-48.
- [5] Bernard, Andrew B.; Jonathan Eaton; J. Bradford Jensen and Sam Kortum, 2003, "Plants and Productivity in International Trade," *American Economic Review*, v. 93, pp. 1268-1290.
- [6] Bernard, Andrew B. and J. Bradford Jensen, 1999, "Exceptional exporter performance: cause, effect, or both?", *Journal of International Economics*, v. 47 no. 1, pp. 1-25.
- [7] Bernard, Andrew B. and J. Bradford Jensen, 2001a, "Why Some Firms Export," *The Review of Economics and Statistics*, v. 86 no. 2.
- [8] Bernard, Andrew B. and J. Bradford Jensen, 2001b, "Who Dies? International Trade, Market Structure, and Industrial Restructuring," *NBER Working Paper # 8327* .
- [9] Bernard, Andrew B. and J. Bradford Jensen, 2002, "Entry, Expansion, and Intensity in the US Export Boom, 1987-1992," *Review of International Economics*, forthcoming.

- [10] Bernard, Andrew B.; J. Bradford Jensen, and Peter K. Schott, 2005, "Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Goods." *Working Paper*.
- [11] Broda, Christian and David E. Weinstein, 2004, "Globalization and the Gains from Variety," *Federal Reserve Bank of New York, Mimeo*.
- [12] Campa, José M., 1998, "Hysteresis in Trade: How Big are the Numbers?" *New York University Working Paper*.
- [13] Campa, José M. and J. Myles Shaver, 2001, "Exporting and Capital Investment: On the Strategic Behavior of Exporters," *New York University Mimeo*.
- [14] Clerides, Sofronis K.; Saul Lach and James R. Tybout, 1998, "Is learning by exporting important? Microdynamic evidence from Colombia, Mexico, and Morocco." *Quarterly Journal of Economics*, v. 113 no. 3, pp. 903-947.
- [15] Dornbusch, Rudiger; S. Fischer, and Paul A. Samuelson, "Comparative Advantage, Trade, and Payments in a Ricardian Model with a continuum of Goods," *American Economic Review*, vol. 67, pp. 823-839.
- [16] Eaton, Jonathan; Sam Kortum, 2002, "Technology, Geography, and Trade," *Econometrica*, vol. 70 no. 5 (September 2002), pp. 1741-1779.
- [17] Eaton, Jonathan; Sam Kortum, and Francis Kramarz, 2004, "An Anatomy of International Trade: Evidence from French Firms," *American Economics Review, Papers and Proceedings*, forthcoming.
- [18] Eaton, Jonathan; Sam Kortum, and Francis Kramarz, 2004, "Dissecting Trade: Firms, Industries, and Export Destinations," *NBER Working Paper 10344*.
- [19] Fazzari, Steven, R.; R. Glenn Hubbard, and Bruce Petersen, 1988, "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, pp. 141-195.
- [20] Fariñas, Jose C. and Ana Martín-Marcos, 2002, "Exporting and Productivity: Firm-Level Evidence for Spanish Manufacturing." *Universidad Complutense, Mimeo*.

- [21] Gertler, Mark and Simon Gilchrist, 1994, "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms." *Quarterly Journal of Economics*, v. 59, pp. 309-340.
- [22] Ghironi, Fabio, and Marc J. Melitz, 2005, "International Trade and Macroeconomic Dynamics with Heterogeneous Firms." *Working Paper*.
- [23] Helpman, Elhanan; Marc Melitz and Yona Rubinstein, 2004, "Trading Partners and Trading Volumes," *Working Paper*.
- [24] Holmstrom, Bengt and Jean Tirole, 1997, "Financial Intermediation, Loanable Funds and the Real Sector." *The Quarterly Journal of Economics*, v. 112 no. 3, pp. 663-91.
- [25] Hoshi, T.; Anil Kashyap and D. Scharfstein, 1994, "The Choice Between Public and Private Debt: An Examination of Post-Regulation Financing in Japan." *MIT Working Paper*.
- [26] Kashyap, Anil and Jeremy Stein, 2000, "What do One Million Observations on Banks Have to Say About the Transmission of Monetary Policy." *American Economic Review*, v. 90 no. 3, pp. 407-428.
- [27] Krugman, Paul. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review*, vol. 70, pp. 950—959.
- [28] Melitz, Marc, 2003, "The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations," *Econometrica*, v. 71 no. 6 (November 2003), pp. 1695-1725.
- [29] Roberts, M. and James Tybout, 1999, "An Empirical Model of Sunk Costs and the Decision to Export." *Policy Research Working Paper Series 1436, The World Bank*.
- [30] Ruhl, Kim, 2003, "Solving the Elasticity Puzzle in International Economics," *University of Minnesota, Mimeo (November 2003)*.
- [31] Stein, Jeremy, 1998, "An Adverse Selection Model of Bank Asset and Liability Management with Implications for the Transmission of Monetary Policy." *RAND Journal of Economics*, v. 29 no. 3, pp. 466-86.

[32] Stiglitz, J. and A. Weiss, "Credit Rationing in Markets with Imperfect Information." *American Economic Review*, v. 71, pp. 393-410.

2.5 Appendix

2.5.1 Proof: monotonicity of $\bar{x}(A, C_d)$

Proposition 10 $\bar{x}(A)$ is increasing in C_d .

Proof. The function $g(C)$ is steeper than $C^{\frac{1}{\sigma-1}}$. This can be seen simply by rearranging the definition of the function $g(\cdot)$ in Eq. (5). We can define the function $\xi(C)$ in the following way:

$$\xi(C) \equiv \frac{g(C)}{C^{\frac{1}{\sigma-1}}} = \frac{\sigma}{\mu} \int_{x>g(C)} x^{\sigma-1} dF_x(x)$$

Since $g(\cdot)$ is increasing in C , $\xi(\cdot)$ is decreasing in C . How steep it is depends on the underlying distribution of productivity F_x . Depending on this distribution, $\xi(\cdot)$ can be arbitrarily steep, or arbitrarily flat.

We can now plug in this new function $\xi(\cdot)$ into the formula for $\bar{x}(A)$ in Eq. (2.9):

$$\bar{x}(A) = \left(C_d + \frac{w^*}{w} C_f - A \right)^{\frac{1}{\sigma-1}} \times \xi(C_d)$$

When C_d increases, the first term of the product goes up, and the second term goes down. The net effect can go either way, depending on how steep the function $\xi(\cdot)$ is.

We can describe the intuition behind each term of the product. When the domestic fixed cost goes up, it eats up the liquidity of all firms. This pushes up the minimum productivity required for entering foreign markets, at any level of A . This is the increase in the first term of the product.

At the same time, due to the increase in the domestic fixed cost, some firms are pushed out of business. Some firms die, competition softens among survivors, and profits increase. Each surviving firm now generates more liquidity from its domestic sales. The minimum productivity required for entering foreign markets falls, at any level of A . This is the reduction in the second term of the product. ■

2.5.2 Proof of proposition 5

Proposition 5 (reminded) *If $\left(\frac{C_d^*}{C_f} + \frac{w^* C_d^*}{w C_d}\right)^{\frac{1}{\sigma-1}} \frac{g(C_d)}{g(C_d^*)} > \frac{\tau w}{w^*}$, then there is a non empty set of liquidity constrained firms (denoted Ω). These firms could profitably export, but are prevented from doing so because they lack sufficient liquidity.*

Proof. All firms below $\bar{x}(A)$ are liquidity constrained, and cannot export no matter how profitable their exporting would be. All firms above \bar{x}_f could profitably export, if they have sufficient liquidity. I want to prove that $\Omega \neq \emptyset$, with $(A, x) \in \Omega$ iff $\bar{x}_f \leq x < \bar{x}(A)$. Firms in Ω could profitably export ($x \geq \bar{x}_f$), but they are prevented from doing so because they are liquidity constrained ($x < \bar{x}(A)$). A necessary and sufficient condition for Ω to be non empty is that $\bar{x}(0) > \bar{x}_f$. I know that:

$$\begin{aligned}\bar{x}_f &= \left(\frac{\tau w}{w^*}\right) \left(\frac{C_f}{C_d^*}\right)^{\frac{1}{\sigma-1}} g(C_d^*) \\ \bar{x}(A) &= \left(\frac{C_d + \frac{w^*}{w} C_f - A}{C_d}\right)^{\frac{1}{\sigma-1}} g(C_d)\end{aligned}$$

so that $\bar{x}(0) > \bar{x}_f$ iff,

$$\left(\frac{C_d^*}{C_f} + \frac{w^* C_d^*}{w C_d}\right)^{\frac{1}{\sigma-1}} \frac{g(C_d)}{g(C_d^*)} > \frac{\tau w}{w^*}$$

If $\left(\frac{C_d^*}{C_f} + \frac{w^* C_d^*}{w C_d}\right)^{\frac{1}{\sigma-1}} \frac{g(C_d)}{g(C_d^*)} > \frac{\tau w}{w^*}$, Ω is non empty, and there are liquidity constrained firms. ■

2.5.3 Proof of proposition 7

Proposition 7 (reminded) *Both a deepening and a widening of financial markets has a positive impact on total trade flows.*

