Unpacking Trade Costs: Theory and Evidence

Inaugural-Dissertation

zur Erlangung des Doktorgrades

an der Wirtschaftswissenschaftlichen Fakultät

der Eberhard-Karls-Universität Tübingen

vorgelegt von

Benjamin Jung

aus Frankfurt/Main

2009

Dekanin:	Professor Dr. Kerstin Pull
Erstberichterstatter:	Professor Dr. Wilhelm Kohler
Zweitberichterstatterin:	Professor Dr. Claudia Buch
Tag der mündlichen Prüfung:	21.07.2009

Acknowledgements

I am grateful to my supervisor, Wilhelm Kohler, for his support. I am fortunate to have Gabriel Felbermayr as advisor and co-author. His guidance and insights has provided essential support for my work.

I greatly appriciate financial support from the Deutsche Forschungsgemeinschaft (DFG) through a PhD grant. Part of this doctoral thesis was written when I was visiting scholar at The Leverhulme Centre for Research on Globalisation and Economic Policy (GEP), University of Nottingham. The visit was generously supported by The Leverhulm Trust.

I am grateful to all my colleagues and friends in Tübingen and Hohenheim with whom I shared all my academic and non-academic moments. They all provided me with great support. I have greatly benefited from comments, criticism, and ideas during conferences, workshops, and research seminars at the Universities of Athens, Champain-Urbana, Göttingen, Graz, Hamburg, Hohenheim, Munich, Nottingham, Paderborn, Tübingen, and Vienna, and at the Institute for Applied Economic Research (IAW). In particular, I am grateful to Alan Deardorff for his discussion of an early draft of my thesis.

Last but not least I would like to thank my family for all their love, patience, and encouragement.

Contents

Chapter 1

Introduction

There is a common perception that trade liberalization unlocks welfare gains. Consequently, politicians have undertaken a huge effort to abolish tariff barriers. Trade impediments, however, may appear in various guises. In their survey article, Anderson and van Wincoop (2004) argue that trade costs other than direct policy measures are important. Among others, they refer to costs associated to regulatory barriers, the lack of enforceable contracts, and information barriers. On total, trade costs of an average rich country amount to an alarming 170% ad valorem tax equivalent.

This doctoral thesis aims at contributing to the ongoing process of unpacking trade costs.[1] Drawing on theoretical and empirical approaches, it analyzes the role of trading frictions to shed light on the *missing trade puzzle* discussed in the literature; see Trefler (1995). The thesis is organized along the lines of different types of trade barriers.

Technical barriers to trade. Trade economists traditionally study the effect of lower variable trade costs. Technical barriers to trade (TBTs) have become increasingly important politically but received less attention yet. Chapter 2 therefore views TBTs as *fixed regulatory costs* related to the entry into export markets. It develops a model with heterogeneous firms, trade in differentiated goods, and variable external economies of scale to sort out the rich interactions between TBT reform, input diversity, firm-level productivity, and aggregate productivity. Moreover, it offers a calibration for 14 industries in order to clarify the

[1] The term 'unpacking trade costs' was coined by Harrigan and Venables (2006).

theoretical ambiguities. Overall, the results tend to suggest beneficial effects of TBT reform but also reveal interesting sectoral variation.

Information barriers to trade. Rauch and Trindade (2002) argue that migration networks reduce trade costs through "provision of market information, matching, and referral services" (p. 118). They find that ethnic Chinese networks of the magnitude observed in Southeast Asia increase bilateral trade by at least 60%. Chapter 3, however, shows that this estimate is upward biased due to omitted variable bias. Moreover, it is partly related to a preference effect rather than to the availability of information. Applying a theory-based gravity model to ethnicity data for 1980 and 1990, and focusing on pure network effects, we find that the Chinese network leads to a more modest amount of trade creation of about 15%. Using new data on bilateral stocks of migrants from the World Bank for the year of 2000, we extend the analysis to all potential ethnic networks. There is, i.a., evidence for a Polish, a Turkish, a Mexican, or an Indian network. While confirming the existence of a Chinese network, its trade creating potential is dwarfed by other ethnic networks.

Lucas (2006) argues that the empirical link between migration and trade is subject to "spurious associations" (p. 373). Chapter 4 therefore aims at sorting out confounding factors.[2] Using newly available panel data on developing countries' diaspora to rich OECD nations in a theory-grounded gravity model, we uncover a robust, causal pro-trade effect. Moreover, we do not find evidence in favor of strong differences across education groups.

Migrant networks are an imperfect measure of available information about a foreign market. Moreover, there are identification problems. Empirically one can hardly separate the information channel from the preference channel. Chapter 5 therefore suggests a different measure of information. We use 15 years of Reuters-Business-Briefing (RBB) data in an empirical gravity model to analyze how business news about potential trading partners' countries affect the volume of bilateral trade. The exercise is motivated by a theoretical model where exporters sink fixed costs before engaging into a trade relationship. They are thus exposed to expropriation risk and need to form expectations. We hypothesize that news on other agents' experiences in the respective country affect exporters' subjective beliefs.

[2] Chapter 4 studies the effect of migration networks in general and does not explicitly separate the information cost channel from the preference channel.

Instrumental variable regressions indicate that cooperative and non-cooperative messages robustly affect trade flows, but that non-cooperative measures matter more strongly.

Incomplete cross-border contracts. Exporters typically use trade intermediaries or establish own foreign wholesales representations. Standard trade models ignore this fact. Chapter 6 studies the trade-off between the impact of imperfect cross-border contracts if exporters opt for a trade intermediary and higher fixed costs of market access if they set up an own wholesale affiliate. The optimal choice of export modes is embedded in a general equilibrium model that shows an interesting sorting pattern of heterogeneous firms according to their degree of competitive advantage. Relative prevalence of trade intermediation depends on the risk of expropriation, the severity of contractual frictions, the degree of heterogeneity, and the elasticity of demand. Trade and wholesale FDI are complements. Econometric analysis confirms these predictions.

Search and matching frictions in business-to-business relationships. Chapter 7 differs from Chapter 6 in assuming that there are search and matching frictions in the relationship between an exporter and a trade intermediary, while contracts are complete. It proposes a general equilibrium business-to-business matching model in which heterogeneous producers may seek a foreign general importer. Importantly, market access costs and the size of the non-tradables sector are endogenously determined.

Trust in bilateral trade relationships. Repeated interactions between trading partners lead to "the accumulation of invisible assets" (De Benedictis and Vicarelli, 2005). One possible interpretation of these asset is trust. Whereas in Guiso et al. (2009) trust is given exogenously, Chapter 8 incorporates the endogenous accumulation of an 'invisible asset' into a standard Anderson and van Wincoop (2003) model of international trade. The invisible asset is nurtured by repeated interactions between trading partners, and affects trade through lower trade costs. This motivates a dynamic gravity specification, and allows to address the following question: How quickly do bilateral trade flows adjust to trade reforms? The answer has important policy implications. If adjustment is fast, potential gains from trade reforms are achieved quickly, whereas in the opposite case it needs a long time to see the full beneficial impact. Our dynamic panel data regressions suggest that on average 20% of the distance to the steady-state trade flows are closed each year. Casting the gravity relation dynamically, there seems to be no role for 'phasing-in' of free trade agreements. Moreover, we detect interesting variation in the adjustment rates across country pairs which differ in geographical and cultural proximity, and across trade in commodity groups.

Finally, Chapter 9 summarizes the main findings. Moreover, it sketches an interesting array of future research.

Chapter 2

Sorting It Out: Technical Barriers To Trade and Industry Productivity[1]

2.1 Introduction

In the last fifty years, import duties like tariffs and non-tariff barriers (NTBs) like quantity and price controls on most relevant manufacturing goods have fallen substantially. In the meantime, however, another type of NTBs, has gained importance: *technical barriers to trade* (TBTs). TBTs require that exporters must customize their goods to meet the importing country's technical norms, its health, safety, or environmental norms, and that firms undergo costly product labeling and conformity assessment procedures. TBTs may serve a multitude of legitimate goals; however, TBTs "are intrinsically protectionist whenever they do not address market failures such as externalities and information asymmetries" (Beghin, 2008).

Maskus et al. (2000) call standards a "priority for debate" within the multilateral trading system. Accordingly, TBSs are an important issue for the European Single Market

[1]This Chapter is based on an article forthcoming in the *Open Economies Review* (doi 10.1007/s11079-009-9114-z). For the working paper version, see Felbermayr and Jung (2008a). The concept for the paper was developed jointly, theoretical analysis and writing were shared equally, and the calibration exercise was carried out by the author of this thesis.

Programme. The stringency of regulatory barriers to trade has increased in many EU countries from 1995-2005 (see Gwartney et al. 2007).

There is robust evidence that TBTs restrict trade. Anderson and van Wincoop (2004) show that non-tariff trade costs appear to be larger than tariffs. Kee et al. (2006) estimate that they amount to a 9% tariff equivalent on average for all goods and to a 40% tariff equivalent for goods that are actually affected. Case studies for the U.S. automotive and pharmaceutical sectors show that in some cases "total compliance costs would have exceeded the value of the market, suggesting that U.S. firms would have pulled out of the market" (Popper et al., 2004). This shows that TBTs may be particularly important for the extensive margin of trade.

There are essentially two ways to address TBT reform: (i) *harmonization*, and (ii) *mutual recognition* of standards. In the first case, standards are unified, so that domestic and imported varieties are subject to the same licensing procedures. This rules out discriminatory practices, but still exposes exporters to additional costs since a separate license is required for each market. In the second case, a firm that has obtained a license in one country can sell in any country, so that exporters face no further compliance costs. Clearly, absent mutual recognition, the duplication of costs faced by exporters amounts to an entry barrier into foreign markets which is entirely *wasteful*.

Within the context of the Single Market Programme, the EU champions *mutual recognition* of technical standards in areas where standards have not been harmonized. Ilzkovitz et al. (2007), however, argue that while about 20% of industrial production and about 26% of intra EU manufacturing trade are covered by mutual recognition, "practical implementation [...] is often hampered by legal uncertainty, administrative hassle and lack of awareness both from the side of the companies and of the Member States' authorities" (p. 61).

Despite the importance of TBTs, the theoretical literature has usually focused on variable trade costs such as transportation costs or tariffs; an exception being recent work by Baldwin and Forslid (2006). Following those approaches, we model TBT as a component of foreign market entry costs, which is related to *regulatory* costs. We deviate from existing papers by using a framework in which TBTs may play an efficiency-increasing role since they restrain excessive entry of monopolistic firms. However, they also affect the degree of product diversity and the equilibrium distribution of firm productivities.

We consider two deregulation scenarios that are inspired by current EU policies; in both entry costs of foreign firms into the domestic market fall. In the first, regulatory costs for domestic firms are reduced equiproportionally so that the relative competitive position of foreigners remains unchanged. For reasons that will become evident we call this case *T-neutral deregulation*. In the second scenario, the reform reduces entry costs of foreign firms only. We call this case *incremental mutual recognition*, since it makes additional licensing of goods for exporting gradually redundant.

We analyze these two scenarios in a model of international trade in differentiated goods with heterogeneous firms. Our setup is essentially the one of Melitz (2003). However, since TBTs are particularly relevant for trade in inputs (Popper et al., 2004) the utility function is reinterpreted as a production function which assembles various input varieties to a final output good. In contrast to standard treatments, we assume that the production function exhibits *variable degrees of external scale economies* (as Egger and Kreickemeier, 2009). This gives us a (second-best) rationale for the existence of TBTs. Moreover, recent literature (e.g., Corsetti et al., 2007) has established how important the size of the scale effect is as a major determinant of the qualitative and quantitative implications of trade liberalization. Finally, empirical work points towards substantial industry variance and generally rejects the implicit numerical choice of the scale effect parameter embodied in the traditional formulation of the Melitz models. As many other authors,[2] we work with a specific productivity distribution (Pareto) to sort out ambiguities and to parameterize the model for simulation purposes.

Our object of interest is productivity of final good producers (industry productivity). In the proposed framework, TBT reform affects the equilibrium *input diversity* (i.e., the mass of imported and domestic varieties) available in an industry, which affects industry productivity through an external effect. TBT reform also modifies the equilibrium productivity distribution of input producers and, hence, their *average productivity*, which also shapes industry productivity. These two forces determine the total effect, with their relative importance given by the external scale elasticity.

[2] Egger and Kreickemeier (2009), Baldwin and Forslid (2006), Helpman et al. (2004), etc.

Incremental mutual recognition changes the *extensive* margin of firm behavior; i.e., it modifies the selection of input producers into exporting and domestic sales. It also affects the *intensive* margin, as additional competitive pressure lowers sales per firm. The two effects lead to reallocation of resources towards medium-productivity new exporters, away from the upper and lower areas of the productivity distribution. The net reallocation effect that drives average productivity of input producers depends on the relative importance of these two countervailing reallocation effects. Also the effect on input variety is theoretically unclear. It depends on industry characteristics; e.g., on the degree of productivity dispersion. It is therefore not surprising that the total effect of TBT reform on industry productivity is a complicated function of model parameters. The contribution of this paper is to analytically sort out those ambiguities.

The theoretical analysis has a couple of interesting implications. First, it may rationalize the low robustness of a positive relationship between trade openness and real per capita income (see Rodríguez and Rodrik, 2000). Both variable and fixed cost trade liberalization lead to a higher volume of trade, thereby increasing openness. However, for similar parameter constellation, the former unambiguously improves productivity while the latter does not. Second, our paper suggests that the productivity effect of lower variable trade costs is importantly conditioned by the existence of fixed costs protection. Indeed, if TBTs are too high, lower transportation costs may turn out to lower industry productivity.

We offer an industry-by-industry calibration exercise in order to compare the *status quo* with a situation where the same technical requirements are applied to domestic producers and importers (harmonization).[3] It turns out that not all industries (e.g., machinery) gain from the reform. Many do gain, but only very modestly.

Our paper is related to recent work on productivity effects of intra-EU variable trade costs reduction under quasi-linear preferences with heterogeneous firms, e.g., Corcos et al., (2007). Our paper differs, since we use a modified Melitz (2003) model as a point of departure and relate TBT to fixed costs of foreign market access. Baldwin and Forslid (2006) provide an excellent discussion of trade policy in the standard Melitz model. They also

[3]Note that this exercise does not describe mutual recognition, since then importers would not face any regulatory costs at all.

address lower market access costs, and find the impact on input diversity to be ambiguous (p. 18). They do not study the effects on aggregate productivity.

The contribution of the present paper is the following. We generalize the standard approach by allowing for variable degrees of external scale economies which provides a potentially efficiency-enhancing role to TBTs. We sort out the full set of intricate implications of TBT reform on industry productivity *theoretically.* Finally, we check the relevance of our theoretical relevance by means of an industry-by-industry calibration exercise.

The remainder of the paper is organized as follows. Section 2 introduces the analytical framework and solves for general equilibrium. Section 3 theoretically derives conditions under which TBT reform increases productivity. Section 4 calibrates the model and simulates a harmonization scenario. Finally, Section 5 concludes.

2.2 Theoretical framework

2.2.1 Demand for inputs

We study a single market (such as the EU) with $n+1$ identical countries. Each country is populated by a representative consumer who has symmetric Cobb-Douglas preferences for final consumption goods produced by H industries. Final output producers in each industry h are perfectly competitive. They assemble their output using a continuum of inputs $q(\omega)$ according to the same constant elasticity of substitution (CES) production function

$$y_h = M_h^{\frac{\eta_h - 1}{\sigma_h - 1}} \left(\int_{\omega \in \Omega_h} q(\omega)^{\frac{\sigma_h - 1}{\sigma_h}} d\omega \right)^{\frac{\sigma_h}{\sigma_h - 1}}, \sigma_h > 1, \eta_h \geq 0. \quad (2.1)$$

The set Ω_h represents the mass of available inputs in industry h, and σ_h is the elasticity of substitution between any two varieties in that industry. M_h is the measure of Ω_h and denotes the degree of *input diversity* (the number of available differentiated inputs).

Higher input diversity may improve the fit of inputs in the production of the final good, which leads to external scale economies. η_h allows to flexibly govern the degree of external scale economies independently of σ_h. Expression (2.1) is analogous to the traditional CES production function for $\eta_h = 1$.[4] For $\eta_h = 0$, there are no external economies of scale. In the standard treatments of the Melitz (2003) or the Krugman (1980) models, the implicit choice of $\eta_h = 1$ links the effect of input diversity on output directly to the elasticity of substitution σ_h. Recent empirical work finds that $0 < \eta_h < 1$, rejecting the standard formulation (Ardelean, 2007).

The optimal demand quantity for each input ω is

$$q(\omega) = \left(\frac{p(\omega)}{P_h}\right)^{-\sigma_h} \frac{R_h/P_h}{M_h^{1-\eta_h}}, \tag{2.2}$$

where R_h is aggregate industry spending on inputs, $p(\omega)$ is the price charged by an input producer to the final output producers, and

$$P_h = M_h^{-\frac{\eta_h - 1}{\sigma_h - 1}} \left(\int_{\omega \in \Omega_h} p(\omega)^{1-\sigma_h} \, d\omega \right)^{\frac{1}{1-\sigma_h}} \tag{2.3}$$

is the price index dual to (2.1). Clearly, demand for variety ω is larger the smaller the price $p(\omega)$ relative to the average price of competing varieties P_h, and the larger real spending R_h/P_h. Higher input diversity M_h affects demand through two channels: indirectly, through its effect on the price level, and, if $\eta_h \neq 1$, directly, through the reduction of relevant real spending $(R_h/P_h)/M_h^{1-\eta_h}$ on each variety ω. Markups over marginal costs are constant in this framework; nevertheless we find it useful to call M_h a competition effect.

2.2.2 Production of inputs

Differentiated inputs are produced by a continuum of monopolistically competitive firms. Each industry draws on a single industry-specific factor L_h, which is inelastically supplied in equal quantities to all industries in all countries. Industry specificity of factors and the Cobb-Douglas utility function make sure that trade reforms generate only within rather than between-industry resource reallocation.

Input producers differ with respect to their productivity index φ; in the following we use

[4] The generalization is already discussed in the working paper version of the Dixit-Stiglitz (1977) paper and has been revived by Benassy (1996). Variants of it have been adopted by Blanchard and Giavazzi (2003), Egger and Kreickemeier (2009), Corsetti et al. (2007), or Felbermayr and Prat (2007).

this index instead of ω to identify firms.[5] They share the same domestic and foreign market entry costs, f_h^d and f_h^x, and the same iceberg variable trade costs $\tau_h \geq 1$. All fixed costs have to be incurred in terms of the industry-specific factor. $T_h \equiv f_h^x / f_h^d$ measures the *competitive disadvantage* of imported relative to domestically produced inputs. To ensure the existence of the selection effect (and in line with empirical evidence) we assume $\tau_h^{\sigma_h - 1} T_h > 1$.

Following Melitz (2003), we assume that firms are *ex ante* identical but face uncertainty regarding their productivity φ. They learn about φ only after sinking the entry cost f_h^e. Not all of those entrants turn out to be productive enough to bear the domestic fixed costs f_h^d. Hence, they remain inactive. Firms with intermediate productivity sell on the domestic market, but cannot recover the additional fixed costs associated to foreign sales, f_h^x. The most productive firms are active on all markets. Under the assumption $\tau_h^{\sigma_h - 1} T_h > 1$, there exist threshold productivity levels $0 < \varphi_h^* < (\varphi_h^x)^*$, which partition the distribution of input producers into inactive firms, purely domestic ones, and exporters.

We characterize the *ex ante* productivity distribution by the Pareto.[6] The c.d.f. is $G_h(\varphi) = 1 - \varphi^{-\gamma_h}$ with support on $[1, \infty)$, where the shape parameter $\gamma_h > \sigma_h - 1$ controls the dispersion of the distribution.[7] Larger values of γ_h characterize industries in which the productivity distribution is skewed towards inefficient input producers. We can then write the probability that a given entrant (that has just paid the entry fee f_h^e) starts production by $p_h^{in} = 1 - G(\varphi_h^*) = (\varphi_h^*)^{-\gamma_h}$. Analogously, $p_h^x = \frac{1 - G[(\varphi_h^x)^*]}{1 - G(\varphi_h^*)} = [\varphi_h^* / (\varphi_h^x)^*]^{\gamma_h}$ is the *ex-ante* (and *ex-post*) probability that one of these successful entrants will export.

Input producers have linear production functions $q(\varphi) = \varphi l_h(\varphi)$, where $l_h(\varphi)$ denotes the employment of the industry-h specific factor in firm φ. Profit maximization of input producers results in the standard rule for determining the ex-factory (f.o.b.) price, i.e. $p_h(\varphi) = w_h / (\rho_h \varphi)$, where $\rho_h = 1 - 1/\sigma_h$. Since the description of technology (2.1) is identical over all countries, we may pick the factor price specific to some industry, w_h, as the numeraire. In the following, we focus on that industry.

[5] This is possible because, in equilibrium, each input is produced by one firm only and the distribution of φ is assumed to have no mass points.

[6] This assumption is not necessary for many properties of the model; see Melitz (2003). However, it allows to understand the importance of industry dispersion to sort out the potentially ambiguous effects of various forms of trade liberalization on industry productivity.

[7] The assumption $\gamma_h > \sigma_h - 1$ makes sure that the equilibrium sales distribution converges.

Optimal demand (7.3) and the pricing rule of input producers imply that revenues earned on the domestic market are given by

$$r_h^d(\varphi) = R_h (P_h \rho_h \varphi)^{\sigma_h - 1} / M_h^{1-\eta_h}. \tag{2.4}$$

By symmetry, producers who find it optimal to sell to a foreign market generate revenues of $r_h(\varphi) = r_h^d(\varphi)\left(1 + n\tau_h^{1-\sigma}\right)$. In turn, profits from selling domestically and exporting to one foreign market are respectively given by

$$\pi_h^d(\varphi) = r_h^d(\varphi)/\sigma_h - f_h^d, \tag{2.5}$$

$$\pi_h^x(\varphi) = \tau_h^{1-\sigma} r_h^d(\varphi)/\sigma_h - f_h^d T_h. \tag{2.6}$$

2.2.3 Industry aggregation

The productivity of final output producers (*industry productivity)* depends on *input diversity* (the number of available inputs), and on the *average productivity* level of input producers. Input diversity has a domestic and an imported component: $M_h = M_h^d + nM_h^x$, where n is the number of identical import (and, by symmetry: export) markets. Since $M_h^x = p_h^x M_h^d$, one can express M_h as $M_h = M_h^d(1 + np_h^x)$.

The average productivity level of domestic input producers, $\tilde{\varphi}_h^d$, is defined as the mean over sales-weighted productivities of all active producers.[8] Using the Pareto assumption,

$$\left(\tilde{\varphi}_h^d\right)^{\sigma_h - 1} = \frac{\int_{\varphi_h^*}^{\infty} \varphi^{\sigma_h - 1} dG_h(\varphi)}{1 - G\left(\varphi_h^*\right)} = \frac{\gamma_h}{\gamma_h - (\sigma_h - 1)} \left(\varphi_h^*\right)^{\sigma_h - 1}. \tag{2.7}$$

Equation (2.7) shows that the endogenously determined entry cutoff productivity level φ_h^* shapes the average productivity of domestically produced inputs. The average over exporters, $\tilde{\varphi}_h^x$, is constructed analogously, and crucially depends on the export cutoff productivity level $(\varphi_h^x)^*$.

Given perfect *symmetry* across countries, the average productivity of inputs used in

[8] See Melitz (2003), p. 1700.

production of the final good, $\tilde{\varphi}_h$, is

$$(\tilde{\varphi}_h)^{\sigma_h - 1} = \frac{1}{1 + np_h^x} \left(\tilde{\varphi}_h^d\right)^{\sigma_h - 1} + \frac{np_h^x}{1 + np_h^x} \left(\frac{\tilde{\varphi}_h^x}{\tau_h}\right)^{\sigma_h - 1}, \tag{2.8}$$

where productivities of foreign firms are adjusted for iceberg transportation costs τ_h, and average productivities of domestic and imported varieties are weighted by their respective shares in total input diversity.

The weighting in (2.8) implies that $q\left(\tilde{\varphi}_h\right) = R_h M_h^{-\frac{\eta_h + \sigma_h - 1}{\sigma_h - 1}} / P_h$. With $\eta_h = 0$ (i.e., in the absence of industry externalities), the output of the average firm is equal to average output $R_h / \left(P_h M_h\right)$. Similarly, applying (2.8) to the industry price index (2.3), one has $P_h = M_h^{-\frac{\eta_h}{\sigma_h - 1}} p\left(\tilde{\varphi}_h\right)$. Hence, if $\eta_h = 0$, the price index is equal to the price chosen by the average firm.

Using optimal pricing of inputs in P_h and recognizing that aggregate productivity A_h is given by $1/P_h$, we are now ready to write the level of aggregate *industry productivity* as a function of average productivity and input diversity

$$A_h = \rho \tilde{\varphi}_h M_h^{\frac{\eta_h}{\sigma_h - 1}}. \tag{2.9}$$

Industry productivity[9] increases as ρ goes up so that markups and the amount of resources used for fixed costs are lower. Industry productivity is directly proportional to average productivity of input producers $\tilde{\varphi}_h$. It depends positively on input diversity M_h as long as η_h is strictly positive. The term $\frac{\eta_h}{\sigma_h - 1}$ is the elasticity of industry productivity with respect to input diversity.[10] The aim of the subsequent analysis is to understand how A_h changes with different types of TBT reform. To do this, we need to endogenize $\tilde{\varphi}$ and M_h. Typically, TBT liberalization moves these two components of industry productivity in opposite directions. Hence, the elasticity $\frac{\eta_h}{\sigma_h - 1}$ will play a crucial role.

[9] This is the *ideal* measure of industry productivity. Gibson (2008) points out that productivity effects induced by Melitz (2003)-type reallocation of market shares within industries are not reflected by data-based measures of productivities, (e.g., value added per worker).

[10] If $\eta_h = 1$, (2.9) is formally equivalent to the expression describing total welfare in Melitz (2003).

2.2.4 General equilibrium

In this section, we solve for the equilibrium values of M_h and $\tilde{\varphi}_h$. The discussion is deliberately brief, since it is close to Melitz (2003) and to Baldwin and Forslid (2006); the only difference comes through $\eta_h \neq 1$. Equilibrium is determined by four conditions.

Zero cutoff profit (ZCP) conditions. The domestic ZCP condition identifies the firm φ_h^* that is indifferent between selling domestically and remaining inactive; the foreign ZCP condition locates the firm $(\varphi_h^x)^*$ that is indifferent between selling domestically and also selling on the n symmetric foreign markets. Formally, the ZCPs are

$$\pi_h^d(\varphi_h^*) = 0, \qquad \pi_h^x[(\varphi_h^x)^*] = 0. \tag{2.10}$$

Using the profit functions derived in (2.5) and (2.6), the zero cutoff profit conditions imply that $r_h^d(\varphi_h^*) = \sigma_h f_h^d$ and $r_h^d[(\varphi_h^x)^*] = \sigma_h f_h^d \tau_h^{\sigma_h - 1} T_h$. Then (2.4) links the export cutoff $(\varphi_h^x)^*$ and the domestic entry entry cutoff φ_h^*[11]

$$(\varphi_h^x)^* = \varphi_h^* \tau_h T_h^{\frac{1}{\sigma_h - 1}}. \tag{2.11}$$

Moreover, using the definition of $\tilde{\varphi}_h^d$ (2.7) and the condition linking the two cutoff productivities (2.11), one can link the average productivity of domestic firms with those of exporters

$$\tilde{\varphi}_h^x = \tau_h T_h^{\frac{1}{\sigma_h - 1}} \tilde{\varphi}_h^d. \tag{2.12}$$

It follows from (2.11) that the probability of exporting conditional on successful entry is given by

$$p_h^x = \tau_h^{-\gamma_h} T_h^{-\frac{\gamma_h}{\sigma_h - 1}}. \tag{2.13}$$

Our results so far allow to express average productivity $\tilde{\varphi}_h$, as defined in (2.8), by using (2.12)

$$\tilde{\varphi}_h = \tilde{\varphi}_h^d \left(\frac{1 + n p_h^x T_h}{1 + n p_h^x} \right)^{\frac{1}{\sigma_h - 1}}. \tag{2.14}$$

[11] Derivations of analytical results are detailed in the working paper version of this article; see Felbermayr and Jung (2008a).

Both the domestic and the foreign market ZCPs can be combined and graphed in $(\varphi_h^*, \bar{\pi}_h)$ −space by using the definition of average profits (defined over active firms, *ex post* perspective) $\bar{\pi}_h = \pi_h^d\left(\tilde{\varphi}_h^d\right) + np_x^x \pi_h^x\left(\tilde{\varphi}_h^x\right)$ and noting that $\tilde{\varphi}_h^d$ and $\tilde{\varphi}_h^x$ are both functions of φ_h^*. It is well known that, given the Pareto assumption, average profits $\bar{\pi}_h$ do not depend on φ_h^*.[12]

Free entry. The free entry condition ensures that *expected* profits (from the ex ante perspective) cover entry costs f_h^e :

$$\frac{p_h^{in}\bar{\pi}_h}{\delta_h} = f_h^e, \tag{2.15}$$

where δ_h is the exogenous Poisson exit rate of producers and $p_h^{in} = (\varphi_h^*)^{-\gamma_h}$ is the likelihood that a random productivity draw allows a producer to at least break even on the domestic market. Clearly, this free entry condition defines an upward-sloping relationship between $\bar{\pi}_h$ and φ_h^*. Equating that condition with the combined ZCP condition discussed above, one can determine the entry cutoff productivity level φ_h^* as a function of exogenous variables only[13]

$$\varphi_h^* = \left[\frac{\sigma_h - 1}{\gamma_h - (\sigma_h - 1)} \frac{f_h^d}{\delta_h f_h^e}\left(1 + np_h^x T_h\right)\right]^{\frac{1}{\gamma_h}}. \tag{2.16}$$

Substituting φ_h^* into the definition of *domestic* average productivity (2.7) one can determine $\tilde{\varphi}_h^d$. Finally, using (2.13) and (2.14) allows to compute the average productivity defined over all input producers, $\tilde{\varphi}_h$. Note that the above analysis has not used any factor market clearing condition; $\tilde{\varphi}_h$ is therefore independent from L_h. Moreover, when solving for $\tilde{\varphi}_h$, input diversity is irrelevant. Input diversity M_h can be found recursively, i.e., *given* $\tilde{\varphi}_h$.

Stationarity condition. The fourth equilibrium condition allows to pin down input diversity. In a stationary equilibrium, in any country, the mass of successful entrants $p_h^{in}M_h^e$ must equal the mass of producers hit by the exit shock $\delta_h M_h^d$. Hence,

$$p_h^{in}M_h^e = \delta_h M_h^d. \tag{2.17}$$

[12] See, e.g., Baldwin and Forslid (2006).

[13] Noting that (2.13) relates p_h^x to exogenous variables.

As shown in Melitz (2003, p. 1704), under stationarity, aggregate revenue R_h is fixed by the size of L_h (and the normalization of the $h-$factor price). This determines equilibrium input diversity by $M_h = R_h/r^d(\tilde{\varphi}_h)$. Using the zero cutoff profit conditions (2.10), we obtain equilibrium industry diversity

$$M_h = \frac{L_h}{\sigma_h f_h^d}\left(\frac{\varphi_h^*}{\tilde{\varphi}_h}\right)^{\sigma_h - 1}. \tag{2.18}$$

2.3 Industry productivity effects of TBT reform

Totally differentiating industry productivity (2.9) yields

$$\hat{A}_h = \frac{\eta_h}{\sigma_h - 1}\hat{M}_h + \widehat{\tilde{\varphi}}_h. \tag{2.19}$$

We use the conventional 'hat' notation to denote an infinitesimally small deviation of a variable from its initial level ($\hat{x} = dx/x$). Any type of trade liberalization has potential implications for the cutoff productivity levels φ_h^* and $(\varphi_h^x)^*$, and hence for productivity averages of domestic and international firms, $\tilde{\varphi}_h^d$ and $\tilde{\varphi}_h^x$, respectively. The productivity level of the average firm $\tilde{\varphi}_h$ is a weighted average over domestic and international firms, with the relative weights potentially being affected by TBT reform, too. Different trade liberalization scenarios may have similar effects on cutoff productivities (the *extensive* margin); yet, they may lead to drastically different patterns of inter-industry resource reallocation along the *intensive* margin, and, hence, different results for industry productivity. The literature has not fully recognized this point yet.

Input diversity adjusts to changes in the entry cutoff $\tilde{\varphi}_h$ and average productivity φ_h^* such that factor markets clear (see equation (2.18)). For the assessment of industry productivity, both effects need to be combined, with the elasticity $\eta_h/(\sigma_h - 1)$ playing a crucial role. Hence, the overall effect of TBT reform on industry productivity works through a number of different mechanisms and is likely to be ambiguous theoretically. In the extreme case where $\eta_h = 0$ (Blanchard and Giavazzi, 2003), (2.19) simplifies substantially as variation in input diversity has no bearing on industry productivity. Also the case where $\eta_h = 1$, typically studied in the literature, turns out offer more clear-cut results. In this special case

(2.19) is formally isomorphic to the description of welfare in the Melitz (2003) model. In the subsequent analysis, we discuss the empirically relevant situation where $\eta_h \in (0,1)$ and focus on TBT reform rather than on the more widely studied case of variable trade cost liberalization.

We assume that fixed market costs f_h^d and f_h^x have two components: fixed *distribution costs*, $\bar{f}_h^d$ and $\bar{f}_h^x$, and fixed regulatory costs, $\tilde{f}_h^d$ and $\tilde{f}_h^x$, that relate to approval and conformity assessment costs. The latter is set by national authorities, but differs from a tax since it does not generate revenue. We define as a TBT reform any policy measure that reduces regulatory costs for foreign firms $\tilde{f}_h^x$.[14] Full-fledged *mutual recognition* of standards, in contrast, would make licensing procedures for imported varieties redundant, hence $\tilde{f}_h^x = 0$. Only in this case do TBTs disappear entirely.

We consider two scenarios of TBT reform. In the first, policy makers lower the burden on foreign firms $\tilde{f}_h^x$, but also adjust regulatory costs for domestic firms $\tilde{f}_h^d$ such that the competitive disadvantage of foreign firms, $T_h \equiv f_h^x/f_h^d$, remains unchanged. We term this case *T-neutral deregulation*. In the second scenario, $\tilde{f}_h^x$ is reduced, while $\tilde{f}_h^d$ remains fixed. Any marginal reduction in $\tilde{f}_h^x$ brings the economy closer to the ideal situation of full mutual recognition. Hence, we call our second scenario *incremental mutual recognition*. Throughout, we assume that distribution-related fixed costs are such that the partitioning of firms into exporters and purely domestic firms is maintained (i.e. $\bar{f}_h^x/f_h^d > \tau_h^{1-\sigma_h}$).

2.3.1 T-neutral deregulation

In this scenario, f_h^x and f_h^d both fall, but T_h remains constant. Therefore, the export probability p_h^x (2.13), which depends on fixed market access costs only through T_h, is fixed. It is also clear, that the domestic φ_h^* and the export cutoff productivity levels $(\varphi_h^x)^*$ move proportionally (see (2.11)). To understand the effect of T-neutral deregulation, note that φ_h^* is determined in $(\varphi_h^*, \bar{\pi}_h)$ −space by the intersection of the ZCP condition and the free entry condition. In the present context, the first is a horizontal line, while the latter is upward-sloping. Domestic deregulation does not affect the free entry locus. However, the

[14] Note that harmonization of standards need not be a TBT reform if regulatory costs increase for foreign firms.

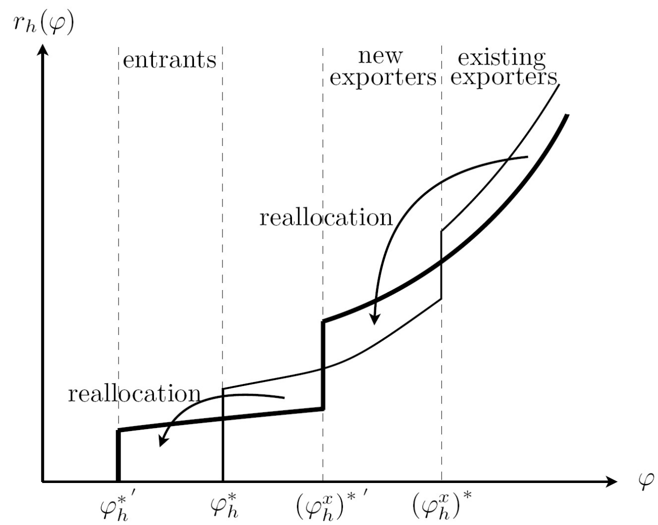


Figure 2.1: Within-industry reallocation of market shares as response to T-neutral deregulation.

ZPC condition shifts downwards, so that φ_h^* falls. The reasoning is as follows. The ZCP locus summarizes combinations of $\bar{\pi}_h$ and φ_h^* for which the marginal firm φ_h^* just breaks even. When fixed costs f_h^d fall, the firm starts to make profits. To restore zero profits, the firm's revenue has to fall. This is achieved by tighter competition: either relative prices have to increase or residual demand has to drop. This is however not limited to firm φ_h^*; profits fall for *all* firms; hence $\bar{\pi}_h$ goes down.

The effect on the cutoff productivities at hand, one can now use Figure 2.3.1 to gain some intuition on the reallocation of market shares that domestic deregulation entails. The figure shows sales $r_h(\varphi)$ per firm as a function of productivity. This locus is upward-sloping as more efficient firms have higher sales (given $\sigma > 1$). Since total sales R_h are pinned down by L_h, r_h can be read as a measure of market share. The sales function changes with T-neutral deregulation.[15] Due to the increase in the number of traded varieties, competition goes up, which means that incumbent exporters and domestic-only firms lose market share (intensive margin).

Since domestic and foreign market entry costs fall in proportion, the probability of exporting, given successful entry, does not change (see equation (2.13)). Moreover, the entry cutoff levels shift proportionally. Hence, the reallocation of market shares towards

[15]Figure 2 in Melitz (2003) which studies the reallocation of market shares as an economy moves from autarky to trade. Our Figure 2.3.1 is similar, but studies incremental trade liberalization.

less productive firms directly translates into a decrease in average productivity. We shall discuss the effect on average productivity, input diversity, and industry productivity in more detail below.

Average productivity of input producers. The change in average productivity is completely driven by the change in the entry cutoff productivity level, i.e.$\widehat{\tilde{\varphi}}_h/\hat{f}_h^d = \hat{\varphi}_h^*/\hat{f}_h^d$ (see (2.14) and (2.7)). Totally differentiating (2.16) yields[16]

$$\frac{\widehat{\tilde{\varphi}}_h}{\hat{f}_h^d} = \frac{1}{\gamma_h} > 0. \tag{2.20}$$

Thus, average productivity declines in response to T-neutral deregulation. The parameter γ_h is inversely related to the degree of productivity dispersion (heterogeneity) in the industry. In the extreme case where $\gamma_h \to \infty$, all firms are identical and there cannot be any selection or reallocation effect (as long as all firms remain exporters or purely domestic). The room of reallocation is bigger as γ_h is smaller and industry heterogeneity is larger. It is therefore natural that the effect of T-neutral deregulation on average productivity is larger the smaller γ_h.

Input diversity. As argued above, lowering fixed market entry costs attracts new input producers to start production and makes it profitable for additional firms to export. The change in input diversity is given by $\hat{M}_h = -\hat{f}_h^d + (\sigma_h - 1)\left(\hat{\varphi}_h^* - \widehat{\tilde{\varphi}}_h\right) < 0$. In the present scenario, the entry cutoff productivity level $\hat{\varphi}_h^*$ and average productivity $\widehat{\tilde{\varphi}}_h$ move proportionally. Hence, the elasticity of input diversity with respect to f_h^d is

$$\frac{\hat{M}_h}{\hat{f}_h^d} = -1. \tag{2.21}$$

Industry productivity. The industry productivity effect combines the input diversity effect and the effect on input producers' productivity. This leads to the following proposition.

---

[16]Recall that changes in the regulatory component directly translate into changes in total market access costs, i.e. $\hat{f}_h^d = \widehat{\tilde{f}}_h^d$.

Proposition 1. ***(T-neutral deregulation)*** *Industry productivity only increases in response to T-neutral deregulation, if the degree of external economies of scale is larger than the inverse dispersion measure of the Pareto*

$$\frac{\eta_h}{\sigma_h - 1} > \frac{1}{\gamma_h}. \tag{2.22}$$

Proof. Follows from using (2.20) and (2.21) in (2.19). □

Hence, the elasticity of aggregate productivity with respect to input diversity has to be sufficiently large in order to overcompensate the loss in average productivity. Note that, in the case of $\eta_h \geq 1$, the above inequality always holds (by the regularity condition $\gamma_h > \sigma_h - 1$). Hence, domestic deregulation always makes the final goods producer more productive. However, this result is not general: in the empirically relevant case, where $\eta_h < 1$, the industry productivity effect is ambiguous.

2.3.2 Incremental mutual recognition

This scenario implies a reduction of T_h with f_h^d held constant. Consider again the determination of the domestic cutoff productivity φ_h^* in $(\varphi_h^*, \bar{\pi}_h)$ −space. The free entry condition does not change as T_h falls. However, the ZPC condition now shifts up, so that φ_h^* goes up. The marginal domestic producer is not an exporter; hence there is no direct effect of the reduction in f_h^x. However, the entry of foreign importers makes competition tougher, revenue per firm goes down, and the φ_h^* firm starts to make losses. To restore zero profits, there must be an upward adjustment of $\bar{\pi}_h$. The number of competitors or their average productivity (or both) have to go down.

Hence, φ_h^* increases while $(\varphi_h^x)^*$ goes down. Figure 2.3.2 provides some intuition on the reallocation of market shares: the emergence of new exporters causes a loss of market share to incumbent exporters and domestic firms. Since new exporters are firms with medium levels of productivity, the net effect on average productivity is ambiguous.

Average productivity of input producers. $\tilde{\varphi}_h$ only increases in response to a cut in T_h if the shape parameter γ_h is large enough. The intuition is straightforward: The larger

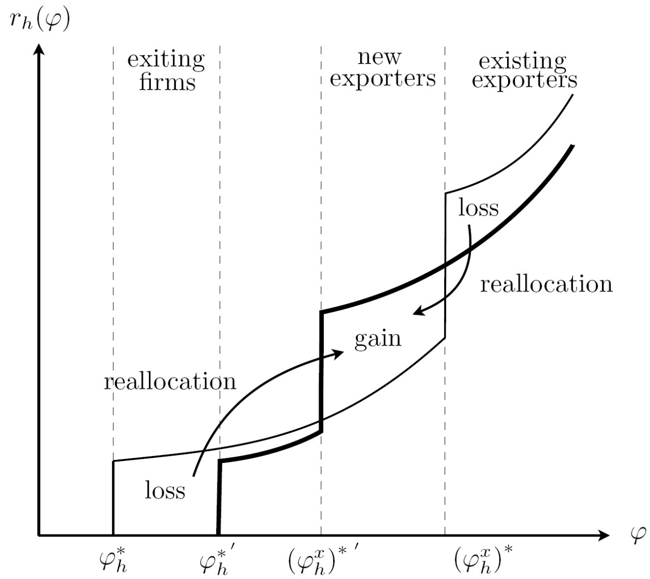


Figure 2.2: Within-industry reallocation of market shares as response to incremental mutual recognition.

the shape parameter γ_h, the more mass is given to low productive firms, thus giving a high potential for reallocation from fairly unproductive, exiting firms to new exporters.

If the initial level of competitive disadvantage of importers is already smaller than 1, there is almost no export selection effect, and $\tilde{\varphi}_h$ never increases in response to a TBT reform regardless of the shape parameter γ_h. We may summarize the result in Lemma 1.

Lemma 1. ***(Average productivity)*** *Fix f_h^d and reduce T_h. Average productivity $\tilde{\varphi}_h$ increases in response to incremental mutual recognition if and only if the the dispersion measure of the Pareto distribution is large enough, i.e.*

$$\frac{\widehat{\tilde{\varphi}}_h}{\hat{T}_h} < 0 \Leftrightarrow \frac{1}{\gamma_h} < \frac{1}{\underline{\gamma}_h} \equiv \frac{1}{\sigma_h - 1}\sqrt{\frac{1}{1 + np_h^x}\frac{T_h - 1}{T_h}} \tag{2.23}$$

Proof. Follows immediately from totally differentiating (2.14). □

At the extensive margin, the least productive firms are forced to exit (selection effect), while new exporters enter (adverse export selection effect). Only if the selection effect is large enough as compared to the adverse selection effect, $\tilde{\varphi}_h$ rises as stated in condition (2.23).

Input diversity. If the productivity distribution is not extremely skewed towards the least productive firms (i.e. if the shape parameter γ_h is sufficiently small), the number of input varieties lost through exposure to trade is overcompensated by additionally imported inputs, resulting in an increase in input diversity.

Lemma 2. ***(Input diversity)*** *Fix f_h^d and reduce T_h. Input diversity M_h increases in response to incremental mutual recognition if and only if the dispersion measure of the Pareto γ_h is sufficiently small, i.e.*

$$\frac{\hat{M}_h}{\hat{T}_h} < 0 \Leftrightarrow \frac{1}{\gamma_h} > \frac{1}{\bar{\gamma}_h} \equiv \frac{1}{\sigma_h - 1} \frac{1}{1 + np_h^x} \frac{T_h - 1}{T_h}. \tag{2.24}$$

Proof. Follows immediately from totally differentiating (2.18). □

Lemma 2 presents a necessary condition (2.24). Note that a simple sufficient condition is $T_h < 1$.

Industry productivity. Using (2.19) and Lemmata 1 and 2, average productivity and input diversity increase in response to a incremental mutual recognition, if the value γ_h is not too extreme, i.e., if

$$\bar{\gamma}_h > \gamma_h > \underline{\gamma}_h. \tag{2.25}$$

Then, industry productivity improves unambiguously regardless the degree of external economies of scale.

However, even if condition (2.25) is violated, industry productivity can actually increase, depending on the degree of external economies of scale. If $\tilde{\varphi}_h$ is falling and M_h rising, the degree of external economies of scale has to be sufficiently large for industry productivity to increase, and *vice versa*. There exists the following trade-off: If the shape parameter γ_h is sufficiently small unproductive firms have little relative mass. Hence, there is little potential for reallocation from the exiting, low-productivity firms to new exporters. Then average productivity declines. In contrast, input diversity increases, since more imported varieties are attracted than domestic ones are forced to exit. If, on the other hand, the shape parameter is γ_h is sufficiently large, the logic reverses, and average productivity increases whereas input diversity declines.

Consider that input diversity decreases in response to incremental mutual recognition, which means a violation of condition (2.24) in Lemma 2. Then, by condition (2.23) average productivity unambiguously rises, and the degree of external economies of scale has to be sufficiently small. The negative diversity effect is always offset for the empirically relevant cases $\eta_h \leq 1$.

Turn now to the case where average productivity declines in response to incremental mutual recognition, i.e. a violation of condition (2.23) in Lemma 1. Then, by (2.24) industry diversity always increases, and $\eta_h / (\sigma_h - 1)$ has to be sufficiently large to generate an increase in industry productivity, which is always true for the special Melitz case ($\eta_h = 1$). These results are summarized in the following Proposition.

Proposition 2. ***(Incremental mutual recognition)*** *Let υ_h^* be the threshold degree of external economies of scale*

$$\upsilon_h^* \equiv \frac{1}{\gamma_h - \bar{\gamma}_h} \left(\frac{\gamma_h}{\sigma_h - 1} - \frac{\bar{\gamma}_h}{\gamma_h} \right).$$

(i) Violation of condition (2.24). A decrease in input diversity in response to incremental mutual recognition is overcompensated by an increase in average productivity, if and only if the degree of external economies of scale is below the threshold value υ_h^*, i.e.

$$\frac{\eta_h}{\sigma_h - 1} < \upsilon_h^*, \tag{2.26}$$

where $\upsilon_h^* > 1/(\sigma_h - 1)$.

(ii) Violation of condition (2.23). A decrease in average productivity in response to incremental mutual recognition is overcompensated by an increase in input diversity, if and only if the degree of external economies of scale is above the threshold value υ_h^*, i.e.

$$\frac{\eta_h}{\sigma_h - 1} > \upsilon_h^*, \tag{2.27}$$

where $0 < \upsilon_h^* < 1/(\sigma_h - 1)$.

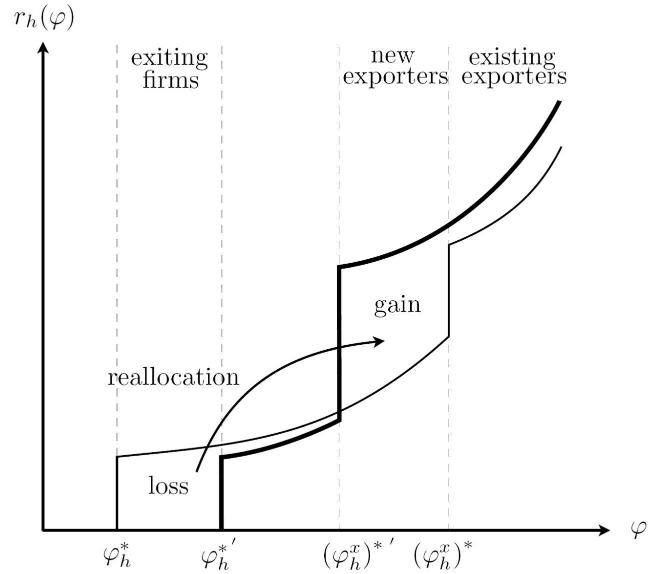


Figure 2.3: Within-industry reallocation of market shares as response to variable trade cost liberalization.

Proof. The conditions follow from equation (2.19). □

2.3.3 Comparing TBT reform to lower variable trade costs

As with incremental mutual recognition, lower variable trade costs induce an upward-shift in the ZCP. The reason for this effect is the same as before. Hence, tariff liberalization (or any reduction in variable trade costs) has similar effects on the cutoff productivity levels as lower T_h with f_h^d fixed. However, lower trade costs on net benefit incumbent exporters, as additional competitive pressure is over-compensated by lower trade costs.[17] It follows, that the direction of market share reallocation is unequivocally towards more productive firms. Note, however, that the sales function depicted in Figure 2.3.3 does not suffice to determine the effect on average productivity, which depends on the masses of firms engaged in exporting relative to purely domestic ones. It turns out that the productivity effect is *a priori* ambiguous and depends on T_h, which governs the size of the selection and export selection effect.

If $T_h > 1$, imported inputs are on average more productive (they have to cover higher

[17]This result holds for all productivity distributions.

fixed market entry costs). This implies lower prices and, in turn, given CES preferences, results in higher expenditure. Thus, more than one domestically produced input has to be displaced in order to import one additional input variety, and input diversity drops.[18] Reallocation of market shares towards more productive firms and the reduced availability of the least productive inputs, result in higher average productivity. If $T_h < 1$, we end up with higher input diversity. It turns out that in this case average productivity actually declines.[19]

For the empirical relevant case $T_h > 1$, input diversity drops at the lower end of the productivity distribution, resulting in an increase in average productivity. As mentioned above, the condition under which average productivity increases is less strict:

Lemma 3. ***(Average productivity)*** *Average productivity increases in response to variable trade cost liberalization if and only if the dispersion measure of the Pareto distribution is small enough, i.e.*

$$\frac{\widehat{\tilde{\varphi}}_h}{\hat{\tau}_h} < 0 \Leftrightarrow \frac{1}{\gamma_h} > -\frac{1}{\sigma_h - 1}\frac{1}{1 + np_h^x}\frac{T_h - 1}{T_h}. \tag{2.28}$$

Proof. Follows from totally differentiating (2.14). □

Condition (2.28) clearly holds if M_h decreases, i.e. $T_h > 1$. However, average productivity also rises if the selection effect is sufficiently large, shifting input production to more productive firms.

Proposition 3. ***(Lower variable trade costs)*** *Let ψ_h^* be the threshold degree of external economies of scale*

$$\psi_h^* \equiv \frac{1}{\sigma_h - 1} + \frac{1}{\gamma_h}\frac{T_h}{T_h - 1}\left(1 + np_h^x\right).$$

(i) Assume $T_h > 1$, so that input diversity decreases and average productivity increases. Then industry productivity goes up if and only if the degree of external economies of scale

[18] A similar explanation has been put forward by Baldwin and Forslid (2006).

[19] $T_h > (<)\ 1$ is a necessary condition for input diversity to decrease (rise), whereas for average productivity to increase (drop) it is a sufficient condition. The necessary condition would be less strict and depend on the skewness of the productivity distribution.

is below the threshold value ψ_h^*, i.e.

$$\frac{\eta_h}{\sigma_h - 1} < \psi_h^*. \tag{2.29}$$

(ii) Assume $T_h < 1$ and a violation of condition (2.28), so that input diversity increases and average productivity decreases. Then industry productivity increases if and only if the degree of external economies of scale is above the threshold value ψ_h^*, i.e.

$$\frac{\eta_h}{\sigma_h - 1} > \psi_h^*. \tag{2.30}$$

Proof. The conditions follow from totally differentiating (2.18). □

Conditions (2.29) and (2.30) always hold if respectively $\eta_h \leq 1$, and $\eta_h \geq 1$. Hence, in the special Melitz case ($\eta_h = 1$), industry productivity always increases in response to variable trade cost liberalization. In contrast, incremental mutual recognition reduces the market shares of existing exporters, thereby inducing reallocation of market shares towards less productive firms, and $T_h > 1$ is not sufficient to guarantee an increase in average productivity.

There are three interesting corollaries that follow from the comparison between TBT reform and variable trade cost reductions. First, in the empirically relevant case $T_h > 1$ and $\eta_h \leq 1$, lower variable trade costs unambiguously improve industry productivity, while the effect of TBT reform is still ambiguous. However, in both situations, total export sales increase.[20] Hence, there is no clear link between increased trade openness and industry (or even economy-wide) productivity measures. This theoretical result may rationalize the low degree of robustness that empirical cross-country analysis of the openness-productivity (or more often: GDP per capita) link suffers from; see, e.g., Rodríguez and Rodrik (2000).

Second, the effect of lower trade costs is conditioned by the importance of competitive disadvantage of foreign firms as measured by T_h. We have seen above, that – if $T_h > 1$ – lower variable trade costs may lead to a fall in industry productivity. In other words:

[20] Total sales abroad are given by $X_h^{cif} = nM_h^x r^x(\tilde{\varphi}_h^x)$. Recall that $M_h^x = p_h^x M_h^d$. Using (2.4), (2.17) and (2.18) one finds that $X_h^{cif} = L_h n p_h^x T_h / (1 + n p_h^x T_h)$, and $\partial X_h^{cif} / \partial T_h < 0$.

industries can be hurt by reductions in tariffs or transportation costs if the degree of fixed-cost protection is too high. This allows the following policy conclusion. Before engaging in variable trade cost reforms, countries should lower TBTs. Only countries with sufficiently low TBTs benefit from the (exogenous) downward trend in transportation costs. Hence, productivity gains from technical progress in transportation can be tapped only if TBTS are low enough.

Third, there seems to be substantial resistance against TBT reforms. Gwartney et al. (2007) argue that the EU25 countries have failed on average to decrease regulatory costs to importers. Our paper allows two interpretations of this result. First, based on efficiency considerations, TBT reform is not desirable per se, at least not under arbitrary parameter constellations. Second, TBT reform – even if it leads to industry productivity gains – inflicts losses to the vast majority of firms due to the implied reallocation of resources towards new exporters – by nature a relatively small fraction out of all domestic firms. Hence, it may not be overly surprising that total resistance against TBT reform is strong, and, in particular, stronger than against lower variable trade cost reductions, which tend to be beneficial for incumbent exporters.

2.4 Numerical exercise at the industry level

In this section, we use estimates of the key parameters from the literature to answer the question how TBT reform affects average productivity of input producers, input diversity, and industry productivity.[21] Since there is substantial cross-industry variation of parameters, we do a separate analysis for 14 industries.

Rather than evaluating the sign of a marginal TBT reform, we quantify the productivity gains and losses associated to the *harmonization* of technical standards. First, we calibrate the *status quo* such that the model is in line with key data for 14 industries (amongst other things, we set T_h to replicate observed export participation rates given τ_h which we take as data). Then, we reduce the policy-related component of fixed foreign market entry costs $\tilde{f}_x$ such that foreign firms face the same licensing costs than domestic firms. Hence, there

[21] For detailed information on the calibration of the degree of external economies of scale, $\eta_h/(\sigma_h - 1)$, and the level of competitive disadvantage of importers, T_h, see the Appendix.

is no more discrimination against foreign firms. Moreover, one can interpret this situation as one where all technology-related entry costs across markets are identical.[22] Then the comparative disadvantage of importers T_h merely reflects regulatory differences. We compare the status quo to a situation where $T_h = 1$, i.a., where all *discriminatory* regulation (beyond sunk entry costs) is scrapped. Note, however, that our scenario stops short from full harmonization which would imply that regulatory costs for exporters are zero if they have already incurred those costs in their domestic market. Since we have no way to decompose technological and policy-related components of fixed costs, our scenario seems the most plausible given data limitations. Table 2.3 reports the calibration results, distinguishing between parameters taken from the data and calibrated numbers. Our scenario is technologically feasible for all industries since the sorting condition $\tau_h^{1-\sigma_h} T_h > 1$ continues to hold, given any level of variable trade costs τ_h.[23]

One could expect that industry productivity unambiguously rises. However, *a priori* this is not the case, since their are industry externalities at work. The decentralized equilibrium does not necessarily feature the efficient degree of input diversity if $\eta_h \neq 1$.[24] Neither do producers internalize the effect of entry on the external economies of scale in the industry nor on the profits of incumbent producers. If $\eta_h < 1$, there is over-supply of varieties, if $\eta_h > 1$ (which is empirically implausible) there is under-supply. Only in the special case where $\eta_h = 1$ does the planner solution coincide with the decentralized equilibrium. High regulatory costs reduce entry and thereby mitigate the distortion due to external economies of scale. However, TBTs are certainly not the first-best policy to cope with oversupply of varieties, since they do not generate any income (unlike entry taxes).

Table 2.1 shows the results of our *global* analysis.[25] Clearly, the cut to $T_h = 1$ is largest for industries with high *status quo* degrees of competitive disadvantage of importers (knowledge-intensive industries such as scientific, measuring, optical, and photographic

[22] In the present model, this is a natural assumption since countries are symmetric.

[23] It *may* be feasible to reduce beyond harmonization and eliminate $\tilde{f}_h^x$ such that T_h goes below unity. However, since we have no data on the components of f_h^d and f_h^x, we cannot calibrate the lowest feasible level of $\tilde{f}_h^d$. Also note that our numerical analysis does not require calibration of f_h^e or δ_h since those parameters drop out when comparing equilibrium outcomes at $T_h = 1$ to those obtained under the benchmark calibration.

[24] The welfare-theoretic results obtained by Benassy (1996) for arbitrary η_h and homogeneous firms continue to hold in the presence of productivity heterogeneity.

[25] Results may differ from a *local* analysis due to non-linearities.

Table 2.1: Productivity gains and losses from harmonization

Industry	(1) $\Delta\%T_h$	(2) $\Delta\%p_h^x$	(3) $\Delta\%\tilde{\varphi}_h$	(4) $\Delta\%M_h^d$	(5) $\Delta\%M_h$	(6) $\Delta\%A_h$
Chemicals	-24.0	54.4	-10.6	-12.0	23.2	0.0
Rubber and plastics	-24.4	73.0	-7.6	-18.7	22.2	3.8
Leather and footwear	-42.9	211.9	-11.5	-35.3	42.6	7.4
Lumber and wood	-62.1	488.2	-12.0	-43.1	61.7	7.5
Paper products	-23.6	70.6	-8.8	-18.5	21.4	6.3
Textile	-47.1	227.4	-13.5	-33.7	48.6	5.0
Apparel	-37.9	253.1	-9.6	-44.1	33.3	15.8
Non-ferrous metals	-29.1	68.2	-10.8	-13.2	29.8	1.2
Machinery except electrical	-51.7	225.7	-16.5	-29.7	61.2	-5.0
Electrical machinery	-45.3	214.7	-15.9	-34.2	48.7	-2.1
Road vehicles	-48.3	151.4	-19.0	-18.5	58.2	-4.0
Transport equipment	-48.0	173.1	-18.4	-24.2	57.4	1.6
Scientific/measuring equip.	-70.8	588.9	-25.0	-41.2	97.0	-4.5
Optical/photographic equip.	-68.2	586.7	-21.8	-44.4	85.8	-0.9

We compare the status quo relative to a situation where regulatory fixed market access costs for foreign firms are as low as for domestic firms. $\Delta\%x = \Delta x/x * 100$.

equipment); see column (1). It is relatively small, on the other hand, for industries with low protection to start with (paper products, chemicals, and rubber and plastics). In any case, reductions in T_h holding variable trade costs fixed, induce more firms to export, thereby implying $\Delta\%p_h^x > 0$; see column (2). Since less productive firms start to export, this comes along with a deterioration of average productivity level ($\Delta\%\tilde{\varphi}_h < 0$); see column (3).

Due to increased competition the least productive input producers are forced to exit, thereby decreasing the mass of firms operating domestically ($\Delta\%M_h^d < 0$); see column (4). On total, however, input diversity rises ($\Delta\%M_h > 0$) since more inputs are imported; see column (5).[26] The increase is relatively large in industries with high initial T_h.

The net effect of harmonization on industry productivity critically depends on $\eta_h/(\sigma_h - 1)$. This ratio turns out to be high for apparel (see Data Appendix), which leads to a large positive effect; see column (6). It is low, however, for knowledge-intensive industries (scientific equipment and machinery). Accordingly, in these industries the rise in input

[26]From a social planner's perspective, there is over-supply of varieties also under T_h^* as $\eta_h < 1$. However, even if $\Delta\%M_h > 0$, the over-supply of varieties relative to the planner's solution is smaller for T_h^* than for T_h.

diversity can not make up for the drop in average productivity of input producers, and the net effect on industry productivity is negative.

2.5 Discussion and conclusions

2.5.1 Discussion

The present model highlights the input diversity and the average productivity effects of TBT reform. These two channels are thought of great interest in recent trade models. However, TBT deregulation can affect outcomes also through additional channels.

Resource saving effect. One may expect a direct *resource saving effect* of lower regulatory costs which may increase the amount of final output per worker in the industry. However, in our model, the resource-saving effect is exactly offset by additional entry so that TBT does not affect productivity through this channel. To see this, let F_h denote sector-h specific resources[27] devoted to fixed costs of entry f_h^e, fixed domestic costs f_h^d, and fixed foreign market costs f_h^x. Making use of the stationarity condition (2.17) and the free entry condition (2.15), one obtains $F_h = L_h/\sigma_h$. Hence, irrespectively of the absolute size of f_h^e,f_h^d, and f_h^x, a constant share of the industry-specific labor force is used for the payment of fixed costs.[28] The result is summarized in Lemma 4.

Lemma 4. *In a stationary equilibrium, the number of workers devoted to fixed costs of entry, domestic regulation, and fixed costs associated with the foreign market is a constant share* $1/\sigma_h$ *of the industry-specific labor force.*

Proof. In the Appendix. □

Pro-competitive productivity gains. Additional entry may reduce the dead-weight loss associated to the existence of monopoly power. In our framework with constant elasticity of substitution between varieties, markups are constant and TBT reform does not lead

[27] Recall: by choice of numeraire, $w_h = 1$.

[28] This result is specific to the CES production function but holds for general productivity distributions. Also, it hinges on free entry of firms. In the Chaney (2007) model, where the number of potential producers is fixed, there would be a resource saving effect. By allowing for free entry, the present paper takes a long-run perspective.

to *pro-competitive productivity gains*. The paper of Melitz and Ottaviano (2008) addresses pro-competitive effects with heterogeneous firms in a model with a linear demand system. However, in that framework, there is no natural role for fixed foreign market access costs (and hence TBT as defined in our paper), since the partitioning of firms into exporters and domestic sellers is achieved by the structure of demand.

Between-industry reallocation. TBT reform potentially induces not only within-industry resource reallocation, but also reallocation between industries. This would require a theoretical framework like the one proposed by Bernard, Redding, and Schott (2007a). However, in that model analytical results are hard to obtain, even for simple setups and with the Pareto assumption. This makes our task of 'sorting out' the intricacies unfeasible. Thus, we have based our analysis on the Melitz framework and relegate the (worthwhile) computational analysis of between-industry reallocation effects to future research.

Learning-by-exporting. Knowledge spillovers from international buyers and competitors may improve the productivity of exporters. The so-called learning-by-exporting hypothesis has been subject of intense empirical research, but has not encountered robust empirical support so far (see the survey by Wagner, 2007). This is why we have refrained from modeling a link between a firm's export status and its productivity.

2.5.2 Conclusions

This paper analyzes the reallocation and industry productivity effects of technical barriers to trade (TBT) reform in a single market with heterogeneous firms and variable and fixed trade costs. The model goes beyond existing versions of the Melitz (2003) model by explicitly parameterizing external scale effects. Our framework allows to disentangle the effect of a TBT reform on average productivity of input producers $\tilde{\varphi}_h$ and input diversity M_h, thereby making the industry productivity effect dependent on the strength of external economies of scale.

We find that – under the parameter constellations obtained in our industry-level calibration exercise – lower TBTs lead to reallocation of market shares from more to less productive firms, potentially negatively affecting industry productivity for a wide range of parameter constellations. However, input diversity usually goes up: To the danger of oversimplification, the aggregate industry-level productivity effect is positive whenever the externality linked to input diversity is strong enough. Our calibration exercise shows that this is indeed the case for most industries. The aggregate effect, however, is sensitive to details of the calibration, while the adverse effect on the productivity of the average firm is fairly robust.

Our analysis has a number of interesting implications: First, while variable trade cost and TBT liberalization both increase the openness of industries, the relation between openness and productivity is unclear. This may rationalize existing empirical results, which are often mixed. Second, whether reductions in variable trade costs improve aggregate productivity depends on the level of TBTs. This interdependence calls for an integrative approach in trade policy. Third, our analysis suggests that TBT reform typically is harder to achieve politically than tariff reform. The reason is that, under a range of parameter constellations, existing exporters would lose market share from lower TBTs but gain from lower variable trade costs.

The present paper suggests an array of interesting extensions. First, we have studies a model of symmetric countries. This is probably defendable on grounds of carving out the general driving forces and sorting out the ambiguities. For a relevant analysis of trade policy, however, a model with asymmetric countries is needed. However, whenever the number of countries goes beyond two, analytical results become hard to come by.

Second, we have treated trade costs as exogenous. It would be interesting to study the strategic setting of TBTs in an asymmetric two-country model. A key challenge is how to deal with the complex adjustment dynamics, that we have ignored in the present model, but which are probably important in any political-economy analysis.

Third, given the ambiguous effects of different types of trade liberalization on aggregate productivity, better estimates of the key parameters governing the model would be highly welcome. This calls for structural estimation and identification of the key parameters in trade models with heterogeneous firms: the shape parameter, the elasticity of substitution, and the degree of external economies of scale.

