University of St.Gallen

Bachelor Thesis

Determinants of bilateral trade flows

Author:
Patrick Steiner
patrick.steiner2@student.unisg.ch
Matriculation number: 12-609-178
Bachelor in Economics (BVWL)

Referee:
Prof. Dr. Reto Föllmi
Department of Economics, SIAW-HSG

November 12, 2015

Abstract

This thesis aims to shed light on the determinants of bilateral trade flows across countries. Following the ideas of Linder, a demand-side approach is being taken next to the prevailing trade theories analysing trade from a supply-side perspective. Recent theoretical works predict patterns of trade in which southern countries export mostly to northern countries and northern countries to a broader set of countries. Patterns of trade will be empirically explored incorporating income differences across countries and their income distribution as determinants. Trade will be split into an intensive and an extensive margin where the focus will be on the latter. Single bilateral trade flows will be aggregated to sectors to tackle sector-specific characteristics. Not only the percentage of markets, but also the percentage of products traded bilaterally will be investigated using stylised facts and Bernoulli quasi-likelihood estimation. Variables for income differences and distributions will be GDP per capita and Gini coefficients respectively. The empirical analysis reveals that GDP p.c. has a significant impact on both the percentage of markets and the percentage of traded products. The effect of inequality is ambiguous with no clear-cut results. The sectoral analysis proofs to be crucial to analyse trade.

JEL: D31, D63, F10, F14, F17, O15

Keywords: international economics, trade flows, patterns of trade, income differences, inequality, income distribution, empirical analysis

Contents

List of Figures

List of Tables

List of abbreviations

AAP	Average adjusted prediction
APE	Average partial effect
BEC	Broad Economic Classification
DCG	Durable Consumer Goods according to BEC
GDP	Gross Domestic Product
GDP p.c.	Gross Domestic Product per Capita
GINI	Gini coefficient, a measure for inequality
H-O	Heckscher-Ohlin model
HS	Harmonized Commodity Description and Coding System
i	Exporting country/origin market
j	Importing country/destination market
mRATE	Number of markets in % from maximum
MTR	Multilateral trade resistance
nm	Number of markets
np	Number of product categories
OLS	Ordinary Least Squares
pRATE	Number of traded product categories in % from maximum
PW	Fractional response model used in Papke & Wooldridge (1996)
RTA	Regional Trade Agreement
WDI	World Development Indicators

1 Introduction

"What improves the circumstances of the greater part can never be regarded as an inconveniency to the whole. No society can surely be flourishing and happy, of which the far greater part of the members are poor and miserable."

- Adam Smith in The Wealth Of Nations (1776, p. 96)

Opinions on inequality differ widely and have been hotly debated since centuries. Not only since the Communist Manifesto written by Karl Marx and Friedrich Engels proclaimed the class struggle in 1848, but already 72 years earlier did Adam Smith in his landmark work "the Wealth of Nations" highlight the need for an equal society. While the upper-class indulges in an extravagant lifestyle, people on the other side of the income distribution struggle to get by financially. With the outbreak of the financial crisis, topics on inequality received more attention culminating in the famous Occupy Wall Street movement in the USA declaring: "We are the 99%!" The protestors gathered to fight against greed, corruption and social as well as economic inequality. While within-country inequality seems to be on the rise, as investigated by recent works including the bestseller "Capital in the 21st Century" written by Thomas Piketty, cross-country inequality takes on the opposite direction. A recent report from the Guardian estimates that the Gini index for global inequality as fallen from 72.2 in 1988 to 70.5 in 2008. Similarly, the share of the global population living in poverty has decreased from more than 50% in 1981 to 17% in 2011 (Roser, 2015).

Along with more integrated economies, economic growth and globalisation, income differences become more visible and tourism allows for first-hand experiences of life in third-world countries. What strikes the tourist's eye are western products sold even in the most remote places in the world. Also the sheer variety of goods available in rich countries is immense. Fresh fruits, cocoa, coffee, clothes and so on are mostly marked with exotic origins. This thesis is devoted to answer questions of what role income differences and the within-country distribution play in explaining these patterns of trade.

A brief example to illustrate the outset. While the Czech Republic (CZE) and Nigeria (NGA) are almost identical in size, both have a GDP of \$154 bn., CZE exports to around 3× as many destinations in the durable consumer good sector as NGA. The import markets are slightly more balanced when CZE imports from 84% of all possible origins and Nigeria from 50%. Also the percentage of traded product categories varies a lot. CZE exports in 83% product categories while Nigeria exports in only 25% in the durable consumer good sector. It can be easily inferred that the Czech Republic trades more extensively than Nigeria. Possible explanations can be found by either income differences, because the Czech Republic's GDP per capita is 13× higher than the one of Nigeria or the within-country income distribution.

The focus of this thesis will be on the extensive margin of trade which is basically a count of markets or products. Other margins such as the intensive margin of trade are also popular in research, but will not be part of this thesis. Cross-country income differences and the within-country income distribution will be linked with the extensive margin of trade. This will be done by a sectoral analysis aggregating bilateral product-category trade flows up to sectors, namely to the consumption, capital, intermediate and the durable consumer good sector.

The thesis starts with a literature review explaining developments of popular trade theories and how the ideas of the extensive margin of trade and income effects emerged. Worth mentioning are the works of Linder (1961), Trefler (1995), Mitra & Trindade (2005), Foellmi et al. (2013), and Hepenstrick & Tarasov (2014) in particular. Stylised facts will be presented for the percentage of trading partners as well as the percentage of traded product categories in order to show graphically the proposed relationship between income and the extensive margin of trade. In a third step, the actual empirical analysis is presented following the works of Papke & Wooldridge (1996) and Santos Silva et al. (2014) in estimating the extensive margins by a fractional response regression. The use of a Bernoulli quasi-likelihood estimator is proposed. Alternatives would include Poisson, Heckman or the method proposed by Helpman et al. (2008), although these methods do not account the doubly-bounded nature of our data. The data is cross-sectional leaving out the time dimension, which will therefore not be treated. Robustness checks will be presented by standard-OLS regression and the limitations of the analysis will be discussed afterwards.

2 Literature review

2.1 Standard trade models: A brief overview of trade theory

Until the late 18th Century, mercantilism was the prevailing economic philosophy of that time. Lofty tariffs on imports combined with a heavily subsidised export industry allowed to accumulate massive surpluses enriching the nation and demonstrating power. The crux is obvious. Adam Smith stated, that not all countries are able to run trade surpluses, as this would also mean that no one runs deficits. Smith puts labour productivity in the centre of attention by establishing the concept of absolute advantage. Trade is therein beneficial for countries specialising in production in which they have a superior productivity over the other and exporting the surplus in specialized goods (Das, 2008, p. 1-2). The Ricardian Theory proposes the model of comparative advantage highlighting the importance of technology differences and enforcing the paradigm of free trade. David Ricardos main contribution was the use of labour productivity ratios instead of absolute values. This allowed poorer countries to participate in the world trade system as their absolute productivity was inferior to the developed ones but in relative terms still allowed for specialisation (Ricardo, 1817; Das, 2008).

The Heckscher-Ohlin (H-O) or Factor Endowment Model, following the Ricardian tradition of com-

parative advantage, then removed technology differences and introduced variable capital to the model years later coming forward with a model stressing the factors of production. Four major theorems have arisen from the famous $2 \times 2 \times 2$ -model (2 goods, 2 countries and 2 factors of production).

First, the Heckscher-Ohlin Theorem states: A capital-abundant country will export mostly the capital-intensive good, while the labor-abundant country will export prevalently the labor-intensive good (Bowen et al., 1987, p. 791). The reason for this is the price advantage a country has for goods using the abundant factor intensively.

Second, the Stolper-Samuelson theorem postulates that "international trade necessarily lowers the real wage of the scarce factor expressed in terms of any good" (Stolper & Samuelson, 1941, p. 66). According to Stolper & Samuelson (1941, p. 73) this does not imply to impose protectionist measures as the benefits of free trade outweigh any gains of real wages through protecting the scarce factor rewards.

Third, factor price equalisation had already been noted by Bertil Ohlin in 1933, but Samuelson (1948, p. 163) criticises that an adequate proof is nowhere to be found. Samuelson (1948, p. 175) subsequently delivers the proof under certain assumptions, e.g. constant returns to scale, law of diminishing marginal utility, land & capital are qualitatively identical inputs, different factor intensity for the goods and all goods move perfectly freely in international trade. Under the conditions previously mentioned, factor prices tend to equalise along with commodity prices. Caveats of the model are being addressed in the last part of his landmark work as the strong conditions do seem to fail in reality. The fuzzy concept of production factors, whether or not knowledge is included, or zero trade frictions are just an example of highly controversial assumptions (pp.181-183).

The fourth theorem explains that at a constant rate of substitution in production, an increase of one factor must lead to an absolute increase in output of the good using that factor intensively as well as a reduction of supply of the other good (Rybczynski, 1955, p. 337-338).

2.2 Other branches of international trade theory and critics of neoclassical models

Although the H-O model dominates in its very essence international trade theory, widespread criticism questioned the workhorse model's real life applicability. The paradox described by Leontief (1953) concludes that although the United States are the most capital-abundant country in the world, mostly labour-intensive goods are being exported contradicting the implications of the H-O model. Possible explanations include that $capital \neq capital$ as human capital or a skilled workforce are not tracked in physical capital proposed by H-O. Also production technologies vary greatly. Although Leontief's paradox had already been declared as outdated by Stern & Maskus (1981, p. 223) for the year 1972, it was still valid in 1958. The composition of trade plays a crucial role as the decline of the relative importance of imports of natural resources contributed to the reversal of the paradox (p. 223). In general, paradoxical outcomes of empirical tests are common for the H-O model (Maskus, 1985).

Bowen et al. (1987, p.805) found, for the world as a whole, the correlation between factor endowment and trade patterns to be tenuous. It is concluded soberly that "the H-O model does poorly, but we do not have anything better". It was tested against models based on weaker assumptions, e.g. nonproportional consumption an technological differences, yet the alternatives yielded to economically unsatisfying parameters. Furthermore, H-O fails to explain the case of missing trade, which is a common phenomena in empirical analysis (Trefler, 1995). The Stolper-Samuelson theorem is even considered to be dangerous often misused as a justification for worldwide trade liberalisation. Theoretically, wages of unskilled labour should have been risen through trade liberalisation in South America as exports of labour-intensive goods surged, but in reality only the inequality-level changed not favouring the unskilled labour-class (Davis & Mishra, 2007).

New trade theories emerged. Krugman (1979) provides more realistic assumptions and focuses on economies of scale instead of technology differences or factor endowments as explanations for patterns of trade. Costless product differentiation and imperfect competition (monopolistic competition) complement the existing H-O model and allow for intra-industry[1] trade. The main insight is that countries export products in what they also have a strong home market for. It is also acknowledged that further extensions regarding demand patterns and zero trade flows are needed (Krugman, 1980, p. 958). Helpman (1981, p. 337) refines the model considering inter-industry trade as result of differences in capital-labour ratios while intra-industry trade takes place between countries with close factor-proportions giving rise to the hypothesis that a country's share of bilateral intra-industry trade is negatively correlated with the absolute difference in bilateral GDP p.c. Moreover, it is proposed that higher average per capita income represents a higher average capital-labour endowment ratio. On the assumption that industries that are capital intensive tend to have relatively more production of differentiated products, countries with higher average capital-labour ratios will experience a greater share of intra-industry specialisation (Helpman & Krugman (1985) in Bergstrand (1990, p. 1217)). This tells us that within north-north trade flows mainly differentiated products are being traded. Evenett & Keller (2002, pp. 310-311) highlight the large differences in factor endowments to be essential for perfect product specialisation in north-south trade flows. The H-O model does predict incomplete product specialisation allowing two-way trade in differentiated, similar goods only in a north-north setting.

The H-O model and the Krugman model are both subject to the assumption of identical technologies. That a firm's productivity vary greatly is self-evident in reality. Melitz (2003, p.1696) implements firm productivity heterogeneity into the Krugman model to adjust the model to recent empirical findings that more efficient firms self-select into export markets. It was shown that trade exposure forces less efficient firms out of the industry and that trade-induced reallocations towards more efficient firms

[1] Inter-industry trade occurs when a country exports in one set of industries and imports in another set of industries; intra-industry trade occurs when there is two-way exporting and importing within the same industry (Bernard et al., 2007, p. 107)

promotes overall productivity. Therewith, individual firm-level productivity will not necessarily improve. Market entry costs play a vital role favouring larger firms over smaller ones (p. 1718). Recent literature stressed the importance of analysing decomposed firm-level data. The highly disaggregated data up to average value per product per firm provides valuable insights into firm behaviours as export firms are fundamentally different from firms serving only the domestic market (Bernard et al., 2007, 2012). Only a few numbers of firms actually export. Larger firms serve more markets and across industries, larger markets are being served by more firms. In this respect, consumers profit from greater variety (Eaton et al., 2004, p. 154). So far, trade models provided possible explanations on absolute values of trade. A larger economy will inherently export a larger amount than a smaller economy. Little has been done to get a grasp on how trade flows are composed of. Although the Krugman model features a prominent extensive margin in differentiated goods, predictions are still ambiguous. Klenow & Hummels (2005, pp. 704-705) deliver a model focusing on the extensive margin including fixed entry costs, providing insights on why larger economies tend to export higher volumes of a given good (intensive margin) to a wider set of countries (extensive margin) and export higher-quality goods (see also Helpman et al. (2008) & Chaney (2008)). Based in this study, Arkolakis (2010) identifies these fixed entry costs as market penetration costs and attributes them a crucial function in forming export patterns. Most products are only exported to a few destinations where export zeros are strongly correlated with distance and importer size (Baldwin & Harrigan, 2011, p. 86).

All the literature presented up to now follows the original neoclassical model of the H-O model. These models follow a general supple-side perspective neglecting demand-side considerations partially or completely. As we will see, models can be interpreted in various ways and supply/demand-side considerations may overlap or complement each other. Two other streams of trade theory are the Linder-Hypothesis and the prominent gravity equation presented in the following paragraphs. The Linder hypothesis states that "the more similar the demand structures of two countries, the more intensive, potentially, is the trade between these two countries (Linder, 1961, p. 94)". The assumption of changing relative demand with income per capita was newly introduced to trade theory implying differing aggregate preferences across countries. So far, the H-O model stressed the importance of differences in production functions and factor endowments rather than centring the theory on demand-side considerations. Counter-intuitive results of the H-O model regarding intra-industry trade lead to Linder's conclusion that other forces might be at work, such as transport costs and economies of scale[2]. In his thinking, goods are bipolar switching from labour-intensive to capital-intensive and in reverse if necessary (p. 84-85). In other words, Capital is stored-up labor and labor is stored-up capital. Furthermore, internal demand, a product being consumed at home, is seen as a necessary condition for a product to become a potential export (Linder, 1961, p. 87). Therefore, the home bias later implemented in several other works along with other assumptions such as the aforementioned

[2]Intra-industry trade implies homogenous factor proportions predicting in theory that no trade would occur.

trade costs and economies of scale were already predicted by Linder, see also Krugman (1979, 1980); Helpman (1981). In this regard, Linder (1961) concludes that "international trade is really nothing but an extension across national frontiers of a country's own web of economic activity (p.88)"[3].
Also the phenomena of "zeros" are mentioned explained by fundamentally different demand patterns so that some countries do not even participate in world trade if there are no primary exports to trade (p. 137). Several overlaps exist between Linder's hypothesis and neoclassical theories and its successors, e.g. that the more capital per head in an economy, the higher GDP p.c. will be (p.101). The conclusion of Helpman (1981, p. 305) is that intra-industry trade depends on income per capita and inter-industry trade from differences in factor endowments, being related to Linder's insight as differentiated products are dominant in capital-abundant countries. Linder's signature is ubiquitous in trade theory although all the insights do not stem from supply-side considerations. Our analysis so far provides a rather blurred picture of current streams of trade theory and the next line manifested in the gravity equation does not relieve this pain. Both supply (H-O model) and demand (Linder) reasons simultaneously contribute the the understanding of patterns of trade. Several other propositions evolved and were tested in gravity equation-like models (Bergstrand, 1990, p. 1228).
First applied by Tinbergen (1962), the gravity equation states that the bigger the economy and the closer the countries (distance) the higher the trade volume (intensive margin). Bacchetta et al. (2012, p. 103) describes the equation by analogy with the Newtonian law of gravity: The more planets are attracted to each other in proportion to their sizes and proximity, the higher are trade volumes in proportion to GDP and distance. The gravity equation looks in its simplest form as the following (The second equation is the gravity equation taken in logs):

$$Trade_{ij} = Constant \times \left(\frac{GDP_i \times GDP_j}{Distance_{ij}} \right)$$

$$\ln(Trade_{ij}) = \beta_0 + \beta_1 \ln(GDP_i) + \beta_2 \ln(GDP_j) - \beta_3 \ln(Distance_{ij}) + \epsilon_{ij}$$

The gravity equation is very versatile in its applications and all major models (H-O model, Linder and Krugman) can be derived from it. Thinking of the gravity equation as a characterisation of all the major theoretical models is a rather futile endeavour as it is not suitable to test theories, but only useful to capture most of the variation in empirical works. That's why gravity is simply consistent with common theories rather than actually explaining patterns of trade Deardorff (1998). Haveman & Hummels (2004) sees the gravity equation as a statistical relationship that can be derived from every model under trade costs and incomplete specialisation. The "puzzle" of zero trade flows is consistent with incomplete specialisation and distance resembling trade frictions.

