Trade Persistence and the Limits of Trade Agreements

Florian Mölders

German Institute for Economic Research (DIW), Mohrenstrasse 58, 10117 Berlin, Germany

Abstract

The trade creating effect of trade agreements reveals a high level of heterogeneity. This paper is concerned with network structures as drivers of this difference in the magnitude. Linking export networks with per-capita income levels, it is claimed that more developed countries are associated with smaller increases in bilateral trade flows if an agreement is present, compared to lower-income countries. This gives rise to a re-assessment of trade agreements and hence of economic policy. While they are a powerful tool for trading partners at the lower end of the per-capita income distribution, they are less so at the upper end.

Key words: Trade agreements, Gravity model, Trade persistence

1. Introduction

The proliferation of free trade agreements (FTAs) over the past decades has been documented quite explicitly (see Frankel (1997) and Bhagwati (2008)). They seem to attain more popularity the less multilateral trade liberalization is perceived to proceed. Countries sign FTAs primarily to give their domestic firms better access to the foreign country's market and therefore to boost bilateral trade volumes. Other benefits include the consolidation of democracy (Liu and Ornleas (2011)) or the promotion of peaceful relation between countries (Martin et al. (2010)). Most ex-post studies on aggregate trade flows assert positive trade effects due to trade agreements, even though the respective magnitudes may differ considerably. This can be derived from Frankel and Wei (1998) who estimate cross-sectional trade elasticities with respect to the European Community (EC), Mercosur and ASEAN. Their estimates suggest that stronger intra-regional trade creating effects are visible for the latter two agreements whereas the effects for the former reveal a negative sign. In a similar vein, Ghosh and Yamarik (2004) present comparative bounds estimates for a similar group of economies. Their results support those by Frankel and Wei (1998) in the way that intra-regional trade is positively affected by Mercosur and ASEAN, whereas NAFTA and the EU reveal negative signs. In contrast to these findings,

Email addresses: `fmoelders@diw.de` (Florian Mölders)

This version: March 2012

Carrère (2006) reports positive effects regarding intra-regional trade for the EU and NAFTA.

In disregard of the trading partners involved, it is difficult to draw meaningful conclusions on whether trade agreements boost intra-regional trade flows. Hence it is of interest whether we could filter out a pattern that points towards the magnitude of the trade-creating effects for a certain group of country-pairs. The literature on the empirical analysis of international trade flows has largely neglected heterogeneous effects of policy instruments.[1] Countries differ in a variety of aspects, from per-capita income levels to level of democratization to geographic characteristics, such as landlockedness. All of these factors are potentially responsible for differences in trade elasticities with respect to a change in variable trade costs. As outlined in section two, Helpman et al. (2008) show that the extensive margin of trade plays a decisive role. Other explanations involve e.g. the composition of exports at the industry level (Orefice and Rocha (2011)).

I describe a theoretical motivation for the source of the heterogeneous effect by assuming endogenous market penetration costs, as introduced by Arkolakis (2010). The implications of his model, which is based at the firm level, are then taken over to the country level. The analysis at the aggregate level is then motivated by Hausmann et al. (2007) who investigate the determinants of a country's export productivity level. They find a strong correlation with the respective level of economic development, approximated with the per-capita income level.

This paper makes use of a panel dataset that allows for the elimination of unobserved time-invariant effects.[2] These effects may well refer to a country pair's tendency to trade more than others. For descriptive purposes, the standard gravity model is first analyzed with a quantile regression that allows identification of potential heterogeneous effects along the trade volume levels. I then proceed with an analysis along the per-capita income levels, motivated by the results of Arkolakis (2010) and Hausmann et al. (2007). The results reveal significant differences for the trade-creating effect of FTAs for trading partners with low compared to high income levels. Larger elasticities are prominent for lower per-capita income quartiles.

The rest of this paper is structured as follows: Chapter two reviews the literature in this field of research. Chapter three describes a theoretical framework in which heterogeneity of trade creating effects of trade agreements may result.

[1] In its most recent *World Trade Report* (WTO (2011): 142), the WTO puts special emphasis on the role of preferential trade agreements in the multilateral trading system. When it comes to the level of economic integration with respect to technical barriers to trade (TBT), the report finds that countries that are similar in their level of economic development exhibit the highest level of integration (*North-North, South-South*).

[2] In a previous version of this paper, I have included information on regressions that made use of the publicly available dataset by Glick and Rose (2002). The results do neither contradict those I get with my data, note do they add to the quality of this paper. The respective results may be available from the author upon request.

Chapter four introduces the econometric model and data in order to analyze the heterogeneity. Chapter five describes the results and chapter six concludes.

2. Related literature

Helpman et al. (2008) provide a gravity model framework that is based on the micro-level of the Melitz (2003) model but can be empirically implemented with country-level data. In order to incorporate the extensive margin of trade, the authors propose a two-stage estimation that builds on the Heckman selection approach. After dividing their dataset according to the income levels of the respective country-pairs, their findings suggest that the elasticity with respect to a change in bilateral distance is highest for trading partners that are referred to as “South-South”, meaning that they are situated at the lower end of economic development. The authors base their result on the effect of the extensive margin of trade, as the effect on the intensive margin is fixed across all country pairs. Their result partly confirms the theoretical predictions outlined in section three, as well as the empirical analysis in sections four and five. One of the differences in the approach by Helpman et al. (2008) from the one presented here is that I rely on interaction terms to uncover potential heterogeneity in the RTA effect and apply dynamic estimation of the gravity model in order to incorporate the unobserved bilateral trade structure.[3]

Eicher and Henn (2011) also look at potential heterogeneous effects of preferential trade agreements (PTAs). Their findings on PTAs suggest that there is substantial heterogeneity in the effect of membership in a trade initiative: developing economies benefit with a trade increase of 214 percent while the industrialized economies’ equivalent points towards a 16 percent increase.[4] Orefice and Rocha (2011) take a look at the industry level of countries’ exports and estimate the effect of the depth of integration on various sectors. Given the depth of a trade agreement (defined by the respective coverage) the authors find that the automobile industry reflects trade elasticities that are up to four times higher than in the textile industry. It is claimed that the level of standardization lowers the impact for the latter industry while the prevalence of large production networks promotes the former. In an analysis on the determinants of the trade-creating effect of trade agreements, Vicard (2011) introduces various interaction

[3] In a previous version of this paper, I have made use of sub-samples to estimated the respective RTA effect on the trade volume. In order to make use of the full sample, I opt for the approach with interaction terms. Nevertheless, the results have been similar to those presented here.

[4] The authors do not find any trade effect coming from membership in the WTO once they control for multilateral resistance, unobserved bilateral heterogeneity and PTAs: “It may well be the case that developing countries reoriented their import activity considerably towards PTA partners after joining PTAs. This reorientation might produce trade creation, but it might also include some trade diversion that redirected trade from WTO trade partners to fellow PTA members. If this is a common pattern among developing countries, such a reorientation would have a negative impact on the WTO estimate for developing countries [...].”

terms in a gravity equation. Similar to the empirical approach in this paper, the author also incudes the level of economic development, approximated by the average and the difference of/in the per-capita income level. The findings do not support the hypothesis of any heterogeneity in the trade elasticity but suggest that North/North, North/South and South/South agreements have similar effects on the bilateral trade volume. Nevertheless, comparisons to the results in section five should be drawn with caution as the methodology and the data coverage differ. Vicard (2011) e.g. makes use of a panel dataset that only includes ten periods from 1955 to 2000.

In sectors that are dominated by trade in homogeneous goods, the average productivity level is lower than in sectors that exhibit a high level of heterogeneity (see Chaney (2008)). Furthermore, the level of economic development (in terms of per-capita GDP) has proven to be positively dependent on the average productivity level of a country's export basket (Rodrik (2006), Hausmann et al. (2007), Del Gatto et al. (2006)).[5] Following the results of Chaney (2008), a decrease in trade barriers will alter both the intensive and extensive margins of trade. The former will increase as existing exporters are faced with lower trading costs. At the same time, more firms enter the foreign market which have previously been excluded from entering due to the prohibitive trade barrier. At the aggregate level, the elasticity of a country's exports with respect to a change in variable trade costs (e.g. via trade liberalization) is determined by the sum of the extensive and intensive margins. His findings suggest that in sectors, which reveal a high level of homogeneity, the elasticity is expected to be higher.