2.6 Appendix A. Data

Appendix

Variable trade costs, τ_h, come from Hanson and Xiang (2004). The parameter η_h is taken from Ardelean (2007). T_h is calibrated to meet export participation rate p_h^x from Eaton et al. (2004), and n is calibrated to meet openness of 40%. Elasticities of substitution, σ_h, are imputed from shape parameters estimated by Corcos et al. (2007), and sales dispersion measures from Helpman et al. (2004).

Table 2.2 shows how we map industry classifications. If two categories of the source classification fall into one class of our classification, we compute means across source classifications, in case of more than two, we use medians across source classifications. Table 2.3 reports the estimates we take from the literature and the calibrated parameters.

Table 2.2: Industry concordance

Industry	No. γ_h	HS η_h	BEA ς_h	SITC τ_h, σ_h	SIC p_h^x
Chemicals	9	28-31	281,283,284,287	51-56	28
Rubber and plastics	10	39	305,308	57,58	30
Leather and footwear	4	64	310	61,85	31
Lumber and wood	5	45	240	63	24
Paper products	6	48	262, 265	64	26
Textile	2	50-52	220	65	22
Apparel	3	61,62	230	84	23
Non-ferrous metals	12	74-81	335	68	33
Machinery except electrical	14	84	354	73	35
Electrical machinery	15	84	351-353,355-366	71,72,74-77	35,36
Road vehicles	17	87	371	78	37
Transport equipment	17	86,88	379	79	37
Scientific/measuring equip.	16	90,91	381	87	38
Optical/photographic equip.	16	90,91	386	88	38

2.7 Appendix B. Guide to calculations

2.7.1 Theoretical framework

Export cutoff productivity. Evaluating (2.5) and (2.6) at φ_h^* and $(\varphi_h^*)^x$ and solving for r_h^d respectively yields $r_h^d(\varphi_h^*) = \sigma_h f_h^d$ and $r_h^d[(\varphi_h^*)^x] = \sigma_h f_h^d \tau_h^{\sigma_h - 1} T_h$. Divding $r_h^d(\varphi_h^*)$ and

Table 2.3: Parameter description: Estimates from the literature and calibrated values

	Data				Calibration		
Industry	τ_h	p_h^x	η_h	γ_h	σ_h	T_h	$\frac{\eta_h}{\sigma_h-1}$
Chemicals	1.09	0.55	0.62	1.81	2.14	1.32	0.54
Rubber and plastics	1.12	0.44	0.70	2.37	2.21	1.32	0.58
Leather and footwear	1.09	0.26	0.63	2.35	2.15	1.75	0.55
Lumber and wood	1.15	0.12	0.56	2.45	2.34	2.64	0.42
Paper products	1.14	0.45	0.78	1.97	1.99	1.31	0.79
Textile	1.11	0.24	0.59	2.25	2.21	1.89	0.49
Apparel	1.09	0.24	0.59	1.80	1.68	1.61	0.86
Non-ferrous metals	1.06	0.53	0.71	2.21	2.46	1.41	0.49
Machinery except electrical	1.06	0.27	0.39	2.35	2.45	2.07	0.27
Electrical machinery	1.06	0.29	0.39	1.93	2.02	1.83	0.38
Road vehicles	1.10	0.33	0.55	2.06	2.48	1.94	0.37
Transport equipment	1.05	0.33	0.65	2.06	2.34	1.92	0.48
Scientific/measuring equip.	1.05	0.13	0.42	1.84	2.18	3.43	0.36
Optical/photographic equip.	1.05	0.13	0.42	1.84	2.10	3.15	0.38

$r_h^d\left[(\varphi_h^*)^x\right]$ using (2.4) then leads to expression (2.11)

$$\begin{aligned} \frac{r_h^d\left(\varphi_h^*\right)}{r_h^d\left[\left(\varphi_h^*\right)^x\right]} &= \left(\frac{\varphi_h^*}{\left(\varphi_h^*\right)^x}\right)^{\sigma_h-1} = \frac{\sigma_h f_h^d}{\sigma_h f_h^d \tau_h^{\sigma_h-1} T_h} \\ &\Rightarrow \quad \left(\varphi_h^*\right)^x = \varphi_h^* \tau_h T_h^{\frac{1}{\sigma_h-1}} \end{aligned}$$

Average sales-weighted productivity level of domestically produced inputs. Under Pareto, the average sales-weighted productivity level of domestically produced inputs (2.7) is given by

$$\begin{aligned} \left(\tilde{\varphi}_h^d\right)^{\sigma_h-1} &= \frac{1}{1-G\left(\varphi_h^*\right)} \int_{\varphi_h^*}^{\infty} \varphi^{\sigma_h-1} g_h\left(\varphi\right) d\varphi \\ &= \gamma_h \left(\varphi_h^*\right)^{\gamma_h} \int_{\varphi_h^*}^{\infty} \varphi^{\sigma_h-\gamma_h-2} d\varphi \\ &= \frac{\gamma_h}{\gamma_h-(\sigma_h-1)} \left(\varphi_h^*\right)^{\sigma_h-1}. \end{aligned} \tag{2.31}$$

Average sales-weighted productivity level of imported inputs. Correspondingly, under Pareto the average sales-weighted productivity level of imported components is

$$(\tilde{\varphi}_h^x)^{\sigma_h - 1} = \frac{1}{1 - G\left[(\varphi_h^x)^*\right]} \int\limits_{(\varphi_h^x)^*}^{\infty} (\varphi)^{\sigma_h - 1} g_h(\varphi)\, d\varphi \tag{2.32}$$
$$= \frac{\gamma_h}{\gamma_h - (\sigma_h - 1)} \left[(\varphi_h^x)^*\right]^{\sigma_h - 1} :$$

Plugging in the entry cutoff productivity level (2.11) and recalling (2.31), we find that $(\tilde{\varphi}_h^x)^{\sigma_h - 1} = \tau_h T_h^{\frac{1}{\sigma_h - 1}} \tilde{\varphi}_h^d$.

Average sales-weighted productivity of inputs Plugging in (2.12) into (2.14), and using $M_h = (1 + n p_h^x) M_h^d$, we obtain the following expression

$$(\tilde{\varphi}_h)^{\sigma_h - 1} = \frac{1}{M_h}\left[M_h^d \left(\tilde{\varphi}_h^d\right)^{\sigma_h - 1} + n M_h^x \left(\tau^{-1}\tilde{\varphi}_h^x\right)^{\sigma_h - 1}\right] \tag{2.33}$$
$$= \frac{1}{1 + n p_h^x}\left[\left(\tilde{\varphi}_h^d\right)^{\sigma_h - 1} + n p_h^x \left(\tau^{-1}\tilde{\varphi}_h^x\right)^{\sigma_h - 1}\right]$$
$$= \frac{1 + n p_h^x T}{1 + n p_h^x}\left[\left(\tilde{\varphi}_h^d\right)\right]^{\sigma_h - 1}. \tag{2.34}$$

Average profit. By using (2.5), (2.6), (2.7), and (2.12) we can compute average profits.

$$\bar{\pi}_h = \pi_h^d\left(\tilde{\varphi}_h^d\right) + n p_h^x \pi_h^x\left(\tilde{\varphi}_h^x\right) \tag{2.35}$$
$$= \frac{r_h^d\left(\tilde{\varphi}_h^d\right)}{\sigma_h} - f_h^d + n p_h^x\left(\frac{\tau_h^{1-\sigma_h} r_h^d\left(\tilde{\varphi}_h^x\right)}{\sigma_h} - f_h^x\right)$$
$$= f_h^d\left[\left(\frac{\tilde{\varphi}_h^d}{\varphi_h^*}\right)^{\sigma_h - 1} - 1\right] + n p_h^x f_h^d\left[\left(\frac{\tau_h^{-1}\tilde{\varphi}_h^x}{\varphi_h^*}\right)^{\sigma_h - 1} - T_h\right]$$
$$= f_h^d \frac{\sigma_h - 1}{\gamma_h - (\sigma_h - 1)} + n p_h^x f_h^x \frac{\sigma_h - 1}{\gamma_h - (\sigma_h - 1)}$$
$$= f_h^d \frac{\sigma_h - 1}{\gamma_h - (\sigma_h - 1)}\left(1 + n p_h^x T_h\right). \tag{2.36}$$

Given equation (2.13), the average profit line is horizontal in the $(\bar{\pi}_h, \varphi_h^*)$ −space.

Entry cutoff productivity level. Plugging in (2.36) and $p_h^{in} = (\varphi_h^*)^{-\gamma_h}$ into the free entry condition (2.15), and using $p_h^{in} = (\varphi_h^*)^{-\gamma_h}$ and (2.13), one can solve for the entry productivity cutoff level

$$\begin{aligned} f_h^e \delta_h \left(\frac{\varphi_h^*}{\varphi_h^0}\right)^{\gamma_h} &= f_h^d \left(\frac{\sigma_h - 1}{\gamma_h - (\sigma_h - 1)}\right)^{\sigma_h - 1} \left[1 + n\tau p_h^x T\right] \\ \Leftrightarrow \varphi_h^* &= \left\{\frac{\sigma_h - 1}{\gamma_h - (\sigma_h - 1)} \frac{f_h^d}{f_h^e \delta_h} \left[1 + n\tau p_h^x T\right]\right\}^{\frac{1}{\gamma_h}}. \end{aligned}$$

Input diversity. Recall that $r_h^d(\varphi_h^*) = \sigma_h f_h^d$. It follows from (2.4) that $r_h^d(\varphi) = \left(\frac{\varphi}{\varphi_h^*}\right)^{\sigma_h - 1} \sigma_h f_d$. Evaluating r^d at $\tilde{\varphi}_h$, and using $R_h = L_h$, input diversity is given by

$$\begin{aligned} M_h &= \frac{R_h}{r^d(\tilde{\varphi}_h)} \\ &= \frac{L_h}{\sigma_h f_h^d} \left(\frac{\varphi_h^*}{\tilde{\varphi}_h}\right)^{\sigma_h - 1}. \end{aligned}$$

Price index. By inserting optimal pricing of monopolists and the definition of average productivity (2.8) in the definition of the price index (2.3) we obtain

$$P_h = M_h^{-\frac{\eta_h}{\sigma_h - 1}} p(\tilde{\varphi}_h) \tag{2.37}$$

$$= M_h^{-\frac{\eta_h}{\sigma_h - 1}} / \rho \tilde{\varphi}_h. \tag{2.38}$$

2.7.2 Industry productivity effects of TBT reforms

T-neutral deregulation

Entry cutoff level. Differentiating the entry cutoff level (2.16) with respect to f_h^d, holding everthing else constant, we obtain

$$\hat{\varphi}_h^* / \hat{f}_h^d = 1/\gamma_h.$$

Export cutoff level. From equation (2.11) it can immediately seen that

$$\widehat{(\varphi_h^x)^*} / \hat{f}_h^d = \hat{\varphi}_h^* / \hat{f}_h^d$$

Average productivity. It follows from (2.14) along with (2.7) that average productivity changes according to

$$\widehat{\tilde{\varphi}}_h / \hat{f}_h^d = \widehat{\tilde{\varphi}}_h^d / \hat{f}_h^d = \hat{\varphi}_h^* / \hat{f}_h^d.$$

Industry productivity. Plugging $\hat{\varphi}_h^*/\hat{f}_h^d$ into (2.19), we find

$$\frac{\hat{A}_h}{\hat{f}_h^d} = -\left(\frac{\eta_h}{\sigma_h - 1} - \frac{1}{\gamma_h}\right).$$

Incremental mutual recognition

Entry cutoff level From totally differentiating (2.16) one obtains

$$\frac{\hat{\varphi}_h^*}{\hat{T}_h} = -\frac{\gamma_h - (\sigma_h - 1)}{\gamma_h (\sigma_h - 1)} \frac{np_h^x T_h}{1 + np_h^x T_h}. \tag{2.39}$$

Export cutoff level Totally differentiating (2.11) and using (2.39) yields

$$\begin{aligned}\frac{\widehat{\left(\varphi_h^x\right)^*}}{\hat{T}_h} &= \frac{\hat{\varphi}_h^*}{\hat{T}_h} + \frac{1}{\sigma_h - 1} \\ &= \frac{1}{\sigma_h - 1}\left[1 - \frac{\gamma_h - (\sigma_h - 1)}{\gamma_h} \frac{np_h^x T_h}{1 + np_h^x T_h}\right].\end{aligned}$$

Since $\left[\gamma_h - (\sigma_h - 1)\right]/\gamma_h < 1$ and $np_h^x T_h / \left(1 + np_h^x T_h\right)$, we have $\widehat{\left(\varphi_h^x\right)^*}/\hat{T}_h > 0$.

Fraction of exporters. Differentiating (2.13) with respect to T_h, we obtain

$$\frac{\hat{p}_h^x}{\hat{T}_h} = -\frac{\gamma_h}{\sigma_h - 1}. \tag{2.40}$$

Average productivity of domestic producers. It follows from (2.7) that

$$\frac{\hat{\tilde{\varphi}}_h^d}{\hat{T}_h} = \frac{\hat{\varphi}_h^*}{\hat{T}_h}. \tag{2.41}$$

Average productivity. Differentiating (2.14), we obtain

$$\frac{\hat{\tilde{\varphi}}_h}{\hat{T}_h} = \frac{\hat{\tilde{\varphi}}_h^d}{\hat{T}_h} + \frac{1}{\sigma_h - 1} \frac{np_h^x T_h}{1 + np_h^x T_h}\left(\frac{\hat{p}_h^x}{\hat{T}_h} \frac{T_h - 1}{T_h} \frac{1}{1 + np_h^x} + 1\right).$$

Using (2.40), (2.41), and (2.39), we get

$$\frac{\hat{\tilde{\varphi}}_h}{\hat{T}_h} = \frac{np_h^x T_h}{1 + np_h^x T_h}\left(\frac{1}{\gamma_h} - \frac{\gamma_h}{(\sigma_h - 1)^2} \frac{T_h - 1}{T_h} \frac{1}{1 + np_h^x}\right). \tag{2.42}$$

Then

$$\frac{\hat{\tilde{\varphi}}_h}{\hat{T}_h} < 0 \Leftrightarrow \frac{1}{\gamma_h^2} < \frac{1}{\underline{\gamma}_h^2} = \frac{1}{(\sigma_h - 1)^2} \frac{T_h - 1}{T_h} \frac{1}{1 + np_h^x},$$

which can never hold if $T_h < 1$.

Input diversity. From inserting (2.39) and (2.42) into $\hat{M}_h/\hat{T}_h = (\sigma_h - 1)\left(\widehat{\tilde{\varphi}}_h/\hat{T}_h - \widehat{\tilde{\varphi}}_h/\hat{T}_h\right)$, we get

$$\frac{\hat{M}_h}{\hat{T}_h} = \frac{np_h^x T_h}{1+np_h^x T_h}\left(\frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\frac{1}{1+np_h^x} - 1\right). \tag{2.43}$$

Then

$$\frac{\hat{M}_h}{\hat{T}_h} < 0 \Leftrightarrow \frac{1}{\gamma_h} > \frac{1}{\bar{\gamma}_h} = \frac{1}{\sigma_h - 1}\frac{T_h - 1}{T_h}\frac{1}{1+np_h^x},$$

which always holds if $T_h < 1$.

Industry productivity. Plugging (2.43) and (2.42) into (2.19), we have

$$\begin{aligned}
\frac{\hat{A}_h}{\hat{T}_h} &= \frac{1}{\sigma_h - 1}\frac{np_h^x T_h}{1+np_h^x T_h}\left(\frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\theta_h(\eta_h - 1) + \frac{\sigma_h - 1 - \eta_h\gamma_h}{\gamma_h}\right) \\
&= \frac{1}{\sigma_h - 1}\frac{np_h^x T_h}{1+np_h^x T_h}\left(\left(\frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\theta_h - 1\right)\eta_h - \frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\theta_h + \frac{\sigma_h - 1}{\gamma_h}\right) \\
&= \frac{1}{\sigma_h - 1}\frac{np_h^x T_h}{1+np_h^x T_h}\left(\frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\theta_h - 1\right)\left(\eta_h - \frac{\frac{\gamma_h}{\sigma_h-1}\frac{T_h-1}{T_h}\theta_h - \frac{\sigma_h-1}{\gamma_h}}{\frac{\gamma_h}{\sigma_h-1}\frac{T_h-1}{T_h}\theta_h - 1}\right) \\
&= \frac{np_h^x T_h}{1+np_h^x T_h}\left(\frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\theta_h - 1\right)\left(\frac{\eta_h}{\sigma_h - 1} - \frac{1}{\sigma_h - 1}\frac{\frac{\gamma_h}{\sigma_h-1}\frac{T_h-1}{T_h}\theta_h - \frac{\sigma_h-1}{\gamma_h}}{\frac{\gamma_h}{\sigma_h-1}\frac{T_h-1}{T_h}\theta_h - 1}\right) \\
&= \frac{np_h^x T_h}{1+np_h^x T_h}\left(\frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\theta_h - 1\right)\left(\frac{\eta_h}{\sigma_h - 1} - \frac{1}{\sigma_h - 1}\frac{\frac{\gamma_h}{\bar{\gamma}_h} - \frac{\sigma_h-1}{\gamma_h}}{\frac{\gamma_h}{\bar{\gamma}_h} - 1}\right) \\
&= \frac{np_h^x T_h}{1+np_h^x T_h}\left(\frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\theta_h - 1\right)\left(\frac{\eta_h}{\sigma_h - 1} - \frac{1}{\sigma_h - 1}\frac{\gamma_h - \frac{\bar{\gamma}_h}{\gamma_h}(\sigma_h - 1)}{\gamma_h - \bar{\gamma}_h}\right) \\
&= \frac{np_h^x T_h}{1+np_h^x T_h}\left(\frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\theta_h - 1\right)\left(\frac{\eta_h}{\sigma_h - 1} - \frac{\frac{\gamma_h}{\sigma_h-1} - \frac{\bar{\gamma}_h}{\gamma_h}}{\gamma_h - \bar{\gamma}_h}\right) \\
&= \frac{np_h^x T_h}{1+np_h^x T_h}\left(\frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\theta_h - 1\right)\left(\frac{\eta_h}{\sigma_h - 1} - \upsilon_h^*\right)
\end{aligned}$$

Then

$$\hat{A}_h/\hat{T}_h < 0 \Leftrightarrow \begin{cases} \bar{\gamma}_h > \gamma_h > \underline{\gamma}_h \\ \gamma_h < \underline{\gamma}_h \text{ and } \quad \frac{\eta_h}{\sigma_h-1} > \upsilon_h^* \\ \gamma_h > \bar{\gamma}_h \text{ and } \quad \frac{\eta_h}{\sigma_h-1} < \upsilon_h^* \end{cases}$$

Comparing tariff and TBT liberalization

Entry cutoff productivity. From totally differentiating (2.16) one obtains

$$\frac{\hat{\varphi}_h^*}{\hat{\tau}_h} = -\frac{np_h^x T_h}{1+np_h^x T_h} < 0.$$

Export cutoff level Accordingly, totally differentiating (2.11) yields

$$\begin{aligned}\frac{\widehat{(\varphi_h^x)^*}}{\hat{\tau}_h} &= \frac{\hat{\varphi}_h^*}{\hat{\tau}_h} + 1 \\ &= \frac{1}{1+np_h^x T_h} > 0.\end{aligned}$$

Fraction of exporters. From differentiating (2.13) we obtain

$$\frac{\hat{p}_h^x}{\hat{\tau}_h} = -\gamma_h. \tag{2.44}$$

Average productivity. From totally differentiating (2.14) we obtain

$$\frac{\widehat{\tilde{\varphi}}_h}{\hat{\tau}_h} = \frac{\widehat{\tilde{\varphi}}_h^d}{\hat{\tau}_h} - \frac{\gamma_h}{\sigma_h - 1}\frac{np_h^x}{1+np_h^x}\frac{T_h - 1}{1+np_h^x T},$$

where we have used (2.44). Using (2.7), we see that $\widehat{\tilde{\varphi}}_h^d/\hat{\tau}_h = \widehat{\varphi}_h^*/\hat{\tau}_h$. Then totally differentiating (2.16) and (2.14) yields respectively

$$\frac{\widehat{\varphi}_h^*}{\hat{\tau}_h} = -\frac{np_h^x T_h}{1+np_h^x T_h} \tag{2.45}$$

and

$$\frac{\widehat{\tilde{\varphi}}_h}{\hat{\tau}_h} = -\frac{np_h^x T_h}{1+np_h^x T_h}\left(1 + \frac{\gamma_h}{\sigma_h - 1}\frac{1}{1+np_h^x}\frac{T_h - 1}{T_h}\right). \tag{2.46}$$

Then

$$\frac{\widehat{\tilde{\varphi}}_h}{\hat{\tau}_h} < 0 \Leftrightarrow \frac{1}{\gamma_h} > -\frac{1}{\sigma_h - 1}\frac{1}{1+np_h^x}\frac{T_h - 1}{T_h}.$$

Input diversity. From totally differentiating (2.18), we get

$$\frac{\widehat{M}_h}{\hat{\tau}_h} = (\sigma_h - 1)\left(\frac{\widehat{\varphi}_h^*}{\hat{\tau}_h} - \frac{\widehat{\tilde{\varphi}}_h}{\hat{\tau}_h}\right).$$

Plugging in (2.45) and (2.46) yields

$$\frac{\widehat{M}_h}{\hat{\tau}_h} = \frac{\gamma_h (T_h - 1)}{1+np_h^x T_h}\frac{np_h^x}{1+np_h^x}. \tag{2.47}$$

Industry productivity. Inserting (2.46) and (2.47) into (2.19) yields

$$\frac{\hat{A}_h}{\hat{\tau}_h} < 0 \Leftrightarrow \eta_h \frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\frac{1}{1+np_h^x} < 1 + \frac{\gamma_h}{\sigma_h - 1}\frac{T_h - 1}{T_h}\frac{1}{1+np_h^x}. \tag{2.48}$$

Let $\psi_h^* \equiv \frac{1}{\sigma_h - 1} + \frac{1}{\gamma_h} \frac{T_h}{T_h - 1} \left(1 + n p_h^x\right)$. Then condition (2.48) implies $\eta_h / \left(\sigma_h - 1\right) < \psi_h^*$ for $T_h > 1$ and $\eta_h / \left(\sigma_h - 1\right) > \psi_h^*$ for $T_h < 1$ respectively.

Chapter 3

Ethnic Networks, Information, and International Trade: Revisiting the Evidence[1]

3.1 Introduction

Simple static and dynamic neoclassical models predict that more liberal international mobility of labor has the potential to unlock huge efficiency gains; see Moses and Letnes (2004) or Klein and Ventura (2007). These effects derive from a more efficient allocation of the global work force over countries. While hiding potentially important adverse effects in sending countries, they also abstract from gains due to increased international integration of markets driven by networks of migrants.

In an influential paper, Rauch and Trindade (2002) – henceforth R&T – use an empirical trade flow model to show that the network formed by the overseas Chinese population has a major trade creating effect. Quantitatively, they find that "*for trade between countries with ethnic Chinese population shares at the levels of prevailing in Southeast Asia, the smallest estimated average increase in bilateral trade in differentiated products attributable to ethnic Chinese networks is nearly 60%*" (p. 116). They argue that this effect is due to the reduced information costs and improved contracting conditions that networks may bring about. Compared to other determinants of bilateral trade, this effect is large. For example, R&T find that the pro-trade effect of colonial ties is only 13.8%.

[1] This Chapter is based on a working paper, see Felbermayr, Jung, and Toubal (2009). The concept for the paper was developed jointly, the empirical analysis was carried out by the author of this thesis, and writing was shared equally.

The paper by R&T is widely cited. In their survey article, Anderson and van Wincoop (2004) devote substantial space to the results of R&T and argue that the ad valorem tariff equivalent of informational costs is about 6 percent. This is higher than the average tariff rate applied worldwide in recent years.[2] Using data for OECD countries, Evans (2003) argues that tariff equivalents implied by R&T are exaggerated. Existing empirical work connected to R&T makes use of a standard gravity framework. However, in the last years, the econometric modeling of bilateral trade flows has improved due to a sequence of major innovations. Most importantly, Anderson and van Wincoop (2003) have derived a testable gravity equation from the standard monopolistic competition trade model. They show that unbiased estimation of parameters requires to take the so called multilateral resistance terms into account: how strongly trade impediments between two countries reduce their bilateral trade depends crucially on the strength of impediments between each of these two countries with all the other countries that they trade with. This argument is clearly important when quantifying the quantitative importance of ethnic networks: how strongly such a network between two countries encourages bilateral trade depends on the costs of alternative trade routes that these two countries entertain.[3]

Besides potential omitted variable bias, the results of R&T may also suffer from misspecification. Santos Silva and Tenreyro (2006) argue and show empirically that log-linear specifications of the gravity equation may lead to inconsistent estimates if the assumed error term does not enter multiplicatively into the relationship. Liu (2007) emphasizes that this critique also applies to Tobit estimation, the estimation technique used by R&T.[4] One way to deal with this problem is to estimate the gravity model by Poisson pseudo maximum likelihood, which robust is to the type of misspecification mentioned above.[5]

As Combes *et al.* (2005) point out, ethnic networks may affect bilateral trade not only through their effect on trade costs (information and contracting costs), but also through preferences: members of ethnic minorities abroad may derive higher utility from goods im-

[2]The WTO World Trade Report (2007) documents that for the US, Canada, and the majority of European countries, the import-weighted average applied tariff rate was 4.1 percent in 2005.

[3]Controlling for multilateral resistance is crucial, e.g., for the correct estimation of border effects – see the discussion in Feenstra (2004) – and, hence, for dealing with the so called border puzzle (McCallum, 1995).

[4]The article is forthcoming in the *Review of International Economics*.

[5]Liu (2007) also shows that Poisson estimation helps addressing the puzzle raised by Rose (2004) that WTO membership does not create trade.

ported from countries that host their ethnic majority. It is therefore difficult to clearly disentangle the trade cost from the preference effect.[6] Separate identification, however, would be welcome, since trade-cost savings from networks free up resources and therefore represent welfare-improving efficiency improvements. The preference channel is not associated to such efficiency gains. The existence of a measurable and sizable trade-cost effect would be another – hitherto neglected – channel through which international migration leads to an improved allocation of resources worldwide.

In this paper we offer three contributions. First, we discuss the identification of the trade-cost channel of networks in a theory-based gravity model. We argue that, excluding the links of ethnic minorities with the ethnic majority country, one may minimize the preference effect and come closer to the pure trade cost effect. Second, we apply a modern approach to the data of R&T. This avoids a number of problems related to the R&T approach; see Baldwin and Taglioni (2006) for a discussion of those issues.[7] Combining the first and the second point, we show that the large trade-creating effect of 60% estimated by R&T is probably two to four times too large. Most of the overestimation comes from the omission of the multilateral resistance terms; the preference channel seems to be less important.

Finally, we extend the analysis beyond R&T. Using data from the World Bank for the year of 2000, we proxy ethnic networks by the stocks of foreign-born individuals. This gives us a more narrow definition of ethnic networks, because only migration in a life-time can constitute an overseas ethnic minority. However, applying R&T's methodology, we find qualitatively similar results than for the years of 1980 and 1990 with the broader definition of ethnicity. Moreover, the World Bank allows to check for the existence of other ethnic (or better: migrant) networks. Besides the Chinese network, we document the existence of an Indian, a Turkish, or a Mexican network, to name only a few. Interestingly, in terms of trade-creating potential, the Chinese network is by far not the most important one.

Our paper is related to the literature as follows. Besides the paper by R&T, which we

[6] Felbermayr and Toubal (2008) is a first attempt to disentangle the trade costs and preference effects of migrants on OECD bilateral trade.

[7] Baldwin et al. (2008) document the quantitative importance of these problems in a study on the effect of the Euro on trade and investment.

take as our starting point, our analysis is very close to Combes *et al.* (2005). That paper studies the role of social and business networks constituted by inter-regional migrants in France. Using a theory-based gravity approach, they find that these regional networks are quantitatively important and that they may contribute toward an explanation of the border puzzle introduced by McCallum (1995). Our paper is also related to a large literature on the direct effect of migration on bilateral trade. Gould (1994), Head and Ries (1998), Girma and Yu (2000), and Wagner *et al.* (2002) study the trade promoting role of immigration into the U.S. or Canada. Dunlevy (2006) and Bandyopadhyay *et al.* (2008) document a pro-trade effect of migration on the exports of US states. While the older literature usually focuses on bilateral trade of one anchor country with many trade partners, Felbermayr and Jung (2008c) extend the analysis to the full matrix of sending and receiving countries and identify a strong causal effect of bilateral migration on bilateral trade between Southern and Northern countries.

The remainder of this paper is structured as follows. Section 2 introduces the theoretical framework and discusses our econometric approach. Section 3 provides a detailed look at the data. Section 4 and 5 contain our results while section 6 offers concluding remarks. The Appendix further details regression results .[8]

3.2 Econometric specification

3.2.1 A theory-based gravity model

We assume the existence of representative household with CES preferences over domestic and imported varieties of some differentiated good. Different to the standard treatment, we use the utility function proposed in Combes *et al.* (2005) which introduces source-country specific weights a_{ij}. These weights capture the particular attachment of country $i's$ household to imports from country j. We may use this slightly modified utility function in the multi-country monopolistic competition model of international trade proposed by Anderson and van Wincoop (2003), henceforth A&vW. Utility maximization under the appropriate aggregate budget constraint, market clearing, and the assumption that iceberg

[8] Our data, program codes, and further results can be downloaded from `http://www.economics2.uni-tuebingen.de/team/felbermayr/fjt08.7z`

trade costs T_{ij} and preference weights a_{ij} are symmetric $(T_{ij} = T_{ji}; a_{ij} = a_{ji})$, the (c.i.f.) value of bilateral imports M_{ij} can be written as

$$M_{ij} = \frac{Y_i Y_j}{Y_w} \left(\frac{T_{ij}}{a_{ij}} \right)^{1-\sigma} \left(\tilde{P}_i \tilde{P}_j \right)^{\sigma-1}, \tag{3.1}$$

where the price indices $\tilde{P}$ solve $\left(\tilde{P}_j\right)^{1-\sigma} = \sum_{i=1}^{C} (Y_i/Y_w)(T_{ij}/a_{ij})^{1-\sigma} \left(\tilde{P}_i\right)^{\sigma-1}$; see Feenstra (2004) for the details of the derivation. A&vW call $\tilde{P}_i$ indices of multilateral resistance because they depend on the trade costs of country i with all countries in the world, the number of which is given by C. The variables Y_i denote GDP of country i, the subindex w refers to the world. The elasticity of substitution in the underlying CES utility function is given by σ. We will be interested by the determinants of T_{ij} in general, and by the cost of obtaining information in particular. Following the literature, we assume that T_{ij} is a log-linear function of its determinants.

The central insight of A&vW is that the volume of trade between i and j depends not only on the trade costs between i and j but on the entire distribution of trade costs between i and j and *all other countries* of the world. How strongly T_{ij} restricts trade between i and j depends on the costs that affect trade with alternative partners. Hence, in the estimation we have to deal with the $\tilde{P}_i$ terms. We also have to decide about the appropriate econometric estimation technique. Finally, in order to make the role of networks explicit, we need to model T_{ij} and a_{ij}. We deal with the first two issues first and relegate the modeling of trade costs and preferences into section 3.2.2.

The multilateral resistance terms $\tilde{P}_i$ are essentially unobserved since they do not correspond to official CPI deflators. A&vW show how one can solve for the $\tilde{P}_i$ terms numerically and use them in an iterative estimation strategy. They demonstrate that the failure to control for multilateral resistance typically biases the absolute value of estimated trade cost variables upwards. R&T recognize the problem of multilateral resistance (without mentioning the issue) by adding an ad-hoc remoteness term to their regressions. Ex ante, it is unclear whether this is sufficient to deal with omitted variable bias. In our regressions, we follow Feenstra (2004) who argues that the use of importer and exporter specific fixed effects in a simple OLS model leads to very similar results than A&vW's strategy but is technically much less demanding. We opt for this strategy, which is now common in virtually all gravity applications. In order to save on degrees of freedom, we do not allow for separate role for importer and exporter fixed effects; rather, we will use country dummies which, nevertheless, fully control for all purely country-specific variables such as the $\tilde{P}_i$ terms; see Baier and Bergstrand (2007) for a similar strategy.[9]

Traditionally, the gravity literature estimates a log-linear version of (8.1). In non-stochastic form, the relationship between the multiplicative constant-elasticity model (8.1) and its log-linear additive formulation is trivial. This does no longer hold if trade flows are measured with error. Santos Silva and Tenreyro (2006) warn that heteroskedastic residuals do not only lead to inefficiency of the log-linear estimator, but also cause inconsistency. This is because of Jensen's inequality which says that the expected value of the logarithm of a random variable is different from the logarithm of its expected value, i.e., $E(\ln M_{ij}) \neq \ln E(M_{ij})$. Then, $E(\ln M_{ij})$ not only depends on the mean of M_{ij}, but also on higher moments of the distribution. Thus, heteroskedasticity in the residuals, which on a first glance only affects efficiency of the estimator, feeds back into the conditional mean of the dependent variable, which, in general, violates the zero conditional mean assumption on the error term needed to guarantee consistency.

To be more precise, consider that the true model can be represented as $M_{ij} = \exp(\beta X_{ij}) + \epsilon_{ij}$. Then, estimating a log-linear model of the form $\ln(M_{ij}) = \beta X_{ij} + \ln(\eta_{ij})$ would imply that $\eta_{ij} = 1 + \epsilon_{ij}/\exp(\beta X_{ij})$. Hence, $E[\eta_{ij}|X_{ij}]$ can only be independent of X_{ij} for the special case $\epsilon_{ij} = \exp(\beta X_{ij})\nu_{ij}$, where ν_{ij} is a random variable statistically independent of X_{ij}. In general, this requirement is violated.

Santos Silva and Tenreyro (2006) solve these problems by estimating the gravity equation multiplicatively (without taking the logarithm on M_{ij}) and allowing for heteroskedasticity. Their proposed estimator is equivalent to the Poisson pseudo-maximum likelihood estimator (PML), the most commonly used conditional mean specification of which is $E(M_{ij}|X_{ij}) = \exp(\beta X_{ij})$. Coefficients can be explained as elasticities if the dependent variable is in level and covariates X_{ij} are in logs. It is worthy to note that country fixed effects can be included

[9] R&T do not overtly address the issue of multilateral resistance. They implicitly deal with it in an ad hoc way by introducing a variable called remoteness.

in the PML model as a control for multilateral resistance terms.

Santos Silva and Tenreyro (2006) justify the hypothesis that conditional variance is proportional to the conditional mean for the Poisson model, although the Poisson regression is consistent even when the variance function is misspecified.[10]

Liu (2007) argues that the problem of inconsistency due to heteroskedasticity also applies to the Tobit estimator, which has been used by R&T. Moreover, they use the log of total bilateral trade $\ln\left[\left(M_{ij}+M_{ji}\right)/2\right]$ as the dependent variable, which would be correct only if the theoretical assumption of perfect symmetry in trade costs $\tau_{ij}=\tau_{ji}$ was to be taken literally and the error terms were symmetric, too.

3.2.2 The role of networks

We now need to formalize the role of ethnic (or: migrants') networks. While there is little doubt that such networks may play an important role in conveying important information about the trading opportunities between countries, there is no apparent consensus in the existing literature as to how such networks are to be defined and modeled.

We define as the k-ethnic network the set of bilateral links between all countries in the world maintained by members of the ethnicity k. In other words, there are as many ethnic networks as there are ethnicities in the world. In our empirical work, we will assume that every ethnicity is associated to exactly one country in the world.[11] Moreover, most of our analysis concentrates on the most sizable ethnic network studied by R&T: that of Chinese.

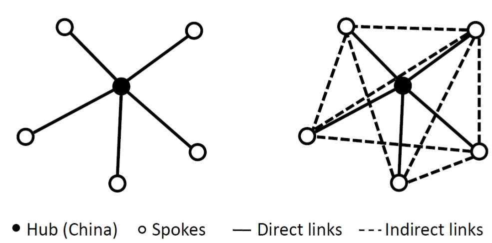


Figure 3.1: Direct and indirect links in a network of countries.

[10]Using nonparametric tests, Henderson and Millimet (2008) recommend estimation of the gravity model in levels. Wooldridge (2002, p.676) emphasizes "while the leading application is to count data, the fixed effect Poisson estimator works whenever the conditional mean assumption holds. Therefore, the dependent variables could be a nonnegative continuous variable, or even a binary response if we believe the unobserved effect is multiplicative...".

[11]Obviously, the number of ethnicities is much larger than the number of independent countries since there are many ethnicities without their own state, e.g., the Kurds. We abstract from this possibility.

Figure 3.1 illustrates a stylized network of six countries. The single star in the left panel depicts the Chinese ethnic network. It illustrates the links between the hub (China), illustrated by a full black circle, and its spokes (other countries), depicted by hollow circles. *Through the hub*, all spokes are linked with each other. The right-hand panel in Figure 3.1 depicts some of the bilateral trade links between the six countries. The solid lines coincide with the ethnic network which will affect trade relationships directly. We therefore use the term *direct* links. Bilateral trade flows between spokes are illustrated by dashed lines. Since the ethnic network affects those flows only through links to the same hub, we talk about *indirect* links. For simplicity, we assume that each ethnic (or migrant) network is associated to a single hub, but this need not be so in all cases. The strength of the link between any spoke i and the hub is measured by the share of individuals with ethnicity k in the total resident population of spoke i, and denoted by s_{ik}.

Clearly, ethnic networks can foster trade along the direct links. They also, however, potentially affect trade between spokes i and j. This is the case, because migrants with ethnicity k residing in spoke i convey information on trading opportunities with migrants of the same ethnicity residing in spoke j, i.e., information about one country is made available through the k-network. Indirect k-ethnic links between spokes are measured by $N_{ij}^k = s_{ik}s_{jk}$, for all i, j, k.[12] The k-ethnic network is then just the vector N^k that collects all elements N_{ij}^k for all i and j.

Following R&T, we assume that trade costs T_{ij} are a function of geographical measures related to transportation costs (distance, adjacency), of variables related to trade policy (membership in regional trade agreements), a variable measuring cultural proximity (common language), and one related to historical ties (joint colonial past). Central to our analysis, T_{ij} also depends on the network variable defined above. We assume that the trade cost function can be linearized. Collecting all variables other than the network into the (row) vector $\mathbf{X}_{ij}$, we may therefore posit $\ln T_{ij} = \boldsymbol{\xi}'_T \mathbf{X}_{ij} - \sum_k \nu_T^k N_{ij}^k$, where $\boldsymbol{\xi}$ is a vector of coefficients, N_{ij}^k measures the strength of the k-ethnic network (CHINSHARE in R&T), and ν_T^k is the associated coefficient measuring the effect of the k-ethnic network on trade costs (expected to be positive). Evidence in favor of $\nu_T^k > 0$ would suggest that the network

[12] Note that $N_{ij}^k \leq N_{kj}^k$, $j, i \neq k$.

lowers informational or contractual costs, thereby encouraging trade through lower total trade costs. This is the *trade cost channel* of networks which R&T focus on in their paper.

Similarly, we may posit that country $i's$ cultural, political, or geographical proximity to country j increases the weight of goods imported from i, so that $\ln a_{ij} = \boldsymbol{\xi}_a' \mathbf{X}_{ij} + \sum_k \nu_a^k N_{ij}^k$, where ν_a^k is expected positive. Evidence for $\nu_a^k > 0$ would be in line with the existence of a *preference effect* of ethnic networks.

Employing these specifications for T_{ij} and a_{ij} in (8.1), and using non-overlapping sets of country dummies $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ to control for all country-specific variables, we have

$$\begin{aligned} M_{ij} &= \exp\left\{\ln\left(Y_i Y_j\right) + (\sigma - 1)\left(\boldsymbol{\xi}_a' - \boldsymbol{\xi}_T'\right)\mathbf{X}_{ij} + \sum_k (\sigma - 1)\left(\nu_a^k + \nu_T^k\right) N_{ij}^k + \boldsymbol{\mu}_i + \boldsymbol{\mu}_j\right\} + \epsilon_{ij}, \\ &= \exp\left\{\ln\left(Y_i Y_j\right) + \bar{\boldsymbol{\xi}}' \mathbf{X}_{ij} + \sum_k \bar{\nu}^k N_{ij}^k + \boldsymbol{\mu}_i + \boldsymbol{\mu}_j\right\} + \epsilon_{ij}. \end{aligned} \tag{3.2}$$

Clearly, the estimated coefficients $\bar{\boldsymbol{\xi}}'$ and $\bar{\nu}^k$ will reflect the elasticity of substitution σ as well as the effect of X or the network on trade costs *and* preferences. In other words, there is a twofold identification problem. First, the identification of the *total* network effect is impossible without external information about σ. Second, the trade cost and the preference channels are typically confounded.

At this point, we want to make two observations. First, R&T run equation (5.14) on different dependent variables: trade in differentiated goods, trade in reference-price goods, and trade in exchange-traded goods. This classification being directly related to the degree of substitutability σ, there are no clear predictions concerning the comparison between parameter estimates $\bar{\boldsymbol{\xi}}'$ and $\bar{\nu}^k$ obtained from these different regressions. For example, even if the trade cost and the preference channel could be separated, for a given strength of the network effect ν_T^k, the estimated coefficient $(\sigma - 1)\,\nu_T^k$ would be large for homogeneous goods since the degree of substitutability is high and low for differentiated goods. The opposite may be true if, for given σ, ν_T^k varies across the groups of goods. However, neither σ nor ν_T^k can be assumed constant over those subaggregates of goods so that the naive comparison of coefficients obtained from different regressions is problematic.

Second, in general, any estimate of $\bar{\nu}^k$ reflects the preference and trade cost effect of the k-ethnic network.[13] However, the following observation may help in the separate identification of the channels. Any ethnic (or migrant) network consists of direct and indirect links. Direct links are those that relate an individual of ethnicity k residing in country i to another individual of the same ethnicity *at the hub*, namely country k. Indirect links, in turn, relate the individual to another one of the same ethnicity in country $j \neq k$. If migrants (or their offspring) have special preferences for goods produced in country k, then direct links will reflect the preference channel along with information channel. The preference channel should, however, not be so important in indirect links, since these do not relate to the country of origin. Rather, indirect links should only reflect the information channel.[14]

There is another advantage of looking at indirect links: it may well be that citizens of country k move to country i (and vice versa) as a response to some positive shock to the trading potential between the two countries. Then, the direct ethnic link $s_{ik}s_{ki}$ would be endogenous to the volume of bilateral trade. In contrast, the indirect links $s_{ik}s_{kj}$ would not be affected.[15]

Summarizing, our econometric approach differs from R&T in the following ways:

1. In all of our specifications, the dependent variable is the log (or level, depending on the model) of *imports* rather than the log (or level) of the arithmetic average over imports and exports. This implies that we have two observations per country pair instead of only one. This increases the degrees of freedom, but requires to control for correlation of error terms within each pair.

2. We control for the multilateral resistance terms and all other country-specific determinants of trade costs, policy, history, etc., by including a complete set of *country*

[13] R&T conceptually decompose the trade cost channel into an 'contractual enforcement' and an 'informational' component. They try to isolate the informational part by distinguishing between differentiated, reference-priced and exchange-traded goods. Their identifying assumption is that network improve contractual enforcement for all categories of goods, but information is only relevant for differentiated goods. Hence, the difference between the network estimates for differentiated and exchange-traded goods reflects information; see also Anderson and van Wincoop (2004). We do not wish to push this interpretation, since equation (5.14) shows that estimated coefficients would also reflect systematic differences in elasticities of substitution across categories of goods, which are essentially unobserved. Rather, we subsume both effects under 'trade costs'.

[14] R&T propose a similar strategy in a section where they measure the strength of networks in levels rather Chinese ethnic population shares.

[15] Clearly, any combinations of $s_{ik}, s_{jk}, s_{ki}, s_{kj}$ would satisfy this criterion.

fixed-effects. This strategy also mitigates spurious correlation concerns driven, e.g., by language etc..

3. Our preferred specification is a *Poisson (pseudo) maximum likelihood* approach with country fixed-effects.

4. Since the comparison of results by commodity group is complicated by a (potentially) varying degree of substitutability, we also show results for aggregate trade,

5. Besides computing the total network effect, as R&T do, we present *direct* and *indirect* effects for the case of measuring the strength of network in shares, where the latter are supposed to be more informative about the pure trade cost channel.

3.3 Data

3.3.1 Trade data

R&T estimate the effect of Chinese ethnic networks on different dependent variables: trade in differentiated goods, trade in reference-priced goods, and trade in exchange-traded goods. This classification requires trade data at the level of the four-digit Standard Industrial Trade Classification (SITC) Revision 2, which can be downloaded from the United Nations Commodity Trade Statistics Database (UN Comtrade).

Since the raw data are incomplete in time, country, and commodity coverage, several attempts have been made to recompile the data, thereby allocating exports to unspecified regions, and correcting for entrepôt trade. The correct identification of trading partners seems to be an important issue, which is also recognized by R&T.

Statistics Canada has constructed the World Trade Database (WTDB), covering the years 1970-1997. Feenstra (2000) concludes that the "method of dealing with entrepôt trade seems to be adaptable to the situation of an entrepôt country as the Netherlands. [...]. It does not seem to cover the case of entrepôt trade countries such as Hong Kong or Singapore" (p. 4). In order to assess the severity of the problem, Feenstra (2000) compares the total value of U.S. imports from China and Hong Kong, respectively, from Statistics Canada and U.S. Census data, and finds these values to be "reasonably close" (p. 7) up until 1983.

R&T have made use of an early version of the WTDB. Unfortunately, this data is no longer distributed by the NBER. A slightly revised version covering the years 1980-1997 is made available by Robert Feenstra, and can be downloaded as UCD-Statistics Canada Trade Data. The data differ from those used by R&T, because they do not contain zero trade flows. Rather, trade flows below 1,000 thousand U.S. dollar are coded as missing.

Robert Feenstra also provides a newer dataset (NBER-UN World Trade Data), covering the years 1962-2000. Data for early years (1962-1983) are taken from UN Comtrade, making adjustments for country codes only. For the latter years (1984-2000), data only cover 72 countries, and are adjusted in several ways. Most importantly, Feenstra et al. (2005) revise Chinese exports shipped through Hong Kong.[16]

In order to take advantage of all the corrections made, we utilize the UCD-Statistics Canada Trade Data for 1980, and the NBER-UN World Trade Data for 1990 and 2000. We restrict our sample to the 63 countries used by R&T. Unfortunately, the 72 reporting countries in the NBER-UN World Trade Data do not completely overlap the 63 countries of interest, such that we do not have the full trade flow matrix.[17]

3.3.2 Migration data

Data on Chinese ethnic networks for 1980 and 1990 is taken from R&T. In order to check the existence of migrant networks, we utilize the World Bank international bilateral migration stock database which is available for 226 countries and territories and is described in detail by Parsons et al. (2007). Rather than including all persons with any Chinese ancestry, the World Bank data comprise migrants which have been born in China and now reside in a foreign country. While the migration data are broken down by receiving country, the data make no reference to the time at which migration has taken place (Parsons et al., 2007, p. 4). It allocates the total outstanding stock of 175.7 million international migrants over sending and receiving countries.

[16] Feenstra et al. (2005) estimate the value-added in Hong Kong on re-exports, and reduce the value of imports from China and increase the value of imports from Hong Kong by this amount. The markup calculation is described in detail in Feenstra et al. (1999), and discussed in Feenstra et al. (2005).

[17] Countries and data availability are listed in the Appendix. In order to come from trade data on four-digit SITC level to trade by commodity group, we make use of the Rauch (1999) classification. In order to save space, we focus on the 'liberal' aggregation which maximizes the number of SITC categories classified as either exchange-trade or reference-priced goods in case of ambiguities. R&T compare results for 'liberal' and 'conservative' aggregation rules and find no qualitative and quantitative difference.

Both the Chinese ethnic and the migration network cover Chinese citizens residing abroad and naturalized citizens of Chinese descent. Whereas the Chinese ethnic network also captures descendants of Chinese parents, people born who have just been born in China without being of Chinese ancestry add to the Chinese migrant network. In any case, the focus of the migrant network is on people who have moved during their lifetime.[18]

The World Bank matrix also allows for examining the role of other networks in a similar way. In absolute values, Mexico is the top sending country with more than 10.1 million of its natives living abroad. However, 92.6% of emigrants go to neighboring countries. This ratio is 18.4% for India, the second largest sending country with about 9.0 million of its natives abroad, 42.2% for China (fourth largest expatriate population, 5.8 million), 6.2% for Turkey (10th largest expatriate population, 3.0 million), or 13.2% for Morocco (12th largest expatriate population, 2.6 million).[19]

3.3.3 Other data

Data on population in 2000 and GDP come from the World Development Indicators (WDI).[20] Data on geographical and cultural proximity like distance, use of a common official language, colonial ties, and common colonizer are taken from the CEPII. Following R&T, we include dummies for common membership in the EEC and EFTA for 1980 and 1990. In 2000, we additionally control for common membership in NAFTA and MERCOSUR, which seem to be the most important regional free trade agreements at that time.

3.4 Results

In this section, we present results for the effect of Chinese networks on trade. The discussion of other potential networks is relegated to the next section. Following R&T, we start with looking at the effect of country pairs trading along the direct *and* indirect links. While this strategy disallows to distinguish between preference and trade cost channel, we proceed with

[18]This criterion is often not met in case of the Former Soviet Union, Yugoslavia, and Czechoslovakia, where the break-ups of former states have "produced" migrants. However, our analysis does not cover these countries. Moreover, Parsons et. al (2007) states that "the return of Hong Kong to Chinese sovereignty in 1997 did not reduce the number of migrants" (p. 9).

[19]This collection reflects the largest sending countries for which we find network effects in our empirical analysis below.

[20]Unfortunately, WDI do not cover Taiwan which is therefore excluded from our analysis. It turns out that the replication of R&T's results does not hinge on the inclusion of Taiwan.

a decomposition of the average effect. In order to make transparent how our estimation strategy impacts on the trade creation of Chinese ethnic networks in Southeast Asia (where the network is quantitatively strong), we also decompose the average effect along the lines of strong and weak networks.

We do not only present the estimated coefficients, but also compute implied trade creation and associated ad valorem tariff equivalents. We do this in order to make our results comparable to the results presented by R&T and A&vW, respectively.

3.4.1 The direct and indirect effect of the Chinese network on aggregate bilateral trade

We start the discussion of our results by looking at aggregate bilateral trade. Hence, the dependent variable records the total value of imports of country i from country j. In later tables, we will disaggregate bilateral trade flows into the groups of exchange-trade, reference-priced, and differentiated goods, as proposed by R&T.

The first three columns in Table 3.1, (A1) to (A3), show the effect of the Chinese ethnic network on the value of bilateral trade, without distinguishing between direct and indirect network links. The list of controls is identical to R&T. This implies that we also use the product of *per capita* GDPs, despite the fact that the standard theoretical derivations of the gravity equation do not allow any role for this variable.[21] Column (A1) replicates R&T for the case of aggregate trade and the year of 1980. The coefficients on standard gravity covariates appear with signs and magnitudes comparable to those found by R&T and other studies: the coefficients on the product of GDPs and distance are close to -1 and 1, respectively. The dummies controlling for common membership in regional trade agreements (EEC, EFTA) yield implausible results (this is common, see Baier and Bergstrand, 2007). Common language and colonial ties have large and significant effects, and the adjacency dummy is not statistically significant.

[21] With non-homothetic preferences, there would be a natural role for per capita income in gravity equations.

Table 3.1: The Chinese network in aggregate trade

Dependent variable: Aggregate trade															
			1980					**1990**					**2000**		
	(A1)	(A2)	(A3)	(A4)	(A5)	(B1)	(B2)	(B3)	(B4)	(B5)	(C1)	(C2)	(C3)	(C4)	(C5)
	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML
CHIN	4.589***	0.853**	0.982**			4.526***	1.259***	1.816***			5.757***	1.711**	3.335*		
	(7.24)	(2.45)	(2.35)			(8.53)	(2.60)	(4.69)			(3.86)	(2.07)	(1.76)		
CHIN*(1-DIR)				1.062**	0.517				0.943**	0.473				102.5***	44.01
				(2.41)	(1.21)				(2.52)	(1.00)				(2.61)	(1.53)
CHIN*DIR				0.747*	1.121**				1.416**	2.037***				1.782**	3.514*
				(1.77)	(2.44)				(2.23)	(5.40)				(2.13)	(1.85)
ln(GDP_i GDP_j)	0.947***	1.135***	0.945***	1.135***	0.945***	0.876***	0.909***	0.859***	0.909***	0.838***	1.013***	1.050***	0.860***	1.049***	0.860***
	(42.72)	(27.43)	(35.36)	(27.42)	(35.36)	(51.36)	(20.73)	(14.49)	(20.73)	(13.79)	(66.62)	(31.26)	(29.69)	(31.27)	(29.71)
ln(PGDP_i PGDP_j)	0.345***	-0.0142	0.0462	-0.0154	0.0528	0.231***	0.00165	0.0148	0.00321	0.0285	0.133***	0.0786*	0.170***	0.0782*	0.169***
	(11.84)	(-0.29)	(1.45)	(-0.31)	(1.57)	(12.17)	(0.04)	(0.30)	(0.08)	(0.57)	(7.42)	(1.70)	(4.28)	(1.69)	(4.24)
ln(DISTANCE)	-1.080***	-1.124***	-0.699***	-1.123***	-0.703***	-0.935***	-0.992***	-0.569***	-0.993***	-0.586***	-1.148***	-1.197***	-0.538***	-1.192***	-0.536***
	(-16.88)	(-19.65)	(-17.41)	(-19.48)	(-17.17)	(-19.47)	(-23.58)	(-15.20)	(-23.49)	(-15.31)	(-25.48)	(-28.38)	(-15.27)	(-28.19)	(-15.09)
ADJACENT	-0.0463	0.0762	0.438***	0.0785	0.431***	0.431**	0.450**	0.794***	0.446**	0.750***	-0.0603	-0.0923	0.355***	-0.0804	0.362***
	(-0.20)	(0.34)	(3.65)	(0.35)	(3.56)	(2.36)	(2.32)	(9.02)	(2.30)	(8.38)	(-0.33)	(-0.47)	(3.34)	(-0.41)	(3.40)
EEC	-0.673***	-1.735***	-0.107	-1.734***	-0.111	-0.109	-0.347**	0.513***	-0.348**	0.494***	-0.147	-0.593***	0.577***	-0.589***	0.577***
	(-3.96)	(-8.82)	(-0.72)	(-8.81)	(-0.74)	(-0.96)	(-2.32)	(4.38)	(-2.33)	(4.22)	(-1.51)	(-5.19)	(4.83)	(-5.16)	(4.83)
EFTA	0.133	0.0338	0.385**	0.0341	0.386**	-0.458***	-0.0563	0.354**	-0.0563	0.365**	0.548	0.981	-0.744**	0.983	-0.745**
	(0.65)	(0.18)	(2.24)	(0.19)	(2.25)	(-2.95)	(-0.31)	(2.39)	(-0.31)	(2.48)	(0.72)	(1.17)	(-1.96)	(1.17)	(-1.96)
NAFTA											0.173	0.626	0.908***	0.632	0.908***
											(0.29)	(0.81)	(4.50)	(0.82)	(4.51)
MERCOSUR											0.0964	0.837	1.874***	0.840	1.871***
											(0.18)	(1.50)	(12.11)	(1.51)	(12.09)
LANGUAGE	0.573***	0.528***	0.237**	0.528***	0.238**	0.496***	0.514***	0.0688	0.514***	0.0769	0.604***	0.640***	0.212**	0.632***	0.206**
	(4.46)	(4.22)	(2.20)	(4.22)	(2.22)	(5.06)	(5.35)	(0.87)	(5.34)	(0.98)	(6.85)	(6.94)	(2.48)	(6.82)	(2.39)
COLOTIE	0.631***	0.657***	0.157	0.656***	0.157	0.523***	0.453***	0.0840	0.453***	0.0753	0.433***	0.437***	0.0883	0.442***	0.0905
	(2.99)	(3.12)	(1.27)	(3.12)	(1.26)	(3.02)	(3.00)	(0.73)	(3.00)	(0.65)	(2.91)	(2.79)	(0.75)	(2.82)	(0.77)
ln(REMOTE)	0.858***					0.710***					1.098***				
	(6.36)					(6.35)					(11.15)				
R2	0.641	0.722	0.904	0.722	0.904	0.719	0.793	0.930	0.793	0.931	0.753	0.817	0.930	0.817	0.930
Trade creation (%)															
CHIN	1.445	0.267	0.307			1.307	0.362	0.522			0.176	0.0523	0.102		
CHIN*(1-DIR)				0.333	0.162				0.271	0.136				3.183	1.354
CHIN*DIR				0.234	0.351				0.407	0.586				0.0545	0.107
Tariff equivalent (%)															
CHIN	0.205	0.0382	0.0439			0.186	0.0517	0.0745			0.0251	0.00747	0.0146		
CHIN*(1-DIR)				0.0475	0.0231				0.0387	0.0194				0.448	0.192
CHIN*DIR				0.0334	0.0502				0.0581	0.0836				0.00778	0.0153

N=2520 in 1980, N=2795 in 1990, and N=3259 in 2000. All regressions include a constant. FE-OLS and FE-PML include country dummies. Observations clustered by (undirectional) country-pair. Robust t statistics in parenthesis. *, **, *** indicate significance at the 1%, 5%, and 10% level, respectively. Trade creation (%) and ad valorem tariff equivalents (%) evaluated at the respective sample means. Elasticity of substitution is eight.

The variable of interest is $CHIN$. The coefficient obtained under OLS without fixed effects in column (A1) yields a point estimate of 4.589 and a robust, cluster-corrected, t-value of more than 7, which is comparable to results for trade by commodity group reported by R&T.[22] That effect amounts to total trade creation of about 1.5%, if assuming that $CHIN$ moves from zero to the sample average.[23] In terms of ad valorem tariff equivalents, the estimated network effect is equivalent to a hypothetical tariff reduction of about 0.2 percentage points.[24] This is much smaller than the headline result of 60% trade creation or, equivalently, 6% tariff equivalent, discussed by Anderson and van Wincoop (2004), which focus on differentiated goods, and relate to the effect of the network when both concerned countries have *large* (i.e., larger than 1%) ethnic chinese populations. Table 3.9 columns (A1) and (B1) replicate the findings by R&T.

Column (A2) includes country-specific fixed effects to deal with multilateral resistance. This changes the usual gravity covariates only modestly, with the exception of common EEC membership and colonial ties. In sharp contrast, the network effect drops to 0.853 and is only about 19% as big as the one obtained without fixed effects. Statistical significance, however, is maintained, with a t-value of 2.45. The amount of trade creation or the tariff equivalents are scaled downwards to 0.3% and 0.04%, respectively.[25] Finally, column (A3) replaces OLS estimation with Poisson (pseudo) maximum likelihood (PML). Compared to (A2), the heteroskedasticity-robust approach does not lead to important further changes and has only minor effects on the accuracy of the estimate.

Columns (A4) and (A5) decompose the total network into direct (involving mainland China) and indirect links (not involving China as a trade partner). The dummy variable DIR takes the value of one if the bilateral relationship involves China and zero otherwise. Using fixed-effects in an OLS model, the direct effect comes with an estimate of 0.747 and the indirect one with 1.062, both estimated at satisfactory (though not excellent) statistical precision. Using the fixed-effects PML model, we do not find any evidence for the indirect

[22]As R&T point out, the Tobit and OLS without fixed effects yield qualitatively and quantitatively comparable results.

[23]The formula employed is $100 \times \left[\exp\left(\bar{\nu} \times \overline{CHIN}\right) - 1\right]$, where $\bar{\nu}$ is the obtained coefficient and $\overline{CHIN}$ the sample mean; see R&T. Summary statistics are shown in the Appendix.

[24]The formula employed is $100 \times \bar{\nu} \times \left[\exp\left(\overline{CHIN}\right) - 1\right] / (\sigma - 1)$. We use the same assumption on σ as Anderson and van Wincoop (2004), i.e. $\sigma = 8$.

[25]This is less than 19% smaller due to the non-linearity of the trade cost function.

effect any more. This finding suggests that the preference channel is probably quantitatively more important than the information channel. However, also the preference channel is associated to a fairly modest amount of trade creation (0.35%) and equivalent to a small tariff (0.05%).

Columns entitled (B1) to (B5) repeat the exercise for the year of 1990. The sample composition and the total number of observations is different, but the estimated coefficients are mostly qualitatively and quantitatively similar to those obtained for 1980. However, looking at our preferred specifications (B3) and (B5), we find a larger role for the Chinese network. The total effect now amounts to trade creation of 0.5% and to a tariff equivalent of 0.07% (both about 70% larger than in 1980). This effect is virtually entirely driven by the direct effect. The overall conclusions from 1980 remain robust: the network effect is dramatically reduced when using fixed effects, its economic significance is small, and the total effect is mostly driven by the direct effect.

The remainder of the table turns to the year 2000 where we use data on bilateral stocks of foreign born individuals rather than on ethnic populations. The network variable, constructed as the product of shares in each of the two trading countries' populations, is smaller than for ethnic populations since the concept of foreign-born status is more narrow than that of ethnicity. It is therefore not surprising that the estimated network coefficients are larger. However, when evaluated at the sample mean, the associated amounts of trade creation or the tariff equivalents are again small. The total effect yields trade creation of 0.1% (column C3) with marginal statistical significance. There is no evidence for an indirect effect, and the direct effect, though statistically significant, is small. Note that the estimated coefficient of the indirect effect obtained under FE-OLS (C4) is huge (102.5). That number, together with the estimated for the direct effect, is not plausible, since it opens an interval that does not encompass the average (total) effect found in column (C2). We may conclude that it is crucial to use the PML strategy since the bias due to potential misspecification of the error structure can be large. However, it is qualitatively not important whether ethnic networks are measured using data on overseas ethnic populations (as for years 1980 and 1990), or on populations of foreign born individuals.

3.4.2 The Chinese network by commodity group

In Table 3.2, we present the trade creation and tariff equivalent results and the significance level of the associated coefficients for different product categories. We only report the estimation results from the FE-OLS and FE-PML specifications.[26]

The upper third of the table refers to the group of exchange-traded goods; the second to the group of reference-priced goods; and the third to differentiated goods. It is natural to suppose that the underlying degrees of substitutability differ in those groups. Since exchange-traded goods are show-cases for homogeneous goods such as steel, corn, or ore, the elasticity of substitution can be expected to be much larger than in the group of differentiated goods. These differences are taken into account when computing the ad valorem tariff equivalents associated to each network coefficient.

According to R&T, one may expect that the network effect should be largest for commodity goods, smaller for reference-priced goods, and minimum for goods traded on organized exchanges. This conjecture does no longer generally hold true when the gravity equation is estimated in a theory-consistent way by including fixed-effects. In columns (A1) and (B1), exchange-traded goods yield the largest trade-creation effects and tariff equivalents.[27] It is, therefore, no longer meaningful to draw on the comparison of the coefficients obtained with different trade categories using the R&T network variable to disentangle the respective roles of the information and contract-enforcement channels as R&T propose to do.

The intuition that migrants convey trade-relevant information on differentiated goods that are not already captured by the price system bears nicely out in column (C1). The network variable in this column is different from the one used by R&T. We are studying the network effect of China-born residents living overseas in 2000. In terms of economic

[26]Trade creation effects and tariff equivalents correspond to columns (A3)-(A5), (B3)-(B5) and (C3)-(C5) of Table 3.7 in Appendix B. Notice that the results in Table 3.7 confirms the pattern that OLS without fixed effects typically overestimates the size of the network effect. Interestingly, this problem is particularly severe in the case of differentiated goods where the mere inclusion of these effects cuts the estimate by at least the factor 5 (and makes it disappear in the year 1990); compare columns (A1) and (A2) of Table 3.7 in Appendix B. Using Poisson has little quantitative effect on the obtained estimates but can have a strong effect on the precision. Similar observations can be made regarding the years 1990 and 2000.

[27]It follows from Table 3.7 in Appendix B, columns (A3) and (B3), that also the points estimates are largest for exchange-traded goods.

Table 3.2: The Chinese network in different commodity groups

	1980			1990			2000		
	(A1)	(A2)	(A3)	(B1)	(B2)	(B3)	(C1)	(C2)	(C3)
	FE-PML	FE-OLS	FE-PML	FE-PML	FE-OLS	FE-PML	FE-PML	FE-OLS	FE-PML
Exchange-traded goods									
Trade Creation (%)									
CHIN	0.935***			0.763***			0.0860		
CHIN*(1-DIR)		0.791***	0.939**		0.682***	0.835**		2.893	3.876
CHIN*DIR		1.076***	0.934***		1.025***	0.753***		0.109***	0.103
Tariff Equivalent (%)									
CHIN	0.0491***			0.0401***			0.00453		
CHIN*(1-DIR)		0.0416***	0.0493**		0.0358***	0.0438**		0.150	0.200
CHIN*DIR		0.0565***	0.0490***		0.0538***	0.0538***		0.00573***	0.00544
Reference-priced goods									
Trade Creation (%)									
CHIN	0.390***			0.326**			0.0500		
CHIN*(1-DIR)		0.931***	0.127		0.584***	-0.201		4.182***	1.575
CHIN*DIR		0.605**	0.427**		0.503**	0.416***		0.0654***	0.0579
Tariff Equivalent (%)									
CHIN	0.0278***			0.0233**			0.00357		
CHIN*(1-DIR)		0.0663***	0.0091		0.0416***	0.0144		0.293***	0.112
CHIN*DIR		0.0432**	0.0305**		0.0359**	0.0297***		0.00467***	0.00413
Differentiated goods									
Trade Creation (%)									
CHIN	0.265			0.590***			0.124*		
CHIN*(1-DIR)		0.314*	0.0330		0.260**	0.104		1.626	1.137
CHIN*DIR		0.156	0.340**		0.282	0.656***		0.0676***	0.128*
Tariff Equivalent (%)									
CHIN	0.0523			0.147***			0.0310*		
CHIN*(1-DIR)		0.0784*	0.00827		0.0649**	0.0260		0.403	0.283
CHIN*DIR		0.0390	0.0850**		0.0706	0.164***		0.0169***	0.0319*

N=2114, N=2127, N=2377 in 1980 for exchange-traded goods, referenced-priced goods and differentiated goods; N=2372, N=2377, N=2533 in 1990 and N=2741, N=2914, N=3025 in 2000, respectively. Each column represents regression results for three different dependent variables: exchange-traded goods, reference-priced goods, and differentiated goods. Only trade creation effects and tariff equivalents are shown; see footnotes 21 and 22 for computational details. Elasticity of substitution is 20, 15, and 5 for exchange-traded goods, referenced-priced goods and differentiated good, respectively. All regressions include the full list of covariates as reported in Table 1, and country dummies (all not shown). Observation clustered by (undirectional) country-pair. *, **, ***, indicate significance of the coefficient at the 1%, 5% and 10% level, respectively.

magnitudes, trade creation and tariff equivalents for the group of differentiated goods are comparable to the ones obtained for aggregate trade.