[3] A very nice example of the origins of trade is as follows: Production of a good is based on invention, we have an additional reason to believe that home market demand is necessary. An invention is, in itself, most likely to have been the outcome of an effort to solve some problem which has been acute in one's own environment. The exploitation of the invention will then, in its first phase, automatically be geared to the home market (Linder, 1961, p. 88)

2.3 Revitalization of demand-side approaches

Economists never really philandered with Linder's demand-side approach. The properties of homothetic preferences seemed to tempting to be abandoned (Krugman, 1980; Melitz, 2003; Chaney, 2008). Even though Helpman's (1981) empirical implications of the role of GDP p.c. overlapped in its predictions with Linder, assumptions on consumer behaviour and demand patterns who do empirically hold remained outnumbered. Homothetic preferences determine demand to be only dependent on relative prices ignoring the effect of income. Hunter (1991, p. 356) concludes that non-homothetic preferences significantly contribute to the volume of bilateral trade flows. Furthermore, supply-side induced trade models tend to over-predict inter-industry trade by over 25% with homothetic preferences. New models incorporating non-homothetic preferences emerged over the last decade taking income into account whether this is for demand reasons or production reasons (Bergstrand, 1990; Bond et al., 2010; Dinopoulos et al., 2011; Markusen, 2013). Newer studies explicitly link the size of the intensive, extensive and the quality margin of trade to GDP p.c. (Hallak, 2010; Hepenstrick & Tarasov, 2014; Markusen, 2013; Hepenstrick & Tarasov, 2014). In a Ricardian model, higher GDP p.c resembles to a higher technology-level which in turn promotes productivity fostering innovation and hence, more variety is being supplied. On the other hand, the effect is twofold. On the demand-side, higher productivity implies higher wages which models the consumer demand to be higher and therefore, more variety is demanded from the consumers (Hepenstrick & Tarasov, 2014, p. 11).

Dividing goods into the categories of divisible goods and indivisible manufactured goods predicts that high per capita incomes promotes variety. Ramezzana (2000, 23-24) shows in this standard monopolistically competitive framework that a higher GDP p.c. leads to more variety in indivisible manufactured goods besides divisible goods. Therefore, holding GDP fixed, a less populous country will trade more goods than a country with a higher population. Switzerland has the same GDP as Columbia, but imports are much more diversified along the extensive margin although trade theories suggest that it must be equal once controlled for GDP (Bernasconi & Wuergler, 2013, pp. 2; 28). Thus, it could be consistently shown that higher income countries trade more intensively and extensively with each other, whether it is because of capital-abundance, technology differences or consumer preferences. Incorporating non-homothetic preferences into the Krugman model, Markusen (2013) formalized the aforementioned empirical facts as a systematic theoretical model bringing theory under one roof. It is also stated, that in higher income countries higher price levels and markups are prevalent.

The presented empirical and theoretical propositions are quite simple if presented on a single firm level. Consider the decision of a market entry. With the introduction of a new product, let us assume

a mobile phone, a manufacturer has to evaluate the market size[4], demand[5] and market fixed/variable costs[6]. In order to be able to export, according to the theory, the manufacturer has already established a strong home market. After the evaluation, the manufacturer decides whether or not to serve the export market which explains why there are "zeros" because markets seem either unattractive or unprofitable. To extend the literature of variable mark-ups, pricing-to-market and income differences, Foellmi et al. (2013, 2-4) introduced international arbitrage (or the potential threat thereof) into economic trade theory. Similar to Ramezzana (2000), consumers have either the choice of consuming a good or not (take it or leave it). Threat of international arbitrage emerges when countries with a lower price level import high-quality/durable consumer or manufactured goods to a lower price than in the origin market and export them again to the richer northern country for a higher price pocketing the price differences as profit. Therefore, some firms chose to export to the south, lowering the price in both markets or decide to abstain from the southern export market and sell the good to a higher price only to the northern home market. Foellmi et al. (2013, 27-28) concludes that almost all northern U.S. exporters sell their products to northern markets but export probabilities to middle and low income countries decrease in importers' GDP p.c.

2.4 Income distribution and the Linder hypothesis revisited

Like most predictions on trade, Linder (1961, p. 96) already proposed that not only the average income per capita plays a vital role in explaining trade, but also its distribution within the country. Although he states that its influence on consumer demand should not be exaggerated. He concludes that "high-income earners in a poor country may demand the same goods as low-income earners in a rich country" and that is why a greater overlapping of demands due to income distribution results in a widened range of potential exports and imports. In a closed-economy, Murphy et al. (1989, p. 560) stress that the role of income distributions in industrialisation is crucial, highlighting evenly distributed incomes in order to create a larger domestic market. Although implications of the effects on trade on income distributions are common in the H-O model, the opposite view receives little attention. To what extend does the within-country income distribution affect patterns of trade?
In a two-good setting with differentiated manufactured consumer goods (Linder goods) and necessities, in which preference for Linder goods rises in income, it could be empirically shown that the theory holds and demand for Linder goods rises with a general rise of income levels and to a lesser extend with a less equal income distribution (Francois & Kaplan, 1996, p. 246/249). Another paper shows

[4]Market size can be measured in GDP. This is important as a high aggregated income of a country can offset possible high fixed costs and therefore justify an entry

[5]We assume here that the GDP p.c of the importer as the average income implies whether consumers are actually able to buy the mobile phone. Higher GDP p.c. are related to high price levels also resembling an attractive market as the manufacturer is able to charge higher prices increasing the markup

[6]The supply chain needs to be not too costly, so the manufacturer has also to include import/export costs in his calculations as well as the distance to serve the market

that an equal and an unequal country engages in trade of luxuries and necessities while the latter only imports the luxuries and the former specialises in trade of necessities. Intra-industry trade increases in inequality as long as it is lower than the partners inequality and inter-industry trade increases only if the partner's inequality is lower. This fact implies a V-shaped relationship with a turning-point at equal income distributions (Mitra & Trindade, 2005, pp. 1269-1270). Imports of luxuries increase with the destination countries' inequality-level while decreasing imports of necessities simultaneously. Estimates have shown that the U.S. would import around 9-13% fewer luxury goods assuming a similar income distribution as Canada (Dalgin et al., 2008, p.747). Fajgelbaum et al. (2011, p. 756) formalized the given facts which are also in line with the empirical findings of Klenow & Hummels (2005) that when countries export goods to a given quality (higher income countries export higher-quality goods), to countries of similar size, more varieties will be exported to countries whose income distribution or income ranking will be similar to its own (overlapping distributions).

The effect of within-country income distributions on patterns of trade could have been empirically shown indicating that not only average incomes but also full income distributions should be considered to measure the effect on aggregate demand (Bernasconi, 2013). Intuitively, this conclusion matches the predictions of Linder. When we consider two countries, similar in size but with different GDP p.c., an uneven income distribution of the poorer country implies that there is still a richer elite being able to afford the richer countries' products and therefore, the extensive margin of trade will be wider than in a similar country with an even income distribution. This may be surprising, as inequality is generally perceived as unfair and to be no greater good for society. In this thesis, we are only interested in the short-term effects on trade. It could be already shown that inequality may have desirable short-term effects on growth, but in the long-run the negative outcomes prevail (Halter et al., 2014).

2.5 Hypotheses: Income differences and income distributions do matter!

While in the academic literature the intensive, extensive or the quality margin of trade had often been mixed, this paper focuses solely on the extensive margin of trade empirically exploring the effect of GDP p.c. and the within-country income distribution on the set of goods exported/imported following the papers of Bernasconi & Wuergler (2013), Bernasconi (2013), Dalgin et al. (2008), Foellmi et al. (2013), Hepenstrick & Tarasov (2014) and Ramezzana (2000). As we will see, the patterns of trade are influenced by an interplay of several determinants. "Zeros" can arise along the extensive margin of trade why it is also attempted to capture those effects. We will have a look at those margins on aggregated market-level and on sectoral level through aggregated HS6-digits products according to Broad Economic Categories (BEC)[7]. The special focus will be the comparison of total products and

[7] Provided and updated according to the Harmonized System (HS) by the United Nations Statistics Division (UNSD)

the durable consumer goods (dcg) sector[8]. The following hypotheses are being conceived from the previous literature review.

Implication 1: Countries with higher GDP p.c. trade more extensively, resulting in a broader set of countries served. They also import from a broader set of countries.

Implication 2: Poorer countries are more attractive export markets for the richer northern countries when, ceteris paribus, inequality is higher in comparison to peers. Therefore, these countries import from a broader set of countries.

Implication 3: The percentage of goods/products traded rises in GDP p.c. "Zeros" arise not only through trade frictions, but also from a lack of demand when destination countries have less purchasing power.

Implication 4: Similar to proposition 2, more goods are demanded in a poor country with more inequality, ceteris paribus, than in a country with a more equal income distribution.

Implication 5: Our findings will be more significant in the durable consumer good sector considered to be meme "luxuries" and therefore, income differences across and within-countries will matter more.

[8]Durable consumer goods, BEC Nr. 61, are assumed to be related to the luxuries or the Linder goods mentioned in the previous literature review. Durable consumer goods, includes commodities which have an expected lifetime of more than one year and are of a relatively high value, such as refrigerators and washing machines, together with other commodities with a useful life of three years or more

3 Data

The data set is derived from several sources and the construction of the dependant variables will be explained in detail. To construct the variables, we make use of the BACI database provided by CEPII created and used in the paper of Gaulier & Zignago (2010). BACI delivers highly disaggregated trade flow data up to the HS6 digit level according to the Harmonized System for product classification, Version 2002 (HS02) for the year 2012. The database covers more than 200 countries with its bilateral trade flows originally collected and stored within the U.N. Comtrade database reconciling F.O.B and C.I.F trade values to assign the same trade flow the same unique value dropping double observations. This method enables more observations per countries since import data is usually more precise (revenue generating). We drop observations from small countries with a population less than 1 mio. and where data on income is not available leaving 125 countries in the data set (all countries & HS classifications numbers listed in *Appendix A*, tables 4 & 6).

Dummy variables for low, lower middle, upper middle and high income countries are being generated according to the U.N. country classification (United Nations, 2014). Moreover, to summarise the single HS6 categories to sectors with similar characteristics, namely agriculture products, capital goods, consumption goods and its subgroup durable consumer goods (dcg), Broad Economic Categories (BEC) identifiers are being added to the data set (United Nations, 2002).

Several methods to model the extensive margin of trade have been proposed (Klenow & Hummels, 2005; Bernasconi & Wuergler, 2013; Hepenstrick & Tarasov, 2014; Santos Silva et al., 2014). We will be following the rather simple count model of Hepenstrick & Tarasov (2014, p. 33) and Santos Silva et al. (2014), because it is more intuitive and its properties are advantageous to match the assumptions made. Two separate data sets will be created.

First, we will count the number of destination markets and generate percentages of maximum markets $mRATE_i$ or origin markets $mRATE_j$ on country-level according to the following specification:

$$mRATE_i = \frac{\sum_{ij_{HS6}} I(q_{ij}(m_j) > 0)}{124(\text{Maximum trading partners})}$$

The indicator variable takes on the value $I = 1$ if the trade flow between exporter i and importer j is positive and hence, the destination market j will be counted. In the end, we have the variable taking on values such as: Germany exports to 100% of the countries (the maximum) or Guinea-Bissau exports to only 27.4% of all possible destinations. Duplicates have been dropped so that there are only 125 countries in the data set. When we count the number of origin countries, the same procedure is being used but in reverse, in order to generate $mRATE_j$. Moreover, several variations of these variables have been created. For instance, the variable $mRATE_i^{dcg|high}$ would mean the percentage of high income countries served in the sector "durable consumer goods". GDP p.c. and GDP (Constant

USD per 2005) and Gini coefficients originally from the World Development Indicators (WDI) had been added as well, as these are the variables of interest. Also fixed trade costs, e.g. landlocked dummies, costs to set up a business or number of documents to import/export are taken from WDI or GeoDist (Worldbank, 2014; Mayer & Zignago, 2011).

Second, we extend the data set to fill in all possible combinations of country pairs leading to $125 \times 124 = 15'500$ observations. As we are also interested in "zeros", this method of extending the data set and treating missing observations as "zeros" is common in the gravity literature and was used in several works (Helpman et al., 2008; Santos Silva & Tenreyro, 2006; Santos Silva et al., 2014; Gómez-Herrera, 2013). Every of these 125 countries in the data set reported at least once in every product category why it is justified to assume that missing observations are "zeros". For every trading pair, the number of products are being counted according to the following formula:

$$pRATE_{ij} = \frac{\sum_{ij_{HS6}} I(q_{ij}(p_{ij}) > 0)}{\text{Total in HS classification}}$$

We start with around 5.6 mio. observations of HS6-digit bilateral product trades. For every positive trade, we count the trades and aggregate to sectors on each trading pair. For instance, Germany trades with Switzerland 4'577 in total products, 143 durable consumer goods, 1116 consumption goods, 603 capital goods and 2'830 intermediate goods. In comparison, the number of products traded with Albania are 1'658 in total, 80 durable consumer goods, 439 consumption goods, 300 capital goods and 910 intermediate goods. Percentages are created using the HS classification as denominator for the number of counted products. This methodology allows to capture the extensive margin of trade as well as the income effect Hepenstrick & Tarasov (2014, p. 33). Also Eaton et al. (2004) advocates for a sectoral analysis instead of just looking at the aggregated total to capture income effects. In addition to the aforementioned fixed trade costs, we add bilateral trading pair specific variable trade costs to the data set e.g. distance, colonial relationships, regional trade agreements (RTA), common language etc. An overview of all the used variables and the respective sources can be found on table 5. Descriptive statistics can be also found in *Appendix B*. On table 7 we can see that the average country exports only to around 63 or 50.8% of all countries in the category of durable consumer goods, but to 84.7% in total. We have also high variation in percentage of countries served. While the maximum is 100%, one country only exports consumption goods to 10% of all countries. Moreover, histograms to the dependant variables $mRATE_i$, $mRATE_j$ and $pRATE_{ij}$ are shown in figure 22, 23 and 24.

4 Stylised facts on the determinants of the extensive margin of trade

4.1 Sectoral analysis of the percentage of origins/destinations

In this section, we will discover empirical evidence to *implication 1* and *implication 2* regarding the extensive margin of trade on market-level. Using the variables $mRATE_i$ and $mRATE_j$, stylised facts will be presented to show graphically the effects of GDP p.c. and the within-country income distribution, measured with the Gini coefficient.

4.1.1 Do per capita incomes shape international patterns of trade along markets?

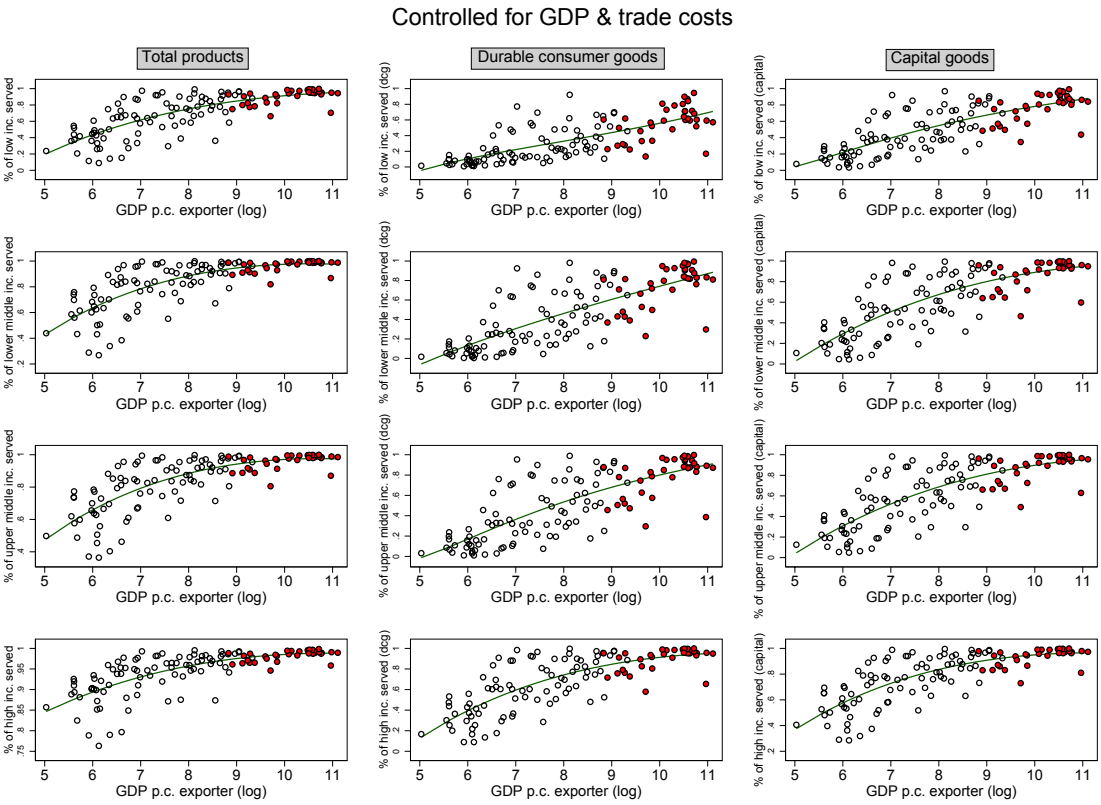

Figure 1: Percentage share of destination countries served

First, to shed light on *implication 1*, figure 1 displays the percentage of destinations served ($\%mRATE_i$) against exporters' GDP p.c. For illustrative purposes, the values of the per capita incomes are in logarithmic form. The percentages of destinations served take on values from 0 to 1. In each column is a different aggregated sector, namely total products, durable consumer goods and capital goods. In each row, different income groups are shown as importers. Hence, on *top right* the percentage of low income countries served in the category "total" is revealed. *Bottom right* would therefore show the percentage of high income countries served in the category "capital goods". As the variable of

interest is GDP p.c. and variation is prevalent, it is necessary to control[9] for GDP and trade costs[10] according to common gravity literature in order to isolate the effect of GDP p.c. on the number of destinations served (see also Bacchetta et al., 2012, p. 105; Head & Mayer, 2013, p. 6).