Another strand of the literature on a potential heterogeneity of the effect of trade-enhancing initiatives is derived from the discussion about the effectiveness of membership in the WTO. Starting with the seminal paper by Rose (2004), there has been a vivid discussion on whether membership in the WTO is beneficial in terms of its trade-creating effects. While Rose neglects heterogeneous effects of GATT/WTO membership on the trade volume, Subramanian and Wei (2007) differentiate between the effect on developing and industrialized countries' trade. While they find only little influence of membership on the imports of developing countries, developed economies reveal a substantial increase in imports. Similar results are found in Chang and Lee (2011) who use matching to derive the heterogeneity in the GATT/WTO trade effect. Both studies have in common that their results point towards a strong trade effect for industrialized (developed) economies while the respective effect for low-income (developing) countries is significantly lower, if existent.

The following section motivates a potential source of the heterogeneous trade-creating effect for changes in the variable trade costs. The section argues that when countries implement trade agreements, trade costs decrease and

[5] Rodrik (2006) points out that the "productivity level of the export basket" of a country is strongly correlated with the per-capita income. Fajgelbaum et al. (2011) derive a demand based explanation of trade patterns by incorporating income heterogeneity in their trade model which leads the richer countries to export relatively more higher-quality products.

cause heterogeneous effects which can be traced back to the firm-level and investments in exporting networks.

3. Theoretical motivation

This section highlights network structures among trading partners as potential drivers of heterogeneous trade creating effects of FTAs. Endogenous network structures and productivity differences are introduced in the model by Arkolakis (2010).[6] He assumes away the fixed market-entry costs that are prevalent e.g. in Melitz (2003), but introduces endogenous market entry (or penetration) costs that reflect the ability of a firm to reach a certain consumer base. In the model, it is essential that the marginal costs of reaching a new consumer in a foreign country are dependent on the existent consumer base of the firm. The firm experiences diminishing returns to scale with respect to the investment in market penetration. In the following, I will summarize the model proposed by Arkolakis (2010) in order to derive implications for the aggregate country level analysis in the proceeding chapter.

The network

The number of potential consumers in the foreign country j is denoted by L_j. In order to reach consumers, the firm in country i invests in a network of size S. The probability of reaching a certain consumer after the firm has invested in the size S of the network is $n(S)$. The key ingredient of the approach lies in the diminishing returns to scale of the investment: doubling the investment in the network size does not lead to an equivalent growth of the potential consumer base. This feature is captured by the probability that a consumer is reached by the network for the first time: $(1 - n(S))^\beta$, with β in $[0, +\infty)$. To illustrate the diminishing returns of the distribution network consider the following example: In order to serve the foreign market, the firm needs to establish a network. Consumers need to be reached by the network before the firm is able to sell. Reaching the first consumer in a city for the first time by setting up a distribution network can be considered easier than reaching consumer 1,000 without simultaneously reaching consumer one again. This diminishing ("wasteful") effect of investment in the distribution network is reflected by the parameter β and is central to the hypothesis on heterogeneous responses to a change in variable trade costs. The higher β, the higher the diminishing returns to scale of a Euro spent on the network.

The consumer

The utility representation of each consumer is fairly standard: in the foreign economy, the consumer maximizes utility according to standard Dixit-Stiglitz preferences, common in a monopolistic competition setting. Utility is derived

[6]Other contributions to the literature that introduce networks in exporting activities include e.g. Krautheim (2012) or Freund (2000)

from consumption of a bundle of goods, each denoted by c_{ijt} for the consumption of a good from country i in country j at time period t. Accordingly, we can write:

$$U_{jt} = \left(\int_0^N c_{ijt}(\phi)^\rho d\phi \right)^{\frac{1}{\rho}}, 0 < \rho < 1, \tag{1}$$

where N denotes the number of varieties that are available to the consumer, $c(\phi)$ the quantity consumed from each of the varieties and ρ refers to the substitutability of the varieties and translates into the elasticity of substitution $\sigma = \frac{1}{1-\rho}$. Solving the maximization problem of the consumer with respect to her available income y_{jt} yields the demand (consumption) for a variety:[7]

$$c_{ijt}(\phi) = \frac{p_{ijt}(\phi)^{-\sigma}}{P_{jt}^{1-\sigma}} y_{jt} \tag{2}$$

Combined with the total number of consumers reached with the distribution network, the total demand that a firm with productivity ϕ from country i faces in country j at time t amounts to:

$$q_{ijt}(\phi) = \underbrace{n_{ijt}(\phi)L_{jt}}_{\text{Number of consumers reached}} \times \underbrace{\frac{p_{ijt}(\phi)^{-\sigma}}{P_{jt}^{1-\sigma}} y_{jt}}_{\text{Demand per consumer}} \tag{3}$$

The firm

In accordance with the literature on firm heterogeneity, firms draw their productivities (ϕ) out of a Pareto distribution such that the firm problem reduces to a maximization problem with respect to the size of the network and the price of the product. We can combine the previous equations into a profit function that consists of three segments ((ϕ) is omitted in the following equation):

[7] P_{jt} denotes the price aggregator prevalent in country j. It is the weighted average price index that consists of prices from all importers to j. I refer to appendix B.1 for more details.

$$\pi_{ijt}(p,n,\phi) = \underbrace{n_{ijt}L_{jt}y_{jt}\frac{p_{ijt}^{1-\sigma}}{P_{jt}^{1-\sigma}}}_{\text{Revenue from sales}} - \underbrace{n_{ijt}L_{jt}y_{jt}\frac{p_{ijt}^{-\sigma}\tau_{ijt}\omega_{it}}{P_{jt}^{1-\sigma}\phi}}_{\text{Production costs}} - \underbrace{\omega_{jt}\frac{L_{jt}}{\psi}\frac{1-(1-(n_{jt}))^{1-\beta}}{1-\beta}}_{\text{Network costs}} \tag{4}$$

Where $(\tau_{ijt}\omega_{it})/\phi$ denotes the per-unit costs of the firm with productivity ϕ. Optimization of (4) with respect to the price of the firm $p_{ijt}(\phi)$ yields: $p_{ijt}(\phi) = (\tilde{\sigma}\tau_{ijt}\omega_{it})/\phi$.[8] The last term in (4) refers to the costs that originate from the distribution network in country j. Revenue from sales are denoted by $r_{ijt} = p_{ijt}q_{ijt}$

Optimization of equation (4) with respect to the size of the network in period t results in the following:[9]

$$n_{ijt}(\phi) = 1 - \left[\frac{y_{jt}\phi^{\sigma-1}(\tilde{\sigma}\tau_{ijt}\omega_{it})^{1-\sigma}\psi P_{jt}^{\sigma-1}}{\omega_{jt}\sigma}\right]^{-1/\beta} \tag{5}$$

Trade agreements lower trade barriers, e.g. tariffs, and thus lead to lower per-unit costs of exporting. Now, what happens when trade barriers are removed? Following Arkolakis (2010), a decrease in trade costs lowers the threshold productivity below which no firm finds it profitable to export: $\partial\bar{\phi}/\partial\tau_{ijt} > 0$. Firms that have only exported small amounts before now export more due to lower trade costs. This intensive margin can be separated into two distinct margins: the first refers to the increase in sales as the prices in the foreign market decrease due to lower transport costs. The second is related to the simultaneous increase in the distribution network that needs to be established in order to export. As an additional investment in the network exhibits decreasing returns to scale, ex-ante smaller exporters can benefit relatively more from a decrease in trade costs and this effect is thus more prominent for firms at the lower end of the productivity distribution.

Aggregation

The aggregation of the firm-level perspective to the country-level is a matter of country-specific productivities. I therefore relate the theoretical results that were based at the firm level to country-level trade flows. Most importantly, a country's individual cumulative density function of the Pareto distribution

[8] With $\tilde{\sigma} = (\sigma)/(1-\sigma)$.