Finally, we distinguish again between direct and indirect network effects, see columns (A2), (A3), (B2), (B3), (C2), and (C3). Across all categories of goods, the FE-OLS tend to yield more statistically significant results than Poisson. Also, the trade-creation effects and tariff equivalents are often smallest for differentiated goods and largest for exchange-traded ones, with the exception of the estimates for the year 2000 (with a more narrow definition of the network). In the latter case, we find convincing evidence only for the direct effect, but not for the indirect one. Moreover, in all cases, the trade creation effects are small and never exceed 1%.

Table 3.3: Strong versus weak network links: Ethnic Chinese and aggregate trade

	1980			1990			2000		
	(A1)	(A2)	(A3)	(B1)	(B2)	(B3)	(C1)	(C2)	(C3)
	FE-PML	FE-OLS	FE-PML	FE-PML	FE-OLS	FE-PML	FE-PML	FE-OLS	FE-PML
Trade Creation (%)									
CHIN*(1-L)	0.817			1.334***			-0.148		
CHIN*L	14.68***			28.79***			3.190		
CHIN*(1-L)*(1-DIR)		1.057	1.340**		1.310**	1.300**		4.455**	0.677
CHIN*(1-L)*DIR		-1.890	-0.528		0.0146	1.199*		-1.055***	-0.136
CHIN*L*(1-DIR)		14.84***	12.84**		14.59	12.71**		274.2***	70.63*
CHIN*L*DIR		6.975	14.82***		18.17**	31.25***		1.947*	3.727*
Tariff Equivalent (%)									
CHIN*(1-L)	0.116			0.189***			0.0212		
CHIN*L	2.064***			3.816***			0.451		
CHIN*(1-L)*(1-DIR)		0.150	0.190**		0.186**	0.185**		0.623**	0.0964
CHIN*(1-L)*DIR		-0.273	-0.0756		0.00209	0.170*		-0.151***	-0.0194
CHIN*L*(1-DIR)		2.085***	1.834**		2.054	1.804**		18.95***	7.672*
CHIN*L*DIR		1.016	2.083***		2.518**	4.102***		0.277*	0.525*

N=2520 in 1980, N=2975 in 1990, N=3253 in 2000. Only trade creation effects and tariff equivalents are shown; see footnotes 21 and 22 for computational details. Elasticity of substitution is eight. All regressions include the full list of covariates as reported in Table 1, and country dummies (all not shown). Observations clustered by (undirectional) country-pair. *, **, ***, indicate significance at the 1%, 5% and 10% level, respectively. Trade creation and ad valorem tariff equivalent at the respective sample means.

3.4.3 Strong versus weak network links: Ethnic Chinese and aggregate trade

Next, Table 3.3 replicates the key findings of R&T for aggregate trade; results for different categories follow in Table 3.4. To do so, we distinguish between strong and weak network links. Strong link are defined as those for which in both trading countries the share of ethnic Chinese exceeds 1% of the population. Weak links are made up by the complementary set. We define by L a dummy that takes the value of 1 in the former case and zero in the latter. We may further distinguish between direct and indirect effects as in Tables 3.1 and 3.2.[28]

In Table 3.3, we augment the R&T standard specification by country specific fixed-effects. In this theory-consistent estimation, strong network links increase trade in 1990 by at most 29% with a tariff equivalent of at most 3.8%; for the years of 1980 and 2000, the effects are smaller or non-existent. Weak networks perform worse.

Compared to the R&T results, we find much smaller network effects when estimated in a theory-consistent framework. However, with aggregate bilateral trade as the dependent variable, there is evidence for a substantial and significant effect when focusing on strong

[28]The estimated coefficients are presented in columns (A3)-(A5), (B3)-(B5) and (C3)-(C5) of Table 3.8 in Appendix B.

links and the broadly defined ethnic network. In 2000, where the more narrow criterion of China-born is used to constitute the foreign Chinese network, we do not find any evidence for a Chinese network anymore, regardless of the intensity of links. This is striking; compare to column (C3) in Table 3.1 where we have found a (marginally) significant average network effect. The reason for this apparent inconsistency may lie in the lack of linearity in the network effect, so that the effects of weak and strong links estimated in Table 3.3 do not average up to the total effect found in Table 3.1.

Table 3.3 further decomposes the network effects into direct and indirect ones. Across all specifications, we find positive effects for the strong network and for both direct and indirect links. Considering the non-linear estimation strategy FE-PML, we find large trade creation effects and associated tariff equivalents for direct and strong links, ranging from 13 to 71% and from 1.8 to 7.7%, respectively. Interestingly, the more narrowly defined migrant network yields stronger effects than the broader ethnic network.

3.4.4 Strong versus weak network links: Ethnic Chinese in different commodity groups

The final step, presented in Table 3.4, looks separately at different categories of goods, but otherwise replicates Table 3.3.[29]

Neither the ranking of estimates across categories of goods nor their absolute magnitudes are robust to the inclusion of country-fixed effects, neither in 1980 or in 1990. Drawing on the preferred estimates (FE-PML), we find that increasing the size of the network from zero to the sample mean for strong links yields trade creation of about 14% for differentiated goods and a tariff equivalent of 3.4%. Trade creation is larger for exchange-traded goods, but the tariff equivalent is lower due to a higher assumed elasticity of substitution (20) for the latter category of goods.

We may now summarize the main results obtained from a theory-consistent view on the

[29]The estimated coefficients are presented in columns (A3)-(A5), (B3)-(B5) and (C3)-(C5) of Table 3.9 in Appendix B. Notice that the results presented in columns (A1), (B1) of Table 3.9 in Appendix B are comparable with R&T. In the OLS specification, we find the intuitive ranking of coefficients across differentiated, exchange-traded, and reference-priced goods. For differentiated goods, we find the headline trade creation of almost 65% that R&T report in the abstract of their paper. The associated tariff equivalent is higher (13%) than the one computed by Anderson van Wincoop (2004) using the results and data of R&T because we are using a lower elasticity of substitution (5 instead of 8).

Table 3.4: Strong versus weak network links: Ethnic Chinese in different commodity groups

	1980			1990			2000		
	(A1)	(A2)	(A3)	(B1)	(B2)	(B3)	(C1)	(C2)	(C3)
	FE-PML	FE-OLS	FE-PML	FE-PML	FE-OLS	FE-PML	FE-PML	FE-OLS	FE-PML
Exchange-traded goods									
Trade Creation (%)									
CHIN*(1-L)	-0.359			-0.424			-1.098**		
CHIN*L	28.77***			27.45***			1.044		
CHIN*(1-L)*(1-DIR)		0.951	1.200		0.631	-0.639		3.581	-0.482
CHIN*(1-L)*DIR		-0.137	-3.538**		0.834	-0.125		-0.435	-1.060*
CHIN*L*(1-DIR)		28.64***	39.64***		28.05***	29.35**		200.9	181.1
CHIN*L*DIR		37.30***	26.72**		43.32***	27.21***		3.693**	1.922
Tariff Equivalent (%)									
CHIN*(1-L)	-0.0189			-0.0224			-0.0581		
CHIN*L	1.405***			1.352			0.0549		
CHIN*(1-L)*(1-DIR)		0.0498	0.0628		0.0331	-0.0337		0.185	-0.0254
CHIN*(1-L)*DIR		-0.00720	-0.190**		0.0437	-0.00658		-0.0229	-0.0561*
CHIN*L*(1-DIR)		1.399***	1.855***		1.378***	1.434**		5.829	5.469
CHIN*L*DIR		1.761***	1.316**		2.005***	1.341***		0.192**	0.101
Reference-priced goods									
Trade Creation (%)									
CHIN*(1-L)	0.329			1.031**			0.717***		
CHIN*L	12.92**			17.01***			2.214*		
CHIN*(1-L)*(1-DIR)		2.982***	0.635		1.652***	1.123*		6.455***	3.820**
CHIN*(1-L)*DIR		0.655	-0.130		-0.263	0.472		-0.630*	0.692**
CHIN*L*(1-DIR)		39.66***	6.268		28.69***	-0.660		433.8***	170.7**
CHIN*L*DIR		23.11**	13.92**		21.41***	19.69***		2.785***	4.158***
Tariff Equivalent (%)									
CHIN*(1-L)	0.0235			0.0733**			0.0511***		
CHIN*L	0.915**			1.189***			0.157*		
CHIN*(1-L)*(1-DIR)		0.210***	0.0452		0.117***	0.0798*		0.447***	0.268**
CHIN*(1-L)*DIR		0.0467	-0.00928		-0.0188	0.0336		-0.0451*	0.0493**
CHIN*L*(1-DIR)		2.517***	0.458		1.909***	-0.0501		12.03***	7.149**
CHIN*L*DIR		1.567**	0.982**		1.469***	1.360***		0.197***	0.292***
Differentiated goods									
Trade Creation (%)									
CHIN*(1-L)	1.180*			1.547**			-0.297		
CHIN*L	13.87**			33.21***			2.687		
CHIN*(1-L)*(1-DIR)		0.934*	1.435*		1.409***	1.323*		3.876**	0.540
CHIN*(1-L)*DIR		-0.405	-0.0711		-0.530	1.780		-0.190	-0.366
CHIN*L*(1-DIR)		13.23	7.675		14.27	11.61		114.9**	54.05
CHIN*L*DIR		5.634	14.81**		12.01	36.29***		2.787***	3.916
Tariff Equivalent (%)									
CHIN*(1-L)	0.293*			0.384**			-0.0743		
CHIN*L	3.428**			7.599***			0.666		
CHIN*(1-L)*(1-DIR)		0.232*	0.356*		0.350***	0.329*		0.951**	0.135
CHIN*(1-L)*DIR		-0.102	-0.0178		-0.133	0.441		-0.0474	-0.0916
CHIN*L*(1-DIR)		3.280	1.951		3.535	2.910		19.23**	10.86
CHIN*L*DIR		1.446	3.645**		3.006	8.203***		0.691***	0.965

N=2114, N=2127, N=2377 in 1980 for exchange-traded goods, referenced-priced goods, and differentiated goods, N=2372, N=2377, N=2533 in 1990, and N=2741, N=2914, N=3025 in 2000, respectively. Each column represents regression results for three different dependent variables: exchange-traded goods, reference-priced goods, and differentiated goods. Only trade creation effects and tariff equivalents are shown; see footnotes 21 and 22 for computational details. Elasticity of substitution is 20, 15 and 5 for exchange-traded, reference-priced, and differentiated goods, respectively. All regressions include the full list of covariates as reported in Table 1, and country dummies (all not shown). Observations clustered by (undirectional) country-pair. *, **, *** indicate significance of the coefficient at the 1%, 5%, and 10% level, respectively.

trade flow implications of the Chinese ethnic network.

1. **Controlling for multilateral resistance is important.** Without doing so, the quantitative importance of the Chinese ethnic network is overestimated, at least by a factor of two. The omitted variable bias is therefore positive, which signals a positive correlation between the degree of multilateral remoteness of both the importer and the exporter and the size of the Chinese network. Besides controlling for the unobserved resistance terms, our fixed-effects estimation also deals with other country-specific and time-invariant determinants of bilateral trade that may correlate with the size of the network. The overall stance of policies toward the rest of the world (e.g., overall trade policy, overall restrictions to migration, etc.) is such a candidate determinant.

2. Poisson estimation (PML) is immune to misspecification of the error term in the empirical form of the gravity equation. It turns out that point estimators of the network coefficients are usually not strongly affected by misspecification bias. However, in several cases the PML affects the estimated standard errors. Usually, **PML makes results more plausible**; however, it also makes it more difficult to find robust network effects.

3. **Direct network links amount to almost all the trade creation** due to ethnic networks. Indirect links are rarely statistically and economically significant. One way to interpret this result is that the preference channel of ethnic networks dominates the trade cost channel. There is also evidence in favor of threshold effects in the sense that network links need to be strong enough to be visible in the data and to matter economically.

4. We do not find overwhelming empirical support for an intuitive ranking of estimated network effects across different categories of goods. Put differently, while we find evidence for a Chinese network effect in aggregate data and for exchange-traded goods, we do not find it for differentiated goods, where the effect is supposed to be strongest. This sheds **doubts on the overall usefulness of R&Ts identification strategy** which distinguishes between the contract enforcement and the information channel of

ethnic networks.

3.5 Other migrant networks

R&T have studied the quantitative implications of the Chinese ethnic network in a traditional gravity framework. We have qualified the picture using more recent econometric techniques. One of the underlying assumptions of this work is that the Chinese network is the most influential amongst the large number of potential other ethnic (or migrant) networks. In this section, we look at a large number of potential networks and, using the same econometric setup than for the Chinese network, test for their existence.

In particular, for any network k, we compute the tariff equivalent of increasing the size of the network (the product of population shares $s_{ik}s_{jk}$ of migrants in i and j coming from country k. We focus on aggregate trade and on the total effect (without differentiating between strong and weak and between direct and indirect links). For each network k, we run a separate regression. Since we have information about the location of individuals born in country k only for the year of 2000, all regressions refer to this year. Detailed results are found in Appendix C.

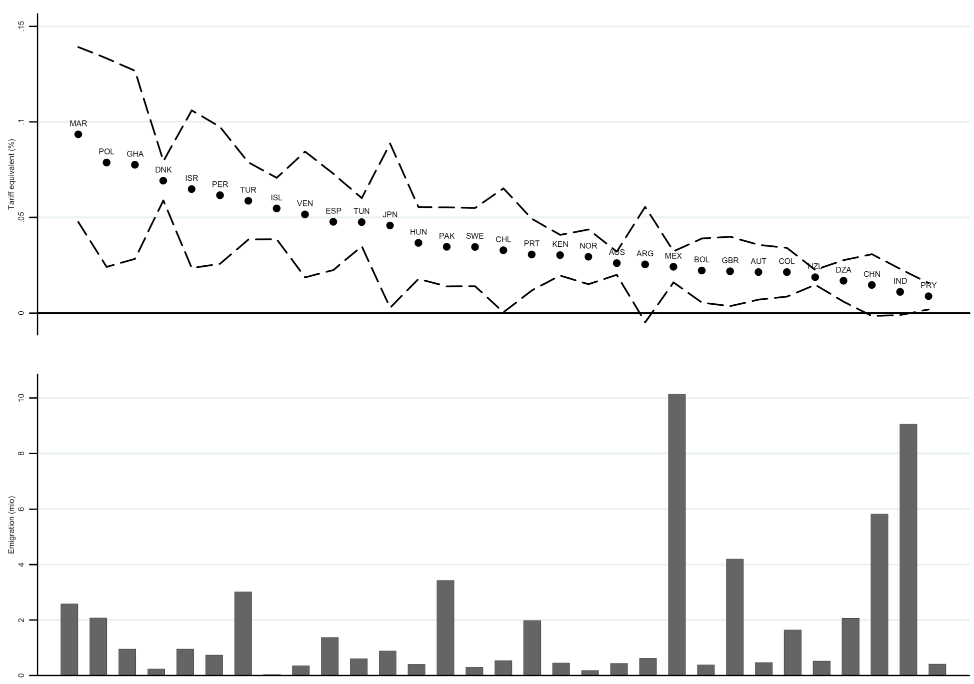


Figure 3.2: Tariff equivalents of different migrant networks (aggregate trade, total effects).

The upper part of Figure 3.2 represents the point estimates obtained for each network from separate regressions as dark circles. It also plots the 1.96 standard deviations band around those coefficients as dashed lines. All estimates shown are statistically significant at least at the 1% level. The figure shows that the Chinese network is not at all the most important one in terms of the trade cost reduction that it entails. The lower part of the figure, which records the total sizes of emigrant networks in million individuals, shows that the Chinese network is also not the largest one in terms of the emigrant population.

The most powerful network seems to be that of Moroccans, of whom about 2 million live abroad. The associated tariff equivalent is close to 0.1%, which is, of course, still extremely small compared to real-life tariffs, or to other estimated trade barriers (compare, e.g., to the border effect identified in Anderson and van Wincoop, 2003). The second and the third most powerful networks are those of the Polish and the Ghanaese, respectively. The largest emigrant stock in the world is the Mexican one, with almost 10 million individuals. That network seems to be relevant for trade creation, albeit at a tariff equivalent inferior to 0.05%. The second largest sending country, India, is associated to an even weaker network, with a measurable yet quantitatively negligible network effect.

3.6 Conclusions

In this paper we have revisited the important work by Rauch and Trindade (R&T, 2002) on the trade-enhancing role of Chinese ethnic networks. Those authors have found that for countries with ethnic Chinese populations shares at the levels prevailing in Southeast Asia, the smallest estimated average increase in bilateral trade in differentiated products attributable to ethnic Chinese networks is nearly 60%. This estimate is obtained by the authors using a traditional gravity model. Recent advances by Anderson and van Wincoop (2003) and Santos Silva and Tenreyro (2006) allow to estimate the network effect in a more theory-consistent and robust way.

Using the econometric techniques proposed in the modern literature, we confirm the existence of a Chinese network effect. However, in terms of magnitudes, the trade creation associated to the network is at most half as big as the one computed by R&T. Moreover, we fail to find the intuitive size ranking of network coefficients across differentiated, reference-priced, and exchange-traded categories of goods. This is not overly surprising since the theory-based gravity model signals that the estimated coefficients confound the elasticity of substitution with the trade-cost elasticity of networks, so that comparing across categories of goods is not an ideal identification strategy. Focusing on indirect network links (i.e., links that relate two trading partners other than China) in order to mitigate endogeneity concerns and to reduce the role of preferences as compared to information, we find that the average network effect is very small (and, indeed, often indistinguishable from zero).

We also investigate other than the Chinese ethnic network. To do so, we use recent data on bilateral stocks of foreign-born individuals provided by the World Bank for the year of 2000 and a total of about 200 countries. Using this data, which implies a more narrow definition of an ethnic network, we conduct a comprehensive quest for the existence of network effects in trade data. Focusing on average effects, we document the existence of a large number of networks. Judging by the obtained size of coefficients and the size of the involved emigrant population, the most relevant are the Moroccan, the Polish, the Turkish, the Pakistan, the Mexican, the British, the Chinese and the Indian networks. However, in all of these cases, the amount of trade creation due to these networks is very small.

3.7 Appendix

3.7.1 List of countries and summary statistics

Table 3.5: List of countries included in the regressions

Country	ISO-Code	Reporter (1990, 2000)	Country	ISO-Code	Reporter (1990, 2000)
Algeria	DZA	YES	Kuwait	KWT	YES
Argentina	ARG	YES	Libyan Arab Jamahiriya	LBY	YES
Australia	AUS	YES	Malaysia	MYS	YES
Austria	AUT	YES	Mexico	MEX	YES
Belgium-Luxembourg	BEL	YES	Morocco	MAR	YES
Bolivia	BOL	NO	Netherlands	NLD	YES
Brazil	BRA	YES	New Zealand	NZL	YES
Canada	CAN	YES	Nigeria	NGA	YES
Chile	CHL	YES	Norway	NOR	YES
China	CHN	YES	Pakistan	PAK	YES
Colombia	COL	YES	Paraguay	PRY	NO
Denmark	DNK	YES	Peru	PER	YES
Ecuador	ECU	YES	Philippines	PHL	YES
Egypt	EGY	NO	Poland	POL	YES
Ethiopia	ETH	NO	Portugal	PRT	YES
Finland	FIN	YES	Saudi Arabia	SAU	YES
France	FRA	YES	Singapore	SGP	YES
Germany	DEU	YES	South Africa	ZAF	YES
Ghana	GHA	NO	Spain	ESP	YES
Greece	GRC	YES	Sudan	SDN	NO
Hong Kong	HKG	YES	Sweden	SWE	YES
Hungary	HUN	YES	Switzerland	CHE	YES
Iceland	ISL	NO	Taiwan	TWN	NO
India	IND	YES	Thailand	THA	YES
Indonesia	IDN	YES	Tunisia	TUN	YES
Iran, Islamic Republic of	IRN	YES	Turkey	TUR	YES
Ireland	IRL	YES	United Kingdom	GBR	YES
Israel	ISR	YES	United States of America	USA	YES
Italy	ITA	YES	Uruguay	URY	NO
Japan	JPN	YES	Venezuela	VEN	YES
Kenya	KEN	NO	Yugoslavia	YUG	YES
Korea, Republic of	KOR	YES			

Note: In 1990 and 2000, bilateral tade flows are only available if at least one of the trading partners is a trade data reporting country. The restriction comes from the NBER-UN World Trade Data.

Table 3.6: Summary statistics

	1980		1990		2000	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Aggregate trade						
Trade	10.55008	2.83535	10.97026	2.55616	11.44128	2.75705
CHIN	0.00313	0.04150	0.00287	0.03893	0.00031	0.00823
CHIN*(1-L)	0.00010	0.00051	0.00014	0.00082	0.00002	0.00014
CHIN*L	0.00303	0.04150	0.00273	0.03893	0.00028	0.00823
N	2520		2795		3253	
Exhange-traded goods						
Trade	9.17629	2.88275	9.66701	2.36043	9.82783	2.51934
CHIN	0.00370	0.04528	0.00333	0.04218	0.00036	0.00896
CHIN*(1-L)	0.00011	0.00055	0.00015	0.00088	0.00002	0.00015
CHIN*L	0.00360	0.04529	0.00318	0.04218	0.00034	0.00896
N	2114		2378		2745	
Reference-priced goods						
Trade	8.84966	2.60053	9.70924	2.33574	10.19298	2.49671
CHIN	0.00369	0.04515	0.00324	0.04154	0.00034	0.00870
CHIN*(1-L)	0.00010	0.00053	0.00015	0.00087	0.00002	0.00015
CHIN*L	0.00359	0.04516	0.00309	0.04154	0.00032	0.00869
N	2127		2453		2916	
Differentiated goods						
Trade	9.24913	2.94396	9.97873	2.70043	10.59338	2.85541
CHIN	0.00330	0.04272	0.00315	0.04088	0.00033	0.00854
CHIN*(1-L)	0.00010	0.00052	0.00015	0.00085	0.00002	0.00014
CHIN*L	0.00320	0.04272	0.00300	0.04088	0.00031	0.00854
N	2377		2533		3025	

3.7.2 Details to Chinese networks (Tables 3.2 to 3.4)

Table 3.7: The Chinese network in different commodity groups (Details to Table 3.2)

Dependent variable: Trade by commodity group															
			1980					**1990**					**2000**		
	(A1)	(A2)	(A3)	(A4)	(A5)	(B1)	(B2)	(B3)	(B4)	(B5)	(C1)	(C2)	(C3)	(C4)	(C5)
	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML
Exchange-traded goods															
CHIN	3.743***	2.635***	2.513***			3.607***	2.474***	2.277***			1.880*	2.272**	2.388		
	(6.86)	(4.27)	(3.23)			(4.47)	(5.17)	(3.91)			(1.85)	(2.06)	(1.40)		
CHIN*(1-DIR)				2.128***	2.522**				2.034***	2.490**				79.21	105.6
				(3.17)	(2.13)				(3.34)	(2.37)				(0.93)	(1.42)
CHIN*DIR				2.890***	2.510***				3.052***	2.244***				3.023***	2.870
				(3.88)	(2.92)				(6.24)	(3.74)				(2.61)	(1.62)
Trade creation (%)															
CHIN	1.396	0.981	0.935			1.212	0.830	0.763			0.0677	0.0819	0.0860		
CHIN*(1-DIR)				0.791	0.939				0.682	0.835				2.893	3.876
CHIN*DIR				1.076	0.934				1.025	0.753				0.109	0.103
Tariff equivalent (%)															
CHIN	0.0731	0.0515	0.0491			0.0635	0.0436	0.0401			0.00356	0.00431	0.00453		
CHIN*(1-DIR)				0.0416	0.0493				0.0358	0.0438				0.150	0.200
CHIN*DIR				0.0565	0.0490				0.0538	0.0395				0.00573	0.00544
Reference-priced goods															
CHIN	3.318***	1.930***	1.054**			3.806***	1.530***	1.003**			4.863***	1.513**	1.466		
	(9.78)	(3.58)	(2.21)			(8.65)	(3.62)	(2.32)			(3.48)	(2.03)	(1.25)		
CHIN*(1-DIR)				2.509***	0.344				1.793***	-0.619				120.1***	45.80
				(4.09)	(0.57)				(3.58)	(-0.71)				(2.85)	(1.01)
CHIN*DIR				1.635**	1.153**				1.544**	1.278***				1.916***	1.696
				(2.13)	(2.39)				(2.57)	(3.36)				(2.64)	(1.37)
Trade creation (%)															
CHIN	1.232	0.715	0.390			1.243	0.498	0.326			0.166	0.0516	0.0500		
CHIN*(1-DIR)				0.931	0.127				0.584	-0.201				4.182	1.575
CHIN*DIR				0.605	0.427				0.503	0.416				0.0654	0.0579
Tariff equivalent (%)															
CHIN	0.0876	0.0510	0.0278			0.0884	0.0355	0.0233			0.0119	0.00369	0.00357		
CHIN*(1-DIR)				0.0663	0.00910				0.0416	0.0144				0.293	0.112
CHIN*DIR				0.0432	0.0305				0.0359	0.0297				0.00467	0.00413
Differentiated goods															
CHIN	4.634***	0.633*	0.802			5.078***	0.872	1.867***			9.186***	2.017***	3.770*		
	(9.69)	(1.84)	(1.63)			(7.73)	(1.23)	(4.03)			(5.46)	(2.61)	(1.80)		
CHIN*(1-DIR)				0.949*	0.100				0.824**	0.330				49.11	34.41
				(1.82)	(0.17)				(2.38)	(0.58)				(1.44)	(1.52)
CHIN*DIR				0.472	1.029**				0.896	2.077***				2.057***	3.890*
				(1.02)	(1.99)				(0.88)	(4.39)				(2.61)	(1.85)
Trade creation (%)															
CHIN	1.540	0.209	0.265			1.611	0.275	0.590			0.302	0.0663	0.124		
CHIN*(1-DIR)				0.314	0.0330				0.260	0.104				1.626	1.137
CHIN*DIR				0.156	0.340				0.282	0.656				0.0676	0.128
Tariff equivalent (%)															
CHIN	0.383	0.0523	0.0662			0.400	0.0687	0.147			0.0754	0.0166	0.0310		
CHIN*(1-DIR)				0.0784	0.00827				0.0649	0.0260				0.403	0.283
CHIN*DIR				0.0390	0.0850				0.0706	0.164				0.0169	0.0319

N=2114, N=2127, N=2377 in 1980 for exchange-traded goods, referenced-priced goods, and differentiated goods, N=2372, N=2377, N=2533 in 1990, and N=2741, N=2914, N=3025 in 2000, respectively. All regressions include the full list of covariates as shown in Table 1, and a constant (all omitted). FE-OLS and FE-PML include country dummies. Observations clustered by (undirectional) country-pair. Robust t statistics in parenthesis. *, **, *** indicate significance at the 1%, 5%, and 10% level, respectively. Trade creation (%) and ad valorem tariff equivalents (%) evaluated at the respective sample means. Elasticity of substitution is twenty, fiveteen, and five for exchange-traded, reference-priced, and differentiated goods, respectively.

Table 3.8: Strong versus weak network link in aggregate trade (Details to Table 3.3)

Dependent variable: Aggregate trade			1980					1990					2000		
	(A1)	(A2)	(A3)	(A4)	(A5)	(B1)	(B2)	(B3)	(B4)	(B5)	(C1)	(C2)	(C3)	(C4)	(C5)
	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML
CHIN*(1-L)	442.9***	4.793	81.16			238.3***	61.22**	94.25***			656.0***	-287.1**	-64.22		
	(5.94)	(0.08)	(1.45)			(4.74)	(2.46)	(3.54)			(4.18)	(-2.06)	(-0.37)		
CHIN*L	4.625***	0.862**	1.292***			4.593***	1.448***	2.354***			5.880***	0.701	3.059		
	(7.28)	(2.30)	(2.87)			(8.94)	(2.96)	(5.94)			(4.39)	(0.77)	(1.47)		
CHIN*(1-L)*(1-DIR)				104.8	132.7**				92.57**	91.90***				1884.1**	291.8
				(1.55)	(2.02)				(2.42)	(3.22)				(2.46)	(0.51)
CHIN*(1-L)*DIR				-190.3	-52.79				1.039	84.77*				-458.4***	-58.79
				(-1.45)	(-0.54)				(0.02)	(1.67)				(-3.37)	(-0.33)
CHIN*L*(1-DIR)				1.305***	1.148**				1.267	1.113**				128.5***	52.05*
				(2.75)	(2.41)				(1.22)	(2.30)				(3.09)	(1.65)
CHIN*L*DIR				0.636	1.304***				1.553**	2.530***				1.879*	3.564*
				(1.39)	(2.71)				(2.08)	(6.72)				(1.94)	(1.71)
Trade creation (%)															
CHIN*(1-L)	4.540	0.0481	0.817			3.406	0.864	1.334			1.529	-0.662	-0.148		
CHIN*L	63.28	9.565	14.68			63.85	16.84	28.79			6.221	0.722	3.190		
CHIN*(1-L)*(1-DIR)				1.057	1.340				1.310	1.300				4.455	0.677
CHIN*(1-L)*DIR				-1.890	-0.528				0.0146	1.199				-1.055	-0.136
CHIN*L*(1-DIR)				14.84	12.94				14.59	12.71				274.2	70.63
CHIN*L*DIR				6.975	14.82				18.17	31.25				1.947	3.727
Tariff equivalent (%)															
CHIN*(1-L)	0.634	0.00687	0.116			0.479	0.123	0.189			0.217	-0.0949	0.0212		
CHIN*L	7.389	1.377	2.064			7.447	2.347	3.816			0.867	0.103	0.451		
CHIN*(1-L)*(1-DIR)				0.150	0.190				0.186	0.185				0.623	0.0964
CHIN*(1-L)*DIR				-0.273	-0.0756				0.00209	0.170				-0.151	-0.0194
CHIN*L*(1-DIR)				2.085	1.834				2.054	1.804				18.95	7.672
CHIN*L*DIR				1.016	2.083				2.518	4.102				0.277	0.525

N=2520 in 1980, N=2795 in 1990, and N=3253 in 2000. All regressions include the full list of covariates as shown in Table 1, and a constant (all omitted). FE-OLS and FE-PML include country dummies. Observations clustered by (undirectional) country-pair. Robust t statistics in parenthesis. *, **, *** indicate significance at the 1%, 5%, and 10% level, respectively. Trade creation (%) and ad valorem tariff equivalents (%) evaluated at the respective sample means. Elasticity of substitution is twenty, fiveteen, and five for exchange-traded, reference-priced, and differentiated goods, respectively.

Table 3.9: Strong versus weak network link in different commodity groups (Details to Table 3.4)

Dependent variable: Trade by commodity group			1980					1990					2000		
	(A1)	(A2)	(A3)	(A4)	(A5)	(B1)	(B2)	(B3)	(B4)	(B5)	(C1)	(C2)	(C3)	(C4)	(C5)
	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML	OLS	FE-OLS	FE-PML	FE-OLS	FE-PML
Exchange-traded goods															
CHIN*(1-L)	271.1***	52.23	-31.87			116.1***	34.46	-27.05			-47.71	-42.80	-432.3**		
	(2.97)	(0.60)	(-0.20)			(2.84)	(1.23)	(-0.56)			(-0.26)	(-0.28)	(-1.96)		
CHIN*L	3.784***	2.754***	2.362***			3.654***	2.587***	2.149***			1.882*	2.180*	0.981		
	(6.88)	(4.16)	(2.58)			(4.49)	(5.09)	(3.45)			(1.84)	(1.83)	(0.54)		
CHIN*(1-L)*(1-DIR)				83.97	105.9				40.03	-40.76				1377.0	-189.2
				(0.83)	(0.59)				(1.08)	(-0.57)				(1.29)	(-0.19)
CHIN*(1-L)*DIR				-12.14	-319.6**				52.82	-7.957				-170.6	-417.1*
				(-0.08)	(-1.96)				(1.54)	(-0.16)				(-1.15)	(-1.90)
CHIN*L*(1-DIR)				2.352***	3.118***				2.191***	2.281**				104.1	97.71
				(3.18)	(2.79)				(3.30)	(2.11)				(1.22)	(1.19)
CHIN*L*DIR				2.961***	2.212**				3.189***	2.132***				3.428**	1.800
				(3.75)	(2.27)				(6.06)	(3.30)				(2.52)	(0.90)
Trade creation (%)															
CHIN*(1-L)	3.103	0.590	-0.359			1.842	0.543	-0.424			-0.122	-0.109	-1.098		
CHIN*L	49.96	34.30	28.77			51.04	33.90	27.45			2.011	2.333	1.044		
CHIN*(1-L)*(1-DIR)				0.951	1.200				0.631	-0.639				3.581	-0.482
CHIN*(1-L)*DIR				-0.137	-3.538				0.834	-0.125				-0.435	-1.060
CHIN*L*(1-DIR)				28.64	39.64				28.05	29.35				200.9	181.1
CHIN*L*DIR				37.30	26.72				43.32	27.21				3.693	1.922
Tariff equivalent (%)															
CHIN*(1-L)	0.161	0.0310	-0.0189			0.0961	0.0285	-0.0224			-0.00642	-0.00576	-0.0581		
CHIN*L	2.251	1.638	1.405			2.298	1.626	1.352			0.105	0.122	0.0549		
CHIN*(1-L)*(1-DIR)				0.0498	0.0628				0.0331	-0.0337				0.185	-0.0254
CHIN*(1-L)*DIR				-0.00720	-0.190				0.0437	-0.00658				-0.0229	-0.0561
CHIN*L*(1-DIR)				1.399	1.855				1.378	1.434				5.829	5.469
CHIN*L*DIR				1.761	1.316				2.005	1.341				0.192	0.101
Reference-priced goods															
CHIN*(1-L)	344.9***	200.9***	30.93			192.7***	55.53**	66.10**			580.1***	-79.00	282.8***		
	(5.67)	(3.72)	(0.57)			(4.42)	(2.09)	(2.14)			(4.21)	(-0.67)	(2.69)		
CHIN*L	3.369***	2.365***	1.146**			3.884***	1.716***	1.367***			4.831***	1.345	2.110*		
	(10.26)	(4.32)	(2.15)			(8.99)	(3.99)	(2.88)			(3.74)	(1.64)	(1.88)		
CHIN*(1-L)*(1-DIR)				276.7***	59.61				105.6***	71.98*				2474.9***	1483.3**
				(3.86)	(0.82)				(2.88)	(1.82)				(4.16)	(2.14)
CHIN*(1-L)*DIR				61.49	-12.23				-16.98	30.32				-249.9*	272.8**
				(0.75)	(-0.16)				(-0.56)	(1.14)				(-1.90)	(2.55)
CHIN*L*(1-DIR)				3.151***	0.573				2.196***	-0.0577				161.4***	95.94**
				(4.45)	(0.78)				(3.97)	(-0.07)				(3.77)	(2.35)
CHIN*L*DIR				1.961**	1.230**				1.689***	1.565***				2.647***	3.925***
				(2.54)	(2.31)				(2.79)	(3.72)				(3.18)	(3.52)
Trade creation (%)															
CHIN*(1-L)	3.731	2.157	0.329			3.035	0.866	1.031			1.477	-0.199	0.717		
CHIN*L	42.92	28.50	12.92			56.22	21.78	17.01			5.142	1.405	2.214		
CHIN*(1-L)*(1-DIR)				2.982	0.635				1.652	1.123				6.455	3.820
CHIN*(1-L)*DIR				0.655	-0.130				-0.263	0.472				-0.630	0.692
CHIN*L*(1-DIR)				39.66	6.268				28.69	-0.660				433.8	170.7
CHIN*L*DIR				23.11	13.92				21.41	19.69				2.785	4.158
Tariff equivalent (%)															
CHIN*(1-L)	0.262	0.152	0.0235			0.214	0.0616	0.0733			0.105	-0.0143	0.0511		
CHIN*L	2.691	1.890	0.915			3.376	1.492	1.189			0.360	0.100	0.157		
CHIN*(1-L)*(1-DIR)				0.210	0.0452				0.117	0.0798				0.447	0.268
CHIN*(1-L)*DIR				0.0467	-0.00928				-0.0188	0.0336				-0.0451	0.0493
CHIN*L*(1-DIR)				2.517	0.458				1.909	-0.0501				12.03	7.149
CHIN*L*DIR				1.567	0.982				1.469	1.360				0.197	0.292
Differentiated goods															
CHIN*(1-L)	448.7***	46.25	113.2*			270.3***	48.67*	101.9**			1252.4***	51.21	-122.1		
	(5.23)	(0.94)	(1.69)			(4.48)	(1.65)	(2.28)			(6.26)	(0.36)	(-0.51)		
CHIN*L	4.669***	0.731**	1.213**			5.180***	1.029	2.490***			9.116***	2.115**	2.526		
	(9.78)	(1.97)	(2.28)			(8.22)	(1.44)	(5.81)			(6.22)	(2.50)	(1.06)		
CHIN*(1-L)*(1-DIR)				89.72*	137.5*				92.89***	87.24*				1561.3**	221.1
				(1.68)	(1.77)				(2.68)	(1.83)				(2.36)	(0.31)
CHIN*(1-L)*DIR				-39.20	-6.860				-35.26	117.1				-77.89	-150.4
				(-0.37)	(-0.09)				(-0.87)	(1.39)				(-0.60)	(-0.60)
CHIN*L*(1-DIR)				1.161**	0.691				1.158***	0.953				72.91**	41.17
				(2.08)	(1.05)				(2.97)	(1.59)				(2.03)	(1.25)
CHIN*L*DIR				0.512	1.290**				0.985	2.688***				2.619***	3.660
				(1.03)	(2.36)				(0.96)	(6.68)				(2.96)	(1.48)
Trade creation (%)															
CHIN*(1-L)	4.760	0.480	1.180			4.156	0.736	1.547			3.097	0.125	-0.297		
CHIN*L	64.86	8.143	13.87			81.60	12.59	33.21			10.04	2.244	2.687		
CHIN*(1-L)*(1-DIR)				0.934	1.435				1.409	1.323				3.876	0.540
CHIN*(1-L)*DIR				-0.405	-0.0711				-0.530	1.780				-0.190	-0.366
CHIN*L*(1-DIR)				13.23	7.675				14.27	11.61				114.9	54.05
CHIN*L*DIR				5.634	14.81				12.01	36.29				2.787	3.916
Tariff equivalent (%)															
CHIN*(1-L)	1.163	0.120	0.293			1.018	0.183	0.384			0.763	0.0312	-0.0743		
CHIN*L	13.19	2.066	3.428			15.81	3.142	7.599			2.404	0.558	0.666		
CHIN*(1-L)*(1-DIR)				0.232	0.356				0.350	0.329				0.951	0.135
CHIN*(1-L)*DIR				-0.102	-0.0178				-0.133	0.441				-0.0474	-0.0916
CHIN*L*(1-DIR)				3.280	1.951				3.535	2.910				19.23	10.86
CHIN*L*DIR				1.446	3.645				3.006	8.203				0.691	0.965

N=2114, N=2127, N=2377 in 1980 for exchange-traded goods, referenced-priced goods, and differentiated goods, N=2372, N=2377, N=2533 in 1990, and N=2741, N=2914, N=3025 in 2000, respectively. All regressions include the full list of covariates as shown in Table 1, and a constant (all omitted). FE-OLS and FE-PML include country dummies. Observations clustered by (undirectional) country-pair. Robust t statistics in parenthesis. *, **, *** indicate significance at the 1%, 5%, and 10% level, respectively. Trade creation (%) and ad valorem tariff equivalents (%) evaluated at the respective sample means. Elasticity of substitution is twenty, fiveteen, and five for exchange-traded, reference-priced, and differentiated goods, respectively.

3.7.3 Details to other migrant networks (Figure 2)

	ARG		AUS		AUT		BEL		BOL		BRA	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	72.71*		140.0***		157.4***		26.60		279.3***		16.66	
	(1.65)		(8.44)		(2.92)		(0.19)		(2.60)		(0.59)	
MIG*(1-DIR)		-20272.3		36489.4**		28891.1*		-116595.0*		188388.8		-3927.4
		(-0.61)		(2.48)		(1.84)		(-1.84)		(1.48)		(-0.06)
MIG*DIR		67.09		153.2***		177.9***		-58.88		304.6***		14.19
		(1.52)		(8.19)		(3.17)		(-0.40)		(2.81)		(0.30)
Trade creation (%)												
MIG	0.178		0.183		0.150		0.0201		0.156		0.0419	
MIG*(1-DIR)		-39.07		60.96		31.66		-58.54		185.9		-9.411
MIG*DIR		0.164		0.200		0.170		-0.0444		0.170		0.0357
Tariff equivalent (%)												
MIG	0.0254		0.0261		0.0214		0.00287		0.0222		0.00599	
MIG*(1-DIR)		-7.079		6.799		3.929		-12.58		15.01		-1.412
MIG*DIR		0.0234		0.0286		0.0242		-0.00635		0.0243		0.00510

	CAN		CHE		CHL		COL		DEU		DNK	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	-537.6***		72.60		191.7**		53.49***		-3.474		362.8***	
	(-6.49)		(0.30)		(1.99)		(3.29)		(-0.22)		(13.23)	
MIG*(1-DIR)		1497948.4		334190.4**		35113.3		7731.8		-277.4		61897.8***
		(0.58)		(2.38)		(0.82)		(0.93)		(-0.31)		(5.67)
MIG*DIR		-382.7		231.0		201.2**		61.01***		-4.590		393.4***
		(-0.06)		(0.90)		(2.07)		(5.48)		(-0.30)		(14.36)
Trade creation (%)												
MIG	-0.614		0.0488		0.230		0.150		-0.0304		0.486	
MIG*(1-DIR)		2.80344e+09		845.1		52.35		24.15		-2.403		128.5
MIG*DIR		-0.437		0.155		0.242		0.171		-0.0402		0.527
Tariff equivalent (%)												
MIG	-0.0879		0.00697		0.0328		0.0214		-0.00435		0.0692	
MIG*(1-DIR)		245.0		32.09		6.015		3.090		-0.347		11.81
MIG*DIR		-0.0626		0.0222		0.0345		0.0244		-0.00575		0.0750

	DZA		ECU		EGY		ESP		ETH		FIN	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	38.49***		92.07		-10.16*		113.3***		21.78		-2.021	
	(3.05)		(0.96)		(-1.70)		(3.72)		(0.55)		(-0.18)	
MIG*(1-DIR)		-5395.6**		-173799.2***		-4600.0		-13007.8**		63118.3		39988.9***
		(-2.24)		(-3.23)		(-1.45)		(-2.10)		(1.57)		(4.95)
MIG*DIR		36.08***		51.97		-10.80*		92.82***		23.73		10.15
		(2.86)		(0.55)		(-1.78)		(2.86)		(0.59)		(0.87)
Trade creation (%)												
MIG	0.118		0.0640		-0.0285		0.334		0.0197		-0.00348	
MIG*(1-DIR)		-15.24		-70.11		-12.11		-31.84		76.83		99.16
MIG*DIR		0.111		0.0361		-0.0303		0.274		0.0214		0.0175
Tariff equivalent (%)												
MIG	0.0169		0.00914		-0.00407		0.0477		0.00281		-0.000497	
MIG*(1-DIR)		-2.362		-17.25		-1.844		-5.475		8.143		9.842
MIG*DIR		0.0158		0.00516		-0.00433		0.0391		0.00306		0.00250

	FRA		GBR		GHA		GRC		HKG		HUN	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	-7.372		9.719**		1614.1***		86.08		7.188		253.7***	
	(-0.53)		(2.35)		(3.09)		(0.71)		(0.04)		(3.82)	
MIG*(1-DIR)		1071.8		233.5*		-284870.3		-75510.0**		-46917.5		97605.1***
		(0.63)		(1.68)		(-0.89)		(-2.06)		(-0.94)		(2.60)
MIG*DIR		-5.109		10.41**		1565.6***		53.67		-29.45		296.3***
		(-0.36)		(2.50)		(2.99)		(0.45)		(-0.17)		(4.71)
Trade creation (%)												
MIG	-0.0299		0.153		0.544		0.121		0.00557		0.257	
MIG*(1-DIR)		4.441		3.737		-61.63		-65.32		-30.46		168.4
MIG*DIR		-0.0207		0.164		0.528		0.0753		-0.0228		0.300
Tariff equivalent (%)												
MIG	-0.00427		0.0218		0.0775		0.0172		0.000795		0.0367	
MIG*(1-DIR)		0.621		0.524		-13.68		-15.13		-5.190		14.10
MIG*DIR		-0.00296		0.0234		0.0752		0.0108		-0.00326		0.0428

	IDN		IND		IRL		IRN		ISL		ISR	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	-20.52*		9.146*		51.68		-558.8		2452.4***		1949.3***	
	(-1.69)		(1.80)		(1.62)		(-1.44)		(6.69)		(3.08)	
MIG*(1-DIR)		196.8		-353.1		-3490.5		14744.0		2182238.5***		-2102390.7**
		(0.20)		(-1.63)		(-0.27)		(1.24)		(5.90)		(-2.17)
MIG*DIR		-20.22		9.516*		50.07		-556.6		2604.1***		1464.7**
		(-1.63)		(1.88)		(1.59)		(-1.42)		(7.19)		(2.20)
Trade creation (%)												
MIG	-0.103		0.0773		0.0623		-0.986		0.384		0.455	
MIG*(1-DIR)		0.992		-2.940		-4.116		29.89		2917.1		-99.25
MIG*DIR		-0.101		0.0804		0.0603		-0.982		0.407		0.341
Tariff equivalent (%)												
MIG	-0.0147		0.0110		0.00889		-0.142		0.0547		0.0648	
MIG*(1-DIR)		0.141		-0.426		-0.601		3.736		48.67		-69.87
MIG*DIR		-0.0145		0.0115		0.00861		-0.141		0.0581		0.0487

Dependent variable: Aggregate trade. Estimation method: Fixed-effect PML. N=3259 in all regressions. All regressions include the full list of covariates as shown in Table 1, a dummy for common colonizer, and a set of country dummies. Observations clustered by (undirectional) country-pair. Robust t statistics in parenthesis. *, **, *** indicate significance at the 1%, 5%, and 10% level, respectively. Trade creation (%) and ad valorem tariff equivalents (%) evaluated at the respective sample means. Elasticity of substitution is eight.

	ITA		JPN		KEN		KOR		KWT		LBY	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	8.101		316.6**		634.5***		-0.628		-695.0**		-825.0	
	(0.94)		(2.09)		(5.57)		(-0.01)		(-2.57)		(-0.22)	
MIG*(1-DIR)		-2266.8*		2885646.7*		235378.1		12987.8		-795301.7*		1158791.7**
		(-1.77)		(1.81)		(1.05)		(0.35)		(-1.88)		(2.14)
MIG*DIR		-1.026		2880.9*		662.2***		10.87		-758.7***		-810.5
		(-0.10)		(1.95)		(5.65)		(0.11)		(-2.77)		(-0.22)
Trade creation (%)												
MIG	0.0606		0.321		0.212		-0.00108		-0.154		-0.0541	
MIG*(1-DIR)		-15.60		4.81877e+14		119.6		25.09		-82.93		113.9
MIG*DIR		-0.00767		2.959		0.222		0.0187		-0.169		-0.0532
Tariff equivalent (%)												
MIG	0.00866		0.0458		0.0303		-0.000155		-0.0221		-0.00773	
MIG*(1-DIR)		-2.423		417.2		11.24		3.198		-25.26		10.86
MIG*DIR		-0.00110		0.417		0.0316		0.00268		-0.0241		-0.00760

	MAR		MEX		MYS		NGA		NLD		NOR	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	98.86***		41.25***		-0.165		164.5		39.52		247.1***	
	(4.00)		(5.83)		(-0.04)		(0.76)		(0.94)		(4.03)	
MIG*(1-DIR)		1699.5		-222.8		619.5		-19237.2		-21607.3		109563.3***
		(1.45)		(-0.35)		(0.72)		(-0.24)		(-1.14)		(14.53)
MIG*DIR		100.2***		40.95***		0.112		156.8		20.53		276.1***
		(3.99)		(5.69)		(0.03)		(0.72)		(0.48)		(4.52)
Trade creation (%)												
MIG	0.656		0.169		-0.000853		0.169		0.0841		0.206	
MIG*(1-DIR)		11.90		-0.910		3.254		-17.88		-36.84		149.2
MIG*DIR		0.665		0.168		0.000581		0.161		0.0437		0.230
Tariff equivalent (%)												
MIG	0.0934		0.0242		-0.000122		0.0241		0.0120		0.0294	
MIG*(1-DIR)		1.606		-0.131		0.457		-2.815		-6.565		13.04
MIG*DIR		0.0947		0.0240		0.0000829		0.0229		0.00624		0.0329

	NZL		PAK		PER		PHL		POL		PRT	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	106.6***		57.32***		441.6***		38.15		140.4***		71.48***	
	(9.38)		(3.29)		(3.37)		(1.64)		(2.83)		(3.20)	
MIG*(1-DIR)		3681.9		-4013.2		-261755.5*		-4827.5		-7911.7		-4227.9**
		(0.18)		(-1.56)		(-1.75)		(-1.33)		(-1.02)		(-2.12)
MIG*DIR		107.1***		54.25***		330.7**		30.70		130.4***		66.52***
		(8.77)		(3.10)		(2.54)		(1.32)		(2.61)		(2.98)
Trade creation (%)												
MIG	0.131		0.243		0.432		0.192		0.552		0.214	
MIG*(1-DIR)		4.629		-15.61		-92.24		-21.55		-26.69		-11.90
MIG*DIR		0.132		0.230		0.323		0.154		0.513		0.199
Tariff equivalent (%)												
MIG	0.0187		0.0346		0.0616		0.0274		0.0787		0.0306	
MIG*(1-DIR)		0.646		-2.425		-36.51		-3.468		-4.436		-1.810
MIG*DIR		0.0188		0.0328		0.0461		0.0221		0.0731		0.0285

	PRY		SAU		SDN		SGP		SWE		THA	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	104.3**		-4253.8*		53.79		0.237		152.9***		-86.29	
	(2.50)		(-1.85)		(1.15)		(0.00)		(3.31)		(-1.38)	
MIG*(1-DIR)		-8971.8		-16021445.5**		-112495.9*		99205.9**		40961.8***		87286.5
		(-0.11)		(-2.23)		(-1.84)		(2.17)		(4.05)		(1.54)
MIG*DIR		103.2**		-5232.5**		49.38		19.16		179.6***		-56.17
		(2.40)		(-2.23)		(1.08)		(0.21)		(3.78)		(-0.86)
Trade creation (%)												
MIG	0.0616		-0.601		0.0236		0.000125		0.242		-0.120	
MIG*(1-DIR)		-5.160		-100.00		-38.97		68.81		91.11		236.5
MIG*DIR		0.0610		-0.739		0.0217		0.0101		0.284		-0.0781
Tariff equivalent (%)												
MIG	0.00880		-0.0862		0.00337		0.0000179		0.0345		-0.0171	
MIG*(1-DIR)		-0.757		-324.5		-7.054		7.480		9.253		17.33
MIG*DIR		0.00871		-0.106		0.00310		0.00144		0.0406		-0.0112

	TUN		TUR		URY		USA		VEN		ZAF	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
MIG	324.7***		61.83***		-157.5*		-197.8***		749.2***		155.0	
	(7.40)		(5.70)		(-1.70)		(-5.03)		(3.07)		(1.62)	
MIG*(1-DIR)		-73522.3***		-2402.9		-5708.7		-3471.1		15129.9		61571.4***
		(-2.68)		(-1.51)		(-0.02)		(-0.40)		(0.07)		(6.01)
MIG*DIR		307.2***		57.36***		-158.1*		-205.3***		752.9***		187.5*
		(6.91)		(5.17)		(-1.65)		(-4.53)		(3.09)		(1.93)
Trade creation (%)												
MIG	0.333		0.412		-0.0571		-0.740		0.362		0.225	
MIG*(1-DIR)		-52.94		-14.75		-2.049		-12.22		7.564		144.0
MIG*DIR		0.315		0.382		-0.0573		-0.768		0.363		0.272
Tariff equivalent (%)												
MIG	0.0475		0.0587		-0.00816		-0.106		0.0516		0.0321	
MIG*(1-DIR)		-10.77		-2.280		-0.296		-1.862		1.042		12.75
MIG*DIR		0.0450		0.0544		-0.00819		-0.110		0.0518		0.0388

Dependent variable: Aggregate trade. Estimation method: Fixed-effect PML. N=3259 in all regressions. All regressions include the full list of covariates as shown in Table 1, a dummy for common colonizer, and a set of country dummies. Observations clustered by (undirectional) country-pair. Robust t statistics in parenthesis. *, **, *** indicate significance at the 1%, 5%, and 10% level, respectively. Trade creation (%) and ad valorem tariff equivalents (%) evaluated at the respective sample means. Elasticity of substitution is eight.

Chapter 4

The Pro-Trade Effect of the Brain-Drain: Sorting Out Confounding Factors[1]

4.1 Introduction

In the perfect-competition aggregate production function framework emigration triggers a static welfare loss for remaining individuals as the marginal productivity of immobile complementary factors declines. Skill-biased emigration (a brain drain) may add a dynamic loss if the source country's average human capital falls.

Docquier and Marfouk (2006) show that the *total* stock of migrants from poor Southern countries in the rich OECD has grown from about 19 million people in 1990 to 31 million in 2000. Moreover, the average rate of *high-skilled emigration* has been 6.6 percent in 1990 and 7.2 in 2000, with higher numbers for least-developed countries.

Theory papers discuss channels which may mitigate this brain drain. Besides remittances, migration prospects may increase the incentives for higher education, so that average human capital in the non-migrant population may actually rise. Moreover, a diaspora may improve access to foreign markets, thereby encouraging international trade or investment. However, Lucas (2006) concludes that "*the empirical evidence on each of these ... channels remains highly controversial. The most systematic portion of this evidence looks at the links between migration and trade, though difficulties eliminating spurious associations remain*"

[1] This Chapter is based on an article forthcoming in the *Economics Letters*. For the working paper version, see Felbermayr and Jung (2008b). The concept for the paper was developed jointly, empirical analysis and writing were shared equally.

(p. 373).

Spurious association arises due to confounding factors that determine both, the volume of bilateral trade and the bilateral stock of migrants. For example, cultural proximity matters for bilateral trade volumes, but may also affect emigration rates. Similar considerations apply for the ease of geographical mobility. If *unobserved* components of cultural and geographical proximity positively affect migration, OLS estimates would suffer from endogeneity bias and overestimate the true effect of migration on trade.

We include the bilateral stock of migrants into a theory-grounded gravity equation. Recent data on the stock of emigrants from poor sending countries comes from Docquier and Marfouk. The data has a time dimension and distinguishes between three different educational classes. The panel nature of the data allows to account for unobserved heterogeneity by differencing out unobserved country-pair specific characteristics. Baier and Bergstrand (2007) have recently shown the advantages of this approach in a comparable gravity context. Moreover, we can perform a regression-based test for strict exogeneity (Wooldridge, 2002).

We report three major results. First, failing to control for unobserved heterogeneity indeed leads to *overestimation*. Second, there is, nevertheless, a statistically and economically significant *causal* effect of migration on trade. Third, low- and high-skilled migrants strongly boost bilateral trade by comparable quantities while medium-skilled migration does not seem to matter.

So far, empirical gravity studies have typically focused on a single anchor country, see the survey of Wagner *et al.* (2002). Dunlevy (2006) and Bandyopadhyay *et al.* (2008) document a pro-trade effect of migration on the exports of US states. Kugler and Rapoport (2007) analyze how emigration into the US fosters capital formation; Docquier and Lodigiani (2009) extend this exercise to a cross-section of host countries. The two latter papers use the same data than ours; however, we seem to be the first to exploit the *temporal and bilateral* dimensions of the data in a theory-grounded South-North gravity model.

4.2 Econometric specification

We augment the theory-based gravity framework described in Feenstra (2004) with the bilateral stocks of migrants. We strive to explain the volume of trade T_{snt} between a (poor) Southern sending country, s, and a (rich) Northern receiving country, n, at time $t \in \{1990, 2000\}$. We investigate the effect of MIG^k_{snt}, the stock of foreign-born residents from s in n by education k ($k \in \{l, m, h\}$, l: low-skilled, m: medium-skilled, and h: high-skilled).

Our gravity equation is

$$\ln T_{snt} = \sum_{k \in \{l,m,h\}} \beta^k \ln MIG^k_{snt} + \boldsymbol{\gamma} \mathbf{PROX'_{sn}} + \boldsymbol{\delta} \mathbf{POL'_{snt}} + \nu_{st} + \nu_{nt} + \varepsilon_{snt}, \tag{4.1}$$

where the vector $\mathbf{PROX_{sn}}$ collects indicators of cultural and geographical proximity, and $\mathbf{POL_{snt}}$ measures time-variant bilateral trade policy. We include a comprehensive set of country-and-time effects ν_{st} and ν_{nt} to control for all source and destination specific determinants, in particular for multilateral resistance terms.[2]

We impose the error structure $\varepsilon_{snt} = c_{sn} + u_{snt}$, where c_{sn} is a dyad-effect and u_{snt} the usual idiosyncratic error term. In the presence of unobserved confounding factors explanatory variables will be correlated with the error term u_{snt} so that OLS is invalid. Following Baier and Bergstrand, we difference equation (4.1) to eliminate c_{sn}. As suggested by Wooldridge (p. 285), in a two-period framework we can test whether the differenced version of (4.1) satisfies the assumption of strict exogeneity $E(\Delta u_{sn} | \boldsymbol{\Delta} \mathbf{X_{sn}}) = 0$, where $\boldsymbol{\Delta} \mathbf{X_{sn}}$ is the vector of first differences of all explanatory variables. We include the *stocks* of foreign-born residents in the differenced version of equation (4.1) and perform an F-test for joint significance. Failing to reject the null would signal that differencing has indeed solved the endogeneity concern.

4.3 Data and empirical results

We use bilateral data on international migration by education for the years 1990 and 2000 collected by Docquier and Marfouk. The trade data has been assembled and provided by

[2] Baltagi *et al.* (2003) explain the importance of country-and-time interactions in panel gravity equations.

Feenstra *et al.* (2005).[3] We focus on a balanced panel of low-income Southern sending countries and high-income Northern receiving countries.[4] Our sample covers more than 92 percent of total South-North migration.

Geographical (distance, contiguity) and cultural covariates (common language, colonial ties) are taken from the CEPII data base. We include dummies for non-reciprocal preferential trade arrangements (NR_PTA_{snt}), preferential trade arrangements (PTA_{snt}), free trade agreements, and customs unions (FTA_{snt}), and the Euro-zone ($EURO_{snt}$). This data comes from Baier and Bergstrand.

Table 1 presents pooled OLS estimations of equation (4.1). Odd numbered columns present the most parsimonious model; even numbered columns include covariates related to cultural proximity. Columns (1) and (2) disallow for elasticities to vary across educational classes. Columns (3) to (8) estimate the pro-trade effect of single educational groups in isolation, while columns (9) and (10) report the unconstrained version of (4.1).

Across all specifications, the elasticity of trade volumes with respect to distance is close to unity. While non-reciprocal trade agreements seem to matter, preferential trade arrangements and free trade agreements fail to show statistical significance. These are standard results which nicely replicate Baier and Bergstrand.

Concerning the link between migration and trade we find the following: First, there is a strong positive association between the *total* bilateral stock of migrants and bilateral trade. The effect remains when considering migrants at different educational levels, see columns (3), (5), and (7). Second, in column (9), where migration of all skill groups is accounted for, we find that the pro-trade elasticity of high-skilled workers is almost four times bigger than that of low-skilled workers. Surprisingly, conditional on the emigration of other skill classes, medium-skilled individuals seem to reduce bilateral trade volumes. Third, including

[3]The dependent variable is the *geometric* average of trade flows between the two countries; see Baldwin and Taglioni (2006).

[4]A country with per capita GDP above the 80th quantile is classified North and South else. This strategy yields the same classification for 1990 and 2000, except for Greece. There is no data for countries from the former USSR, Yugoslavia, and Czechoslovakia. The obtained sample is similar to that used by Beine *et al.* (2008).
We average the bilateral trade data over the periods 1988-1990 and 1998-2000 to reduce measurement error and increase data availability. This has no importance for our results.
The Feenstra *et al.* data does not distinguish between missing and zero trade flows. Hence, we cannot empirically distinguish between the intensive and the extensive margin of trade.

controls for cultural proximity almost reduces the effects by half; compare odd and even numbered columns. Hence, ignoring cultural proximity as a common determinant of both trade and migration leads to upward biases estimates. However, the unexpected negative effect of medium-skilled migrants remains, see Column (10). While these results go beyond the literature in showing the effect of skill structure in a fairly comprehensive sample of North-South trade relations, they may still suffer from endogeneity bias.

Table 2 presents our preferred specification where confounding factors are differenced out. It also presents the outcome of a regression-based F-test on strict exogeneity. Since all p-values are above 0.1, we cannot reject strict exogeneity in all specifications at conventional levels of significance. Hence, we interpret our estimates as the *causal effect* of migration on trade.

The following results stand out. First, the positive link between migration and trade remains intact for the total stock of migrants as well as for low- and high-skilled migrants, but turns insignificant for medium-skilled migrants; see columns (1) to (4). Second, comparing even numbered columns of Table 1 (which include additional measures of cultural proximity) and results presented in Table 2, we find that OLS *always* overestimates the effect of migration on trade, signaling the presence of endogeneity bias. However, that bias is much smaller when the OLS model includes measures of cultural proximity than when it does not. Third, column (5) shows that the partial effect of medium-skilled migrants on trade is now statistically insignificant compared to the corresponding OLS estimates in column (9)-(10) of Table 1. Here, OLS actually seems to underestimate the true effect. While the results in column (5) suggest that the skill-composition of migration does matter – since medium-skilled migrants do not appear to promote trade – we cannot formally reject the hypothesis that the pro-trade elasticity of low-skilled migrants equals the one of high-skilled.

We conclude with three remarks. First, the pro-trade effect of migration is quantitatively important. A one-percent increase of the bilateral stock of migrants raises bilateral trade by 0.11 percent (column (1), Table 2). Since the mean bilateral migrant population in our sample is 27,000 persons and the mean North-South trade volume is 665 mio dollar in year 2000, our estimate implies that one additional migrant creates about 2,700 dollar in additional trade.[5] Hence, the pro-trade effect of emigration is a powerful driver in overturning welfare losses from emigration. Second, medium-skilled migrants do not foster trade. This may have to do with the low overlap between educational classes and occupational groups: medium-skilled workers may be predominantly employed in the non-tradeable sector. Moreover, the skill-distribution of migrants is often bi-modular, with relatively little mass on medium-skilled workers. Third, there are two interesting avenues for further research. Our empirical strategy provides consistent estimates of the *average* elasticity of migration on trade (see Feenstra), leaving the analysis of potential systematic differences across country pairs to future work. Moreover, one would have to establish that a diaspora creates trade not exclusively through its effect on the preferences of the representative consumer in the receiving country, but also through lower trade costs. This would complete the case that the pro-trade effect of a diaspora can mitigate or even overturn the emigration loss.

[5] $0.11\times\ 1/27,000 \times 665$ mio dollar $\approx 2,700$ dollar.

Table 1: The pro-trade effect of migrants – pooled OLS

Dependent variable: Geometric average of bilateral trade flows	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Stock of migrants from South in North										
Total	0.220***	0.120***								
	(0.015)	(0.017)								
Low-skilled			0.173***	0.091***					0.089***	0.063**
			(0.013)	(0.014)					(0.032)	(0.032)
Medium-skilled					0.204***	0.106***			−0.176***	−0.138**
					(0.015)	(0.016)			(0.057)	(0.057)
High-skilled							0.251***	0.151***	0.336***	0.226***
							(0.016)	(0.019)	(0.044)	(0.046)
Geographical and cultural proximity										
Distance	−0.984***	−1.057***	−1.041***	−1.091***	−1.007***	−1.074***	−0.959***	−1.030***	−0.964***	−1.033***
	(0.052)	(0.054)	(0.052)	(0.053)	(0.052)	(0.053)	(0.051)	(0.054)	(0.051)	(0.054)
Contiguity		0.493		0.466		0.499		0.521		0.511
		(0.336)		(0.341)		(0.343)		(0.332)		(0.330)
Common language		0.189***		0.246***		0.209***		0.139**		0.140**
		(0.068)		(0.067)		(0.068)		(0.070)		(0.071)
Colonial ties		0.875***		0.911***		0.915***		0.834***		0.815***
		(0.092)		(0.091)		(0.091)		(0.091)		(0.091)
Trade policy										
NR_PTA	0.340***	0.417***	0.311***	0.414***	0.327***	0.416***	0.355***	0.412***	0.358***	0.416***
	(0.076)	(0.077)	(0.077)	(0.077)	(0.077)	(0.078)	(0.076)	(0.077)	(0.076)	(0.077)
PTA	0.075	0.091	0.076	0.088	0.074	0.090	0.096	0.104	0.105	0.112
	(0.095)	(0.094)	(0.096)	(0.094)	(0.095)	(0.094)	(0.094)	(0.093)	(0.094)	(0.093)
FTA	−0.002	0.049	0.015	0.057	0.002	0.053	0.019	0.062	0.032	0.071
	(0.161)	(0.150)	(0.163)	(0.150)	(0.160)	(0.149)	(0.154)	(0.147)	(0.154)	(0.147)
EURO	−0.417	−0.357	−0.372	−0.325	−0.346	−0.315	−0.415	−0.370	−0.450	−0.399
	(0.346)	(0.283)	(0.335)	(0.276)	(0.331)	(0.275)	(0.314)	(0.273)	(0.309)	(0.270)
RMSE	0.793	0.767	0.801	0.769	0.798	0.769	0.786	0.764	0.784	0.763
R^2	0.892	0.899	0.890	0.899	0.890	0.899	0.894	0.900	0.895	0.900

Notes: NR_PTA: non-reciprocal preferential trade arrangements, PTA: preferential trade arrangements, FTA: free trade agreements, customs unions, $EURO$: common use of the Euro. All variables in logs, except for dummies. Balanced sample of 1195 dyads. Robust standard errors in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1%, respectively. All regressions include country-and-time effects, and a constant.

Table 2: The pro-trade effect of migrants – differenced model

Dependent variable: Geometric average of bilateral trade flows					
	(1)	(2)	(3)	(4)	(5)
Stock of migrants from South in North					
Total	0.112***				
	(0.043)				
Low-skilled		0.076**			0.078**
		(0.032)			(0.040)
Medium-skilled			0.042		−0.095
			(0.037)		(0.065)
High-skilled				0.098***	0.131**
				(0.036)	(0.056)
Trade policy					
NR_PTA	−0.262	−0.254	−0.253	−0.266	−0.273
	(0.303)	(0.302)	(0.301)	(0.301)	(0.303)
PTA	0.210**	0.200**	0.218**	0.215**	0.203**
	(0.099)	(0.099)	(0.100)	(0.100)	(0.100)
FTA	0.500***	0.499***	0.501***	0.500***	0.499***
	(0.134)	(0.134)	(0.134)	(0.136)	(0.136)
EURO	0.377***	0.395***	0.380***	0.363***	0.391***
	(0.106)	(0.102)	(0.110)	(0.109)	(0.102)
Regression-based F-test for strict exogeneity					
p-value	0.425	0.557	0.373	0.197	0.201
Wald test for equality of $\hat{\beta}^h$ and $\hat{\beta}^l$					
p-value					0.400
RMSE	0.615	0.615	0.616	0.614	0.613
R^2	0.559	0.558	0.557	0.559	0.562

Notes: We cannot reject strict exogeneity of migration, and equality of $\hat{\beta}^h$ and $\hat{\beta}^l$. See Table 1 for further notes.