As one can infer from figure 1, there is a strong positive relationship between an exporters GDP p.c. and the number of destinations served across all sectors and destination income groups. Therefore, it can therefore be infered that countries with a higher income per capita export to a broader set of countries confirming the notion of *implication 1.* Explanations for this phenomena is linked to higher productivity due to better technology and more capital-abundance associated with higher GDP p.c. (Helpman et al., 2008; Hepenstrick & Tarasov, 2014; Bernasconi & Wuergler, 2013).

Especially durable consumer goods export probabilities seem to increase with importers' income. While upper to lower middle income countries appear to export durable consumer goods to high income countries, these origins opt out of most low income markets indicated by a straighter, flatter regression line. As an additional supporting fact, the 28 low income countries export on average to 15% of the other low income destinations, to 12% lower middle income markets and 13% upper middle countries, but to a relatively high value of 37% high income destinations. A simpler version of the durable consumer goods markets can be seen in figure 2. On the X-axis the income groups are reported while the different lines are separated by destination income groups. As with figure 1, the low number of export markets for durable consumer goods for low income origins can also be explained by an absence of production facilities. The same applies to capital goods. Both sectors require capital, technology and know how to produce them. Moreover, a lack of demand in the home market due to low GDP p.c. offers little incentive to innovate. In contrast, high income countries possess all these characteristics why they are competitive in these sectors and are able to export to a broader set of countries.

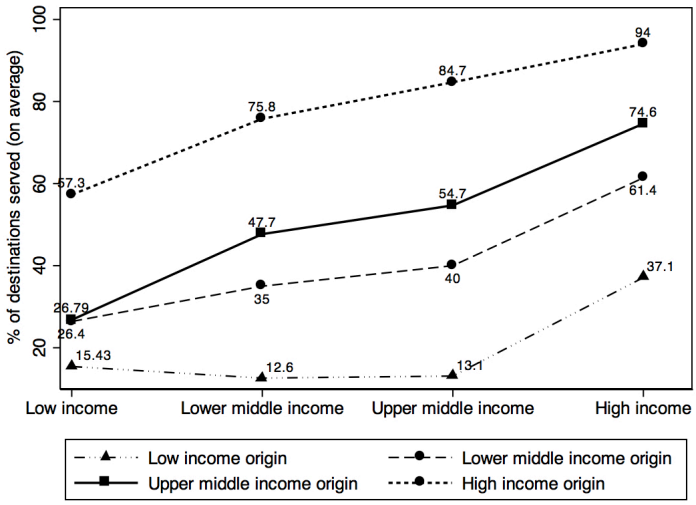

Figure 2: Average percentage of destinations served by income groups

[9]Further treatment of the used estimation method "fractional response regression" can be found in section 5.1.1

[10]As this is no bilateral data set, we can not use "distance" as a proxy for trade costs. Instead, we use the fixed export/import costs of the origin/destination country to account for trade frictions. Therefore, the variables "cost to export/import ($ per container in logs)" is being used as alternative.

In figure 25 in *Appendix C* the whole setup is switched from exports to imports and extended with the sectors consumption & intermediate goods. The steep increasing concave curves on the bottom row signal that most countries across all income groups maintain trade relationships with high income countries as origins. Only in the intermediate goods sector emerge somewhat more balanced patterns of trade for low income countries (*top right*). Not only do high income countries import from low income countries, but also lower and upper middle income countries do import from a broader set of low income countries. Additionally, as we include consumption goods aggregating all goods intended for household consumption, a similar pattern of trade can be observed as the one described for intermediate goods. Intuitively, these patterns resemble typical supply chains observed in reality. Unsophisticated products requiring low-level technology, raw materials or non-processed food such as fruits, vegetables etc. are normally imported from southern[11] countries. Although Linder (1961, p. 88) stated, that countries with similar demand patterns engage in trade, we clearly see that the set of possible trading partners clearly increases in GDP p.c. Low income countries do not really export to a broader set of low income countries, nor do they maintain significantly more trading relationships among each other in the import market. The statements made are also related to the notion of Helpman (1981, p. 305) that countries with differing income levels engage more in inter-industry trade and intra-industry trade prevails within similar income levels. Though, we will get back to this later.

One other possible narrative of the observed patterns can be that higher income countries have more means to establish supply chains. Meaning that they do have a higher reach. It is hard to imagine that low income countries can afford to build massive infrastructure projects just to export products to more countries or import luxuries for a very small elite. Therefore, as mostly intermediate and consumption goods are exported to a broader set of countries from low income origins, it can be deduced that higher income countries invest in infrastructure to funnel raw materials, intermediate goods etc. to the processing facilities in the north. On the other side of the coin, we can see that durable consumer goods are exported less frequently to low and lower middle income countries. This confirms somewhat the Linder hypothesis where demand actually matters, as these countries have not enough purchasing power for more sophisticated products and therefore the exporters abstain from these markets in the first place (Foellmi et al., 2013, 27-28).

In figure 3, the aforementioned assumptions are shown. The X-axis is marked with exporters' GDP p.c. and the Y-axis is the north/south ratio of served destinations. Northern countries include high income countries while southern countries summarises the remaining destinations (low, lower middle & upper middle income countries). A north/south ratio equal to 1 implies that for every served southern destination a northern destination will be served as well. A lower north/south ratio means

[11] The author is aware of the politically incorrect term of "southern" countries. Nevertheless, low income, lower middle income and upper middle income countries are summarised in this group while high income countries are called "northern" countries

practically speaking an improvement of the extensive margin, because trade will not only be directed towards northern countries, but also more diversified to southern countries. That's why countries with a lower GDP p.c. tend to have a higher north/south ratio serving mostly high income countries. The effect can be seen in all product categories, although less variation is seen for intermediate goods. Especially durable consumer goods are heavily affected by unexplained variation.

To summarise this section, a strong positive correlation between the percentage of destinations served and income differences across all product categories could have been shown. Moreover, similar effects can be observed for the percentage of origins. *implication 1* is for now a valid claim. Northern countries export to a broader set of countries while southern countries export mostly to northern countries. Whether these effects are statistically and also economically significant will be explored in section 5.

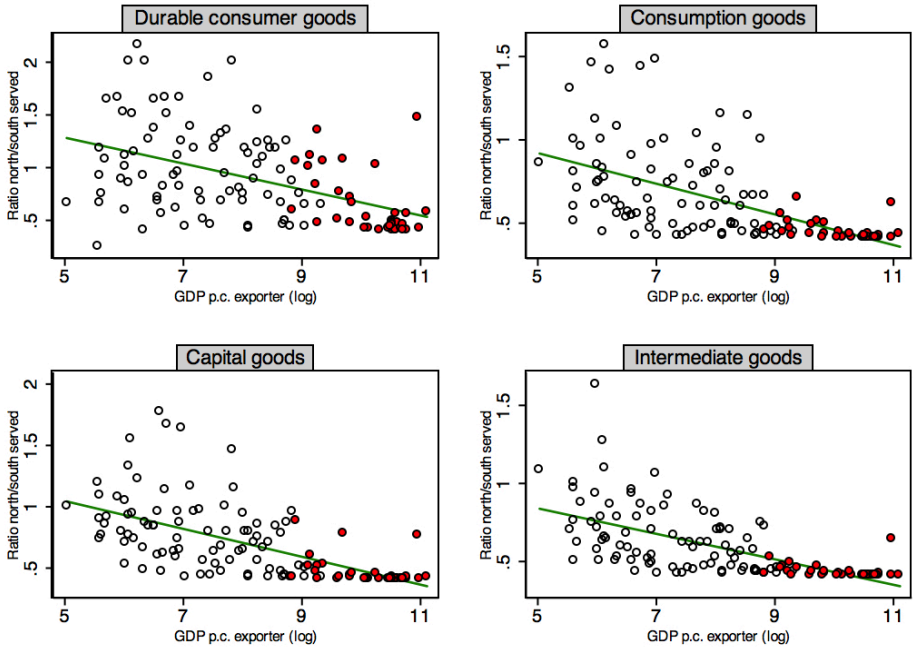

Figure 3: Exporter GDP p.c. and ratio north/south destinations served

4.1.2 Market diversification and income distributions

In this section, stylised facts supporting *implication 2* will be presented. In the literature review, it was already stated that countries with a more unequal within-country income distribution are able to extend the margin of trade and import more products from a broader set of countries. Unanimously, non-homothetic preferences are marked as a crucial precondition to include income distributions into trade theory (Francois & Kaplan, 1996; Mitra & Trindade, 2005; Bertola et al., 2006; Dalgin et al., 2008). Similarly, also Linder products/luxuries etc. are named as the only products affected. Consumers increase consumption with higher income and are also able to buy the higher quality/luxurious products. Purchasing power can be higher for certain groups within a country if there is higher inequality (Francois & Kaplan, 1996, p. 244). Durable consumer goods include products which are

more sophisticated (e.g. refrigerators, mobile phones, washing machines etc.) and therefore mostly produced or distributed through companies in northern countries. While in a country with low income levels most people cannot afford these kind of apparatus, a certain inequality can shift demand to a level resembling the demand patterns of high income countries. Hence, the upper class of this country would show similarities to the lower class in a high income country demanding similar goods.

As a proxy of inequality, we make use of the most common and popular, named the Gini coefficient (or index). A coefficient of zero can express perfect equality meaning everybody has the same income while a coefficient of 100% is perfect inequality as only one person earns all the income within a country (or in the world). It is worthy to mention that Gini coefficients perform poorly in capturing extreme values of either side of the income distribution. We therefore use a decile ratio from the highest/lowest decile to capture extreme values as a robustness check (Dalgin et al., 2008, p. 758). In Figure 4, Gini coefficients are plotted against GDP per capita. In figure 26 the top/bottom decile ratio is shown as well for illustrating the difference[12]. First of all, a non-linear relationship can be deduced. High income countries have, on average, a Gini index of 33% while the other income groups move around 40%, regardless of GDP p.c. Matching the presented facts with theory implies that it is

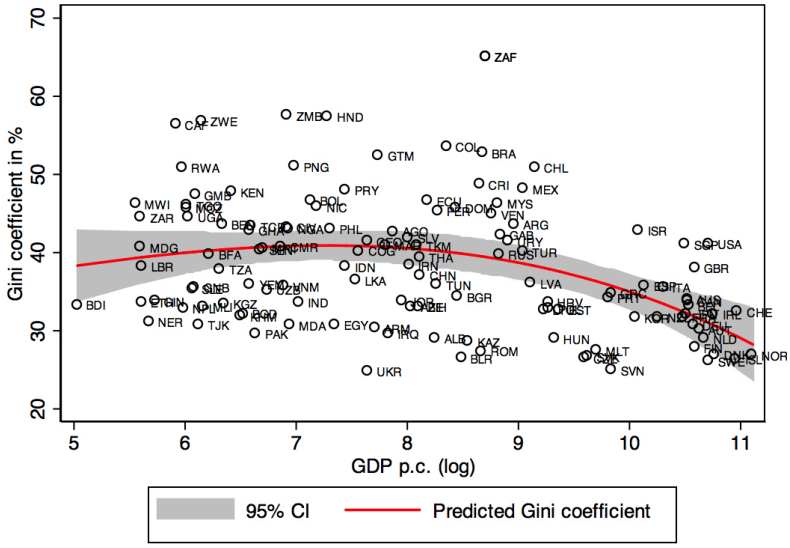

Figure 4: GDP p.c. and Gini coefficients

most likely that only some upper middle income countries can improve or extend their percentage of origins for durable consumer goods. Overlapping income distributions with high income countries are almost impossible to achieve for countries from low and lower middle income groups as their income differences are inherently too high. In figure 5 the percentage of origins of total products $mRATE_j$ and durable consumer goods $\%nmi_j^{dcg}$ are plotted against the importers Gini coefficient controlled for GDP, GDP p.c (logs) and trade costs[13]. Like in figure 1, we make use of a fractional response regression to predict $\%nm_j$ & $\%nmi_j^{dcg}$.

[12] In *Appendix C*, a brief discussion and comparison of the decile ratio and the Gini coefficient is being provided
[13] Cost to import (in $ per container)

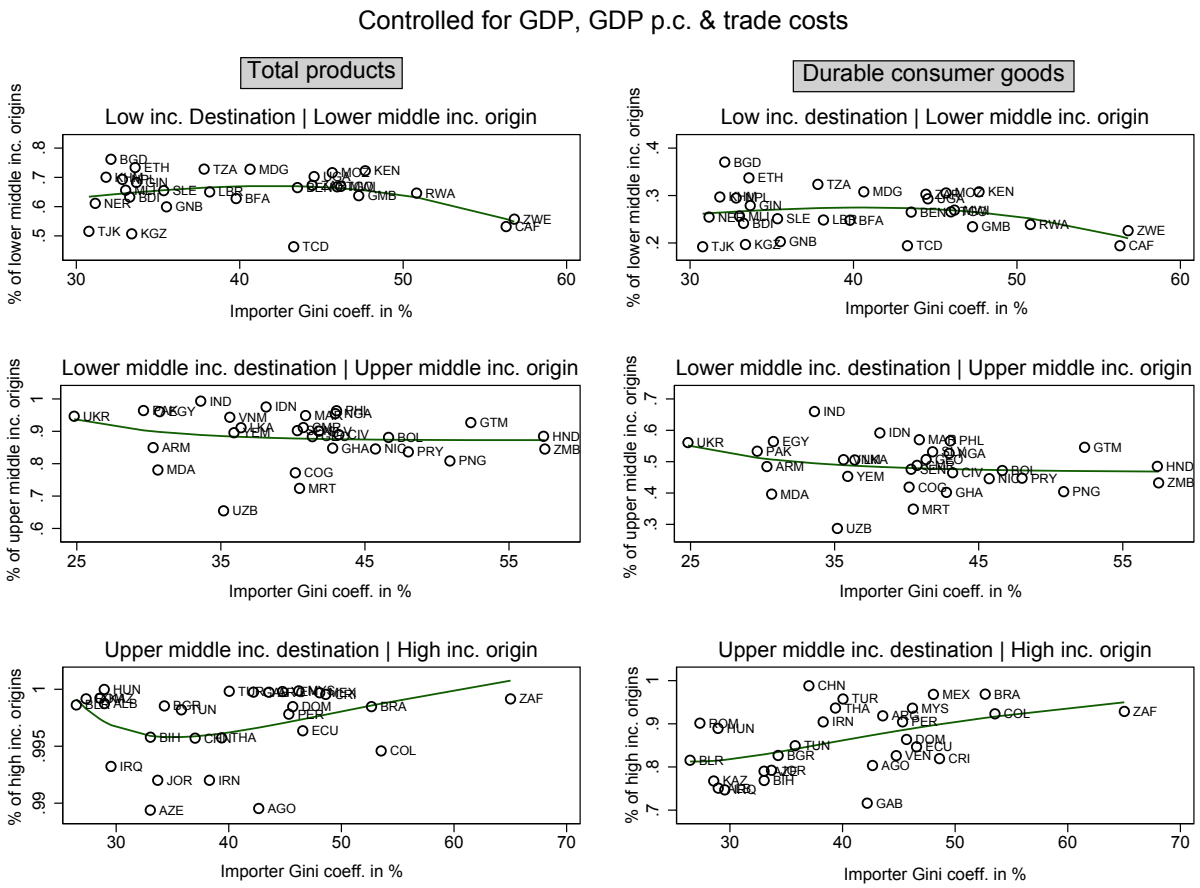

Figure 5: Gini coefficient and the effect on % of origins

The figure reveals very insightful information on the nature of the patterns of trade when inequality is present. As predicted in theory, upper middle income countries do profit from an unequal within-country income distribution. Although we controlled for GDP, GDP p.c. and trade costs, there is still variation left resulting in effect on the number of origins. Especially on the *bottom right* we can observe a linear positive relationship between an importer's Gini coefficient and the number of high income origins. On the contrary, the low and lower middle income countries seem to be negatively affected by higher inequality for total products and durable consumer goods as well. This is not only true for their trade flows between them and their next higher income group, but also for general imports from high income origins (see figure 28 in *Appendix C*). Therefore, *implication 2* seems to be true as well, but only on the condition that upper middle income countries (can also be interpreted as emerging economies) are being considered. Again, it seems like low income and lower middle income countries do not participate extensively in trade and are more or less dependent on few trading partners. In contrast to upper middle income countries, low income countries lack purchasing power and even inequality does not affect or improve the percentage of origins. Furthermore, some signs that *implication 5* is relevant can also be observed. The regression line for durable consumer goods is significantly steeper for upper middle income countries indicating that a separation of luxuries/dcg and total goods is necessary (Dalgin et al., 2008, p. 758; Eaton et al., 2004, p. 154).