[9] See appendix B.2 for a derivation of equation (5).

has to be introduced: $G(\phi) = 1 - \phi^{-k}$.[10] Keeping in mind that exporters with low levels of productivity exhibit the largest increase in export flows due to a decrease in variable trade cost, we can derive implications for bilateral trade flows between two countries, based on their respective cumulative density functions.

The productivity *cdf* drives the number of firms in country i that are active in exporting to country j. Given this distribution, total exports from i to j can be written as follows:

$$T_{ijt} = N_{Eit} \int_{\bar{\phi}}^{\infty} r_{ijt}(\phi) dG(\phi), \tag{6}$$

$$T_{ijt} = N_{jit} \int_{\bar{\phi}}^{\infty} r_{ijt}(\phi) d \frac{G(\phi)}{1 - G(\bar{\phi})}, \tag{7}$$

where N_{Eit} reflects the total number of firms active in country i and N_{ijt} the number of firms serving the market j from i. This is due to $N_{ijt} = N_{Eit}(1 - G(\bar{\phi}))$. Additionally, from equation (6) we can derive that the total number of firms from country i that export to country j is equal to $N_{Eit} dG(\phi)$.

In order to get to the effect of a decrease in trade barriers on the bilateral trade volume, equation (7) needs to be differentiated with respect to τ. Following Chaney (2008) and Arkolakis (2010) this is most intuitively conducted by splitting the equation into its components: the intensive and extensive margin of bilateral trade. The intensive margin refers to changes in the volume of consumption by a consumer that is already served. Furthermore, we can control for the effect of a change in variable trade costs on the size of the distribution network. The extensive margin refers to the part of the aggregate trade volume that is sourced from any additional firm that is incited to enter the foreign market due to a decrease in trade barriers. The respective aggregate elasticity gives:

$$\frac{\partial ln T_{ijt}}{\partial ln \tau_{ijt}} = -k \tag{8}$$

This result can be found in Chaney (2008) and Arkolakis (2010). But note that the interpretation in both papers differs: While Chaney stresses the importance of the extensive margin as the driver of the difference in the response to a trade liberalization, Arkolakis refers to the *new consumers margin*. What does this tell us about a potential heterogeneous effect of trade liberalization with respect to the characteristics of a country's firms? The elasticity is higher the

[10] k denotes the distribution parameter which governs the shape of the Pareto distribution together with the minimum value.

larger the shape parameter of the Pareto distribution. This parameter governs the level of heterogeneity of firms across an industry/country. A high value of k denotes a high level of homogeneity and thus a higher density of low-productive firms.

I rely on a crucial assumption about the average productivity level of a country's exports, in order to simplify the result on bilateral trade flows in equations (7) and (8): I abide by the empirical evidence of Hausmann et al. (2007) who relate the productivity level of a country's export basket to its per-capita income levels and find a strong evidence for a positive correlation.[11] Note that there are expected to be potential outliers whenever a country has a low per-capita income level but its average productivity level of its exports is high. This may result from a single or only a few commodities that are exported and demand high productivity levels but play only a minor role in explaining the per-capita income of the economy. In a similar vein, Del Gatto et al. (2006) estimate the average productivity levels or European countries and show that they are closely related to the prevalent per-capita income levels in the respective countries.

Hausmann et al. (2007) calculate an index (EXPY) that refers to the embedded productivity of a country's export basket. This index relates the share of any sector in total exports to a, previously determined, productivity level of the respective sector (PRODY). PRODY is derived from the per-capita income levels of all countries which export in a sector (l):

$$PRODY_l = \sum_i \frac{(x_{il}/X_i)}{\sum(x_{il}/X_i)} Y_i \tag{9}$$

For exporting firms, EXPY reflects an approximation of the productivity level as it refers to the average productivity of exporting firms in country i. Countries that are subject to a higher prevalence of more homogeneous firms (higher k in equation (8)) reveal a distribution function $G(\phi)$ that generates exporters which are more concentrated towards the threshold level $\bar{\phi}$ (see Corcos et al. (2012)). In these countries, a higher proportion of firms that export are considered small exporters which exhibit a higher trade elasticity with respect to a change in the variable trade cost, e.g. via a trade agreement. This is a crucial assumption as it abstracts from the firm-level heterogeneity and follows a representative firm perspective for country i.

If the average productivity level of exporting firms is highly correlated with a country's per-capita income level, a change in the level of variable trade costs (e.g. in the case of FTAs) should have a non-homogeneous effect on the aggregate bilateral trade volume along the respective level of economic development. This conclusion is drawn from the Pareto shape parameter k, which refers to a country's level of heterogeneity, and hence to a country's prevalent average level

[11]The authors find a correlation of about 0.80-0.83.

of productivity.

To satisfy the predictions of a gravity model, bilateral exports should be decomposable in the following way:

$$ln(T_{ijt}) = \gamma ln\tau_{ijt} + A_{it} + B_{jt} + v_{ijt} \tag{10}$$

A_{it} and B_{jt} denote country level fixed effects and v_{ijt} reflects the error term. The country fixed effects are included to control for any unobservable influence that is persistent and influences a country's export activity. Differentiating (10) with respect to the trade costs τ_{ijt} yields the elasticity γ which, according to the argumentation outlined above, should depend on the level of heterogeneity of a country's exporting firms, $\gamma(k)$, approximated by the average productivity of the exporting firms. In the following sections, I will use the level of economic development of a country to control for the average productivity level (following Hausmann et al. (2007)), both are highly correlated.[12]

The following hypothesis summarizes the theoretical motivation for the firm and country level perspective:

H 1. *Trade agreements have a stronger trade-creating effect for country-pairs at the lower end of economic development (in terms of per-capita GDP).*

Dynamics

If the networks (which are a prerequisite for exporting activities in the model outlined in this section) reveal persistence in the sense that an investment in period t will prevail in $t+1$, a dynamic structure of the model and adjustment of the econometric approach are potential alternatives. If e.g. a fraction of the network ($0 < \varphi < 1$) is available in the next period, a high level of trade persistence may be expected, as investment in the exporting network is only necessary in the amount of $1 - \varphi$.[13]

Eichengreen and Irwin (1998) have stressed that bilateral trade evolves in a dynamic, rather than static, manner. The authors refer to regional trade agreements (RTAs) that reveal a higher than expected trade-creating effect. This persistence in trading relationships over time may be interpreted as the result of historical ties between any two economies. If not controlled for this factor, the respective coefficient of the RTA variable will overestimate the real effect of the agreement as it incorporates the trade persistence. Eichengreen and Irwin (1998) propose the inclusion of the lagged value of the dependent variable in order to control for the previous period's level of trade. As expected, the resulting coefficient of the lagged value in the authors' estimation is highly significant and lowers the magnitude of other explanatory variables' coefficients considerably.

[12]Del Gatto et al. (2006) derive that (while excluding Germany in the calculation) the correlation between the per-capita income and the productivity levels amounts to 0.88

[13]Costs that represent maintenance of distribution chains or a reorientation of the marketing towards new customers are intuitive examples that justify φ to be smaller than unity.

The persistence in bilateral trade flows may originate from various sources, such as common cultural ties or network structures. Inclusion of the lagged trade flow in a regression analysis takes care of any unobservable factor of current bilateral trade that was present in the foregoing period already. Furthermore, the coefficients of other explanatory variables necessarily decrease if exports in $t-1$ e.g. explain 90 percent of exports in t. Nevertheless, we may expect that a fraction of the unobservable factors are represented by the network structures and may potentially lower the heterogeneity of a trade-creating effect. Having a large network in period $t-1$ potentially causes large trade flows in t:

$$T_{ijt} = f(\underbrace{n_{ijt-1}}_{+}), \tag{11}$$

where n_{ijt-1} denotes the network size of the foregoing period. In order to control for the unobserved and time-variant "network effects", the most applicable way would then be to make use of the lagged dependent variable as an instrument. T_{ijt-1} would then be highly correlated with the unobserved n_{ijt-1}[14] but by construction not with the current trade flow T_{ijt}. This follows from the assumption that todays' bilateral trade is altered by trade in period $t-1$ by nothing else but n_{ijt-1}.[15]

At this point it is worth looking back at what has been claimed so far and what this implies for the empirical analysis. Firms have to establish a distribution network in order to sell their product to the consumer. This network investment exhibits decreasing returns to scale which govern the main implication of the model. It leads smaller exporters to gain substantially more from a decrease in trade costs as they can more easily reach new consumers due to low per-consumer network costs. This leverage effect is more prominent for low productivities, in contrast to firms that trade at a high level (high productivities) that have to invest relatively more in order to sell an additional unit to the foreign country. At the aggregate level, this translates into a potential heterogeneous response to trade liberalization, depending on the country's density of homogeneous firms. I will expect higher elasticities for trading partners with high levels of homogeneity. This hypothesis will be tested in the following sections.