4.4 Appendix

The appendix contains a detailed description of the data sources, along with a complete list of countries included, summary statistics, and results from regressions where we restrict our sample to countries which classify migrations by the foreign-born concept.

Data sources

Stock of foreign-born residents by educational level: Docquier and Marfouk (2006)

`siteresources.worldbank.org/INTRES/Resources/Dataset_BD_DocquierMarfouk.xls`

Bilateral trade flows: NBER-United Nations trade data, Feenstra et al. (2005)

`www.internationaldata.org/data/undata/undata.html`

Geographical and cultural proximity: CEPII Institute, Paris

`www.cepii.fr/anglaisgraph/bdd/distances.htm`

Trade policy dummies: Baier and Bergstrand (2007)

`web.mac.com/baier_family/iWeb/Site%202/Data.html`

Summary statistics

	1990			2000		
	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Geometric average of trade flows*	281	33	1215	666	50	4107
Stock of migrants from S in N	16552	1084	95364	27245	1960	201778
- Low-skilled	7167	233	54149	9748	394	96125
- Medium-skilled	3729	207	25522	7638	513	72350
- High-skilled	4999	280	24555	9397	593	48470
NR_PTA	0.13	0	0.34	0.14	0	0.35
PTA	0.03	0	0.18	0.11	0	0.32
FTA	0.23	0	0.42	0.25	0	0.43
$EURO$	0	0	0	0.01	0	0.08

Trade flows in millions of dollar. NR_PTA: non-reciprocal preferential trade arrangements, PTA: preferential trade arrangements, FTA: free trade agreements, customs unions, and common markets, $EURO$: common use of the Euro.

Countries included – North

ISO	Country	ISO	Country	ISO	Country
AUS	Australia	ESP	Spain	ITA	Italy
AUT	Austria	FIN	Finland	JPN	Japan
BEL	Belgium/Luxembourg	FRA	France	NLD	Netherlands
CAN	Canada	GBR	United Kingdom	NOR	Norway
CHE	Switzerland	GRC	Greece	NZL	New Zealand
DEU	Germany	IRL	Ireland	SWE	Sweden
DNK	Denmark	ISL	Iceland	USA	United States

Countries included – South

ISO	Country	ISO	Country	ISO	Country
AFG	Afghanistan	GTM	Guatemala	OMN	Oman
AGO	Angola	GUY	Guyana	PAK	Pakistan
ALB	Albania	HND	Honduras	PAN	Panama
ARG	Argentina	HTI	Haiti	PER	Peru
BDI	Burundi	HUN	Hungary	PHL	Philippines
BEN	Benin	IDN	Indonesia	PNG	Papua New Guinea
BFA	Burkina Faso	IND	India	POL	Poland
BGD	Bangladesh	IRN	Iran	PRT	Portugal
BGR	Bulgaria	IRQ	Iraq	PRY	Paraguay
BHR	Bahrain	JAM	Jamaica	ROM	Romania
BLZ	Belize	JOR	Jordan	RWA	Rwanda
BOL	Bolivia	KEN	Kenya	SAU	Saudi Arabia
BRA	Brazil	KHM	Cambodia	SDN	Sudan
BRB	Barbados	KIR	Kiribati	SEN	Senegal
CAF	Centr. Afr. Rep.	KNA	Saint Kitts and Nevis	SLE	Sierra Leone
CHL	Chile	KOR	Korea	SLV	El Salvador
CHN	China	LAO	Laos	SOM	Somalia
CIV	Cote d'Ivoire	LBN	Lebanon	SUR	Suriname
CMR	Cameroon	LBR	Liberia	SYC	Seychelles
COG	Congo Rep. of the	LBY	Libya	SYR	Syria
COL	Colombia	LKA	Sri Lanka	TCD	Chad
COM	Comoros	MAC	China Macao SAR	TGO	Togo
CRI	Costa Rica	MAR	Morocco	THA	Thailand
CUB	Cuba	MDG	Madagascar	TTO	Trinidad and Tobago
CYP	Cyprus	MEX	Mexico	TUN	Tunisia
DJI	Djibouti	MLI	Mali	TUR	Turkey
DOM	Dominican Republic	MLT	Malta	TWN	Taiwan
DZA	Algeria	MMR	Burma (Myanmar)	TZA	Tanzania
ECU	Ecuador	MNG	Mongolia	UGA	Uganda
EGY	Egypt	MOZ	Mozambique	URY	Uruguay
ETH	Ethiopia	MRT	Mauritania	VEN	Venezuela
FJI	Fiji	MUS	Mauritius	VNM	Vietnam
GAB	Gabon	MWI	Malawi	WSM	Samoa
GHA	Ghana	MYS	Malaysia	YEM	Yemen
GIN	Guinea	NER	Niger	ZAF	South Africa
GMB	Gambia The	NGA	Nigeria	ZAR	Congo, Dem. Rep.
GNB	Guinea-Bissau	NIC	Nicaragua	ZMB	Zambia
GNQ	Equatorial Guinea	NPL	Nepal	ZWE	Zimbabwe

Results from our restricted sample

The classification of immigrants is not harmonized across OECD countries. Germany, Greece, Italy, and Japan report migrants by the concept of citizenship rather than by country of birth. Thus, the respective naturalization policies may influence our results. Our destination-and-time effects perfectly control for non-discriminatory naturalization policies. However, they do not suffice to capture discriminatory policies.

We restrict our sample to countries which employ the foreign-born concept, and repeat our empirical exercise. Tables A and B respond to Tables 1 and 2 in the paper, and present the results of the pooled OLS regressions and our differenced model, respectively. The results are qualitatively and quantitatively similar, though the negative elasticity of medium-skilled migrants remains in our preferred specification, see column (5) of Table B.

Table A: The pro-trade effect of migrants – pooled OLS (Restricted sample)

Dependent variable: Geometric average of bilateral trade flows	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Stock of migrants from South in North										
Total	0.237***	0.120***								
	(0.016)	(0.020)								
Low-skilled			0.185***	0.088***					0.086**	0.058*
			(0.014)	(0.015)					(0.036)	(0.035)
Medium-skilled					0.219***	0.103***			−0.189***	−0.157**
					(0.016)	(0.019)			(0.062)	(0.061)
High-skilled							0.275***	0.160***	0.374***	0.259***
							(0.017)	(0.023)	(0.048)	(0.052)
Geographical and cultural proximity										
Distance	−0.925***	−1.043***	−0.986***	−1.079***	−0.952***	−1.064***	−0.902***	−1.008***	−0.909***	−1.011***
	(0.057)	(0.058)	(0.057)	(0.056)	(0.057)	(0.057)	(0.056)	(0.058)	(0.056)	(0.058)
Contiguity		0.634**		0.597**		0.657**		0.716***		0.740***
		(0.264)		(0.255)		(0.256)		(0.256)		(0.273)
Common language		0.128*		0.188***		0.153**		0.073		0.070
		(0.073)		(0.070)		(0.072)		(0.074)		(0.076)
Colonial ties		0.960***		1.007***		1.014***		0.898***		0.880***
		(0.093)		(0.090)		(0.091)		(0.093)		(0.093)
Trade policy										
NR_PTA	0.425***	0.479***	0.400***	0.478***	0.410***	0.479***	0.435***	0.469***	0.446***	0.482***
	(0.081)	(0.081)	(0.082)	(0.081)	(0.083)	(0.082)	(0.081)	(0.081)	(0.080)	(0.081)
PTA	0.065	0.082	0.071	0.081	0.067	0.083	0.083	0.093	0.093	0.105
	(0.097)	(0.095)	(0.098)	(0.096)	(0.097)	(0.095)	(0.095)	(0.094)	(0.095)	(0.094)
FTA	0.052	0.053	0.063	0.055	0.063	0.057	0.086	0.075	0.096	0.084
	(0.181)	(0.164)	(0.184)	(0.165)	(0.178)	(0.162)	(0.170)	(0.160)	(0.169)	(0.158)
EURO	−0.519	−0.460	−0.439	−0.405	−0.418	−0.405	−0.514	−0.490	−0.557	−0.535
	(0.447)	(0.352)	(0.427)	(0.337)	(0.423)	(0.340)	(0.404)	(0.345)	(0.393)	(0.334)
RMSE	0.813	0.787	0.822	0.789	0.821	0.789	0.805	0.783	0.802	0.781
R^2	0.885	0.892	0.882	0.892	0.883	0.892	0.887	0.894	0.888	0.894

NR_PTA: non-reciprocal preferential trade arrangements, *PTA*: preferential trade arrangements, *FTA*: free trade agreements, customs unions, and common markets, *EURO*: common use of the Euro. All variables in logs, except for dummies. Sample restricted to countries which classify migrants by country of birth. Balanced sample of 1102 country pairs. Robust standard errors in parentheses. *, **, and *** denote significance at 10, 5, and 1 percent levels, respectively. All regressions include comprehensive sets of country-and-time effects for source and destination, and a constant (all not shown).

Table B: The pro-trade effect of migrants – differenced model (Restricted sample)

Dependent variable: Geometric average of bilateral trade flows	(1)	(2)	(3)	(4)	(5)
Stock of migrants from South in North					
Total	0.099**				
	(0.050)				
Low-skilled		0.070*			0.083*
		(0.037)			(0.045)
Medium-skilled			0.007		−0.124*
			(0.045)		(0.071)
High-skilled				0.099**	0.144**
				(0.047)	(0.068)
Trade policy					
NR_PTA	−0.202	−0.194	−0.198	−0.212	−0.224
	(0.307)	(0.306)	(0.306)	(0.306)	(0.308)
PTA	0.222**	0.211**	0.231**	0.226**	0.213**
	(0.101)	(0.101)	(0.102)	(0.101)	(0.101)
FTA	0.553***	0.550***	0.552***	0.551***	0.551***
	(0.145)	(0.144)	(0.145)	(0.147)	(0.147)
EURO	0.366***	0.401***	0.388***	0.338***	0.428***
	(0.113)	(0.115)	(0.117)	(0.114)	(0.128)
Regression-based F-test for strict exogeneity					
p-value	0.227	0.403	0.210	0.0781	0.108
Wald test for equality of $\hat{\beta}^h$ and $\hat{\beta}^l$					
p-value					0.430
RMSE	0.635	0.635	0.636	0.635	0.633
R^2	0.535	0.535	0.533	0.536	0.539

NR_PTA: non-reciprocal preferential trade arrangements, PTA: preferential trade arrangements, FTA: free trade agreements, customs unions, and common markets, EURO: common use of the Euro. All variables in logs, except for dummies. Sample restricted to countries which classify migrants by country of birth. Balanced sample of 1102 country pairs. Robust standard errors in parentheses. *, **, and *** denote significance at 10, 5, and 1 percent levels, respectively. All regressions include comprehensive sets of country effects for source and destination, and a constant (all not shown). In all specifications, we cannot reject strict exogeneity of the included migration variables at 5 percent level of significance; see Wooldridge (2002, p. 285) for a detailed discussion of the test, and Baier and Bergstrand (2007) for a recent application. According to the Wald test, $\hat{\beta}^h$ and $\hat{\beta}^l$ are not statistically different.

Chapter 5

Public Information, Bayesian Updating, and Bilateral Trade[1]

5.1 Introduction

Much recent theoretical and empirical work in international trade deals with 'unpacking' trade costs; see Anderson and van Winccop (2004) for an extensive survey. Trade costs come under many guises, and their cumulative effect is believed to be large. In this paper we analyze a situation where exporters have to incur irreversible fixed costs whenever they want to engage in a trade relationship. These sunk costs may lock exporters into a holdup situation. Depending on exporters' beliefs about the likelihood of expropriation by business partners, exporters may not sink the fixed costs at all, which reduces potentially beneficial trade flows. In the presence of imperfect information, exporters form private beliefs by looking at publicly available information on whether business peers from their own country encounter cooperative or non-cooperative behavior in in some foreign market.

We draw on a new data set that codes news published in the Reuters Business Briefing into different forms of cooperative and non-cooperative events. The data set contains information about both the initiator's and the addressee's country of residence. It goes beyond conventional measures on the quality of diplomatic relations between two countries in that it primarily draws on business news rather than on diplomatic activities. The dyadic nature of the data and its availability over a time span of fifteen years (1990-2004) make it particularly attractive for use in a standard gravity model of bilateral international trade.

[1]This Chapter is based on a working paper, see Felbermayr and Jung (2006). The concept for the paper was developed jointly, empirical analysis and writing were equally shared.

Using this framework, we find evidence for our initial hypothesis, namely, that public information on cooperative or non-cooperative events substantially affects bilateral trade flows. Our findings are robust to the inclusions of proxies of institutional quality, relating amongst other things to the enforcement of laws. Interestingly, the effect is stronger for cooperative than for non-cooperative events and weaker in the subsample of OECD countries.

Our work is related to two recent strands of literature. First, Casella and Rauch (2003) and Rauch and Trindade (2004) propose theoretical models of international trade under imperfect information, using a stochastic matching process that brings together domestic and foreign firms. In this context, better information improves the matching process and increases the likelihood of successful matches, therefore boosting trade flows. Second, a growing empirical literature testifies to the importance of information costs. For example, Combes et al. (2005) show that social and business networks matter substantially for trade between French regions. Rose (2007) and Nitsch (2005) find evidence for trade-promoting diplomacy and allude to the role that embassies or state visits play for the flow of information across countries.[2]

The present paper takes a somewhat different stance on the role of information in international trade. We explicitly focus on the formation of trade relationships between exporters and importers rather than on the entry of exporters into some foreign market. In our model, the availability and content of information does not play any special role for the matching of trade partners nor does it affect the process that makes potential trade opportunities accessible to firms. Rather, we focus on the question: Given that the identity of two potential trading partners is already known, under what circumstances will the exporter sink the necessary fixed costs to make the trade relationship operative? The type of information needed for the exporter's decision concerns the likelihood of cooperative behavior by the importer. We believe that this informational friction gains relative importance in a world where new information technology and lower transportation costs render the sheer matching of trade partners ever less expensive. Our perspective also leads more straightforwardly to

[2]Our paper is loosely related to the large political science literature on the interplay between international political relations and the intensity of economic transactions, see Pollins (1989), Polachek (1980), or Reuveny (2001) for recent contributions. Martin et al. (2008) provide an interesting economic model that underpins the mechanisms highlighted in the earlier political science papers and apply it to the *Covariates of War* data base.

normative conclusions, which are more difficult to draw in the case of Nitsch's or Rose's trade promotion stories.

Our argument applies in particular to trade in differentiated goods, where trade-relation specific sunk costs may be more likely than with homogeneous goods traded on anonymous markets. It applies best to situations where the frequency of interaction between trade partners is rather low, so that potential exporters have to rely on public news rather than on their own experience when they form expectations. In that sense, our model is best suited to analyzing trade trade in homogeneous goods between advanced and less developed countries and amongst the group of less advanced countries.

Our paper is closely related to Guiso et al. (2009) who analyze the situation where beliefs on the transaction partners' probability of cooperation are essentially customary, i.e., static, in nature and depend on the availability of ´trust' in the respective relationship. They use survey data from the Eurobarometer project to proxy for the stock of trust and find that their proxy matters for several types of international transactions, including bilateral trade, foreign direct investment and portfolio investment. However, they do not address how customary beliefs are formed. Moreover, the Eurobarometer survey data draws on main street opinions rather than on the more relevant thinking of business people. It is clear that opinions of these two sets of agents may diverge strongly, as debates on trade liberalization, migration policy, or genetically modified food show.[3] Finally, the Eurobarometer data excludes non-European countries, where informational frictions may matter most.

The remainder of the paper is organized as follows. Section 2 sketches a simple model of international trade under the presence of informational imperfections and shows how we implement that framework empirically. Section 3 provides a heuristic overview over the data set and discusses how we adjust the data to make it applicable in a gravity model of international trade. Section 4 contains our main results and discusses robustness checks, involving potential endogeneity and measurement problems. Section 5 concludes.

[3]Guiso et al. (2009) recognize this fact and show that the Eurobarometer data resembles a survey among European managers. However, it is not clear whether the same notion of *trust* is applied in both surveys.

5.2 Model and empirical strategy

5.2.1 A heuristic derivation of the gravity equation under sunk costs

There is increasing consensus among theorists and empiricists of international trade that exporting involves irreversible fixed costs; see Melitz (2003) for a prominent example. Evans (2003) assumes that exporters rather than importers bear those costs. This is a reasonable assumption since exporters usually need to abide by technical and other standards of the importer's country. We incorporate this fact into our analysis, too, but recognize that the presence of sunk costs exposes the exporter to the risk of expropriation by the importer whenever prices are not determined by a competitive process characterized by free entry. To capture this argument, we assume that exporters sink the fixed cost before negotiation about a possible transaction surplus takes place.

Our model captures the essence of the interaction between sunk costs, incomplete information, and imperfect markets in a parsimonious way; see Anderlini and Felli (2001, 2004) for a fundamental theoretical characterization of this problem. We posit the following timing. First, a potential exporter draws the address of a potential importer and the *total* size of the transaction surplus from some distribution. Second, the exporter decides whether or not to pay the fixed costs. Third, both partners to the transactions bargain over the distribution of the surplus. The second step is crucial, since the exporter has to form expectations about the importers behavior in the third stage of the game. We assume that the true probability of cooperative behavior of the importer is unknown to the exporter. Hence, only information on actual transactions can be brought to bear on the exporter's expectations. The larger the number of cooperative events between transaction partners in the exporter's and the importer's markets, the more likely the exporter is in attributing a high probability of cooperation to her potential partner.

We start by assuming that the surplus generated by a trade with some foreign transaction partner is $V > 0$ while the surplus available on the domestic market is normalized to zero. This assumption reflects perfect competition in the domestic market, while imperfect competition on the foreign markets leads to rents.[4] Moreover, we posit a negative relation-

[4] It is possible to introduce non-sunk market entry costs and allow for free entry. Then the term V would be understood as an ex ante rent which is exactly absorbed by free entry.

ship between the size of the surplus and iceberg trade costs, so that the surplus generated by a transaction between an exporter in country i and an importer in country j is given by

$$V_{ij} = V \frac{1}{1 + \tau D_{ij}}, \tag{5.1}$$

where D_{ij} may denote the geographical distance between i and j or any other country pair specific variable trade cost, and $\tau > 0$ is the iceberg transport cost rate.

The irreversible fixed cost $c_{ij} \geq 0$ is also assumed to have a dyadic dimension. However, it is crucial to understand the cost term as specific to a relationship between a given exporter in country i and an importer in country j. The cost c_{ij} is not a beachhead cost that is paid to gain access to market j; rather its payments opens up the possibility to enter into a relationship with a certain trading partner. Importantly, since c_{ij} is sunk, the presence of those costs constitutes a bilateral monopoly between the exporter and the importer.

For the sake of simplicity, we assume that c_{ij} is drawn from some known distribution function with c.d.f. $G(c_{ij})$, which has standard properties and is defined over R^{+}. It is perfectly admissible in the present setup that c_{ij} positively correlates to D_{ij}. Finally, once c_{ij} is sunk, the surplus V_{ij} is distributed according to a Nash Bargaining rule, with $\beta \in (0, 1)$ the 'fair' share of the surplus accruing to the exporter. However, the exporter knows that there is the possibility that a given importer does not play according to rules and appropriates the entire surplus. In that case, the exporter does not recover the sunk costs. We assume that the exporter expects the probability of shirking by the importer to be equal to $q_{ij} \in [0, 1]$.

Putting these elements together, in the third stage of the game, the exporter expects the following payoff

$$E[\pi_{ij}] = (1 - q_{ij}) \beta V_{ij} - c_{ij}. \tag{5.2}$$

A rational exporter pays the fixed cost only if $E[\pi_{ij}] \geq 0$. Hence, the probability of a potential exporter to pay c_{ij} and engage in a transaction with an importer in country i is

$$\Pr(\pi_{ij} > 0) = G\left[(1 - q_{ij}) \frac{\beta V}{1 + \tau D_{ij}}\right]. \tag{5.3}$$

Finally, we have to describe the probability that a prospective country i exporter draws an address in country j in the first stage of the process. It appears natural to posit that the probability to draw a partner in country j, p_{ij}, depends on the economic size of country j relative to the rest of the world net of country i. Writing Y_k for the GDP level of country k, we have

$$p_{ij} = \frac{Y_j}{\sum_{k \neq i} Y_k}, \tag{5.4}$$

which is a number lying in the interval $[0, 1]$. Then, the expected value of exports from country i to j follows

$$E[X_{ij}] = p_{ij} \Pr(\pi_{ij} > 0), \tag{5.5}$$

where we have normalized the value of exports to a hypothetical partner without fixed and variable trade costs to unity.

It is straightforward, to derive the following comparative statics result from equation (5.5).

Proposition 4. *Exports from country i to j are positively affected by an increase in each country's GDP, by the total trade surplus available in the transaction, $\bar{V}$, and by the exporter's bargaining power, β. Exports are negatively affected by an increase in the conjectured shirking probability of the importer, q_{ij}, and in geographical distance between i and j, D_{ij}. Moreover, the effect of an increase in q_{ij} is stronger the larger the elasticity of the cumulative distribution function $G(.)$.*

Proof. Substituting (5.3) and (5.4) into equation (5.5), and log-linearizing around some initial equilibrium, we have

$$\hat{X}_{ij} = \hat{Y}_j + \frac{Y_i}{\sum_{k \neq i} Y_k} \hat{Y}_i + \frac{\partial G}{\partial c_{ij}} \frac{c_{ij}}{G} \left[\hat{V} + \hat{\beta} - \frac{q_{ij}}{1 - q_{ij}} \hat{q}_{ij} - \frac{\tau}{\tau + 1/D_{ij}} \hat{D}_{ij} \right]. \tag{5.6}$$

The claims in the above expression follow immediately.■

Equation (5.6) shares a number of key properties with the type of gravity equations based on models of monopolistic competition, see Feenstra (2004). First, the volume of trade declines in a convex manner with distance. Second, both trade partners' GDPs enter with positive coefficients. The effect of exporter GDP is stronger the larger the share of the exporter GDP relative to the GDP of the rest of the world. The term $(\partial G/\partial c_{ij})(c_{ij}/G)$ measures the curvature of the c.d.f. $G(.)$. The larger that expression, the stronger the probability of a beneficial transaction increases when the terms of the deal improve for the exporter.

5.2.2 The interplay of private expectations and public information

In the above framework, q_{ij} denotes the exporter's prior belief about the importer's probability to shirk. We hypothesize that the prospective exporter uses available news about the experience of other players of his country in country j to form the belief q_{ij}. This is particularly important, when own experience is unavailable. Introducing calender time as another subscript t, we may assume that a prospective exporter draws her belief on the importer's likelihood of shirking from some distribution, i.e.,

$$q_{ijt} \sim Q\left(q_{ijt}\right), q_{ijt} \in [0,1]. \tag{5.7}$$

The c.d.f. $Q\left(q_{ijt}\right)$ has the usual properties. However, its moments depend on the the quality and content of available information on how the exporters' country mates experiences in the importer's country. In order to be more precise, define the following variables

$$\bar{I}^s_{ijt} = (1-\delta)\,\bar{I}^s_{ijt-1} + I^s_{ijt}, s \in \{g, b\}, \tag{5.8}$$

$$\bar{I}_{ijt} = \bar{I}^g_{ijt} + \bar{I}^b_{ijt}. \tag{5.9}$$

A bar over the variable I indicates the stock of information, while its absence implies a flow magnitude. I^g_{ijt} and I^b_{ijt} denote the (possibly weighted) sum of cooperative (good) and non-cooperative (bad) events that occur between country i and country j at time t and their barred counterparts indicates the respective stocks. The variable $\bar{I}_{ijt}$ is the sum of all events, cooperative and non-cooperative, that are held in collective memory at time t. We assume that information is lost with the rate $\delta \in [0,1]$.

How does the stock of public information affect the formation of an individual belief? This is a complicated question, and a large game theoretical literature deals with it. At the danger of oversimplification, we assume that the distribution $Q\left(q_{ijt}\right)$ has mean $\mu_{ijt} \equiv \bar{I}^g_{ijt}/\left(\bar{I}^g_{ijt} + \bar{I}^b_{ijt}\right)$ and variance $\sigma_{ijt} \equiv \sigma\left(\bar{I}_{ijt}\right)$, where σ is a positive, decreasing function with $\lim_{s\to 0}\sigma(s) = \infty$ and $\lim_{s\to\infty}\sigma(s) = 0$. Clearly, the larger the share of cooperative events, the higher the likelihood that an exporter draws a belief that leans towards cooperation, too. The more events are present in collective memory, the more compressed the distribution from which the exporter draws her belief. In the extreme, where $\bar{I}_{ijt}$ tends to infinity, the distribution shrinks to a single point with mass unity at the importer's true likelihood to shirk. If $\bar{I}_{ijt}$ is small, no public information is available, and the exporter can only guess so that the expected deviation of her belief and the true probability of shirking will be large.

The proposed framework straightforwardly lends to a welfare perspective. The larger the expected deviation of q_{ijt} from the true shirking probability, the stronger the tendency towards inefficiently low or high export flows X_{ijt}. If exporters underestimate the true shirking probability, they will engage in transactions where the danger of expropriation is too high. However, in that situation, transactions do take place, new information is generated and the mean of the subjective distribution function μ_{ijt} moves closer to the true shirking probability. If exporters overestimate the true shirking probability, they will tend not to engage into a trade. New information is generated at a lower rate and μ_{ijt} tends to remain unchanged. Hence, if exporters start out with pessimistic priors, there will be a lack of new information, and actual trade volume will be inefficiently low in a persistent way.

If a trade transaction takes place, $\bar{I}_{ijt}$ goes up, either through I^g_{ijt} or through I^b_{ijt}. We postulate a positive relationship

$$I^s_{ijt} = \left(\gamma^s \ln T^s_{ijt}\right)^{\beta^s} N^s_{ijt}, s \in \{g, b\}, \beta^s > 0, \tag{5.10}$$

where N^s_{ijt} is the component of the information flow at time t that is unrelated to trade transactions, $\gamma^s \in [0, 1]$ is the share of transactions that have been characterized by behavior of type s, and ε^s_{ijt} summarizes all influences other than those triggered by trade that affect I^s_{ijt}. It follws from (5.8) that

$$\bar{I}^s_{ijt} = (1-\delta)^t \bar{I}^s_{ij0} + \sum_{k=1}^{t} (1-\delta)^{k-1} \left(\gamma^s \ln T^s_{ijt-k+1}\right)^{\beta^s} N^s_{ijt-k+1}. \tag{5.11}$$

Since by (5.7), the mean and variance of the c.d.f. $Q(q_{ijt})$ depend on $\bar{I}^s_{ijt}$, the realized bilateral trade volume T_{ijt} at time t depends on past trade values and on current as well as on past trade-unrelated news N^s_{ijt}, appropriately discounted by the factor $(1-\delta)$.

The are at least three interesting observations to be made. First, the process exhibits path dependency. Imagine a situation where some exporter is the first to consider trading with an importer in country j. Since her initial prior is essentially arbitrary, she engages in a beneficial trade with probability fifty percent. Imagine that the true probability of shirking in the population of country j is fifty percent, too, and let our exporter be unlucky so that her counterpart ends up to be non-cooperative. In that case, the trade takes place, but other potential exporters from country i will now draw a realization of q_{ijt} with $E[q_{ijt}] > 0.5$. Thus, it will be harder for the next possible exporter to conclude that $E[\pi_{ijt}] \geq 0$. Hence, less trades will actually take place, slowing down the process of information generation. On the other hand, if our exporter had been lucky, the next potential exporters would expect $E[q_{ijt}] < 0.5$ and the likelihood that they find it attractive to engage in trading increases.

Second, the formulation features the existence of a positive externality: larger countries have an advantage, because the relevant amount of information about trading partners in different countries will be larger and, accordingly, the beliefs will be more accurately tracking the importer's true likelihood of shirking.

Third, from the observers (the trade analyst's) perspective, the observed volume of exports of country i to j is not a sufficient statistic for the degree of cooperation of importers in country j. The reason is that trades take place once the sunk cost is paid, regardless of whether the exporter is cheated on or not. Only if the number of events in collective history grows very large, will the information conveyed by the volume of exports allow conclusions on the degree of cooperation of an importer. Hence, a larger trade volume increases the precision of our exporters' estimate of the true shirking probability, but not the expected prior μ_{ijt} itself.

5.3 Data and empirical strategy

This section argues that the Reuters Business Briefing (RBB) data set contains useful information that can be brought to bear on the empirical relationship implied by our theoretical framework. First, we offer a description of the RBB data, discuss the construction of an aggregate index, and provide a heuristic exploration of the data. Second, we discuss our empirical model and address econometric issues.

5.3.1 Exploiting data from business briefings

How can we hope to measure the availability and quality of information available in one country about the other, and how can we sort out the bias of this information, i.e., whether it informs about cooperative or non-cooperative events that involve the two countries? Clearly, one can turn to news archives which are usually maintained by press agencies or newspapers. But how can one make sense of this gigantic stock of news, sort out the relevant data from the irrelevant?

Business leaders rely on timely and accurate information about what is going on in the markets that they are active in. While large companies sometimes provide the required information services in house, there is also a profitable market for them. After World War II, specialized corporate information services emerged, with two products leading the field: Dow Jones Interactive and Reuters Business Briefing. In 1999, Dow Jones & Company and Reuters merged their corporate news branches and founded Factiva.

The principle of business briefings is that they contain condensed and filtered information targeted specially to high-level executives in the private sector. The data set that we use draws on the Reuters Business Briefing (RBB) data which – with some modifications and under another name – are still sold after the merger. This data is attractive for our purposes precisely because business people can be expected to draw on them when they form priors about what they have to expect when they interact with transaction partners in other countries. Moreover, in that data, the formidable task of filtering relevant information is achieved in a market environment, since the editors of business briefings have to meet their customers' demand in order to be successful.

Still, using business briefings to construct a data set that can be used in an empirical analysis of bilateral trade flows is a formidable task. Fortunately, political scientists have produced such data. In a large research project, The Kansas Event Data System (KEDS) Project has developed a program that allows automated coding of the Reuters Business Briefing data to generate event data with a (potentially) dyadic dimension[5]. The data

[5]Note that the data has been collected for the purpose of forecasting military and diplomatic conflicts between countries. The data therefore contains much more information than what we can make use of in the present context. See below for more details.

construction tool is discussed at length in King and Lowe (2003) who argue that machine coding is equivalent to human coding in terms of bias and better in terms of efficiency. Their data set consists of daily news that are extracted from RBB or from a precompiled database (1990-05/2003), Factiva's World News (06/2003-08/2003), and Reuters World News (09/2003-2004). The availability of information has increased throughout the period due to technological change. Especially the waves of intranet (1994) and extranet (1997) led to a rise in news considered in the data files. The data is organized according to the typology of Integrated Data for Events Analysis (IDEA) that distinguishes about 200 types of events[6] and contains detailed information on the initiator and the addressee. It contains about 2.5 millions events in 1990-1994, more than 4 millions in 1995-2000, and about 3.5 millions in 2000-2004. Data and documentation are available from Gary King's homepage at Harvard (http://gking.harvard.edu/events/).

In order to better understand the organization of the data, consider the following entries in the data set:

Event date	Event form	Initiator name	Initiator sector	Initiator level	Addressee name	Addressee sector	Addressee level
9-Mar-00	NEGO	USA	BUSI	ORGA	GER	BUSI	ORGA
9-Mar-00	NEGO	GER	BUSI	ORGA	USA	BUSI	ORGA
9-Mar-00	DISC	GER	BUSI	ORGA	USA	BUSI	ORGA
9-Mar-00	DISC	USA	BUSI	ORGA	GER	BUSI	ORGA
9-Mar-00	DISC	GER	BUSI	ORGA	USA	BUSI	ORGA
9-Mar-00	DISC	USA	BUSI	ORGA	GER	BUSI	ORGA
9-Mar-00	AGAC	GER	BUSI	ORGA	USA	BUSI	ORGA

For each event, we know the type of the event (event form) and details about the initiator and the addressee. Hence, the data is dyadic. Initiator/addressee sector and level give more precise information about the identity. In the example above, the German telecoms giant, Deutsche Telekom, undertakes the attempt of making an acquisition in the US telecommunication market. Both companies are codes as business organizations (ORGA), and the agents are representatives of that business organizations (BUSI). On 9-Mar-00, Reuters and

[6]See Bond et al. (2003) for a description of the coding scheme.

the Wall Street Journal both report that negotiations (NEGO) between Deutsche Telekom and Qwest still go on. Initiator and addressee are not separable, so one record in the news leads to two entries. As the negotiations are in the verge of failing, it is worth posting that the discussions (DISC) are still going on, until in the late evening an agreement (AGAC) is achieved. It this case "Deutsche Telekom" declares that it is not longer interested in an acquisition of Qwest.

As pointed out above, the data spans the years 1990 to 2004, and contains a total of about 10 million events. We are interested primarily in events that involve two different countries. We suppress about 8 million data points for which initiator name and addressee name coincide. King and Lowe (2003) classify events into cooperative, non-cooperative, and neutral. We drop all neutral events, which reduces the number of events roughly by another 0.9 million. Moreover, we drop all events that have to do with warfare. The reason is that for those situations we usually have no trade data. Finally, we focus only on events where the players's countries of residence are likely to coincide with their nationality. This leads us to drop, i.a., International Organizations such as the WTO or NATO. The final number of events considered in our study is approximately 240,000.

We consider the following seven types of cooperative events and eight types of non-cooperative events. *Agreement* contains the acceptation of invitations and proposals as well as the formation of any kind of alliance. Meetings, consultations, negotiations and mediations are summarized to *Consultations*. *Endorsement* covers the ratification of an agreement, the laud of someone's practice or policy and the expression of condolences, regrets, and forgiveness. *Grant* contains the extension of an invitation or the shelter to a victim as well as the initiation or improvement of relations and relaxing economic sanctions. *Promise* captures all promises of economic, financial and non-material support and the promise to mediate. *Reward* is the extension of economic and non-economic assistance, while *Yield* captures the give-up of locations and possessions. Blaming and rebuking is summarized to *Accusation*, *Complain* contains informally and formally protests, while *Denial* means the denial of accusations. *Demand* covers demand for information, policy and material support as well as demand for discussions, negotiations or mediations. *Demonstration* captures all kind of sit-ins and picketing. Rejection of proposals, settlements, requests and law is summarized to *Rejection*. *Sanction* is the reduction of routines and planned activities, the terminations of aid and the halting of discussions, negotiations and mediations as well as the break of relations. *Threat* covers the threat of sanctions, reducing aid, halting discussions, negotiations and mediations as well as issuing an ultimatum.

Cooperative and non-cooperative events respectively are summed up over all types, sectors and levels to construct proxies for the annual cooperation share between the public and the business sectors of two countries. We use equal weights for all events, because the weight of news is implicitly given by editor's selection of news into the RBB.

5.3.2 A heuristic exploration of the RBB data

Before we go on, we need to convince the reader that the RBB data reproduces stylized facts of cross-countries relations worldwide. In order to do so, we provide some detail on the degree of cooperation that the U.S. displays with respect to its partner countries, and vice versa. We also try to explain cooperative and non-cooperative behavior in a simple regression analysis.

Table 5.1: **Cooperation with the U.S.**

	U.S. is initiator		**U.S. is addressee**	
World region	$\frac{I^g_{ij}}{I^g_{ij}+I^b_{ij}}$	$I^g_{ij} + I^b_{ij}$	$\frac{I^g_{ij}}{I^g_{ij}+I^b_{ij}}$	$I^g_{ij} + I^b_{ij}$
Africa	0.77	8	0.80	7
Asia (excl. Japan)	0.83	48	0.77	48
Australia/New Zealand	0.84	59	0.83	87
Canada	0.84	242	0.77	301
Cuba	0.70	74	0.46	87
Europe	0.87	55	0.86	56
Japan	0.81	557	0.84	629
Middle East Asia	0.77	95	0.72	79
Western Hemisphere	0.86	21	0.80	17
Total	0.83	42	0.80	40

Table 5.2 provides an overview of U.S. relations with the main world regions. The table differentiates between events in which the U.S. has been the initiator and those, where the U.S. has been the addressee. The cells in the table contain averages of the degree of cooperation, $I^g_{ij}/\left(I^g_{ij}+I^b_{ij}\right)$, and the 'size' of information flows $I^g_{ij} + I^b_{ij}$. Averages are computed over years and countries. For example, in 77 percent of all events, the U.S.

behaves cooperatively with an average African partner country. However, the sum of the cooperation and the non-cooperation index averaged over years and countries, $I_{ij}^g + I_{ij}^g$, is 8, indicating an average volume of information that is much below the sample mean of 42.

Two facts stand out from Table 5.2. First, European countries enjoy the highest degree of cooperation by the U.S., while the value for Cuba is lowest. However, cooperation shares are consistently above 50 percent, which is a well noted feature of the data (King and Lowe, 2003). Second, the total volume of news is highest for Japan, followed by Canada. European countries have a rather low volume of 55. These results are driven by the aggregation scheme used for computing I_{ij}^g and I_{ij}^b, and by the fragmentation of countries in each world region. If we were to consider the volume of information flows between the U.S. and the European Union (EU 15), we could expect to find a volume of the magnitude 55×15 which would be the highest for all world regions. Similar observations apply for the case where the U.S. is the addressee. Overall, the U.S. is slightly more often an initiator than an addressee (42 versus 40), and tends to behave somewhat more cooperatively (83 versus 80).

Table 5.2 shows the conditional correlation between our cooperation and non-cooperation indices and a host of economic, geographical, and institutional variables. We have run the following regression,

$$I_{ijt}^s = \zeta_0 + \boldsymbol{\zeta}_1'\mathbf{X}_{jt} + \boldsymbol{\zeta}_2'\mathbf{X}_{it} + \boldsymbol{\zeta}_3'\mathbf{X}_{ijt} + \upsilon_t + u_{ijt},\ s \in \{g, b\} \tag{5.12}$$

where I_{jit}^s measures the extent of cooperation and non-cooperation between two countries, respectively, where the initiator is indexed by subscript j and the addressee by subscript i. The vectors $\mathbf{X}_{jt}$ and $\mathbf{X}_{jt}$ contain characteristics of the initiator and the addressee, respectively: the natural logarithms of their GDPs and per capita incomes. The vector $\mathbf{X}_{ijt}$ has dyadic dimension and contains the natural logarithm of geographical distance, and a host of dummy variables, that measure whether the countries are in the same regional trade agreement, whether their chief government executives are share the same political background (right, left, center)[7], whether the countries share a common border or a common language, or whether they have similar entries in the ICRG corruption index.[8] We control

[7]Database of Political Institutions, DPI; see the Appendix for details.

[8]The International Country Risk Guide (ICRG) corruption index is meant to capture the likelihood that

Table 5.2: **Explaining cooperative behavior**
Dependent variable: Degree of cooperation/non-cooperation

Variable	**Full sample**		**OECD sample**	
	Cooperation	**Conflict**	**Cooperation**	**Conflict**
ln GDP target	1.714***	0.377***	9.313***	1.960***
	(0.17)	(0.045)	(1.89)	(0.40)
ln GDP source	1.743***	0.373***	8.846***	1.583***
	(0.17)	(0.046)	(1.77)	(0.37)
ln GDP/cap. target	-0.436***	-0.0978***	0.508	0.0877
	(0.11)	(0.032)	(1.20)	(0.28)
ln GDP/cap. source	-0.315***	-0.112***	1.245	-0.165
	(0.10)	(0.037)	(1.21)	(0.28)
Institutional similarity	-0.600***	-0.0746**	-3.295***	-0.667***
	(0.13)	(0.030)	(1.06)	(0.23)
Both same RTA	-2.516***	-0.560***	-0.961	0.148
	(0.39)	(0.10)	(1.99)	(0.51)
Same party	1.006***	0.168**	2.678**	0.498*
	(0.28)	(0.079)	(1.16)	(0.28)
ln Geographical distance	-0.699***	-0.128***	-0.944	-0.192
	(0.15)	(0.038)	(0.96)	(0.22)
Contiguity	8.397***	2.815***	9.446	3.956**
	(1.67)	(0.58)	(7.25)	(1.95)
Common language	3.488***	0.709***	20.19**	4.034**
	(0.60)	(0.14)	(9.12)	(1.90)
Year fixed effects	YES	YES	YES	YES
Observations	131457	131457	7637	7637
R^2	0.10	0.05	0.23	0.18
Number of clusters	12760	12760	812	812
F-statistic	21.62	10.17	4.864	3.358
RMSE	15.12	4.646	40.28	9.453

Institutional similarity: absolute difference in ICRG Index. *Same party*: Dummy, takes value 1 if chief executive party of same wing (right, left, center). Each regression includes a constant and time dummies (all not shown). Variance-covariance matrix adjusted for serial correlation within country clusters. Robust standard errors in parentheses.
*** significant at 1%, ** at 5% and * at 1% level.

for time effects in a non-parametric way by introducing year dummies v_t and adjust the variance-covariance matrix for serial correlation within country clusters (the same pairs ij are observed several times) and use the Huber-White heteroskedasticity robust standard errors.

We run the regression for two samples: the full sample, comprising 12,760 country pairs

government officials will demand special payments, and the extent to which illegal payments are expected throughout government tiers as ranked by panels of international experts. In our context, the index should measure the similarity between two countries' political institutions.

and 131,457 observations, and a subsample consisting of OECD countries only. The latter contains 812 country pairs and a total of 7637 observations. The following results seem noteworthy. First, the qualitative pattern of the intensity of cooperative and non-cooperative behavior depends in a symmetric way on the explanatory variables. For example, contiguity increases the cooperation index by 8.4 points and the conflict index by 2.8 points, geographical distance decreases the extent of cooperation, but does not have much bearing on non-cooperation. The same is true for two countries that are in the same regional trade agreement (RTA). Interestingly, institutional similarity, measured by the absolute distance in two countries' ICRG scores, reduces cooperative and non-cooperative events, but in line with intuition, the latter are affected much more strongly. Two countries in which the ruling political party has the same background (right, left, centrist), have more cooperative events. However, there are also more non-cooperative events, but, again, the magnitude of the effect is much smaller. The existence of meaningful conditional correlations between the RBB indices and political, geographical, and economic variables suggests that the RBB data correlate to salient features of cross-country relations, that go beyond business ties.

5.3.3 Empirical strategy

Using $E[q_{ijt}] = \mu_{ijt} = \bar{I}^g_{ijt} / \left(\bar{I}^g_{ijt} + \bar{I}^b_{ijt} \right)$ in expression (5.5), and assuming a constant elasticity of the c.d.f. $G(c)$, we can write the volume of exports from i to j in the following way

$$\ln X_{ijt} = \alpha_0 \ln \beta + \alpha_1 \ln V + \phi \left(1 - E[q_{ijt}]\right) + \ln Y_{jt} - \ln \sum_{k \neq i} GDP_k + \tau \ln D_{ij} + u_{ijt}, \quad (5.13)$$

where u_{ijt} is an error term whose properties have to be discussed below. Our simple theoretical model would suggest that the term $\alpha_0 \ln \beta + \alpha_1 \ln V$ is just a constant α. While we do not have any reason to suspect that the exporter's 'fair' bargaining power β should vary across importers or time, the same is unlikely to hold for V, the total joint available from the transaction. The surplus is likely to depend also on the exporter's market size, Y_{it}, as well as on its per capita income of both countries. Moreover, the price competitiveness should also matter. Turning to trade costs, the theoretical model only talks about geographical distance, which is of course time-invariant. In reality, the evolution of technology and the stance of trade policy should be taken into account as well.

B&B specification

These considerations motivate us to include a non-parametric time trend and a host of controls that are meant to capture the stance of commercial policy and the quality of institutions. We include measures of the population size for both countries, which accounts jointly with the GDP terms for per capita income. Finally, to account for price competitiveness, we follow Baier and Bergstrand (2001) and include price level data for both countries.[9] This leaves us with a first specification, that–for mnemonic reasons is called B&B specification:

$$\begin{aligned}\ln X_{ijt} = \alpha + \phi RBB_{ijt} + \tau \ln D_{ij} + \gamma_1 \ln Y_{jt} + \gamma_2 \ln Y_{it} + \gamma_3 POP_{it} + \gamma_4 POP_{jt} \\ + \pi_1 P_{it} + \pi_2 P_{jt} + \boldsymbol{\xi}_1 \mathbf{POL}'_{it} + \boldsymbol{\xi}_2 \mathbf{POL}'_{jt} + \boldsymbol{\xi}_3 \mathbf{POL}'_{ijt} + \theta_t + \upsilon_{ijt}, \quad (5.14)\end{aligned}$$

where we use the variable RBB_{ijt} to proxy for $1 - E\left[q_{ijt}\right]$; see below for details. POP_{it} refers to the population of country i, P_{it} is its price level and **POL** is a collection of policy or institutional variables, which may be dyadic in nature, e.g., the incidence of regional trade areas, or apply to the exporter or importer separately, e.g., the quality of institutions. The coefficient of interest will be ϕ.

FE specification

We estimate a fixed effects model that we term FE specification. Here, we replace all exporter- or importer-specific price and policy variables by a comprehensive set of fixed effects. This regression is somewhat closer to the standard specification used in gravity equations of bilateral trade; see Feenstra (2004) for a survey or Combes *et. al.* (2005) for

[9] However, in contrast to our work, Baier and Bergstrand work with first differences, since they are ultimately interested in a decomposition of world trade *growth*.

a recent application:

$$\ln X_{ijt} = \alpha + \phi RBB_{ijt} + \tau \ln D_{ij} + \gamma_1 \ln Y_{jt} + \gamma_2 \ln Y_{it} + \gamma_3 POP_{it} + \gamma_4 POP_{jt} + \nu_i + \nu_j + \theta_t + u_{ijt}, \tag{5.15}$$

where ν_i and ν_j are comprehensive sets of exporter and importer dummies, respectively. This specification has the advantage of lower data requirements[10] so that the sample of countries and years covered is larger. However, the disadvantage is that the time dimension of the price- and policy variables is usually relegated into time fixed effects, which are identical for the same sample. Note that the dependent variable in both models is the natural logarithm of bilateral exports. This implies that country pairs with zero trade drop from the sample. One could argue that either a semi-log specification, or a Tobit-type regression methodology should therefore be preferred to our model, see Felbermayr and Kohler (2004) for a discussion. However, we follow the established literature and drop zero-trade observations.

Purging and IV Strategy

We have argued that the RBB data contains useful information that can be brought in as a proxy for μ_{ijt}. However, a number of problems arise. First, and most importantly, there is a strong endogeneity problem. Successful transactions will produce good news, and this will in turn drive up the RBB index. In order to sort out this simultaneity problem, exploiting data with a highly disaggregated time dimension would help. While the RBB data is available on a daily basis, trade data is not.[11]

A second problem arises due to the fact that single transactions (of whatever type they may be) often are not independent from each other: one event triggers the next so that it is difficult to exactly identify what an event is. Most event studies share this feature. We are not particularly worried by this fact, since the reappearance of the same event may indicate that this event is more important relative to the others and should therefore obtain a higher weight. However, if a non-cooperative event triggers a cooperative event later in the same year, we may have a problem, since the measure μ_{ijt} is reduced by these events, while it is

[10] In particular, price data is difficult to obtain for a large sample of countries from 1990-2004.

[11] In recent years, monthly and quarterly data are available also for bilateral trade flows.

more natural to argue that the net effect of these effects should cancel out.

For these reasons, we purge the cooperation and conflict data before using it in the regression models (5.14) and (5.15). More precisely, we adjust the cooperative (non-cooperative) RBB measure so that it is orthogonal to the most important drivers of bilateral trade, namely GDP levels and geographical distance, and an aggregate measure of non-cooperative (cooperative) behavior. We include time dummies and run the following regressions using OLS (correcting the variance-covariance matrix for within-group correlation) and a zero-inflated Poisson regression when applicable:

$$I^g_{ijt} = \alpha^g_{ij} + \beta^g_1 \ln Y_{jt} + \beta^g_2 \ln Y_{it} + \delta^g \ln D_{ij} + \xi I^b_{ijt} + v^g_t + u^g_{ijt}, \tag{5.16}$$

$$I^b_{ijt} = \alpha^b_{ij} + \beta^b_1 \ln Y_{jt} + \beta^b_2 \ln Y_{it} + \delta^b \ln D_{ij} + \xi I^g_{ijt} + v^b_t + u^b_{ijt}. \tag{5.17}$$

Table 5.9 in the Appendix provides the results of some of these regressions, with either the aggregate cooperation measure or the underlying event forms as dependent variables. The Appendix also gives summary statistics for the dependent and independent variables used in our regressions. Note that whenever we purge single event forms, we use a zero-inflated Poisson model, since we deal with count data that features a large amount of zero entries. However, once the different events are aggregated, the data becomes continuous. Hence, we use OLS. The key insight from this table is that larger countries (as measured by GDP) have a larger amount of information, both cooperative and non-cooperative, and that geographical remoteness reduces our RBB measure. In the following, we use the residuals from the above equations as 'purged' measures of cooperative and non-cooperative behavior.

Finally, we recognize that even the purged measures do not appropriately account for *simultaneity* bias. For this reason, when we run our regressions (5.14) and (5.15), we instrument the purged cooperation indices by their lags. We find that two lags are enough to consistently achieve first-stage F-statistics that satisfy Stock and Watson's rule of thumb (F-statistic above 10). We have experimented with other lag structures and have also tried differences; these variations do not seem to make much of a difference. Since we use constructed variables from our 'purging' equations (5.16) and (5.17) in our final regressions, we have to account for sampling error stemming from these auxiliary regressions. Hence, we bootstrap the standard errors in our IV regressions. In order to be on the safe side, we choose 200 replications. Note that we have chosen $\delta = 1$ in (5.8).

5.4 Results and robustness checks

5.4.1 Results

B&B Specification

Table 5.3 presents the results of *instrumental variables* regressions for model (5.14), differentiating between the full sample, and the OECD subsample. The detailed description of the variables and their respective sample statistics can be found in Table in the Appendix. We report two main models, where the degree of cooperation between the two partners is measured by the the aggregate cooperation index and the non-cooperation (conflict), respectively. We choose this setup in order to avoid multicollinearity problems. In all regressions, we include an extensive set of dummies on the membership of the exporter, importer, or both, is regional trade agreements but do not report the results to save space.

Signs and sizes of the regression coefficients reported in Table 3 resemble strongly the usual output (see, i.a., Feenstra, 2004), with all relevant variables entering at high levels of statistical significance. Both, the GDP measures and geographical distance enter with elasticities close to unity and bear the right sign. Contiguity and common language also matter with the expected large effects (taking exponents, contiguity increases exports by a factor 1.8 and common language by a factor 2.4). The price level variables enter also with high significance levels and the expected pattern of signs. The WTO dummies tend to have the 'wrong' signs, consistent with the findings of Rose (2004), who used essentially the same data source for the trade data. Corruption of either trade partner tends to reduce exports, but the effect is not statistically significant in all samples and specifications.

Turning to our variables of interest, the degree of cooperation between two countries, we find that cooperative behavior of the importer towards some exporter spurs its imports from that source country, and does so in a statistically significant way (with P-values below one percent). The economic size of the effect is non-negligible neither. In the full sample, a one standard-deviation increase in the cooperation index drives trade up by $100 \times 23.11 \times$

Table 5.3: **Regressions results: B&B Specification**
Dependent variable: ln Exports, IV regression

	Cooperation		**Conflict**	
Variable	**Full**	**OECD**	**Full**	**OECD**
Cooperation Index	0.00258***	0.00247***	-0.0263***	-0.0716
	(0.00044)	(0.00024)	(0.0073)	(0.13)
ln GDP Exporter	1.264***	1.198***	1.284***	1.371***
	(0.019)	(0.032)	(0.020)	(0.31)
ln GDP Importer	0.857***	0.677***	0.874***	0.798***
	(0.016)	(0.035)	(0.016)	(0.22)
ln Population Exporter	-0.261***	-0.388***	-0.253***	-0.405***
	(0.019)	(0.036)	(0.020)	(0.079)
ln Population Importer	-0.000254	0.200***	0.00936	0.208***
	(0.018)	(0.039)	(0.017)	(0.048)
ln Price level Exporter	-0.0922***	0.195***	-0.0782***	0.284
	(0.028)	(0.027)	(0.029)	(0.19)
ln Price level Importer	0.196***	0.164***	0.208***	0.241**
	(0.025)	(0.033)	(0.024)	(0.10)
ln Geographical distance	-1.224***	-1.110***	-1.232***	-1.272***
	(0.010)	(0.017)	(0.012)	(0.27)
Contiguity	0.607***	0.226***	0.732***	0.476
	(0.051)	(0.052)	(0.053)	(0.48)
Common language	0.881***	0.533***	0.909***	0.682***
	(0.026)	(0.035)	(0.025)	(0.19)
Corruption Index Exporter	-0.109***	0.0438***	-0.105***	0.0622
	(0.010)	(0.011)	(0.0100)	(0.051)
Corruption Index Importer	-0.0112	-0.0502***	-0.00967	-0.0373
	(0.0100)	(0.012)	(0.0098)	(0.045)
Exporter in WTO	0.143***	-0.455***	0.157***	-0.372***
	(0.039)	(0.037)	(0.040)	(0.14)
Importer in WTO	-0.0962***	-0.498***	-0.0905**	-0.388**
	(0.035)	(0.053)	(0.036)	(0.16)
Year fixed effects	YES	YES	YES	YES
RTA dummies	YES	YES	YES	YES
Observations	50615	6384	50615	6384
R^2	0.74	0.87	0.74	0.76
RMSE	1.764	0.799	1.770	1.080

Each regression includes a constant, various dummies for membership in RTAs and time dummies (all not shown). Variance-covariance matrix adjusted for serial correlation within country clusters. Robust standard errors in parentheses. *** significant at 1%, ** at 5% and * at 1% level.

$0.0026 = 5.97$ percent. Turning to the OECD sample, the effect is even stronger, with a one standard deviation increase leading to a trade gain of $100 \times 46.35 \times 0.0025 = 11.44$ percent.

Non-cooperative behavior has strong statistical support in the full sample, but fails to have an effect distinguishable from zero in the OECD sample. In the full sample, a one standard-deviation increase in the index leads to a reduction of trade of approximately $100 \times 6.28 \times -0.0263 = 16.52$ percent. Hence, non-cooperative behavior by the importer towards some country seems to strongly deter imports from that country. Hence, non-cooperative behavior by the importer seems to reduce imports more than the extent to which cooperative behavior increases them.

The fact that non-cooperative behavior does not seem to matter in the OECD sample is interesting. Clearly, the identification of a significant negative effect depends on the between-group variance between the groups of OECD and non-OECD countries. The OECD sample is very special in several respects. First, it contains only rich countries with fairly high degrees of openness. Second, institutions are well developed in those countries so that exporters can expect that potential non-cooperative behavior by importers can be swiftly and efficiently tackled using legal instruments. Third, OECD countries closely cooperate in a number of important economic and military networks, such as the EU or NATO. This may reduce the economic importance of single firms' non-cooperative behavior since it happens in an environment that is generally geared towards cooperation.

FE Specification

Table 5.4 reports the results of running the fixed effects specification (5.15). Note that this specification draws on a somewhat different sample that the B&B specification, since several variables (in particular the data on price levels) that was used before is now replaced by fixed effects.

As in most applications of gravity-type trade models, introducing importer and exporter fixed effects has significant effects on the size of the estimated coefficients. For example, the GDP coefficients are now longer in the neighborhood of unity. Some researchers constrain these coefficients to unity (Anderson and Van Wincoop), however, we refrain from doing so. Except from the income variables (which can be expected to correlate strongly with the

Table 5.4: **Regressions results: FE Specification**
Dependent variable: ln Exports, IV regressions

	Cooperation		**Conflict**	
Variable	**Full**	**OECD**	**Full**	**OECD**
Cooperation Index	0.000138	0.00227***	-0.0292***	-0.0384
	(0.00058)	(0.00026)	(0.0048)	(0.039)
ln GDP Exporter	0.142***	0.440***	0.143***	0.460***
	(0.022)	(0.10)	(0.030)	(0.11)
ln GDP Importer	0.297***	0.811***	0.297***	0.828***
	(0.032)	(0.097)	(0.023)	(0.12)
ln Population Exporter	-0.352**	-5.052***	-0.329**	-4.790***
	(0.18)	(1.29)	(0.17)	(1.31)
ln Population Importer	-0.219*	-1.084	-0.198	-1.467
	(0.13)	(1.11)	(0.14)	(1.26)
ln Geographical Distance	-1.589***	-1.119***	-1.598***	-1.250***
	(0.010)	(0.019)	(0.0097)	(0.11)
Contiguity	0.497***	0.313***	0.555***	0.456***
	(0.034)	(0.053)	(0.039)	(0.14)
Common language	0.824***	0.429***	0.826***	0.474***
Year fixed effects	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES
Observations	135184	6384	135184	6384
R^2	0.74	0.90	0.74	0.87
RMSE	1.905	0.710	1.906	0.810

Each regression includes a constant, various dummies for membership in RTAs and time dummies (all not shown). OECD sample includes additional controls for price level and ICRG Index. Variance-covariance matrix adjusted for serial correlation within country clusters. Robust standard errors in parentheses. *** significant at 1%, ** at 5% and * at 1% level.

fixed effects), the other covariates, such as geographical distance, contiguity, or common language, do not change much compared to the B&B specification.

Turning to our variables of interest, we find that the cooperation index still affects bilateral exports when the sample is restricted to OECD countries, but no longer does so in the full sample, which is in stark contrast to our results from the B&B specification. In contrast, in line with the B&B specification, non-cooperative behavior matters for the full sample, but not in the OECD sample. This pattern of results is interesting and suggests a large deal of asymmetry between rich and poor countries: news on cooperative behavior facilitates trade in the OECD sample, while there is no evidence for this effect when poor and emerging countries are included into the sample. However, non-cooperative news does not significantly reduce trade among rich OECD countries, but does so once the sample is extended to include poor countries as well. Turning to the quantities, a one standard-deviation change in the cooperation index increases bilateral trade among OECD countries by about 9.53 percent, while a one standard-deviation change in the conflict variable reduces trade by about 12.69 percent in the full sample. We conclude that the overall effect of news on trade volumes seems to be large.

5.4.2 Robustness checks

This subsection addresses an important issue that arises in relation to the construction of our cooperation and non-cooperation measures. Both have been constructed using Goldstein (1992) weights. However, at least in principle, it is possible that the results depend on that specific aggregation scheme. Rather than proposing yet another necessarily arbitrary aggregation method–for example, unweighted sums or averages– we use the ingredients used in the aggregate indices separately and check whether they matter for bilateral trade flows. The basic insight from this section is that the results tend to hold also on the disaggregated level and do not seem to depend on the specific aggregation scheme implied by the Goldstein (1992) weights.

Table 5.5: **Robustness Checks–B&B Specification**
Dependent variable: ln Exports, IV regressions
Effects of a one standard-deviation change in independent variable

	Cooperation			Conflict	
Event	**Full**	**OECD**	**Event**	**Full**	**OECD**
Index	**0.0597*****	**0.1144*****	**Index**	**-0.1653*****	**-0.7752**
	(0.0103)	(0.0110)		(0.0459)	(1.4075)
Agreement	0.1906***	0.2482***	Accusation	-0.1167***	-0.1663***
	(0.0290)	(0.0416)		(0.0425)	(0.0568)
Consultation	0.0607***	0.2749***	Complaint	-0.0756	-0.0688
	(0.0158)	(0.0491)		(0.1695)	(0.2121)
Endorsement	0.1146***	0.1578***	Demand	0.1840	-0.4261
	(0.0268)	(0.0527)		(0.2072)	(0.2905)
Grant	0.2268***	0.3073***	Demonstration	-0.8075*	-0.3025
	(0.0640)	(0.0911)		(0.4780)	(0.4911)
Promise	0.0866**	0.0939	Denial	-0.1949**	0.1625
	(0.0432)	(0.0973)		(0.0878)	(0.3897)
Reward	0.1563***	0.2129**	Rejection	0.0714	-1.8593
	(0.0432)	(0.1007)		(0.2426)	(1.2798)
Yield	0.1385***	0.2541***	Sanction	0.5025	-0.3039
	(0.0351)	(0.0888)		(0.7555)	(0.4538)
			Threat	-0.3657	-0.0658
				(0.8783)	(0.9845)

Index refers to the aggregate cooperation/non-cooperation measure. Each regression includes all those covariates also present in the regression results shown in Table 3, including a constant, various dummies for membership in RTAs and time dummies. Variance-covariance matrix adjusted for serial correlation within country clusters. Robust standard errors in parentheses.
*** significant at 1%, ** at 5% and * at 1% level.

B&B Specification

Table 5.5 provides an overview of the regression coefficients obtained when the ingredients of the aggregate measure are used in regressions of type I instead of the aggregate measure itself. For comparison reasons, we provide the aggregate measure in the first line of the table. The left block refers to the seven cooperative event forms that hide behind the aggregate index; the right block reports the eight non- cooperative events. Note that every entry in that table corresponds to a different regression. The only difference among these regressions is the exact definition of the cooperation/non-cooperation variable. The numbers shown are to be interpreted as the effects that a one standard-deviation change of the respective indices implies for bilateral trade volumes.

The results are quite striking. We find that all cooperative events used in the computation of the aggregate index enter positively and with precisely estimated coefficients, in both, the full and the OECD samples. The only exception to this rule is 'Promise' which does not turn out significant in the OECD sample.

FE Specification

The last set of robustness checks is reported in Table 5.6. This table is constructed similarly to Table 5.5. However, the results are more ambivalent. While disaggregation does not seem to undo the general direction of the estimates for cooperative events in the OECD sample, it seems to matter in the full sample, where the aggregate measure has not shown up significantly different from zero. Interestingly, all ingredients of the aggregate measure have the right sign and are estimated with high precision, indicating that news on cooperative behavior systematically increases bilateral trade flows. 'Agreements', the largest single event form boosts trade in a substantial way: a one standard-deviation increases in the number of such events leads to an increase in bilateral trade volumes by 9.61 percent. With non-cooperative events few disaggregated events seem to work out. Only the events 'demonstration' and 'denial' are significantly different from zero and bear the right signs. The size of the coefficient attached to 'demonstration' is quite large and shows that trade relations are severely disrupted when there are demonstrations in one country against the other country. Such demonstrations occur not too rarely in our data, and they also occur

Table 5.6: **Robustness Checks–FE Specification**
Dependent variable: ln Exports, IV regressions
Effects of a one standard-deviation change in independent variable

	Cooperation			Conflict	
Event	**Full**	**OECD**	**Event**	**Full**	**OECD**
Index	**0.0021**	**0.0953*** **	**Index**	**-0.1269*** **	**-0.3779**
	(0.0089)	(0.0109)		(0.0209)	(0.3857)
Agreement	0.0961***	0.2208***	Accusation	-0.0466	-0.0687
	(0.0190)	(0.0372)		(0.0311)	(0.0552)
Consultation	0.0834***	0.2479***	Complaint	0.0906	-0.0138
	(0.0137)	(0.0445)		(0.2182)	(0.2603)
Endorsement	0.0918***	0.1466***	Demand	0.1597	-0.1366
	(0.0185)	(0.0489)		(0.1296)	(0.3140)
Grant	0.1489***	0.2928***	Demonstration	-0.7939**	0.0077
	(0.0392)	(0.0897)		(0.3950)	(0.6052)
Promise	0.1025**	0.0861	Denial	-0.1299***	0.1485
	(0.0436)	(0.0902)		(0.0450)	(0.3386)
Reward	0.1303***	0.1965**	Rejection	0.1656	-1.4487
	(0.0214)	(0.0897)		(0.2007)	(2.3623)
Yield	0.1777***	0.2463***	Sanction	0.4298	-0.1589
	(0.0510)	(0.0863)		(0.7093)	(0.4417)
			Threat	-0.0585	0.6515
				(0.2697)	(1.0602)

Index refers to the aggregate cooperation/non-cooperation measure. Each regression includes all those covariates present in the regression results shown in Table 3, including a constant, various dummies for membership in RTAs and time dummies. OECD sample includes additional controls for price level and ICRG Index. Variance-covariance matrix adjusted for serial correlation within country clusters. Robust standard errors in parentheses. *** significant at 1%, ** at 5% and * at 1% level.

between OECD countries (e.g., during the Iraq war).

5.5 Conclusions

In this paper, we argue that successful cross-border transactions in international trade require ex ante information on the trade partner's likelihood to behave according to standards of fairness. This information, being private by nature, may become public trough players' revealed behavior. We hypothesize both in our theoretical framework and in our empirical exercise that trade partners take into account the experiences of other players from their home country in the export country under scrutiny.

The innovation of this paper is to use an original data base extracted from Reuters Business Briefing and to apply it in a gravity-type framework. We construct aggregate indices of cooperative and non-cooperative behavior, using weights attached to a large array of different bilateral events proposed by Goldstein (1992). We purge this index so that it is independent from country sizes (reflected by GDP) and geographical distance and instrument it by its first two lags. This instrument turns out to work sufficiently well, both in terms of its correlation to the contemporaneous indices and its overall goodness of fit. We use two broad estimation strategies: (i) a Baier and Bergstrand (2001) type specification where we use price level data to proxy for countries' price indices, and (ii) a fixed effect estimation.

We find that cooperative behavior of importers induces trade flows while non-cooperative behavior deters them. However, the trade-reducing effect of non-cooperative behavior is stronger than the trade-creating effect of cooperative behavior. It also turns out that news about non-cooperative events do not affect bilateral trade volumes in the OECD sample, which we conjecture is due to the high overall degree of cooperation that should be expected in that specific subsample. In general, the results are particularly robust for cooperative events. The reason for this pattern may lie in the fact that our data features many more cooperative than non-cooperative events.

There are a couple of extensions that one might consider in future work. First, endogeneity issues loom large in our paper, even if we are confident that using lags does contribute towards a solution of the problem. However, other instrumentation strategies may be available. Second, the Reuters Business Briefing data assigns the day and even the hour of each event it records. Since bilateral trade data is available in monthly frequency at least for the last years, it seems promising to work with higher frequencies to better identify the time pattern of the news-trade nexus. Third, it would be highly desirable to test our theoretical argument on more disaggregated trade data, since the informational effects highlighted in the present should matter more strongly for differentiated goods as compared to homogeneous goods.