4.2 Product variety in bilateral trade: The extensive margin of trade extended

This section aims to provide a deeper understanding of the extensive margin of trade on bilateral level. As it was noted in section 3, the data set will be extended to a dyadic setup resulting in 15'500 observations where each country trades with each other or not. Hence, not trading partners will be counted but the number of product categories. This is just another definition of the extensive margin of trade as it can be investigated on several layers. This setup allows to compare the origin and destination countries' characteristics in more detail. The extensive margin of trade can be investigated and defined on several levels, e.g. on country-level, sectoral-level, firm-level and even on single product-level accounting for every single trade flow category (Santos Silva et al., 2014, p. 67). In figure 6, the number of markets and the number of traded product categories for imports and exports are shown and plotted against each other. There is strong evidence for a positive relationship for both import and export markets that traded product categories (and therefore variety) increase with markets served. What strikes the reader's eye is the positive non-linear relationship which is skewed towards a preference for more export markets (left) and more variety for imports (right).

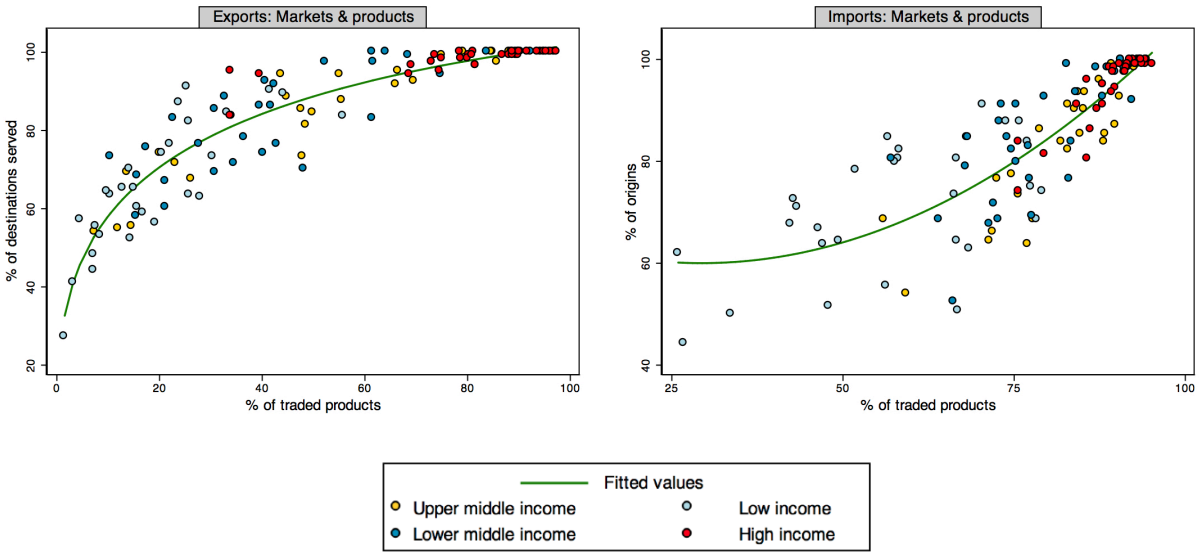

Figure 6: Relationship between markets & traded products

Furthermore, high income and upper middle income countries all range in the upper section of both markets and products. It can be inferred that the number of markets served has to be extended in order to export or import more products. In this section, we will have a look at *implication 3*, *implication 4*, and *implication 5*. The analysis is extended not only to export diversification (more destinations served), but also to a more extensive trade relationship resulting in more traded product categories.

4.2.1 The role of per capita incomes in bilateral trade

The H-O model predicts consistently high values for trade relationships, which are not observed in reality. This fact became known as the "case of the missing trade" (Trefler, 1995, p. 1030). Countries do not always export labour-intensive goods to capital-abundant countries and vice versa. Some patterns of trade seem to be rather arbitrary. Linder (1961, p. 137) explains the emergence of zero trade flows with differing demand patterns. The taste for certain goods depends heavily on the income-level and some goods are not demanded in certain parts of the world at all. From a supply-side perspective, it is also evident that some firms abstain from markets promising only small margins or result even in a loss for the firm. This assumption is tightly linked to the notion of price-to-market when firms charge a price according to the purchasing power of a market (Atkeson & Burstein, 2008, p. 1998). Moreover, when introducing international arbitrage (or even just the threat), firms also abstain from the market and rather choose to serve only the home market (Foellmi et al., 2013). Also trade costs play a role in determining patterns of trade as some firms can just not afford to enter the market in certain countries due to trade barriers (Arkolakis, 2010). In figure 7, zero trade flows between two countries for total products and durable consumer goods are plotted against the income group of the destination.

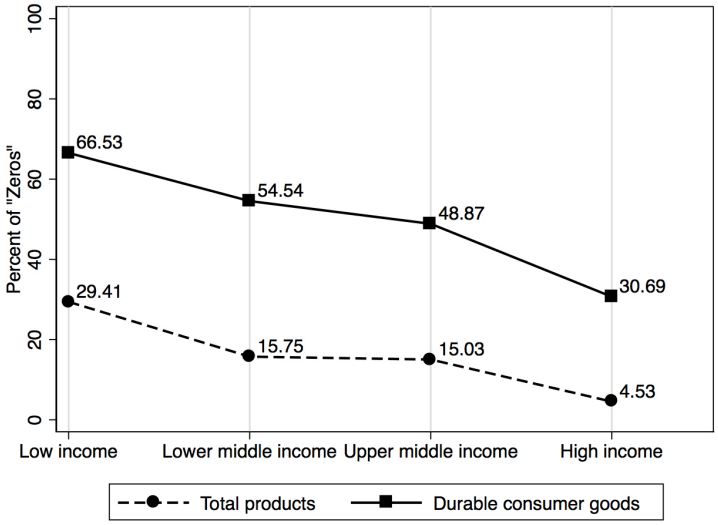

Figure 7: Percent of "zeros" in bilateral trade by destination income group

It is depicted that 66.52% of all trade flows to low income countries will be "zeros" for durable consumer goods while it is 29.41% for total products. Generally, there are significantly less zero trade flows for total products than for durable consumer goods. "Zeros" decrease in GDP p.c. or in this case, with income group. High income countries have up to 30% less "zeros" in durable consumer goods than low income countries and 18% less than upper middle income countries. The results from figure 7 can be traced back to all the aforementioned facts. Low income countries do both lack the purchasing power as mentioned by Linder, or firms just abstain from the market as there is a risk of arbitrage.

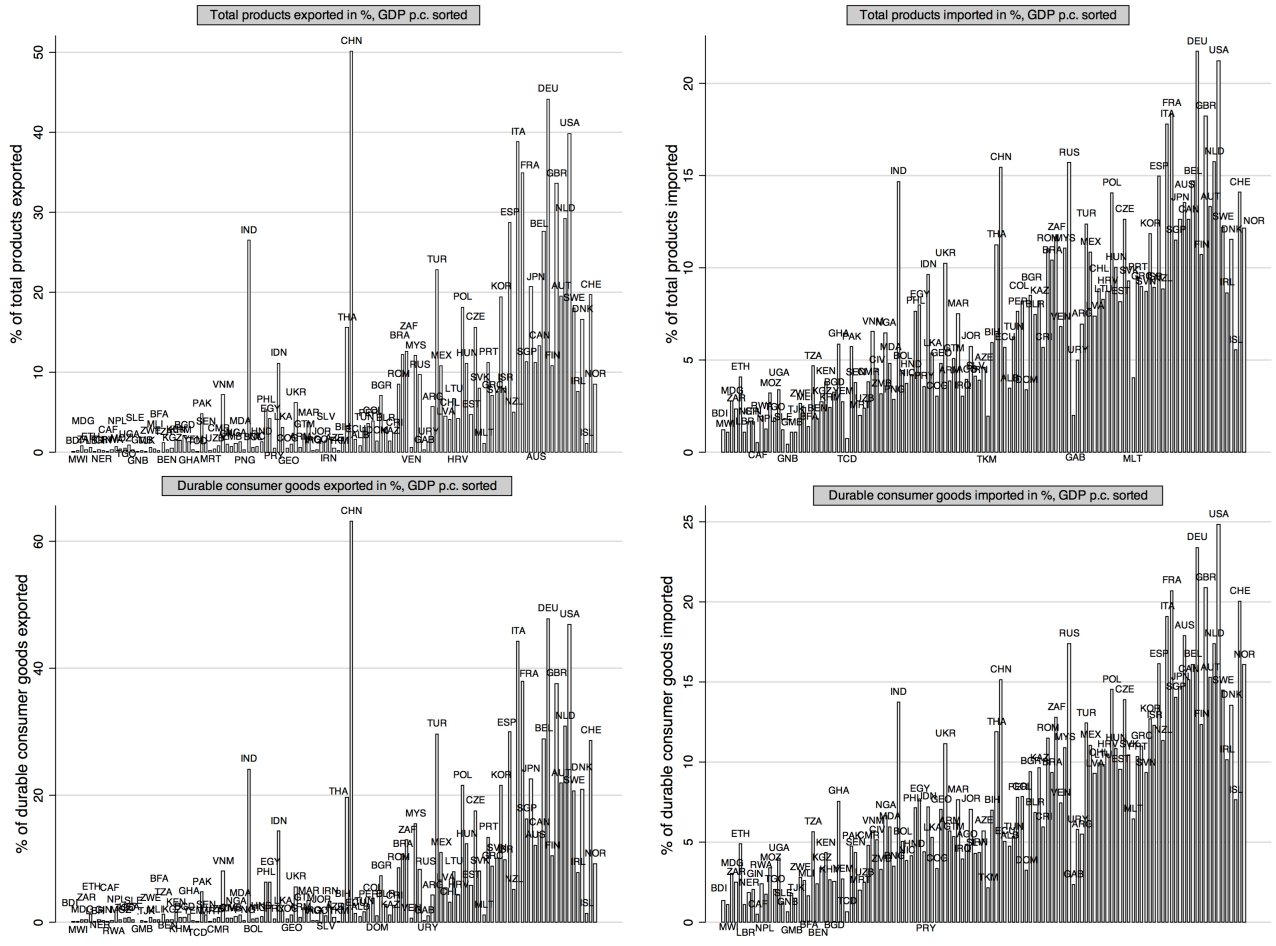

Figure 8: Percentage of traded products (exports/imports), GDP p.c. sorted ascending from left to right

Having shown that some countries do not engage in trade with each other at all, in figure 8 the countries with positive bilateral trade flows are illustrated. On top, the category "total products" is shown while on the bottom the focus is on the durable consumer good sector. To the left are the exporters and the the right the importers. All figures are ranked by GDP p.c. ascending from left to right. First, on the right-hand side, the emerging pattern is as expected. The extensive import margin of trade increases in GDP p.c. Unanimously, more variety of products is demanded in higher income countries. Both the United States and Germany import in most product categories with around 25% on average. The results lead to no surprise. Indeed, Hepenstrick & Tarasov (2014) and Bernasconi & Wuergler (2013) predicted this outcome. Whether it is explained by technology differences, higher wages resulting in more purchasing power or homothetic-preferences in several theoretical models comes second.

On the export-side to the left, the picture looks rather bleak. The world market appears to be dominated by high income countries, producing and exporting in most product categories. Mind that these percentages have nothing to do with the intensive margin of trade, respectively the volume of trade. What we can infer is that higher GDP p.c. leads to more product variety to be exported. The fact hat the percentage of product categories exported does not only depend on the exporter's size

(GDP or population) demonstrates the examples of Switzerland, Sweden, Denmark, Czech Republic etc. All differ in the aforementioned characteristics, but have in common a high GDP p.c. exporting around 20% of total products and durable consumer goods. Outstanding is the position of China, India, Thailand and on a smaller scale Indonesia. China exports in 50% of all product categories and in 60% of durable consumer goods crowning itself to be the world leader in export product diversification even though being an upper middle income country. Also India, Thailand and Indonesia stand out from the crowd having way above average percentage of products with around 20%. The reason for these outliers can be traced back to their population size profiting from heavy investments from the western world over the last 30 years. Furthermore, Krugman (1980, p. 958) already noted that countries export products for which there is strong demand in the home market. From a western investors' perspective it makes therefore sense to invest in low income countries first where there is a potential big home market. Also through heavy state subsidies, firms were able to force other companies out of the home market taking on dominant positions in their industry. This notion is also related to the ideas of Melitz (2003) when exporting firms hold dominant positions in their respective home market. Low and lower middle income countries have a negligible role in international trade. Low income countries export in astonishingly low 0.5% product categories in total and durable consumer goods. Similarly, lower middle income countries export only in 3%.

We can infer from figure 29 in *Appendix C* that low and lower middle income countries export mostly, with around 2/3, in the consumption goods (excl. durable consumer goods) sector and, to a lesser extend, in the intermediate goods sector. As already mentioned in the previous section, this fact can be explained by their abundance in raw materials, natural resources and strong agriculture sector etc. providing the simple products for the industry to process and for consumers more variety in general. On the other hand, import patterns look more balanced what we can deduce from figure 30, as most countries import from all three sectors the proportionally equal amount proportionally to their GDP p.c. Additionally, the percentage of imported and exported product categories reveals a considerable difference. While higher income countries export in 50% more product categories than they import with a high proportion of capital and intermediate goods, low and lower middle income seem to specialise in only one sector but import from all three. Hence, GDP p.c. is decisive in predicting the sectoral composition of trade flows.

To sum up, there is strong evidence in favour of *implication 3*. Zeros decrease in GDP p.c., the percentage of exported and imported product categories increase in GDP p.c. in absolute values in all sectors (although proportions differ with GDP p.c.).

4.2.2 Inequality and the percentage of traded product categories

To establish the relationship between inequality and the percentage of traded product categories is not quite obvious. Remember that Linder (1961, p. 96) warned not to overestimate the effect of inequality on product demand stressing the rather marginal small effect on the shift of demand to meme a similar income-level of countries with higher GDP p.c. As more inequality in a country is present, the small rich elite demands a similar variety of goods as the middle-class from a high income country. Although he defends the importance of inequality as a determinant of the extensive margin of trade, it remains a puzzle to completely isolate the effect. While on market-level upper middle income countries with higher inequality, ceteris paribus, import durable consumer goods from a wider set of countries, *implication 4* proposes a similar effect on the extensive margin of trade on sectoral-level, especially in durable consumer goods.

While Bernasconi & Wuergler (2013) investigated the overlaps of income distributions, we stick to the methods applied in the paper of Dalgin et al. (2008) using the Gini index or the decile ratio to capture a destinations' inequality and therefore modelling more purchasing power to buy more upscale products. With higher income also demand for variety rises boosting the extensive margin of trade (Mitra & Trindade, 2005). When we take a look at figure 9, on the left, the average percentage of imported durable consumer goods from northern countries for upper middle destinations are shown sorted ascending by Gini and to right by the decile ratio. On the left, we see a small cluster with high Gini coefficients on the right-side of the figure but no clear patterns are visible. Interestingly, on the right figure some countries like Hungary, Turkey, Romania etc. shifted to the right and a clear pattern can be seen that inequality may have a positive impact on the percentage of products traded in the durable consumer good sector. The caveats of the Gini coefficient comes to the surfaces as it seems crucial in this example to capture extreme values on the income distribution. The decile ratio proofs a better choice in this example. Of course, China again enjoys an outstanding position and should not be treated as representative for this income group.