[14] Note that this correlation is expected to be higher for trading partners at the upper levels of the trading volume distribution.

[15] Nevertheless, the unobserved and time-invariant part of the error term is common to both T_{ijt} and T_{ijt-1} which leads the lagged variable to be correlated with the error term, causing the dynamic panel bias (Nickell (1981)). In the empirical part of this paper, I will further approach this issue.

4. Econometric Method and Data

In this section, I first outline the econometric approach and then describe the dataset that is employed. As mentioned before, I focus on potential heterogeneity in gravity estimates of the variable denoting trade agreements. To do so, I need a tool that reflects this approach. Quantile regression has the advantage of pointing towards heterogeneous effects along the distribution of the dependent variable.[16] It can be used for various specifications of the regression equation as long as the dependent variable is continuous. In my case, the specification is described by the gravity equation of international trade flows. The dependent variable denotes the bilateral trade flow from country i to j in period t. In a first step, the evolution of the effect of the treatment variable (RTA) for the upper and lower quantiles is of special interest. In a second step, I analyze the behavior of this treatment variable with respect to per-capita income levels of the trading-partners, as an approximation for the average prevalent productivity levels of the trading partners. I investigate whether trade-creating effects of trade initiatives reveal any heterogeneity across the level of economic development, as was suggested in the previous section.

4.1. Gravity

The empirical literature on international trade flows is dominated by its workhorse: the gravity model. First introduced by Tinbergen (1962), the gravity model was rapidly recognized as an excellent tool for the analysis of bilateral trade flows. Even though the model reflected an excellent fit to the data, a theoretical underpinning was needed. The contributions of Anderson (1979), Bergstrand (1985) and Anderson and Wincoop (2003) have added substantially to the economic foundations of the model. Its high explanatory power and sound theoretical underpinnings make it a useful tool for uncovering bilateral trade flows' determinants. While the basic economic relationships are easily described by the model (positive influence of countries' GDPs and a negative influence of bilateral distance), the challenge over the past decade was to establish an econometric approach that would match the increasing use of panel data. The combination of time-series and cross-sectional data allows to draw considerably more information out of the data than is the case in cross-sectional estimates. Examples that stress the use of the associated fixed- and random-effects models are Baier and Bergstrand (2002), Baier and Bergstrand (2007) and Magee (2008).[17] The use of fixed-effects eliminates all unobserved, time-invariant effects that may alter a country-pairs' bilateral trade. This potential source of endogeneity has been stressed by Baier and Bergstrand (2007) who advocate

[16] See Angrist and Pischke (2008) for a comprehensive summary of quantile regression.

[17] Magee makes use of a Poisson estimator in a gravity model. This method (introduced by Santos Silva and Tenreyro (2006)) has two advantages: First, it allows to include zero-trade values and hence observations that have previously either been dropped from the data or transformed. Second, Poisson is more robust in the case of heteroskedastic error terms.

the fixed-effects model when estimating the effects of trade agreements. One considerable disadvantage of this approach is that only time-invariant factors are controlled for. Unobserved time-variant factors that simultaneously affect bilateral trade as well as the establishment of trade agreements will remain in the error term and may cause an endogeneity bias.

The equation that is estimated as a benchmark takes the following form:

$$\begin{aligned} X_{ijt} = \beta_0 + \beta_1 ln(D_{ij}) + \beta_2 Border_{ij} + \beta_3 Language_{ij} \\ +\beta_4 ln(GDPAB_{ijt}) + \beta_5 ln(PCGDPAB_{ijt}) \\ +\beta_6 RTA_{ijt} + d_t + A_i + B_j + v_{ijt}, \end{aligned} \quad (12)$$

where X denotes exports from country i to j in period t. The explanatory variables on the right hand side are fairly standard in a gravity setting.[18] My focus lies on the behavior of the coefficient β_6, because it denotes the increase in trade flows due to the existence of a trade agreement between the trading partners. Additionally, I introduce interaction terms in section five. As I work in a panel environment, the error term might consist of more than just the idiosyncratic part and carries country-pair specific information that influences both the dependent as well as some of the independent variables. This sort of endogeneity can be controlled for using fixed effects that eliminate any time-invariant effect in the regression equation.

A variety of econometric specifications are available when analyzing the gravity model. The panel structure of the data demands for a proper treatment of the idiosyncratic and time-invariant, country-pair specific part via fixed effects estimation if a specification test (Hausman test) rejects the random effects model. Furthermore, I estimate the model using Poisson regressions, as proposed by Santos Silva and Tenreyro (2006). The advantage here lies in the treatment of the error term that might exhibit heteroskedasticity. In this case, Poisson will be favored as it does not rely on the homoskedasticity assumption of the error term and is more robust in the presence of heteroskedasticity.[19] Even if the proportionality condition ($E(Y|X) \propto Var(Y|X)$) does not hold strictly, Poisson is more adequate than models that rely on homoskedasticity assumption (OLS). In addition, I am able to include zero observations for bilateral trade flows with Poisson. This importance of this property increases with the information involved in the zero-observations. The fact that some countries do not engage in bilateral trade carries valuable information on the determinants of the dependent variable. The share of these zero-observations in the data is substantial, as

[18] D_{ij} denotes bilateral distance, $Border_{ij}$ and $Language_{ij}$ indicate whether the country-pair shares a common border and language, $GDPAB_{ijt}$ refers to the product of the country-pair's GDPs, $PCGDPAB_{ijt}$ denotes per-capita income (as product of the country-pair's per-capita GDP) and RTA_{ijt} denotes common membership in a trade agreement. A_i and B_j refer to the country fixed effect of country i and j respectively, d_t denotes time dummies and v_{ijt} the error term that includes both the time-invariant and the idiosyncratic part.

[19] See Winkelmann (2008) for further justification of the Poisson model in non-count models.

the time-span dates back to 1950. Out of all the available entries in the dataset (773,214), 48.8 percent are zero.[20] In subsequent econometric specifications (System GMM estimations), zero entries are dropped from the estimation as no logarithm of zero exists.

In order to control for potential dynamics in the gravity model, I incorporate the lagged value of the dependent variable in the regression equation. Here, I follow the theoretical motivation of chapter three that mentioned a dynamic adjustment in order to incorporate a potential persistence in bilateral trade flows. After this modification, the estimated equation then looks as follows:

$$X_{ijt} = \beta_0 + \beta_1 lnX_{ijt-1} + \beta_{(-)} I_{ijt} \tag{13}$$

$\beta_{(-)}$ and I_{ijt} denote the remaining coefficients and variables on the right hand side of (11).

In order to uncover potential heterogeneous trade-creating effects, I separate the dataset into four quantiles according to the distribution of the per-capita income variables.[21] This method allows us to select the sample conditional on per-capita income quartiles and to introduce interaction terms that simultaneously label the country-pair's income quartile and common membership in a trade agreement. The data have been classified into quartiles (or income groups) such that the evolution over time (countries become richer) does not bias the classification. Country-pairs that have e.g. been in the first quartile with their per-capita income of the first year of the time-span may well end up in the fourth in the latter years.[22]

Dynamics in the gravity model can be modeled via the Arellano and Bond (1991) or the Blundell and Bond (1998) (System GMM) estimators. Both incorporate the lagged trade value on their respective right hand side. Computational complexity is introduced in the specification as the lagged trade value, X_{ijt-1}, is correlated with the error term in period t by definition. Furthermore, this dynamic panel bias is weakened for panels with a "large" time-dimension.[23] As mentioned in the previous chapter, I separate the dataset in four quarters according to their rank in the per-capita income distribution and then compute interaction terms that reflect a country-pair's income quartile in the full sample

[20] This share is subject to substantial changes over time. While in 1950 there are almost 50 percent of zeros (out of all observable export flows), this share decreases as global trade flows increase. By 2006, 33.3 percent of all observed trading relations did report non-zero trade flows.