Proof. A deepening of the financial markets corresponds to a relaxation of the liquidity constraint of constrained firms (a reduction of A). A widening of financial markets corresponds to a reduction in the number of liquidity constrained firms (a reduction of θ). Differentiating

the expressions for total trade and missing trade in Eqs (2.12) and (2.11), I get,

$$\begin{aligned}\frac{\partial T_{\text{total}}}{\partial \theta} &= -L \int_{\bar{x}(A)}^{\bar{x}_f} r_f(x) dF_x(x) \\ \frac{\partial T_{\text{total}}}{\partial A} &= -\theta T \frac{\partial \bar{x}(A)}{\partial A} r_f(\bar{x}(A)) \frac{\partial F_x}{\partial x}(\bar{x}(A))\end{aligned}$$

From Eq. (2.9), I get $\frac{\partial \bar{x}(A)}{\partial A} = -\left(\frac{1}{\sigma-1}\right) \frac{\bar{x}(A)}{C_d + \frac{w^*}{w} C_f - A} < 0$, which insures that a relaxation of the liquidity constraint has a positive impact on trade. Therefore,

$$\frac{\partial T_{\text{total}}}{\partial \theta} < 0 \text{ and } \frac{\partial T_{\text{total}}}{\partial A} > 0$$

■

2.5.4 Proof of proposition 9

Proposition 9 (reminded) *If competition is soft (σ close to 1), an appreciation of the exchange rate will have a positive impact on exports.*

Proof. I can rewrite total exports in Eq. (2.12) in the following way,

$$T_{\text{total}} = \int_{x \geq \bar{x}(A)} r_f(x) L dF_x(x) + (1 - \theta) \int_{\bar{x}_f}^{\bar{x}(A)} r_f(x) L dF_x(x)$$

Differentiating this equation with respect to w , and using Leibnitz rule, and applying Lebesgues' theorem of monotone convergence to insure existence of the integrals, I get,

$$\begin{aligned}\frac{\partial T_{\text{total}}}{\partial w} &= \int_{x \geq \bar{x}(A)} \frac{\partial r_f(x)}{\partial w} L dF_x(x) + (1 - \theta) \int_{\bar{x}_f}^{\bar{x}(A)} \frac{\partial r_f(x)}{\partial w} L dF_x(x) \\ &\quad - (1 - \theta) \frac{\partial \bar{x}_f}{\partial w} r_f(\bar{x}_f) L \frac{\partial F_x(\bar{x}_f)}{\partial x} \\ &\quad - \theta \frac{\partial \bar{x}(A)}{\partial w} r(\bar{x}(A)) L \frac{\partial F_x(\bar{x}(A))}{\partial x}\end{aligned}$$

Using the formulas for $\partial r_f(x)/\partial w$, $\partial \bar{x}_f/\partial w$, and $\partial \bar{x}(A)/\partial w$ from the previous proof, I can

rewrite this as,

$$\begin{aligned} \frac{\partial T_{\text{total}}}{\partial w} = & - \left(\frac{\sigma - 1}{w} \right) T_{\text{total}} \\ & - (1 - \theta) \frac{\bar{x}_f r_f(\bar{x}_f)}{w} L \frac{\partial F_x(\bar{x}_f)}{\partial x} \\ & + \left(\frac{\theta}{\sigma - 1} \right) \frac{w^* C_f}{w} \left(\frac{C_d}{C_d + \frac{w^*}{w} C_f - A} \right)^{\frac{1}{\sigma-1}} \frac{\bar{x}(A) r_f(\bar{x}(A))}{w} L \frac{\partial F_x(\bar{x}(A))}{\partial x} \end{aligned}$$

The first two terms in the sum are negative (the loss of competitiveness of existing exporters, and the exit of non constrained exporters). The last term is positive (the entry of liquidity constrained exporters).

The first two terms are bounded, and the last term converges to infinity as σ converges to 1. In other words, provided that the loss in competitiveness is not too severe (σ close to 1), the entry of liquidity constrained exporters will dominate the exit of low productivity unconstrained exporters, and the reduction in exports by existing exporters. ■

Chapter 3

The Dynamic Impact of Trade

Opening: Productivity Overshooting with Heterogeneous Firms

Summary 11 *In this chapter, I build a dynamic model of trade with heterogeneous firms which extends the work of Melitz (2003). As countries open up to trade, they will experience a productivity overshooting. Aggregate productivity increases in the long run, but it increases even more so in the short run. When trade opens up, there are too many firms, inherited from the autarky era. The most productive foreign firms enter the domestic market. Competition is fierce. The least productive firms that are no more profitable are forced to stop production. Not only do the most productive firms increase their size because they export, but the least productive firms stop producing altogether. Aggregate productivity soars. As time goes by, firms start to exit because of age. Competition softens. Some less productive firms resume production. This pulls down aggregate productivity. The slower the exit of firms, the larger this overshooting phenomenon. This model also predicts that the price compression that accompanies trade opening may be dampened in the long run. It also predicts that inequalities should increase at the time when a country opens up to trade, and then gradually recede in the long run.*

3.1 Introduction

There are strong empirical evidence that trade opening induces massive reallocations of the factors of production, not only between sectors, but even more so between individual firms within a given sector. In a series of papers, Bernard and Jensen (1999, 2001a, 2001b, 2002) acknowledge the importance of those reallocations between firms, and the importance of heterogeneity between firms with regard to exports. Bernard, Eaton, Jensen and Kortum (2003) show that differences in productivity among firms may explain the patterns of international trade. Exporters are more productive, larger, and more capital intensive than non exporters. Using a panel of French firms, Eaton, Kortum and Kramarz (2004a, 2004b) document the large differences in size and productivity between exporters and non exporters, and between multiple countries exporters and single country exporters. The reallocation of the factors of production from low productivity firms towards high productivity exporters when trade is opened up accounts for large variations in aggregate volatility.

There is little understanding of the transitional dynamics following trade opening. Empirically, we do observe that the short run impact of trade opening is much larger than the long run impact. The exit of firms in import competing sectors following a trade liberalization is typically large, but short lived. Many adjustments that are typical of trade liberalization episodes seem to be stronger in the short run than in the long run.

I build a dynamic model of international trade with heterogeneous firms that helps explain the dynamics of productivity after a country opens up to trade. As countries open up to trade, they will experience a productivity overshooting. Aggregate productivity increases in the long run, but it increases even more so in the short run. When trade opens up, there are too many firms, inherited from the autarky era. The most productive foreign firms enter the domestic market. Competition is fierce. The least productive firms that are no more profitable are forced to stop production. Not only do the most productive firms increase their size because they export, but the least productive firms stop producing altogether. Aggregate productivity soars. As time goes by, firms start to exit because of age. The situation improves. Some less productive firms resume production. This pulls down aggregate productivity. The slower the exit of firms, and the more competitive the economy, the larger this overshooting phenomenon. This model also predicts that the price compression that accompanies trade opening may be

dampened in the long run. It also predicts that inequalities should increase at the time when a country opens up to trade, and then gradually recede in the long run.

This model extends the pioneering work on trade with heterogeneous firms of Marc Melitz (2003). Melitz describes the long run impact of trade opening. He only considers the steady state properties of such a model with heterogeneous firms. By considering the transitional dynamics, I am able to see how differences between the mass of firms in the short run and in the long run may lead to a non monotonic response of aggregate productivity along the transition towards the new steady state. Hopenhaym (1992) also describes the reallocation of production between firms in an dynamic model, but considers only long run predictions. Eaton and Kortum (2002) develop a model of international trade with heterogeneous firms which extends the framework of Dornbusch, Fischer and Samuelson (1977) to a multi country setting. They only describe the steady properties of this model.

The model most closely related to this is Ghironi and Melitz (2005). They analyze the transitional dynamic in a model of trade with heterogeneous firms. They find that firm heterogeneity may explain systematic departures from purchasing power parity, and provides microfoundations for the Harrod-Balassa-Samuelson effect. In their model, they impose by assumption that no firm will be forced to exit when trade is opened up. The only dynamic comes from the entry and exit of firms into the export market, and the exogenous natural death among firms. I adopt a different formalization that allows me to account for the massive exit of firms at times of trade liberalizations. I am still able to describe the transition towards a steady state where some firms are allowed to resume production.

The mechanism generating productivity overshooting is intimately related to the overshooting model of Rudi Dornbusch (1976). Technically, it takes more time for the number of firms to adjust than for the relative size of firms. This is what explains the difference between short run and long run adjustments. In the short run, the number of firms cannot adjust discretely beyond a certain point. It only evolves sluggishly, as firms start dying. Therefore, in the short run, the relative size of firms must adjust in order to clear the labor market. The size of less productive firms shrinks whereas that of more productive firms increases. This shift of mass towards the most productive firms explains the large increase of aggregate productivity in the short run. As time goes by, the number of firms gradually adjusts, and the size of less pro-

ductive firms increases faster than that of more productive ones. Aggregate productivity falls towards its long run steady state.

The remainder of the paper is organized as follows. Section 3.2 introduces a simple dynamic model of trade with heterogeneous firms. Section 3.3 describes the transitional dynamics when trade is opened up. Section 3.4 concludes.

3.2 A simple model of trade with heterogeneous firms

I build up a simple model of trade with heterogeneous firms based on Melitz (2003). For simplicity, I use similar notations as the ones used by Melitz, and I add new ones when necessary.

The world is comprised of two identical countries, home and foreign. I will only consider symmetrical equilibria. Each country is populated with a mass L of workers. Those workers produce goods, earn wages and dividends, and consume. For simplicity, I assume that all workers own a single share in a mutual fund. The mutual fund owns all domestic firms, collects all their profits, invests in new firms when optimal, and redistributes all remaining profits to the workers. There are no international capital markets. Perfect competition on the labor market, and identical ownership in the mutual fund allow me to consider that everything is as if all decisions were undertaken by a representative consumer.