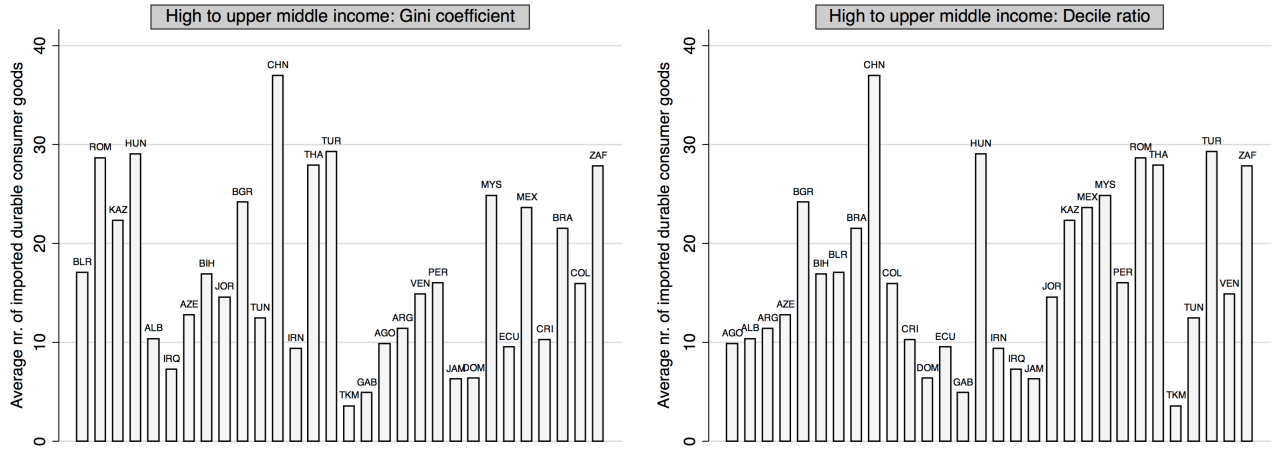

Figure 9: The percentage of imported durable consumer goods and inequality

Trying to replicate the effect for lower middle income countries was rather futile as seen in figure 31. We will have a more profound look on the biggest exporters in the world economy hoping to disentangle several effects on a single-country-level in order to answer both implication 3 & 4.

4.2.3 Country examples: United States, Germany, China & India examined

Four of the biggest exporters deserve a deeper look into their export behaviour, namely the United States, Germany, China and India. For the effect of income differences (GDP p.c.), the whole sample will be considered while for the effect of the income distribution (Gini), the sample will be restricted to north-south trade flows for USA and DEU and southern flows for China and India. The rationale behind the restrictions is simple. First, the income distribution matters mostly for southern countries importing from a northern origin to enhance purchasing power and reach an overlap of income distributions. Second, India and China will export anyway to northern countries, as these countries are already on a higher income-level. But within the southern country group, income differences exist as well, why it is worthwhile to shed light on inequality-effects as well.

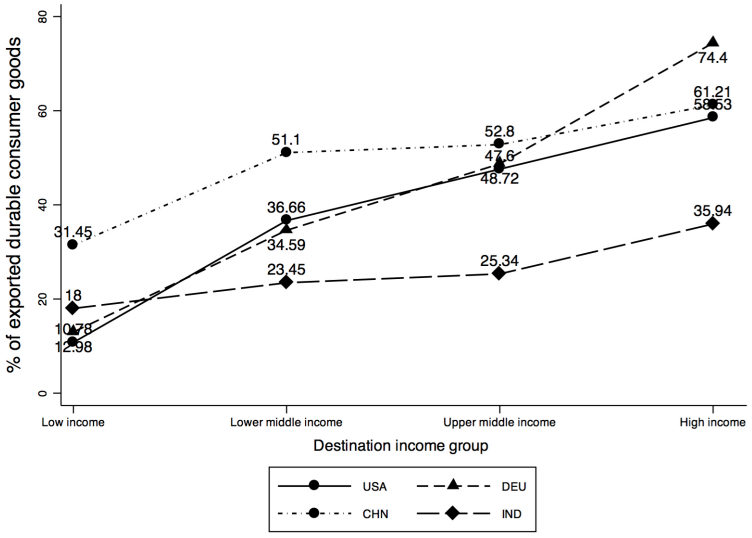

Figure 10: Comparison of % exported products: USA, DEU, CHN & IND

In figure 10, the % of exported products in the durable consumer good sector is plotted against the destinations' income group similar to figure 2. As we can see, the pattern is not as straightforward. Even though the percentage of products traded increases in income, China exports in more product categories across all income groups than the USA, which is rather surprising. Similarly, Germany lacks behind as well in serving southern destinations in comparison to China. Another surprising fact is that India exports in more export categories to low income countries compared to USA & DEU. Germany is leader in exporting durable consumer goods to high income countries with around 74%. We will consider next all four countries in detail. Using a fractional response model, we will also isolate the effect of income differences and distributions controlling for GDP and trade costs[14], for

[14]Included in trade costs are the distance between the capitals in logs and dummies for regional trade agreements and colonial relationships.

the GDP p.c. model and for GDP, GDP p.c. and trade costs for the Gini model. All figures are conceived using the same methods and they only differ in the specific country values.

The United States (USA)

In figure 11 on the left-hand side, the destinations of US-products are depicted showing the percentage of durable consumer goods exported from the U.S. and the importer's GDP p.c. On the right-hand side figure, the same is shown with the importer's Gini coefficient while both use weighted GDP to show the size of an economy graphically. It is evident that with rising GDP p.c. the demand for more variety emerges, which was already theoretically predicted by Bernasconi & Wuergler (2013), Hepenstrick & Tarasov (2014), and with a higher extensive margin leading to less "zeros" by Foellmi et al. (2013). What strikes us is that the size of a destination seems to allow countries with lower GDP p.c. than the US to compensate for that, for example destinations like China, India, Indonesia, Turkey, Mexico, South Africa etc. On the right-hand side figure, we see that the extensive margin decreases in Gini, but not in the aforementioned countries. Interestingly, right on the bottom is Iran, which is the only country in this section of the graph with a considerable size of the economy. Of course, we can explain its standing with the harsh sanctions imposed by the USA. What is additionally inferred from figure 11 is that all these countries are hegemons in their area of influence, whether it is Mexico in Central America, Argentina and Brazil in South America, South Africa in Africa or Turkey in the Middle East. Hence, there is a lot of other variation in this figure why we isolate the effects of GDP (size) and trade costs (incl. distance) to get a clearer notion of the effect of income differences and income distributions.

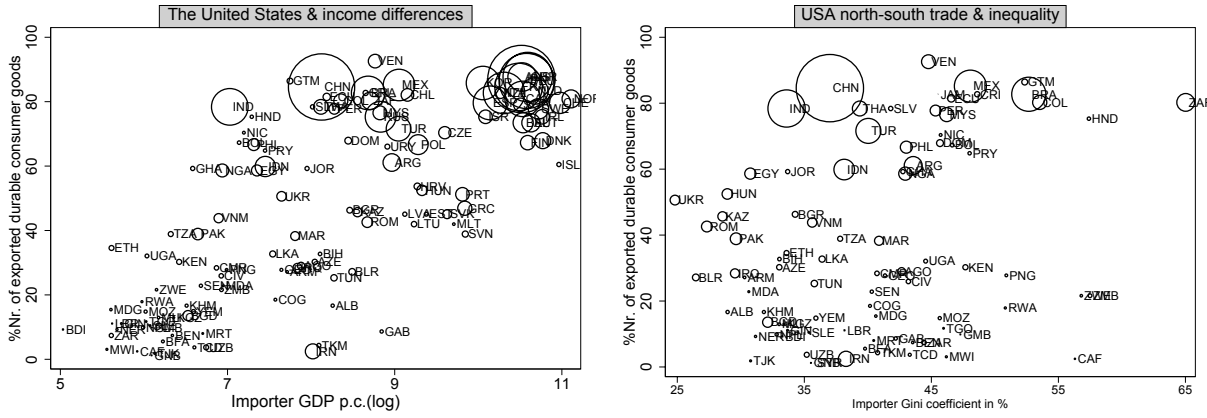

Figure 11: USA and its destinations for durable consumer goods $pRATE_{ij}^{dcg}$

The effect of GDP p.c. in figure 12 is as expected. Higher GDP p.c. can lead to moderate higher percentages of traded product categories in the sector of durable consumer goods. The effect of income distributions is twofold. The red line is restricted to upper middle income countries while the green line represents lower middle and low income countries. As predicted in Dalgin et al. (2008), all southern countries have increasing extensive margins with rising inequality, although the effect is stronger for upper middle income countries.

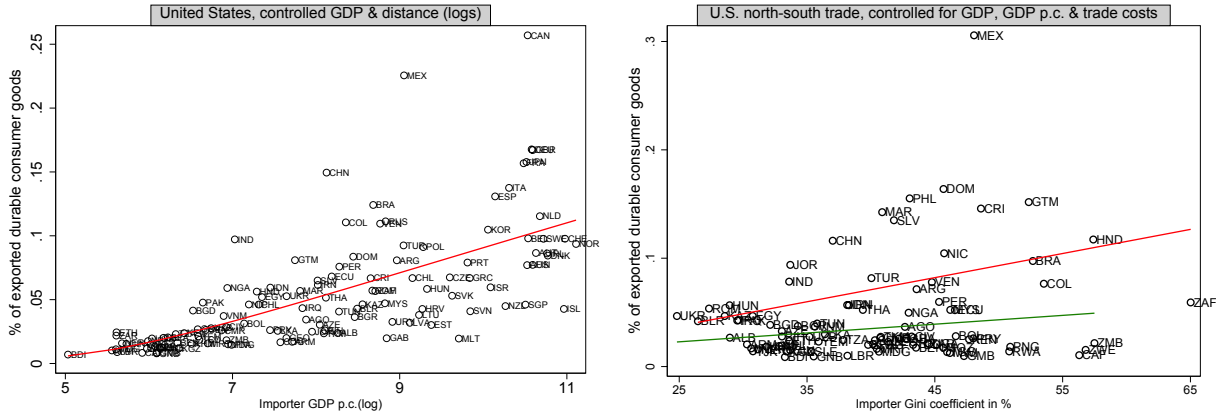

Figure 12: Percentage of traded durable consumer goods and inequality in USA

Germany (DEU)

Germany has a different background than the USA as well as a differing geographical location. Being a member of the European Union, its "natural" trading partners are already set. Germany exports in 74.4% product categories on average to high income destinations marking the highest value within the durable consumer good sector. On the other side, it imports in 52.8% product categories implying intra-industry trades. Mitra & Trindade (2005, p. 1270) already noted that when two countries share similar inequality-levels, intra-industry trades will be more dominant, which is valid for the EU, as inequality is lower on average. While the left graph does not reveal any new insights, the effect of inequality seems to tend more to the negative than to the positive, contradicting common theories.

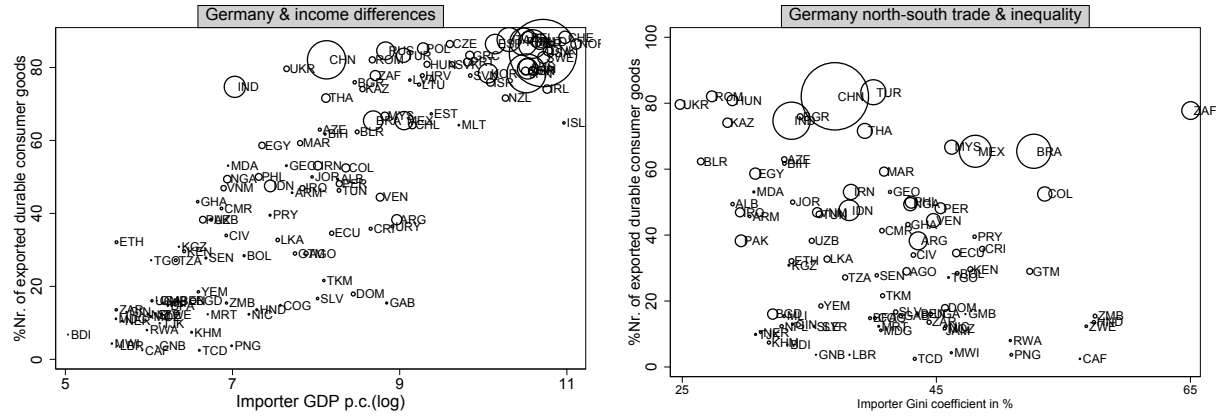

Figure 13: Germany and its destinations for durable consumer goods $pRATE_{ij}^{dcg}$

When looking at figure 14, the regression curve on the left is even steeper than in the U.S. suggesting a more profound effect of GDP p.c. on the extensive margin. While in the U.S., after controlling for GDP and trade costs, destinations can improve their extensive import margin by only around 10%, it is possible to improve the same margin in Germany by around 30%. On the right, our assumption that the extensive margin decrease in inequality seems to be true for all income groups. After controlling for GDP, GDP p.c. and trade costs, there is little variation left of what inequality would be able

to explain. The high extensive margin for Hungary, Bulgaria and Romania can be traced back to the EU-membership while Turkey has historically close ties to Germany (large diaspora of Turkish immigrants).

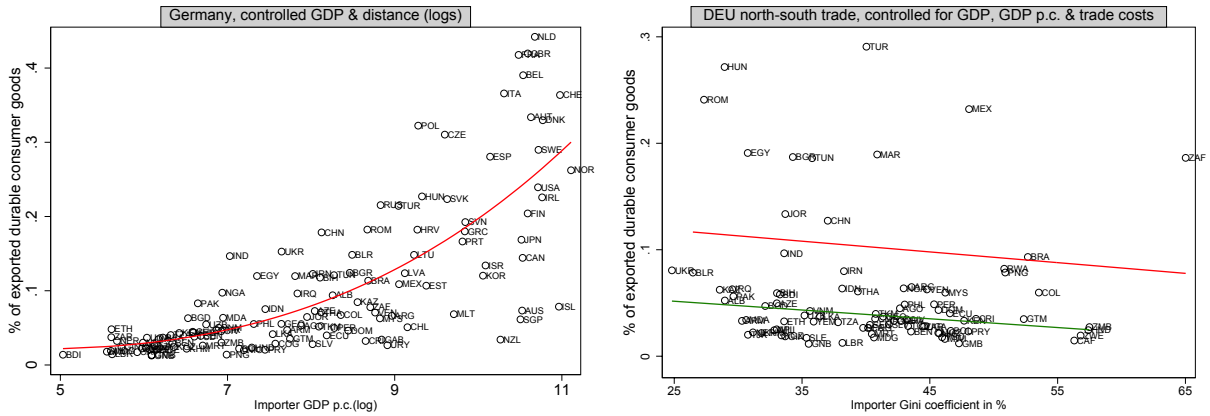

Figure 14: Percentage of traded durable consumer goods and inequality in DEU

China (CHN)

Also for China in the left of figure 15, we have the already known pattern for the effect of GDP p.c. on the extensive import margin of trade. Larger economies combined with higher GDP p.c. seem to import more varieties. On the right, the pattern looks similar as in the U.S for upper middle income countries (large circles) implying a moderate improvement of the extensive margin with higher inequality. The opposite seems to be true for smaller economies with high inequality.

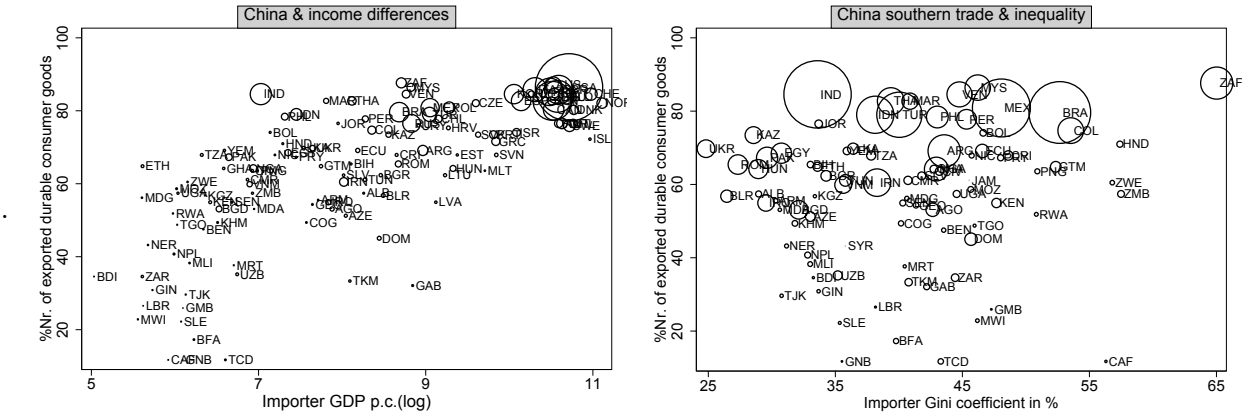

Figure 15: China and its destinations for durable consumer goods $pRATE_{ij}^{dcg}$

Income differences seem to play a minor role in comparison to the preceding countries, shown in figure 16. The regression line is considerably flatter. As we have seen in figure 10, CHN exports even to low income countries in around 38% product categories in the durable consumer good sector and 50% to lower & upper middle income countries. The high values reflect China's ability to reach even remote regions and provides a variety of goods to poorer countries. What is not tracked in the extensive margin of trade is the quality of the goods, as already investigated by Hallak (2010). Even though it

is controlled for "luxuries" using the durable consumer good sector, the per unit value of the products is not tracked in the extensive margin of trade. Therefore, some durable consumer goods exported from high income countries might be just too expensive for low income countries while China produces lower quality goods exporting them to all countries. China's imitations of northern products such as smartphones from Huawei etc. are well known.