[21] I further classify the countries into "poor" and "rich", according to their per-capita income levels.

[22] It should be considered that this method splits up otherwise strung together time-series. This has to be kept in mind when it comes to the econometric specification, especially whether to neglect or take into account a potential dynamic panel bias, as the time dimension of a series in considerably shortened.

[23] See Alvarez and Arellano (2003) for an analysis of the asymptotic properties of panel estimators.

in each year as well as common membership in a trade agreement. Additionally, I classify countries into "poor" and "rich", by dividing the sample by per-capita income in any period. The sub-sampling of the data according to an exogenous variable is not subject to the sample-selection bias (Heckman (1979)). In the next section, I perform various regressions with different specifications and in a first step include the lagged value of the dependent variable to control for the stated dynamics in the model. I then proceed by estimating a corresponding model using the System GMM estimator (proposed by Blundell and Bond (1998)).

4.2. Data

I make use of aggregate trade data as the research question points towards an assessment of trade liberalization measures that are not aimed at any particular sector. I further analyze panel data that convey information about both the cross-sectional and time-series type to include the time-dynamics of bilateral trade flows. The data covers the period from 1950-2006.

The trade data is sourced from IMF's *Direction of Trade Statistics*, and it refers to unidirectional exports. The dependent variable is denominated in current US Dollars in my dataset. The values for the respective economies' GDPs are in constant US Dollars, drawn from the *Penn World Tables.* CEPII data was the source for the geographic characteristics (such as common border and language, as well as the distance between countries). Taking into account the unbalanced panel structure in the dataset, this gives a total of 773,214 observations for the bilateral trade flow. The high number of observations results from a long time span (1950-2006) and the fact that unidirectional exports are used. I make use of information on regional trade agreements from Head et al. (2010) who source their information mainly from WTO data.

5. Results

In this section, I build my empirical analysis on the argumentation outlined in section three. I show that a closer investigation of the data with respect to the level of economic development reveals a high level of heterogeneity regarding common membership in a trade agreement. Following the theoretical considerations on productivity differences and export networks, I analyze the potential heterogeneity along the per-capita income distributions. I incorporate the dynamic structure as well as panel specific estimation methods (fixed effects) in the regressions to control for unobservable time-invariant factors.[24]

[24] As noted before, I have estimated the same equations with the data by Glick and Rose (2002). I do not include them in the analysis as the results do not add to the quality of the paper.

5.1. Heterogeneity across the trade volume

For descriptive reasons, the heterogeneity of the elasticity with respect to a trade agreement is first described by a quantile regression. Instead of solving for the conditional expectation function (as would be the case in a standard OLS estimation), the minimization problem of the conditional quantile function is solved according to the following function and yields the coefficients for the respective quantiles τ:[25]

$$\beta_\tau = arg_{(b)}min \sum \rho_\tau(Y_i - X_i(b)), \tag{14}$$

where β solves for the minimum value of the quantiles. ρ_τ denotes the "check"- function that represents an asymmetric loss function and (compared to the loss in squared errors for the minimization of least squares) gives smaller weight to outliers.[26]

Special focus lies on the estimates of the RTA variable. The coefficients of this variable along the trade volume distribution are displayed in Figure 1. It reflects the behavior of the coefficient along the trade volume distribution, as well as the OSL estimate and the respective confidence interval. The coefficients reveal a high level of significance throughout the quantile estimates. Other gravity variables such as the countries' GDPs and distance mirror a fading magnitude in their effect on the bilateral trade volume but remain significant at the 1%-level throughout all quantiles. RTAs have a higher elasticity for trading partners at the lower end of the trade volume distribution. For the results of five quantiles (0.1, 0.25, 0.5, 0.75 and 0.9) I refer to the table in the appendix.

Most interestingly, distance and RTA refer to variable trade costs. These costs seem to have a fading effect the higher the bilateral trade volume is. It should be noted that the quantile regression approach does not directly refer to the theoretical considerations outlined in section three. There, heterogeneous effects with respect to a change in the variable trade costs were motivated by referring to the average productivity level of a country's exports. The quantile regression does not pick up on this derivation but focuses on the dependent variable, bilateral exports. Nevertheless, from the regression that is plotted in Figure 1, we observe a monotonically decreasing effect of liberalization measures. Countries which do not trade intensively, compared to others, can expect to gain relatively more from a common trade agreement.

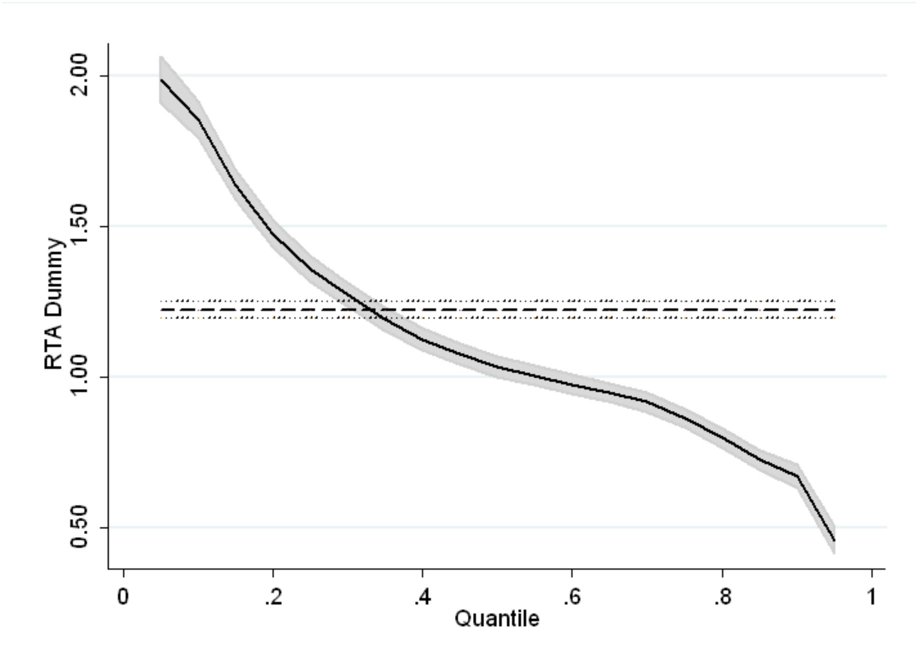
5.2. Heterogeneity across the per-capita income

In the following, I classify the data according to an exogenous variable in the gravity estimates: the per-capita income level. I am interested in whether the potential trade-creating effect of a trade agreement differs for less developed economies compared to developed ones. In a first step, this is conducted by

[25] Following Koenker and Hallock (2001).

[26] The function is defined as $\rho_\tau(u) = \tau * |u|$ if $u > 0$ and $\rho_\tau u = (1 - \tau) * |u|$ if $u \leq 0$.

Figure 1: Quantile regression, RTA effect



Note: Explanatory variables according to table 1 included.

including an interaction term that consists of the RTA and the product of the countries' per-capita income levels.[27] In a second step, I divide the interaction terms along the classifications *poor-poor*, *poor-rich*, *rich-poor* and *rich-rich*. The respective thresholds for the countries are calculated as being above or below the per-capita income median in any respective year. In the second approach, the income group, in which an exporting country is classified, is observed. This is more in line with the theoretical considerations outlined in section three where the potential heterogeneity was derived from the average productivity level of the exporting country.

Additionally, I investigate whether the dynamics capture a part of the heterogeneity in the per-capita context.[28] The results of the regressions are described in Tables 1 and 2. As a benchmark I employ a Poisson pseudo-maximum likelihood estimator (PPML) (column (1)) as proposed by Santos Silva and Tenreyro (2006). This method neglects the panel structure of the data and therefore does not include fixed effects. A fixed effects Poisson estimation is applied in column

[27]I have estimated the same regressions using the sum of the per-capita income levels instead. This method has lead to similar results in the majority of regressions.