5.6 Appendix. Additional Tables

Table 5.7: Summary statistics - RBB Data

	Specification I				**Specification II**			
	Full sample		**OECD sample**		**Full sample**		**OECD sample**	
Event	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Cooperation Index	4.309799	23.11825	13.05835	46.35197	2.295607	15.31968	11.17856	42.04317
Agreement	.4494913	2.621801	1.60213	5.733813	.2261806	1.695114	1.343236	5.09897
Consultation	.9561197	5.524738	2.549969	10.06418	.5283613	3.706781	2.181306	9.079153
Endorsement	.1050282	.7461611	.3338033	1.415856	.0539487	.5100613	.2907468	1.320515
Grant	.0742073	.48743	.2196115	.8560961	.0422683	.3592879	.1924612	.8155532
Promise	.0177023	.193655	.0421366	.2982179	.0099124	.1437317	.0364021	.2734515
Reward	.0629063	.4655031	.1842105	.8848268	.0343014	.3234398	.1614608	.8153685
Yield	.0051763	.0900736	.0169173	.1680089	.0028554	.0660023	.0157351	.1628762
Conflict Index	.8651467	6.28505	2.561372	10.82717	.456142	4.345155	2.221606	9.840286
Accusation	.2081992	2.090813	.4821115	2.458699	.109297	1.428491	.4179897	2.253804
Complaint	.0334762	.38731	.1003759	.7307999	.0177181	.2844269	.0868013	.6696224
Demand	.2377635	1.995568	.7276316	3.617637	.1251968	1.436468	.6317755	3.340914
Demonstration	.0315776	.5432658	.0774279	.7606487	.0160574	.3832602	.0703382	.7555384
Denial	.013336	.1631959	.0411967	.2742729	.0069609	.1157263	.0351104	.2514788
Rejection	.209108	1.786102	.7136591	3.498854	.1081785	1.228302	.6124941	3.179005
Sanction	.0954895	1.04796	.2964756	1.813322	.0538385	.7628552	.2613668	1.656749
Threat	.0361968	.645696	.1224937	1.214101	.018895	.4599939	.1057304	1.165656
N	50,615		6,384		135,184		8,516	

Sample I: Regressions with fixed effects, Sample II: Regressions with controls for price level, corruption and RTAs

Indeces: Goldstein weighted aggregates

Table 5.8: Summary statistics

Event	Full sample Mean	Full sample Std. Dev.	OECD sample Mean	OECD sample Std. Dev.	Data source
ln Exports	16.37815	3.482576	19.93024	2.206921	IMF Direction of Trade Statistics
ln GDP Exporter	25.00215	1.970429	26.30838	1.489534	IMF International Financial Statistics
ln GDPImporter	24.87071	2.055084	26.32108	1.480834	
ln Population Exporter	2.874013	1.522181	2.760725	1.466934	
ln Population Importer	2.837264	1.506566	2.778317	1.460406	
ln Price level Exporter	3.703087	.7800766	4.244035	.6474992	Penn World Tables,
ln Price level Importer	3.684679	.785925	4.24311	.6469306	IMF International Financial Statistics[a]
ln Geographical Distance	8.620402	.8976228	8.08758	1.119735	CEPII, Database Distances
Contiguity	.0318088	.1754923	.058584	.2348628	
Common language	.1592018	.3658678	.0928885	.2902988	
Corruption Index Exporter	-3.681897	1.325956	-4.670829	1.168509	International Country Risk Guide (ICRG)
Corruption Index Importer	-3.62286	1.339505	-4.663793	1.174135	
Dummy: Similarity	.2372024	.4253716	.3795426	.4853111	Database on Political Institutions[b]
N	50,615		6,384		

[a]: PPP for 2000 from Penn World Tables and growth rates from IMF International Financial Statistics

[b]: Dummy, takes value 1 if chief executive party of same wing (right, left, center).

Table 5.9: ZIP and OLS results

	ZIP				OLS	
Variables	**Agreement**	**inflate**	**Accusation**	**inflate**	**Cooperation**	**Conflict**
ln GDP exporter	0.364***	-0.315***	0.0690**	-0.309***	0.460***	0.00637
	(0.021)	(0.019)	(0.035)	(0.023)	(0.037)	(0.0098)
ln GDP importer	0.361***	-0.340***	0.0397	-0.324***	0.522***	-0.00885
	(0.019)	(0.018)	(0.027)	(0.022)	(0.036)	(0.010)
ln Distance	-0.190***	0.683***	-0.140*	0.486***	-0.482***	-0.0146
	(0.028)	(0.032)	(0.082)	(0.048)	(0.093)	(0.030)
Conflict	0.0103***	-0.206***			2.272***	
	(0.00097)	(0.018)			(0.26)	
Cooperation			0.00367***	-0.0605***		0.213***
			(0.00068)	(0.0047)		(0.015)
Observations	243277	243277	243277	243277	243277	243277
R^2	.	.	.	.	0.52	0.50
Clusters	20865		20865		20865	20865
Chi^2	947.6		118.3			
Zero observations	230378		239914			
F-statistic					209.9	197.3
RSME					8.528	2.614

Each regression includes a constant, and time dummies (all not shown). Variance-covariance matrix adjusted for serial correlation within country clusters. Robust standard errors in parentheses. *** significant at 1%, ** at 5% and * at 1% level.

Table 5.10: **Coding scheme. Cooperative events**

Description	**Goldstein**	**Description**	**Goldstein**
Agreement		**Promise**	
Agree	4.8	Promise	4.7
Agree or accept	3	Promise policy	4.5
Agree to	3	Promise material	5.2
Agree to mediation	3	Promise economic	5.2
Agree to negotiate	3	Promise military	5.2
Agree to	3	Promise	5.2
Collaborate	4.8	Assure	2.8
Consult		Promise to mediate	4.7
Consult	1.5	**Reward**	
Discussion	1	Reward	7
Mediate talks	1	Extend economic	7.4
Engage in	1	Extend military	8.3
Travel to meet	1.9	Extend	7.6
Host a meeting	2.8	Rally support	7.6
Endorsement		**Yield**	
Endorse	3.5	Yield	1.1
Praise	3.4	Yield to order	.6
Empathize	3.4	Yield position	.6
Apologize	2.2		
Forgive	3.5		
Ratify a decision	3.5		
Grant			
Grant	2.2		
Extend invitation	2.5		
Provide shelter	2.2		
Evacuate victims	2.2		
Improve relations	5.4		
Ease sanctions	2.9		
Observe truce	2.2		
Relax censorship	2.2		
Relax	2.2		
Relax curfew	2.2		
Ease economic	2.2		
Release or return	1.9		
Return, release	1.9		
Return, release	1.9		

Table 5.11: **Coding scheme. Non-Cooperative events**

Description	**Goldstein**	**Description**	**Goldstein**
Accusation		Political flight	4
Accuse	2.8	Disclose	4
Criticize or	2.2	Break law	4
Complaint		**Sanction**	
Complain	2.4	Sanction	4.5
Informally	2.4	Armed force	7
Formally complain	2.4	Reduce routine	2.2
Denial		Reduce or stop aid	5.6
Deny	1	Reduce or stop	5.6
Demand		Reduce or stop	5.6
Demand	4.9	Reduce or stop	5.6
Demand information	4.9	Reduce or stop	5.6
Investigate human	4.9	Halt discussions	3.8
Demand policy	4.9	Halt negotiation	3.8
Demand aid	4.9	Halt mediation	3.8
Demand mediation	4.9	Break relations	7
Demand withdrawal	4.9	Strikes and	4.5
Demand ceasefire	4.9	**Threat**	
Demand meeting	4.9	Threaten	6.4
Demand rights	4.9	Non-specific	4.4
Demonstration		Sanctions threat	5.8
Demonstrate	6.9	Threaten to halt	5.8
Protest	5.5	Threaten to halt	5.8
Protest	5.2	Threaten to reduce	5.8
Protest procession	5.2	Threaten to	5.8
Protest defacement	5.2	Threaten to reduce	5.8
Protest altruism	6.9	Give ultimatum	6.9
Rejection		Other physical	6.4
Reject	4		
Reject proposal	4		
Reject ceasefire	4		
Reject	4		
Reject settlement	4		
Reject request for	4		
Reject proposal to	4		
Reject mediation	4		
Refuse to allow	4		
Impose	4		
Impose censorship	4		
Veto	4		
Defy norms	4		

Chapter 6

Trade Intermediation and the Organization of Exporters[1]

6.1 Introduction

The international business literature (e.g., Peng and Ilinitch, 1998) stresses that firms typically require an own sales affiliate or a trade intermediary in the foreign country to become successful exporters. The optimal organizational choice between these two major *export modes* is an important issue for firms' internationalization strategies. However, conventional trade models assume that exporters sell directly to foreign end-clients.

While that assumption may be an innocuous assumption for many important questions, it is plainly wrong empirically. And it bars a more profound understanding of international trade costs broadly defined, which are estimated to be substantial despite recent progress in transportation and communication technologies.[2]

In this paper, we suggest a simple theoretical framework in which exporters face a choice of *how to export* to foreign markets. That is essentially an organizational choice as domestic producers can either provide distribution services abroad through an own foreign sales affiliate, or through a specialized firm: a trade intermediary. In our model, due to incomplete cross-border contracts, intermediation entails a distortion that leads to lower export revenues. This is a disadvantage compared to the use of an own sales affiliate; however, intermediaries make capital investments of producers in the foreign country redundant,

[1] This Chapter is based on a working paper, see Felbermayr and Jung (2008c). The concept for the paper was developed jointly, theoretical analysis and writing were equally shared, and the empirical analysis was carried out by the author of this thesis.

[2] For example, Anderson and van Wincoop (2004) report that retail and wholesale distribution costs are equivalent to an *ad valorem* tax of 55 percent, thereby dwarfing other types of trade costs.

thereby offering savings in fixed distribution costs.

Our paper innovates along three lines: First, we cast the above choice of export modes in a model where producers differ with respect to their *competitive advantage* and derive a sorting pattern of monopolistic firms over different export modes. Second, as the organizational choice of producers is built into a multi-country *general equilibrium* trade model with trade-cost asymmetries, we are able to derive empirical predictions that take aggregate income and price effects in due account. Third, we use US census data to provide a rough *empirical check* of our results.

We focus on the *endogenous emergence* of trade intermediaries as important institutions in the operation of real-world international business. Trade intermediaries enjoy easier access to foreign markets due to better local knowledge and the exploitation of economies of scope. However, new advances in the literature on the boundaries of the firm (e.g., Antras and Helpman, 2004) stress the lack of enforceable contracts in international transactions. The interaction with an intermediary exposes the producer to a hold-up problem. Prices and quantities are determined in a game between producers and intermediaries: the optimal response of the producer is to restrict output for the export market, which drives up consumer prices. Hence, contractual imperfections act differently than standard iceberg-type trade costs, as they do not induce loss in transit. The trade-off between fixed-cost savings and lower revenue pins down the producers' optimal organizational mode of exporting.

Facing a hold-up problem, producers may wish to internalize sales activities by setting up a foreign sales affiliate. Internalization forgoes the fixed-cost savings available with intermediation, but avoids relationship-specific distortions. We embed this trade-off in a somewhat generalized Melitz (2003) model, where firms differ with respect to the idiosyncratic components of variable distribution costs or preferences as well as with respect to their labor productivity. We derive an interesting sorting pattern: firms with low distribution costs, strong brand reputation, and high productivity internalize foreign sales activities, while those with medium realizations of those variables prefer to use trade intermediaries. The relevant firm characteristics correlate with firm size, so that the paper predicts selection of firms along their sizes.

Empirical relevance. There is massive systematic and anecdotal evidence on the importance of trade intermediaries. Peng, Zhou and York (2006) find for the U.S. that more than 45% of export sales in 68 out of 97 product categories are handled by export intermediaries (see their Tables 4 and 5 for 1998, pp. 296f).[3] Our model also stresses the role of foreign wholesale affiliates. Buch, Kleinert, Lipponer and Toubal (2005) report that about 39% of all German subsidiaries abroad are active as wholesellers. They make up a substantial share – about 30% – of the total stock of German foreign direct investment (FDI). In the proposed model, to the extent that exports are channeled through sales affiliates, FDI and trade are complements. This may rationalize the finding discussed by Neary (2008) that bilateral FDI stocks seem to decrease with bilateral distance, much as bilateral trade does. Note that Neary's observation is in conflict with the standard *proximity-concentration* view of FDI, where trade and investment are substitutes.

Bernard, Jensen, and Schott (2008) show that about 13% of all U.S. exporters sell to own foreign affiliates (related parties), and that their combined export market share is about 30%. It follows that the average firm exporting to own wholesale affiliates has larger export revenue than the average firm selling through an alternative channel. This pattern is fully consistent with the evidence (e.g., surveyed by Helpman, 2006) that only the largest firms engage in FDI. Business literature (e.g., Ellis, 2000) suggests that exports that are not channeled via intermediaries or an own foreign subsidiary make up only a negligible share of total exports. Those 'direct' exports typically stem from unsolicited orders, often through direct contact of a buyer via phone or mail. This is is an opportunity that almost any producer with some international visibility has at some point in time; hence, the number of concerned firms may be fairly large while the role of direct exports for total trade is small. Bringing the available evidence together, it follows that firms sort according to their size in line with the pattern predicted in this paper.[4] To the best of our knowledge, the present paper is the first to theoretically explain this sorting pattern.

Besides the predicted sorting pattern, our framework has additional testable implica-

[3]Trabold (2002) provides evidence for France.

[4]Available firm-level data usually does not provide information on the mode of serving a foreign market ('directly', through an intermediary, or via an own affiliate). Hence, direct evidence for the proposed sorting pattern is (currently) still impossible to obtain.

tions. Thanks to the general equilibrium nature of our model, we can derive structural relationships that can be tested econometrically in a consistent way. First, the prevalence of sales through trade intermediaries relative to sales through own affiliates does neither depend on geographical distance between two countries nor on their respective market sizes. Second, relative prevalence decreases in the strength of contractual imperfections (which may be good/sector-specific) but it increases in the (country-specific) risk of expropriation in the foreign country. Third, relative prevalence increases as firms become more homogeneous in terms of their underlying characteristics (productivity, quality, tradability of goods). Fourth, when systematic trade costs go up, both the stock (and flow of) wholesale FDI between two countries and the volume of bilateral trade (in both modes) fall. Hence, in this scenario, trade and FDI appear as complements. Using the ratio of related-party over non-related-party trade reported by the US Census at the industry level as a proxy for the relative prevalence of intermediation we find support in favor of our hypotheses.

Related literature. Our work is related to at least three important strands of literature. First, as in Grossman and Helpman (2002) or Antras and Helpman (2004, 2008) we allow for incomplete contracts to affect the boundaries of firms.[5] Bernard, Jensen, Redding, and Schott (2008) provide empirical evidence that products' revealed contractability plays a role in explaining the intra-firm share of imports.[6] We set up a theoretical model to reproduce the set of stylized facts discussed above, and assume that contractual imperfections arise when legal entities of two different countries (a producer and the foreign trade intermediary) interact, but that the relation between wholesale agents and retailers is free from frictions. There is an interesting literature that analyzes falling trade costs trade when producers and retailers interact strategically (see, e.g., Raff and Schmitt, 2005 or 2006). We abstract on strategic issues and rather focus on the endogenous emergence of trade intermediaries, the sorting of producers across export modes and the role of contractual frictions as determinants of trade costs.

Second, a growing number of papers model the distribution to foreign markets in more

[5] Whereas their focus is on a sourcing decision which involves the location of input production, we analyze the pattern of sourcing distribution services.

[6] The paper is marked "Preliminary".

detail. Rauch (1999) or Krautheim (2007a) analyze the role of networks; Arkolakis (2008) models exporters' marketing decisions on foreign markets. However, while Spulber (1996) provides a detailed discussion of the importance of intermediation,[7] trade economists have not paid much attention to this issue. There are, however, exceptions. Feenstra and Hanson (2004) study the role of Hong-Kong as a center of trade intermediation. Rauch and Watson (2003) model the emergence of network intermediation in international trade. More related to our approach, Schröder, Trabold and Trübswetter (2005) discuss trade intermediation in a simple two-country monopolistic competition model of international trade. However, in their partial equilibrium model, firms are identical, and there is no endogenous choice of export mode.

Third, a number of recent papers discusses the endogenous sorting of firms into different modes of serving foreign markets. In Helpman, Melitz, and Yeaple (2004) [henceforth: HMY] firms either produce locally and export to a foreign market, or they engage in horizontal FDI and produce abroad. Krautheim (2007b) develops an interesting generalization of that model, allowing firms to use an additional mode of selling to the foreign market, namely via *export supporting FDI*. His model allows to address the facts discussed by Neary (2008), but does not address trade intermediation. More importantly, he exogenously imposes the cost structure that drives his result, whereas in the present paper variable revenue from exporting via intermediaries is determined endogenously.

Structure of the paper. The remainder of the paper is organized as follows. Section 7.3 introduces the model and solves the game between the trade intermediary and the producer. Section 6.3 derives the key propositions of the paper: it shows how firms sort into different export modes according to their attributes and derives predictions on the relative prevalence of either export modes and the trade-FDI relationship in general equilibrium. Section 6.4 provides tentative empirical evidence, and Section 7.6 concludes. Proofs of our results, intermediate steps of calculations, and a number of tables are contained in the Appendix.

[7] Various models of intermediation are presented in Spulber (1999).

6.2 Model setup

In this section we describe a model with heterogeneous firms akin to Melitz (2003) in which we introduce the endogenous emergence of trade intermediaries. Besides its focus on firms' choice of export mode, our model differs from existing treatments in that it allows for a broader characterization of firm heterogeneity. Our general equilibrium approach has the advantage that it generates structural relationships between observable variables, thereby allowing for consistent econometric validation.

The world consists of N countries, indexed $j = 1, ..., N$, who may differ according to the size of their labor forces. In each country, heterogeneous firms produce varieties of a differentiated good and interact under conditions of monopolistic competition. We allow for exogenous firm turnover, so that in a stationary environment at each instant of time a measure $\bar{\delta} > 0$ of firms dies and enters. Firm death is the only source of discounting.

6.2.1 Demand structure

Each country j is populated by a representative household who inelastically supplies L_j units of labor to a perfectly competitive labor market. Preferences are a CES aggregate of differentiated goods, each indexed by ω:

$$U_j = \left[\int_{\omega \in \Omega_j} \left[\zeta(\omega) \, x_{ij}(\omega) \right]^{\rho} \mathrm{d}\omega \right]^{1/\rho}. \tag{6.1}$$

The parameter $\rho \in (0,1)$ describes the degree of substitutability between any pair of varieties. Ω_j is the set of available varieties in country j. The quantity $x_{ij}(\omega)$ denotes consumption of a variety produced in country $i, i = 1, ..., N$. Our specification slightly generalizes the standard CES case in that it adds the parameter $\zeta(\omega) \geq 0$ which captures the brand reputation of variety ω as perceived by the household.[8] The larger $\zeta(\omega)$, the bigger is the contribution of variety ω to overall utility.[9]

Each variety is produced by a single firm. Despite the existence of operational profits

[8] Combes, Lafourcade, and Mayer (2005) introduce a similar weighting factor in their representation of utility.

[9] In principle, our setting allows to read equation (7.1) as a CES production function of a competitive final output good producer. Then, we study trade in inputs rather than in final goods. The predictions of the model do not hinge on the interpretation. This conceptual flexibility facilitates the empirical exercise of Section 4, since the data do not allow to dissect trade in final goods from trade in inputs.

of successful firms, ex ante expected profits are driven to zero by free entry. In equilibrium, aggregate operational profits are exactly matched by firms' total setup costs. Thus, labor is the only source of income, and the budget constraint reads

$$w_j L_j \geq \int_{\omega \in \Omega_j} p_{ij}(\omega) x_{ij}(\omega) \, \mathrm{d}\omega, \tag{6.2}$$

where w_j denotes the wage rate that is determined endogenously in equilibrium.

Maximizing (7.1) subject to (7.2), we find the following demand function for a variety ω from country i

$$x_{ij}(\omega) = H_j \frac{\zeta(\omega)^{\sigma-1}}{p_{ij}(\omega)^{\sigma}}, \tag{6.3}$$

where $H_j \equiv w_j L_j P_j^{\sigma-1}$ and $P_j = \left(\int_0^{n_j} [p_{ij}(\omega)/\zeta(\omega)]^{1-\sigma} d\omega\right)^{1/(1-\sigma)}$. P_j is the price index dual to (7.1), n_j is the measure of the set Ω_j and $\sigma \equiv 1/(1-\rho) > 1$ is the elasticity of substitution between varieties.

6.2.2 Product heterogeneity and exporting via own wholesale affiliates

Monopolistically competitive producers differ with respect to a vector of characteristics $\{\zeta(\omega), \tau(\omega), a(\omega)\}$, where $\zeta(\omega)$ is the taste parameter introduced above, $a(\omega) > 0$ denotes the labor input requirement for producing one unit of variety ω, and $\tau(\omega) \geq 1$ refers to variety-specific variable distribution costs of the iceberg type, which measure the ease at which a variety is brought to the consumer (marketability). Realistically, we assume that this cost occurs regardless of whether a good is traded internationally or not. However, in international transactions, total variable trade costs are $\tau_{ij}(\omega) = \bar{\tau}_{ij}\tau(\omega)$, where $\bar{\tau}_{ij} \geq 1$ accounts for transportation costs from country i to country j and may be thought of as a function of distance. We refer to $\bar{\tau}_{ij}$ as to the *systematic* component of trade costs, and of $\tau(\omega)$ as the *idiosyncratic* component.[10]

Firm ω's variable cost function in country i is given by $c_i(\omega) = y(\omega) a(\omega) w_i$ where $y(\omega)$ is the quantity of output. Regarding their cost structure, firms do not differ across countries. We map the vector of firm characteristics $\{\zeta(\omega), \tau(\omega), a(\omega)\}$ into a scalar measure of effective firm-level productivity $\Phi(\omega) \equiv \zeta(\omega)/[a(\omega)\tau(\omega)]$. It turns out that $Q \equiv \Phi^{\sigma-1}$ is a measure of *competitive advantage* which fully characterizes firm behavior.

[10] Bergin and Glick (2007) also discuss variety-specific trade costs.

Following the structure of the entry process introduced by Hopenhayn (1992) and simplified in Melitz (2003), prospective entrants are uncertain about their respective values of Φ. Only after entry, which requires sinking the cost f^E, is Φ revealed and remains constant afterwards. We assume that Φ follows the Pareto distribution. More precisely, we let the c.d.f. be $G(\Phi) = 1 - \Phi^{-k}$, with a shape parameter $k > \max\{2, \sigma - 1\}$ and the support $[1, +\infty)$.[11] Note that we need not restrict in any way the stochastic processes that govern the components of $\Phi(\omega)$.

Along with variable distribution costs $\tau(\omega)$, there are also fixed distribution costs. These costs are associated to warehousing, the maintenance of customer relations, or regulatory burdens. Without loss of generality, given perfect capital markets, we can express investment costs as flow costs. Flow fixed distribution costs are expressed in terms of labor and are given by $f_j = f w_j$, where f is the labor requirement that is constant over all countries. We assume, that a firm from country i has to pay f_i when selling to its home market, but that the cost of an own foreign representation is given by $f_{ij} = \phi_{ij} f_j$, with $\phi_{ij} > 1$ for $i \neq j$, and $\phi_{ii} = 1$, so that firms' fixed distribution costs in the foreign country are higher than in the home economy. In contrast, trade intermediaries are assumed to originate in country j so that they enjoy cheaper access to foreign markets than foreign producers. Whenever $i \neq j$, we call f_{ij} wholesale FDI (henceforth: WFDI).[12]

The fact that producers face higher fixed distribution costs abroad may have two reasons. First, trade costs may simply have a firm-specific fixed component which is larger in foreign markets due to additional costs associated to linguistic, legal or informational issues. Second, ϕ_{ij} may represent the higher foreign expropriation risk (e.g., because of ill-defined property rights). To see this, let δ_{ij} denote the Poisson rate of expropriation and assume that $\delta_{ii} = 0$ for the sake of simplicity. Then, ϕ_{ij} would be equal to $\left(\bar{\delta} + \delta_{ij}\right)/\bar{\delta}$

[11] The Pareto assumption has been made in a large number of related papers (e.g., Helpman, Melitz, and Yeaple (2004), Chaney (2008), Helpman, Melitz, and Rubinstein (2008), Bernard, Redding, and Schott (2006)). The Pareto allows for closed form solutions. The assumption $k > 2$ makes sure that the variance of the productivity distribution is well-defined, and $k > \sigma - 1$ guarantees that the equilibrium distribution of firm sizes has a finite mean. The shape parameter can be interpreted as an inverse measure of productivity dispersion.

[12] In principle, the sales representative could also be located domestically. However, our preferred interpretation allows to view f_{ij} as wholesale FDI. Krautheim (2007b) uses the term *export-supporting FDI* instead of WFDI. Essentially, this is just a reinterpretation of the fixed costs of exporting in the original Melitz (2003) model.

which is a strictly increasing function of δ_{ij}. Hence, expropriation risk works just as a higher depreciation rate on foreign assets.

We want to understand how differences in terms of competitive advantage Q across producers determine their choice of foreign market entry mode: through wholly owned foreign sales affiliates or through trade intermediaries. For that purpose, we first briefly show how domestic profits and profits achieved through foreign sales affiliates depend on Q. Discussion of profits through intermediation is less standard and discussed in more detail in the next section.

Domestic sales. Operating profits from *domestic sales* are $\tau(\omega) \cdot H_i [\tau(\omega) p(\omega)]^{-\sigma} \zeta(\omega)^{\sigma-1} \cdot [p(\omega) - a(\omega) w_i] - f_i$. The presence of the term $\tau(\omega)$ reflects the presence of non-zero variable distribution costs also for domestic sales. The term $H_i [\tau(\omega) p(\omega)]^{-\sigma} \zeta(\omega)^{\sigma-1}$ describes a household's level of demand for a variety ω. The term $[p(\omega) - a(\omega) w_i]$ refers to the per unit margin of the price over marginal cost. Monopolistic producers in country i set their ex factory price as $p_i(\omega) = a(\omega) w_i / \rho$ so that domestic profits per period are

$$\pi_i^D(Q) = w_i^{1-\sigma} B_i Q - f_i. \tag{6.4}$$

Domestic profits are an increasing function of competitive advantage Q. The components of Q (marketability, brand reputation, and productivity) do not matter separately for profits. The term $B_i \equiv (1-\rho) H_i \rho^{\sigma-1}$ is an aggregate magnitude which captures the size of the market and is taken as exogenous by producers. Domestic profits increase in the size of the home market B_i reflecting larger demand at constant profit margins; obviously, they fall in fixed costs of production, f, and in the domestic wage rate w_i.

Foreign sales through own affiliates. The monopolist generates non-negative profits from *exporting via an own affiliate*, if export revenues suffice to cover additional variable production costs and the annuitized costs of foreign investment $\phi_{ij} f_j$.[13] Profits from exporting through an own sales affiliate are $\tau_{ij}(\omega) \cdot H_j [\tau_{ij}(\omega) p(\omega)]^{-\sigma} \zeta(\omega)^{\sigma-1} \cdot [p(\omega) - a(\omega) w_i] -$

[13] Recall the assumption of perfect capital markets.

$\phi_{ij} f_i$. Using the monopolist's optimal pricing rule, this gives

$$\pi_{ij}^{F}(Q) = (w_i \bar{\tau}_{ij})^{1-\sigma} B_j Q - \phi_{ij} f_j, \tag{6.5}$$

where the systematic part of trade costs $\bar{\tau}_{ij}^{1-\sigma}$ appears as an additional determinant of variable profits, along with the foreign measure of market size B_j and the costs of investing abroad, $\phi_{ij} f_j$. Again, profits increase in the degree of competitive advantage Q and market size B_j; they fall in effective unit costs $w_i \bar{\tau}_{ij}$, the expropriation risk ϕ_{ij} and the fixed costs of maintaining the foreign distribution network f_j.

6.2.3 Trade intermediation

Assumptions. An intermediary is "*...an economic agent who purchases from suppliers for resale or who helps sellers and buyers to meet and transact*" (Spulber, 1996). We view trade intermediaries as wholesale agents that facilitate transactions between producers and consumers from different countries. Trade intermediaries benefit from *economies of scale* since there are fixed distribution costs. Being incorporated in the foreign country, they have the same fixed distribution costs than producers in the same country, $f_j = f w_j$.[14] We do not explicitly model a retail sector; our assumption of variable trade costs accruing also for domestic sales capturing in a parsimonious way the cost of retailing when there are no specific contractual or strategic interactions between wholesellers and retailers (which we rule out in this paper).

Our model accommodates trade intermediaries that have *diversified product portfolios*. Under general circumstances, the pricing and the product range choice of intermediaries interact in a complicated way due to a cannibalization effect. However, under monopolistic competition, intermediaries do not internalize the effect of an additional variety on demand of the other varieties, such that pricing and product-range decisions are independent; see Bernard, Redding, and Schott (2006) for a related model of multi-product producers. We may also reconcile our model with *economies of scope*. When fixed costs of distribution depend on the number of varieties sold, intermediaries determine their product range such

[14] The intermediary's specific knowledge could also translate into lower variable (distribution) costs. However, the largest portion of variable distribution costs such as transportation services, taxes, etc. are the same across export modes.

that those costs are minimized. Assuming an interior solution to this problem, we may think of f as the minimum fixed distribution cost. As intermediaries are identical in our model, they all share the same fixed costs.[15]

Finally, we assume that producers and intermediaries cannot write enforceable *cross-country* contracts on quantities and prices and that the variety to be exchanged features some export market specificity. This might be the case if the product has to meet some specific technical standards that prevent it from being fully 'recycled'.[16] The lack of *ex ante* contracts exposes the producer to potential hold-up: the intermediary can deny the order *ex post*, i.e., after production has taken place. This assumption is crucial in that it provides an endogenous rationale for lower variable revenues when the producer opts for the intermediated export mode. Variants of this assumption have been used by Helpman and Grossman (2002) or in Antras and Helpman (2004, 2008) in the context of vertical relations between final goods and intermediate inputs producers (outsourcing).

The game between producers and intermediaries. As in Antras and Helpman (2004), there is an infinitely elastic supply of intermediaries in every country. Each producer P who finds it optimal to search for a trade intermediary M, makes a take-it-or-leave-it offer, which specifies an upfront-fee for participation $T(\omega)$ in the relationship that has to be paid by M. This fee can be positive or negative, and may be interpreted as a franchising fee paid by M to P or as a down-payment of P to M towards financing fixed foreign distribution costs. There is full information on product characteristics ω, so that prospective intermediaries would know that a variety offered by some producer is already sold by another intermediary. In that case, both intermediaries would see their operative profits driven to zero by Bertrand competition and would thereby not be able to recover T. It follows that all producer-dealer relationships in equilibrium involve *exclusive dealership* arrangements in that each producer is matched to at most one intermediary in every market.

With the supply of M infinitely elastic, M's profits from the relationship net of the

[15] In Felbermayr and Jung (2008a) we study fixed market access costs which depend on the tightness of the matching market between producers and intermediaries.

[16] This 'recycling' process may be, of course, a metaphor for many things: sales in the foreign market may require market-specific adjustments, so that selling a shipment elsewhere requires undoing these changes; one could also think about a situation where, in case of disagreement, a shipment needs to be shipped back from the foreign country to the producer, thereby causing additional transportation costs.

participation fee in equilibrium are equal to its outside option, which we have set zero. Hence, $T(\omega)$ will indeed differ across varieties: the higher the competitive advantage of a variety, the larger the fee that the producer can extract from the trade intermediary. However, while perfect competition for producers leaves trade intermediaries without rents ex post, they can still hold up the producers. Due to the lack of enforceable contracts, the producer cannot be sure to receive adequate payment for the output delivered to the trade intermediary. The latter can refuse delivery until the price is low enough. We assume that the countervailing incentives of producers and intermediaries are sorted out via the usual asymmetric Nash bargaining process, where $\bar{\beta}_{ij} \in [0,1]$ is the bargaining power of a producer from country i with an intermediary located in country j. At the bargaining stage, the producer is particularly vulnerable since production costs are sunk at the time of bargaining. If bargaining fails, the producer can recycle the goods that were meant for exports, thereby partly recovering a fraction $\lambda_{ij} \in [0,1]$ of the inputs used in production.[17]

We may summarize the sequencing of the game between the trader M and the producer P. First, the producer P effectively auctions an exclusive dealership relationship with a trade intermediary. Second, if some M has accepted the offer, P decides about the quantity $\tau_{ij}(\omega)\, x_{ij}^{M}(\omega)$ to produce for the purpose of exports.[18] Finally, P delivers the goods to M, M sells the goods, and P and M bargain about sharing of revenues (and, thereby, implicitly about a transaction price).

As usual, the game is solved by backward induction. The joint surplus generated on the foreign market is given by

$$J_{ij}(\omega) = p_{ij}^{M}(\omega)\, x_{ij}^{M}\left[p_{ij}^{M}(\omega)\right] - \tilde{\pi}_{ij}^{P}(\omega) - f_j, \tag{6.6}$$

where $x_{ij}^{M}\left[p_{ij}^{M}(\omega)\right]$ is the level of foreign demand at a c.i.f. price $p_{ij}^{M}(\omega)$ and $f_j = f w_j$ is fixed foreign costs of distribution incurred by M.

The producer's outside option $\tilde{\pi}_{ij}^{P}(\omega)$ is the amount of the numeraire input that firm ω

[17] Note that λ_{ij} measures how specific the product is to the respective export market.

[18] $x_{ij}^{M}(\omega)$ is the quantity demanded by foreign consumers, which implies the production of $\tau_{ij}(\omega)\, x_{ij}^{M}(\omega)$ units due to loss in transit.

can recover when bargaining fails

$$\tilde{\pi}_{ij}^{P}(\omega) = \lambda_{ij}\tau_{ij}(\omega)\, x_{ij}^{M}(\omega)\, a(\omega)\, w_i, \tag{6.7}$$

where $\tau_{ij}(\omega)\, x_{ij}^{M}(\omega)$ is the amount of production required to deliver the quantity x_{ij}^{M} to the foreign market. If $\lambda_{ij} = 0$, there is no alternative use for the goods delivered to the foreign market; if $\lambda_{ij} = 1$, production can be entirely and costlessly unwinded.

The Nash solution of the bargaining problem between the producer and the intermediary requires that M receives a pay-off $\left(1 - \bar{\beta}_{ij}\right) J_{ij}(\omega)$, while the producer gets $\bar{\beta}_{ij} J_{ij}(\omega) + \tilde{\pi}_{ij}^{P}(\omega)$. Predicting its share of the surplus at the bargaining stage, the producer chooses the optimal quantity to supply to the intermediary. She solves

$$\max_{x_{ij}^{M}(\omega)} \bar{\beta}_{ij} J_{ij}(\omega) + \tilde{\pi}_{ij}^{P}(\omega) - x_{ij}^{M}(\omega)\, \tau_{ij}(\omega)\, a(\omega)\, w_i \tag{6.8}$$

subject to the demand function (7.3). The quantity choice of the producer finally determines the price that the consumer in the foreign country ends up paying. The following lemma states that price.

Lemma 5 (Pricing behavior)**.** *The c.i.f. price charged for imports from country i into the foreign market j is given by*

$$p_{ij}^{M}(\omega) = \frac{\tau_{ij}(\omega)\, a(\omega)\, w_i}{\beta_{ij}\rho}, \tag{6.9}$$

where $\bar{\beta}_{ij} \leq \beta_{ij} = \beta_{ij}\left(\bar{\beta}_{ij}, \lambda_{ij}\right) \equiv \bar{\beta}_{ij} / \left[1 - \lambda_{ij}\left(1 - \bar{\beta}_{ij}\right)\right] \leq 1.$

The foreign price is determined as effective marginal costs $\tau_{ij}(\omega)\, a(\omega)\, w_i$ multiplied by a markup $1/(\beta_{ij}\rho) \geq 1$. The markup $1/\rho$ usually arises in a model with monopolistic competition and CES preferences. However, it is magnified by an additional factor $1/\beta_{ij}$ that arises due to the export market specificity of the product and lack of enforceable contracts, and that is endogenously pinned down by the parameters governing the bargaining process and by the ease at which products can be recycled. At the bargaining stage, the producer appropriates only a share $\bar{\beta}_{ij}$ of the surplus, and therefore optimally restricts the output below the level that would be optimal without intermediation.

If the intermediary has no clout in the bargaining stage, i.e., if $\bar{\beta}_{ij} = 1$, or if the producer can recycle the output meant for exports at no costs, i.e., if $\lambda_{ij} = 1$, then the additional markup vanishes, i.e., $1/\beta_{ij} = 1$.[19] In the case where output is totally specialized for the respective foreign market, i.e., if $\lambda_{ij} = 0$, the additional markup factor is only driven by the bargaining power, $1/\beta_{ij} = 1/\bar{\beta}_{ij}$. In the limiting case where the producer has no bargaining power, we have $\beta_{ij} \to 0$ regardless of the recycling rate. Moreover, $\partial\beta_{ij}\left(\bar{\beta}_{ij}, \lambda_{ij}\right)/\partial\bar{\beta}_{ij} > 0$ for all $\lambda_{ij} \in [0, 1)$ and $\partial\beta_{ij}\left(\bar{\beta}_{ij}, \lambda_{ij}\right)/\partial\lambda_{ij} > 0$ for all $\bar{\beta}_{ij} \in (0, 1)$.

If λ_{ij} and $\bar{\beta}_{ij}$ both lie strictly below unity, then $p_{ij}^M(\omega) > p_{ij}^F(\omega)$. Grossman and Helpman (2002) use a similar setup in the context of outsourcing with homogeneous firms. They relate the pricing rule (6.9) to the *double marginalization* problem that appears in vertical relationships of monopolistic firms. Both, higher trade costs $\bar{\tau}_{ij}$ and contractual frictions imply a higher consumer price. However, there is a crucial difference between iceberg-type trade costs and the effect of frictions $1/\beta_{ij}$. The former drives up the c.i.f. price as the delivery of a good to a foreign market requires the use of specific services which require resources in proportion to the price of the good. In contrast, contractual frictions drive up the c.i.f. price because producers optimally reduce supply, thereby moving up the demand schedule. Trade costs are a technological feature while the double marginalization phenomenon is due to imperfect markets.

Finally, potential intermediaries compete for contracts with producers, so that they end up bidding their entire ex post profits $\left(1-\bar{\beta}_{ij}\right)J_{ij}(\omega)$ as participation fees $T(\omega)$. The profits that a producer P makes on the foreign market using an intermediary are given by the optimal value of (6.8) plus the participation fee $T(\omega)$ that the producer receives. The producer's pay-off from the bargaining stage, plus income from the participation fee, minus variable production costs, all evaluated at the optimal price $p_{ij}^M(\omega)$, give her total profit from exporting via a trade intermediary as

$$\pi_{ij}^M(Q) = \left(\frac{w_i\bar{\tau}_{ij}}{\tilde{\beta}_{ij}}\right)^{1-\sigma} B_j Q - f_j. \tag{6.10}$$

Given our assumptions, we can replace the firm index ω with Q. The term $\tilde{\beta}_{ij} = \tilde{\beta}_{ij}\left(\beta_{ij}, \sigma\right) \equiv \left[\beta_{ij} + \left(1-\beta_{ij}\right)\sigma\right]^{\frac{1}{\sigma-1}}\beta_{ij} \in \left[\beta_{ij}, 1\right]$ is endogenously determined as a function of bargaining parameters $\beta_{ij}\left(\bar{\beta}_{ij}, \lambda_{ij}\right)$ and the elasticity of substitution σ. We have $\tilde{\beta}_{ij}(0, \sigma) = 0$,

[19] Note that domestic sales are nested by our model of intermediation with $\bar{\tau}_{ii} = \lambda_{ii} = 1$, or alternatively (and less realistically), $\bar{\beta}_{ii} = 1$.

$\tilde{\beta}_{ij}(1,\sigma) = 1$, $\lim_{\sigma\to 1}\tilde{\beta}_{ij}\left(\beta_{ij},\sigma\right) = \beta_{ij}e^{1-\beta_{ij}}$, and $\lim_{\sigma\to+\infty}\tilde{\beta}_{ij}\left(\beta_{ij},\sigma\right) = \beta_{ij}$. Moreover, $\partial\tilde{\beta}_{ij}\left(\beta_{ij},\sigma\right)/\partial\beta_{ij} > 0$ for all $\beta_{ij} \in [0,1)$ and $\sigma > 1$, and $\partial\tilde{\beta}_{ij}\left(\beta_{ij},\sigma\right)/\partial\sigma < 0$.[20]

Hence, incomplete contracts reduce the slope of the profit function $\pi_{ij}^{M}(Q)$ in a similar way than an increase in iceberg trade costs $\bar{\tau}_{ij}$ would. For given Q, the variable component of profits is always smaller when the producer chooses a trade intermediary than when the producer establishes an own wholesale affiliate.

Despite the fact that the producer does not directly lay out the fixed cost expenditure f_j in the foreign market, those costs are nevertheless entirely deducted from the producer's profit. This is due to the fact that the producer extracts all profits from the intermediary when setting the participation fee T. Hence, fixed distribution costs are fully rolled-over from the intermediary to the producer.

6.3 The choice of export modes

6.3.1 Sorting of firms

Firms partition endogenously into different modes along their degree of competitive advantage. The weakest firms do not even take up domestic production as they generate insufficient revenue to cover fixed domestic distribution costs f_i. The firm that is exactly indifferent between serving the domestic market or not is identified by the condition $Q_i^D = w_i^{\sigma-1}f_i/B_i$. Firms may export and do so using different export modes. The producer that is indifferent between exporting through a trade intermediary and selling on the domestic market only is given by the condition $\pi_{ij}^{M}\left(Q_{ij}^{M}\right) = 0$, which gives

$$Q_{ij}^{M} = \left(\frac{w_i\bar{\tau}_{ij}}{\tilde{\beta}_{ij}}\right)^{\sigma-1} f_j B_j^{-1}. \tag{6.11}$$

Finally, the producer with competitive advantage Q_{ij}^{F} achieves identical profits from serving the foreign market in either export modes: $\pi_{ij}^{M}\left(Q_{ij}^{F}\right) = \pi_{ij}^{F}\left(Q_{ij}^{F}\right)$. This indifference condition translates into

$$Q_{ij}^{F} = \left(w_i\bar{\tau}_{ij}\right)^{\sigma-1}\left(\frac{\phi_{ij}-1}{1-\tilde{\beta}_{ij}^{\sigma-1}}\right) f_j B_j^{-1}. \tag{6.12}$$

[20] The latter follows from $\frac{x-1}{x} < \ln x$, where $x = \beta_{ij} + \left(1-\beta_{ij}\right)\sigma$.

Clearly, for $Q_{ij}^F > 0$ we require that $\phi_{ij} > 1$; since variable revenues are lower when the firm chooses the intermediary, there must be an off-setting gain in lower fixed market access costs for the intermediated regime to be viable. As β_{ij} rises, $\tilde{\beta}_{ij}$ goes up as the producer can appropriate a larger portion of the surplus, and the cut-off level Q_{ij}^F moves rightwards, reflecting the loss of attractivity of establishing an own affiliate.

We may now state the condition under which intermediaries and own wholesale affiliates coexist within some bilateral trade relationship ij.[21]

Lemma 6 (Existence of intermediaries)**.** *If the inequality*

$$\tilde{\beta}_{ij}^{1-\sigma} < \phi_{ij}$$

holds, then a strictly positive non-overlapping mass of producers from country i exports to country j in each of the two available export modes (intermediation and own foreign sales affiliates).

The condition required for complete partitioning is fairly intuitive: trade intermediation only arises as a viable alternative to wholesale FDI if the distortion associated to it, $\left(\tilde{\beta}_{ij}^{1-\sigma}\right)$, is small enough relative to the cost savings that the avoidance of FDI implies $\left(\phi_{ij}\right)$. If $\tilde{\beta}_{ij} < 1$, for any finite ϕ_{ij}, there is a positive mass of firms that wish to establish a foreign sales affiliate. Note the role of the elasticity of substitution between different varieties: if σ is very small, even a small (effective) cost disadvantage implied by intermediation reduces export revenue by a large amount, making wholesale FDI comparably attractive.

We have $\partial\tilde{\beta}_{ij}/\partial\bar{\beta}_{ij} > 0$ and $\partial\tilde{\beta}_{ij}/\partial\lambda_{ij} > 0$. Hence, when the bargaining power of the producer $\bar{\beta}_{ij}$ is higher in some export market or her outside option better, the loss of revenue implied by intermediation is smaller.

Note that the partitioning of producers into different export modes neither depends on variable transport costs $\bar{\tau}_{ij}$ nor on the (endogenous) size of the export market, as proxied by B_j, or the wage rate. These variables affect both modes in similar fashion.

Figure 1 relates the firms' sorting pattern to their degree of competitive advantage.[22]

We can now use Figure 1 and state the first proposition of our paper.

[21] The following lemma does not suffice to make sure that there always exists a positive measure of firms that do not serve the foreign market at all, i.e, that $Q_i^D < Q_{ij}^M$. This inequality holds for some firms if $\frac{w_j}{w_i}\left(\frac{\bar{\tau}_{ij}}{\tilde{\beta}_{ij}}\right)^{\sigma-1} B_i > B_j$, where w_i, w_j, B_i, and B_j are endogenous objects which can be solved using the labor market clearing and balanced trade conditions for all countries. As the focus of the present paper is not on whether firms export but rather on how they do it, we refrain from determining these objects. We can

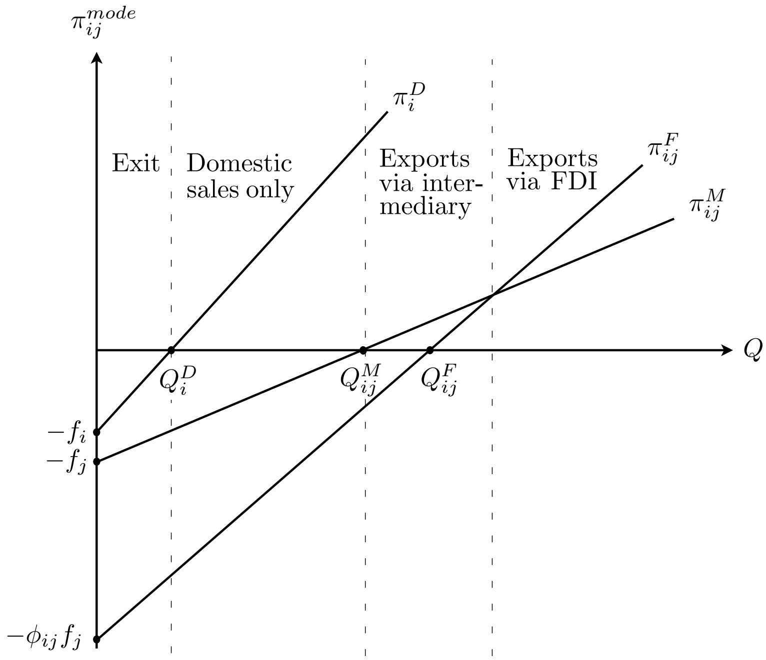


Figure 6.1: Firms sort into different export modes according to competitive advantage Q.

Proposition 5. *Under the condition stated in Lemma 6, producers endogenously select into export modes as a function of their degree of competitive advantage. Firms with high levels of productivity, easily tradable variants, or a strong brand reputation, establish own subsidiaries, while those with intermediate values of the above characteristics search for intermediaries. Firms with low values of the above characteristics do not export.*

Proof. Directly follows from Lemma 6 and Figure 1. □

Proposition 6 implies that – as an export market grows in size (i.e., B_j increases) – medium-size producers switch from using trade intermediaries to establishing own foreign sales affiliates, while smaller firms start exporting via an intermediary.

6.3.2 The prevalence of export modes

Sales of firm Q in either mode are simple log-linear functions of firms' competitive advantage

$$s_{ij}^M(Q) = \sigma \left(\frac{w_i \bar{\tau}_{ij}}{\beta_{ij}}\right)^{1-\sigma} B_j Q \quad \text{and} \quad s_{ij}^F(Q) = \sigma \left(w_i \bar{\tau}_{ij}\right)^{1-\sigma} B_j Q \ . \tag{6.13}$$

derive our main theoretical results without solving for w_i, w_j, B_i, and B_j.

[22]This picture is related to Figure 1 in Helpman, Melitz and Yeaple (2004) [HMY], where the sorting of firms into exporters and firms producing abroad also involves a trade-off between fixed and variable costs, in their case the proximity-concentration trade-off. In the present context, the trade-off is between variable revenue and fixed costs of foreign market access. And, importantly, the slope of the profit functions shown in Figure 1 is endogenously determined as a function of the producers' bargaining power $\bar{\beta}_{ij}$, the technology parameter λ_{ij}, and the elasticity of substitution σ.

Clearly, in each mode, sales are larger the greater is the degree of competitive advantage (Q), the smaller are systematic transportation costs ($\bar{\tau}_{ij}$) and the more income the foreign market has (B_j). Sales per firm also increase in σ as the markup goes down and the firm has to sell larger quantities to amortize fixed costs. The more severe contractual imperfections $1/\beta_{ij}$ are, the lower sales per firm channeled through intermediaries, whereas exports per firm via wholesale affiliates in not affected by the contracting environment.[23]

We can compute the value of total export sales of country i to country j that are facilitated by trade intermediaries S_{ij}^M.[24] With Φ distributed according to the Pareto distribution, as assumed above, aggregate export sales of country i to country j via intermediaries are given by

$$S_{ij}^M = \Psi_{ij}\beta_{ij}^{\sigma-1}\left[\tilde{\beta}_{ij}^{k-(\sigma-1)} - \left(\frac{1-\tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij}-1}\right)^{\bar{k}}\right], \tag{6.14}$$

where $\bar{k} \equiv \frac{k}{\sigma-1} - 1$ is a constant, and Ψ_{ij} is a variable that will turn out constant over export modes.[25] Looking at the first order effect only, intermediated exports from i to j increase when both countries involved are larger or systematic trade costs $\bar{\tau}_{ij}$ are smaller. Intermediated exports also fall in f, the fixed costs that any foreign market presence entails.

Contractual frictions $1/\beta_{ij}$ affect intermediated export sales in several ways. First, taking wages and market size as given, for any firm, a lower degree of contractual imperfections increases sales through intermediaries, see (6.13). Second, contractual imperfections affect the selection of producers into the intermediated distribution mode. As $1/\beta_{ij}$ goes down, more firms find it optimal to export through trade intermediaries and either choose to establish an own sales affiliate abroad or stop exporting to market j completely; see (6.11) and (6.12). Hence, a reduction of contractual imperfections has a positive effect on total intermediated export sales both on the *intensive* and on the *extensive* margin. Ignoring general equilibrium effects, the derivative of S_{ij}^M with respect to β_{ij} is positive.[26]

Similarly, we can derive total exports of i into j through own wholesales affiliates S_{ij}^F.

[23] These observations relate to *direct* effects only; $\sigma, \bar{\tau}_{ij}, \beta_{ij}$ also affect sales through B_j.

[24] We have $S_{ij}^M = M_i^E \int_{\Phi_{ij}^M}^{\Phi_{ij}^F} s_{ij}^M(\Phi)\, dG(\Phi)$, where M_i^E is the mass of entrants in country i.

[25] The term Ψ_{ij} is endogenously determined and given by $\Psi_{ij} = \frac{\sigma k}{k-(\sigma-1)} M_i^E B_j^{\frac{k}{\sigma-1}} (w_i\bar{\tau}_{ij})^{-k} (fw_j)^{-\bar{k}}$

[26] This follows immediately from the considerations on the intensive and extensive margin above.

Evaluating $S_{ij}^F = M_i^E \int_{\Phi_{ij}^F}^{\infty} s_{ij}^F(\Phi)\, dG(\Phi)$, we have

$$S_{ij}^F = \Psi_{ij} \left(\frac{1 - \tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij} - 1} \right)^{\bar{k}}. \tag{6.15}$$

A rise in $1/\beta_{ij}$ now does not affect sales of each single exporter in the FDI mode directly, see (6.13). Total sales to affiliates, however, increase as some firms switch from using intermediaries to establishing own affiliates so that the cut-off value Q_{ij}^F falls; see (6.12).

Without explicitly solving for Ψ_{ij}, we can can now state the following proposition on the relative prevalence of export modes:

Proposition 6. *If the condition stated in Lemma 6 holds and if firms' degree of competitive advantage follows the Pareto distribution, the prevalence of export sales via trade intermediaries relative to sales through affiliates, $\chi_{ij} \equiv S_{ij}^M / S_{ij}^F$, is*

$$\chi_{ij} = \beta_{ij}^{\sigma-1} \left[\left(\frac{\phi_{ij} - 1}{\tilde{\beta}_{ij}^{1-\sigma} - 1} \right)^{\bar{k}} - 1 \right].$$

This measure increases in the risk of expropriation ϕ_{ij} and decreases in the severity of contractual problems $1/\beta_{ij}$. It is independent from the size of the export market as given by B_j, the wage rates in either country, and from transportation costs $\bar{\tau}_{ij}$. It decreases in the degree of dispersion of competitive advantage $1/k$ and falls in the elasticity of substitution σ. Moreover, χ_{ij} decreases in the dispersion of domestic sales, given by $1/[k - (\sigma - 1)]$.

Not surprisingly, when the strength of contractual imperfections increases (i.e., β_{ij} drops) intermediation becomes more expensive relative to the use of an own wholesale affiliate; hence relative prevalence of intermediation (χ_{ij}) falls. On the other hand, sales through intermediaries are more prevalent if the protection of property against expropriation is low (i.e., ϕ_{ij} is high).

More interestingly, χ_{ij} does not depend on the systematic component of transportation costs $(\bar{\tau}_{ij})$. This is due to the fact that sales in both distribution modes are affected by systematic transportation costs in the same way. Approximating $\bar{\tau}_{ij}$ with bilateral geographical distance, it follows that the relative prevalence of intermediation does not depend on geographical distance. This is a prediction of our framework that is testable given adequate data. Also, relative prevalence χ_{ij} increases as firms become more homogeneous $(\bar{k} \to \infty)$. In the extreme case, the distribution of Q has a mass point at the lower bound of its support (here: normalized to unity). If the condition in Lemma 6 is met, most firms cluster in the neighborhood of the lower bound of the support and therefore export through intermediaries. As $\bar{k}$ falls, more firms find it optimal to establish own subsidiaries and χ_{ij} falls.

6.3.3 The trade-FDI relation

Total exports of country i to country j over all possible distribution modes are $X_{ij} \equiv S_{ij}^M + S_{ij}^F$. Hence,

$$X_{ij} = \Psi_{ij} \beta_{ij}^{\sigma-1} \tilde{\beta}_{ij}^{k-(\sigma-1)} \left(\left(\frac{\tilde{\beta}_{ij}^{1-\sigma} - 1}{\phi_{ij} - 1} \right)^{\bar{k}} \left(\beta_{ij}^{1-\sigma} - 1 \right) + 1 \right). \tag{6.16}$$

This expression is the bilateral trade flow (gravity) equation associated to our model. It is a somewhat generalized version of the one derived by Chaney (2008): Due to the presence of an extensive margin (selection of firms into exporting), the *first order effect* of transportation costs $\bar{\tau}_{ij}$, which in the expression above is part of Ψ_{ij} (see footnote 24) has an elasticity of k and not $|1-\sigma|$ as in models with homogeneous firms. This result of Chaney still holds in the present context, where firms have a choice of export mode since transportation costs affect both modes similarly.

Several researchers have used measures of contractual frictions in gravity models of bilateral trade without offering a structural derivation of the estimated relationship. Equation (6.16) offers a foundation for such econometric practice. However, since contractual frictions interact with the underlying distribution of firms' competitive advantage, and thereby affect trade volumes both on an extensive and an intensive margin, the gravity equation (6.16) turns out non-linear in frictions and therefore not trivial to estimate.

Moreover, we can express the total volume of foreign direct investment of firms headquartered in country i into wholesale affiliates in country j as $F_{ij} = \phi_{ij} f_j M_i^E \left[1 - G\left(\Phi_{ij}^F \right) \right] = \phi_{ij} f_j M_i^E \left(\Phi_{ij}^F \right)^{-k}$, where the second equality follows from the Pareto assumption. The stock of wholesale FDI from country i invested in country j is[27]

$$F_{ij} = \phi_{ij} \Psi_{ij} \rho \frac{\bar{k}}{k} \left(\frac{1 - \tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij} - 1} \right)^{\frac{k}{\sigma-1}}. \tag{6.17}$$

[27]Note that $\left(\bar{\delta} + \delta_{ij} \right) F_{ij}$ is the flow of FDI in a stationary state.

This expression is related to (6.15) and features similar comparative statics. In particular, the elasticities of FDI and exports, both taken with respect to the systematic component of transportation costs, are identical. The comparative statics with respect to the cost of FDI ϕ_{ij} is more involved. A higher cost of wholesale FDI has a negative effect on the extensive margin (fewer firms engage in wholesale FDI), but a positive effect on the intensive margin (firms who export through an affiliate have to invest more). However, ignoring general equilibrium effects through Ψ_{ij} one can show that the extensive margin dominates. On the other hand, higher cost of wholesale FDI also reduce total exports.

More generally, one can state the general equilibrium relationship between total exports and the stock of bilateral wholesale FDI as follows:

Proposition 7. *Wholesale FDI (F_{ij}) and total exports (X_{ij}) are identically affected by changes in the systematic component of trade costs so that they appear as complements.*

Proof. Totally differentiating (6.16) and (6.17) immediately reveals that $\frac{\hat{F}_{ij}}{\hat{\bar{\tau}}_{ij}} = \frac{\hat{X}_{ij}}{\hat{\bar{\tau}}_{ij}}$ (where we use the usual hat notation, i.e. $\hat{x} = dx/x$.) Hence, any change in $\bar{\tau}_{ij}$ makes WFDI and exports move in tandem. □

In contrast to the results reported in the present paper, statements on the comparative statics of WFDI or exports with respect to the risk of expropriation ϕ_{ij} and contractual imperfections $1/\beta_{ij}$ would require a solution of the model in general equilibrium. However, models with asymmetric countries and bilateral trade costs are close to the frontier of analytical tractability, see HMY. Moreover, we do not expect additional testable implications from such an exercise, whose theoretical exploration we relegate to future research.

6.4 Empirical evidence

The present paper has a number of predictions that are – in principle – testable empirically. In this section, we show that existing evidence using US Census data on related versus non-related party trade and data on German foreign plants is compatible with the main results of the present paper. Clearly, it would be worthwhile to use detailed firm-level data. We are not aware of a data set that contains at the same time information about firms' characteristics, their choice of export modes by export country, and export volumes. The data situation is improving quickly; hence, we are confident that the predictions of the present paper can be put to systematic econometric scrutiny soon. For the time being, we are confident that the available data plus other empirical results do allow a rough yet meaningful empirical check.

6.4.1 Relative prevalence

In official data, exports to wholesale affiliates for the purpose of selling to foreign consumers appear as within-firm trade. Hence, one may use data on related-party and non-related party exports, as collected by the US Census, to see whether sectoral and cross-country variation in the incidence of intermediation is in line with the predictions of our paper.[28] While in the theoretical part of the paper we mainly focus on trade in final goods, exports to affiliates not only include final output goods but also intermediate inputs.[29] Unfortunately, this problem is common to the literature. The empirical analysis in HMY relies on export data from Feenstra (1997) that do not distinguish between final goods and imports either. However, as we have pointed out in footnote 8, our setting is flexible enough to nest also trade in inputs without altering the testable implications of the model.

Compared to HMY, we discuss a different issue (the choice of export mode versus the choice of location of production) and stress a different mechanism (contractual imperfections versus concentration-proximity). We can use a similar empirical strategy on different data, namely US census data, to assess the predictions of the model. While HMY study sales *of* foreign affiliates versus export sales, our dependent variable relates export sales to intermediaries versus those *to* foreign affiliates. Hence, our exercise is not subject to the criticism, that it is essentially unknown where (and by whom) products sold by foreign affiliates have been produced.

Relative prevalence and product characteristics. We strive at checking the signs of the partial derivatives as derived in Proposition 6. For that purpose, we show unconditional and conditional correlations. Starting with the cross-industry perspective, we need

[28] A detailed description on the firm-level version of the data can be found in Bernard, Jensen, and Schott (2007b). Strictly spoken, non-related party trade also comprises exports directly to the consumer.

[29] Using BEA data on majority owned US affiliates, Borga and Zeile (2004) find that finished goods make up only about 20% of exports to affiliates. While we could back out the exports of finished goods to related parties, we cannot do so for non-related party trade.

comparable proxies for the relative prevalence of export modes, the dispersion of sales, and contractual imperfections.

Data on related party and non-related party trade is taken from the US Census.[30] To make the export data comparable to the dispersion measures reported by HMY,[31] we aggregate the US Census data from the 6-digit NAICS level to match the BEA 3-digit industry classification.[32] Since exports on that very disaggregated level in either mode may be driven by only very few transactions, we average over the years 2000 to 2003. Finally, we restrict our analysis to the (wide) sample of countries as considered in HMY.[33] This choice makes sure that we focus on countries with strictly positive exports in both modes.

According to Proposition 6, we expect the following signs in our analysis. The relative prevalence of export sales decreases in the dispersion of the sales distribution. Variation in the dispersion measure derives from variation in the shape parameter of the Pareto distribution k, and from variation in the elasticity of substitution σ. However, the latter is not only related to sales dispersion, but also affects the strength of the intermediation impeding effect of contractual problems $1/\beta_{ij}$. Thus, as stated in Proposition 6, the relative prevalence of export modes decreases in σ. Moreover, there might be a countervailing effect running via the recycling rate λ_{ij}. If goods are very specialized (σ low), λ_{ij} may be lower so that producers are more vulnerable and contractual problems $1/\beta_{ij}$ are stronger. There is then a negative correlation between $1/\beta_{ij}$ and σ. The total effect of σ is therefore unclear.

Our theory differs from the proximity-concentration trade-off as proposed in HMY with regard to the role of variable trade costs: While they drive the decision between exporting and producing abroad, they should not play a role for the relative prevalence of export modes. In order to check this prediction of our model, we add the freight rate as a control.

Data on σ and freight rates are taken from Hanson and Xiang (2004). We match their product classification into the BEA industry classification and are left with 27 manufactur-

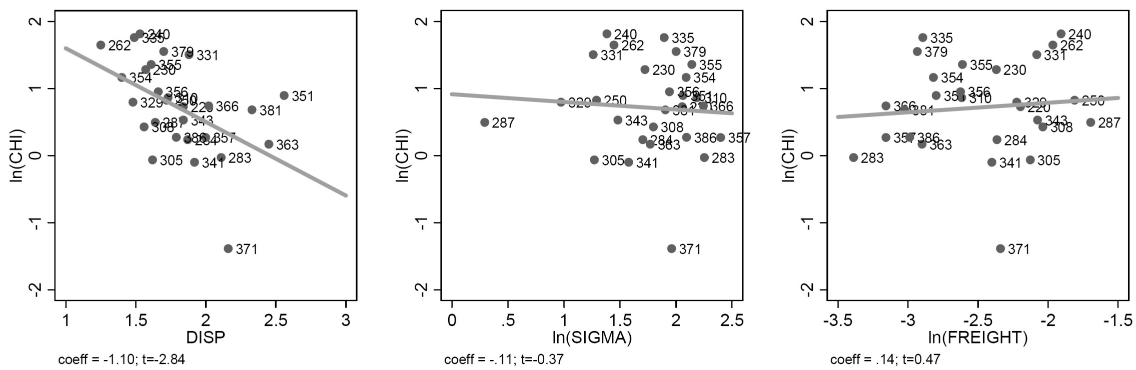
[30]The data can be downloaded from `http://sasweb.ssd.census.gov/relatedparty`.

[31]They construct measures on the basis of different data sources for the US and Europe for 52 BEA 3-digit manufacturing industries.

[32]The correspondence table can be found in Appendix D.

[33]The 38 countries in the sample are Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, Colombia*, Denmark, Finland*, France, Germany, Greece*, Hong Kong, Indonesia*, Ireland, Israel*, Italy, Japan, Malaysia*, Mexico, Netherlands, New Zealand, Norway, Peru, Philippines, Portugal*, Singapore, South Africa*, South Korea, Spain, Sweden, Switzerland, Taiwan, Thailand*, Turkey, United Kingdom, and Venezuela. An asterisk indicates countries excluded in the narrow sample.

Figure 6.2: Unconditional correlations. Cross-industry variation



ing industries (see Appendix D for the industry concordance). It turns out that for Motor Vehicles (BEA industry 371) exports via an own affiliates is the most prevalent mode, while Wood and Lumber (240) Non-Ferrous Metals (336), and Pulp and Paper (262) are prevalently exported via intermediaries; see Appendix B for a detailed description of the data).

The analysis comes in two parts. First, we aggregate over all countries in our sample and plot the log of relative prevalence of export modes separately against industry dispersion, the log of elasticity of substitution, and the log of freight rate (see Figure 6.2). As expected, there is a negative unconditional correlation between the relative prevalence and industry dispersion. However, there is no clear correlation between χ and σ. This confirms our hypothesis that contractual problems reduce the prevalence of intermediated exports. Moreover, as predicted by our model, variable transport costs do not drive the relative prevalence of export modes. Hence, the concentration proximity argument indeed does not play a role in determining the choice of export modes.

Second, in order to get closer to an empirical test of the relationship proposed in Proposition 6, we run a regression of the type

$$\ln \chi_{sj} = \alpha_0 + \alpha_1 DISP_s + \alpha_2 \ln \sigma_s + \alpha_2 \ln FREIGHT_s + \nu_j + u_{sj}, \qquad (6.18)$$

where s denotes an industry and j a partner country. With respect to the estimation strategy, we follow HMY. First, we include country fixed effects to control for unobserved heterogeneity, e.g. systematic trade costs (distance from the US), market size, and investment risk. Moreover, the country fixed effect controls for multilateral resistance, thereby addressing the issue that some countries like the Netherlands, Hong Kong, or Singapore may act as intermediaries in entrepôt trade. Second, in order to address potential endogeneity bias in the US dispersion measure, we instrument the US measure by those of four European countries (also estimated by HMY). Finally, we cluster standard errors within BEA industries to control for correlation of residuals due to omitted industry characteristics. The regression is based on a balanced sample of $27 \times 38 = 1,026$ observations. Details on this regression are found in Appendix C.