3.2.1 Demand

Each worker is endowed with one unit of labor that she supplies inelastically. I normalize wages (equal in both country) to one, and express all prices in terms of wages. Workers share the same intertemporal utility. They consume a CES aggregate of different goods in each period. If they consume a quantity $q_t(\varphi)$ of variety φ in period t , and all varieties in the set Φ_t , they derive a utility, $U_0 \equiv E_0 [\sum_{t=0}^{+\infty} \beta^t C_t]$, with $C_t = \left(\int_{\varphi \in \Phi} q_t(\varphi)^{\frac{\sigma-1}{\sigma}} d\varphi \right)^{\frac{\sigma}{\sigma-1}}$. The elasticity of substitution between any two varieties is constant and equal to σ , β is a subjective discount factor. For simplicity, I will consider the limiting case where $\beta \xrightarrow{\beta < 1} 1$, so that everything is as if $\beta = 1$, but the intertemporal utility is still well defined. The price P_t of one unit of the

composite good C depends on the price of each variety: $P_t = \left(\int_{\varphi \in \Phi_t} p_t(\varphi)^{1-\sigma} d\varphi \right)^{\frac{1}{1-\sigma}}$. Given those isoelastic preferences, the representative consumer will spend a fraction of its income on each differentiated variety. How much of each variety she consumes depends on the price of this variety relative to the price of others. The information about the price of all other varieties is summed up in the price of the composite consumption good P_t . If total expenditure on differentiated goods is R_t , the representative consumer spends $r_t(\varphi) = R_t (p_t(\varphi) / P_t)^{1-\sigma}$ on each variety φ .

3.2.2 Production and trade

Labor is the only factor of production. Production is done under increasing returns to scale. Each firm must pay an overhead cost each period. This fixed per period cost is identical for all firms. Firms are heterogeneous in terms of productivity. The marginal cost of production is constant for each firm, but differs across firms. Each firm draws a random labor productivity shock φ , meaning that the unit labor requirement is equal to $1/\varphi$. For simplicity, the productivity of a firm is fixed upon entry, and does not evolve over the life-span of the firm. The cost of producing q units of goods for a firm with productivity φ is $c(\varphi) = f + q/\varphi$.

When trade is allowed between the two countries, there two types of trade barriers. A fixed cost, and a variable cost. In order to enter the foreign market, a firm must pay a fixed cost f_x ¹. For simplicity, I assume that this fixed cost is paid each period. Having a sunk entry cost into the foreign market in addition to a fixed per period cost would not change the dynamics fundamentally. It would only slow down the adjustments of trade flows following the opening to trade between the two countries. The variable trade cost is a traditional "iceberg" transportation cost. If 1 unit of good is shipped between the two countries, only a fraction $1/\tau$ arrives. The larger τ , the more expensive transportation.

Each firm is a monopolist for its own variety. If it does export, it is allowed to charge different prices in each market. Given the production technology, and the technology of transportation

¹Outside of steady state, and unlike in Melitz (2003), it does matter whether the fixed cost of exporting is paid once and for all or at the beginning of each period. Those two formulations won't be equivalent anymore. I assume the cost is paid each period, which simplifies greatly the computation of the transitional dynamics. I believe results would not change qualitatively if I opted for the other formulation.

for international trade, and given that demand is isoelastic, firms will charge a constant mark-up over marginal cost. A firm with productivity φ , if it does survive, will charge a price $p_d(\varphi)$ on the domestic market; if it does export, it will charge a price $p_x(\varphi)$ on the foreign market:

$$p_d(\varphi) = \frac{1}{\rho\varphi} \text{ and } p_x(\varphi) = \frac{\tau}{\rho\varphi} \quad (3.1)$$

with $\frac{1}{\rho} = \frac{\sigma}{\sigma-1} > 1$ the mark-up charged by each firm.

I define the distribution of firm productivity at time t by $\mu_t(\cdot)$, with all firms above productivity φ_t^* selling on the domestic market, and all firms with a productivity $\varphi_x^*(t)$ exporting. Both the sequence of μ_t 's and of $\{\varphi^*, \varphi_x^*\}_t$'s will be determined in equilibrium. Plugging the prices set by each individual firm from Eq. 3.1, aggregate prices at time t are defined by,

$$(\rho P_t)^{1-\sigma} = \int_{\varphi_t^*}^{+\infty} \varphi^{\sigma-1} \mu_t(\varphi) d\varphi + \int_{\varphi_{x,t}^*}^{+\infty} \left(\frac{\varphi}{\tau}\right)^{\sigma-1} \mu_t(\varphi) d\varphi \quad (3.2)$$

I can therefore compute quantities sold by each firm, both at home and abroad, profits earned by firms, and from these profits, I know the set of firms that are able to survive, and the set of firms that are able to export². At time t , a firm with productivity φ , if it does produce at all, produces for the domestic market a quantity $q_{d,t}(\varphi) = \frac{R_t}{P_t} (P_t \rho \varphi)^\sigma$, so that its total domestic sales are $r_{d,t}(\varphi) = R_t (P_t \rho \varphi)^{\sigma-1}$, and the total profits it earns from selling domestically are $\pi_{d,t}(\varphi) = \frac{r_{d,t}(\varphi)}{\sigma} - f$. If this firm is able to export, it produces for the foreign market a quantity $q_{x,t}(\varphi)_{f.o.b.} = \frac{R_t}{P_t} \left(\frac{P_t \rho \varphi}{\tau}\right)^\sigma$ (or including the shipping cost, $q_{x,t}(\varphi)_{c.i.f.} = \frac{\tau R_t}{P_t} \left(\frac{P_t \rho \varphi}{\tau}\right)^\sigma$), so that its total foreign sales are $r_{x,t}(\varphi) = R_t \left(\frac{P_t \rho \varphi}{\tau}\right)^{\sigma-1}$, and the total profits it earns from exporting are $\pi_{x,t}(\varphi) = \frac{r_{x,t}(\varphi)}{\sigma} - f_x$.

Assumption 12 *In any period, a firm may decide not to produce any quantity. In such a case, it does not have to incur the overhead cost. In other words, a firm that wants to survive does*

²Note that I use the productivity φ as the identity of a firm. Literally, there is zero mass of firms with a productivity exactly equal to φ . In this continuum setting, we can say that the number of firms with a productivity φ is equal to $\mu(\varphi) d\varphi$. So potentially, there are "more than one" single firm with such a productivity. Formally, each of these firms has a different identity (each of them produces a unique differentiated variety). However they all behave in exactly the same way. They are indistinguishable from their actions. Hence I can safely abuse language and identify a firm by its productivity, φ .

not have to earn negative profits in order to stay in business³.

As in Melitz, I can define thresholds for domestic production, and for exports. No firm will produce quantities if it means earning non positive profits, and no firm will export if it means earning non positive profits from exporting. I define φ_t^* as the productivity of the least productive firm earning just non negative profits from domestic sales: $\pi_{d,t}(\varphi_t^*) = 0$. By the same token, $\varphi_{x,t}^*$ as the productivity of the least productive firm earning non negative profits from exporting: $\pi_x(\varphi_{x,t}^*) = 0$. I solve for the thresholds φ_t^* and $\varphi_{x,t}^*$. The conditions defining those threshold are the zero cutoff profit conditions (domestic and foreign):

$$\varphi_t^{*\sigma-1} = \frac{\sigma f}{R_t} \times (\rho P_t)^{1-\sigma} \quad (ZCP_t)$$

$$\varphi_{x,t}^{*\sigma-1} = \frac{\sigma f_x}{R_t} \times \left(\frac{\rho P_t}{\tau} \right)^{1-\sigma} = \left(\tau^{\sigma-1} \frac{f_x}{f} \right) \times \varphi_t^{*\sigma-1} \quad (ZCP_{x,t})$$

Any firm with a productivity below φ_t^* will not produce for the domestic market, and no firm with a productivity below the threshold $\varphi_{x,t}^*$ will export⁴. From those productivity thresholds, I can define the probability of exporting, conditional on survival: $p_{x,t} = \frac{P(\varphi > \varphi_{x,t}^*)}{P(\varphi > \varphi_t^*)}$.

3.2.3 Entry and exit of firms

The distribution of firms at any point in time is the result of a history of entry and exit of firms.

Entry is done in the following way. An entrepreneur may decide to start up a firm. In order to do so, she must pay a sunk entry cost f_e . Once this cost is paid, she receives a productivity shock φ , drawn from a random distribution with c.d.f. $g(\cdot)$ and p.d.f. $G(\cdot)$ defined over the support⁵ $[\varphi_{\min}, +\infty)$. Any firm that does not expect to earn positive profits in the future exits.

³This assumptions insures that the distribution of productivity among surviving firms is stationary. It greatly simplifies the description of the dynamic adjustments after opening to trade. I discuss relaxing this assumption (and forcing firms to pay overhead costs each period in order to stay in business) in appendix 3.5.1.

⁴I assume that $\frac{f_x}{f} \tau^{\sigma-1} > 1$, in other words, trade barriers are always sufficiently high so that only a subset of firms export: $\varphi_t^* < \varphi_{x,t}^*$. If this condition were violated, I would impose $\varphi_t^* = \varphi_{x,t}^*$, all active firms would export, and every result would carry through.