[28]I expect the per-capita income to be exogenous in the outcome equation in order to conform to exogenous sample selection.

(2) before I proceed with System GMM estimates in column (3).

The results in Table 1 strongly suggest that the trade creating effect of a common trade agreement is unevenly spread across the per-capita income levels of the trading partners. The coefficient of 1.706 signals a highly significant trade creating effect of RTAs. The interaction terms reveal more detailed information on whether this effect is homogeneous along all country pairs. Compared to countries for which the common per-capita income levels are situated at the bottom of the distribution, the higher income countries reflect a trade-creating effect of only 30.21 percent,[29] which is more than 400 percent lower than for the former country group. All interaction terms are highly significant and the respective magnitudes are decreasing in the income levels. The hypothesis that the interaction terms do not reflect heterogeneity in the trade creating effect was rejected.[30]

Column two differs with respect to the estimation technique as it includes country-pair as well as importer and exporter fixed effects. Qualitatively, the results are similar to those obtained in column one. We observe a highly significant trade creating effect of a common RTA (trade increases by 78.43 percent) and interaction terms that are either insignificant or negative. The hypothesis (*H1*) is again rejected. The System GMM estimates involve a more careful treatment of the panel dynamics and take into account the potential endogeneity of the lagged trade flow. Adding dynamics to the estimation leads to lower coefficients as a significant amount of the explanatory power is incorporated in the lagged trade flow. A one percentage change in the export volume of the foregoing period changes the current trade flow by 0.8 percent. Nevertheless, the RTA effect remains positive, yet less significant. An interesting finding is that by controlling for dynamics in the regression analysis, the differences in the trade creating effects become insignificant. This result supports the assumption that the lagged trade flow may absorb part of the effect that leads to the heterogeneous outcome.[31]

Table 2 plots the respective results for the estimation that makes use of different interaction terms, namely the income classification of the exporter and importer. At first sight, the PPML estimates are a mirror-image of those presented in Table 1. We observe a highly significant trade increase for members of a trade agreement, but less so for countries which are both classified to be high income. For "poor" exporters, the largest trade increase is observable if they export to a country that is situated on the same side of the income-level.

[29] Calculated as $e^{1.706-1.442} - 1$.

[30] *H1* denotes the respective test which tests for the hypothesis that all interaction terms are equal to zero. Confirmation of the hypothesis would signal that the difference with respect to the control group (lowest per-capita income quartile) is statistically insignificant.

[31] For the dynamic panel regressions I have to assign (exogenous) instruments as well as potential endogenous and predetermined variables. Strictly exogenous variables: Distance, Border, Language, Year-dummies. Potentially predetermined/endogenous variables: Lagged Trade, RTA, Quart2RTA–Quart4RTA. Lags of order 3–6 were used as instruments. A Sargan test refers to the validity (exogeneity) of the instruments.

Table 1: Differences in response to an RTA? Interaction terms (income quartiles)

Trade	PPML	Poisson FE	System GMM
Lag Trade			0.805***
			(0.020)
Distance	-0.466***		-0.385***
	(0.015)		(0.035)
Border	0.601***		-0.270**
	(0.037)		(0.128)
Language	0.442***		0.497***
	(0.024)		(0.050)
GDPs	0.792***	0.281***	0.464***
	(0.005)	(0.096)	(0.046)
PCGDPs	0.332***	0.848***	-0.034
	(0.008)	(0.118)	(0.101)
RTA	1.706***	0.579***	0.548**
	(0.065)	(0.089)	(0.247)
Quart2RTA	-0.611***	0.111	-0.482**
	(0.059)	(0.078)	(0.196)
Quart3RTA	-0.974***	-0.212**	-0.545*
	(0.066)	(0.123)	(0.280)
Quart4RTA	-1.442***	-0.205	-0.777**
	(0.063)	(0.145)	(0.391)
N	773,214	700,806	360,794
IM/EX FEs	No	Yes	No
Year FEs	Yes	Yes	Yes
H1	0.000	0.002	0.104
Sargan- test	—	—	0.905

Notes: * p<.1, ** p<.05, *** p<.01. *H2* refers to the hypothesis: no difference in coeff. of Quart2RTA-Quart4RTA. Quart(2-4)RTA denotes interaction terms of RTA with income quartile. PPML refers to Poisson pseudo-maximum-likelihood estimation. Standard errors have been bootstrapped. IM/EX FEs refer to Importer and Exporter, Year FEs to Year fixed effects and Poisson FE includes country-pair fixed effects

Table 2: Differences in response to an RTA? Interaction terms (Rich, Poor)

Trade	PPML	Poisson FE	System GMM
Lag Trade			0.805***
			(0.020)
Distance	-0.463***		-0.382***
	(0.015)		(0.034)
Border	0.616***		-0.256**
	(0.036)		(0.125)
Language	0.439***		0.503***
	(0.025)		(0.049)
GDPs	0.791***	0.294***	0.462***
	(0.005)	(0.095)	(0.044)
PCGDPs	0.328***	0.832***	-0.033
	(0.008)	(0.106)	(0.092)
RTA	1.299***	0.601***	0.439**
	(0.057)	(0.056)	(0.194)
RTA*PR	-0.670***	-0.269**	-0.610***
	(0.059)	(0.122)	(0.217)
RTA*RP	-0.593***	-0.483**	-0.441**
	(0.072)	(0.081)	(0.208)
RTA*RR	-1.032***	-0.209**	-0.604*
	(0.052)	(0.099)	(0.319)
N	773,214	700,806	360,794
IM/EX FEs	No	Yes	No
Year FEs	Yes	Yes	Yes
H1	0.000	0.000	0.004
Sargan- test	—	—	0.951

Notes: * p<.1, ** p<.05, *** p<.01. *H2* refers to the hypothesis: no difference in coeff. of RTA*PR-RTA*RR. RTA*PR etc. denotes interaction terms of RTA with income classification. PPML refers to Poisson pseudo-maximum-likelihood estimation. Standard errors have been bootstrapped. IM/EX FEs refer to Importer and Exporter, Year FEs to Year fixed effects and Poisson FE includes country-pair fixed effects

The addition of various fixed effects does not alter these results, and similar to column one, the hypothesis of a homogeneous trade creating effect of RTAs is rejected. According to the results in Table 2, the lowest trade-creating effect of RTAs is observed for countries from the upper half of the per-capita income distribution that export to partners which are considered less developed. In this respect, the elasticity reveals an increase of only 12.52 percent. System GMM again lowers the magnitudes of most of the coefficients but adds relatively little to the analysis in this setting. It is noticeable that homogeneity of the trade effect is now rejected, contrary to the result in Table 1. This is, at least partly, due to the different design of the interaction terms.

Differences in the trade-creating effect that were observed in the estimates can also be interpreted against the background that the predominant trade initiatives may be more effective for trade with homogeneous goods and hence for those trading partners at the lower end of the per-capita income distribution. The magnitude of the trade creation by a liberalization measure should then be mirrored by a respective analysis on the composition of the trade basket along the per-capita income distribution.

The results may also be interpreted against the background of a higher sensitivity to trade barriers in vertical production networks (see Yi (2003)). Countries with lower per-capita income levels are often involved in production networks that lead to more intense cross-border trade due to the multiple production processes. Further, the results confirm the estimates of Helpman et al. (2008) who also differentiate their dataset according to the income-per capita levels, but focus on the heterogeneity with respect a change in bilateral distances. They classify the country-pairs into three sub-categories: trade among South-South, trade among North-South and trade among North-North countries. Felbermayr and Jung (2011) make the point that countries for which the Pareto shape parameter (k) is higher, reflecting a high mass of less productive firms, there is substantial potential for reallocation towards more productive exporters. Trade liberalization would in this respect leave a bigger opportunity for increases in the exporting activity.

6. Conclusion

Against the background of the theoretical motivation for distribution networks, this paper has claimed that the driver of the heterogeneity in the trade creating effect of trade agreements may be investments in the trading relationship between country-pairs. Particular focus was put on the role of the level of economic development as a link to the average productivity level of a country. The results presented here are largely in line with a number of studies (e.g. Helpman et al. (2008) and Eicher and Henn (2011)).