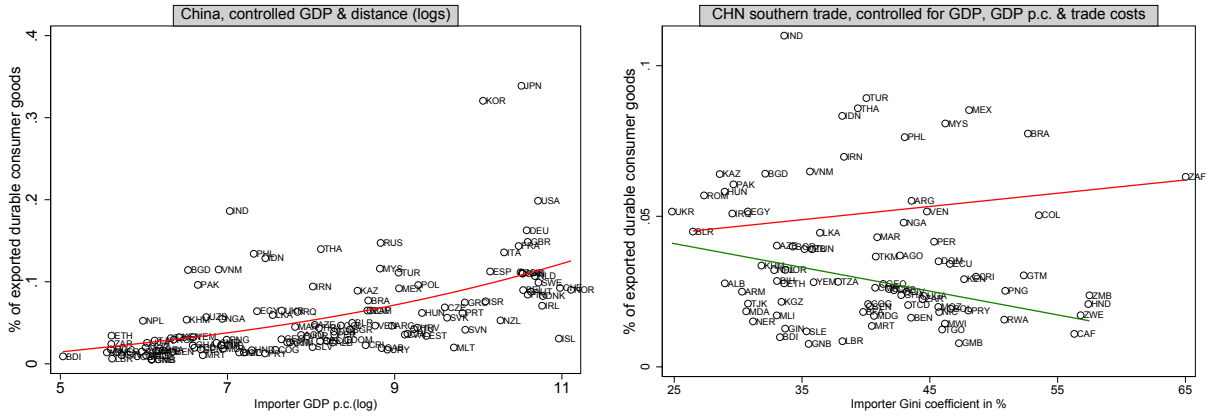

Figure 16: Percentage of traded durable consumer goods and inequality in CHN

On the right figure, upper middle income countries demand more variety with rising inequality (red line) as predicted by theory. On the other hand, the extensive import margin decreases in inequality for low & lower middle income countries (green line). Unfortunately, there is no simple explanation for the divergence. As we can see that the effect as a whole move only in between a range of a few percent, it can also just be a biased estimation as there is also still a lot of variation for the green line regression. The effect for upper middle income countries seems to be more straightforward.

India (IND)

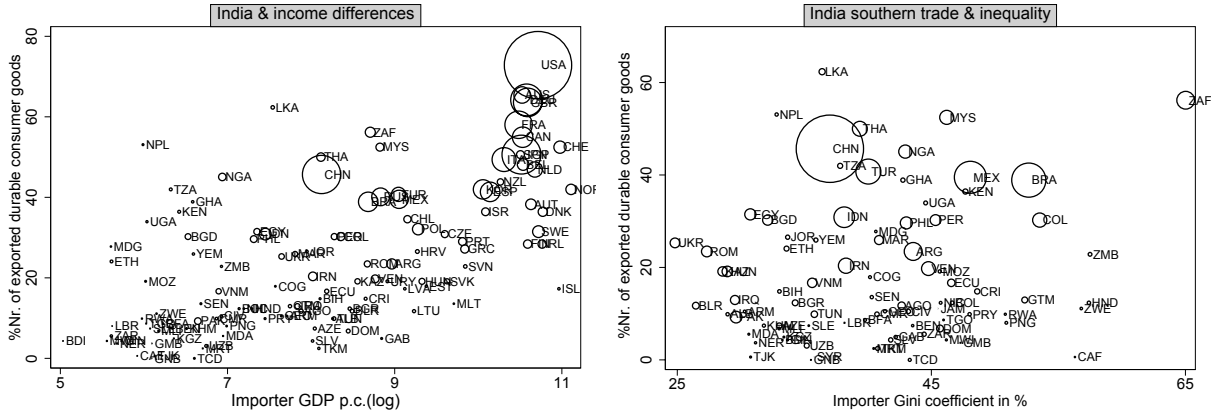

Figure 17: India and its destinations for durable consumer goods $pRATE_{ij}^{dcg}$

For the last country, India, we can confirm the aforementioned patterns of trade. Apart from countries of greater influence, everything seems to be in line with our predictions. Similar to China, we can see

in figure 18 to the right an ambiguous effect of the income distribution on the extensive margin of trade. Except from single outliers variation is high an the red regression line is flat while the green one is negative, again conflicting with our theoretical assumptions. In figure 17 to the left, some countries, which we can identify as neighbouring countries, step out of line. Nonetheless, the same countries still stand out even after controlling for trade costs and distance supporting our hegemony theory.

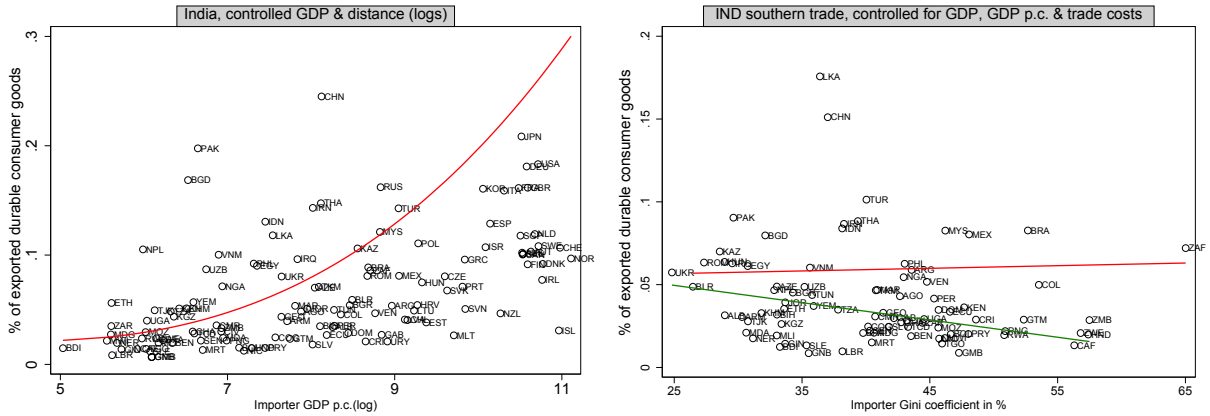

Figure 18: Percentage of traded durable consumer goods and inequality in IND

Wrap-up

Even though we controlled for proximity and gravity factors like size, distance, RTA etc., every country has outliers which are either neighbours or countries in the origin's area of influence supporting our hegemony assumption. USA trades with Mexico & Canada unexplainably high amounts of goods, while the same is true for China, Japan and Korea. Having a look at India reveals Sri Lanka, Pakistan, Bangladesh and Nepal as outliers. On the other hand, Germany seems to be completely in line with the predictions, mainly because members of the EU all have similar income levels and distributions. Moreover, we were able to show a positive effect of an importer's GDP p.c. on the extensive import margin of trade for all four countries. The results for the effect of the income distribution is rather ambiguous. While the trading partners of USA are in line with the theoretical predictions of *implication 4*, Germany, China and India destinations do not seem to behave as predicted. The effect of more inequality seems to impact the extensive margin rather negatively. In the following section, we will extend our analysis to the whole data set, as these was only a first exploratory analysis with rather contradictory results for giving an answer to *implication 4* and make a comparison to total products (*implication 5*).

5 Empirical analysis

5.1 Estimation problem and identification

Various estimation techniques have been proposed in previous works to estimate margins of trade and the gravity equation. It is essential to distinguish between models estimating the intensive, extensive and the quality margin of trade. Log-linear models using ordinary least squares (OLS) have long been treated as standard, estimating all three margins in a similar fashion (Dalgin et al., 2008; Eaton et al., 2004; Klenow & Hummels, 2005). Recent works cast doubt on the suitability of OLS for estimating trade flows especially in the presence of heteroskedasticity or "zeros". In particular, the log-linearisation of the DV leads to a loss of information because zero trade flows will be ignored.[15] Taking the logs and using OLS can lead to severely biased and inconsistent estimates as shown in the works of Gómez-Herrera (2013), Santos Silva & Tenreyro (2006) and Santos Silva et al. (2014).

To overcome the issue of "zeros", estimation techniques like a Poisson, a modified version of Poisson (PPML) or a Tobit regression have been considered, although with limited success (Gómez-Herrera, 2013, p. 1101). A two-step estimation method is considered superior accounting for the sample selection bias described by Heckman (1979, p. 160). To incorporate "zeros", a probit model is being used to estimate whether or not two countries trade with each other. In a second step, if countries actually engage in trade, the volume (or in our case the extensive product margin) is estimated with a standard log-linearised OLS model. Helpman et al. (2008) refined the heckman two-step estimation method incorporating firm heterogeneity, creating the HMR model. The empirically solid gravity equation estimated with HMR has the big advantage that market entry is simulated in the first step, and in the second, the decision of exporting volume is being made.

First and foremost, when estimating models it is crucial to consider the nature of the data. Our dependant variable (DV) is originally derived from a count whether or not a country exports in certain product categories or markets (also imports). After the product counts are aggregated to sectors, fractions have been calculated why the DV takes on values between [0,1]. In figures 22, 23 and 24 in *appendix B*, the histogram of the DVs are shown. It can be inferred from figure 22 and 23 that there are no "zeros" for $mRATE$ except in the durable consumer good sector. Another picture can be drawn from figure 24. In the DV $pRATE_{ij}$, up to 30% are "zeros" while the whole distribution is heavily skewed towards zero as already presented in figure 7.

Various estimation methods have been presented, but which one is the right choice? Well, none of the aforementioned methods appear to be adequate for a doubly-bounded dependent variable with a sizeable number of "zeros". In this thesis, we follow an estimation method first proposed by Papke & Wooldridge (1996) using Bernoulli quasi-likelihood estimators (this method is following abbreviated

[15] Log(0) is not defined and will be therefore ignored or treated as "missing" by common statistical software

by "PW"). Robustness checks will be conducted among others with a slightly modified version of PW allowing for flexible specification proposed by Santos Silva et al. (2014). For binary data, the log-likelihood function usually follows a Bernoulli-distribution as in equation (1).

$$lnL = \sum_{m=1}^{N} w_m y_m ln\{G(x'_m\beta)\} + w_m(1-y_m)ln\{1-G(x'_m\beta)\} \tag{1}$$

We assume an independent sequence of observations $\{(y_m, x_m) : m = 1, 2...N\}$, where $0 < y_m < 1$ while$m$ can be either $i = exporter$, $j = importer$ or $ij = bilateral$. In this respect, y_m is either mRATE or pRATE while x_m are the covariates of the explanatory variables. N is the sample size. For $G(\cdot)$ is usually a cumulative distribution function (cdf) chosen, taking on values $0 < G(z) < 1$. For our model, we decide for the standard normal cdf. The weights w_m are set equal to1. The following objective function is being maximised as follows:

$$l_m(b) = \sum_{m=1}^{N} y_m ln\{\Phi(x'_m b)\} + (1-y_m)ln\{1-\Phi(x'_m b)\} \tag{2}$$

or the optimisation problem in short,

$$\max_b \sum_{m=1}^{N} l_m(b) \tag{3}$$

Equation (4) shows the functional form of the estimated model. $\Phi(x'_m\beta)$ can be interpreted as the probability that a randomly drawn product category will be exported from origin i to destination j or a randomly drawn market will be served etc.

$$E\left(y_m|x_m\right) = \Phi(x'_m\beta) \tag{4}$$

While a linear model assumes partial effects to be constant, quasi-likelihood estimators only assume the conditional mean to be correctly specified stated in equation (4). Therefore, $\widehat{\beta}$ is consistent and $\sqrt{N}$-asymptotically normal ($n \to \infty$) regardless of the underlying distribution of y_m conditional on x_m (Papke & Wooldridge, 1996, p. 621-622).

The three equations estimated in this thesis are based on the gravity equation proposed by Tinbergen (1962). Additionally, next to the standard variables such as GDP for size and trade costs (originally bilateral distance), per capita incomes as well as Gini coefficients for equation (6) and (7) are incorporated according to the implications outlined in *section 2.5*.

$$E\left(mRATE_i|x_i\right) = \Phi\left(\beta_1 + \beta_2 log(GDPp.c._i) + \beta_3 log(GDP_i) + \gamma_1\tau_i\right) \tag{5}$$

$$E\left(mRATE_j|x_j\right) = \Phi\left(\beta_1 + \beta_2 log(GDPp.c._j) + \beta_3 log(GDP_j) + \beta_4 GINI_j + \beta_5 GINI_j^2 + \gamma_1\tau_j\right) \tag{6}$$

In equation (5), Gini coefficients are not needed because demand-side effects are subject to the demand-perspective in equation (6). For trade costs τ, fixed-country trade costs landlocked & cost to import/export in logs are being used. As proposed by Papke & Wooldridge (1996, p. 623), robust standard errors are used.

Equation (7) is in a dyadic setup in contrast to the previous equations. Hence, both importer and exporter effects are estimated simultaneously. Also next to the country specific fixed trade costs τ_i & τ_j, bilateral trade costs τ_{ij}[16] are added. This equation allows for estimation supply-side (GDP_i, GDP p.c.$_i$) and demand-side effects (GDP_j, GDP p.c.$_j$ & Gini).

$$\begin{aligned} E(pRATE_{ij}|x) = \Phi \big(&\beta_1 + \beta_2 \log(\text{GDPp.c.}_i) + \beta_3 \log(\text{GDPp.c.}_j) + \beta_5 \log(\text{GDP}_i) \\ &+ \beta_6 \log(\text{GDP}_j) + \beta_7 GINI_j + \beta_7 GINI_j^2 + \alpha_1 \tau_{ij} + \alpha_2 \tau_i + \alpha_3 \tau_j + \gamma_1 MTR_i + \gamma_2 MTR_j \big) \end{aligned} \tag{7}$$

The multilateral trade resistance term (MTR), measuring "remoteness", is a proxy for the difficulties not only a country faces in bilateral trade, but in a multilateral trade setting (Head & Mayer, 2000, p. 292). In equation (8) it can be seen that remoteness refers to the weighted distance where the weights are the partners' countries trading shares of world GDP. Since Anderson & van Wincoop (2003) proposed implementing country-fixed effects with dummy variables, this methodology seems to be rather outdated. Indeed, most works use dummy variables for each country now.

$$MTR_i = \sum_j \frac{distance_{ij}}{\left(GDP_j / GDP_{world}\right)} \tag{8}$$

In a cross-sectional data set where country-fixed effects like GDP and GDP p.c. are the variables of interest, country dummy variables cannot be included because the dummies would offset the effect of GDP etc.. Therefore, the MTR-proxy is used (Bacchetta et al., 2012, p. 110).

What is more, because we used the link-function $G(\cdot)$ to ensure that $\widehat{y_m}$ will be in the interval $[1, 0]$, the $\widehat{\beta}$-coefficients only tell us about the sign (direction) of the effect, but nothing of the size of the effect. In fact, the estimates depict $\frac{\beta}{\sigma}$, meaning that magnitudes are in standard deviation units of the errors. That is why average partial effects (APE) described by equation (9) are used. Taking the derivatives enables to interpret the effect to standard OLS-interpretations with one difference. The marginal effect is based on averages so that we obtain the effect $\beta\phi(\beta\bar{x})$.

$$\frac{\partial E(y_m|x_m)}{\partial x_m} = \beta\phi(\beta x_m) \tag{9}$$

Inferencing with paired or dyadic data as we have in our data set poses another challenge. Although Papke & Wooldridge (1996, p. 623) proofed the estimated variance to be asymptotically true using robust standard errors, we have to loose the assumption of independence across observations in a dyadic set up. The explanation is straightforward. Countries itself can be seen as independent from each other, but the bilateral trade flows within the countries are not (country-specific effects). Although two-way cluster robust standard errors are proposed by Cameron & Miller (2014), clustering only on the importer j is being used in this thesis to tackle intragroup correlation [17].

[16] τ_{ij} is a vector with the variables distance (logs) an the dummies RTA, common language and colonial relationship

[17] Unfortunately, a two-way cluster robust standard error is only available in the user-written STATA command '*ivreg2*' which is used in the robustness checks. As shown in the works of Gómez-Herrera (2013) and Santos Silva et al. (2014), OLS is biased which makes inference pointless anyway.