⁵The only condition on the distribution of productivity shocks is that the $(\sigma - 1)^{\text{th}}$ moment of G is defined, or that the integral $\int_{\varphi_{\min}}^{+\infty} \varphi^{\sigma-1} g(\varphi) d\varphi$ converges. This property ensures that the total size of the economy is finite. The choice of φ_{\min} is purely arbitrary. The assumption that the support for this distribution is unbounded from above simplifies notations greatly, but is only a notational assumption. It is perfectly admissible within this model that there is zero mass above a certain threshold φ_{\max} ($g(\varphi) = 0, \forall \varphi > \varphi_{\max}$).

In addition, all surviving firms have an exogenous probability δ of dying each period.

I assume free entry. If a firm with productivity φ earns total profits $\pi_t(\varphi)$ in period t (domestic profits, which may be equal to zero at some points in time, plus, for some firms, export profits), then the free entry condition at date t states that if there are firms that enter at time t , the value of entering, $v_{e,t}$, must equal the cost of entering, f_e ,

$$v_{e,t} = \int_{\varphi_{\min}}^{+\infty} \left(\sum_{s=t}^{+\infty} (1-\delta)^{s-t} \pi_s(\varphi) \right) g(\varphi) d\varphi = f_e \quad (FE_t)$$

If $f_e > \int_1^{+\infty} (\sum_{s=t}^{+\infty} (\beta(1-\delta))^{s-t} \pi_s(\varphi)) \mu_t(\varphi) d\varphi$, no firm enters. Free entry prevents the other inequality to ever happen. Condition FE_t holds as long as a strictly positive number of firms do enter. Note that it will be crucial to define the profit function $\pi_t(\cdot)$. This function is potentially complex along the transition following opening to trade. Firms may not earn any profit over some period of time, and then start earning positive profits. After some point in time, they may even start earning some extra profits from exporting.

The free entry condition and the general equilibrium will determine how many firms enter each period. Call $M_{e,t}$ the number of new entrants at time t . I must now impose that the labor market clears. Labor is used for investment (to cover the sunk entry cost of new entrants), and for production. The labor allocated to production is used both to pay for the fixed costs (fixed overhead cost for domestic production plus fixed trade barrier if the firm exports), and to cover the variable cost of production. The total labor used for investment⁶ is $f_e \times M_{e,t}$. If there is a total mass M_t of firms operating at time t , and a fraction $p_{x,t}$ of those firms are exporters, the total workforce used to cover overhead costs is $(f + p_{x,t}f_x) \times M_t$. Since each firm charges a constant mark-up over marginal cost, it can easily be proven that if total expenditures on differentiated goods is R_t , the total workforce used for producing differentiated goods (variable cost only) is ρR_t . The labor market clearing condition at each point in time during the transition requires,

$$L = f_e \times M_{e,t} + (f + p_{x,t}f_x) \times M_t + \rho R_t \quad (LMC_t)$$

⁶It is possible that at some point in time, there is no entry, so that $M_{e,t} = 0$.

I will see in the next two sections how the free entry characterizes the autarky and trade steady states. I will then see how entry of new firms take place along the transition from the autarky steady state to a new trade steady state.

3.2.4 Autarky steady state

In this section and the next, I recall Melitz (2003) computation of the steady state of this economy, both under autarky, and under trade. I will denote the autarky steady state by the time subscript $t = -\infty$, and the trade steady state by the time subscript $t = +\infty$.

In the steady state, there is as much entry as exit. All firm with a productivity below $\varphi_{-\infty}^*$ exit immediately upon receiving their productivity draw. Labor market clearing insures that total expenditure on differentiated goods is exactly equal to L . All profits are used to invest into starting up new firms, so that the mutual fund's finances are balanced. So in the autarky steady state, $(LMC_{-\infty}) \Leftrightarrow R_{-\infty} = L$.

Following Melitz (2003), I can define a special average productivity among active firms measure:

$$\tilde{\varphi}(\varphi^*) \equiv \left(\frac{\int_{\varphi^*}^{+\infty} \varphi^{\sigma-1} g(\varphi) d\varphi}{\int_{\varphi^*}^{+\infty} g(\varphi) d\varphi} \right)^{\frac{1}{\sigma-1}}$$

It will be useful to determine the total mass of firms, $M_{-\infty}$. Using the special average notation, the following accounting identity holds: $R_{-\infty} = M_{-\infty} \times r(\tilde{\varphi}_{-\infty})$. Since expected profits are constant over time, I can define average profits conditional on survival, $\bar{\pi}_{-\infty} = \pi_{d,-\infty}(\tilde{\varphi}_{-\infty})$. The steady state autarky equilibrium is defined by the zero cutoff profits condition (which defines $\varphi_{-\infty}^*$), the free entry condition, and the labor market clearing condition:

$$\begin{cases} (ZCP_{-\infty}) \\ (FE_{-\infty}) \\ (LMC_{-\infty}) \end{cases} \Leftrightarrow \begin{cases} \bar{\pi}_{-\infty} = f \left(\left[\frac{\tilde{\varphi}(\varphi_{-\infty}^*)}{\varphi_{-\infty}^*} \right]^{\sigma-1} - 1 \right) \\ \bar{\pi}_{-\infty} = \frac{\delta f_e}{P(\varphi > \varphi_{-\infty}^*)} \\ M_{-\infty} = \frac{L}{\sigma(\bar{\pi}_{-\infty} + f)} \end{cases}$$

The zero cutoff profit condition and the free entry condition define two schedules of $\bar{\pi}_{-\infty}$ as

a function of $\varphi_{-\infty}^*$, which have a unique intersection⁷. In every period, $\delta M_{-\infty}$ firms die from attrition, and $M_{e,-\infty} = \delta M_{-\infty} / P(\varphi > \varphi_{-\infty}^*)$ firms are created, among which only those firms with a productivity above $\varphi_{-\infty}^*$ survive, and replace the deceased firms.

In the next section, I describe the steady state that the economy will reach after trade between the two countries is opened up.

3.2.5 Trade steady state

In the trade steady state, there must also be as much entry as there is exit. Upon receiving their productivity shock, incumbents with a productivity below $\varphi_{+\infty}^*$ exit immediately. Among survivors, all firms with a productivity above $\varphi_{x,+\infty}^*$ export. The average profits that a firm earns, conditional on surviving, are the sum of profits earned domestically, and profits earned from exporting: $\bar{\pi}_{+\infty} = \pi_{d,+\infty}(\tilde{\varphi}_{+\infty}) + p_{x,+\infty}\pi_x(\tilde{\varphi}_x)$, where $p_{x,+\infty} = \frac{P(\varphi > \varphi_{x,+\infty}^*)}{P(\varphi > \varphi_{+\infty}^*)}$ is the probability of exporting, conditional on survival. The stationary assumption implies that the labor market clearing condition is the same in the trade steady state as in the autarky steady state: $(LMC_{+\infty}) \Leftrightarrow R_{+\infty} = L$. The trade steady state is defined by the two zero cutoff profits conditions (which define $\varphi_{+\infty}^*$ and $\varphi_{x,+\infty}^*$), the free entry condition, and the labor market clearing condition:

$$\begin{cases} (ZCP_{x,+\infty}) \\ (ZCP_{+\infty}) \\ (FE_{+\infty}) \\ (LMC_{+\infty}) \end{cases} \Leftrightarrow \begin{cases} \varphi_{x,+\infty}^* = \tau \left(\frac{f_x}{f} \right)^{\frac{1}{\sigma-1}} \varphi_{+\infty}^* \\ \bar{\pi}_{+\infty} = f \left(\left[\frac{\tilde{\varphi}(\varphi_{+\infty}^*)}{\varphi_{+\infty}^*} \right]^{\sigma-1} - 1 \right) + p_{x,+\infty} f_x \left(\left[\frac{\tilde{\varphi}(\varphi_{x,+\infty}^*)}{\varphi_{x,+\infty}^*} \right]^{\sigma-1} - 1 \right) \\ \bar{\pi}_{+\infty} = \frac{\delta f_e}{P(\varphi > \varphi_{+\infty}^*)} \\ M_{+\infty} = \frac{L}{\sigma(\bar{\pi}_{+\infty} + f + p_x f_x)} \end{cases}$$

As is the case in the autarky steady state, $\delta M_{+\infty}$ firms die from attrition each period, and are replaced by new entrants whose productivity is above $\varphi_{+\infty}^*$. One can easily prove that $\bar{\pi}_{+\infty} > \bar{\pi}_{-\infty}$, $\varphi_{+\infty}^* > \varphi_{-\infty}^*$ and $M_{+\infty} < M_{-\infty}$.

There are two important properties of the trade steady state to be noted. First, there are fewer firms when trade is allowed than under autarky, $M_{+\infty} < M_{-\infty}$. This is a crucial

⁷See Melitz (2003) for a formal proof of this statement.

prediction of the Melitz model of trade with heterogeneous firms. In autarky, firms that have access to a given technology are the only ones to compete for the home market. A given number of those firms can survive. As trade opens up, some highly productive firms are able to export. Those firms increase the toughness of competition. Fewer firms can survive. Those high productivity firms eat up an even larger share of the market. Mechanically, since those firms are more productive, fewer firms are required to service the entire market.