The use of a large panel dataset has uncovered heterogeneity in gravity estimates, in particular for the variable denoting the trade-creating effect of RTAs. I confirm what has previously been stated in the literature: an overestimation of the trade-creating effect of trade agreements. Additionally, I find that the overestimation effect is increasing from lower to upper quantiles of the trade distribution. Moreover, trade initiatives reveal stronger trade effects for country pairs at the lower end of the income distribution (*H1*). While trade agreements are a powerful tool for trading partners along all levels of economic development, stronger effects are observed for those at the lower end of the per-capita income distribution. The addition of a dynamic structure in the empirical analysis partly absorbs the heterogeneous effect of change in the variable trade costs.

These results may also be founded on determinants other than distribution networks. One might think about the design of a trade agreement which certainly adjusts to the trading partners involved. If trade negotiations only focus on homogeneous goods or a few sectors, the impact of the implementation may be stronger as the agreement is better tailored towards the export structures of the trading partners. This would imply that the complexity of successfully implementing a trade agreement rises with the level of heterogeneity in the respective negotiating economies. Analyzing the validity of this argument may be the subject of further research.

References

Alvarez, J., Arellano, M., 2003. The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators. Econometrica 71, 1121–1159.

Anderson, J., 1979. A Theoretical Foundation for the Gravity Equation. American Economic Review 69, 106–116.

Anderson, J., Wincoop, E.v., 2003. Gravity with Gravitas: A Solution to the Border Puzzle. American Economic Review 93, 170–192.

Angrist, J., Pischke, J.S., 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Arellano, M., Bond, S., 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.

Arkolakis, C., 2010. Market Penetration Costs and the New Consumers Margin in International Trade. The Journal of Political Economy 118, 1151–1199.

Baier, S., Bergstrand, J., 2002. On the Endogeneity of International Trade Flows and Free Trade Agreements. Mimeo.

Baier, S., Bergstrand, J., 2007. Do free trade agreements actually increase members' international trade? Journal of International Economics 71, 72–95.

Bergstrand, J., 1985. The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. Review of Economics and Statistics 67, 474–481.

Bhagwati, J., 2008. Termites in the Trading System - How Preferential Agreements Undermine Free Trade. Oxford University Press.

Blundell, R., Bond, S., 1998. Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics 87, 115–143.

Carrère, C., 2006. Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model. European Economic Review 50, 223–247.

Chaney, T., 2008. Distorted Gravity: The Intensive and Extensive Margins of International Trade. American Economic Review 98, 1707–1721.

Chang, P.L., Lee, M.J., 2011. The WTO Trade Effect. Journal of International Economics 85, 1.

Corcos, G., Del Gatto, M., Mion, G., Ottaviano, G., 2012. Productivity and Firm Selection: Quantifying the "nee" Gains from Trade. forthcoming in The Economic Journal .

Del Gatto, M., Mion, G., Ottaviano, G., 2006. Trade Integration, Firm Selection and the Costs of Non-Europe. CEPR Working Paper 5730.

Eichengreen, B., Irwin, D.A., 1998. The Role of History in Bilateral Trade Flows, in: Frankel, J. (Ed.), The Regionalization of the World Economy. University of Chicago Press, pp. 33 – 62.

Eicher, T.S., Henn, C., 2011. In search of WTO trade effects: Preferential trade agreements promote trade strongly, but unevenly. Journal of International Economics 83, 137–153.

Fajgelbaum, P., Grossman, G., Helpman, E., 2011. Income Distribution, Product Quality, and International Trade. Journal of Political Economy 119, 721–765.

Felbermayr, G., Jung, B., 2011. Sorting It Out: Technical Barriers to Trade and Industry Productivity. Open Economies Review 22, 93–117.

Frankel, J., 1997. Regional Trading Blocs in the World Economic System. Institute for International Economics.

Frankel, J., Wei, S.J., 1998. Regionalization of World Trade and Currencies: Economics and Politics, in: Frankel, J. (Ed.), The Regionalization of the World Economy. University of Chicago Press, pp. 189 – 226.

Freund, C., 2000. Different Paths to Free Trade: The Gains from Regionalism. Quarterly Journal of Economics 115, 1317–1341.

Ghosh, S., Yamarik, S., 2004. Are regional trading arrangements trade creating?: An application of extreme bounds analysis. Journal of International Economics 63, 369–395.

Glick, R., Rose, A., 2002. Does a currency union affect trade? The time-series evidence. European Economic Review 46, 1125–1151.

Hausmann, R., Hwang, J., Rodrik, D., 2007. What you export matters. Journal of Economic Growth 12, 1–25.

Head, K., Mayer, T., Ries, J., 2010. The erosion of colonial trade linkages after independence. Journal of International Economics 81, 1–14.

Heckman, J.J., 1979. Sample Selection Bias as a Specification Error. Econometrica 47, 153–161.

Helpman, E., Melitz, M., Rubinstein, Y., 2008. Estimating Trade Flows: Trading Partners and Trading Volumes. The Quarterly Journal of Economics 123, 441–487.

Koenker, R., Hallock, K.F., 2001. Quantile Regression. Journal of Economic Perspectives 15, 143–156.

Krautheim, S., 2012. Heterogeneous firms, exporter networks and the effect of distance on international trade. forthcoming in Journal of International Economics .

Liu, X., Ornleas, E., 2011. Free Trade Agreements and the Consolidation of Democracy. MPRA Working Paper 24014. Mimeo.

Magee, C., 2008. New measures of trade creation and trade diversion. Journal of International Economics 75.

Martin, P., Mayer, T., Thoenig, M., 2010. The Geography of Conflicts and Free Trade Agreements. CEPR Working Paper 7740.

Melitz, M., 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica 71, 1695–1725.

Nickell, S., 1981. Biases in Dynamic Models with Fixed Effects. Econometrica 49, 1417–1426.

Orefice, G., Rocha, N., 2011. Deep integration and production networks: an empirical analysis. Staff Working Paper ERSD-2011-11.

Rodrik, D., 2006. What's So Special about China's Exports? China & World Economy 14.

Rose, A.K., 2004. Do We Really Know That the WTO Increases Trade? American Economic Review 94.

Santos Silva, J.M.C., Tenreyro, S., 2006. The Log of Gravity. The Review of Economics and Statistics 88.

Subramanian, A., Wei, S.J., 2007. The WTO promotes trade, strongly but unevenly. Journal of International Economics 72.

Tinbergen, J., 1962. Shaping the World Economy: Suggestions for an International Economic Policy. Twentieth Century Fund, New York.

Vicard, V., 2011. Determinants of successful regional trade agreements. Economics Letters 111.

Winkelmann, R., 2008. Econometric analysis of count data. Springer.

WTO, 2011. World Trade Report 2011 – The WTO and preferential trade agreements: From co-existence to coherence. World Trade Organization, Geneva.

Yi, K.M., 2003. Can Vertical Specialization Explain the Growth of World Trade? Journal of Political Economy 111.

A. Appendix

Table 3: Quantile regression estimates

Trade	Q 0.1	Q 0.25	Q 0.50	Q 0.75	Q 0.90
Distance	-0.974***	-0.927***	-0.811***	-0.728***	-0.661***
	(0.009)	(0.006)	(0.005)	(0.005)	(0.006)
Border	0.588***	0.565***	0.696***	0.635***	0.744***
	(0.038)	(0.028)	(0.021)	(0.020)	(0.026)
Language	0.715***	0.788***	0.828***	0.894***	0.927***
	(0.016)	(0.012)	(0.009)	(0.009)	(0.011)
GDP	0.892***	0.920***	0.871***	0.792***	0.716***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
PCGDP	0.278***	0.266***	0.230***	0.186***	0.169***
	(0.004)	(0.003)	(0.002)	(0.002)	(0.003)
RTA	1.855***	1.361***	1.036***	0.864***	0.672***
	(0.030)	(0.022)	(0.016)	(0.016)	(0.020)
constant	-28.78***	-29.10***	-26.78***	-22.99***	-19.75***
	(0.153)	(0.115)	(0.092)	(0.091)	(0.122)
N	410,728	410,728	410,728	410,728	410,728

Notes: * p<.1, ** p<.05, *** p<.01
Controlled for year effects.