In the preferred specification, as expected, the sign of α_1 is negative ($\widehat{\alpha}_1 = -1.8$, t-value: -2.58), while $\widehat{\alpha}_2$ and $\widehat{\alpha}_3$ are not significantly different from 0. This result is in line with our prediction that contractual problems–that are negatively correlated with σ–hampers exports via intermediaries. It also confirms our hypothesis that relative prevalence is not driven by variable trade costs. Our results are robust to restricting the sample to the 27 countries used by HMY and referred to as the narrow sample.[34]

Relative prevalence and destination country characteristics. The US Census data also allow a rough impression on the relative prevalence of export modes with respect to destination country characteristics like geographical distance, country size, and the degree of property rights protection. For that purpose, we include geographical distance of country j to the USA, $DIST_j$, and the size of population, POP_j in our dataset.[35] Moreover, we include a measure, $RISK_j$ from the International Country Risk Guide (ICRG) that captures the risk of expropriation and confiscation of productive assets by the state or other actors in country j. According to our model, firms use an intermediary instead of setting up an own wholesale affiliation if this risk is high.[36]

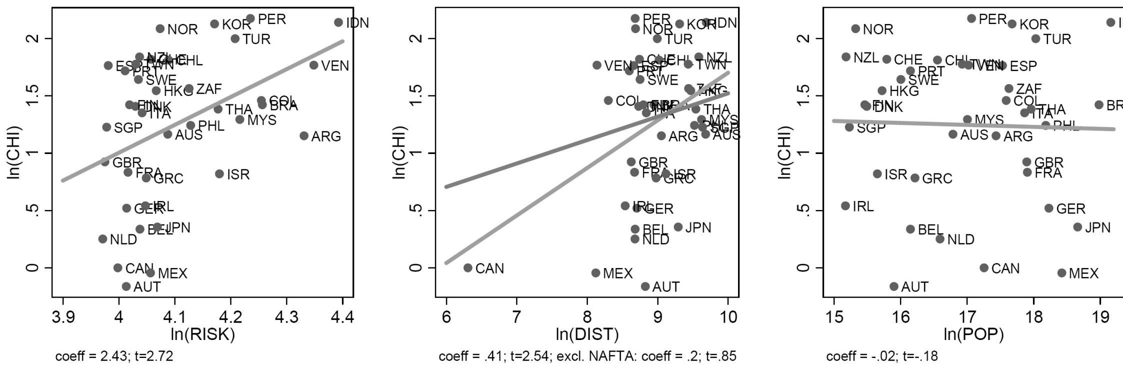
We aggregate the trade data over all of our 27 industries and find that Austria and the NAFTA trading partners Canada and Mexico are mainly served via own affiliates, while countries like Peru, Indonesia, and South Korea are prevalently exported to via intermedi-

[34]The coefficient of dispersion slightly increases in absolute values ($\widehat{\alpha}_1 = -1.9$).

[35]The distance data is provided by Centre d'Etudes Prospectives et d'Informations Internationales (CEPII) in Paris on their website `www.cepii.fr/francgraph/bdd/distances.htm`; population (for year 2000) comes from the World Development Indicators (WDI) provided by the World Bank.

[36]The variable *investment profile* in the ICRG attaches high scores to low risk. However, we invert the values to obtain our measure $RISK_j$.

Figure 6.3: Unconditional correlations. Cross-country variation



aries; see Appendix B for detailed summary statistics.[37]

Turning to the unconditional effects of our explanatory variables, Figure 6.3 suggests that higher risk of confiscation or expropriation ($RISK_j$), leads to relatively higher exports via intermediaries. This result is well in line with our theory. Relative prevalence and geographical distance seem to correlate positively, while our model would predict that distance as a proxy for variable trade costs does not affect the prevalence of export modes. However, the result is obviously driven by trade within NAFTA (Canada and Mexico). If we omit NAFTA trading partners, there is no significant relation. Finally, Figure 6.3 shows that market size, measured in terms of population, does not affect the choice of export mode as predicted by our model.

We also run a regression of the type

$$\ln \chi_{sj} = \beta_0 + \beta_1 \ln RISK_j + \beta_2 \ln DIST_j + \beta_3 NAFTA_j + \beta_4 \ln MR_j + \beta_5 \ln POP_j + \nu_s + u_{sj}, \tag{6.19}$$

where MR_j measures the multilateral resistance of country j, which is related to the price index P_j in our model.[38] Details on this regression are in Appendix C. As expected, the relative prevalence increases in $RISK$ ($\widehat{\beta}_1 = 1.943$, t-value: 2.16). Moreover, we find that distance and multilateral resistance have no significant impact on the relative prevalence of export modes.

[37] The high level of exports to Austria channeled through sales affiliates may point to its role as re-exporter to neighboring countries. We will control for that by including a measure of multilateral resistance into the cross-country regressions.

[38] Following the literature, multilateral resistance is the GDP weighted distance of country j to all other countries than the US.

Interestingly, while in the plot of Figure 6.3 market size measured by POP_j seemed unrelated to the relative prevalence of intermediation, it appears significant in the conditional regression where χ_{sj} decreases in country size ($\widehat{\beta}_5 = -0.22$, t-value: -2.71). This means that firms tend to serve larger markets via own sales affiliates and smaller markets via intermediaries. This is not in line with the predictions of our model, where larger markets attract more exports through a proportional expansion of both modes.

One can rationalize the empirical finding by assuming a relation between fixed foreign market access costs and country size. If those costs depend positively on country size, e.g., because more sales agents need to be hired, but increase less strongly in the case of wholesale FDI, e.g., because the foreign firm loses some of its initial cost disadvantage relative to domestic firms as it grows larger, then relative prevalence of intermediation declines in population. Another explanation may involve the fact that firms are risk averse, so that the hold-up problem implicit in the intermediated mode becomes fiercer when the profits at stake are larger (which is the case in a larger country). Finally, the Pareto assumption, while found realistic in many empirical studies of firm size distributions, may not be adequate. In this case, the expression describing relative prevalence is no longer independent (amongst other things) from country size.

Our empirical exercise is in line with a number of predictions of the model, supporting our view that the choice of export mode reflects a trade-off between the costs of contractual frictions in the case of intermediation and the cost of FDI in the case of internalization. The fact that the data show a negative relationship between the prevalence of intermediation and market size suggests an interesting research agenda in which fixed costs of foreign market access are modeled with more detail.

6.4.2 The FDI-distance nexus

Citing various empirical studies, Neary (2008) points out that – in the cross-section of countries – distance is negatively correlated to the stock and flow of foreign direct investment (FDI). In other words, destination countries that are farther away from the source country receive less FDI. This seems inconsistent with the concentration-proximity trade-off in the standard model of horizontal FDI, where producers overcome the cost of distance by establishing foreign production plants. Buch, Kleinert, Lipponer, and Toubal (2005) provide a nice example for German data; controlling for destination country size and similarity with the source country, they find that distance reduces sales of foreign affiliates, regardless of whether these subsidiaries are active in production or wholesale activities. HMY present evidence that is in line with concentration-proximity, but focus on cross-industry variance in trade costs.

Our paper predicts that sales of wholesale affiliates decline in distance, exactly for the same reason that this is true for traditional exports. Also, the number of firms engaged in running foreign wholesale affiliates declines with distance. Hence, the very rudimentary theory of FDI present in our paper fits the data. If wholesale activities are sufficiently important for total FDI, then it may well be that aggregate data exhibit a negative relationship between affiliate sales and distance as well. In Germany, for example, in 2005, about 27 percent of the total value of foreign FDI was in the wholesale sector.[39]

6.5 Conclusions

In this paper, we have discussed the choice between two different modes of exporting to a foreign market: a producer can either use a foreign trade intermediary, who enjoys a fixed cost advantage but – due to the lack of enforceable cross-country contracts – exposes the producer to a hold-up problem, or they can establish an own wholesale affiliate, avoiding the threat of hold-up at the cost of increased investment. This trade-off produces an interesting sorting pattern of producers into the two export modes. Firms with high perceived quality of their products, low variable production costs, and strong marketability of goods prefer to establish affiliates; firms with low realizations of those characteristics prefer to use trade intermediaries. The reason is that contractual frictions reduce variable revenues proportionally, while the fixed-cost disadvantage of affiliates does not depend on sales. Hence, firms with high sales opt for wholesale subsidiaries in the foreign country.

Importantly, in our model, variable trade costs are endogenously determined in the game between the producer and the intermediary. However, the contractual frictions are

[39] The total stock of FDI used as a base does not include the financial sector (in particular holding companies and off-shore investment vehicles).

not isomorphic to the usual iceberg-type trade costs, since they do not lead to a loss of output. Rather, they imply an additional restriction of production by monopolistically competitive firms, so that the markup goes up. Hence, our model warns against modeling differences across modes as exogenous differences in iceberg-type variable trade costs.

Under the assumption of the Pareto distribution, we show that the relative prevalence of intermediation does not depend on transportation costs between the source and the destination country, on market size or on wage rates. It increases with the risk of expropriation of foreign assets and in the degree of heterogeneity of producers. It falls with the severity of contractual problems and the elasticity of substitution between varieties.

Our paper is related to Helpman, Melitz, and Yeaple (2004). While we discuss a different issue (the choice of export mode versus the choice of location of production) and stress a different mechanism (contractual imperfections versus concentration-proximity), we can use a related empirical strategy on US census data to assess the predictions of the model. We find that most predictions of our theory are in line with the data.

We close the paper with a brief outlook on further research. First, while capturing an important trade-off between contractual frictions and the cost of internalization through wholesale FDI, our model of endogenously arising trade intermediation is only a first pass at a complex issue. In order to improve our understanding of trade costs further, one may want to develop a more realistic model of multi-product trade intermediaries. Second, our empirical analysis draws on sectoral data; a firm-level analysis would be preferable. As soon as data on firms' choices of export modes becomes available, one can put a wider array of implications of our model to a test.

6.6 Appendix A. Proofs and detailed derivations

Proof of Lemma 5 (pricing behavior). The producer maximizes her expected profits from exporting via a trade intermediary subject to the demand function to choose her optimal quantity to supply in the match. Using optimal demand to substitute out the c.i.f price and inserting (6.7) the solves

$$\begin{aligned}
& \max_{x_{ij}^{M}(\omega)} \bar{\beta}_{ij} J_{ij}(\omega) + \tilde{\pi}_{ij}^{P}(\omega) - x_{ij}^{M}(\omega)\tau_{ij}(\omega)a(\omega)w_i \\
= & \max_{x_{ij}^{M}(\omega)} \bar{\beta}_{ij} p_{ij}^{M}(\omega) x_{ij}^{M}(\omega) + \left[\left(1-\bar{\beta}_{ij}\right)\lambda_{ij} - 1\right] x_{ij}^{M}(\omega)\tau_{ij}(\omega)a(\omega)w_i \\
= & \max_{x_{ij}^{M}(\omega)} \bar{\beta}_{ij}(H_j)^{1/\sigma}\zeta(\omega)^{(\sigma-1)/\sigma}\left[x_{ij}^{M}(\omega)\right]^{(\sigma-1)/\sigma} + \left[\left(1-\bar{\beta}_{ij}\right)\lambda_{ij} - 1\right] x_{ij}^{M}(\omega)\tau_{ij}(\omega)a(\omega)w_i
\end{aligned}$$

The first order condition is

$$\rho\bar{\beta}_{ij}(H_j)^{\frac{1}{\sigma}}\zeta(\omega)^{\frac{\sigma-1}{\sigma}}\left[x_{ij}^{M}(\omega)\right]^{-\frac{1}{\sigma}} = \left[\left(1-\bar{\beta}_{ij}\right)\lambda_{ij} - 1\right]\tau_{ij}(\omega)a(\omega)w_i.$$

Substituting $x_{ij}^{M}(\omega)$ yields the pricing rule stated in Lemma 5

$$p_{ij}^{M}(\omega) = \frac{\tau_{ij}(\omega)a(\omega)w_i}{\beta_{ij}\rho},$$

where $\beta_{ij} = \beta_{ij}\left(\bar{\beta}_{ij},\lambda_{ij}\right) = \bar{\beta}_{ij}/\left[1-\lambda_{ij}\left(1-\bar{\beta}_{ij}\right)\right] \geq \bar{\beta}_{ij}$.

Comparative statics related to Lemma 5. The additional markup is inverse proportional to the degree of contractual imperfections β_{ij}. $\beta_{ij}\left(\bar{\beta}_{ij},\lambda_{ij}\right)$ is increasing in the bargaining power $\bar{\beta}_{ij}$ and the recycling rate λ_{ij}

$$\begin{aligned}
\frac{\partial\beta_{ij}\left(\bar{\beta}_{ij},\lambda_{ij}\right)}{\partial\bar{\beta}_{ij}} &= \frac{1-\lambda_{ij}}{\left(1-\lambda_{ij}\left(1-\bar{\beta}_{ij}\right)\right)^2} > 0, \\
\frac{\partial\beta_{ij}\left(\bar{\beta}_{ij},\lambda_{ij}\right)}{\partial\lambda_{ij}} &= \frac{\left(1-\bar{\beta}_{ij}\right)\bar{\beta}_{ij}}{\left(1-\lambda_{ij}\left(1-\bar{\beta}_{ij}\right)\right)^2} > 0.
\end{aligned}$$

The term $\tilde{\beta}_{ij} = \tilde{\beta}_{ij}\left(\beta_{ij},\sigma\right) \equiv \left[\beta_{ij} + \left(1-\beta_{ij}\right)\sigma\right]^{\frac{1}{\sigma-1}}\beta_{ij} \geq \beta_{ij}$ is closely related to our measure of contractual imperfections β_{ij}. We have $\tilde{\beta}_{ij}(0,\sigma) = 0$ and $\tilde{\beta}_{ij}(1,\sigma) = 1\tilde{\beta}_{ij}$ is strictly increasing in β_{ij} for $\beta_{ij} \in (0,1)$

$$\frac{\partial\tilde{\beta}_{ij}\left(\beta_{ij},\sigma\right)}{\partial\beta_{ij}} = \frac{\tilde{\beta}_{ij}}{\beta}\left(1 - \frac{\beta_{ij}}{\beta_{ij} + \left(1-\beta_{ij}\right)\sigma}\right) > 0,$$

since $\beta_{ij}/\left[\beta_{ij}+\left(1-\beta_{ij}\right)\sigma\right]<1$.

The derivative with respect to σ is given by

$$
\begin{aligned}
\frac{\partial\tilde{\beta}_{ij}}{\partial\sigma} &= \tilde{\beta}_{ij}\left(-\frac{\ln\left[\beta_{ij}+\left(1-\beta_{ij}\right)\sigma\right]}{\left(\sigma-1\right)^{2}}+\frac{1-\beta_{ij}}{\left[\beta_{ij}+\left(1-\beta_{ij}\right)\sigma\right]}\right)\\
\frac{\partial\tilde{\beta}_{ij}}{\partial\sigma} &= \frac{\tilde{\beta}_{ij}}{\sigma-1}\frac{\beta_{ij}+\left(1-\beta_{ij}\right)\sigma-1-\left[\beta_{ij}+\left(1-\beta_{ij}\right)\sigma\right]\ln\left[\beta_{ij}+\left(1-\beta_{ij}\right)\sigma\right]}{\left(\sigma-1\right)\left[\beta_{ij}+\left(1-\beta_{ij}\right)\sigma\right]}<0.
\end{aligned}
$$

$\tilde{\beta}_{ij}$ is strictly decreasing in σ, since $\frac{x-1}{x}<\ln x$, where $x=\beta_{ij}+\left(1-\beta_{ij}\right)\sigma$.

Moreover, $\tilde{\beta}_{ij}$ is well behaved in the limiting cases

$$
\begin{aligned}
\lim_{\sigma\to1}\tilde{\beta}_{ij}\left(\beta_{ij},\sigma\right) &= \beta_{ij}\exp\left[\lim_{\sigma\to1}\left(\frac{\ln\left(\beta_{ij}+\left(1-\beta_{ij}\right)\sigma\right)}{\sigma-1}\right)\right]\\
&= \beta_{ij}\exp\left[\lim_{\sigma\to1}\left(\frac{1-\beta_{ij}}{\beta_{ij}+\left(1-\beta_{ij}\right)\sigma}\right)\right]\\
&= \beta_{ij}e^{1-\beta_{ij}},\\
\lim_{\sigma\to\infty}\tilde{\beta}_{ij}\left(\beta_{ij},\sigma\right) &= \beta_{ij}\exp\left[\lim_{\sigma\to\infty}\left(\frac{\ln\left(\beta_{ij}+\left(1-\beta_{ij}\right)\sigma\right)}{\sigma-1}\right)\right]\\
&= \beta_{ij}\exp\left[\lim_{\sigma\to1}\left(\frac{1-\beta_{ij}}{\beta_{ij}+\left(1-\beta_{ij}\right)\sigma}\right)\right]\\
&= \beta_{ij}.
\end{aligned}
$$

Proof of Lemma 6 (Existence of intermediation. The cutoff Q_{ij}^{M} immediately follows from rearranging (6.10)

$$
Q_{ij}^{M}=\left(\frac{w_i\bar{\tau}_{ij}}{\tilde{\beta}_{ij}}\right)^{\sigma-1}f_jB_j^{-1}.
$$

Q_{ij}^{F} is determined by solving $\pi_{ij}^{M}\left(Q_{ij}^{F}\right)=\pi_{ij}^{F}\left(Q_{ij}^{F}\right)$ for Q_{ij}^{F}

$$
\begin{aligned}
\left(w_i\bar{\tau}_{ij}\right)^{1-\sigma}B_jQ_{ij}^{F}-\phi_{ij}f_j &= \left(\frac{w_i\bar{\tau}_{ij}}{\tilde{\beta}_{ij}}\right)^{1-\sigma}B_jQ_{ij}^{F}-f_j\\
Q_{ij}^{F} &= \left(w_i\bar{\tau}_{ij}\right)^{\sigma-1}\left(\frac{\phi_{ij}-1}{1-\tilde{\beta}_{ij}^{\sigma-1}}\right)f_jB_j^{-1}\\
&= Q_{ij}^{M}\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right)
\end{aligned}
$$

Sorting exists, if Q_{ij}^F is strictly larger than Q_{ij}^M :

$$\begin{aligned}
(w_i\bar{\tau}_{ij})^{\sigma-1}\left(\frac{\phi_{ij}-1}{1-\tilde{\beta}_{ij}^{\sigma-1}}\right) f_j B_j^{-1} &> \left(\frac{w_i\bar{\tau}_{ij}}{\tilde{\beta}_{ij}}\right)^{\sigma-1} f_j B_j^{-1} \\
\phi_{ij}-1 &> \tilde{\beta}_{ij}^{1-\sigma}\left(1-\tilde{\beta}_{ij}^{\sigma-1}\right) \\
\phi_{ij} &> \tilde{\beta}_{ij}^{1-\sigma}.
\end{aligned}$$

Derivations of equations (6.14) **and** (6.15) **(Export sales per mode).** Sales per firm from exporting via a trade intermediary are given by

$$\begin{aligned}
s_{ij}^M(\omega) &= p_{ij}^M(\omega)\, x_{ij}^M\left[p_{ij}^M(\omega)\right] \\
&= H_j\left[\frac{p_{ij}^M(\omega)}{\zeta(\omega)}\right]^{1-\sigma} \\
s_{ij}^M(Q) &= \sigma\left(\frac{w_i\bar{\tau}_{ij}}{\beta_{ij}}\right)^{1-\sigma} QB_j.
\end{aligned}$$

Using $Q = \Phi^{(1-\sigma)}$ and the Pareto distribution, total exports via intermediaries can be calculated as

$$\begin{aligned}
S_{ij}^M &= M_i^E\sigma\left(w_i\frac{\bar{\tau}_{ij}}{\beta_{ij}}\right)^{1-\sigma} B_j k\int_{\Phi_{ij}^M}^{\Phi_{ij}^F}\Phi^{\sigma-k-2}d\Phi \\
&= M_i^E\left(w_i\frac{\bar{\tau}_{ij}}{\beta_{ij}}\right)^{1-\sigma} B_j\frac{\sigma k}{k-(\sigma-1)}\left[\left(\Phi_{ij}^M\right)^{\sigma-k-1}-\left(\Phi_{ij}^F\right)^{\sigma-k-1}\right] \\
&= M_i^E\sigma\left(w_i\frac{\bar{\tau}_{ij}}{\beta_{ij}}\right)^{1-\sigma} B_j\frac{\sigma k}{k-(\sigma-1)}\left(\Phi_{ij}^F\right)^{\sigma-k-1}\left[\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right)^{\frac{k-(\sigma-1)}{\sigma-1}}-1\right] \\
&= M_i^E\left(w_i\bar{\tau}_{ij}\right)^{-k} B_j^{\frac{k}{\sigma-1}} f_j^{-\bar{k}}\frac{\sigma k}{k-(\sigma-1)}\beta_{ij}^{\sigma-1}\left(\frac{\phi_{ij}-1}{1-\tilde{\beta}_{ij}^{\sigma-1}}\right)^{\frac{\sigma-k-1}{\sigma-1}}\left[\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right)^{\frac{k-(\sigma-1)}{\sigma-1}}-1\right] \\
&= \Psi_{ij}\beta_{ij}^{\sigma-1}\left[\tilde{\beta}_{ij}^{k-(\sigma-1)}-\left(\frac{1-\tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij}-1}\right)^{\bar{k}}\right].
\end{aligned}$$

The last expression is equivalent to (6.14) in the text. Analogously, sales per firm from exporting via a wholesale affiliate take the form

$$\begin{aligned}
s_{ij}^F(\omega) &= p_{ij}(\omega)\, x_{ij}\left[p_{ij}(\omega)\right] \\
s_{ij}^F(Q) &= \sigma\left(w_i\bar{\tau}_{ij}\right)^{1-\sigma} QB_j,
\end{aligned}$$

and

$$\begin{aligned} S_{ij}^F &= M_i^E \sigma \left(w_i \bar{\tau}_{ij}\right)^{1-\sigma} B_j k \int_{\Phi_{ij}^F}^{\infty} \Phi^{\sigma-k-2} d\Phi \\ &= M_i^E \left(w_i \bar{\tau}_{ij}\right)^{1-\sigma} B_j \frac{\sigma k}{k-(\sigma-1)} \left(\Phi_{ij}^F\right)^{\sigma-k-1} \\ &= \Psi_{ij} \left(\frac{1-\tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij}-1}\right)^{\bar{k}}. \end{aligned}$$

which corresponds to (6.15) in the text. Note that

$$\begin{aligned} \bar{k} &= \frac{k}{\sigma-1} - 1 \\ \frac{d\bar{k}}{\sigma} \frac{\sigma}{\bar{k}} &= -\frac{\sigma}{\sigma-1} \frac{k}{k-(\sigma-1)} < 0. \end{aligned}$$

Proof of Proposition 6 (Relative prevalence). The relative prevalence of export modes $\chi_{ij} \equiv S_{ij}^M / S_{ij}^F$ follows immediately from (6.14) and (6.15)

$$\begin{aligned} \chi_{ij} &= \frac{\beta_{ij}^{\sigma-1} \left[\tilde{\beta}_{ij}^{k-(\sigma-1)} - \left(\frac{1-\tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij}-1}\right)^{\bar{k}}\right]}{\left(\frac{1-\tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij}-1}\right)^{\bar{k}}} \\ &= \beta_{ij}^{\sigma-1} \left[\tilde{\beta}_{ij}^{k-(\sigma-1)} \left(\frac{\phi_{ij}-1}{1-\tilde{\beta}_{ij}^{\sigma-1}}\right)^{-\bar{k}} - 1\right] \\ &= \beta_{ij}^{\sigma-1} \left[\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right)^{\bar{k}} - 1\right]. \end{aligned}$$

Comparative statics results are derived as follows:

$$\frac{d\chi_{ij}}{d\phi_{ij}} \frac{\phi_{ij}}{\chi_{ij}} = \bar{k} \frac{\phi_{ij}}{\phi_{ij}-1} \frac{\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right)^{\bar{k}}}{\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right)^{\bar{k}} - 1} > 0$$

$$\frac{d\chi_{ij}}{d\beta_{ij}} \frac{\beta_{ij}}{\chi_{ij}} = (\sigma-1) \left[1 + \frac{\frac{\tilde{\beta}_{ij}^{1-\sigma}}{\tilde{\beta}_{ij}^{1-\sigma}-1}}{\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right)^{\bar{k}} - 1} \bar{k} \frac{d\tilde{\beta}_{ij}}{d\beta_{ij}} \frac{\beta_{ij}}{d\tilde{\beta}_{ij}}\right] > 0,$$

since $\frac{d\tilde{\beta}_{ij}}{d\beta_{ij}}\frac{\beta_{ij}}{d\tilde{\beta}_{ij}} > 0$ and $\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1} > 1$ (Lemma 2).

$$\frac{d\chi_{ij}}{d\bar{\tau}_{ij}}\frac{\bar{\tau}_{ij}}{\chi_{ij}} = 0$$

$$\frac{d\chi_{ij}}{d\sigma}\frac{\sigma}{\chi_{ij}} \equiv \zeta = \sigma\ln\beta_{ij} + \bar{k}\frac{\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right)^{\bar{k}}\left(\frac{d\bar{k}}{d\sigma}\frac{\sigma}{\bar{k}}\ln\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right) + \frac{\tilde{\beta}_{ij}^{1-\sigma}}{\tilde{\beta}_{ij}^{1-\sigma}-1}\left(\sigma\ln\beta_{ij} + (\sigma-1)\frac{d\tilde{\beta}_{ij}}{d\sigma}\frac{\sigma}{\tilde{\beta}_{ij}}\right)\right)}{\left(\frac{\phi_{ij}-1}{\tilde{\beta}_{ij}^{1-\sigma}-1}\right)^{\bar{k}} - 1} < 0$$

since $\frac{d\bar{k}}{d\sigma}\frac{\sigma}{\bar{k}} < 0, \frac{d\tilde{\beta}_{ij}}{d\sigma}\frac{\sigma}{\tilde{\beta}_{ij}} < 0, \ln\beta_{ij} < 0$.

Once you assume $\lambda_{ij} = f(\sigma)$ with $f\prime > 0$, $\frac{d\beta_{ij}}{d\sigma}\frac{\sigma}{\beta_{ij}} > 0$, and we have

$$\frac{d\chi_{ij}}{d\sigma}\frac{\sigma}{\chi_{ij}} = \zeta + (\sigma-1)\frac{d\beta_{ij}}{d\sigma}\frac{\sigma}{\beta_{ij}},$$

the sign of which is unclear.

Derivation of equation (6.17). The stock of wholesale FDI is given by

$$\begin{aligned}
F_{ij} &= \phi_{ij} f_j M_i^E \left[1 - G\left(\Phi_{ij}^F\right)\right] \\
&= \phi_{ij} f_j M_i^E \left(\Phi_{ij}^F\right)^{-k} \\
&= \phi_{ij} f_j M_i^E \left((w_i\bar{\tau}_{ij})^{\sigma-1}\left(\frac{\phi_{ij}-1}{1-\tilde{\beta}_{ij}^{\sigma-1}}\right) f_j B_j^{-1}\right)^{-\frac{k}{\sigma-1}} \\
&= \phi_{ij} M_i^E (w_i\bar{\tau}_{ij})^{-k} B_j^{\frac{k}{\sigma-1}} f_j^{-\bar{k}} \left(\frac{1-\tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij}-1}\right)^{\frac{k}{\sigma-1}} \\
&= \phi_{ij}\Psi_{ij}\frac{k-(\sigma-1)}{\sigma k}\left(\frac{1-\tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij}-1}\right)^{\frac{k}{\sigma-1}} \\
&= \phi_{ij}\Psi_{ij}\rho\frac{\bar{k}}{k}\left(\frac{1-\tilde{\beta}_{ij}^{\sigma-1}}{\phi_{ij}-1}\right)^{\frac{k}{\sigma-1}}.
\end{aligned}$$

6.7 Appendix B. Summary statistics

Characteristics of U.S. exports by industry

BEA	Description	χ	DISP	SIGMA	FREIGHT
371	Motor Vehicles	0.2	2.2	7.1	0.10
305	Rubber	0.9	1.6	3.6	0.12
341	Metal Cans, Fabricated Metal	0.9	1.9	4.9	0.09
283	Drugs	1.0	2.1	9.5	0.03
363	Household Appliances	1.2	2.5	5.9	0.06
386	Optical and Photographic Equipment	1.3	1.8	8.1	0.05
357	Computers	1.3	2.0	11.0	0.04
284	Soap and Cleansing Products	1.3	1.9	5.5	0.09
308	Miscellaneous Plastics	1.5	1.6	6.0	0.13
287	Agricultural Chemicals	1.6	1.6	1.3	0.18
343	Heating and Plumbing Equipment	1.7	1.8	4.4	0.13
381	Scientific and Measuring Equipment	2.0	2.3	6.7	0.05
220	Textiles	2.1	1.8	7.8	0.11
366	Audio,Video,Communications Equipment	2.1	2.0	9.4	0.04
329	Stone, Minerals, and Ceramics	2.2	1.5	2.7	0.11
250	Furniture	2.3	1.7	3.6	0.16
310	Leather	2.4	1.7	8.9	0.07
351	Engines and Turbines	2.5	2.6	7.9	0.06
356	General Industrial Machinery	2.6	1.7	7.0	0.07
354	Metalworking Machinery	3.2	1.4	8.1	0.06
230	Apparel	3.6	1.6	5.6	0.09
355	Special Industrial Machinery	3.9	1.6	8.5	0.07
331	Ferrous metals	4.5	1.9	3.5	0.12
379	Other Transport Equipment	4.7	1.7	7.4	0.05
262	Pulp and Paper	5.2	1.3	4.3	0.14
335	Non-Ferrous metals	5.8	1.5	6.7	0.06
240	Wood and Lumber	6.1	1.5	4.0	0.15

Note: Ranked by prevalence of export sales via related relative to non-related parties

Characteristics of U.S. exports by destination

ISO	Name	χ	RISK	DIST	POP
AUT	Austria	0.9	55.3	6.8	8.1
MEX	Mexico	1.0	57.8	3.4	100.0
CAN	Canada	1.0	54.5	0.5	31.1
NLD	Netherlands	1.3	53.0	5.9	16.1
BEL	Belgium	1.4	56.7	5.9	10.3
JPN	Japan	1.4	58.5	10.9	127.0
GER	Germany	1.7	55.4	6.0	82.4
IRL	Ireland	1.7	57.3	5.1	3.9
GRC	Greece	2.2	57.3	7.9	11.0
ISR	Israel	2.3	65.4	9.1	6.3
FRA	France	2.3	55.5	5.8	59.7
GBR	United Kingdom	2.5	53.3	5.6	59.2
ARG	Argentina	3.2	76.0	8.5	37.3
AUS	Australia	3.2	59.6	16.0	19.6
SGP	Singapore	3.4	53.4	15.4	4.1
PHL	Philippines	3.5	62.1	13.7	78.7
MYS	Malaysia	3.6	67.8	15.1	24.3
ITA	Italy	3.9	56.9	6.9	57.3
THA	Thailand	4.0	65.2	13.9	63.2
DNK	Denmark	4.1	56.3	6.2	5.4
FIN	Finland	4.1	55.7	6.6	5.2
BRA	Brazil	4.1	70.6	7.7	175.0
COL	Colombia	4.3	70.5	4.0	43.5
HKG	Hong Kong	4.7	58.4	13.0	6.7
ZAF	South Africa	4.8	61.9	12.6	45.3
SWE	Sweden	5.2	56.5	6.3	8.9
PRT	Portugal	5.6	55.2	5.4	10.3
ESP	Spain	5.8	53.6	5.8	41.1
VEN	Venezuela	5.9	77.4	3.4	24.7
TWN	Taiwan	5.9	56.3	12.5	22.5
CHL	Chile	6.1	59.7	8.3	15.5
CHE	Switzerland	6.2	57.8	6.3	7.2
NZL	New Zealand	6.3	56.7	14.5	3.9
TUR	Turkey	7.4	67.2	8.1	67.8
NOR	Norway	8.0	58.8	5.9	4.5
KOR	South Korea	8.4	64.8	11.1	47.5
IDN	Indonesia	8.5	80.9	16.2	209.0
PER	Peru	8.8	69.1	5.9	25.8

Note: Ranked by prevalence of export sales via related relative to non-related parties; RISK takes scores between 0 (lowest) and 100 (highest); DIST in thousands; POP in millions;

6.8 Detailed econometric results

Table 6.1: Cross-industry regressions

	OLS (1)	IV (2)	OLS (3)	OLS (4)	OLS (5)	IV (6)	OLS (7)	IV (8)	OLS (9)	OLS (10)	IV (11)
$DISP$	-0.751*	-1.760***			-0.723*	-1.815***	-0.560	-1.872***		-0.541	-1.768***
	(0.424)	(0.599)			(0.422)	(0.638)	(0.426)	(0.713)		(0.432)	(0.660)
$\ln SIGMA$			-0.220		-0.0575	0.188			0.358	0.328	0.261
			(0.256)		(0.224)	(0.258)			(0.320)	(0.308)	(0.331)
$\ln FREIGHT$				0.458*			0.281	-0.131	0.738*	0.545	0.106
				(0.261)			(0.260)	(0.320)	(0.405)	(0.401)	(0.411)
RMSE	1.033	1.061	1.055	1.039	1.034	1.063	1.027	1.070	1.034	1.024	1.058
R^2	0.299	0.232	0.269	0.292	0.299	0.229	0.308	0.219	0.299	0.314	0.236
Partial R^2		0.423				0.420		0.351			0.353

N = 1026. Cluster-robust standard errors in parentheses. ***, **, * denote statistical significance at the ten, five, one percent levels of significance, respectively. All regressions include country fixed effects and a constant (not shown). In columns (2), (6), (8), and (11), the US dispersion measure $DISP$ is instrumented by four dispersion measures for European countries; see HMY. Partial R^2 refers to the first-stage regression.

Table 6.2: Cross-country regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\ln RISK$	1.411*			1.366	1.985**		1.943**
	(0.814)			(0.822)	(0.884)		(0.898)
$\ln DIST$		-0.138		-0.118		-0.127	-0.0963
		(0.133)		(0.123)		(0.139)	(0.126)
$\ln POP$			-0.186*		-0.224***	-0.184*	-0.222***
			(0.0923)		(0.0815)	(0.0921)	(0.0809)
$NAFTA$	-1.322***	-1.801***	-1.470***	-1.557***	-1.158***	-1.716***	-1.352***
	(0.202)	(0.280)	(0.132)	(0.316)	(0.184)	(0.281)	(0.320)
$\ln MR$	-0.0161	-0.155	-0.257**	-0.0395	-0.119	-0.277**	-0.137
	(0.116)	(0.0979)	(0.108)	(0.115)	(0.109)	(0.105)	(0.110)
RMSE	1.000	1.006	0.991	0.999	0.977	0.990	0.977
R^2	0.337	0.330	0.350	0.339	0.368	0.352	0.369

N = 1026. Cluster-robust standard errors in parentheses. ***, **, * denote statistical significance at the ten, five, one percent levels of significance, respectively. All regressions include industry fixed effects and a constant (not shown). Method of estimation: OLS.

6.9 Industry concordance

BEA	Description	SITC	NAICS
220	Textiles	65	313, 314
230	Apparel	84	315
240	Wood and Lumber	63	321
250	Furniture	82	337
262	Pulp and Paper	64	3221
283	Drugs	54	3254
284	Soap and Cleansing Products	55	3256
287	Agricultural Chemicals	56	3253
305	Rubber	62	3262
308	Miscellaneous Plastics	57,58	3261
310	Leather	61	316
329	Stone, Minerals, and Ceramics	66	3271
331	Ferrous metals	67	3211
335	Non-Ferrous metals	68	3314
341	Metal Cans, Fabricated Metal	69	3321,3324,3363
343	Heating and Plumbing Equipment	81	3323,3334
351	Engines and Turbines	71	3336
354	Metalworking Machinery	73	3335
355	Special Industrial Machinery	72	3332
356	General Industrial Machinery	74	3339
357	Computers	75	3341
363	Household Appliances	77	3352
366	Audio,Video,Communications Equipment	76	3342,3343
371	Motor Vehicles	78	3361
379	Other Transport Equipment	79	3362,3364-3366,3369
381	Scientific and Measuring Equipment	87	3345
386	Optical and Photographic Equipment	88	333314,333315

Chapter 7

Trade Intermediation Versus Wholesale FDI in General Equilibrium[1]

7.1 Introduction

Firms wishing to export their products to foreign markets either require a *local foreign partner*, who acts as a 'general importer' or a trade intermediator. Or they need deep pockets to set up an *own sales representation.* The choice of export modes plays a key role in strategic management decisions and has received considerable attention in the academic business literature.

A series of articles in the *Journal of International Business Studies* has highlighted the overall importance of trade intermediation, and its relative prevalence across sectors (see, e.g., Peng and Ilinitch, 1998, Peng and York, 2001, and Trabold, 2002). There is also evidence on the huge importance of trade intermediation in history (Greif, 1993) and for small specialized economies such as Hong-Kong or Singapore (Feenstra and Hanson, 2004; Feenstra, Hanson, and Lin, 2004). On the other hand, Kleinert and Toubal (2005, 2006) document the empirical importance of wholesale affiliates as a specific form of foreign direct investment. Fryges (2007) reports that sizeable shares of firms select into different export modes. Recently, starting with Rauch (1999), there is a growing literature on the role of formal and informal networks for the determination of bilateral trade volumes. Empirical

[1]This Chapter is based on an article forthcoming in Workshop Proceedings, see Felbermayr and Jung (2009). The concept for the paper was developed jointly, theoretical analysis and writing were equally shared.

evidence presented by Rauch and Trindade (2002) and Combes et al. (2005) lends support to the idea that the international matching of buyers and sellers involves important frictions.[2]

Despite the strong empirical evidence, trade intermediation and wholesale affiliates do not play any role in canonical trade models. The older literature ignores trade costs altogether; the new trade models pioneered by Krugman (1979) have taken variable trade costs serious. Only very recently, Melitz (2003) models fixed costs of foreign market access ('beachhead costs'; see Baldwin, 1988), which can be interpreted as foreign direct investment in wholesale affiliates. However, his model does not allow for trade intermediation as an alternative mode of exporting.[3]

In this paper we model the choice between the indirect (intermediated) and the direct (through own sales affiliate) export modes. In the first mode, producers save on fixed market access costs but loose discretion over pricing in the foreign market to their partner. Moreover, searching for a partner is costly and takes time. In the second mode, producers have to set up a foreign affiliate. The advantage of that mode is that they retain control over the consumer price of their product. We model the search-and-matching process between business firms (*business-to-business (B2B) matching*) using a matching function approach familiar from the labor market literature Pissarides (2000). This approach has been introduced into international economics by Grossman and Helpman (2002), who focus on vertical supply chains. In that setup, search costs are a function of the tightness of the market, which, in turn, depends on the endogenous decisions of both, producers and general importers, to search for a partner.

We embed the export mode choice in a general equilibrium trade model with heterogeneous firms à la Melitz (2003). We offer a slight generalization of Melitz, by allowing firms to differ in terms of the tradability of their goods, their strength of brand name, and their productivity. This framework allows to reproduce important stylized facts on the importance of trade intermediation relative to own affiliates for heterogeneous firms.

---

[2] Egan and Mody (1993), Hakansson (1982), and Turnbull and Cunningham (1981) provide descriptive studies on bilateral buyer-seller links in international trade. They report suggestive evidence on highly collaborative, long-lasting trade relationships between producers and intermediators in the manufacturing sector. Schröder et al. (2005) offer a partial equilibrium model of trade intermediation.

[3] There are a number of papers in the industrial organization tradition that study the choice of export modes in partial equilibrium (e.g., Raff and Kim, 2005). However, these models do not allow drawing conclusions on aggregate variables. Nor do they easily lend to empirical verification.

Our approach is formally related to Helpman et al. (2004), who study horizontal FDI in a model of the proximity concentration tradeoff. It differs since multinational enterprises do not engage in foreign production. Rather, the focus is on the matching between producers and general importers, and on the relative prevalence of export modes.[4]

Matching between producers and specialized importers is not immediate. This fact has a crucial implication: when parties finally match, they are locked into a *bilateral monopoly* situation which makes them vulnerable to *hold-up* from the other partner. We assume that the only commitment that producers can make is to engage in *exclusive dealership arrangements*. Otherwise, as in Grossman and Helpman (2002), no enforceable contracts exist. Hence, the price at which the producer sells to the general importers is determined through bilateral Nash bargaining. While the general importer has full discretion to set the price in the foreign market, the producer decides about the supplied quantity. The outcome of that game is that trade intermediation drives up the consumer price in the foreign market. The additional markup is given by the inverse of the producer's bargaining power and measures how strongly the producer's quantity decision reaches through to the foreign consumer price. Hence, variable profits are lower when exporting involves a general importer.

The rate at which producers and firms match depends on market tightness, i.e., the number of searching general importers relative to the number of searching producers. Tightness is driven by producers' and general importers' endogenous decisions to engage into costly search. As in all matching approaches, the matching friction involves a departure from first best, since there is an uninternalized search externality: entry of general importers (producers) drives up the expected cost of general importers (producers) to find a partner.

The mechanism studied in this paper is a promising candidate to square theory with models, see the work of Alessandria (2004) and Drozd and Nosal (2007) in international real business cycle models, as well as Reed and Trask (2006) in a homogeneous firms trade

[4] Our framework is also related to recent work by Rauch and Watson (2003) and Casella and Rauch (2002), who stress the importance of B-2-B relationships. Compared to those papers, our model is dynamic, features heterogeneous firms, allows for firms to differ with respect to their preferred foreign export mode, and determines the number of general importers and exporters endogenously. Most importantly, our model endogenizes foreign market access costs, since the cost of searching for a foreign general importer is endogenous.

model. It also provides a point of departure for a companion paper (see Felbermayr and Jung, 2008c).

The main result of the present paper is that in equilibrium, producers are endogenously selected into the two export modes according to the characteristics of their products. Firms with high levels of *productivity*, easily *tradable variants*, or strong *brand reputation*, establish own subsidiaries. Firms with intermediate values of the above characteristics choose to search for general importers. Along the steady state, only a fraction of those firms actually is matched and produces for the export market. Intermediation helps producers with good product characteristics to save on fixed foreign market access costs; however, this translates into lower overall export sales, thereby–at least partly–rationalizing the missing trade puzzle.

Moreover, related to the last observation, we find that institutional change may lead to a lower aggregate productivity, since exporters that switch from the direct to the indirect mode achieve smaller export sales, thereby contributing less to per capita GDP, and since relatively unproductive firms start exporting, drawing weight in the calculation of average GDP.

The remainder of the paper falls into four sections. Section 7.2 gives a short overview over stylized facts, while Section 7.3 introduces the analytical framework and derives a first lemma on the pricing behavior under trade intermediation. Section 7.4 shows the conditions under which a strictly positive share of the total mass of producers export through trade intermediation. Holding aggregate variables constant, it uses a graphical device to discuss the equilibrium sorting of firms obtained in our model. Section 7.5 sketches the free entry conditions of producers and general importers, and discusses theoretical extensions. Finally, Section 7.6 concludes.

7.2 Stylized facts

In this section we discuss a few striking stylized facts. Statistical information on the importance of different export modes is difficult to obtain. However, combining information from the MIDI Database entertained at the German Bundesbank, export sales data from the German Statistical Office, and data from a survey undertaken by the ZEW, a German research institute, we are able to sketch the broad picture. The key fact is that direct contact of a producer in one country with the end user in another country is quantitatively not important. Similar patterns exist in the U.S. (Bernard, Jensen, and Schott, 2008), or in France (Trabold, 2002).

Figure 7.1 shows the distribution of German manufactured goods export sales over different export modes. Sales via own affiliates in foreign countries amount to over 50% of total exports, with sales via foreign intermediators accounting for another 40%. The residual is direct exports that does not involve foreign direct investment nor a foreign general importer. There are a number of empirical problems, since total export sales by goods provided by the statistical office cannot exactly be mapped into the classification of sectors provided by the Bundesbank. In figure 7.1 we choose to present the conservative case, where producer-to-consumer exports are most likely overestimated.

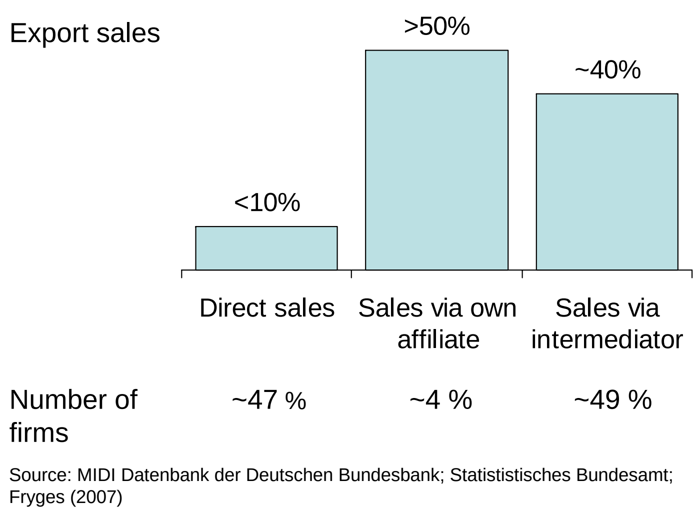


Figure 7.1: Relative prevalence of export modes, Germany, 2003

Figure 7.1 also reports the share of actively exporting firms in each mode. This information draws on survey results presented in Fryges (2007). Most producers export either through an intermediator (49%) or directly to the final client in the foreign country (47%). Only 3% engage in FDI. At first glance, these results seem to contradict our findings on shares in in total export volumes. However, taking the data at face value, they imply that the largest share of exports is undertaken by a small number of firms. There is large empirical evidence that this is actually the case (Bernard, Jensen, and Schott, 2008).

Fryges (2007) documents another important fact, namely that the number of firms that maintains own sales affiliates in foreign countries has increased between 1997 and 2003. This finding comes from a survey of German firms, but it has been replicated in an independent study for the United Kingdom. While in general the number of firms per se is not indicative of the total export volume channeled through some export mode, the fact that own affiliates are the prevailing choice for large firms suggests that also the share of exports channeled through affiliates has increased over time.

The implications of figure 7.1 can be summarized as follows: (i) Direct sales from the producer to a foreign end client amount to less than 10% of German exports, and are therefore quantitatively negligable. Exporters require either an own foreign sales affiliation or a foreign partner. Moreover, the share of exports through own affiliates has increased over time. (ii) It follows that fixed costs of foreign market access must have important aggregate implications, since the largest share of exports involves some type of fixed costs. (iii) A few firms make up a large share of total export sales. This points to a strong degree of heterogeneity amongst exporters.

In 2005, the stock of outward FDI of the entire German manufacturing sector amounted to a total of 223 billion Euro. About half that sum (104 billion) was invested in some foreign affiliate active in the manufacturing sector. Some 32% (71 billion) was parked in holding companies, or financial affiliate. The remaining 17% (38 billion Euro) were held in affiliate trading companies. Taking out holding companies and the finance sector, German manufacturing firms held about a quarter (27%) of their total FDI in companies classified in the trading sector. While that number includes also investment into foreign purchasing units, it is largely dominated by sales representations, as vertical FDI makes up only a small share of total German outward FDI.

Looking at the sectoral distribution of the quantitative importance of FDI into sales affiliates, one finds that the share of FDI invested in sales affiliates relative to total non-finance investment is highest in the mechanical engineering sector (about 36% on average over 2002-2005) and the automotive sector (34% on average), while it is rather low in the chemical (18 percent) or the electric power equipment industries (11%). Figure 7.2 illustrates the cross-sectoral pattern and also shows that over 2001-2005 that pattern was fairly stable.

Figure 7.2: German outward FDI in sales affiliates in percent of industry (non-finance) total

Regarding the geographical dimension of German outward FDI, the Bundesbank publication allows to distinguish between the stock of FDI invested in the USA, EU 25, and the rest of the world. Taking averages over the reported 2002-2005 time period, the share of investment in trade affiliates in total FDI of the manufacturing sector (again, excluding finance), amounts to about 27% for the EU 25, 26% for the US, and again 27% for the rest of the world.

We may summarize: A substantial share of total outward foreign direct investment (FDI) goes into the establishment or acquisition of foreign sales affiliates. There is little variation across the US, Europe, and the rest of the world, but significant sectoral variation.

Facts 1 and 2 establish the importance and relative prevalence of own sales affiliates. Empirical information on the role of general importers is more difficult to find. Trabold (2002) is amongst the rare studies that offer quantitative information. His empirical analysis draws on French customs data. His findings can be summarized as follows: Import intermediation by general importers is most prevalent (i) the farther away in terms of geography and culture an export market market is, and (ii) the lower the marketing-intensity of a product is. Moreover, (iii) the share of total exports that involve import intermediation has been falling during the 1980ies.

Our model can reproduce the stylized facts highlighted above. It is, however, also consistent with the broader evidence on the importance of networks, and search externalities discussed in the introduction.

7.3 Model setup

We study a model with two symmetric countries. Following Helpman et al. (2004), in each country there are two active sectors: a perfectly competitive numéraire sector, with unit labor input coefficients and costless tradability; and a differentiated goods sector, with heterogeneous firms operating under conditions of monopolistic competition.

7.3.1 Demand structure

Each country i is populated by a representative household, which inelastically supplies L units of labor to a perfectly competitive labor market. The household derives utility from consuming z units of the numéraire good, and a basket of differentiated goods. We assume that preferences are separable over those two items, with an upper Cobb-Douglas nest, and the basket of differentiated goods a Dixit-Stiglitz aggregate:

$$U = (1-\mu)\ln z + \frac{\mu}{\rho}\ln \int_{\omega\in\Omega_i} \left[\zeta\left(\omega\right) x\left(\omega\right)\right]^{\rho} \mathrm{d}\omega. \tag{7.1}$$

The household spends the share $0 < \mu < 1$ on differentiated goods and the remainder on the numéraire. The set of available varieties in country i is given by Ω_i, with ω denoting a generic variety.[5] The parameter $0 < \rho < 1$ describes the degree of substitutability of any each pair of varieties. However, unlike in the standard Dixit-Stiglitz representation, consumers may attach different weights $\zeta\left(\omega\right) \geq 0$ to different varieties, reflecting the fact that varieties may contribute asymmetrically to overall utility. We refer to $\zeta\left(\omega\right)$ as to the strength of variety $\omega' s$ brand name or the reputation of the producer. It may also be held to denote quality. In any case, a higher value of $\zeta\left(\omega\right)$ means that the respective variety yields a higher contribution to utility.[6]

The only source of income for the household is from wages, which we can normalize to unity in all countries thanks to our assumptions on the numéraire sector. Hence, the budget constraint reads

$$L \geq z + \int_{\omega\in\Omega_i} p\left(\omega\right) x\left(\omega\right) \mathrm{d}\omega. \tag{7.2}$$

Maximizing (7.1) subject to (7.2), we find the following demand function for a variety ω from country j

$$x\left(\omega\right) = H\frac{\zeta\left(\omega\right)^{\sigma-1}}{p\left(\omega\right)^{\sigma}}, \tag{7.3}$$

where $H \equiv \mu L/\left(\int_0^n \left[\zeta\left(\omega\right) p\left(\omega\right)\right]^{1-\sigma} d\omega\right)$ is proportional to country $i's$ market size L, n is the measure of the sets Ω_i and Ω_j, and $\sigma \equiv 1/\left(1-\rho\right) > 1$ is the elasticity of substitution

[5] Note that the set of available varieties differs across countries, since fixed costs of exporting prevents some varieties from being traded.

[6] Combes et al. (2005) offer a similar formulation of preferences. However, their ζ is constant across varieties imported from a given country.

between varieties.[7]

7.3.2 Heterogeneous production firms and export modes

Firms in the differentiated goods sector differ with respect to a vector of characteristics $\{\zeta(\omega), \tau(\omega), a(\omega)\}$, where $a(\omega) > 0$ denotes the marginal cost of producing variety ω, and $\tau(\omega) \geq 1$ refers to variety-specific variable distribution costs of the iceberg type, which occur regardless of whether a good is traded internationally or not. Whenever one unit of a variety is to be delivered to a foreign partner, $\tau(\omega)$ units of that good have to leave the gates of the producer's factory. We see $\tau(\omega)$ as a short-hand way to introduce marketing and distribution costs that arise when a good is sold. There is no reason to assume that those costs are zero for transactions when the producer and the consumer happen to reside in the same country. However, in international transactions, total variable trade costs are $\tilde{\tau}(\omega) = \bar{\tau}\tau(\omega)$, where $\bar{\tau} \geq 1$ accounts for transportation costs and may be thought of as a function of distance. We refer to $\bar{\tau}$ to the systematic component of trade costs, and of $\tau(\omega)$ as the idiosyncratic component. Note that the systematic component magnifies the idiosyncratic part; hence, more marketing-intensive goods are also more expensive to deliver to foreign markets. The importance of that source of heterogeneity has been recently emphasized by Bergin and Glick (2007).[8]

Producers are also heterogeneous with respect to their marginal costs of production, $a(\omega)$. With the wage rate normalized to unity, $a(\omega)$ is equal to the labor requirement for one unit of output. Heterogeneity along this line has been shown to be empirically relevant, and is core in much recent work following Melitz (2003). For producing $y(\omega)$ units, the firm ω faces incurs total production costs $c(\omega) = a(\omega) y(\omega) + f^D$, where f^D denotes the fixed costs of production..

In much of our analysis, we can summarize the vector of characteristics $\{\zeta(\omega), \tau(\omega), a(\omega)\}$ in a single scalar $A(\omega) \equiv a(\omega) \tau(\omega) / \zeta(\omega)$, since $A(\omega)$ is a sufficient statistic to describe firm behavior (see details below). Higher values of $A(\omega)$ are equivalent to higher marginal costs of production, lower tradability, and a lower degree of brand reputation. Following

[7]Note that by symmetry both sets Ω_i and Ω_j have the same measure n.

[8]However, in contrast to our formulation, his model has zero trade costs for deliveries within a same country.

Melitz (2003), the entry of producers requires payment of a cost f^E. Only after paying the entry fee do firms learn about their characteristics $A(\omega)$. We assume that $A(\omega)$ follows some c.d.f. $G(A)$. We can then rank firms with respect to their realization of A. The advantage of our broader definition of firm heterogeneity relative to the focus in the literature on productivity is that empirical evidence suggests that productivity (or, closely related to it, firm size) are poor predictors of exporting behavior once one controls for unobserved firm characteristics (such as $\zeta(\omega)$ or $\tau(\omega)$), see Fryges (2006).

A key object of the present paper is to understand the sorting of firms into different export modes along their A-dimension. The first mode–direct exports–requires the setup of a sales representation in the foreign country, which implies some additional fixed investment f^F. This is the situation studied by Melitz (2003). The investment f^F has been referred to by Baldwin (1988) as *beachhead costs*, and usually turns up in FDI statistics under the guise of wholly owned sales affiliates.[9]

The second export mode–indirect exports–requires a match with a specialized trade intermediator, which we call general importer (GI). GIs know the foreign market better than the foreign producer. Hence, fixed costs of market entry are lower for the GI. However, the producer has to invest into costly search for a GI and–once matched–looses control on the consumer price of its output. Along the A dimension, we focus on the empirically relevant case where producer with the lowest realizations of A (low marginal costs, high reputation, high tradability) chose the direct export mode, producers with lower-intermediate realizations go for the indirect export mode, producers with upper-intermediate realizations do not find it optimal to export in either mode, and producers with the highest values of A quit the market upon drawing their vector of characteristics. Before turning to a detailed description of the of the indirect export mode, we briefly discuss the monopolists' pricing problem for domestic and indirect export sales.

Operating profits from *domestic sales* are $\tau(\omega) \cdot H [\tau(\omega) p(\omega)]^{-\sigma} \zeta(\omega)^{\sigma-1} \cdot [p(\omega) - a(\omega)] - f^D$. The first part in that expression, $\tau(\omega)$, reflects the fact that domestic sales of x require

[9] The empirical literature on foreign direct investment (FDI) stresses the importance of wholesale affiliates (Kleinert and Toubal, 2006). Somewhat surprisingly, this fact has not provoked theoretical research; in theoretical models, FDI relates to foreign production activities carried out by some multinational firm (see Helpman, 2006). Our paper offers a theory of FDI into wholesale affiliates.

$\tau(\omega) x$ units of the respective variety to be produced. The second part, $H[\tau(\omega) p(\omega)]^{-\sigma} \zeta(\omega)^{\sigma-1}$, gives the level of demand that the household has for a variety ω with c.i.f. price $\tau(\omega) p(\omega)$. The third part, $[p(\omega) - a(\omega)]$, refers to the per unit margin of the price over marginal cost. To maximize profits, the firm sets the f.o.b. price $p(\omega) = a(\omega)/\rho$, where $1/\rho > 1$ is the markup over marginal costs. With our choice of preferences, the f.o.b. price does not depend on $\zeta(\omega)$. Inserting the optimal price in the monopolist's objective function, domestic profits can be written as

$$\pi^D(A) = BA^{1-\sigma} - f^D, \tag{7.4}$$

where it becomes apparent that profits depend only on $A(\omega)$ and not independently on the different components of $A(\omega)$. In the following we drop the dependence of A on ω since it is sufficient to know A in order to identify a specific producer. We follow Helpman, Melitz, and Yeaple (2004) and write profits in terms of $B \equiv (1-\rho) H_i \rho^{\sigma-1}$, which is an aggregate magnitude, that involves the endogenous price index and exogenous parameters. Clearly, profits from domestic sales decline in A since $1-\sigma$ is a negative number. They rise in B, which captures the size of the market, and fall in fixed costs of production, f^D.

The monopolist generates non-negative profits from *direct exporting*, if export revenues suffice to cover additional variable production costs and foreign investment f^F. The objective function now is $\tilde{\tau}(\omega) \cdot H[\tilde{\tau}(\omega) p(\omega)]^{-\sigma} \zeta(\omega)^{\sigma-1} \cdot [p(\omega) - a(\omega)] - f^F$. Maximum profits from direct exporting are

$$\pi^F(A) = B(\bar{\tau}_{ij} A)^{1-\sigma} - f^F, \tag{7.5}$$

where the systematic part of trade costs (independent from A), $\bar{\tau}$, appears as an additional determinant of variable profits, along with the foreign measure of market size B and the costs of investing abroad, f^F. Clearly, foreign profits are lower the higher the systematic component of trade costs.

7.3.3 Trade intermediation and general importers

Our slight generalization of the notion of firm heterogeneity apart, the setup discussed in section 7.3.2 above is the same as in Melitz (2003). In this section, we model the endogenous emergence of a new type of firms that misses in most standard trade models:

trade intermediators or, using our preferred term, general importers. Following Spulber (1999, p. 3), an intermediator is "...*an economic agent who purchases from suppliers for resale or who helps sellers and buyers to meet and transact*." We focus on the first function of a GI and on the matching problem between the GI and the producer of a certain variety. The second function refers to the activity of trade brokerage, where the intermediator confines to matching producers and consumers and does not incur any entrepreneurial risk. Trade brokers are empirically elusive institutions that are difficult to model.[10]

We can think of the GI as a firm that is located in a foreign market and has superior knowledge of local market conditions, legal institutions, idiosyncratic consumer preferences, etc. Hence, we assume that the GI has lower fixed costs of market access, f^M, than the direct exporter would have $\left(f^F\right)$. Without loss of generality, we may set $f^M = 0$, but refrain from doing so for the time being.[11]

A key complication when using a GI is that relationship-specific investment is needed. This comes in terms of search costs. Conceptually, search costs are essential to allow for a meaningful sorting of firms along the A dimension; if a producer would have free access to GI's comparative advantage (low market access costs), every active producer would use that opportunity. We model the emergence of GIs in equilibrium as an explicit trade-off between costs and benefits. In particular, we assume that both GIs and producers have to search for foreign varieties to import, and that this search is costly. Search costs arise due to the participation at international trade fairs, correspondence and direct contact to potential partners, etc. Search costs are endogenous, as they depend on the number of searching firms and GIs. When a search is successful, GIs and producers find themselves in a bilateral monopoly situation which endows the GI with market power that allows to recoup the search costs.

We assume that all firms are *single product firms*. While this is in line with most recent trade models, this assumption is not very realistic. In reality, many GI's have diversified

---

[10]The *raison d'être* of trade brokers is the existence of asymmetric information. This is an interesting issue in itself, which we take up in Felbermayr and Jung (2008c).

[11]One could also think that the GI's specific knowledge of the foreign market translates into lower variable (distribution) costs. While this is a theoretical possibility, it is clear the largest portion of variable distribution costs consists in tariffs and transportation costs, which in principle are the same across export modes. However, one could allow for the idiosyncratic component of trade costs $\tau(\omega)$ to differ across export modes. We discuss this extension in the conclusions.

product portfolios, possibly originating from different countries. In principle, the GI should take this fact into account when deciding about which price to charge to consumers, at least if the different goods are substitutes. If the GI in some country j controls a sufficiently large share of the market, it would internalize the cannibalization effect induced by additional varieties and charge a higher markup (Feenstra and Ma, 2007). In turn, this constitutes an incentive for GIs to expand. Apart from the pricing issue, multiproduct GIs may also benefit from economies of scope. The endogenous emergence of multi-product GIs is certainly worth to look at. However, it also lends to a number of additional complications, so that in the present paper we rule this possibility out.

To endogenize search costs, we follow the standard practice in search and matching models of unemployment (Pissarides, 2000) and assume the existence of a matching function. This approach has been fruitfully applied by Grossman and Helpman (2002) in a model of vertical supply chains. Our model differs in that we study exporting rather than sourcing behavior and allow for heterogeneous firms. Let n^S be the number of producers searching for an opportunity to export, and n^G the corresponding number of GIs searching for an opportunity to import goods. As long as they are unmatched, producers and GIs incur per-unit-of-time search costs c^P and c^G, respectively. At each instant, $N(n^S, n^G) \leq \min\left\{n^P, n^G\right\}$ trade relationships are formed, where $N(.,.)$ is linear-homogeneous, as well as increasing and strictly concave in both arguments.

We model GIs as *ex ante* identical; moreover, since producers differ with respect to their characteristics A, GIs are *ex post* heterogeneous. Firms' heterogeneity does not have any bearing on search costs, so that the rate at which a searching producer is matched with a GI does not depend on A. With our assumptions on the matching technology, matching rates depend only on the degree of *market tightness* $\theta \equiv n^G/n^P$, i.e., the number of searching GIs relative to searching producers. Exploiting the properties of $N(.,.)$, we can write the rate at which a producers are matched to a GI as $\eta(\theta) \equiv n^M(1, \theta)$ and the rate at which GIs are matched to producers as $\eta(\theta)/\theta$. Clearly, the concavity of $N(.,.)$ implies that $\eta(\theta)$ strictly increases in θ while $\eta(\theta)/\theta$ falls. This illustrates the standard search externality associated to entry of producers and GIs on their respective peers.

The empirical work of Besedes and Prusa (2006) suggests that in trade relations there is a substantial amount of turnover. We introduce this fact into our analysis by allowing for some exogenous separation rate $\delta^G > 0$. Moreover, to ensure convergence to an ergodic equilibrium distribution of productivities, we require an exogenous death shocks for producers, δ^P. If δ^G and δ^P are independent, the total rate of match destruction is $\delta \equiv \delta^P + \delta^G$.[12]

7.3.4 The game between producers and general importers

We consider a framework where no enforceable contracts can be written ex ante. Producers and GIs can credibly commit to a single promise: to stick to exclusive dealership arrangements. Expected search costs are $c^P/\eta(\theta)$ from the producer perspective and $c^G\theta/\eta(\theta)$ from the perspective of a generic GI. When a match happens to be formed, these costs are sunk. This implies that both parties find themselves in a situation of bilateral monopoly. Without this commitment, intermediated trade can only be an equilibrium outcome under very special circumstances. Otherwise, we follow Grossman and Helpman (2002) or Antras and Helpman (2004), assuming the bargaining over the joint surplus of a match to be an asymmetric Nash problem, where $\beta \in [0, 1]$ is the bargaining power of a producer.

The game implies the following staging: First, the producer decides about the quantity of output to provide to the GI. Second, both parties bargain about the joint surplus from selling the good at the foreign market at price $p^G(\omega)$. As usual, the game is solved by backward induction.

Denoting the joint surplus by $J(\omega)$, we have $J(\omega) = p^G(\omega)\, x\left[p^G(\omega)\right] - f^M$. At the time of the bargain, variable production costs (which also account for transportation costs) have already been incurred, so that they do not turn up in the *ex post* surplus. The Nash bargaining results in a sharing of the joint surplus according to the two parties' relative bargaining powers, where the producer appropriates $\beta J(\omega)$, and the general importer $(1-\beta) J(\omega)$.

Predicting its share of the surplus at the bargaining stage, the producer choses her optimal quantity to supply to the GI. She solves

$$\max_{x(\omega)} \beta J(\omega) - a(\omega)\, \tilde{\tau}(\omega)\, x\left[p^G(\omega)\right]$$

[12] Time is continuous. Hence, destruction rates and rates of match creation take values on the entire real line. The matching rates refer to the rate by which a match occurs in the next infinitesimally short time period. The death rates δ^P and δ^G relate to the survival rate into the next infinitesimally short time period.

subject to the demand function (7.3), taking into account that in order to supply a quantity x to the GI, she has to produce $\tilde{\tau}(\omega)\,x$ units of her variety, where $\tilde{\tau}_{ij}$ denotes the total iceberg transportation costs from shipping abroad. Plugging in the expression for $J(\omega)$, and using the inverse demand function derived from (7.3), the first order condition of the producer implies a pricing rule $p^G(\omega) = a(\omega)\,\tilde{\tau}(\omega)\,/\,(\beta\rho)$. Importantly, the standard markup $1/\varphi$ is magnified by an additional factor $1/\beta$ which is endogenously pinned down by the parameter governing bargaining between the producer and the GI.

Lemma 7. *The price charged for imports by a general importer (GI) is given by*

$$p^G(\omega) = \frac{1}{\beta\rho} a(\omega)\,\tilde{\tau}_{ij}(\omega)\,, \tag{7.6}$$

with $(\beta\rho)^{-1} > 1$ *the total markup over effective marginal costs.*

Proof. See the Appendix. □

As in Grossman and Helpman (2002), the consumer price indicated in Lemma (7) reflects the presence of *double marginalization*: the price paid by the foreign consumer is driven up by the usual markup $1/\rho$ earned by the GI, and by the markup $1/\beta$ that results from Nash bargaining. Note that the additional distortion depends on β : the larger the producer's bargaining power, the closer (7.6) comes to the price obtained if the producer would sell directly to the foreign market, i.e., $a(\omega)\,\tilde{\tau}_{ij}(\omega)\,/\rho$. Also note that the bargained transaction price is independent from the market tightness θ_{ij}, which is a direct corollary from the fact that both parties' outside options are driven to zero on the one hand by free entry of GIs and on the other hand by the absence of any alternative use of the output quantity delivered by the producer to the market j.

The value of the joint surplus can be obtained by substituting (7.6) into the definition of $J(\omega)$:

$$J(A) = \sigma B \left(\bar{\tau}\beta^{-1}A\right)^{1-\sigma} - f^M. \tag{7.7}$$

The joint surplus is larger the bigger the size of the export market adjusted for transportation costs $\bar{\tau}^{1-\sigma}B$, and the smaller the match-specific fixed costs f^M. The surplus is larger the stronger the producer's bargaining power β : the closer β is to unity, the smaller is the detrimental effect of double marginalization. Clearly, higher marginal costs, lower tradability and lower brand reputation also reduce the surplus, since they translate into a higher value of A.

Similarly, we can now express the additional profits from *selling abroad through a general importer* by inserting $p^G(\omega)$ into the producer's objective function:

$$\pi^{MP}(A) = \beta^\sigma B(\bar{\tau}A)^{1-\sigma} - \beta f^M. \tag{7.8}$$

Note that we use the superscript MP to make clear that only matched producers have access to those profits. When talking about producers' choice of export modes, we will have to link $\pi^{MP}(A)$ to the additional profits that a producer expects to make when engaging into the costly search for a partner.

Comparing (7.8) to $\pi_{ij}^F(A)$, the profits of direct exporting to the foreign market, it is clear that the term $B(\bar{\tau}A)^{1-\sigma}$ appears in both expressions. But, since $\beta^\sigma < \beta < 1$ for given distance-adjusted market size $B\bar{\tau}^{1-\sigma}$ and firm characteristics A, intermediated exporting (7.8) involves lower variable profits than direct exporting (7.5). However, fixed costs of direct exporting have to be shouldered by the producer alone, while fixed costs (if any) are shared by both parties in the indirect mode.