The second prediction, which is the central claim of the Melitz model, is that average productivity increases after trade is opened up. The most productive firms from abroad enter the domestic market, and push the least productive firms out of business. At the same time, the most productive domestic firms, who have access to the export market, benefit disproportionately more from the possibility of exporting. Not only does the productivity threshold for survival go up, but the share of the most productive firms increases compared to that of the least productive ones.

In the next section, I describe the transition towards the new steady state after both countries symmetrically open up to trade.

3.3 The dynamics of trade opening: productivity overshooting

In this section, I describe the transition of the economy from autarky towards trade. I first describe intuitively the forces driving the transitional dynamics, before turning to the formal derivation of those dynamics.

If firms are heterogeneous in terms of productivity, only a subset of firms, the most productive, are able to overcome trade barriers. The presence of these high productivity exporters, along with the upward shift in average productivity, implies that fewer firms can survive under trade than under autarky. In a sense, in the autarky steady state, domestic firms are alone to satisfy the entire domestic demand, and many firms must operate. When trade is opened up, there are "too many" firms. So during the transition, the mass of firms must shrink.

There is a fundamental asymmetry between creation and destruction of firms, due to the presence of sunk entry costs. Because I have assumed free entry, as soon as there is some potential for profits, there will always be some firms entering. Exit on the other hand may

take time. If the mass of firms at every level of productivity must shrink in order to reach the new steady state, because high productivity firms only die at a slow pace, the transition will be slow. Because existing firms have already paid the sunk entry cost, they are far less vulnerable than potential entrants. They will exit if and only if they do not expect to earn any positive profits at any point in the future.

At the moment when trade is opened up, the most productive firms start exporting. They eat up part of the domestic demand, and push many low productivity firms out of business. The least productive among those firms will never be able to generate positive profits ever again in this globalized world, and exit immediately. Competition is at its fiercest right at the time when trade is opened up. So upon opening to trade, there is a spike of destruction of firms. This implies a large increase in average productivity: only the most productive firms can survive in this new environment.

Because the world has inherited an "overcrowded" economy, there won't be any entry of firms for some period of time. During this transitional period, the natural death eats up the total mass of firms. Competition gradually softens. Firms with a low productivity that had been on hold until then can start producing again. Exporting becomes easier as the mass of firms shrinks. The mass of firms gradually shifts towards less productive firms that can more easily survive now.

Eventually, when the mass of firms has shrunk sufficiently, the new trade steady state is reached, with a lower average productivity than at the time of trade opening. Once this state is reached, entry starts again. One important property of this model is that the dynamics towards the trade steady state only take a finite amount of time. This crucially depends on assumption 12, which guaranties a stationary distribution of firm productivity.

In the next three sections, I derive formally the transitional dynamics, and their properties.

3.3.1 Transitional dynamics

Before opening to trade, the economy was in an autarky steady state, defined in section 3.2.4. Trade opens up at time $t = 0$. The opening is unexpected. Because of the dynamic nature of the model, there may be multiple rational expectations equilibria. I will only consider a class of dynamic equilibria, those that converge towards a steady state.

It will be useful to define the following alternative measure of the mass of firms: $\tilde{M}_t = M_t/P(\varphi > \varphi_t^*)$. This measure corresponds to the mass of firms per unit of density at each level of productivity. It also corresponds to an ideal total number of firms, which includes those firms that cannot survive (with a productivity below φ_t^*). Depending on the value of this alternative measure of mass in the trade steady state, there are two possible transitional paths. If $\tilde{M}_{+\infty} \geq \tilde{M}_{-\infty}$, the transition will be immediate. The economy jumps to the new trade steady state within one period. If $\tilde{M}_{+\infty} < \tilde{M}_{-\infty}$ on the other hand, the transition towards the new steady state takes a finite time.

The reason for this is simple. If there are more firms per level of productivity in the trade steady state than in the autarky steady state ($\tilde{M}_{+\infty} \geq \tilde{M}_{-\infty}$), competition is softer after trade is opened than it will be in the steady state, average profits are higher than in the steady state. I know that in the trade steady state, the discounted stream of profits is exactly equal to the sunk entry cost. Hence the appeal of extra profits will attract an influx of new firms, in order to restore the free entry condition (FE_t). Those firms are spread all over the distribution of productivity, so that the mass of firms at each level of productivity jumps immediately to its steady state level.

If on the other hand, there are fewer firms per level of productivity in the trade steady state than in the autarky steady state ($\tilde{M}_{+\infty} < \tilde{M}_{-\infty}$), competition is tougher after trade is opened than it will be in the steady state, average profits are lower. Since in the steady state, profits are just enough to cover the sunk entry cost, lower average profits implies that no firm will enter as long as $\tilde{M}_t \geq \tilde{M}_{+\infty}$. The natural death process gradually erodes the mass of firms, until the mass per level of productivity reaches its steady state level. From that point onward, the economy is in the trade steady state, and entry resumes in order to offset death from attrition.

I now turn to the formal proof of these statements. First, I define the criterion for fast or slow convergence towards the steady state. This criterion depends on whether the economy is overcrowded after trade is opened up or not.

Criterion 13 (Overcrowding)

$$\tilde{M}_{+\infty} < \tilde{M}_{-\infty}$$

Remarks: If criterion 13 is met, there are more firms at every level of productivity in the

autarky steady state than in the trade steady state (the economy is "overcrowded"), then the convergence towards the new steady state will take some finite amount of time, and there is overshooting in productivity. If criterion 13 is not met, then the economy immediately adjusts to its new trade steady state. See appendix 3.5.2 for the full functional form of this criterion.

Proposition 14 *If criterion 13 is not satisfied, that is if $\tilde{M}_{+\infty} \geq \tilde{M}_{-\infty}$, there is an influx of firms upon trade opening: $M_{e,0} = (\tilde{M}_{+\infty} - \tilde{M}_{-\infty}) + \delta\tilde{M}_{-\infty}$. From period $t = 1$ onward, the economy is in the trade steady state defined in section 3.2.5. The productivity thresholds immediately jump to their steady state values, $\varphi_0^* = \varphi_{+\infty}^*$, and $\varphi_{x,0}^* = \varphi_{x,+\infty}^*$.*

Proof. It is sufficient to prove that when trade is allowed and $M_{e,t} = 0$, $\tilde{M}_t < \tilde{M}_{+\infty} \Rightarrow P(\varphi > \varphi_t^*) \bar{\pi}_t > P(\varphi > \varphi_{+\infty}^*) \bar{\pi}_{+\infty}$.

Assume, as proven in appendix 3.5.2, that this property holds.

If $\tilde{M}_0 = \tilde{M}_{+\infty}$, the economy has already reached its new steady state, and no further adjustment occurs.

If the mass of firms per level of productivity were strictly lower than its trade steady state value, $\tilde{M}_0 < \tilde{M}_{+\infty}$, total expected profits would be larger than the sunk entry cost:

$$\begin{aligned} v_{e,0} &= \sum_{t \geq 0} (1 - \delta)^t \times P(\varphi > \varphi_0^*) \bar{\pi}_t \\ &> v_{e,+\infty} = \sum_{t \geq 0} (1 - \delta)^t \times P(\varphi > \varphi_{+\infty}^*) \bar{\pi}_{+\infty} = f_e \end{aligned}$$

Free entry prevents the occurrence of such an imbalance. So there will be a net entry of new firms that increases the mass of firms to its steady state level. In order to reach the trade steady state mass of firms, those firms destroyed by attrition must be replaced by new entrants: $\delta\tilde{M}_{-\infty}$ must enter. In addition, the ideal mass of firms (\tilde{M}) must be increased to $\tilde{M}_{+\infty}$, from $\tilde{M}_{-\infty}$. So the total number of entrants is $M_{e,0} = (\tilde{M}_{+\infty} - \tilde{M}_{-\infty}) + \delta\tilde{M}_{-\infty}$. From $t = 1$ onward, we are in the trade steady state, and new firms enter only to replace attrition deaths: $M_{e,t} = M_{e,+\infty} = \delta\tilde{M}_{+\infty}$.

Along this path, expected profits and survival thresholds are constant, $P(\varphi > \varphi_0^*) \bar{\pi}_0 = \dots = P(\varphi > \varphi_{+\infty}^*) \bar{\pi}_{+\infty}$, so that the free entry condition (FE_t) is satisfied for all $t \geq 0$. ■

Proposition 15 *If criterion 13 is satisfied, that is if $\tilde{M}_{+\infty} < \tilde{M}_{-\infty}$, there a finite length of time $T \geq 0$ such that as long as $t < T$, no firm enters. At time $t = T$, the trade steady state defined in section 3.2.5 is reached. T is uniquely defined as the minimum integer such that $\tilde{M}_T \leq \tilde{M}_{+\infty}$.*

Proof. Assume, as proven in appendix 3.5.2, that when $M_{e,t} = 0$, $\tilde{M}_t > \tilde{M}_{+\infty} \Rightarrow P(\varphi > \varphi_t^*) \bar{\pi}_t < P(\varphi > \varphi_{+\infty}^*) \bar{\pi}_{+\infty}$.

I need to prove that

$$\left\{ \begin{array}{l} M_{e,0} = M_{e,1} = \dots = M_{e,T-1} = 0 \\ M_{e,T} = \left(\tilde{M}_{+\infty} - \tilde{M}_{T-1} \right) + \delta \tilde{M}_{T-1} \\ M_{e,T+1} = \dots = M_{e,+\infty} = \delta \tilde{M}_{+\infty} \end{array} \right\}$$

is a rational expectations equilibrium. To do so, I prove that if firms expect that the economy will follow this path, indeed their entry decisions will follow these patterns.