Table 4: Variable definitions

Variable	Definition	Source
Exports (X_{ijt})	Exports from (i) to (j)	IMF, Direction of Trade Statistics
Distance (D_{ij})	Great circle distance between (i) and (j)	CEPII
$Border_{ij}$	Dummy variable indicating a common border between (i) and (j)	CEPII
$Language_{ij}$	Dummy variable indicating a common language for (i) and (j)	CEPII
$GDPAB_{ijt}$	Product of (i) and (j) gross domestic products, in constant US-Dollars	Penn World Tables
$PCGDP_{ijt}$	Product of (i) and (j) per-capita GDPs, in constant US-Dollars	Penn World Tables, IMF World Economic Outlook
RTA_{ijt}	Dummy variable indicating a common trade agreement between (i) and (j)	Head et al. (2010)
$Quart2RTA_{ijt}$	Interaction term denoting that (i) and (j) both belong to the second per-capita income quartile and share a common trade agreement	Penn World Tables, IMF World Economic Outlook, Head et al. (2010)
$Quart3RTA_{ijt}$	Interaction term denoting that (i) and (j) both belong to the third per-capita income quartile and share a common trade agreement	Penn World Tables, IMF World Economic Outlook, Head et al. (2010)
$Quart4RTA_{ijt}$	Interaction term denoting that (i) and (j) both belong to the fourth per-capita income quartile and share a common trade agreement	Penn World Tables, IMF World Economic Outlook, Head et al. (2010)

Table 5: Countries in Sample

Afghanistan	Denmark	Laos	Senegal
Albania	Djibouti	Latvia	Seychelles
Algeria	Dominica	Lebanon	Sierra Leone
Angola	Dominican Rep.	Lesotho	Singapore
Antigua & Barbuda	Ecuador	Liberia	Slovakia
Argentina	Egypt	Libya	Slovenia
Armenia	El Salvador	Lithuania	Solomon Isl.
Australia	Equatorial Guinea	Luxembourg	Somalia
Austria	Eritrea	Macao	South Africa
Azerbaijan	Estonia	Macedonia	Spain
Bahamas	Ethiopia	Madagascar	Sri Lanka
Bahrain	Fiji	Malawi	St. Kitts & Nevis
Bangladesh	Finland	Malaysia	St. Lucia
Barbados	France	Maldives	St. Vincent
Belarus	Gabon	Mali	Sudan
Belgium	Gambia	Malta	Suriname
Belize	Georgia	Mauritania	Swaziland
Benin	Germany	Mauritius	Sweden
Bermuda	Ghana	Mexico	Switzerland
Bhutan	Grenada	Moldova	Syria
Bolivia	Guatemala	Mongolia	Tajikistan
Bosnia & Herzeg.	Guinea	Morocco	Tanzania
Botswana	Guinea-Bissau	Mozambique	Thailand
Brazil	Guyana	Namibia	Togo
Brunei	Haiti	Nepal	Tonga
Bulgaria	Honduras	New Zealand	Trinidad & Tobago
Burkina Faso	Hong Kong	Nicaragua	Tunisia
Burundi	Hungary	Niger	Turkey
Cambodia	Iceland	Nigeria	Turkmenistan
Cameroon	India	Norway	Uganda
Canada	Indonesia	Oman	Ukraine
Cape Verde	Iran	Pakistan	United Arab Em.
Central Afr. Rep.	Iraq	Palau	United Kingdom
Chad	Ireland	Panama	United States
Chile	Israel	Papua New Guinea	Uruguay
China	Italy	Paraguay	Uzbekistan
Colombia	Jamaica	Peru	Vanuatu
Comoros	Japan	Philippines	Venezuela
Congo, Rep.	Jordan	Poland	Vietnam
Costa Rica	Kazakhstan	Portugal	Yemen
Cote d'Ivoire	Kenya	Qatar	Zambia
Croatia	Kiribati	Russia	Zimbabwe
Cuba	Korea	Rwanda	
Cyprus	Kuwait	Sao T. & Princ.	
Czech Rep.	Kyrgyz Rep.	Saudi Arabia	

B. Appendix

B.1. Price aggregator

The price aggregator, introduced in equation (3) is a weighted average of the prices prevalent in country j. It depends on the number of importers (J_h) and their respective size of the distribution network:

$$P_{jt}^{1-\sigma} = \sum_{h=1}^{N} J_h \int_0^{\infty} p_{hjt}(1-\phi)^{1-\sigma} n_{hjt}(\phi) g_{ht} d\phi$$

B.2. Optimal size of the network

The third term in equation (5)

$$\pi_{ijt}(p,n,\phi) = \underbrace{n_{ijt} L_{jt} y_{jt} \frac{p_{ijt}^{1-\sigma}}{P_{jt}^{1-\sigma}}}_{\text{Revenue from sales}} - \underbrace{n_{ijt} L_{jt} y_{jt} \frac{p_{ijt}^{-\sigma} \tau_{ijt}\omega_{it}}{P_{jt}^{1-\sigma}\phi}}_{\text{Production costs}} \\ - \underbrace{\omega_{jt} \frac{L_{jt}}{\psi} \frac{1-(1-(n_{jt}^*))^{1-\beta}}{1-\beta}}_{\text{Network costs}},$$

describes the cost that are due to the investment in the distribution network. But what is the optimal size of the network that the firm choses? I can solve the equation with respect to n to get the expression in euqation (6):

$$\frac{\partial \pi_{ijt}}{\partial n_{jt}} = L_{jt} y_{jt} \frac{(\tilde{\sigma}\tau_{ijt}\omega_{it}/\phi)^{1-\sigma}}{P_{jt}^{1-\sigma}} - L_{jt} y_{jt} \frac{(\tilde{\sigma}\tau_{ijt}\omega_{it}/\phi)^{-\sigma}\tau_{ijt}\omega_{it}}{P_{jt}^{1-\sigma}\phi} = (1-\beta)\omega_{jt} \frac{L_{jt}}{\psi} \frac{(1-n_{jt})^{-\beta}}{1-\beta}$$

$$\Leftrightarrow \frac{y_{jt}}{P_{jt}^{1-\sigma}} (\tilde{\sigma}\tau_{ijt}\omega_{it}/\phi)^{1-\sigma} - \frac{y_{jt}}{P_{jt}^{1-\sigma}} (\tilde{\sigma}\tau_{ijt}\omega_{it}/\phi^{-\sigma}(\phi))^{-1}(\tau_{ijt}\omega_{it}) = (1-\beta)\omega_{jt} \frac{L_{jt}}{\psi} \frac{(1-n_{jt})^{-\beta}}{1-\beta}$$

$$\Leftrightarrow \frac{y_{jt}(\tilde{\sigma}\tau_{ijt}\omega_{it})^{1-\sigma} - y_{jt}(\tilde{\sigma}\tau_{ijt}\omega_{it})^{-\sigma}\tau_{ijt}\omega_{it}}{P_{jt}^{1-\sigma}\phi^{1-\sigma}} = (1-\beta)\omega_{jt} \frac{L_{jt}}{\psi} \frac{(1-n_{jt})^{-\beta}}{1-\beta}$$

$$\Leftrightarrow \frac{y_{jt}(\tau_{ijt}\omega_{it}\tilde{\sigma})^{1-\sigma}(1-\frac{1}{\tilde{\sigma}})}{P_{jt}^{1-\sigma}\phi^{1-\sigma}} = \frac{\omega_{it}(1-n_{jt})^{-\beta}}{1}$$

$$\Leftrightarrow \frac{y_{jt}\phi^{\sigma-1}(\tau_{ijt}\omega_{it}\tilde{\sigma})^{1-\sigma}\psi P_{jt}^{1-\sigma}}{\omega_{it}\sigma} = (1-n_{jt})^{-\beta}$$

$$\Leftrightarrow n_{ijt}^*(\phi) = 1 - \left[\frac{y_{jt}\phi^{\sigma-1}(\tilde{\sigma}\tau_{ijt}\omega_{it})^{1-\sigma}\psi P_{jt}^{\sigma-1}}{\omega_{jt}\sigma}\right]^{-1/\beta}$$