7.4 Choice of export modes with given market tightness

7.4.1 Zero cutoff profit conditions

Firms select endogenously into different export modes. However, as in the standard Melitz (2003) model, the presence of fixed production costs implies that some firms with the highest realizations of A will choose not to start production at all, and some firms with high values of A prefer to sell only on the domestic market. Finally, firms willing to export face a choice between direct exporting, which is fixed cost intensive but yields high unit revenues, and indirect exporting via a GI, which saves fixed costs but involves lower unit revenues. Hence, we expect that firms with intermediate realizations of A_i prefer indirect exports and those with lowest A sell directly through own sales affiliates. Under conditions to be made explicit below, there is a unique sorting of firms along their A characteristics, with all possible regimes being active in equilibrium. Firms with realizations $A > A^D$ have so high marginal costs, low brand reputation and tradability, that their revenue generated from the domestic market cannot suffice to cover the fixed costs of production. *A fortiori*, they cannot find it optimal to export, neither. Firms with characteristics $A_{ij}^{SP} < A \leq A^D$ produce only for the domestic market. Either way of serving the foreign market involves too high entry costs and too little revenue. Firms with characteristics $A^F < A \leq A^{SP}$ find it optimal to start searching for a GI. At any point in time, a fraction of those firms will be matched and therefore generating export revenues in top of domestic income. Firms with $A \leq A^F$, that is the best firms (with lowest marginal costs, highest tradability and strongest brand names) establish own sales affiliates.[13] Note that the same firm can find it optimal to serve different markets using different modes.

The thresholds A^D, A^{SP}, and A^F are determined by a series of indifference conditions, which, given the sorting described above, can be described by *zero cutoff profit conditions*. The marginal firm A_i^D that finds entry into operations worthwhile is defined by setting domestic profits (7.4) zero:

$$\left(A^D\right)^{1-\sigma} = \frac{f^D}{B}. \tag{7.9}$$

That threshold A_i^D is lower the higher f^D and the lower B, reflecting the fact that higher fixed costs and smaller market sizes make it harder for firms with bad (i.e., high) realizations to survive.

The value of A below which firms find it worthwhile to search for producers (and ultimately be matched to a GI) is slightly more involved to pin down, because of the inherently dynamic nature of the search and matching process: searching for a GI involves an uncertain investment, as the duration of costly search is uncertain. Hence, the producer has to trade off immediate search costs against future profits from foreign sales. Denote the value of a producer that searches for a GI by V^{SP} and the value of a matched producer by V^{MP}. Then, we can establish the following system of value equations:

$$\delta^P V^{SP}(A) = -c^P + \eta(\theta)\left[V^{MP}(A) - V^{SP}(A)\right], \tag{7.10}$$

$$\delta^P V^{MP}(A) = \pi^{MP}(A) + \delta^G\left[V^{SP}(A) - V^{MP}(A)\right]. \tag{7.11}$$

Since δ^P is the only source of discounting from the producer's perspective, $\delta^P V^{SP}$ is

[13] To break ties, we assume that firms that are indifferent between two regimes, chose the next highest (in terms of the ranking of regimes discussed above).

the flow return to searching. That return has to be equal to the flow costs of searching $-c^P$ and the expected capital gain when the search has been successful. That gain $\left[V^{MP}(A) - V^{SP}(A)\right]$ occurs with Poisson rate $\eta(\theta)$ so that equation (7.10) follows. In turn, the flow value of a matched producer $\delta^P V^{MP}$ is given by the flow profits of selling through a GI, $\pi^{MP}(A)$ and the expected capital loss of being separated from the GI, $\delta^G \left[V^{SP}(A) - V^{MP}(A)\right]$.

We can solve for V_{ij}^{SP} from the system (7.10) and (7.11), which yields an expression for the flow value of a searching producer:

$$\delta^P V^{SP}(A) = s(\theta)\,\pi^{MP}(A) - [1 - s(\theta)]\,c^P, \tag{7.12}$$

where the term $s(\theta) \equiv \eta(\theta)/\left[\delta + \eta(\theta)\right]$ denotes the average fraction of time that a producer expects to be matched and earning profits π^{MP} and $1 - s(\theta)$ is the fraction of time that she is searching and hence incurring search costs c^P. We determine the producer, who is just indifferent between engaging into searching for a GI and concentrating on exclusively domestic sales, by the condition $V^{SP}\left(A^{SP}\right) = 0$. Using the expression for profits $\pi^{MP}(A)$, (7.8) in (7.12), we obtain the zero cutoff profits condition for entry into search as

$$\left(A^{SP}\right)^{1-\sigma} = \frac{\bar{\tau}^{\sigma-1}}{\beta^{\sigma} B}\left[\frac{c^P}{\eta(\theta)} + \beta f^M\right]. \tag{7.13}$$

The effective fixed costs of foreign market access consist of two terms: expected total search costs $c^P/\eta(\theta)$ and the producer's share of match-specific fixed costs βf^M. The threshold A^{SP} is lower the higher the sum of those fixed costs is; i.e., the marginal searching producers needs to exhibit lower marginal costs, higher tradability and a stronger brand name. If the distance-adjusted market size $\bar{\tau}^{1-\sigma}B$ goes up, the threshold goes up. Similarly, when the size of the double marginalization distortion, captured by β, falls (i.e., β goes up), the threshold rises, and the marginal searching producer can features a worse realization of A.

Finally, we determine the remaining cutoff level, A^F, by solving $V^{SP}\left(A^F\right) = V^F\left(A^F\right)$. The marginal direct exporter is exactly indifferent between searching for a GI or establishing her own subsidiary. Equating (7.12) and (7.5), and using (7.8) one gets

$$\left(A^F\right)^{1-\sigma} = \frac{\bar{\tau}^{\sigma-1}}{B}\frac{f^F - [1 - s(\theta)]\,c^P}{1 - \beta^{\sigma} s(\theta)}. \tag{7.14}$$

Again, higher distance-adjusted market size $\bar{\tau}^{1-\sigma}B$ allows for firms with worse (i.e., higher) realizations of A to select into direct exporting. The higher the term $f^F - [1 - s(\theta)]\, c^P$, the higher are the opportunity costs of direct exporting relative to the next best alternative, and the lower the maximum realization of A can be. Also, the lower β, the larger is the double marginalization problem that arises in the indirect export mode, and the lower the threshold A^F becomes.[14]

7.4.2 Equilibrium sorting of firms over export modes

Before turning to a full general equilibrium analysis with θ and B endogenous, it is worthwhile to illustrate the sorting of firms over different regimes as a function of their characteristics $A^{1-\sigma}$ in Figure 7.3, which is a modified version of figure 1 in Helpman, Melitz, Yeaple (2004). Expressing flow profits as annuities using the producers' discount rate, we associate an 'expected profit line' $\delta^P V^{mode}$ to each mode, where *mode* either takes the value D (domestic sales only), SP (search for a GI) and F (direct exports through an own affiliate). Note that for modes D and F we have $\delta^P V^{mode} = \pi^{mode}$; this is however not true for the SP mode. The figure plots (7.4), (7.5), and (7.12), taking aggregate variables B and θ taken as constant.

The lines differ with respect to their respective intercepts (representing fixed costs) and slopes (representing net revenues for unit productivity). In the figure, the flow profits (7.4) associated to purely domestic operations have an intercept of $-f^D$ and slope B_i. Expected additional (on top of the profits from the home market) flow profits of searching for a GI involve expected fixed costs consisting of the producer's share in match-specific fixed costs and expected search costs, $f^G \equiv s(\theta)\,\beta f^M + [1 - s(\theta)]\, c^P$, and a slope $B\bar{\tau}^{1-\sigma}\beta^\sigma s(\theta)$. Finally, additional profits (7.5) from direct export sales involve fixed costs f^F and a slope $B_j\bar{\tau}^{1-\sigma}$. Clearly, the slope of the $\delta^P V^{SP}$ line is smaller than the one of the $\delta^P V^F$ line due to the existence of double marginalization, $\beta^\sigma < 1$ and due to the fact that positive sales revenue accrues only if the producer is actually matched to a GI, which is not always the case. The $\delta^P V_{ij}^D$ line is steepest: compared to the other regimes, marginal net revenues are

[14] For (7.14) to be well defined, i.e., $\left(A_{ij}^F\right)^{1-\sigma} > 0$, we need that $f^F - [1 - s(\theta_{ij})]\, c^{SP} > 0$. This implies $\delta c^{SP} / [\delta + \eta(\theta)] < f^F$, a condition that will be verified in Lemma 8 below.

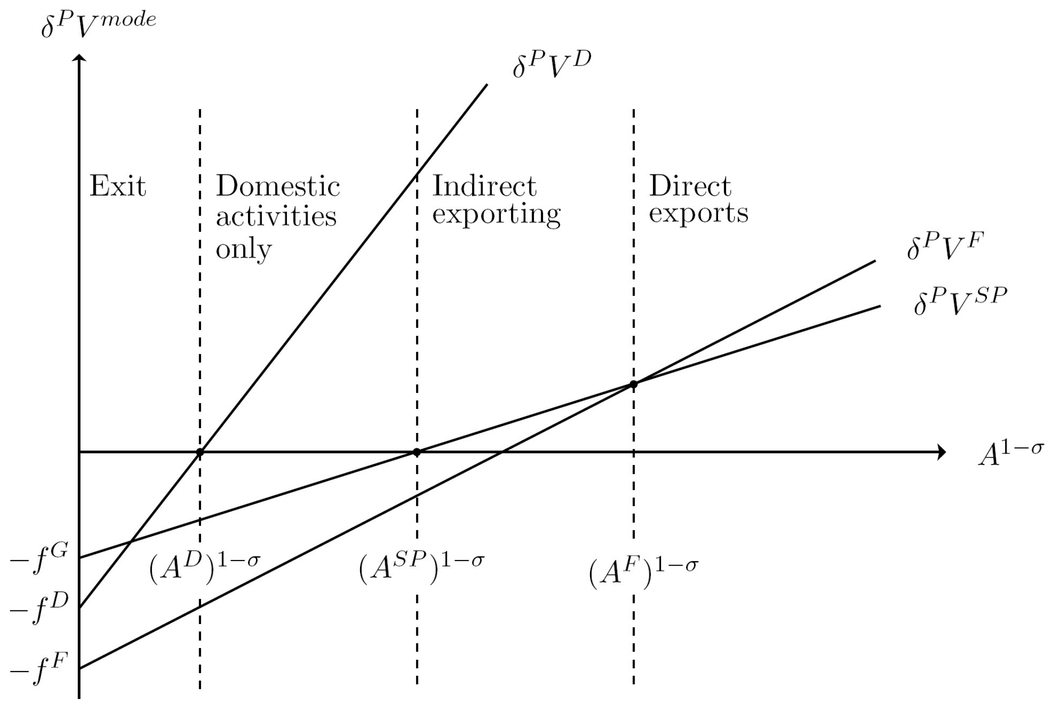


Figure 7.3: Equilibrium sorting for given tightness

higher as there are no transportation costs.

For given θ, a non-zero mass of firms is active in each of the three regimes (D, SP, F) if the hypothesized ranking $\left(A^D\right)^{1-\sigma} < \left(A^{SP}\right)^{1-\sigma} < \left(A^F\right)^{1-\sigma}$ holds. This requires that the effective fixed costs of searching for a GI lie in a bracket between the fixed production costs f^D and the costs of establishing an own foreign sales affiliate f^F.

Lemma 8. *For given market tightness θ, a partial sorting equilibrium exists if the following condition holds*

$$\bar{\tau}^{1-\sigma} f_D < \beta^{-\sigma} \left[\beta f^M + \frac{\delta c^P}{\eta(\theta)} \right] < f^F.$$

That is, strictly positive non-overlapping masses of producers find it optimal to sell domestically only and to sell both domestically and in the foreign market. Among exporters, there are strictly positive, non-overlapping masses of producers that search for a general importer and that own foreign sales subsidiaries.

Proof. See the Appendix. □

This lemma follows directly from using the definitions of $\delta^P V^D, \delta^P V^F$, and $\delta^P V^{SP}$ in Figure 7.3. Note that for a segmentation of firms into non-exporters and owners of own sales affiliates, it is enough that $\bar{\tau}^{1-\sigma} f_D < f^F$, which is exactly the respective condition in Melitz (2003). Also as in Melitz, we do not require the existence of variable trade costs

$\bar{\tau} > 1$; neither the sorting of firms into exporters and non-exporters, and the sorting of exporters into direct and indirect exporters hinges on $\bar{\tau}$. The only reason to allow for $\bar{\tau} > 1$ is for the purpose of conducting comparative statics.

Lemma 8 has a fairly intuitive interpretation. The term in square brackets amounts to the expected effective costs of accessing the foreign market j through a GI, since βf^M are match-specific fixed costs to be borne by the producer, and $\delta c^P / \eta(\theta)$ are the expected, annuitized search costs. The term $\beta^{-\sigma}$ that premultiplies effective expected search costs is related to the elasticity of expected profits of a searching producer with respect to $A^{1-\sigma}$. Hence, the lemma requires that adjusted expected costs of market access in the intermediate mode should neither be too larger nor too small. Clearly, we can restate Lemma 8 in terms of market tightness θ. If θ is high, producers find GIs quickly, expected search costs fall, and so do total effective GI-mediated access costs. However, as long as $f^M > 0$, indirect exporting remains viable, at least for some combinations of parameters, even if θ approaches infinity. However, if θ falls to zero, search costs become infinite and so do GI-mediated access costs: indirect exporting is no longer feasible. Hence, from the producers' perspective, Lemma 8 implies a lower bound for θ. However, for high θ, fewer GIs find it optimal to enter, which puts an upper bound on the equilibrium θ.

Note the difference of the proposed theory to the *proximity-concentration* model in Helpman et al. (2004). There, the sorting of firms into foreign direct investment and exports depends crucially on systematic transportation costs. In their model, as transportation costs fall, exporting becomes more attractive relative to local production. This is an empirically counter-factual implication (Neary, 2008), that our model does not have. Rather, a change in systematic transportation (distance) costs does not directly affect the sorting of firms into different export modes, but would has indirect implications through the market tightness (see below). However, since we allow firms to differ with respect to the genuine tradability of their varieties, we can make statements on how the idiosyncratic (variety specific) transportation costs affect the sorting of firms. We do so in the following proposition.

Proposition 6. *Under the condition stated in Lemma 8, producers endogenously select into export modes according to their product characteristics. Firms with high levels of productivity, easily tradable variants, or a strong brand reputation, establish own subsidiaries, while those with intermediate values of the above characteristics search for general importers.*

Firms with low values of the above characteristics do not export.

Proof. Directly follows from Lemma 8 and Figure 7.3. □

Figure 7.4 looks at the comparative statics of an increase in θ. From (7.12), both the slope and the intercept of the $\delta^P V^{SP}(A)$ line change. The reason is that a higher θ implies a higher matching rate for producers. Hence, the fraction of time that any producer is actually matched goes up. This leads to a stronger marginal effect of a change in $A^{1-\sigma}$: as firms have better characteristics, their export profits rise faster if they are more frequently matched. Hence, the slope of the (7.12) line is steeper if θ goes up. The effect on the intercept, however, is ambiguous. On the one hand, a higher θ rises the fraction of time in which a firm with characteristics $A^F < A \leq A^{SP}$ is matched and hence paying its share of match-specific costs βf^M. On the other hand, a higher θ also means that the firm finds itself less frequently paying search costs c^P. Whether the first effect dominates the latter depends on the sign of $\beta f^M - c^P$. Since $f^M = 0$ is perfectly compatible with a meaningful equilibrium but $c^P = 0$ is not, we set $f^M = 0$ in the following analysis.

Corollary 8. *If $f^M = 0$, an increase in market tightness θ makes indirect exporting more attractive relative to both, the purely domestic mode, and direct exports through own affiliates. That is, the lower cutoff in the indirect exports mode, $\left(A^{SP}\right)^{1-\sigma}$, falls while the upper cutoff, $\left(A^F\right)^{1-\sigma}$, rises.*

Proof. See the Appendix. □

7.4.3 Intermediation, the missing trade puzzle, and other implications

We can use figure 7.4 to discuss a number of interesting implications that result from the option of producers to export via GIs. To that end, we compare the standard Melitz (2003) model, in which intermediation is not a feasible option, to a model where that latter option exists. Lemma 8 suggests that there are several ways to render indirect exporting an option which is always dominated either by non-exporting or by exporting through own affiliates: either β is too small, or c^P and/or f^M are too high, or θ is too low. In all those cases, the intercept of the $\delta^P V^{SP}(A)$ line in figure 7.4 is so large (in absolute values), that the cutoff level $\left(A^{SP}\right)^{1-\sigma}$ does not exist. We focus on the case of a reduction in search costs c^P, either through technological change (the improvement of information and

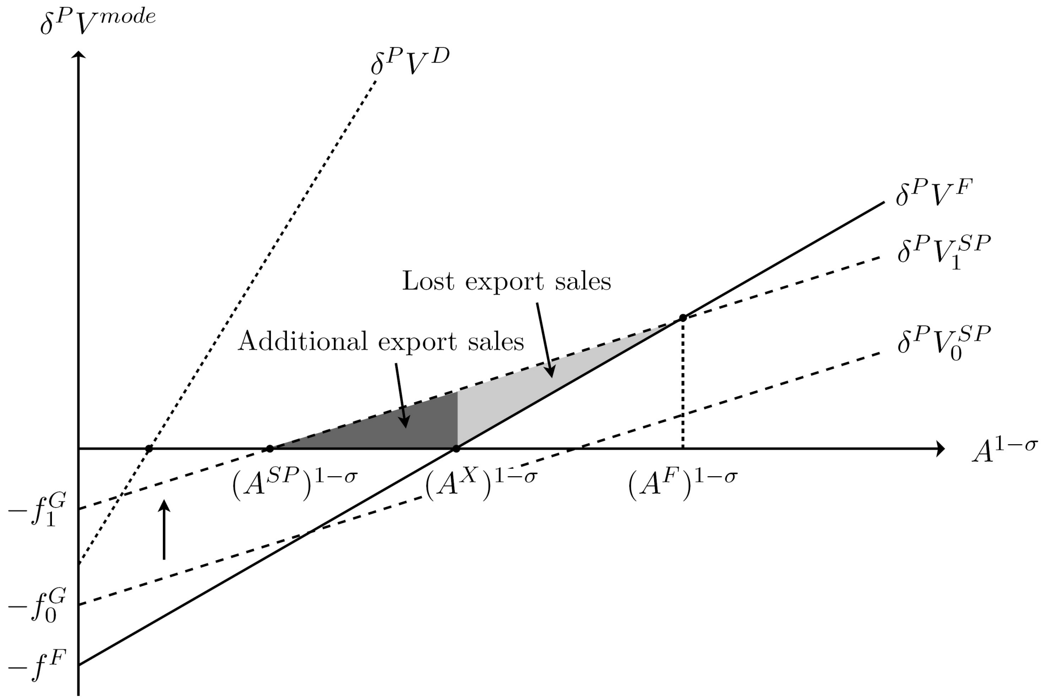


Figure 7.4: Increasing tightness and equilibrium sorting

communication technologies) or through measures of indirect trade promotion (e.g., through the construction and public maintenance of trade fairs, or trade missions in consulates or embassies).[15] There is ample empirical evidence for both facts, see Cummins and Violante (2002) and Rose (2007).

In figure 7.4, if c^P is prohibitively high, only three regimes exist: firms with the lowest values of A export, firms with intermediate values of A are active only domestically, and firms with the highest A never take up operations. Hence, the cutoff $\left(A^D\right)^{1-\sigma}$ is not affected by the parameter c^P. However, if c^P is prohibitively high, the exporting cutoff $\left(A_0^F\right)^{\sigma-1}$ is determined by the condition $\delta^P V^F\left(A_0^F\right) = 0$. This is the case where the $\delta^P V^{SP}(A)$ line cuts the x-axis.

When c^P falls, the intercept of the $\delta^P V^{SP}(A)$ starts to fall in absolute values, and at some point indirect exporting becomes an option for firms. This has two consequences. First, the 'best' firms (those with high $A^{1-\sigma}$) that have not exported before start selling abroad. This generates additional exports. Second, the 'worst' firms that have been export-

[15] Any change in c^P triggers an adjustment in θ if it is not offset otherwise. However, there exists a scalar λ such that $dc^P = \lambda dc^G$ for which θ remains constant even in full general equilibrium.

ing through an own affiliate before now prefer to use the GI instead. This switch of mode is optimal for producers: they give up some variable revenue, but in turn save fixed market entry costs (associated to FDI). Holding $A^{1-\sigma}$ constant, firms achieve higher export sales in the direct relative to the indirect mode. Hence, the switch into indirect exporting leads to a contraction of trade. The overall effect of the fall in c^P on total export values–new firms take up exporting, while switchers export less–is a priori ambiguous. In contrast to received wisdom, ignoring the existence of GIs and the mechanism discussed in this paper, the effect of technological or institutional change on trade can be smaller (and, theoretically, negative).

Another implication of the existence of GIs is that variance in c^P (or any other exogenous determinant of the $\delta^P V^{SP}(A)$ line) affects the exporting behavior of different types of firms differently. Business surveys reveal that there is sizeable cross-country variance in the export behavior of firms of given productivity. For example, while in Germany medium-sized companies are very active exporters, in France this is much less the case: only 5 percent of all small and medium sized firms in France export, while that number is 18 percent in Germany (The Economist, Feb 8th, 2007). On the other hand, large firms seem to achieve higher international sales in France than in Germany. Our model can relate this empirical fact to cross-country heterogeneity in the drivers of the expected fixed costs of exports through GIs. Exporters that for some reason face high expected costs of market access through GIs have less exporting firms, but those that export are on average more productive and, hence, larger.

Finally, and related to the last observation, we can use our model to make claims on the aggregate productivity of countries. Closing down $\tau(\omega)$ and $\zeta(\omega)$ heterogeneity, the emergence of GI intermediated exports makes large exporters that switch from the direct to the indirect mode achieve smaller export sales. Therefore, they contribute less to per capita GDP (which is proportional to a measure of average productivity). On the other hand, some relatively small firms that have preferred to sell domestically only, now find it optimal to export. They receive additional weight in the calculation of average GDP. Again, the overall effect is ambiguous. However, there is the possibility that the emergence of GIs actually lowers the aggregate productivity level. In other words, export promotion need not be good for GDP even if there are more exports. *A fortiori*, a welfare perspective that accounts for resources used in foreign market access, delivers an even bleaker picture.

7.5 Closing the model

In the above discussion, we have treated θ and real income level B as given. However, θ is itself an important endogenous variable, since it reflects the entry of GIs and producers into searching mode. Moreover, free entry of both GIs and producers is crucial to close the model: the free entry conditions hold in expectations so that entry occurs until expected profits are zero.

7.5.1 Free entry of GIs

Free entry of GIs implies that in an equilibrium situation, the expected gains from starting a new GI firm are just zero. That condition pins down the equilibrium number of GIs. When GIs decide to start searching for a foreign producer, they incur search costs. They are matched according to the matching technology described above, with $\eta(\theta)/\theta$ the Poisson arrival rate of a successful match. However, any GI faces *ex ante uncertainty* since the characteristics of the producer that it will ultimately be matched to are known only when the match has occurred. Clearly, since the size of the joint surplus is strictly decreasing in A, a GI is strictly better off with a partner featuring a lower A.

The value equations of a GI can be written as

$$\delta^G E[V^{SG}] = -c^G + \frac{\eta(\theta)}{\theta}\left(E[V^{MG}] - E[V^{SG}]\right), \qquad (7.15)$$

$$\delta^G E[V^{MG}] = (1-\beta)\, E\left[J(\omega)\right] + \delta^P \left(E[V^{SG}] - E[V^{MG}]\right), \qquad (7.16)$$

where $E[V^{SG}]$ denotes the expected value of a searching GI and $E[V^{MG}]$ that of a matched GI. As with producers, there is no discounting other than through the exogenous separation rate δ^G, which measures the rate at which a match is broken and the GI goes out of business. Equation (7.15) shows that the expected flow return to searching consists of a flow search costs $-c^G$, and a positive capital gain $E[V^{MG}] - E[V^{SG}]$, which materializes when the GI switches from searching to being matched. This happens with Poisson rate $\eta(\theta)/\theta$. Equation (7.16) shows that the expected flow value of a matched GI consists of the GI's share of the joint surplus generated in the match, $(1-\beta)\, E\left[J\left(\omega\right)\right]$, and the capital loss $E[V^{SG}] - E[V^{MG}]$, which happens when the producer is hit by an exogenous exit shock δ^P.

Free entry implies that the GIs' ex ante value of searching for a producer $E[V^{SG}]$ is zero. Using equation (7.15), this implies that the expected value of a matched GI $E\left[V^{MG}\right]$ just equals expected search costs of a GI $c^G\theta/\eta(\theta)$. Moreover, it follows from equation (7.16), that the expected value of a matched GI is equal to the GI's share of the joint surplus, appropriately discounted $E[V^{MG}] = \frac{1-\beta}{\delta} E\left[J\left(A\right)\right]$. Thus, the free entry condition for GIs is given by

$$\frac{c^G}{\eta(\theta)/\theta} = \frac{1-\beta}{\delta} E\left[J\left(A\right)\right]. \tag{7.17}$$

This condition equates the expected search costs of a GI on the left-hand-side with the present value of the share of the expected surplus that accrues to the GI.

Note that the GIs' entry decision is formally isomorphic to the producers decision whether or not to pay the fixed costs that reveal their characteristics A. However, while the producers draw from a sampling distribution $G\left(A\right)$, GIs sample the characteristics of their partners from a distribution that is endogenously truncated by the producers' decisions whether or not to search for a GI. Producers who have drawn characteristics $A \leq A^F$ find it optimal to establish a foreign sales representation. Firms with characteristics $A > A^D$ do not find it worthwhile to take up operations at all: their entry fee is simply foregone. In contrast, GIs always find it optimal to start cooperating with the producer $A \in \left[A^F, A^{SP}\right]$ that they have been randomly matched with. The reason for this is straightforward. A necessary and sufficient condition for producers to search for a GI is that their share of the surplus is larger than expected search costs, i.e., $\beta J\left(A\right) \geq \delta^P c^P/\eta\left(\theta\right) > 0$. GIs, in turn, take up cooperation with their producer if their share of the ex post surplus is non-negative, i.e. $(1-\beta)\, J\left(A\right) \geq 0$. Hence, the producers' condition is also sufficient for GIs not to refuse cooperation with a randomly matched producer. Search specific fixed costs f^M are collectivated in the bargaining process and are therefore paid by both parties in the match. It follows that in a rational expectations equilibrium, the criterion of producers to enter into searching for a GI, and of GIs not to reject a successfully matched producer, coincide. We summarize this finding in the following lemma.

Lemma 9. *In equilibrium, a general importer never finds it optimal to reject a producer once a match has occurred.*

Proof. In the text. □

At this point, the crucial assumption that producers can credibly commit to *exclusive dealership arrangements* becomes clear. The problem without such an arrangement is that producers have an incentive to sell to more than one GI, since competition among GIs would allow them to sell larger quantities to the foreign market. However, if one variety is sold by at least two importers, they would enter into Bertrand competition. This would annihilate any ex post profits so that GIs' would never find it worthwhile to start searching for a producer in the first place. Hence, the mode of exporting through a GI can only exist if producers can credibly commit to *exclusive dealership arrangements*, that grant the GI the exclusive right to sell the producers specific variety in the foreign market.

Under Pareto,

7.5.2 Free entry of producers

Free entry of producers ensures equality between the present value of average profit flows of a potential entrant and the entry costs f^E. Recall that the value of a searching producer consists of two components: a first that collects profits from exporting when being matched to a GI, and a second that comprises search costs, occurring regardlessly of the characteristics A.Then, the fee entry condition can be expressed as

$$\delta^P f^E = \int_0^{A^D} \pi^D(A)\, dG(A) + \int_0^{A^F} \pi^F(A)\, dG(A) \tag{7.18}$$

$$+s(\theta) \int_{A^F}^{A^{SP}} \pi^{MP}(A)\, dG(A) - (1 - s(\theta)) \left(G\left(A^{SP}\right) - G\left(A^F\right)\right) c^P, \tag{7.19}$$

where the first and second integral of the above expression reflect, respectively, the expected profits of domestic operations and from exporting through an own subsidiary, and the remaining expressions capture the value of a searching producers.

7.5.3 Definition of a steady state

The system of equilibrium conditions (7.9), (7.13), (7.14), (7.17) and (7.19) implies the equilibrium cutoffs A^D, A^{SP}, A^F, the equilibrium market tightness θ, and the equilibrium real income level B. Assume that all components of A are random realizations from independent distribution functions following the Pareto law. Then, A is also Pareto distributed. More precisely, we let the c.d.f. $G(A) = A^k$, with a shape parameter k and the support $(0, 1]$.[16] Under our Pareto assumption, the expected surplus is independent of B, which immediately leads to recursivity. The next lemma characterizes the solution of the general equilibrium.

Lemma 10. *If A follows the Pareto distribution with shape parameter $k > \sigma - 1$, the zero cutoff profit conditions plus the free entry condition of GIs, solve for the equilibrium cutoff points A^D, A^{SP} and A^F as well as for the market tightness θ independently from $\bar{\tau}$ and B. The value of B then adjusts such that the free entry condition of producers is met.*

Proof. In the Appendix. □

Given Lemma 10, in order to prove existence of the equilibrium, it is sufficient to substitute the zero cutoff profit conditions (7.9), (7.13), and (7.14) into the GIs' free entry condition and search for the values of θ that solves that equation. Since expected search search costs are increasing in θ, for uniqueness it is sufficient to show that the expected surplus is increasing in θ. While our simulations suggest uniqueness of the equilibrium, it is hard to prove it formally, since the expected surplus is a fairly complicated function of the market tightness.

7.5.4 Discussion

The model is close to the frontier of analytical tractability. Hence, theoretical extensions require to restrict the analysis to certain channels, thereby reducing complexity in some elements and enriching the setting in some other areas. This has be done in Felbermayr and Jung (2008c).

[16]The Pareto assumption has been made in a large number of related papers (e.g. Helpman, Melitz, and Yeaple (2004), Helpman, Melitz, and Rubinstein (2008), Bernard, Redding, and Schott (2006).

7.6 Conclusions

This paper provides a general equilibrium framework with heterogeneous firms, in which trade in goods may occur in an indirect mode, via specialized general importers, or directly, via producers' sales affiliates in foreign countries. We therefore offer a theoretical explanation for a key stylized fact, namely, the existence of trade intermediation. This fact has not been explored systematically in the recent trade literature.[17]

In our extension of the Melitz (2003) model, producers have the option to search for foreign general importers and use them as trade intermediaries or access the foreign market through an own sales affiliate. Relative to the second option, the first option saves fixed costs but requires sharing profits with the intermediary. Importantly, our model partly endogenizes trade costs, since expected the expected costs of searching for a general importer are endogenous in the model and determined by the entry decisions of both producers and importers. Hence, our framework contributes towards a better understanding of trade costs that are not covered by tariffs or transportation costs and that may differ systematically across countries.

Compared to the received literature, we broaden the notion of firm heterogeneity and allow firms to differ with respect to the degree of tradability of their goods, the strength of their brand names, and their marginal costs of production. Our key result shows that exporting via a general importer is an attractive way to access foreign markets when firm characteristics lie in an intermediate range.

Another central result is that the effect of institutional change, such as improving the access to trade fairs, on the volume of trade can theoretically be negative, since some firms that have been exporting through a sales affiliate may find it optimal to use the GI instead, thereby giving up variable revenue, but saving fixed market entry costs. Moreover, our model can relate cross-country heterogeneity in export behavior to the drivers of expected fixed costs. Finally, we find that the emergence of GIs may lower the aggregate productivity level.

The present paper is close to the frontier of analytical tractability. Hence, theoretical

[17] There are, of course, some notable exceptions, e.g., Schröder et al. (2005).

extensions require to reduce complexity in some elements, and enrich the model in some other areas. We believe that there are two main avenues of developing the model further. First, general importers usually are multi-product firms. This is true for producers, too, but the incentives to develop product portfolios is stronger for GIs. Eckel and Neary (2006) and Feenstra and Ma (2007) offer promising frameworks to tackle this extension. Second, we have not modeled the rich incentive problems that arise when a general importer has to exert effort to sell a producer's goods to a foreign market. A formalization of that issue is promising since the fruits of investment in marketing and sales promotion would be shared with the producer. Third, and related to the second potential extension, in the present paper, we have restricted our analysis to the case where contracts are not enforceable altogether. A natural extension lies in a more flexible approach, where the degree of contractability is variable. In reality there is a rich panoply of different arrangements between producers and foreign retailers, ranging from licensing to franchising agreements. All this alternative forms of interaction involve some way of solving the double marginalization problem inherent in our analysis. We believe that bringing the rich industrial organization literature into a model of our type could further cast light on the structure of trade costs between two countries.

Regarding empirical analysis, the present paper would motivate a formal econometric study that analyzes the choice of export modes in the presence of heterogeneous firms. As firm level data becomes more widely available for a larger array of countries and a richer set of variables, empirical analysis of our mechanism should become viable in the close future.

7.7 Appendix. Guide to calculations

Proof. Lemma 7.
The problem of the producer is

$$\max_{x(\omega)} \beta J(\omega) - a(\omega)\tilde{\tau}(\omega) x\left[p^G(\omega)\right] \tag{7.20}$$

subject to $x(\omega) = H\frac{\zeta(\omega)^{\sigma-1}}{p(\omega)^{\sigma}}$, where $J(\omega) = p^G(\omega) x\left[p^G(\omega)\right] - f^M$. The first order condition

$$\frac{\sigma-1}{\sigma}\beta H^{\frac{1}{\sigma}}\zeta(\omega)^{\frac{\sigma-1}{\sigma}} x(\omega)^{-\frac{1}{\sigma}} = a(\omega)\tilde{\tau}(\omega) \tag{7.21}$$

implies $p^G(\omega) = a(\omega)\tilde{\tau}(\omega)/(\rho\beta)$..

□

Proof. Lemma 8.
We need to establish the parameter restriction that ensures that for given θ ensures a interior solution to the equilibrium sorting problem. We can write the flow profits associated to each mode of operation, *mode* $\in \{D, SP, F\}$ as the following set of equations:

$$\begin{aligned} \delta^P V^{SP}(A) &= s(\theta)\beta^\sigma B(\bar{\tau}A)^{1-\sigma} - \left\{s(\theta)\beta f^M + [1-s(\theta)]c^P\right\} & (7.22)\\ \pi^F(A) &= B(\bar{\tau}A)^{1-\sigma} - f^F & (7.23)\\ \pi^D(A) &= BA^{1-\sigma} - f^D, & (7.24) \end{aligned}$$

We establish a lower and an upper bound, $\underline{f}$ and $\bar{f}$, respectively, to the expected fixed costs of the search mode SP. First, to pin down $\underline{f}$, we search for the intercept of $\delta^P V^{SP}(A)$ that solves $\delta^P V^{SP}(A_D) = 0$. That condition yields $s(\theta)\beta^\sigma B(\bar{\tau}A)^{1-\sigma} - \underline{f} = B\left(A^D\right)^{1-\sigma} - f^D$. Recognizing from (7.9) that $\left(A^D\right)^{1-\sigma} = f_D/B_i$, we find the lower bound

$$\underline{f} = s(\theta)\beta^\sigma \bar{\tau}^{1-\sigma} f^D.$$

The upper bound is found by finding the intercept $\bar{f}$ for which $\delta^P V^{SP}\left(\tilde{A}\right) = 0$ with $\tilde{A}$ determined by the condition $\pi^F\left(\tilde{A}\right) = 0$. We have $s(\theta)\beta^\sigma B\left(\bar{\tau}\tilde{A}\right)^{1-\sigma} - \bar{f} = 0$. Recognizing from (7.5) that $\tilde{A} = \bar{\tau}^{\sigma-1} f^F/B^j$, we find the upper bound

$$\bar{f} = s(\theta)\beta^\sigma f^F.$$

Collecting results, the condition on the intercept of 7.22 $-\underline{f} < s(\theta)\beta f^M + [1-s(\theta)]c^P < \bar{f}$ can be written as

$$\bar{\tau}^{1-\sigma} f_D < \beta^{-\sigma}\left[\beta f^M + \frac{\delta c^P}{\eta(\theta)}\right] < f^F, \tag{7.25}$$

where we have made use of the definition $s(\theta) \equiv \eta(\theta)/[\delta + \eta(\theta)]$. Condition (7.25) is the one that appears in Lemma 8. □

Proof. Corollary 8.
Consider how an increase in θ affects the $\delta^P V^{SP}(A)$ locus (7.22): first, the locus becomes

steeper since $s'(\theta) > 0$; second, the locus shifts up (down) if $\beta f^M < (>) c^P$. Focusing on the case where $f^M = 0$, the locus always shifts up.
Using 'hats' to denote proportional changes, the cutoff levels A_{ij}^{SP} and A_{ij}^F change as follows:

$$\hat{A}_{ij}^{SP} = \frac{\gamma}{\sigma - 1}\hat{\theta}_{ij}, \tag{7.26}$$

where γ is the elasticity of the matching function with respect to the number of searching GIs. Similarly, we have

$$\hat{A}_{ij}^{F} = -\frac{\gamma}{\sigma - 1}\frac{\delta}{\delta + \eta(\theta)}\beta^{\sigma}\hat{\theta}_{ij} < -\hat{A}_{ij}^{SP}, \tag{7.27}$$

where the inequality follows from the fact that both $\delta / [\delta + \eta(\theta)]$ and β^{σ} are strictly smaller than unity. □

Proof. Lemma 10.
Consider again the GI's share of the expected surplus. Using (7.7) and the Pareto assumption, we find an expression for the expected surplus

$$E[J(A)] = \frac{k\sigma B\left(\bar{\tau}\beta^{-1}\right)^{1-\sigma}}{k - (\sigma - 1)}\frac{\left(A^{SP}\right)^{k-(\sigma-1)} - \left(A^{F}\right)^{k-(\sigma-1)}}{\left(A^{SP}\right)^{k} - \left(A^{F}\right)^{k}}. \tag{7.28}$$

The independence of expected surplus of the demand level B and the homogeneous part of the trade costs $\bar{\tau}$ directly follows from inserting the cutoff profit conditions (7.9), (7.13), and (7.14) into (7.28). The in dependence of θ of B and $\bar{\tau}$ immediately follows from the free entry condition (7.17). □

Chapter 8

On the Importance of Adjustment Dynamics For Bilateral Trade Flows[1]

8.1 Introduction

In the last fivty years, politicians have undertaken a huge effort to liberalize trade. For assessing the outcome of trade reforms it is important to know how quickly bilateral trade flows adjust. If adjustment is fast, potential gains from trade reforms are achieved quickly. In the opposite case, however, it takes a long time to see the full beneficial impact.

Static gravity models implicitly assume trade volumes to be at their steady-state levels. There is empirical evidence, however, for a dynamic relationship. Eichengreen and Irwin (1998) find cross-sectional evidence. De Grauwe and Skudelny (2000) and Egger (2001a) present first dynamic panel data evidence.

Recent years have witnessed an increasing number of dynamic gravity applications, e.g., Micco et al. (2003) and De Nardis et al. (2008) on the dynamic effect of the Euro on bilateral trade, or Moser et al. (2008) and Martinez-Zarzoso et al. (2009) on the effect of German export promotion. These studies, however, do not explicitly consider the speed of adjustment.[2] Moreover, test statistics reveal misspecification problems, e.g. a significant Hansen test in De Nardis et al. (2008) signals invalidity of the instruments; see Roodman (2009).

Various theoretical explanations have been put forward for a dynamic trade relationship.

[1]This Chapter is based on a working paper, see Jung (2009).
[2]An important exception is Egger (2001b).

Baldwin (1988) and Dixit (1989) argue that sunk costs of market entry and exit lead to hysteresis in trade. Moreover, consumers and final good producers tend to prefer products and inputs they are already familiar with. Persistence in habits also result in highly path-dependent trade relationships.

Recent empirical work emphasizes the role of trust for economic activity. Using survey data on bilateral trust between European countries, Guiso et al. (2009) find that a one standard deviation increase in importer's trust toward the exporter boosts exports by 10%. The idea goes back to Arrow (1972), who argues that trust has first order economic effects. More generally, De Benedictis and Vicarelli (2005) point out that "*trade relationships between countries are affected [...] by the accumulation of invisible assets such as political, cultural and geographical factors*" (p. 9). The 'invisible asset', however, is not exogenously given, but created by repeated interactions between trading partners.[3]

In this paper, we offer the following contributions. First, drawing on the neoclassical growth model, we incorporate the endogenous accumulation of an 'invisible asset' into a standard Anderson and van Wincoop (2003) model of international trade. We do this to motivate a dynamic specification of the gravity equation. The notion of 'invisible asset' accommodates various interpretations. It can be *trust* that is created on the basis of what trading partners know about each other. Similarly, *relational* capital is nurtured by good trade relationships, and enhances future trade. One can also think of *knowledge* about the foreign market, which becomes, at least partly, known to all other firms. In either case, there arises a pure externality which positively affects trade through lower trade costs.

Second, we apply a *dynamic* gravity approach which differs from previous studies along the following lines. We account for *time-varying* multilateral resistance terms by including country-and-time effects into our regressions.[4] Moreover, as common in the growth literature, we take data for every five years for the period 1980-2000 in order to tackle business cycle fluctuations. Our sample comprises information on potentially 96×95 country pairs rather than OECD countries only.

The estimated coefficient of the lagged endogenous variable lies in the expected range;

[3] *The Economist* concludes that "reciprocal ties become strongest between people who meet and trade frequently". (October 18th-24th 2008, p. 84).

[4] See Anderson and van Wincoop (2003) for the theoretical explanation.

see Bond (2002). This points at consistent estimation. Moreover, the standard test statistics jointly signal validity of our estimation strategy. We use the results to back out how fast trade volumes reach their steady-state volumes. It turns out that trade flows adjust faster than per capita income. On average, trade flows take 3.5 years to close half the way from their initial positions to the steady state, while the growth literature finds per capita income to take twice as long. The adjustment rate is lower, however, than previous dynamic gravity estimates would imply. This is mainly due to the theoretically motivated inclusion of country-and-time effects to account for multilateral resistance terms.

Third, we use the information on the adjustment rate to evaluate the short- and long-run effect of trade reforms. We address the formation of free trade agreements (FTA) because FTAs aim at reducing "*policy controlled barriers to the flow of goods, services, capital, labor,etc.*" (Baier et al., 2008, p. 461) between two or more countries. Moreover, the last twenty years have witnessed an enormous increase in FTAs, which gives us sufficient variation in the data.[5] We also consider the pro-trade effect of the WTO, and revisit the role of the introduction of the Euro in a dynamic setting.

Fourth, we compare our results to those obtained by Baier and Bergstrand (2007). They introduce a dynamic element in their static regressions by including lagged FTA dummies to account for 'phasing-in' of FTAs. Our regressions show that there is no additional role for 'phasing-in' once the gravity equation is cast as a dynamic model.

Finally, we use our framework to highlight different patterns in the adjustment rate. We find that geographically proximate country pairs need longer to close the gap to the steady state, whereas culturally proximate countries adjust faster. Moreover, we detect interesting variation across trade in commodity groups. The adjustment rate is higher for exchange-traded goods. On impact, FTA formation creates more trade in exchange-traded than in differentiated goods. These two observations jointly imply that common FTA membership leads to more trade creation in exchange-traded goods in the long-run.

The paper is related to the literature as follows. There exist a number of empirical studies analyzing the impact of trust on trade. In a sample of European countries, Guiso et al. (2009) find that lower bilateral trust leads to less trade between two countries. The

[5] See Baier and Bergstrand (2007) for a complete list of FTAs.

setting of our analysis differs, since it does not draw the measure of trust from survey data, but endogenously models trust accumulation as a by-product of trade. This also allows for a larger country coverage in the empirical part. We find trust to be more important for differentiated goods than for exchange-traded goods also in a dynamic setting.

The econometric approach is related to the literature on empirical growth models, e.g., Bond, Hoeffler, and Temple (2001). Recent dynamic gravity applications usually do not control consistently for multilateral resistance terms, suffer from unresolved specification problems, or focus on a single anchor country. In De Nardis et al. (2008), e.g., the Hansen tests signal poor instruments. Moser et al. (2008) and Martinez-Zarzoso et al. (2009) analyze the role of *German* exports promotion in a dynamic setting.

The remainder of the paper is organized as follows. Section 2 introduces the accumulation of an 'invisible asset' into a multi-country monopolistic competition model of international trade to motivate a dynamic specification of the gravity equation. Section 3 discusses the estimation approach and the data. Results from our dynamic panel data regressions are presented in Section 4.[6] The final section concludes.

8.2 A simple dynamic gravity model

In this section, we derive a dynamic version of the gravity model. The dynamics may enter through endogeneity in trade costs or preferences or both. To be more precise, consider, e.g., the lack of trust to be a barrier to trade. Trust, a prediction of reliance on an action, is created on the basis on what trading partners know about each other. Thus, trade nurtures trust. Exchange of goods also generates knowledge about other trading opportunities, which reduces trade costs, and therefore leads to more trade. Moreover, by having good relationships with buyers or suppliers, relational capital is accumulated and maintained. Hence, existing trade relationships reduce trade costs, which, in turn, foster trade.

Trade flows may also endogenously determine preferences. Consumers become accustomed to products they spend income on. Hence, in the future, they prefer goods they are already familiar with, and habits turn out to be highly persistent. The same logic applies to final output good producers, who find it more convenient to source inputs from suppliers

[6] Our data, program codes, and further results can be downloaded from <link to be added>.

they are already in a relationship with.

In our model, trade becomes endogenous through the introduction of an 'invisible asset', which accumulates as trade takes place. This formulation generalizes the concepts outlined above, and accommodates both, endogenous trade costs, and endogenous habit formation.

8.2.1 The theory-based static gravity model

We assume the existence of a representative household with CES preferences over domestic and imported varieties of some differentiated good. Different to the standard treatment, we use the utility function proposed in Combes et al. (2005) which introduces source-country specific weights a_{ijt}. These weights capture the particular attachment of country $i's$ household to imports from country j at time t. We may use this slightly modified utility function in the multi-country monopolistic competition model of international trade proposed by Anderson and van Wincoop (2003). Utility maximization under the appropriate aggregate budget constraint, market clearing, and the assumption that iceberg trade costs T_{ijt} and preference weights a_{ijt} are symmetric $(T_{ijt} = T_{jit}; a_{ijt} = a_{jit})$, the (c.i.f.) value of bilateral imports M_{ijt} can be written as

$$M_{ijt} = \frac{Y_{it}Y_{jt}}{Y_{wt}} \left(\frac{T_{ijt}}{a_{ijt}}\right)^{1-\sigma} \left(\tilde{P}_{it}\tilde{P}_{jt}\right)^{\sigma-1}, \tag{8.1}$$

where the price indices $\tilde{P}_t$ solve $\left(\tilde{P}_{jt}\right)^{1-\sigma} = \sum_{i=1}^{C} (Y_{it}/Y_{wt}) (T_{ijt}/a_{ijt})^{1-\sigma} \left(\tilde{P}_{it}\right)^{\sigma-1}$; see Feenstra (2004) for the details of the derivation. Anderson and van Wincoop (2003) call $\tilde{P}_{it}$ indices of multilateral resistance because they depend on the trade costs and preference weights of country i with all countries in the world, the number of which is given by C. The variable Y_{it} denotes GDP of country i, the subindex w refers to the world. The elasticity of substitution in the underlying CES utility function is given by σ.

The central insight of Anderson and van Wincoop (2003) is that the volume of trade between i and j depends not only on the trade costs and preference weights between i and j but on the entire distribution of trade costs between i and j and *all other countries* of the world. How strongly T_{ijt}/a_{ijt} restricts trade between i and j depends on the costs that affect trade with alternative partners. Hence, in the estimation we have to deal with the $\tilde{P}_{it}$ terms. The multilateral resistance terms $\tilde{P}_{it}$ are essentially unobserved since they do not correspond to official CPI deflators. Anderson and van Wincoop (2003) show how one can solve for the $\tilde{P}_{it}$ terms numerically and use them in an iterative estimation strategy. They demonstrate that the failure to control for multilateral resistance typically biases the absolute value of estimated trade cost variables upwards.

In our regressions, we follow Baier and Bergstrand (2007) and use country dummies which fully control for all purely country-specific variables such as the $\tilde{P}_{it}$ terms.

8.2.2 Endogenous trade costs and preference weights

We will now augment the theoretical gravity equation (8.1) in order to make transparent how trade history shapes present bilateral trade. We introduce a state variable S_{ijt}. It measures the stock of an invisible asset, and affects both, trade costs and preference weights. The higher the stock of the invisible asset between two countries, the lower the trade costs, and the higher the preference weights. Since trade cost T_{ijt} and preference weights a_{ijt} enter equation (8.1) similarly, we set $a_{ijt} = 1$. Thus, in the following we focus on the role of endogenous trade costs. Our approach, however, also extends to habit formation.

To be more precise, we assume trade costs to be decreasing and concave in the stock of the invisible asset. If the the stock of the invisible asset goes towards 0, trade costs become prohibitively high. On the other hand, trade costs cannot fall short of unity.

The invisible asset is nurtured in proportion to the bilateral trade volume. However, it also diminishes as time goes by. Hence, the net increase in the invisible asset at any point in time equals asset accumulation less depreciation. Then, the asset grows at rate

$$\dot{S}_{ijt}/S_{ijt} = \gamma M_{ijt}\left(S_{ijt}\right)/S_{ijt} - \delta, \tag{8.2}$$

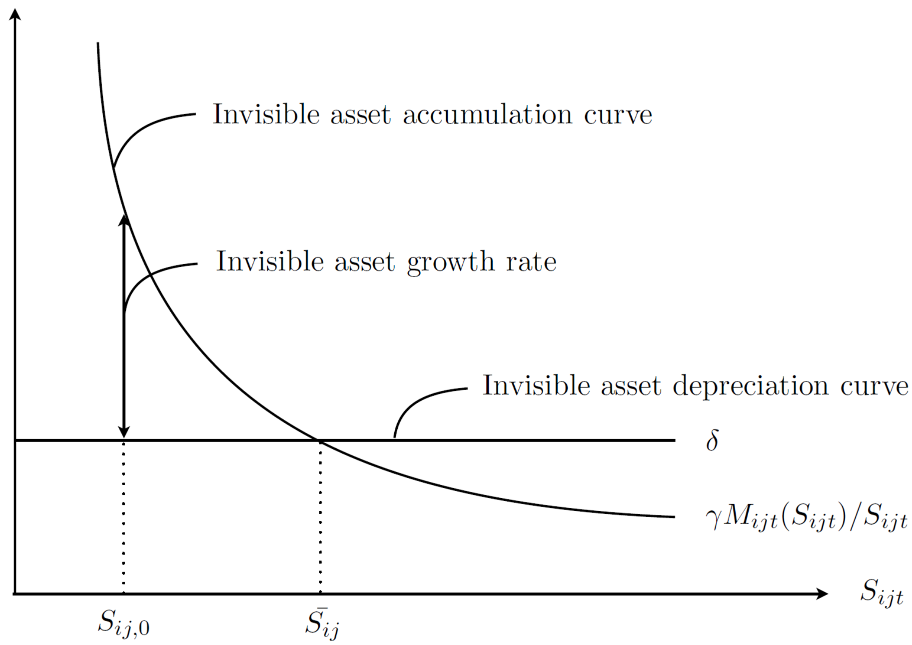
where the dot denotes differentiation with respect to time, the exogenous parameter $0 < \gamma \leq 1$ measures to which extent bilateral trade feeds back to asset accumulation, and $0 < \delta \leq 1$ denotes the depreciation rate.[7]

Figure 8.1 visualizes asset accumulation and depreciation graphically. The invisible asset accumulation curve, $\gamma M_{ijt}(S_{ijt})/S_{ijt}$, is downward sloping. It asymptotes to infinity

[7] Note that equation (8.2) is similar to the fundamental equation in the neoclassical growth model, where capital is accumulated and depreciated; see, e.g., Barro and Sala-i-Martin (2003). For convenience, we express the model in continuous time, whereas in the empirical we use data in discrete time.

at $S_{ijt} = 0$, and approaches 0 as S_{ijt} tends to infinity.[8] The invisible asset depreciation curve is a horizontal line at δ. The vertical distance between the two curves equals the growth rate of the invisible asset; see equation (8.2). The invisible asset reaches its steady-state level in a situation where asset accumulation by economic transactions exactly offsets asset depreciation. It is found at the intersection point of the two curves.

Figure 8.1: Invisible asset accumulation and depreciation



Since asset accumulation is assumed to be proportional to the volume of trade, we proxy its level by the lagged volume of trade, which yields a dynamic formulation of the gravity equation. Besides the stock of the invisible asset, trade costs also depend on other factors. Time-varying bilateral trade policy is proxied by a dummy FTA_{ijt} that takes one if countries i and j are in a free trade agreement at time t. Indicators of geographical and cultural proximity are collected in a vector $\mathbf{PROX}_{ij}$.

Taking logs of equation (8.1), and accounting for the endogeneity of trade costs, we obtain the empirical dynamic gravity equation

$$\ln M_{ijt} = \alpha \ln M_{ij,t-1} + \boldsymbol{\xi}\mathbf{POL_{ijt}} + \boldsymbol{\zeta}\mathbf{PROX}_{ij} + v_{it} + v_{jt} + \varepsilon_{ijt}, \tag{8.3}$$

where the country-and-time effects v_{it}, v_{jt} capture all time-varying country-characteristics

[8]Technically spoken, the trade cost function fulfills the inada conditions.

like GDP (Y_t) and the multilateral resistance terms ($\tilde{P}_t$). ξ equals $\bar{\xi}(\sigma - 1)$, where $\bar{\xi} > 0$ is the vector of semi-elasticities of trade costs with respect to FTA or WTO membership, and the common use of the Euro. Similarly, the vector $\boldsymbol{\zeta}$ collects (semi-)elasticities of different proximity measures. The coefficient α measures the degree of trade flow persistence. α smaller than one signals that the invisible asset, and hence trade, indeed adjust to their steady-state values. A change in trade policy affects trade flows directly through ξ, and indirectly through adjustment of the invisible asset, which is captured by α.[9] The error term ε_{ijt} can be decomposed into a country-pair specific component v_{ij}, and an idiosyncratic part u_{ijt}.

8.3 Dynamic panel data estimation and data

This section tackles issues that arise in the context of dynamic panel data estimation. It also discusses data availability.

8.3.1 How to estimate a dynamic gravity equation

The modern empirical growth literature argues that there are some problems with dynamic estimations. It is well known, that estimating dynamic gravity equation (8.3) by ordinary least squares (OLS) or least square dummy variables (LSDV) leads to dynamic panel bias (Nickell, 1981) due to the endogeneity of lagged trade. OLS and LSDV, however, provide useful guidance in the empirical analysis. Bond (2002) points out that $\hat{\alpha}^{OLS}$ and $\hat{\alpha}^{LDSDV}$ are respectively biased upwards and downwards. Hence, the true α^{true} must fall into this bracket.

In order to deal with endogeneity, several strategies have been proposed to instrument the lagged dependent variable:

Two-stage least squares (2SLS). Anderson and Hsiao (1981, 1982) propose a two-stage least square estimator. In a first step, the country-pair specific fixed effect v_{ij} is eliminated by a first-differentiating transformation. Unfortunately, Δu_{ijt} is correlated with $\ln M_{jit-1}$. Consistent estimates of α, however, can be achieved if instrumental variables are

[9] Note that α and ξ are only *average* effects. Moreover, when computing FTA effects, we omit FTA effects through endogenous changes in multilateral resistance terms; see footnote 11 in Baier and Bergstrand (2007) for this argument in a static setting.

available that are both correlated with $\Delta \ln M_{ijt-1}$ and orthogonal to Δu_{ijt}. Besides the second lag of the difference, $\Delta \ln M_{ijt-2}$, the second lag of the level $\ln M_{ijt-2}$ turn out to fulfill these requirements. They are both correlated with $\Delta \ln M_{ijt-1}$ but uncorrelated with Δu_{ijt}, as long as the u_{ijt} themselves are not serially correlated. Estimators using either (or both) of these instruments obtain consistent results as long as the number of country-pairs is sufficiently large, whereas consistency does not hinge on the number of periods T. Instrumenting with the second lag of the level requires one period less than instrumenting with the second lagged difference, which can be a huge advantage in short panels. Hence, we opt for this strategy. The 2SLS estimator, however, is inefficient, which is usually reflected in large standard errors.

Difference GMM (Diff-GMM). In order to increase efficiency, Arellano and Bond (1991) consider further moment restrictions. They assume that the error terms have the standard structure, i.e., $E[v_{ij}] = 0, E[u_{ijt}] = 0, E[v_{ij}u_{ijt}]$. Moreover, errors have to be serially uncorrelated, $E[u_{ijt}u_{ijs}] = 0, s \neq t$, and initial conditions have to predetermined, $E\left[\ln M_{ij,0}\eta_{ijt}\right] = 0$. Then, the standard linear first-differenced GMM estimator exploits $m = 0.5(T-1)(T-2)$ moment restrictions. Lagged levels dated $t-2$ and earlier are used as instruments for the equations in first differences.

The difference GMM estimator yields consistent estimates if $N \to \infty$ and T is fixed. It suffers, however, from substantial finite sample bias that is in particular increasing in the number of instruments used. If the estimate for α is close to one or below the LSDV estimate, the Diff-GMM estimate is likely to be downward biased due to weak instruments.

System GMM (Sys-GMM). The system GMM estimator proposed by Blundell and Bond (1998) relies on the additional assumption that the means of the ln M_{ijt} series are constant through time for periods $1..T$, while varying across individuals. This can be expressed as $E[v_{ij}\Delta \ln M_{ij,2}] = 0$, which leads to $T-2$ further moment conditions $E[u_{ijt}\Delta \ln M_{ij,t-1}] = 0$. The additional moment conditions of lagged first-differences can be used as instruments for the level equation. This leads to a reduction in finite sample bias and gains in precision in cases where the autoregressive parameter α is only weakly identified in levels.

In our GMM estimations, we use the second and, where available, the third lag of the dependent variable in levels as instruments for the difference equation. Similarly, we use the second and, where available, the third lag of the dependent variable in differences as instruments for the level equation. In order to limit the number of instruments, we "collapse" the set of instruments into a single column.[10]

We treat the trade policy measure, common membership in a free trade agreement (FTA), as exogenous.[11]

It is crucial for the validity of the GMM approaches, that the instruments are exogenous. The validity of the instruments can be tested using the Sargan-Hansen test of overidentifying restrictions. The test is common to all panel data estimations which use instrumental variables, and where the number of instruments exceeds the number of regressors. To be more precise, the null hypothesis of the test is that the excluded instruments are valid, i.e., instruments are uncorrelated with the error term and correctly excluded from the estimated equation. The test statistic is distributed chi-squared with $L-K$ degrees of freedom, where L is the number of excluded instruments and K is the number of regressors. A rejection casts doubt on the validity of the instruments. We present two versions of the test. The *Sargan* statistic is the minimized value of the one-step GMM criterion function. Unfortunately, it is not robust to heteroskedasticity or autocorellation. Hence, for one-step, robust estimation also the *Hansen* statistic is of interest, which is the minimized value of the two-step GMM criterion function. The Hansen test is robust; however, it can be greatly weakened by instrument proliferation.

Moreover, we test for autocorrelation in the error terms, see Arellano and Bond (1991). This is important because autocorrelation in the error term renders some of the instruments invalid. Autocorrelation of order one, AR(1), is expected in first differences. The reason is that $\Delta u_{ijt} = u_{ijt} - u_{ij,t-1}$ should correlate with $\Delta u_{ij,t-1} = u_{ij,t-1} - u_{ij,t-2}$, since they share the term $u_{ij,t-1}$. In order to check for AR(1) in levels, we look for AR(2) in differences, which signals a relationship between the $u_{ij,t-1}$ in Δu_{ijt} and $u_{ij,t-2}$ in $\Delta u_{ij,t-2}$. To summarize ideas, we expect serial autocorrelation of order one, but have to rule out serial

[10]Recent dynamic panel data applications adapt this strategy, e.g., Caldern et al. (2002), Beck and Levine (2004), or Carkovic and Levine (2005). See Roodman (2009) on the superiority of collapsed instruments.

[11]Using a Wooldridge test, Baier and Bergstrand (2007) cannot reject exogeneity of the FTA dummy in the presence of country-pair fixed effects. We tackle time-invariant country-pair characteristics by a first-differentiating transformation.

autocorrelation of order two.

8.3.2 Data

We proxy bilateral trade policy by common membership in a regional free trade agreements. Baier and Bergstrand (2007) provide information on FTA membership for a sample of 96×95 country-pairs covering the years 1960, 1965, ..., 2000.[12] Data on geographical and cultural proximity like distance, contiguity, colonial ties, and use of a common language come from the CEPII database. GDP data are not required, since GDPs as well as multilateral resistance terms are covered by country-and-time effects.

Aggregate bilateral trade flows stem from the Direction of Trade Statistics (DoTS). Zero (and missing) trade flows are excluded.[13] This implies, that we only consider trade relationships, where the initial level of trust suffices to observe positive trade flows.

We aim at constructing a balanced sample which covers as many country pairs as possible. A balanced sample is preferable in the context of dynamic panel data estimation.[14] Therefore, we restrict our analysis to a balanced sample covering the years 1980-2000, which leaves us with $N = 4301$ country pairs.[15] As common in the empirical growth literature, we use five-yearly data. We do this to tackle business cycle fluctuations, which may also be reflected in trade flows. Hence, we end up with $T = 5$ periods.

We will also analyze trade by commodity group which requires data at the level of four-digit Standard Industrial Trade Classification (SITC), Revision 2. We draw on the NBER-UN World Trade Data as distributed by Feenstra et al. (2005), which in principle cover the years 1962-2000. Data for early years (1962-1983), however, are taken from UN Comtrade, making adjustments for country codes only, whereas for the latter years (1984-2000), data only cover 72 countries, but have undergone several adjustments. Most importantly, raw data have been recompiled to allocate exports to unspecified regions, and to correct for entrepôt trade. In order to avoid this structural break in the trade data to affect the dynamic panel data estimations, we restrict our trade-in-commodity sample to

[12] The bilateral trade policy dataset is kindly made available at http://web.mac.com/baier_family/iWeb/Site%202/Data.html. Table 8.7 in the Appendix lists the countries included.

[13] For the sake of data availability, we consider fob trade flows.

[14] Diff-GMM and Sys-GMM set missing entries in the instrument matrix to zero.

[15] The total number of country pairs decreases substantially if longer periods are covered.

the years 1985-2000 ($T = 4$).

In order to aggregate trade data on four-digit SITC level to trade by commodity group, we make use of both the 'liberal' and 'conservative' Rauch (1999) classifications. The conservative classification minimizes the number of three- and four-digit commodities that are classified as either exchange traded or reference priced, whereas the liberal classification maximizes them. Conservatively aggregated exchange-traded goods and liberally aggregated differentiated goods contain no SITCs that switch classifications when moving from the conservative to the liberal classification. For brevity, we only show results for 'liberal' aggregation, and put the corresponding results for 'conservative' aggregation into the Appendix.

8.4 Results from a dynamic gravity approach

In this section, we empirically address the following issues. First, we want to answer the question how quickly bilateral trade flows adjust to their steady-state levels. We compare our results to previous findings. Second, we use this information to back out the long-run effects of FTA formation, WTO membership, and common use of the Euro. Third, we compare our results to estimates from static regressions, where a lagged FTA dummy is introduced. We do this, because we want to examine whether the lagged FTA effects reported in Baier and Bergstrand (2007) stem from exogeneneously undertaken second rounds in trade cost reductions, or are endogenously caused by the induced trust accumulation. Fourth, we analyze the role of geographical and cultural proximity for adjustment patterns. Finally, we repeat the exercise for trade in different commodity groups.

8.4.1 Adjustment dynamics of bilateral trade flows

In a first step, we purely focus on the validity of the estimation approach and on the adjustment dynamics. We therefore estimate the dynamic gravity equation (8.3) using different estimators and restricting the vectors $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ to zero.[16]

Table 8.1 shows the results. Column (1) presents the estimates obtained from a 2SLS approach, where the second lag of the trade value in levels is used as an instrument for

[16] We will show below that the adjustment rate is qualitatively and quantitatively not affected by the inclusion of additional trade policy controls.

Table 8.1: Adjustment dynamics of bilateral trade flows

Dependent variable: Bilateral trade					
	(1)	(2)	(3)	(M1)	(M2)
	2SLS	Diff-GMM	Sys-GMM	OLS	LSDV
Lagged trade	0.356***	0.357***	0.375***	0.670***	0.063***
	(12.50)	(13.69)	(13.18)	(78.91)	(3.91)
Adjustment rate	0.207	0.206	0.196		
Half-life time	3.349	3.365	3.536		
AR(1) test (p-value)	0.000	0.000	0.000		
AR(2) test (p-value)	0.382	0.389	0.365		
Sargan test (p-value)		0.780	0.024		
Hansen test (p-value)		0.861	0.605		

Balanced sample of $N = 3921$ dyads, and $T = 5$ periods. Robust t-values in parentheses. *** indicates significance at 1%. All regressions include a comprehensive set of country-and-time effects. 2SLS estimator uses 2^{nd} lag in levels to instrument lagged endogenous variable in differences. One-step GMM results. 2^{nd} and 3^{rd} lag in levels used as instrument for first differences equation. 2^{nd} and 3^{rd} lag in differences used as instruments for level equation. Set of instruments collapsed into a single column. OLS includes controls for distance, contiguity, common language, and colonial ties (all not shown). Adjustment rate quantifies how much of distance to steady state is closed each year. Half-time life time reports time needed to close half the gap to steady state. See footnotes 21 and 22 for computational details.

the lagged endogenous variable; see Anderson and Hsiao (1981, 1982). The F-statistic of the first stage regression is about 20, which is a sufficient fit; see Staiger and Stock (1997). Columns (2) and (3) respectively show Diff-GMM and Sys-GMM estimation results.

Before turning to the estimates, we show that the validity of our dynamic estimation approaches cannot be rejected. First, Sargan and Hansen tests of overidentifying restrictions do not cast doubt on the specification.[17] Arellano-Bond tests for AR(1) and AR(2) in the (differenced) residuals also do not show invalidity of instruments. As an additional check, Columns (M1) and (M2) in Table 8.1 provide the results of estimating equation (8.3) by OLS and LSDV, respectively. The true effect of the lagged endogenous variable lies in the range of the LSDV and OLS estimates; see Bond (2002). It turns out that 2SLS, Diff-GMM and Sys-GMM estimates indeed fall in the respective range, which proves consistency of our estimates.

In all three dynamic models, the lagged endogenous variable turns out to be highly significant. The point estimate is remarkable stable across specifications. It amounts to

[17]The Sargan test for the Sys-GMM specification should be statistically insignificant, too. It will turn out below that this is indeed the case once we include trade policy controls.

0.36 with a t-value of around 13. Hence, trade flows indeed adjust to their steady-state levels.