As long as no entry takes place, the mass of firms evolves according to the following law of motion,

$$\tilde{M}_t = (1 - \delta) \tilde{M}_{t-1} \tag{M_t}$$

or $\tilde{M}_t = (1 - \delta)^t \tilde{M}_{-\infty}$. The mass of firms steadily declines. Define T as the minimum integer such that $\tilde{M}_T \leq \tilde{M}_{+\infty}$. Since $\tilde{M}_{-\infty} > \tilde{M}_{+\infty} > 0$, since \tilde{M}_t is strictly decreasing in t and converges towards zero, T is uniquely defined⁸. By definition of T , I know that $\tilde{M}_0 > \dots > \tilde{M}_{T-1} > \tilde{M}_{+\infty} \geq \tilde{M}_T$.

I now prove recursively that:

”at $t = T - 1$, no firm enters if agents expect the equilibrium path to be followed after $t = T - 1$ ”

”at $t = T - 2$, no firm enters if agents expect the equilibrium path to be followed after $t = T - 2$ ”

”...”

”at $t = 0$, no firm enters if agents expect the equilibrium path to be followed after $t = 0$ ”

⁸I consider the interesting case where $T > 0$. If $T = 0$, the adjustment to the new steady state is immediate. It is always possible to change the unit of observation for time, and hence to reduce δ sufficiently so that $T > 0$.

From $\tilde{M}_{T-1} > \tilde{M}_{+\infty} \Rightarrow P(\varphi > \varphi_t^*) \bar{\pi}_t < P(\varphi > \varphi_{+\infty}^*) \bar{\pi}_{+\infty}$, I know that $P(\varphi > \varphi_{T-1}^*) \bar{\pi}_{T-1} < P(\varphi > \varphi_{+\infty}^*) \bar{\pi}_{+\infty}$. If agents expect the equilibrium path to be followed after $t = T - 1$, from the free entry condition in the trade steady state ($v_{e,+\infty} = f_e$), I know that,

$$\begin{aligned} v_{e,T-1} &= \sum_{t \geq T-1} (1 - \delta)^t P(\varphi > \varphi_t^*) \bar{\pi}_t \\ &= P(\varphi > \varphi_{T-1}^*) \bar{\pi}_{T-1} + (1 - \delta) v_{e,+\infty} < f_e \end{aligned}$$

So if agents expect the equilibrium path to be followed after $t = T - 1$, no firm enters at $t = T - 1$.

By the same reasoning, if agents expect the equilibrium path to be followed after $t = T - 2$, $P(\varphi > \varphi_{T-2}^*) \bar{\pi}_{T-2} < P(\varphi > \varphi_{+\infty}^*) \bar{\pi}_{+\infty}$, and no firm enters at $t = T - 2$.

No firm enters until $t = T$, and from that point onward, the economy is in steady state.

Along the transition, the equilibrium will be determined by the zero cutoff profit conditions, the labor market clearing condition, and the law of motion for the mass of firms. Once the steady state is reached, the law of motion of the mass of firms is replaced by the free entry condition:

$$\begin{aligned} 0 \leq t < T & \left\{ \begin{array}{l} (ZCP_t) \\ (ZCP_{x,t}) \\ (M_t) \\ (LMC_t) \end{array} \right. \\ t \geq T & \left\{ \begin{array}{l} (ZCP_{x,+\infty}) \\ (ZCP_{+\infty}) \\ (FE_{+\infty}) \\ (LMC_{+\infty}) \end{array} \right. \end{aligned}$$

■

As can be guessed from the description of the dynamic evolution of average profits and the mass of firms, the productivity threshold jumps upon opening to trade, and then gradually falls towards its steady state level. I describe this phenomenon of productivity overshooting in the

next section.

3.3.2 Productivity overshooting

In this section, I will only consider the interesting case where criterion 13 is met, so that transitional dynamics are not collapsed into one single period.

As trade opens up, there are too many firms. Competition is fierce, profits are low. Many firms are forced to stop producing, since they could not even cover their overhead costs. Only the most productive firms are still active, and the most productive among them are exporters. As time goes by, since no firm enters during the transition, firms at every level of productivity start dying. For the survivors, the situation gradually improves. This means that some firms can start exporting. Those new exporters experience a discrete increase in the volume of their sales, and their employment. As the productivity threshold for exports falls, mass is shifted towards those less productive exporters. At the same time, some firms that had stopped producing altogether resume their production. Some of the varieties that had disappeared when the most productive foreign exporters had entered start being produced again. The threshold for domestic sales falls.

Aggregate productivity falls down for two reasons. First, those firms that resume production have a low productivity, and they pull down average productivity. At the same time, the mass of sales is shifted gradually away from the high productivity exporters towards the new lower productivity exporters. So as more and more firms die, aggregate productivity falls as well. Eventually, the economy reaches its steady state when productivity stays constant.

Following trade opening, aggregate productivity increases in the long run. This is the main prediction of the Melitz model. I predict that in the short run, productivity will increase more than in the long run. This rapid increase in productivity followed by a gradual deterioration of productivity is what I call productivity overshooting. The forces driving this overshooting in productivity are very similar in spirit to the forces driving exchange rate overshooting in the Dornbusch model. In Dornbusch (1976), the reason why exchange rate overshoots is because exchange rates can adjust much faster than domestic prices. In this model, productivity overshoots because it takes time for firms that are already here to die. Because those existing firms will not die right away, something else must adjust in the meantime to offset the imbalances

created by the sudden entry of foreign exporters. This is done through a temporary reduction in the share of low productivity firms. Some of these less productive firms stop producing altogether for some time (until sufficiently many high productivity firms have died, and the situation has sufficiently improved). Some other firms do not stop producing, but they reduce their production relative to higher productivity firms (which start exporting). This adjustment, the reduction in the share of less productive firms, can happen much faster than forcing existing firms out. So in the short run, this will be the only variable of adjustment. In a sense, one could say that the extensive margin of productivity (how many levels of productivity can be active) can adjust much faster than the intensive margin of productivity (how many firms there are at each level of productivity). This reduction in the share of less productive firms in the short run will cause productivity to increase substantially. In the longer run, as firms die and the mass of firms shrinks, this increase in aggregate productivity is dampened. The death of firms allows to the share of less productive firms to increase again, partially offsetting the short run productivity gains.

Note that some firms will exit definitively at the time when trade is opened up. Those are the firms that will not survive, even once the new steady state is reached. If criterion 13 does not hold, there will actually be no overshooting, and productivity directly jumps to its steady state level. This is because I do allow for immediate exit of firms when trade is opened up. Firms have to pay a fixed overhead production cost each period, and therefore some firms, after trade is opened up, know that they will never be profitable again, and exit immediately. If I remove the assumption of a fixed overhead cost, as is done in a similar setting by Ghironi and Melitz (2005), adjustments will be even more sluggish, and productivity overshooting will potentially be much larger. In such a setting, the prediction that the total mass of firms must be reduced still holds. There is no possibility of discretely adjusting this mass, so the adjustment will take much more time than in the current model. I believe however that the prediction that opening up to trade will have a sudden negative impact on the least productive domestic firms is a plausible feature. It is interesting to see that even when some firms exit, transition towards the new steady state may still take some time, and we may observe a phenomenon of productivity overshooting.

I will not go into the details of computing the average productivity of firms (weighted by

the size of sales, or by employment) in the economy along the transition. I will only prove that along the transition, as the mass of firms shrinks, the productivity threshold for survival falls. It is intuitive to see that as the productivity threshold falls towards its new steady state level, aggregate productivity also falls. The following proposition proves formally the productivity overshooting triggered by trade opening.

Proposition 16 (Overshooting) *If criterion 13 is met, there is overshooting in productivity,*

$$\varphi_0^* > \varphi_1^* > \dots > \varphi_T^* = \dots = \varphi_{+\infty}^* > \varphi_{-\infty}^*$$

Proof. We already know from Melitz (2003) that productivity is higher in the trade steady state than in the autarky steady state: $\varphi_{+\infty}^* > \varphi_{-\infty}^*$

If criterion 13 is met, I know that the mass of firms per level of productivity, \tilde{M}_t , will gradually fall from its autarky level, towards its trade steady state level (at a constant rate $(1 - \delta)$):

$$\tilde{M}_{-\infty} > \tilde{M}_0 > \tilde{M}_1 > \dots > \tilde{M}_T = \dots = \tilde{M}_{+\infty}$$

See appendix 3.5.2 for a proof that $\tilde{M}_t > \tilde{M}_s \Rightarrow \varphi_t^* > \varphi_s^*$. ■

It is important for this overshooting phenomenon to happen that the opening to trade is unexpected. Looking at the transitional dynamics from an ex ante point of view, we can see further justifications for this overshooting in productivity. There are too many firms at the time when trade is opened up. The profitability conditions worsen in such a way that no firm has any incentive to enter during the transition towards the new steady state. This implies that entrepreneurs who started up their company before the opening to trade had formed the wrong expectations regarding their future expected stream of profits. Expected profits will fall below their steady state level for some time. Had those entrepreneurs known in advance about this opening to trade, they would not have started up their company⁹. This means that there was overinvestment in autarky.