5.2 Results cross-sectional analysis

5.2.1 Percentage of markets

Table 1: Regression analysis on percentage of destinations served over sectors (mRATE_i)

	(1) mR_i^{Total}	(2) mR_i^{DCG}	(3) $mR_i^{Consumption}$	(4) $mR_i^{Capital}$	(5) $mR_i^{Intermediate}$
GDP p.c.$_i$ (log)	0.0192***	0.0194**	0.00988	0.0511***	0.0230***
	(3.14)	(2.20)	(1.28)	(5.82)	(3.17)
GDP$_i$ (log)	0.0472***	0.102***	0.0771***	0.0748***	0.0605***
	(8.66)	(15.16)	(12.00)	(10.47)	(10.27)
Landlocked	0.0537**	0.110***	0.0765***	0.0748*	0.0665**
	(1.98)	(3.44)	(2.61)	(1.91)	(2.16)
Cost to export (log)	-0.104***	-0.164***	-0.161***	-0.120***	-0.113***
	(-4.83)	(-6.70)	(-6.34)	(-4.25)	(-4.49)
Observations	123	123	123	123	123

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(i) Average partial effects (APE) $\frac{dy}{dx}$ of fractional probit response model

(ii) Robust standard errors

In table 1, we get to the bottom of *implication 1*. The estimates generally have the expected sign for GDP, GDP p.c. & costs to export, but a positive effect of landlocked seems rather surprising. To answer *implication 1*, it is worthwhile to examine GDP p.c. in more detail. The magnitude seems to be small at first sight with a 10% increase in GDP p.c. extending the destinations served by only 0.2% in total, and the DCG sector. The APE doubles for the capital good sector to 0.5%, which can be seen that for capital goods (heavy machinery, ships, aircrafts) more human capital as well as manufacturing capabilities are needed. This fact indicates that higher GDP p.c. can be indeed used as a proxy for capital already proposed by Helpman (1981). Therefore, a strong home market for capital goods can lead to more destinations served in this sector. The economical significance of GDP p.c. can take on high magnitudes considering that e.g. Ukraine has an average GDP p.c. of \$2'095 while the U.S. GDP p.c. is as high as \$45'000. Bear in mind that APE considers an average country across all variables. For the effect GDP on the rate of destinations served, the magnitudes are also as expected being in the range of 0.6-1% for a 10% increase of GDP. Size seems particularly important in the durable consumer good sector. According to our estimates, a landlocked country exports 5.4% more in total and around 11% more in the durable consumer good sector.

The H_0 hypothesis that the effect of the dummy landlocked is zero can be rejected on all conventional significance levels. A theoretically derived explanation for this cannot be provided. GDP p.c. is significant on all conventional levels in all sectors except in consumption goods while GDP is highly statistically significant as well in all sectors. That GDP p.c. seems to be irrelevant to determine the percentage of destinations served in the consumption goods is not particularly surprising. Consumption goods include products like fruits, vegetables or other products also abundant in lower income

countries. Therefore, these countries are also able to export these products to more countries. To sum up, we have found strong evidence in favour of *implication 1* that countries with higher GDP p.c. will export to a broader set of countries.

In table 2, the picture is flipped up-side down, when we consider the percentage of origins of the imported products (mRATE$_j$). Apart from the already known variables, we further include the Gini coefficient in % to account for the within-country income distribution similarly applied in the work of Dalgin et al. (2008). *Implication 2* states that if Gini rises, ceteris paribus, the percentage of import markets has to increase as well.

Table 2: Regression analysis on percentage of origins over sectors (mRATE$_j$)

	(1) mR_j^{Total}	(2) mR_j^{DCG}	(3) $mR_j^{Consumption}$	(4) $mR_j^{Capital}$	(5) $mR_j^{Intermediate}$
GDP p.c.$_j$ (log)	0.00349	0.0315***	0.0165*	0.0107	0.000471
	(0.38)	(3.66)	(1.84)	(1.13)	(0.05)
GDP$_j$ (log)	0.0447***	0.0545***	0.0445***	0.0592***	0.0605***
	(7.34)	(9.75)	(7.42)	(9.89)	(10.27)
Gini$_j^2$	-0.00111	0.0000962	-0.00233**	0.00105	-0.00122
	(-1.01)	(0.09)	(-2.10)	(0.96)	(-1.05)
Landlocked	0.00368	0.0251	-0.000355	0.0289	0.00837
	(0.14)	(1.06)	(-0.01)	(1.09)	(0.29)
Cost to import (log)	-0.0416**	-0.0163	-0.0268	-0.0224	-0.0451**
	(-2.08)	(-0.98)	(-1.51)	(-1.15)	(-2.06)
Observations	123	123	123	123	123

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(i) Average partial effects (APE) $\frac{dy}{dx}$ of fractional probit response model

(ii) Robust standard errors

(iii) Gini is originally quadratic, but due to APE now depicted to be linear

When we consider demand-side factors, GDP p.c. is only relevant in the durable consumer good sector, and on a smaller scale in the aggregated consumption goods sector. A rise in GDP p.c. by 10% leads to an increase of ca. 0.32% in the percentage of origins. The coefficient is statistically significant on a 1%-level and considering massive differences in GDP p.c. across countries, also economically significant. This fact supports Linder's notion that income matters. We can therefore conclude that *implication 1* is similarly valid so that countries with higher GDP p.c. import from a broader set of countries in the durable consumer good sector. Moreover, *implication 5* predicted that this effect will be present predominantly in the durable consumer good sector.

To investigate the effect of within-country income distributions in order to answer *implication 2*, we explore the variable of interest, the Gini coefficient. Because non-linearity is assumed (see figure 4), the quadratic term of the Gini is somewhat cumbersome to interpret. To avoid incomprehensible calculations, the average adjusted predictions (AAP)[18] are used. AAPs are the predicted conditional

[18]For all AAP-graphs, the user-written Stata command *mcp* has been used. The official Stata command *marginsplot* is less effective

means for $\bar{x}$-values. Hence, the average conditional means for $mRATE_j$ conditioned on average countries are predicted. We can infer from the figures 32 in *appendix C* that the effect is convex with a turning point at a Gini coefficient of around 45%. Assuming an average country with $\bar{x}_j$ in all variables except Gini, we can deduce from the figures that higher inequality can have a beneficial effect on $mRATE_j$. Depicted in figure 4, mostly countries in the upper middle income group would profit from these findings (high income countries do not have higher Gini values than 50%). The economic effect can also be seen as surprisingly high with a difference in the $\widehat{mRATE_j}$ of around 10%. Of course, these finding have to be enjoyed cautiously. The 95% confidence intervals indicates a wide range, especially after the turning point casting doubt on the statistical significance of Gini. Thus, we can not really tell whether Gini is to incorporate or not, but to exclude it from future analysis completely would be overhasty as well.

In this analysis, it could have been clearly shown that GDP p.c. does matter in the durable consumer goods for importers as well as in all sectors except consumption goods for exporter. We clearly favour *implication 1* to *implication 2*, although there is also some evidence that inequality does improve after a certain threshold the percentage of origins $mRATE_j$. In table 14 in *appendix D*, the analysis is extended to only high income origins leading to similar results.

5.2.2 Percentage of product categories

So far, the extensive margin of trade on country-level had been considered counting only trading pairs. In the following analysis, the data set is extended to a dyadic setup. The number of goods are counted and aggregated to bilateral trade flows exhibiting the extensive margin of trade. Using a fractional probit response model, it is attempted to find evidence supporting *implication 3*, *implication 4* and *implication 5*.

As shown in table 3, using PW for estimation, most variables do indeed have the expected sign. The trade cost variables τ_i & τ_{ij} behave as predicted from common theory and are significant on all conventional levels (Santos Silva et al., 2014; Bacchetta et al., 2012). One deviation are the country-specific fixed trade costs τ_i, measured in cost to import (logs). In this case, our proxy may have been inappropriate. Nevertheless, we still stick to the model because the signs are all in all correct (except consumption).

A very interesting effect is the one of colonial relationships. Country-pairs with a mutual history of colonialism engage in more extensive trade relationships trading in around 2.5-5% more product categories. Especially consumption goods seem to be influenced by common history. Most likely the direction is south-north providing more variety for high income consumers, although this is just an educated guess.The variables of interest will be, similar to the preceding analysis, depicted graphically as AAPs for the durable consumer good sector. For the other sectors, standard analysis is considered to be sufficient.

Table 3: Regression analysis on percentage of product categories over sectors (pRATE_{ij})

	(1) pR_{ij}^{Total}	(2) pR_{ij}^{DCG}	(3) $pR_{ij}^{Consumption}$	(4) $pR_{ij}^{Capital}$	(5) $pR_{ij}^{Intermediate}$
GDP p.c.$_i$ (log)	-0.000994	0.000250	-0.00663***	0.00776***	-0.000619
	(-1.39)	(0.27)	(-7.91)	(7.63)	(-0.91)
GDP p.c.$_j$ (log)	0.00107	0.00748***	0.00984***	-0.00181	-0.00192
	(0.80)	(4.20)	(6.11)	(-1.02)	(-1.43)
GDP$_i$ (log)	0.0358***	0.0397***	0.0383***	0.0443***	0.0342***
	(68.12)	(62.00)	(60.99)	(62.31)	(67.57)
GDP$_j$ (log)	0.0172***	0.0156***	0.0159***	0.0216***	0.0174***
	(18.98)	(14.04)	(13.84)	(18.22)	(19.37)
Gini_j^2	0.000162	0.000229	-0.0000161	0.000457	0.000207
	(0.61)	(0.71)	(-0.05)	(1.31)	(0.87)
Distance (log)	-0.0432***	-0.0447***	-0.0482***	-0.0511***	-0.0413***
	(-25.64)	(-21.45)	(-22.67)	(-22.32)	(-27.47)
RTA	0.0152***	0.0127***	0.0205***	0.0157***	0.0138***
	(5.57)	(3.77)	(6.06)	(3.94)	(5.51)
Common language	0.0213***	0.0284***	0.0242***	0.0287***	0.0188***
	(7.69)	(7.83)	(7.27)	(7.02)	(7.49)
Colonial relationship	0.0329***	0.0327***	0.0487***	0.0433***	0.0259***
	(5.39)	(4.85)	(6.11)	(4.93)	(4.96)
Cost to export (log)	-0.0352***	-0.0485***	-0.0472***	-0.0396***	-0.0303***
	(-34.19)	(-34.94)	(-33.32)	(-30.54)	(-31.06)
Cost to import (log)	-0.00408	-0.00104	0.000338	-0.00231	-0.00627***
	(-1.46)	(-0.31)	(0.09)	(-0.60)	(-2.60)
MTR$_i$	3.24e-12	-7.32e-11**	3.05e-11	-2.90e-11	2.23e-12
	(0.14)	(-2.34)	(1.00)	(-0.95)	(0.10)
MTR$_j$	3.40e-10***	3.29e-10***	3.38e-10***	4.32e-10***	3.38e-10***
	(8.13)	(5.31)	(6.26)	(7.32)	(8.89)
Observations	13806	13806	13806	13806	13806

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(i) Average partial effects (APE) $\frac{dy}{dx}$ of fractional probit response model used by Papke & Wooldridge (1996)

(ii) Variance estimator clustered on importer (j) to account for intragroup correlation. Countries itself are independent observations, but within the countries the bilateral trade flows are correlated.

(iii) Gini is originally quadratic, but due to APE now linear

First, we will have a look at GDP, accounting for size in the gravity literature. In figure 19, a strong effect on pRATE_{ij} is revealed. Especially exporter's GDP seems to be pivotal as a determinant of the extensive margin of trade in the durable consumer good sector. Not only is GDP highly statistical significant on all conventional levels, but also its magnitude is immense. A rising GDP, up to the high income group (log average = 26.3, max 30.28), will significantly improve the percentage of products traded bilaterally. An average country with a size in the 50% percentile can improve the extensive margin by 50% when having the size of a country in the 95% percentile, all other things at means. The importer's GDP effect is less pronounced, though a significant improvement is visible. The effect of GDP is predominant regardless of the sector. The magnitude as well as the statistical significance is similar across all sectors. The results support the home bias assumption as well as Linder's considerations that bigger markets have more variety to offer.

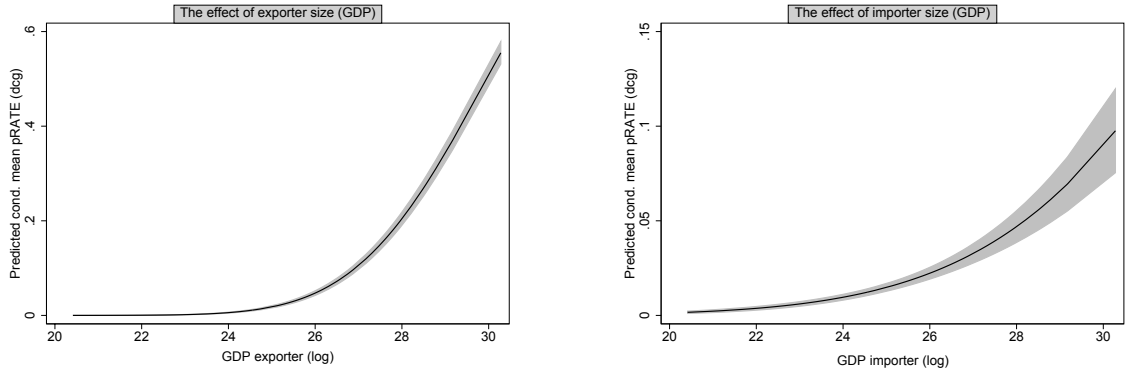

Figure 19: Average adjusted predictions for Size (GDP in logs)

Implication 3 states that exporter's and importer's GDP p.c. improves the extensive margin of trade and more variety will be traded bilaterally owing to supply and demand-side considerations. While we can see in figure 20 to the right a positive effect of GDP p.c. on demanded product variety, the supply-side effect looks rather bleak in the durable consumer good sector. Higher GDP p.c. in origins countries do not significantly improve traded variety. Moreover, the improvement of the APP of pRATE_{ij} is only around 2.5%, although being highly statistically significant.

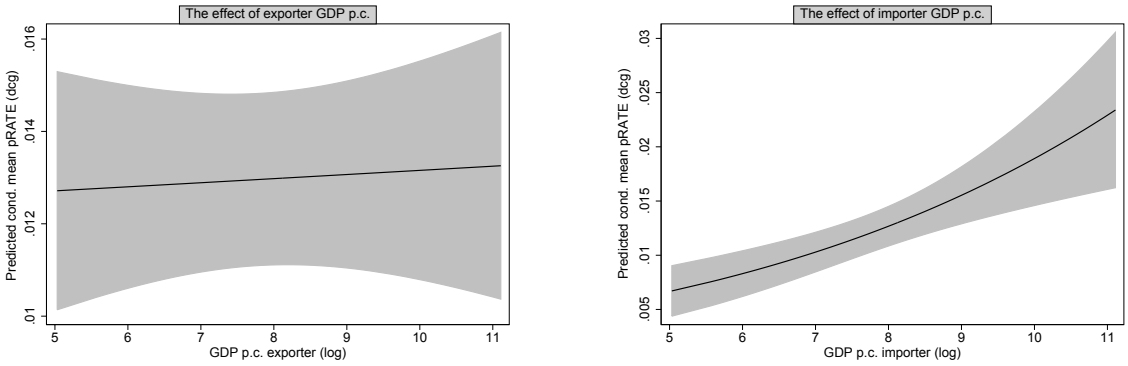

Figure 20: Average adjusted predictions for GDP p.c. in logs

More interesting are the effects of GDP p.c._i in the consumption and capital good sector. While in the latter the effect is positive, exporter's GDP p.c. from the exporter tends to lower the extensive margin in the consumption goods sector. GDP p.c._i is highly statistically significant on all conventional levels in both sectors. The counter-intuitive result is not surprising at a second look. As already mentioned, consumption goods include various goods of lower quality, low prices or just raw materials & agricultural products. Most of them are traditionally exported from southern to northern countries explaining the negative magnitude. In contrast, the production of capital goods presume a high level of development, capital and knowledge. Higher GDP p.c. had already been proposed as a proxy measure for capital. On the demand-side, GDP p.c._j has a positive effect on variety demanded supporting the Linder hypothesis. Also the inclusion of non-homothetic preferences proposed by Hunter (1991) is therefore justified.

We now examine *Implication 4* incorporating inequality within a country. In figure 21 we estimated the AAPs for the effect of the Gini coefficient on the pRATE_{ij} for an average country in the DCG and the capital good sector. Predicted effect is positive across both sectors. Coefficients are not significant in any sector why it is not surprising that in the capital & the durable consumer good sector the economical effect is likewise rather small. The difference of high to low inequality countries only accounts for a meagre 1-2% improvement of the extensive margin of trade. In addition, the range of the 95% confidence interval increases in Gini coefficients. Hence, the statistical significance cannot be confirmed. After this analysis, mixed feelings arise regarding the validity of *Implication 4*.