How rapidly do bilateral trade volumes approach their steady-state values? The coefficient on the lagged endogenous variable allows to quantify the speed of adjustment. To get an idea of what determines the adjustment rate, let the elasticity of trade costs with respect to the level of trust be denoted by $-\kappa$.[18] Then, the growth rate of trade can be written as

$$\frac{\dot{M}_{ijt}}{M_{ijt}} = \kappa(\sigma - 1)\delta \left[\left(\frac{M_{ijt}}{\bar{M}_{ij}} \right)^{\frac{\kappa(\sigma-1)-1}{\kappa(\sigma-1)}} - 1 \right]. \tag{8.4}$$

If the trade volume falls short of its steady-state value $\bar{M}_{ij}$, the growth rate is positive. The growth rate depends on the distance of the present trade volume to its steady-state volume: It becomes smaller as trade rises, and approaches zero as trade reaches its steady-state value.[19]

Interestingly, the growth rate does not depend on the rate at which the asset is accumulated, γ. This result reflects two effects that exactly cancel out: On the one hand, a higher rate γ leads to more asset accumulation, and therefore a higher growth rate of trade. On the other hand, it also raises the steady-state level of the invisible asset, and thereby lowers its average product in the neighborhood of the steady-state. This, in turn, reduces the growth rate of trade.

Approximately, the growth rate of trade can also be expressed as

$$\dot{M}_{ijt}/M_{ijt} \cong - \left[\kappa(\sigma - 1) - 1\right] \delta \ln(M_{ijt}/\bar{M}_{ij}). \tag{8.5}$$

This is an differential equation in $\ln(M_{ijt})$, the solution of which is

$$\ln(M_{ijt}) = \left(1 - e^{-\lambda\tau}\right) \ln(\bar{M}_{ij}) + e^{-\lambda\tau} \ln(M_{ij,t-\tau}), \tag{8.6}$$

where $\lambda = \kappa(\sigma - 1)\delta$, and τ is the time span between two observations.

Notice that equation (8.6) is closely related to the dynamic gravity equation (8.3). In the dynamic regressions, we proxy the steady-state value of trade by the dyad fixed effects

[18] $\kappa > 0$ guarantees that trade cost are inversely related to trust, and $\kappa < (\sigma - 1)^{-1}$ ensures concavity.

[19] Note the similarity to the neoclassical growth model.

ν_{ij}, and the country-and-time effects, ν_{it}, ν_{jt}.[20] The coefficient on the lagged endogenous variable obtained from the regression can be used to compute λ. Interestingly, λ quantifies the speed at which bilateral trade volumes adjust to their steady-state values. To see this, we compute the derivative of the growth rate (8.5) with respect to the trade volume (in logs)

$$\lambda = -\frac{d(\dot{M}_{ijt}/M_{ijt})}{d\left[ln(M_{ijt})\right]} = \kappa(\sigma-1)\delta\left(\frac{Mijt}{\bar{M}_{ijt}}\right)^{\frac{\kappa(\sigma-1)-1}{\kappa(\sigma-1)}}. \tag{8.7}$$

If bilateral trade is close to the steady-state, $M_{ijt} = \bar{M}_{ij}$, the adjustment rate can be proxied by $\lambda = \kappa(\sigma-1)\delta$. The true adjustment rate, however, depends on the distance to the steady-state.

Table 8.1 reports the convergence rate obtained from approximation around the steady-state. It turns out that $\lambda = 0.20$ per year.[21] Then, 20% of the gap between the present trade volume and its steady-state value vanishes in a year. The half-life time of adjustment, which is the time for half the initial gap to be eliminated, is about 3.5 years.[22] It takes about seven years to close three-quarters of the gap to vanish. Figure 8.2 visualizes the adjustment path based on the adjustment rate obtained from the Diff-GMM estimator. The solid line refers to the regression presented in Table 8.1.

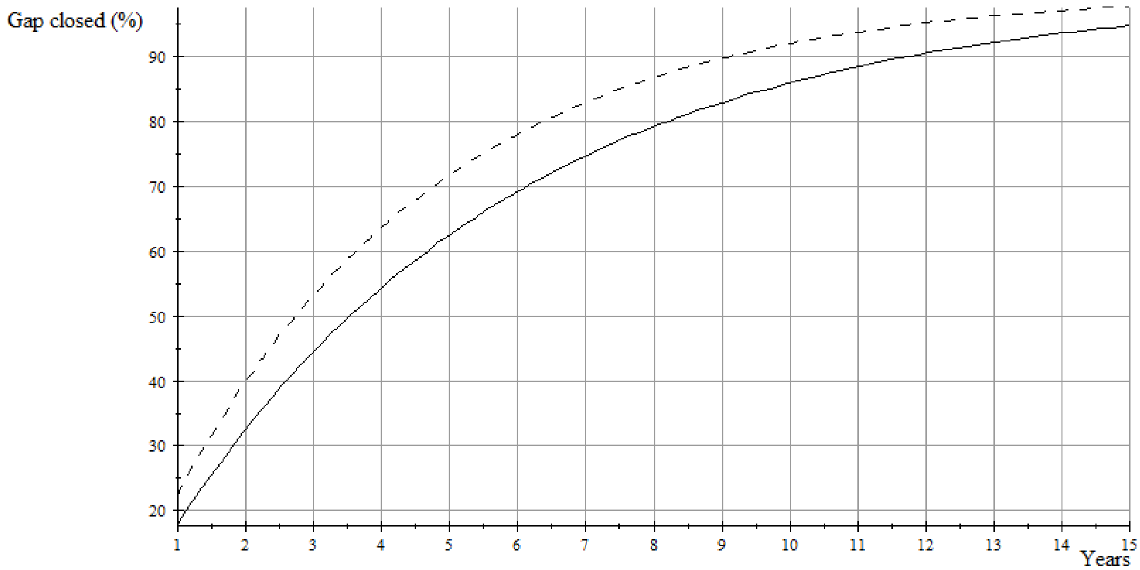
How does our finding compare to previous estimates? The basic specification in Egger (2001a) implies an adjustment rate of 47% per year. His sample covers 15 European countries, and the period 1986-1996. Micco et al. (2003) present dynamic panel estimation results for a sample of trade relationships between 22 developed countries. Using a period comparable to ours, 1980-2002, their estimates imply a convergence rate of about 55% a year, which is much higher than our result. Their estimates, however, are likely to be biased for the following reasons. First, they use yearly data, which are confounded by business cycle fluctuations. Second, they do not account for multilateral resistance terms. When it comes to the shorter period (1992-2002), the Arellano-Bond test for AR(2) in differences signals invalidity of the instruments. The misspecification is also reflected in the cofficient of the lagged endogenous variable, which would imply an adjustment rate of 137%.

[20] Below, we will also include time-varying trade policy controls.

[21] Note that $\hat{\alpha} = e^{-\lambda\tau}$ with $\tau = 5$ implies $\lambda = -\ln(\hat{\alpha})/5$.

[22] Solving $e^{-\lambda\tau^*} = 0.5$ for τ^* yields $\tau^* = -\ln(0.5)/\lambda$.

Figure 8.2: Adjustment to steady-state trade flows



Adjustment dynamics are derived from the formula $\ln(1-y) = \lambda t$, where y denotes the fraction of the gap closed at time t, and λ is the adjustment rate. Adjustment rate obtained from Diff-GMM regressions. Solid and dashed line refer to respectively full and OECD sample.

The estimates reported in De Nardis et al. (2008) also suffer from misspecification. In all regressions, the Hansen tests signal weak instruments. The implied convergence rate would be about 29%. Moser et al. (2008) and Martnez-Zarzoso et al. (2009) focus on German exports. Using the estimates reported by the former, the adjustment rate is about 71%,, whereas the latter find adjustment rates of 39%.

Results of previous studies are summarized in Table 8.2. In order to make our results comparable in terms of sample composition, we restrict our sample to OECD countries; see dashed line in Figure 8.2.[23] The adjustment rate in the OECD sample amounts to 26%. Hence, the within OECD adjustment rate is slightly higher but still below adjustment rates implied by previous studies.

The gravity approach has also been applied by Egger and Merlo (2007) to stocks of outward foreign direct investments. The sample mainly consists of OECD countries and covers the period 1980-2001. The p-values of the Hansen tests cast doubt on the validity of the test; see Roodman (2009). Columns (a_1) and (a_2) of their Table 1 correspond to an adjustment rate of 78%.

[23]Regression results for the OECD sample are shown in Table 8.8 in the Appendix.

Table 8.2: Sample composition and implied adjustment rates. Previous studies

Author(s)	Countries	Period	D^{Exp}	D^{Imp}	λ	τ^*
Egger (2001a)	EU15	1986-1996			47%	1.5
Micco et al. (2003)	Industrial	1980-2002			55%	1.3
Micco et al. (2003)	Industrial	1992-2002			137%	0.5
De Nardis et al. (2008)	EU15	1988-2004	YES	YES	29%	2.4
De Nardis et al. (2008)	EU15	1993-2004	YES	YES	33%	2.1
Moser et al. (2008)	Industrial*	1992-2003		YES	71%	1.0
Martinez-Zarzoso et al. (2009)	All*	1962-2005		YES	39%	1.8

*: Moser et al. (2008) and Martinez-Zarzoso et al. (2009) consider German exports. All studies use data on an annual basis and include time dummies. D^{Exp} and D^{Imp} denote sets of exporter and importer fixed effects, respectively. λ and τ^* are adjustment rate and half-life time, respectively.

For per capita income, Caselli et al. (1996) find an adjustment rate of about 10% per year, which implies a half-life of 7 years. Hence, trade flows take about half the time to close half the way from their initial position to their steady-state levels as compared to per capita income. This result seems plausible, since the adjustment of trade does not face barriers like the installation of physical capital.

8.4.2 Long-run trade policy effects

In this section, we use our approach to compute long-run trade policy effects. First, we consider FTA formation. Second, we follow Subramanian and Wei (2007) and interact FTA and WTO dummies. To be more precise, we construct a dummy which is set to one if both trading partners are member of an FTA and both member of the WTO at time t. Moreover, a dummy is set to one if both are in an FTA but at least one is not member of the WTO. Finally, we crate a dummy for common usage of the Euro.

Table 8.3: Long-run trade policy effects

Dependent variable: Bilateral trade	(A1) Diff-GMM	(A2) Sys-GMM	(B1) Diff-GMM	(B2) Sys-GMM	(C1) Diff-GMM	(C2) Sys-GMM	(D1) Diff-GMM	(D2) Sys-GMM
Lagged trade	0.371***	0.384***	0.371***	0.384***	0.358***	0.376***	0.372***	0.385***
	(14.11)	(13.37)	(14.10)	(13.36)	(13.74)	(13.21)	(14.15)	(13.39)
FTA	0.290***	0.314***						
	(4.65)	(4.53)						
FTA and both WTO			0.272***	0.296***			0.272***	0.296***
			(4.38)	(4.27)			(4.38)	(4.26)
FTA and at least one not WTO			0.542**	0.565**			0.543**	0.564**
			(2.21)	(2.26)			(2.21)	(2.26)
Euro					0.134**	0.152**	0.135**	0.147**
					(2.04)	(2.18)	(2.04)	(2.11)
Adjustment rate	0.198	0.191	0.198	0.191	0.206	0.196	0.198	0.191
Half-life time	3.501	3.623	3.492	3.620	3.371	3.541	3.503	3.627
Long-run effects								
FTA	0.461	0.510						
FTA and both WTO			0.432	0.481			0.433	0.480
FTA and at least one not WTO			0.862	0.917			0.864	0.917
Euro					0.209	0.244	0.215	0.238
AR(1) test (p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) test (p-value)	0.382	0.368	0.377	0.363	0.397	0.374	0.383	0.369
Sargan test (p-value)	0.485	0.140	0.482	0.133	0.763	0.0288	0.468	0.146
Hansen test (p-value)	0.663	0.740	0.661	0.735	0.851	0.617	0.651	0.742

Balanced sample of $N = 3921$ dyads, and $T = 5$ periods. Robust t-values in parentheses. *** and ** indicate significance at respectively 1% and 5%. All regressions include a comprehensive set of country-and-time effects. One-step GMM results. Trade policy controls treated exogenous. 2^{nd} and 3^{rd} lag in levels used as instrument for first differences equation. 2^{nd} and 3^{rd} lag in differences used as instruments for level equation. Set of instruments collapsed into a single column. Adjustment rate quantifies how much of distance to steady state is closed each year. Half-time life time reports time needed to close half the gap to steady state. See footnotes 21 and 22 for computational details. Long-run effects obtained from deviding short-run effect through one minus coefficient on lagged endogenous variable.

Table 8.3 reports the results obtained from Diff-GMM estimation (oddly numbered columns) and Sys-GMM estimation (evenly numbered columns). The following results stand out from Table 8.3. First, the dynamics relationship remains intact. Adjustment rates are of similar size across all specifications.

Second, the FTA estimate appears to be statistically significant as well. It amounts to approximately 0.3 with a t-value close to 4.5; see columns (A). Hence, the effect of FTA formation *on impact* corresponds to $exp(0.31) - 1 = 36\%$ trade creation. In the long-run, however, the short-run effect is almost doubled, and amounts to $exp(0.51) - 1 = 66\%$.[24] The long-run effect is substantially higher than the impact effect, because the FTA formation is a positive trade shock which induces asset accumulation. A higher stock of the invisible asset, in turn, implies lower trade costs. The long-run effect comprises both, the impact effect, and the indirect effect that runs through asset accumulation.

Third, the FTA effect is unevenly distributed across country pairs where both partners are members of the WTO and those where at least one partner is not. To be more precise, the effect is about twice as strong for the latter group, and amounts to $exp(0.9) - 1 = 145\%$ trade creation in the long run; see columns (B).

Fourth, the Euro significantly fosters trade; see columns (C). The long-run effect of 0.20 implies approximately 23% trade-cration. This finding is slightly above the estimate presented in De Nardis et al. (2008). They compute 17% trade creation in the long-run due to the the introduction of the Euro. Including all trade policy controls in one regressions leaves the conclusions unchanged; see columns (D).

8.4.3 'Phasing in' of FTAs versus endogenous invisible asset accumulation

Baier and Bergstrand (2007) argue that FTAs are 'phased-in' over a period of up to ten years. If trade cost reductions take place in several rounds, one needs lagged FTA dummies to capture the total effect of FTA formation on trade. The contemporaneous FTA takes the effect on impact of the first round on trade, and the five-year lagged FTA dummy is associated to the effect of further trade cost reductions within the first five years after the entry into force, and so on. In order to examine the effect of FTA formation on trade, one

[24] The long-run effect is obtained by dividing the short-run effect through one minus the effect of the lagged endogenous variable.

has to sum up the coefficients.

Static regression results obtained from a first-difference estimator (FD) are reported in Columns (1) to (3) of Table 8.4. They are qualitatively similar to those reported by Baier and Bergstrand (2007) in their Table 6. The point estimates, however, differ slightly because our sample is shorter. Besides the contemporaneous FTA dummy, the five-year lagged FTA dummy has a significant impact on today's bilateral trade, while the ten-year lagged dummy has not; see columns (2) and (3).

These findings can be interpreted in favor of the 'phasing-in' hypothesis, as done by Baier and Bergstrand (2007). They allow, however, for an alternative interpretation, which is observational equivalent. If trade costs were endogenous, the invisible asset would comprise information on the history of trade policy. Then, lagged FTA dummies would also significantly affect today's trade, but for different reasons. In case of 'phasing-in', the positive impact stems from second-round trade cost reductions, and reflects exogenous dynamics. In the case of asset accumulation, however, significant coefficients signal endogeneity of trade costs.

In order to assess whether the 'phasing-in' channel is active once asset accumulation is accounted for, we add a lagged FTA dummy, and re-estimate the dynamic gravity equation (8.3).[25] If the the five-year lagged FTA dummy entered significantly, this would signal FTA has a positive effect on trade *conditional* on the existence of asset accumulation.

Columns (4) and (5) of Table 8.4 present the regression results obtained from dynamic estimation. For both, the Diff-GMM and Sys-GMM estimator, the lagged FTA dummy turns out to be economically and statistically insignificant, while the other coefficients remain stable; compare columns (4) and (5) of Table 8.3 with columns (2) and (3) of Table 8.1. Thus, exogenous second-round trade cost reductions appear to have no effect on bilateral trade once the gravity model is cast dynamically.

Interestingly, static and dynamic regressions yield similar estimates of the long-run effect. For the static regressions, the figure refers to the trade creation after five years, whereas for the dynamic regressions we have computed state-steady effects. The reason for the coincidence is that trade flows converge rather fast, such that after five years 63% of

[25]Obviously, we do not lose observations when we include the lagged FTA dummy.

Table 8.4: Exogenous versus endogenous trade cost reductions

Dependent variable: Bilateral trade					
	(1) FD	(2) FD	(3) FD	(4) Diff-GMM	(5) Sys-GMM
Lagged trade				0.371***	0.383***
				(14.44)	(13.68)
FTA (contemporaneous)	0.287***	0.290***	0.250***	0.291***	0.307***
	(5.97)	(5.74)	(4.19)	(4.77)	(4.77)
FTA (1st lag)		0.117**	0.134**	−0.023	−0.014
		(2.16)	(2.25)	(−0.35)	(−0.21)
FTA (2nd lag)			0.080		
			(1.32)		
FTA cumul./long-run effect	0.287	0.407	0.384	0.463	0.497
AR(1) test (p-value)				0.000	0.000
AR(2) test (p-value)				0.383	0.369
Sargan test (p-value)				0.496	0.126
Hansen test (p-value)				0.672	0.734

Balanced sample of $N = 3921$ dyads, and $T = 5$ periods. Robust t-values in parentheses. *** and ** indicate significance at respectively 1% and 5%. All regressions include a comprehensive set of country-and-time effects. FD refers to first-difference estimator. One-step GMM results. FTA treated exogenous. 2^{nd} and 3^{rd} lag in levels used as instrument for first differences equation. 2^{nd} and 3^{rd} lag in differences used as instruments for level equation. Set of instruments collapsed into a single column. Cumulative FTA effect sums up significant FTA coefficient estimates. Long-run FTA effect obtained from deviding short-run effect through one minus coefficient on lagged endogenous variable.

the distance to the steady state is closed.[26]

8.4.4 Proximity shapes convergence patterns

So far, we have focused on *average* adjustment dynamics. This subsection aims at highlighting how geographical and cultural proximity shapes these dynamics.

Proximity potentially affects adjustment dynamics through two channels. First, proximity is a determinant of the initial invisible-asset stock. The geographically or culturally closer any two countries, the higher their initial stock of the invisible asset. Anything else equal, this would imply that proximate countries have a shorter distance to their steady-state position, and thus a lower adjustment rate; see equation (21).

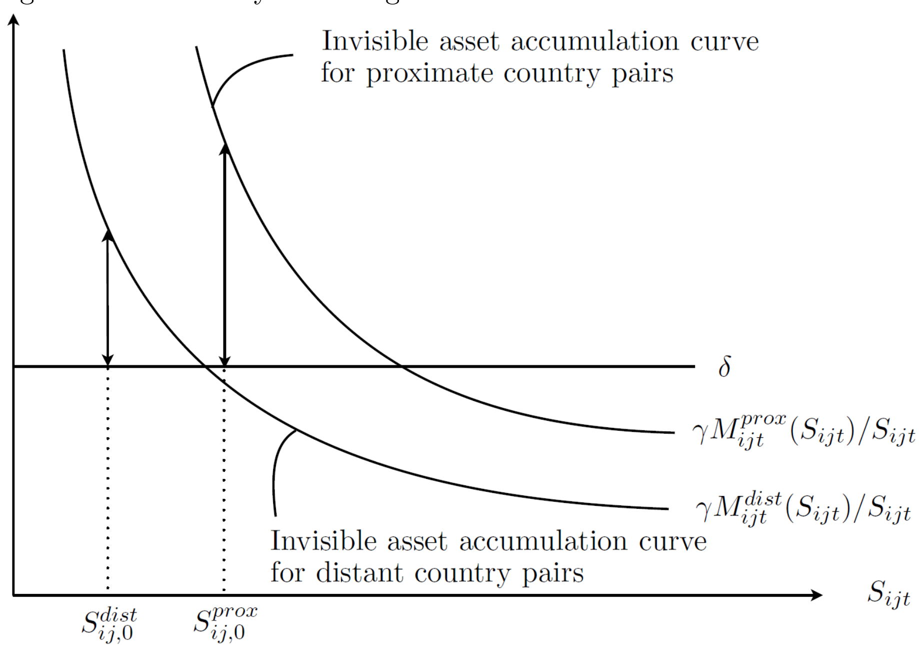
Second, however, proximate countries have lower trade costs. This is reflected by a shift of the invisible asset accumulation curve to the upper right. It also implies a higher steady-state level of trade (and the invisible asset). Hence, proximate countries can have a

[26]In order to obtain this result, evaluate $1 - e^{-\lambda t}$ at $\lambda = 0.2$ and $t = 5$.

longer distance to the steady-state position, which would imply a higher adjustment rate.

Given these two countervailing effects, the net effect of proximity on adjustment dynamics is *a priori* ambiguous. Figure 8.3 depicts a situation where the shift of the asset accumulation curve is strong. Then, proximate countries with a high initial stock $S_{ij,0}^{prox}$ can have a longer distance to their steady-state position, and thus a higher adjustment rate than distant countries with a low initial stock $S_{ij,0}^{dist}$. If the shift of the asset accumulation curve were only modestly, the original sorting would prevail: Geographically or culturally distant countries would have a longer distance to their steady-state positions than proximate countries, and thus higher adjustment rates.

Figure 8.3: Proximity and the growth rate of invisible asset accumulation



In order to make transparent how proximity shapes adjustment rates, we augment the standard dynamic gravity equation (8.3) with interactions between the lagged endogenous variable and proximity measures like geographical distance, contiguity, and common language. A smaller coefficient of the lagged endogenous variable is associated with a higher adjustment rate.

Clearly, the interaction terms are endogenous. Hence, we instrument them, using the same lag structure as for lagged trade. Interacting trade flows with geographical distance measured by the great circle distance would simply rescale a country pair's trade flows,

Table 8.5: How proximity shapes adjustment dynamics

Dependent variable: Bilateral trade	(1) Distance	(2) Contiguity	(3) Language	(4) All
Lagged trade	0.408***	0.380***	0.387***	0.416***
	(14.42)	(13.70)	(13.73)	(15.53)
Lagged trade × Distance	−0.023***			−0.025***
	(−3.15)			(−3.27)
Lagged trade × Contiguity		0.086***		0.084***
		(3.58)		(3.38)
Lagged trade × Language			−0.046***	−0.059***
			(−3.66)	(−4.66)
FTA	0.339***	0.307***	0.320***	0.329***
	(4.89)	(4.52)	(4.75)	(4.89)
Euro	0.152**	0.125*	0.145**	0.141**
	(2.19)	(1.83)	(2.10)	(2.04)
AR(1) test (p-value)	0.000	0.000	0.000	0.000
AR(2) test (p-value)	0.352	0.363	0.407	0.383
Sargan test (p-value)	0.0135	0.297	0.349	0.168
Hansen test (p-value)	0.171	0.780	0.678	0.510

Balanced sample of $N = 3921$ dyads, and $T = 5$ periods. Robust t-values in parentheses. ***, **, and * indicate significance at respectively 1%, 5%, and 10% . All regressions include a comprehensive set of country-and-time effects. One-step Sys-GMM results. Trade policy controls treated exogenous. 2^{nd} and 3^{rd} lag in levels used as instrument for first differences equation. 2^{nd} and 3^{rd} lag in differences used as instruments for level equation. Set of instruments collapsed into a single column.

and the respective lags turn out to be weak instruments. We therefore create a dummy *Distance* that takes one if a country pair's distance is above the median sample distance, and zero else.

Columns (1) to (3) in Table 8.5 show the Sys-GMM results for different interaction terms separately, while in column (4) all interaction terms of interest are included. It turns out from columns (1) and (2) that geographically close pairs tend to have a significantly longer half-life time, i.e. they need more time to close half the way from their initial position to the steady state. This finding suggests that the shift of the asset accumulation curve is only modest.

When it comes to cultural proximity, however, the pattern looks different. Trade between country pairs in which both countries have a common language seem to have a higher adjustment rate; see column (3). This implies that culturally proximate countries have longer distance to the steady state, and refers to the situation depicted in Figure 8.3.

Conditional results in column (4) are qualitatively and quantitatively similar to the results obtained from separate regressions; compare column (4) to columns (1) to (3).

8.4.5 Trade flows of exchange-traded goods adjust faster

According to our theoretical considerations, the adjustment rate is not only affected by the elasticity of trade costs with respect to trust and the depreciation rate of the invisible asset, but also determined by the elasticity of trade with respect to trade costs.[27] By the CES utility function, the latter elasticity is characterized by the elasticity of substitution between varieties, σ.

Following Rauch (1999), we consider three different commodity groups, namely exchange-traded, referenced-priced, and differentiated goods. Exchange-traded goods are show-cases for homogeneous goods such as steel, corn, or ore. Thus, they are characterized by a high elasticity of substitution, $\sigma \gg 1$, whereas the elasticity of substitution between differentiated varieties is close to unity.

Hence, for exchange-traded goods small reductions in trade costs lead to a strong increase in trade. For differentiated goods, however, a large reduction in trade cost is required to raise trade.

The elasticity of trade costs with respect to the stock of the invisible asset may work in the opposite direction: If the influence of trust on trade costs of exchange-traded goods is small, it requires a high stock of the asset to decrease trade costs. Since it is also unclear how the asset depreciation rates differ across commodity groups, the net effect of invisible asset accumulation on trade for different commodity groups is *a priori* ambiguous.

[27] Recall that around the steady state the adjustment rate is given by $\lambda = \kappa(\sigma - 1)\delta$.

Table 8.6: Adjustment dynamics by commodity group (Liberal aggregation)

Dependent variable: Bilateral trade by commodity group									
	Exchange-traded goods			Reference-priced goods			Differentiated goods		
	(A1)	(A2)	(A3)	(B1)	(B2)	(B3)	(C1)	(C2)	(C3)
Lagged trade	0.234***	0.248***	0.251***	0.345***	0.363***	0.357***	0.304***	0.323***	0.316***
	(5.35)	(5.72)	(5.92)	(7.66)	(8.13)	(8.23)	(7.62)	(8.09)	(8.18)
FTA (cont.)		0.324***	0.327***		0.188***	0.185***		0.222***	0.217***
		(2.94)	(2.97)		(2.72)	(2.69)		(3.35)	(3.32)
FTA (1st lag)			0.0735			−0.0722			−0.0752
			(0.80)			(−1.10)			(−1.06)
Adjustment rate	0.290	0.279	0.277	0.213	0.203	0.206	0.238	0.226	0.230
Half-life time	2.387	2.489	2.506	3.254	3.421	3.362	2.911	3.066	3.012
AR(1) test (p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sargan test (p-value)	0.185	0.267	0.278	0.864	0.719	0.770	0.559	0.632	0.572
Hansen test (p-value)	0.217	0.301	0.314	0.872	0.733	0.781	0.581	0.651	0.592

Balanced sample of $N = 1649$, $N = 1768$, and $N = 1837$ dyads for respectively exchange-traded, reference-priced, and differentiated goods, and $T = 4$ periods. Robust t-values in parentheses. ***, **, and * indicate significance at respectively 1%, 5%, and 10%. All regressions include a comprehensive set of country-and-time effects. Diff-GMM results. FTA treated exogeneous. 2^{nd} and 3^{rd} lag in levels used as instrument for first differences equation. 2^{nd} and 3^{rd} lag in differences used as instruments for level equation. Set of instruments collapsed into a single column. Adjustment rate quantifies how much of distance to steady state is closed each year. Half-time life time reports time needed to close half the gap to steady state. See footnotes 21 and 22 for computational details. Long-run effects obtained from deviding short-run effect through one minus coefficient on lagged endogenous variable.

In order to make transparent differences in adjustment dynamics across commodity groups, we repeat our dynamic gravity estimation for exchange-traded, reference-priced, and differentiated goods separately.[28] Table 8.6 presents the results.[29] Columns (A1) and (C1) show that the adjustment rate is higher for exchange-traded (29.7%) than for differentiated goods (22.1%). Hence, the adjustment rates seem to reflect differences in the elasticity of substitution. Accordingly, exchange-traded goods need less time to close the distance to the steady-state.[30]

Also the effect of FTA formation on impact differs across commodity groups; see columns (A2), (B2), and (C2). It is highest for trade in differentiated goods, where it corresponds to a trade creation of $e^{0.324} - 1 = 38\%$. Notice that this result is comparable to the finding for aggregate trade.

The long-run effects of FTA comprise both, the effect on impact and the further endogenous reduction of trade costs due to invisible asset accumulation. For exchange-traded goods, the long-run effect amounts to $e^{0.43} - 1 = 54\%$ trade creation, wheras for differentiated goods the corresponding effect is only $e^{0.33} - 1 = 39\%$ trade creation.[31]

Table 8.6 also contrasts further exogenous with endogenous trade cost reductions after FTA formation. Columns (A3), (B3), and (C3) report results from a dynamic regression where we add a five-year lagged FTA dummy. As in the case of aggregate bilateral trade as dependent variable, we can make two observations. First, the inclusion of the lagged FTA dummy does not affect estimated coefficients of lagged trade and contemporaneous FTA dummy. Second, for all commodity groups the lagged FTA dummies turn out to be insignificant. Hence, we do not find evidence for ‘phasing-in’ of FTAs once accounting for invisible asset accumulation.

[28] Recall that in order to avoid a structural break in the trade data, we narrow our sample to the period 1985-2000. With $T = 4$, however, we cannot perform an Arellano-Bond test for AR(2).

[29] Results for the alternative ‘conservative’ aggregation of goods is shown in Table 8.9 in the Appendix.

[30] The corresponding half-life times are respectively 2.489 years and 3.066 years.

[31] The trade creation is smallest for reference-priced goods (35%). This commodity group, however, comprises residual SITC categories, which can neither be assigned to exchange-traded nor differentiated goods.

8.5 Concluding remarks

This paper has introduced an 'invisible asset' into a standard Anderson and van Wincoop (2003) model of international trade, which endogenously accumulates as a by-product of existing trade relationships. Then, trade costs become endogenous, and a dynamic gravity estimation is required. Our dynamic panel data estimates include country-and-time effects are consistent. Standard tests do not cast doubt on the validity of the instrumentation strategy. Our results are robust across the use of different dynamic panel data estimators.

We use our approach to answer the question how rapidly bilateral trade flows adjust to their steady-state levels. This information is needed to assess how long it takes to fully achieve potential gains induced by trade reforms.

As expected, trade flows are persistent. Our theory-consistent estimations, however, suggest that the adjustment rate is lower than those implied by previous studies. Nevertheless, trade flows adjust faster to steady-state levels than per capita income: Whereas it takes seven years to close half the gap from initial position to steady-state per capita income, trade flows only need half the time.

Our results imply that common membership in an FTA increases bilateral trade by 36% on impact. In the long run, however, this effect is almost doubled. The reason is that trade impediments are further reduced endogenously as a by-product of the trade creation on impact. We show that lagged FTA dummies, which enter static regressions significantly, reflect the endogeneity of trade costs rather than a 'phasing-in' of FTAs.

We also investigate the convergence rate for different groups of countries and across commodities. We find geographical close countries to faster close the gap to the steady state. Moreover, trade in differentiated goods needs longer to achieve trade potentials after trade reform, whereas trade in exchange-traded goods converges faster. Along with a higher effect of FTA formation on trade on impact, this implies that FTA formation leads to more trade creation in exchange-traded goods than in differentiated goods in the long-run.

8.6 Appendix

Table 8.7: List of countries

Albania	Germany	Niger
Algeria	Ghana	Nigeria
Angola	Greece	Norway
Argentina	Guatemala	Pakistan
Australia	Guinea-Bissau	Panama
Austria	Guyana	Paraguay
Bangladesh	Haiti	Peru
Belgium	Honduras	Philippines
Bolivia	Hong Kong	Poland
Brazil	Hungary	Portugal
Bulgaria	India	Romania
Burkina Faso	Indonesia	Saudi Arabia
Cameroon	Iran	Senegal
Canada	Ireland	Sierra Leone
Chile	Israel	Singapore
China	Italy	Spain
Colombia	Jamaica	Sri Lanka
Congo	Japan	Sudan
Congo, Dem. Republic	Kenya	Sweden
Costa Rica	Korea, Republic of	Switzerland
Cote d'Ivoire	Madagascar	Syrian Arab Republic
Cyprus	Malawi	Thailand
Denmark	Malaysia	Trinidad and Tobago
Dominican Republic	Mali	Tunisia
Ecuador	Mauritania	Turkey
Egypt	Mauritius	Uganda
El Salvador	Mexico	United Kingdom
Ethiopia	Morocco	United States of America
Finland	Mozambique	Uruguay
France	Netherlands	Venezuela
Gabon	New Zealand	Zambia
Gambia	Nicaragua	Zimbabwe

Table 8.8: Adjustment dynamics of bilateral trade flows. OECD sample

Dependent variable: Bilateral trade			
	(1)	(2)	(3)
	2SLS	Diff-GMM	Sys-GMM
Lagged trade	0.267***	0.267***	0.267***
	(2.79)	(2.79)	(2.63)
Adjustment rate	0.264	0.264	0.264
Half-life time	2.623	2.623	2.621
AR(1) test (p-value)	0.002	0.002	0.003
AR(2) test (p-value)	0.302	0.302	0.302
Sargan test (p-value)		0.0913	0.149
Hansen test (p-value)		0.167	0.381

Balanced sample of $N = 8791$ dyads, and $T = 5$ periods. Robust t-values in parentheses. *** indicates significance at 1%. All regressions include a comprehensive set of country-and-time effects. 2SLS estimator uses 2^{nd} lag in levels to instrument lagged endogenous variable in differences. One-step GMM results. 2^{nd} and 3^{rd} lag in levels used as instrument for first differences equation. 2^{nd} and 3^{rd} lag in differences used as instruments for level equation. Set of instruments collapsed into a single column. OLS includes controls for distance, contiguity, common language, and colonial ties (all not shown). Adjustment rate quantifies how much of distance to steady state is closed each year. Half-time life time reports time needed to close half the gap to steady state. See footnotes 21 and 22 for computational details.

Table 8.9: Adjustment dynamics by commodity group (Conservative aggregation)

Dependent variable: Bilateral trade by commodity group									
	Exchange-traded goods			Reference-priced goods			Differentiated goods		
	(A1)	(A2)	(A3)	(B1)	(B2)	(B3)	(C1)	(C2)	(C3)
Lagged trade	0.234***	0.248***	0.251***	0.345***	0.363***	0.357***	0.304***	0.323***	0.316***
	(5.35)	(5.72)	(5.92)	(7.66)	(8.13)	(8.23)	(7.62)	(8.09)	(8.18)
FTA (cont.)		0.324***	0.327***		0.188***	0.185***		0.222***	0.217***
		(2.94)	(2.97)		(2.72)	(2.69)		(3.35)	(3.32)
FTA (1st lag)			0.0735			−0.0722			−0.0752
			(0.80)			(−1.10)			(−1.06)
Adjustment rate	0.290	0.279	0.277	0.213	0.203	0.206	0.238	0.226	0.230
Half-life time	2.387	2.489	2.506	3.254	3.421	3.362	2.911	3.066	3.012
AR(1) test (p-value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sargan test (p-value)	0.185	0.267	0.278	0.864	0.719	0.770	0.559	0.632	0.572
Hansen test (p-value)	0.217	0.301	0.314	0.872	0.733	0.781	0.581	0.651	0.592

Balanced sample of $N = 1649$, $N = 1768$, and $N = 1837$ dyads for respectively exchange-traded, reference-priced, and differentiated goods, and $T = 4$ periods. Robust t-values in parentheses. ***, **, and * indicate significance at respectively 1%, 5%, and 10%. All regressions include a comprehensive set of country-and-time effects. Diff-GMM results. FTA treated exogeneous. 2^{nd} and 3^{rd} lag in levels used as instrument for first differences equation. 2^{nd} and 3^{rd} lag in differences used as instruments for level equation. Set of instruments collapsed into a single column. Adjustment rate quantifies how much of distance to steady state is closed each year. Half-time life time reports time needed to close half the gap to steady state. See footnotes 21 and 22 for computational details. Long-run effects obtained from deviding short-run effect through one minus coefficient on lagged endogenous variable.

Chapter 9

Concluding remarks

This doctoral thesis contributes to the process of 'unpacking' trade costs. This chapter summarizes the main findings. Moreover, it gives an outlook on further research.

9.1 Main findings

Technical barriers to trade. Technical barriers to trade may play an *efficiency-increasing* role. This result arises in a model where the degree of external scale effects can be parameterized independently of the elasticity of substitution. Then, there is a second-best rationale for the existence of TBTs since they can be used as an instrument to restrain excessive entry of monopolistic firms. The effect of TBT reform on industry productivity depends on the parameter constellation and is *a priori* ambiguous. However, the numerical exercise suggests that for a wide range of parameter constellations the input diversity effect is large enough to offset the negative effect on average productivity of input producers. Then, industry productivity increases.

The theoretical finding has important implications. First, it may rationalize the observation that there is no robust empirical link between openness and productivity. The model predicts the relation between openness and productivity is ambiguous, while variable trade cost and TBT liberalization both increase the openness of industries. Second, the level of TBTs determines whether reductions in variable trade costs improve aggregate productivity. This interdependence calls for an integrative approach in trade policy. Third, our analysis suggests that TBT reform typically is harder to achieve politically than tariff reform. The reason is that, under a range of parameter constellations, existing exporters would lose market share from lower TBTs but gain from lower variable trade costs.

Information barriers to trade. Using newly available panel data on developing countries diaspora to rich OECD nations in a theory-grounded gravity model, we uncover a robust, causal effect of bilateral migration on bilateral trade. We also confirm the effect of Chinese ethnic networks on bilateral trade. In terms of magnitudes, however, trade creation associated to the Chinese ethnic network is at most half as big as the one computed by Rauch and Trindade (2002). Trade creation effects of indirect network links, which are less confounded by the preference channel, are on average very small and, indeed, often indistinguishable from zero.

Drawing on data obtained from Reuters-Business-Briefings, we find that cooperative behavior of importers induces trade flows while non-cooperative behavior deters them. However, the trade-reducing effect of non-cooperative behavior is stronger than the trade-creating effect of cooperative behavior.

Incomplete cross-border contracts. In the presence of incomplete cross-border contracts, variable trade costs become endogenous. Due to the hold-up problem, the quantity produced for a foreign market is reduced. Moreover, a double marginalization problem arises. The model predicts a sorting of exporters along their comparative advantages, which is confirmed by the data. Comparative advantage of a firm covers not only productivity but also brand reputation and marketability of a variety.

Search and matching frictions in business-to-business relationships. The introduction of search and matching frictions endogeneizes both variable trade costs and fixed market entry costs. The model presented in Chapter 7 predicts a sorting of firms along their comparative advantages, but is otherwise close to the analytical frontier.

Trust in bilateral trade relationships. We confirm the empirical finding that trade history matters for present trade flows. This evidence is in line with the idea that trust accumulates when trade takes place. Trade flows close on average 20% each year of the gap from their initial positions to steady-state values. Casting the gravity relation dynamically, there is no additional role for 'phasing-in' of free trade agreements.

9.2 Future research

To close this thesis, we briefly sketch an array of interesting future research.

Asymmetric trade costs. The standard Melitz (2003) model limits the analysis to the symmetric case. Hence, it requires variable trade costs and fixed market entry costs to be identical across trading partners. This, of course, serves the purpose of carving out the general driving forces. Empirically testable hypotheses and relevant trade policy advice, however, should be derived from a model with *asymmetric* trade costs. Recent theoretical advances address this issue for special classes of utility functions. Melitz and Ottaviano (2008) use quasi-linear preferences but do not introduce fixed market entry costs. Behrens et al. (2008) also allow for multiple asymmetric countries in a setting with a very specific utility function and under the assumption that firm productivities are Pareto.

Strategic setting of trade costs. While this thesis goes some way in endogeneizing variable or fixed trade costs, it does not consider the optimal setting of tariffs and regulatory costs. This, however, would give guidance in the recent debate on the new protectionism. Hence, it would be interesting to study the strategic setting of variable and fixed trade costs in an asymmetric two-country model. Recent research on tariffs in models with heterogeneous firms does not consider strategic setting of fixed market entry costs at all and suffer from various limitations. Jørgensen and Schröder (2005) treat tariffs as symmetric and exogenous, while Demidova and Rodríguez-Clare (2007) study a small economy. Cole (2008) pushes income changes onto a numeraire sector in order to rule out strategic interactions.

Structural estimation and identification of key parameters. The theoretical analysis of trade liberalization scenarios may uncover ambiguities. Simulation exercises, however, require estimates of key parameters. Helpman et al. (2004) do not dissect elasticity of substitution and dispersion of firms' productivities. Estimates of the latter for various industries provided by Corcos et al. (2007) seem to conflict with some of the model assumptions. Recent theoretical advances help to improve the empirical fit of the productivity distribution; see Arkolakis (2008), and Eaton, Kortum, Kramarz (2008). After all, this calls for structural estimation and identification of the key parameters in trade models with heterogeneous firms: productivity dispersion, elasticity of substitution, and degree of external economies of scale.

Bibliography

[1] Anderlini, L., and L. Felli (2001). Costly Bargaining and Renegotiation. *Econometrica* **68**: 377-411.

[2] Anderlini, L., and L. Felli (2004). Transaction Costs and the Robustness of the Coase Theorem. *Economic Journal* **116**: 223-245.

[3] Anderson, T.W., and C. Hsiao (1981). Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association* **76**: 598-606.

[4] Anderson, T.W., and C. Hsiao (1982). Formulation and Estimation of Dynamic Models Using Panel Data. *Journal of Econometrics* **18**: 47-82.

[5] Anderson, J.E. and E. van Wincoop (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review* **93**(1): 170-92.

[6] Anderson, J.E. and E. van Wincoop (2004). Trade Costs. *Journal of Economic Literature* **42**: 691-751.

[7] Antras, P. and E. Helpman (2004). Global Sourcing. *Journal of Political Economy* **112**(3): 552-580.

[8] Antras, P., and E. Helpman (2008). Contractual Frictions and Global Sourcing. In: Helpman, E., D. Marin, and T. Verdier. *The Organization of Firms in a Global Economy.* Harvard University Press.

[9] Ardelean, A. (2007). How Strong is Love of Variety? Mimeo: Purdue University.

[10] Arellano, M. and S. Bond (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and An Application to Employment Equations. *The Review of Economic Studies* **58**: 277-297.

[11] Arkolakis, C. (2008). Market Penetration Costs and the New Consumers Margin in International Trade. *NBER Working Papers* 14214.

[12] Arrow, K. (1972). Gifts and Exchanges. *Philosophy and Public Affairs* **1**: 343-362.

[13] Baier, S.L., and J.H. Bergstrand (2001). The Growth of World Trade: Tariffs, Transport Costs and Income Similarity. *Journal of International Economics* **53**: 1-27.

[14] Baier, S.L., and J.H. Bergstrand (2007). Do Free Trade Agreements Actually Increase Members' International Trade? *Journal of International Economics* **71**: 72-95.

[15] Baier, S.L., J.H. Bergstrand, P. Egger, and P.A. McLaughlin (2008). Do Economic Integration Agreements Actually Work? Issues in Understanding the Causes and Consequences of the Growth of Regionalism. *The World Economy* **31**(4): 461-497.

[16] Baldwin, R. E. (1988). Hysteresis in Import Prices: The Beachhead Effect. *American Economic Review* **78**(4): 773-785.

[17] Baldwin, R.E., V. DiNino, L. Fontagné, R.A. De Santis and D. Taglioni (2008). Study on the Impact of the Euro on Trade and Foreign Direct Investment. European Union, *Economic Papers* 321.

[18] Baldwin, R.E., and R. Forslid (2006). Trade Liberalization with Heterogeneous Firms. *NBER Working Papers* 12192.

[19] Baldwin, R.E., and D. Taglioni (2006). Gravity for Dummies and Dummies for Gravity Equations. *NBER Working Papers* 12516.

[20] Baltagi, B.H, P. Egger, and M. Pfaffermayr (2003). A Generalized Design for Bilateral Trade Flow Models. *Economics Letters* **80**(3): 391-397.

[21] Bandyopadhyay, S., C.C. Coughlin, and H.J Wall (2008). Ethnic Networks and US Exports. *Review of International Economics* **16**(1): 199-213.

[22] Barro, R.J., and X. Sala-i-Martin (2003). *Economic Growth*. Cambridge. MIT Press.

[23] Beck, T., and R. Levine (2004). Stock Markets, Banks, and Growth: Panel Evidence. *Journal of Banking and Finance* **28**(3): 423-442.

[24] Beghin, J. (2008). Non-tariff Barriers. In: S.N. Durlauf and L.E. Blume (eds). *The New Palgrave Dictionary of Economics*. Second Edition. Palgrave Macmillan.

[25] Behrens, K., G. Mion, Y. Murata, J. Südekum (2008). Trade, Wages, and Productivity. *IZA Discussion Paper* 3682.

[26] Beine, M., F. Docquier, and A. Marfouk (2008). Brain Drain and Human Capital Formation in Developing Countries: Winners and Losers. *The Economic Journal* **118**: 631-652.

[27] Benassy, J.-P. (1996). Taste for Variety and Optimum Production Patterns in Monopolistic Competition. *Economics Letters* **52**: 41-47.

[28] Bergin, P. R., and R. Glick (2007). Tradability, Productivity, and Understanding International Economic Integration. *Journal of International Economics* **73**(1): 128-151.

[29] Bernard, A., J.S. Redding, and P.K. Schott (2006). Multi-product Firms and Product Switching. *NBER Working Papers* 12293.

[30] Bernard, A., J.S. Redding, and P. K. Schott (2007a). Comparative Advantage and Heterogeneous Firms. *Review of Economic Studies* **74**(1): 31-66.

[31] Bernard, A., J.B. Jensen, and P.K. Schott (2007b). Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Goods. In: T. Dunne, J.B. Jensen and M.J. Roberts (eds.) *Producer Dynamics: New Evidence from Micro Data*. University of Chicago Press.

[32] Bernard, A., J.B. Jensen, J.S. Redding, and P.K. Schott (2008). Intra-firm Trade and Product Contractibility. Mimeo: Tuck School of Business at Dartmouth.

[33] Bernard, A., J.B. Jensen, and P.K. Schott (2008). Transfer Pricing by U.S.-Based Multinational Firms. *CES Working Paper* CES-WP-08-29.

[34] Besedeš, T. and T.J. Prusa (2006). Ins, Outs, and the Duration of Trade. *Canadian Journal of Economics* **39**(1): 266-295.

[35] Blanchard, O. and F. Giavazzi (2003). Macroeconomic Effects of Regulation and Deregulation in Goods and Labor Markets. *The Quarterly Journal of Economics* **118**(3): 879-907.

[36] Blundell, R. and S. Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics* **87**: 115-143.

[37] Cole, M. (2008). Optimal Tariffs, Tariff Jumping, and Heterogeneous Firms. Mimeo: University of Oregon.

[38] Bond, S. (2002). Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice. *Portuguese Economic Journal* **1**(2): 141-162.

[39] Bond, D., J. Bond, O. Churl O., J.C. Jenkins and C.L. Taylor (2003). Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development. *Journal of Peace Research* **40**(6): 733-745.

[40] Bond, S., A. Hoeffler, and J. Temple (2001). GMM Estimation of Empirical Growth Models. *CEPR Discussion Paper* 3048.

[41] Borga, M., and W.J. Zeile (2004). International Fragmentation of Production and the Intrafirm Trade of U.S. Multinational Companies. *BEA Working Papers* 0013. Bureau of Economic Analysis.

[42] Buch, C., J. Kleinert, A. Lipponer, and F. Toubal (2005). Determinants and Effects of Foreign Direct Investment: Evidence from German Firm-Level Data. *Economic Policy* **20**: 53-110.

[43] Calderon, C.A., A. Chong, and N.V. Loayza (2002). Determinants of Current Account Deficits in Developing Countries. *Contributions to Macroeconomics* **2**(1): Art. 2.

[44] Carkovic, M., and R. Levine (2005). Does Foreign Direct Investment Accelerate Economic Growth? In: T.H. Moran, E.M. Graham, and M. Blomström. Does Foreign Direct Investment Promote Development? Washington, DC: Institute for International Economics and Center for Global Development.

[45] Casella, A., and J. E. Rauch (2002). Anonymous Markets and Group Ties in International Trade. *Journal of International Economics* **58**: 19-48.

[46] Chaney, T. (2008). Distorted Gravity: The Intensive and Extensive Margins of International Trade. *American Economic Review* **98**: 1707-1721.

[47] Combes, P-P., M. Lafourcade and T. Mayer (2005). The Trade Creating Effects of Business and Social Networks: Evidence from France. *Journal of International Economics* **66**: 1-29.

[48] Corcos, G., M. Del Gatto, G. Mion, and G.I.P. Ottaviano (2007). Productivity and Firm Selection: Intra- vs International Trade. *CORE Discussion Paper* No. 2007/60.

[49] Corsetti, G., P. Martin, and P. Pesenti (2007). Productivity, Terms of Trade and the 'Home Market Effect'. *Journal of International Economics* **73**: 99-127.

[50] Cummins, J., and G. Violante (2002). Investment-Specific Technical Change in the U.S. (1947-2000): Measurement and Macroeconomic Consequences. *Review of Economic Dynamics* **5**(2): 243-284.

[51] De Benedictis, L., and C. Vicarelli (2005). Trade Potentials in Gravity Panel Data Models. *Topics in Economic Analysis & Policy* **5**(1): Art. 20 .

[52] De Grauwe P., and F. Skudelny (2000). The Impact of EMU on Trade Flows. *Weltwirtschaftliches Archiv/Review of World Economics* **136**: 381-402.

[53] Demidova, S., and A.Rodríguez-Clare (2007). Trade Policy under Firm-Level Heterogeneity in a Small Economy. *NBER Working Papers* 13688.

[54] De Nardis, S., R. de Santis, and C. Vicarelli (2008). The Euro's Effects on Trade in a Dynamic Setting. *The European Journal of Comparative Economics* **5**(1): 73-85.

[55] Dixit, A. (1989). Hyteresis, Import Penetration, and Exchange Rate Pass-Through. *Quarterly Journal of Economics* **104**(2): 205-228.

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

[57] Docquier, F., and E. Lodigiani (2009). Skilled Migration and Business Networks. Fortcoming in *Open Economies Review.*

[58] Docquier, F., and A. Marfouk (2006). International Migration by Education Attainment, 1990-2000. In: C. Ozden and M. Schiff (eds). International Migration, Brain Drain and Remittances. 151-199. New York: Palgrave Macmillan.

[59] Dunlevy, J.A. (2006). The Influence of Corruption and Language on the Protrade Effect of Immigrants: Evidence from the American States. *The Review of Economics and Statistics* **88**(1): 182-186.

[60] Eaton, J., S. Kortum, and F. Kramarz (2004). Dissecting Trade: Firms, Industries, and Export Destinations. *American Economic Review* **94**(2): 150-154.

[61] Eaton, J., S. Kortum, and F. Kramarz (2008). An Anatomy of International Trade: Evidence from French Firms. *NBER Working Papers* 14610.

[62] Eckel, C., and P. Neary (2006). Multi-product Firms and Flexible Manufacturing in the Global Economy. *CEPR Discussion Paper* 5941.

[63] Egan, M.L. and A. Mody (1993). Buyer Seller Links in Export Development. *World Development* **20**(3): 321-334.

[64] Egger, P. (2001a). European Exports and Outward Foreign Direct Investment: A Dynamic Panel Data Approach. *Weltwirtschaftliches Archiv/Review of World Economics* **137**(3): 427-449.

[65] Egger, P. (2001b). European Integration in Trade and FDI: A Dynamic Perspective. *CESifo Forum* **2**(2): 30-35.

[66] Egger, H. and U. Kreickemeier (2009). Firm Heterogeneity and the Labour Market Effects of Trade Liberalisation. *International Economic Review* **50**(1): 187-216.

[67] Egger, P., and V. Merlo (2007). The Impact of Bilateral Investment Treaties on FDI Dynamics. *The World Economy* **30**(10): 1536-1549.

[68] Ellis, P. (2000). Social Ties and Foreign Market Entry. *Journal of International Business Studies* **31**(3): 443-469.

[69] Evans, C.L. (2003). The Economic Significance of National Border Effects. *American Economic Review* **93**(4): 1291-312.

[70] Feenstra, R.C. (1997). U.S. Exports, 19721994: With State Exports and Other U.S. Data. *NBER Working Papers* 5990.

[71] Feenstra, R.C. (2000). World Trade Flows, 1980-1997. University of California, Davis.

[72] Feenstra, R. (2004). *Advanced International Trade. Theory and Evidence.* Princeton: Princeton University Press.

[73] Feenstra, R.C., W. Hai, W.T. Woo, and S. Yao (1999). Discrepancies in International Data: An Application to China-Hong Kong Entrepôt Trade. *American Economic Review* **89**(2): 338-343.

[74] Feenstra, R.C., and G.H. Hanson, (2004). Intermediaries in Entrepôt Trade: Hong Kong Re-Exports of Chinese Goods. *Journal of Economics and Management Strategy* **13**: 3-36.

[75] Feenstra, R.C., G.H. Hanson, and S. Lin (2004). The Value of Information in International Trade: Gains to Outsourcing through Hong Kong. *Advances in Economic Analysis & Policy* **4**(1): Article 7.

[76] Feenstra, R.C., R.E. Lipsey, H. Deng, A.C. Ma, and H. Mo (2005). World Trade Flows: 1962-2000. *NBER Working Papers* 11040.

[77] Feenstra, R. C. and H. Ma (2007). Multiproduct Firms Under Heterogeneous Costs and Monopolistic Competition. Mimeo: University of California, Davis.

[78] Felbermayr, G., and B. Jung (2006). Business News, Bayesian Updating, and Bilateral Trade. Mimeo: University of Tübingen.

[79] Felbermayr, G., and B. Jung (2008a). Sorting It Out: Technical Barriers to Trade and Industry Productivity. *FIW Working Paper* 14.

[80] Felbermayr, G., and B. Jung (2008b). The Pro-Trade Effect of the Brain Drain: Sorting Out Confounding Factors. *Hohenheimer Diskussionsbeiträge* 302/2008.

[81] Felbermayr, G., and B. Jung (2008c). Trade Intermediaries, Incomplete Contracts, and the Choice of Export Modes. *Tübingen Discussion Paper* 317.

[82] Felbermayr, G. and B. Jung (2009b). Modeling Trade Intermediation in General Equilibrium. Forthcoming in: P. Mooslechner and E.Gnan (editors): "International Trade & Domestic Growth:Determinants, Linkages and Challenges". Workshop Proceedings, Austrian National Bank, Vienna.

[83] Felbermayr, G. B. Jung, and F. Toubal (2009). Ethnic Networks, Information, and International Trade: Revisiting the Evidence. *Hohenheimer Diskussionsbeiträge* 303/2009.

[84] Felbermayr, G., and W. Kohler (2004). Exploring the Intensive and Extensive Margins of World Trade. *Review of World Economics* **142**(4): 642-674.

[85] Felbermayr, G., and J. Prat (2007). Product Market Regulation, Firm Selection and Unemployment. *IZA Discussion Paper* No. 2754l.

[86] Felbermayr, G., and F. Toubal (2008). Revisiting the Trade-Migration Nexus: Evidence from New OECD data. Mimeo: Paris.

[87] Fryges, H. (2006). Hidden Champions–How Young and Small Technology-Oriented Firms Can Attain High Export-Sales Ratios. *ZEW Discussion Paper* 06-045.

[88] Fryges, H. (2007). Change of Sales Modes in International Markets: Empirical Results for German and British High-Tech Firms. *Progress in International Business Research* **1**: 139-185.

[89] Gibson, M.J. (2008). Trade Liberalization, Reallocation, and Productivity. Mimeo: Washington State University.

[90] Girma, S., and Z. Yu (2000). The Link between Immigration and Trade: Evidence from the UK. *Review of World Economics* **138**(1).

[91] Goldstein, J.S. (1992). A Conflict-Cooperation Scale for WEIS Events Data. *Journal of Conflict Resolution* **36**: 369-385.

[92] Gould, D.M. (1994). Immigrant Links to the Home Country: Empirical Implication for US Bilateral Trade Flows. *Review of Economics and Statistics* **68**(4): 811-848.

[93] Greif, A. (1993). Contract enforceability and economic institutions in early trade: the Maghribi traders' coalition. *American Economic Review* **83**(3): 525-548.

[94] Grossman G.M., and E. Helpman (2002). Integration Versus Outsourcing in Industry Equilibrium. *Quarterly Journal of Economics* **117**(1): 85-120.

[95] Guiso, L., P. Sapienza, and L. Zingales (2009). Cultural Biases in Economic Exchange. Forthcoming in *Quarterly Journal of Economics* **124**(3).

[96] Gwartney, J., R. Lawson, R.S. Sobel, and P.T. Leeson (2007). *Economic Freedom of the World 2007 Annual Report*. Fraser Institute, Vancouver.

[97] Hakansson, H. (1982). International Marketing and Purchasing of Industrial Goods. IMP Project Group, Wiley and Sons.

[98] Hanson, G., and C. Xiang (2004). The Home-Market Effect and Bilateral Trade Patterns. *American Economic Review* **94**(4): 1108-1129.

[99] Harrigan, J., and A.J. Venables (2006). Timeliness and Agglomeration. *Journal of Urban Economics* 59(2): 300-316.

[100] Head, K., and J. Ries (1998). Immigration and Trade Creation: Econometric Evidence from Canada. *Canadian Journal of Economics* **31**(1): 47-62.

[101] Helpman, E. (2006). Trade, FDI, and the Organization of Firms. *Journal of Economic Literature* **44**: 589-630.

[102] Helpman, E., M.J. Melitz, and Y. Rubinstein (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. *The Quarterly Journal of Economics* **123**(2): 441-487.

[103] Helpman, E., M.J. Melitz, and S.R. Yeaple (2004). Export Versus FDI with Heterogeneous Firms. *American Economic Review* **94**(1): 300-316.

[104] Henderson, D.J., and D.L. Millimet (2008). Is Gravity Linear? *Journal of Applied Econometrics* **23**(2): 137-172.

[105] Hopenhayn, H. (1992). Entry, Exit, and Firm Dynamics in Long Run Equilibrium. *Econometrica* **60**(5): 1127-1150.

[106] Ilzkovitz, F., A. Dierx, V. Kovacs, and N. Sousa (2007). Steps Towards a Deeper Economic Integration: The Internal Market in the 21st Century. European Economy. *Economic Papers* **271**, European Commission.

[107] Jørgensen, J.G., and P.J.H. Schröder (2006). Tariffs and Firm-Level Heterogeneous Fixed Export Costs. *Contributions to Economic Analysis & Policy* **5**(1): 1495-1495.

[108] Jung, B. (2009). On the Importance of Adjustment Dynamics For Bilateral Trade Flows. Mimeo: University of Hohenheim.

[109] Kee, H.L., A. Nicita, and M. Olarreaga (2006). Estimating Trade Restrictiveness Indices. *World Bank Policy Research Working Paper* 3840.

[110] King, G., and W. Lowe (2003). An Automated Information Extraction Tool for International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design. *International Organization* **57**: 617-642.

[111] Klein P., and G. Ventura (2007). TFP Differences and the Aggregate Effects of Labor Mobility in the Long Run. *The B.E. Journal of Macroeconomics* **7**(1)(Contributions): Article 10.

[112] Kleinert, J., and F. Toubal (2005). A Structural Model of Export versus Affiliates Production. *Tübingen Discussion Paper* 288.

[113] Kleinert, J., and F. Toubal (2006). Dissecting FDI. Mimeo: *Hitotsubashi University Tokyo*.

[114] Krautheim, S. (2007a). Gravity and Information: Heterogeneous Firms, Exporter Networks and the 'Distance Puzzle'. Mimeo: European University Institute.

[115] Krautheim, S. (2007b). Export-Supporting FDI. *EUI Working Papers* ECO 2007/24.

[116] Krugman, P.R. (1979). Increasing Returns, Monopolistic Competition, and International Trade. *Journal of International Economics* **9**(4): 469-479.

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

[118] Kugler, M., and H. Rapoport (2007). International Labor and Capital Flows: Complements or Substitutes? *Economic Letters* **94**: 155-162.

[119] Liu, X. (2007). GATT/WTO Promotes Trade Strongly: Sample Selection and Model Specification. Mimeo: Kennesaw State University.

[120] Lucas, R.E.B. (2006). Migration and Economic Development in Africa: A Review of Evidence. *Journal of African Economies* **15**: 337-395.

[121] Martin, P., T. Mayer, and M. Thoenig (2008). Make Trade Not War? *Review of Economic Studies* **75**(3): 865-900.

[122] Martinez-Zarzoso, I., F. Nowak-Lehmann D., S. Klasen, and M. Larch (2009). Does German Development Aid Promote German Exports? Forthcoming in *German Economic Review*.

[123] Maskus K.E., J.S. Wilson, and T. Otsuki (2000). Quantifying the Impact of Technical Barriers to Trade. A Framework for Analysis. *World Bank Policy Research Working Paper* 2512.

[124] McCallum, J. (1995). National Borders Matter: Canada-U.S. Regional Trade Patterns. *The American Economic Review* **85**(3): 615-623.

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

[126] Melitz, M.J., and G.I.P. Ottaviano (2008). Market Size, Trade, and Productivity. *Review of Economic Studies* **75**(1): 295-316.

[127] Micco, A., E. Stein, and G. Ordonez (2003). The Currency Union Effect on Trade: Early Evidence from EMU. *Economic Policy* **18**(37): 315-356.

[128] Moser, C., T. Nestmann, and M. Wedow (2008). Political Risk and Export Promotion: Evidence from Germany. *The World Economy* **31**(6): 781-803.

[129] Moses, J., and B. Letnes (2004). The Economic Costs to International Labor Restrictions: Revisiting the Empirical Discussion. *World Development* **32**(10): 1609-1626.

[130] Neary, P. (2008). Trade Costs and Foreign Direct Investment. *International Review of Economics and Finance*. DOI:10.1016/j.iref.2008.06.004.

[131] Nickel, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica* **49**(6): 1417-1426.

[132] Nitsch, V. (2005). State Visits and International Trade. *CESIFO Working Paper* 1582.

[133] Parsons, C.R., R. Skeldon, T.L. Walmsley, and L.A. Winters (2007). Quantifying International Migration: A Database of Bilateral Migrant Stocks. *World Bank Policy Research Working Paper* 4165.

[134] Peng, M.W., and A.Y. Ilinitch (1998). Export Intermediary Firms: A Note on Export Development Research. *Journal of International Business Studies* **29**(3): 609-620.

[135] Peng, M.W., and A. Y. York (2001). Behind Intermediary Performance in Export Trade: Transactions, Agents, and Resources. *Journal of International Business Studies* **32**(2): 327-346.

[136] Peng, M.W., Y. Zhou, and A. York (2006). Behind the Make or Buy Decisions in Export Strategy: A Replication with Extension of Trabold. *Journal of World Business* 41: 289-300.

[137] Pissarides, C. A. (2000). *Equilibrium Unemployment Theory*. Cambridge, MA.: MIT Press.

[138] Polachek, S.W. (1980). Conflict and Trade. An Economics Approach to Political International Interactions. *Journal of Conflict Resolution* **24**: 55-78.

[139] Pollins, B.M. (1989). Conflict, Cooperation, and Commerce: The Effects of International Political Interactions on Bilateral Trade Flows. *American Journal of Political Science* **33**(3): 737-761.

[140] Popper, S.W., V. Greenfield, K. Crane, and R. Malik (2004). Measuring Economic Effects of Technical Barriers to Trade on U.S. Exporters. National Institute of Standards and Technology.

[141] Raff, H., and Y.-H. Kim (2005). Optimal export policy in the presence of informational barriers to entry and imperfect competition. *Journal of International Economics* **49**: 99-123.

[142] Raff, H., and N. Schmitt (2005). Endogenous vertical restraints in international trade. *European Economic Review* **49**: 1877-1889.

[143] Raff, H., and N. Schmitt (2006). Exclusive Dealing and Common Agency in International Markets. *Journal of International Economics* **68**: 485-503.

[144] Rauch, J.E. (1999). Networks Versus Markets in International Trade. *Journal of International Economics* **48**: 7-35.

[145] Rauch, J.E., and V. Trindade (2002). Ethnic Chinese Networks in International Trade. *Review of Economics and Statistics* **84**(1): 116-130.

[146] Rauch, J.E., and Watson, J. (2003). Starting Small in an Unfamiliar Environment. *International Journal of Industrial Organization* **21**(7): 1021-1042.

[147] Reed, R.R., and Trask, K.A. (2006). Decentralized International Exchange. *Canadian Journal of Economics* **39**(2): 516-543.

[148] Reuveny, R. (2001). Bilateral Import, Export, and Conflict/Cooperation Simultaneity. *International Studies Quarterly* **45**: 131-158.

[149] Rodríguez, F. and D. Rodrik (2000). Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence. In: B.S. Bernanke and K. Rogoff (eds.). *NBER Macroeconomics Annual 2000*. MIT Press.

[150] Roodman, D.M. (2009). A Note on the Theme of Too Many Instruments. *Oxford Bulletin of Economics and Statistics* **71**(1): 135-158.

[151] Rose, A. (2004). Do We Really Know That the WTO Increases Trade? *American Economic Review* **94**(1): 98-114.

[152] Rose, A. (2007). The Foreign Service and Foreign Trade: Embassies as Export Promotion. *World Economy* **30**(1): 22-38.

[153] Santos Silva, J.M.C., and S. Tenreyro (2006). The Log of Gravity. *Review of Economics and Statistics* **88**(4): 641-658.

[154] Schröder, P., J.H. Harold, H. Trabold, and P. Trübswetter (2005). Intermediation in Foreign Trade: When do Exporters Rely on Intermediaries? *Applied Economics Quarterly* **51**: 2005.

[155] Spulber, D. (1996). Market Microstructure and Intermediation. *Journal of Economic Perspectives* 10: 135-152.

[156] Spulber, D. (1999). Market Micro-structure: Intermediaries and the Theory of the Firm. Cambridge, UK: Cambridge University Press.

[157] Staiger, D., and J.H. Stock (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica* **65**(3): 557-86.

[158] Subramanian, A., and S.-J. Wei (2007). The WTO Promotes Trade, Strongly but Unevenly. *Journal of International Economics* **72**: 151-175.

[159] Trabold, H. (2002). Export Intermediation. An Empirical Test of Peng and Ilinitch. *Journal of International Business Studies* **33**(2): 327-344.

[160] Trefler, D. (1995). The Case of the Missing Trade and Other Mysteries. *American Economic Review* **85**(5): 1029-1046.

[161] Turnbull, P.W. and M.T. Cunningham (1981). International Marketing and Purchasing. Palgrave Macmillan.

[162] Wagner, D., K. Head, and J. Ries (2002). Immigration and the Trade of Provinces. *Scottish Journal of Political Economy* **49**(5): 507-525.

[163] Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

[164] WTO (2007). *WTO World Trade Report*. Geneva.