In other words, the "technology" of production has improved, in an ex ante expectation sense, because there has been overinvestment in improving the available technology (creation

⁹I describe qualitatively what happens when the opening to trade is announced in advance in appendix 3.5.3.

of too many firms). This is why less investment (payments to start up new firms) is required each period in the trade steady state than in the autarky steady state. Since the economy had overinvested while in autarky, it is natural that the opening to trade creates a temporary spike in productivity. This is only temporary. As the economy converges towards its new steady state, there is disinvestment (no creation of new firms while old ones die). Thanks to the available "better technology" offered by trade, the aggregate productivity in the new steady state is higher than in autarky. But it is lower than during the transitional period, which was a time of abundance of "capital".

3.3.3 Price compression and inequality

As trade opens up, the dynamics of productivity are parallel to movements in prices and inequalities. Such a model with heterogeneous firms is perfectly fit for describing the impact of trade opening on aggregate prices as well as inequalities (between firms, in the absence of any formalization of wage bargaining).

As trade opens up, prices are driven down by the entry of high productivity foreign exporters. Literally speaking, since the model imposes that mark-ups are constant, no single firm changes the price it sets. However, low productivity/ high price firms exit, and since the share of high productivity/ low price firms increases, aggregate prices fall when trade is opened up. This fall in aggregate prices is exactly symmetrical to average productivity increases. In this model, productivity is directly a mirror image of prices. This fall in aggregate prices happens in the long run.

But exactly in the same way as there is overshooting in productivity, there will be "overcompression" of prices in the short run. In the short run, only the most productive firms are active. So prices are pushed down substantially. As firms exit, and as competition softens along the transition path towards the new steady state, less productive firms resume production. Those firms charge a higher price than their competitors. So average prices will "undershoot", in the sense that there is a sudden drop of prices in the short run, and a gradual price increase in the medium run.

Melitz (2003) points out that in the presence of firm heterogeneity, inequalities between firms increase. This is due to the fact the more productive firms benefit from trade, whereas

the situation of less productive firms worsens. So inequalities increase in the long run. But once again, inequalities increase more in the short run than in the long run. There is overshooting in inequalities. In the short run, only the very best firms benefit from the access to foreign markets, whereas many less productive firms have to incur a temporary loss in profits (which goes all the way to zero profits for some time for firms with a productivity $\varphi_{+\infty}^* < \varphi < \varphi_0^*$). In the short run, inequalities soar. In the medium run, as some firms exit, inequalities among survivors recede, until the new steady state is reached. Inequalities will always be higher in a trade regime than in autarky, but less so in the long run than in the short run. It takes some time for the economy to adapt to the new regime.

3.4 Conclusion

In this chapter, I have shown that when a country opens up to trade, it will experience an sudden increase in productivity, followed by a gradual reduction towards a steady state. Opening up to trade does lead to a permanent increase of productivity, but there is an even larger increase in the short run. I call this phenomenon productivity overshooting. The reason for this productivity overshooting is simple. Because trade allows a more efficient allocation of factors of production, there will be fewer firms in a globalized world than in a world composed of countries in autarky. After countries open up to trade, the number of firms will have to fall. However, in the presence of sunk entry costs, such a reduction in the number of firms takes time. In the short run, competition is fierce, and many firms cannot profitably produce. Only the most productive firms are active. Aggregate productivity soars. As the number of firms gradually falls (as existing firms age), competition softens, and low productivity firms are able to resume production. Aggregate productivity gradually falls as the world economy converges towards its new steady state. Along with this productivity overshooting, I expect to observe price compression in the short run, followed by a gradual increase in aggregate prices, as well as an overshooting in inequalities.

Bibliography

- [1] Bernard, Andrew B.; Jonathan Eaton; J. Bradford Jensen and Sam Kortum, 2003, "Plants and Productivity in International Trade," *American Economic Review*, v. 93, pp. 1268-1290.
- [2] Bernard, Andrew B. and J. Bradford Jensen, 1999, "Exceptional exporter performance: cause, effect, or both?", *Journal of International Economics*, v. 47, pp. 1-25.
- [3] Bernard, Andrew B. and J. Bradford Jensen, 2001a, "Why Some Firms Export," *The Review of Economics and Statistics*, v. 86 no. 2.
- [4] Bernard, Andrew B. and J. Bradford Jensen, 2001b, "Who Dies? International Trade, Market Structure, and Industrial Restructuring," *NBER Working Paper # 8327* .
- [5] Bernard, Andrew B. and J. Bradford Jensen, 2002, "Entry, Expansion, and Intensity in the US Export Boom, 1987-1992," *Review of International Economics*, forthcoming.
- [6] Dornbusch, Rudiger, 1976, "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, v. 84, pp. 1161-76.
- [7] Dornbusch, Rudiger; S. Fischer, and Paul A. Samuelson, "Comparative Advantage, Trade, and Payments in a Ricardian Model with a continuum of Goods," *American Economic Review*, vol. 67, pp. 823-839.
- [8] Eaton, Jonathan; Sam Kortum, 2002, "Technology, Geography, and Trade," *Econometrica*, vol. 70 no. 5 (September 2002), pp. 1741-1779.
- [9] Eaton, Jonathan; Sam Kortum, and Francis Kramarz, 2004, "An Anatomy of International Trade: Evidence from French Firms," *American Economics Review, Papers and Proceedings*, forthcoming.

- [10] Eaton, Jonathan; Sam Kortum, and Francis Kramarz, 2004, "Dissecting Trade: Firms, Industries, and Export Destinations," *NBER Working Paper 10344*.
- [11] Ghironi, Fabio, and Marc J. Melitz, 2005, "International Trade and Macroeconomic Dynamics with Heterogeneous Firms." *Working Paper*.
- [12] Helpman, Elhanan; Marc Melitz and Yona Rubinstein, 2004, "Trading Partners and Trading Volumes," *Working Paper*.
- [13] Hopenhaym, Hugo, 1992, "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, v.60, pp. 621-653.
- [14] Krugman, Paul. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review*, vol. 70, pp. 950—959.
- [15] Melitz, Marc, 2003, "The Impact of Trade on Aggregate Industry Productivity and Intra-Industry Reallocations," *Econometrica*, v. 71 no. 6 (November 2003), pp. 1695-1725.

3.5 Appendix

3.5.1 Non stationary productivity distribution and dynamic adjustments

Absent assumption 12, the distribution of productivity among firms during the transition may be non stationary.

The reason is the following. When trade opens up, the productivity threshold for domestic sales soars. In this paper, I have described what happens when existing firms have the option of not producing at all during some period, without going out of business (that is without having to pay the sunk entry cost, get a productivity draw from a lottery, all over again). Therefore, the exit decision was simple. All firms with a productivity above $\varphi_{+\infty}^*$ stay in business. because they know that in the long run (once the trade steady state is reached), they will be able to generate some strictly positive profits. However, some of these firms have to stay idle for some time. If those firms had to pay a fixed cost each period in order to stay in business, they would have to gauge the cost of staying in business (a given number of periods with negative profits), against the expected benefits from staying in business (expected positive profits starting up in the future). The least productive among those firms will exit. It is easy to prove that there is a range of productivity $[\varphi_{+\infty}^*; \bar{\varphi}_0]$ over which firms exit immediately.

Let us now move along the transition. Firms die (there were too many firms to start with), the productivity threshold for domestic sales goes down. Eventually, this threshold goes down below $\bar{\varphi}_0$. Then, with each new cohort of entrants, some firms with a productivity below $\bar{\varphi}_0$ will survive. But until that point, there were no firms with such a low productivity, whereas there is a strictly positive number of firms with a productivity above $\bar{\varphi}_0$ that have survived. Because of the nature of the technology, there will be as many new firms with $\varphi > \bar{\varphi}_0$ as firms with $\varphi < \bar{\varphi}_0$ created (up to some relative probability). But since to start with, the stock of firms with $\varphi > \bar{\varphi}_0$ is strictly positive, and there are no firms with $\varphi < \bar{\varphi}_0$, there is an imbalance. The distribution of productivity among existing firms is no longer given by $g(\cdot) \times (Constant)$. The distribution is tilted towards high productivity firms. There are disproportionately many firms with a high productivity. Those are the firms that survived the opening to trade.

In such a configuration, the steady state cannot be reached in a finite number of periods. The disequilibrium of the distribution towards high productivity firms gradually fades away, as

those old firms who survived since the time when trade was opened for the first time die, and are replaced with new cohorts.

However, if the distribution of firms is non stationary, it won't be possible to determine even one equilibrium path. Iterative numerical methods cannot be applied successfully. In order to know how many firms enter each period from say t onward, I must guess a path for the future beyond t . But what equilibrium is reached at time t depends on the current distribution of firms, which is inherited from the past. So I must also guess what has happened in the past. The equilibrium solution is both forward looking and backward looking. The only possibility is to assume that the steady state is reached at a given point in time, say $t = T$ large (and assume that adjustments beyond that point are negligible), guess an entire path for the entry of firms $\{M_{e,0}^{(1)}; \dots; M_{e,T}^{(1)}\}$, solve for the equilibrium in each period given this path of entry has been followed in the past and is expected to be followed in the future, and extract what would be the optimal path entry of forward looking agents, $\{M_{e,0}^{(2)}; \dots; M_{e,T}^{(2)}\}$, and so on until this algorithm converges. The problem of such an algorithm is that it is computationally demanding, since each iteration (from $\{M_{e,0}^{(n)}; \dots; M_{e,T}^{(n)}\}$ to $\{M_{e,0}^{(n+1)}; \dots; M_{e,T}^{(n+1)}\}$) requires to solve the entire sequence of T equilibria simultaneously. Moreover, in such a model, the strategy of a firm may potentially be complex: some firms may decide to incur negative profits for some time because they expect to earn positive profits in the future. The exit decision of firms is therefore a forward looking decision.

3.5.2 Criterion 13 and proofs of propositions 14, 15 and 16

Criterion 13 (reminded) The full functional form of this criterion¹⁰ is:

¹⁰With Pareto distributed shocks, that is if $P(\varphi > \varphi^*) = (\varphi^*/\varphi_{\min})^{-\gamma}$, with $\gamma > 1$ some scaling parameter, this criterion is not met, and there is an exact equality. This implies that with Pareto distributed shocks, the adjustment towards the new steady state is immediate. There are as many high productivity firms ($\varphi > \varphi_{+\infty}^*$) in autarky and in the steady state. There are as many firms that enter (and die from attrition plus immediate exit) each period in both regimes. The only difference is that under the trade regime, fewer firms survive. But, with Pareto distributed productivity shocks, the distribution among survivors is exactly identical under trade and under autarky.

$$\frac{\int_{\varphi_{-\infty}^*}^{+\infty} (\varphi^{\sigma-1} - \varphi_{-\infty}^{*\sigma-1}) dG(\varphi)}{\int_{\varphi_{-\infty}^*}^{+\infty} \varphi_{-\infty}^{*\sigma-1} dG(\varphi)} > \frac{f \frac{\int_{\varphi_{+\infty}^*}^{+\infty} (\varphi^{\sigma-1} - \varphi_{+\infty}^{*\sigma-1}) dG(\varphi)}{\int_{\varphi_{+\infty}^*}^{+\infty} \varphi_{+\infty}^{*\sigma-1} dG(\varphi)} + p_{x,+\infty} f_x \frac{\int_{\varphi_{x,+\infty}^*}^{+\infty} (\varphi^{\sigma-1} - \varphi_{x,+\infty}^{*\sigma-1}) dG(\varphi)}{\int_{\varphi_{x,+\infty}^*}^{+\infty} \varphi_{x,+\infty}^{*\sigma-1} dG(\varphi)}}{f \frac{\int_{\varphi_{+\infty}^*}^{+\infty} \varphi^{\sigma-1} dG(\varphi)}{\int_{\varphi_{+\infty}^*}^{+\infty} \varphi_{+\infty}^{*\sigma-1} dG(\varphi)} + p_{x,+\infty} f_x \frac{\int_{\varphi_{x,+\infty}^*}^{+\infty} \varphi^{\sigma-1} dG(\varphi)}{\int_{\varphi_{x,+\infty}^*}^{+\infty} \varphi_{x,+\infty}^{*\sigma-1} dG(\varphi)}}$$

I now turn to the missing parts in the proofs leading to the productivity overshooting proposition. I will first prove proposition 16: a reduction in the mass of firms leads to a fall in the productivity threshold for domestic sales. This will allow me to prove the missing part in the proof of propositions 14 and 15: a reduction in the mass of firms per level of productivity is equivalent to an increase in expected per period profits.

Proposition 17 (i) *Under the trade regime, and when $M_{e,t} = 0$,*

$$\tilde{M}_t > \tilde{M}_s \Rightarrow \varphi_t^* > \varphi_s^*$$

(ii) *Under the trade regime, and when $M_{e,t} = 0$,*

$$\tilde{M}_t < \tilde{M}_s \Leftrightarrow P(\varphi > \varphi_t^*) \bar{\pi}_t > P(\varphi > \varphi_s^*) \bar{\pi}_s$$

Proof. (i) The equilibrium along the transition is determined by four conditions: the zero cutoff profit conditions, (ZCP_t) and $(ZCP_{x,t})$, the labor market clearing condition, (LMC_t) , and the law of motion for the mass of firms which determines the value of \tilde{M}_t at each point along the transition, (M_t) .

Since the distribution of firms is stationary, I know that $\mu_t(\varphi) = \tilde{M}_t g(\varphi)$. I can therefore compute the price index along the transition path. Plugging the zero cutoff profits condition for exporters, $(ZCP_{x,t})$, into the price index in Eq. (3.2), and using the definition of μ_t along the transition, and rearranging, I get,

$$(\rho P_t)^{1-\sigma} = \frac{\varphi_t^{*\sigma-1} M_t}{f} \left(f \left(\frac{\tilde{\varphi}(\varphi_t^*)}{\varphi_t^*} \right)^{\sigma-1} + p_{x,t} f_x \left(\frac{\tilde{\varphi}(\varphi_{x,t}^*)}{\varphi_{x,t}^*} \right)^{\sigma-1} \right)$$

Plugging this price index into the zero cutoff profit condition for domestic sales, (ZCP_t), I get,

$$R_t = \sigma M_t \left(f \left(\frac{\tilde{\varphi}(\varphi_t^*)}{\varphi_t^*} \right)^{\sigma-1} + p_{x,t} f_x \left(\frac{\tilde{\varphi}(\varphi_{x,t}^*)}{\varphi_{x,t}^*} \right)^{\sigma-1} \right)$$

Plugging this into the labor market clearing condition, (LMC_t), I can solve for the aggregate demand, R_t , and rearranging, I get an implicit relationship between the mass \tilde{M}_t and the productivity threshold φ_t^* ,

$$L = M_t [(\sigma - 1) \bar{\pi}_t + \sigma (f + p_{x,t} f_x)]$$

or $L = \tilde{M}_t \left[f \int_{\varphi_t^*}^{+\infty} \left((\sigma - 1) \left(\frac{\varphi}{\varphi_t^*} \right)^{\sigma-1} + 1 \right) dG(\varphi) + \int_{\varphi_{x,t}^*}^{+\infty} \left((\sigma - 1) f \left(\frac{\varphi}{\tau \varphi_t^*} \right)^{\sigma-1} + f_x \right) dG(\varphi) \right]$

$\left((\sigma - 1) \left(\frac{\varphi}{\varphi_t^*} \right)^{\sigma-1} + 1 \right)$ and $\left((\sigma - 1) f \left(\frac{\varphi}{\tau \varphi_t^*} \right)^{\sigma-1} + f_x \right)$ are positive and increasing in φ_t^* and $\varphi_{x,t}^*$, respectively. $\varphi_{x,t}^*$ is increasing in φ_t^* . So as \tilde{M}_t falls along the transition towards the trade steady state, φ_t^* must fall in order to preserve the equilibrium on the labor market.

So I have proven that along the transition, φ_t^* increases as the mass of firms \tilde{M}_t falls.

(ii) I will now prove that as φ_t^* increases, $P(\varphi > \varphi_t^*) \bar{\pi}_t$ decreases.

Plugging the zero cutoff profit conditions (ZCP_t) and ($ZCP_{x,t}$) into the formula for average profits, I get:

$$\bar{\pi}_t = f \left(\left(\frac{\tilde{\varphi}_t}{\varphi_t^*} \right)^{\sigma-1} - 1 \right) + p_{x,t} f_x \left(\left(\frac{\tilde{\varphi}_{x,t}}{\varphi_{x,t}^*} \right)^{\sigma-1} - 1 \right)$$

Multiplying by the probability of survival, I get,

$$P(\varphi > \varphi_t^*) \bar{\pi}_t = f \int_{\varphi_t^*}^{+\infty} \left[\left(\frac{\varphi}{\varphi_t^*} \right)^{\sigma-1} - 1 \right] dG(\varphi) + f_x \int_{\varphi_{x,t}^*}^{+\infty} \left[\left(\frac{\varphi}{\varphi_{x,t}^*} \right)^{\sigma-1} - 1 \right] dG(\varphi)$$

Since $\left[\left(\frac{\varphi}{\varphi_t^*} \right)^{\sigma-1} - 1 \right]$ and $\left[\left(\frac{\varphi}{\varphi_{x,t}^*} \right)^{\sigma-1} - 1 \right]$ are non negative and decreasing in φ_t^* and $\varphi_{x,t}^*$, respectively, and since $\varphi_{x,t}^*$ is increasing in φ_t^* , $P(\varphi > \varphi_t^*) \bar{\pi}_t$ is decreasing in φ_t^* . ■

3.5.3 Expected opening to trade

If the opening to trade is announced in advance, productivity overshooting will be dampened. Indeed, the very reason why there is overshooting in productivity is because the economy has overaccumulated capital while in autarky, in the sense that too many firms have been started. The number of firms must go down to its steady state level, and the absence of entry during the transition explain why some less productive firms have to temporarily stop their production, and therefore why aggregate productivity overshoots in the short run. If the opening to trade is expected, the economy will start disinvesting in advance. Fewer firms are created each period, so that new firms are not sufficient to replace exit each period. There is net exit of firms even before trade is opened up. This is because rational expectation agents do know that their investment (creation of a new firm) would not be worthwhile at least for some time. The speed at which the economy disinvests obviously depends on the rate at which existing firms die (the coefficient δ), the impatience of entrepreneurs (the discount factor β), as well as the horizon at which trade will be opened up. However, unless trade opening has been too early, there will still be a period of productivity overshooting. There are too many of those firms that had been created before trade opening were announced, and those firms will have to disappear.

We do observe that even those trade agreements that have been have been negotiated long in advance, still have a large impact at the time when they are implemented.