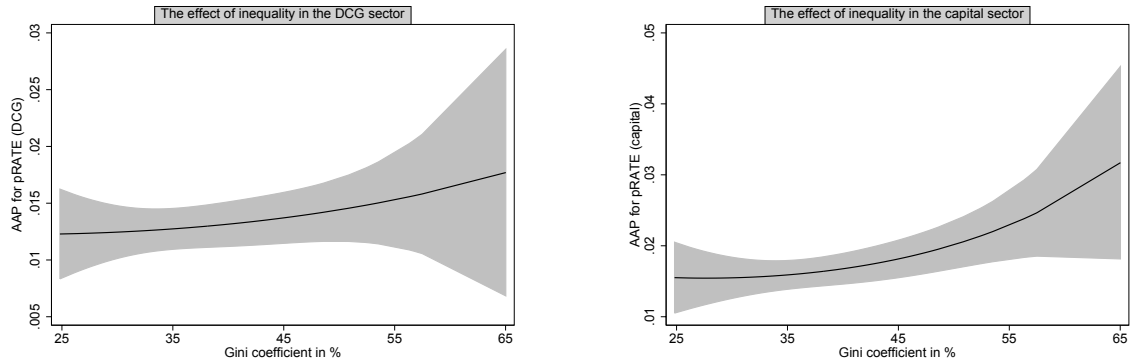

Figure 21: Average adjusted predictions for inequality

5.3 Robustness checks and critical reflection

Several estimation have been conducted in addition to test the presented models on validity. First things first, a model incorporating a non-linear Gini coefficient have been estimated in table 12, *appendix D*. Without a non-linear specification the APEs are slightly overestimated, though the signs are identical. Changing the proxy for inequality from the Gini coefficient to the decile ratio leads to similar results as shown in table 13 in *appendix D*. Also inference does not change.

There is still continuing research regarding methods to estimate bilateral trade margins, which are also divided into sectors. So far, no panacea have been found although PW estimates are considered superior to most comparable methods, namely OLS, Poisson, Heckman etc. (Santos Silva et al., 2014, p. 73). The flexible Bernoulli pseudo maximum likelihood method in Santos Silva et al. (2014, p. 69) is similar to our PW method, but with one difference. In equation (1) the weights w_m are now flexible and also maximised to achieve the correct specification while in PW we assumed the correct specification. Comparing the values from the flexible estimation in table 15 in *appendix D* with the estimates of PW in table 3, we can infer that our initially assumed specification was correct. Only for GDP p.c.$_i$ & $Gini_j$ PW estimates are slightly underestimated while the sign of $Gini_j$ in the consumption good sector even switches to the positive. Significance levels do not change at all.

Standard OLS estimates shown in table 16 in *appendix D* are exposed to a heavily upward bias. Zeros

are ignored in these estimates why the effect of GDP p.c.$_j$ is 10$\times$ the size of our PW estimates. As we have seen in figure 7, "zeros" are strongly correlated with GDP p.c. This fact violates the crucial assumption of OLS that $E(u|x) = 0$, meaning that the error term is uncorrelated to the explanatory variables. When zeros are left out in the estimation, this is exactly what causes an upward bias. In addition, the upper boundary is ignored completely which would be the same for all panel models. A two-way cluster robust standard error is used to account for intragroup correlation in the OLS regression. Little changes for significance levels affirming our assumption that cluster-robust standard errors only for importers are enough.

In this thesis, the Gini coefficient as a proxy for inequality seems to be attacked. It was difficult to find strong evidence favouring the assumption the inequality can improve the extensive margin of trade. Although our results did not really differ from the estimates including the decile ratio, this method is still questionable. Recent studies like Bernasconi (2013) or Martinez-Zarzoso & Vollmer (2011) attempt to model inequality as an overlap of income distributions. These bilateral methods are most likely superior to our method with just incorporating demand-side inequality. For future studies it is therefore suggested to follow the previous examples rather than the example of Dalgin et al. (2008). Another neglected issue is the time dimension. Halter et al. (2014) states that inequality can promote growth in the short-run, but is harmful in the long-run. Therefore, this cross-sectional analysis has to be read with utmost caution. The picture that is made is a one-time shot and variables change over time, especially per capita incomes.

6 Concluding remarks

After a broad literature review, five implications for the effects of income differences and distributions on the patterns of trade have been derived. The conceived implications have been investigated in a rigorous, fact-based fashion. Evidence have been found in favour of these implications. Our findings are summarised as follows:

- *Implication 1*: Exporter's GDP p.c. improves the percentage of markets served except in the consumption good sector. A significant statistical relationship has been proven. Considering the immense income differences across countries, the relatively low average partial effects can also have an economic significance. On the demand-side, only in the consumption and the durable consumer good sector a statistical and economical significance for importer's GDP p.c. has been found.

- *Implication 2*: The linear as well as the non-linear specification of the Gini coefficient revealed little evidence that inequality can enhance the percentage of origins. No significant relationship could have been observed. Although the turning point at 45% indicates an improvement in the capital and durable consumer goods sector, the economic effect lies only in the range of 1-2%.

- *Implication 3*: The analysis confirms a statistically significant and positive effect for importer GDP p.c. in the consumption and durable consumer good sector, while ambiguous results have been obtained in the other sectors. These findings seem to confirm the Linder hypothesis and are of great interest, even though the economic effect is rather small. Regarding exporter GDP p.c., a statistically significant relationship have been found only in the consumption and capital good sector. While in the latter the effect of GDP p.c$_i$ is positive, we obtain negative APEs in the consumption good sector.

- *Implication 4*: Similarly to *Implication 2*, there is only a marginal increase of the extensive margin for Gini coefficients of 45%. A lot of variance is still present and no statistical relationship could have been observed, although the signs point in the right direction.

- *Implication 5*: The last implication, that described effects are more pronounced in the durable consumer good sector, can be confirmed for import's GDP p.c. and the percentage of traded product categories. Furthermore, the same effect has been observed for the percentage of origins. Hence, the demand-side perspective should not be neglected and non-homothetic preferences seems to capture reality much better, as demand for certain goods indeed depend on income per capita.

We presented an extensive framework including both supply and demand-side considerations. Higher GDP p.c. seems to promote variety in trade along products and markets. Richer countries have more extended trade networks both in import and export markets and trade in more product categories. Exceptions like China and India reach the same extensive margin of trade merely because of their economic mass. The weak evidence that inequality improves the extensive margin seems to be pivotal for future policy. Long-term growth may be negatively affected by severe inequality and politics should rather focus on GDP per capita. Of course, our sectoral analysis proofed to be superior to a total-product analysis. This method allowed to differ between various sectors as also diverse, sector-specific effects have been observed.

Caveats of our analysis are first and foremost the proxies for inequality. The Gini coefficient may not be adequate in capturing the whole complexity of inequality. Recent research focus more on individual bilateral overlaps of the respective income distributions (Bernasconi, 2013; Martinez-Zarzoso & Vollmer, 2011). Moreover, it would be interesting to track inequality and its effects over time similarly to the work of Halter et al. (2014), but specifically investigating the extensive margin of trade.

When Adam Smith called for a more equal society, he had specifically England in mind. As the world grows together, rich countries have to understand that their fate is inherently bound to the fate of less fortunate countries. Economic growth seems to be the panacea, although we have to keep in mind that quality growth measured in GDP p.c. is of utmost importance as well. Also the 99% within a country, living in the best houses, enjoying the best education should always be aware that wealth is ephemeral.

References

Anderson, J. E., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, *105*, 709–751.

Arkolakis, C. (2010). Market Penetration Costs and the New Consumers Margin in International Trade Costas Arkolakis. *Journal of Political Economy*, *118*(6), 1151–1199.

Atkeson, A., & Burstein, A. (2008, nov). Pricing-to-Market, Trade Costs, and International Relative Prices. *American Economic Review*, *98*(5), 1998–2031.

Bacchetta, M., Beverelli, C., Cadot, O., Fugazza, M., Helble, J.-M., Nicita, A., & Piermartini, R. (2012). A Practical Guide to Trade Policy Analysis. *World Trade Organization: United Nations Conference on Trade and Development*.

Baldwin, R., & Harrigan, J. (2011). Zeros, Quality, and Space: Trade Theory and Trade Evidence. *American Economic Journal: Microeconomics*, *3*(2), 60–88.

Bergstrand, J. H. (1990). The Heckscher-Ohlin-Samuelson Model, The Linder Hypothesis and the Determinants of Bilateral Intra-Industry Trade. *The Economic Journal*, *100*(403), 1216–1229.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2007). Firms in International Trade. *Journal of Economic Perspectives*, *21*(3), 105–130.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2012). The Empirics of Firm Heterogeneity and International Trade. *Annual Review of Economics*, *4*(1), 283–313.

Bernasconi, C. (2013). *Similarity of Income Distributions and the Extensive and Intensive Margin of Bilateral Trade Flows (UZH Working Paper Series Nr. 115).* Retrieved from `www.econ.uzh.ch/eiit/Papers[accessed:31.08.2015]`

Bernasconi, C., & Wuergler, T. (2013). *Per Capita Income and the Quality and Variety of Imports (UZH Working Paper Series).* Retrieved from `www.econ.uzh.ch/eiit/Papers[accessed:30.06.2015]`

Bertola, G., Foellmi, R., & Zweimüller, J. (2006). *Income Distribution in Macroeconomic Models.* Princeton: Princeton University Press.

Bond, E. W., Iwasa, K., & Nishimura, K. (2010). A dynamic two country HeckscherOhlin model with non-homothetic preferences. *Economic Theory*, *48*(1), 171–204.

Bowen, H. P., Leamer, E. E., & Sveikauskas, L. (1987). Multicountry, Multifactor Tests of the Factor Abundance Theory. *American Economic Review*, *77*(5), 791–809.

Cameron, A. C., & Miller, D. L. (2014). *Robust Inference for Dyadic Data (Working Paper UC Davis).* Retrieved from cameron.econ.ucdavis.edu/research[Accessed:10.09.2015]

Chaney, T. (2008, aug). Distorted Gravity: The Intensive and Extensive Margins of International Trade. *American Economic Review*, *98*(4), 1707–1721.

Dalgin, M., Trindade, V., & Mitra, D. (2008). Inequality, Nonhomothetic Preferences, and Trade: A Gravity Approach. *Southern Economic Journal*, *74*(3), 747–774.

Das, M. (2008). Absolute and comparative advantage. In W. A. Darity (Ed.), *International encyclopedia of the social sciences* (2nd ed., pp. 1–2). Detroit: Macmillan.

Davis, D. R., & Mishra, P. (2007). Stolper-Samuelson is dead: And other crimes of both theory and data. In A. Harrison (Ed.), *Globalization and poverty* (pp. 87–107). Chicago: University of Chicago Press.

Deardorff, A. (1998). *Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World? (NBER Working Paper No. W5377).* Retrieved from www.nber.org/chapters/c7818.pdf[accessed:15.06.2015]

Dinopoulos, E., Fujiwara, K., & Shimomura, K. (2011). International Trade and Volume Patterns under Quasilinear Preferences. *Review of Development Economics*, *15*(1), 154–167.

Eaton, J., Kortum, S., & Kramarz, F. (2004). Dissecting Trade: Firms, Industries, and Export Destinations. *American Economic Review*, *94*(2), 150–154.

Evenett, S. J., & Keller, W. (2002). On Theories Explaining the Success of the Gravity Equation. *Journal of Political Economy*, *110*(2), 281–316.

Fajgelbaum, P., Grossman, G. M., & Helpman, E. (2011). Income Distribution, Product Quality, and International Trade. *Journal of Political Economy*, *119*(4), 721–765.

Foellmi, R., Hepenstrick, C., & Zweimueller, J. (2013). *International arbitrage and the extensive margin of trade between rich and poor countries (UZH Working Paper Series).* Retrieved from www.econ.uzh.ch/eiit/Papers[accessed:09.10.2014]

Francois, J. F., & Kaplan, S. (1996). Aggregate Demand Shifts, Income Distribution, and the Linder Hypothesis. *The Review of Economics and Statistics*, *78*(2), 244–250.

Gaulier, G., & Zignago, S. (2010). *International trade price indices (CEPII Working Paper Nr. 2010-23).* Retrieved from www.banque-france.fr[accessed:20.08.2015]

Gómez-Herrera, E. (2013). Comparing alternative methods to estimate gravity models of bilateral trade. *Empirical Economics*, 1087–1111.

Hallak, J. C. (2010). A Product-Quality View of the Linder Hypothesis. *Review of Economics and Statistics*, *92*(3), 453–466.

Halter, D., Oechslin, M., & Zweimueller, J. (2014). Inequality and growth: The neglected time dimension. *Journal of Economic Growth*, *19*(1), 81–104.

Haveman, J., & Hummels, D. (2004). Alternative hypotheses and the volume of trade: The gravity equation and the extent of specialization. *Canadian Journal of Economics*, *37*(1), 199–218.

Head, K., & Mayer, T. (2000). Non-Europe: The magnitude and causes of market fragmentation in the EU. *Weltwirtschaftliches Archiv*, *136*, 284–314.

Head, K., & Mayer, T. (2013). Gravity Equations: Toolkit, Cookbook, Workhorse. In Gopinath, Helpman, Rogoff, & Elsevier. (Eds.), *Handbook of international economics* (4th ed.).

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *The Econometric Society*, *47*(1), 153–161.

Helpman, E. (1981). International trade in the presence of product differentiation, economies of scale and monopolistic competition:: A Chamberlin-Heckscher-Ohlin approach. *Journal of International Economics*, *11*(3), 305–340.

Helpman, E., Melitz, M., & Rubinstein, Y. (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. *The Quarterly Journal of Economics*, *123*(2), 441–487.

Hepenstrick, C., & Tarasov, A. (2014). *Per capita income and the extensive margin of bilateral trade (LMU Working Paper).* Retrieved from `www.iwb.econ.uni-muenchen.de/forschung/veroffentlichungen[accessed:02.04.2015]`

Hunter, L. (1991, may). The contribution of nonhomothetic preferences to trade. *Journal of International Economics*, *30*, 345–358.

Klenow, P. J., & Hummels, D. (2005). The Variety and Quality of a Nation's Exports. *American Economic Review*, *95*(3), 704–723.

Krugman, P. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of International Economics*, *9*(4), 469–479.

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

Leontief, W. (1953). Domestic Production and Foreign Trade: The American Capital Position Re-Examined. *Proceedings of the American philosophical Society*, *97*(4), 332–349.

Linder, S. B. (1961). *An Essay on Trade and Transformation.* Uppsala: Almqvist & Wiksells Boktryckeri AB.

Markusen, J. R. (2013, jul). Putting per-capita income back into trade theory. *Journal of International Economics*, *90*(2), 255–265.

Martinez-Zarzoso, I., & Vollmer, S. (2011). *Bilateral Trade Flows, Inequality and Income-Distribution Similarity (Working Paper).* Retrieved from `http://works.bepress.com/inma{_}martinez{_}zarzoso/22/[accessed:10.10.2015]`

Maskus, K. E. (1985). A test of the Heckscher-Ohlin-Vanek theorem: The Leontief commonplace. *Journal of International Economics*, *19*, 201–212.

Mayer, T., & Zignago, S. (2011). *Notes on CEPIIs distances measures: The GeoDist database (CEPII Working Paper Nr. 2011-25).* Retrieved from `www.cepii.fr[accessed:10.08.2015]`

Melitz, M. (2003). The Impact of Trade in Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, *71*(6), 1695–1725.

Mitra, D., & Trindade, V. (2005, nov). Inequality and trade. *Canadian Journal of Economics*, *38*(4), 1253–1271.

Murphy, K. M., Shleifer, A., & Vishny, R. (1989). Income distribution, market size, and industrialization. *The Quarterly Journal of Economics*, *104*(3), 537–564.

Papke, L. E., & Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*, *11*(6), 619–632.

Ramezzana, P. (2000). *Per capita income demand for variety, and international trade: Linder reconsidered (LSE Working Paper).* Retrieved from `eprints.lse.ac.uk/20184[accessed:04.05.2015]`

Ricardo, D. (1817). *On the Principles of Political Economy and Taxation* (3rd ed.). Kitchener: Batoche Books.

Roser, M. (2015). *Income inequality: poverty falling faster than ever but the 1% are racing ahead.* Retrieved from `http://www.theguardian.com/news/datablog/2015/mar/27/income-inequality-rising-falling-worlds-richest-poorest[accessed:10.09.2015]`

Rybczynski, T. M. (1955). Factor endowments and relative commodity prices. *Economica*, *22*(88), 336–341.

Samuelson, P. A. (1948). International Trade and the Equalisation of Factor Prices. *The Economic Journal*, *58*(230), 163–184.

Santos Silva, J., & Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, *88*(5), 641–658.

Santos Silva, J., Tenreyro, S., & Wei, K. (2014). Estimating the extensive margin of trade. *Journal of International Economics*, *93*(1), 67–75.

Stern, R. M., & Maskus, K. E. (1981). Determinants of the Structure of U.S. Foreign Trade, 1958-76. *Journal of International Economics*, *11*, 207–224.

Stolper, W., & Samuelson, P. (1941). Protection and real wages. *Review of Economic studies*, *9*(1), 58–73.

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

Trefler, D. (1995). The Case of the Missing Trade and Other Mysteries: Reply. *American Economic Review*, *85*(5), 1029–1046.

United Nations. (2002). *Complete HS and SITC conversion and correspondence tables.* Retrieved from `http://unstats.un.org/unsd/trade/conversions/HSCorrelationandConversiontables.htm[accessed:15.07.2015]`

United Nations. (2014). *Country classification.* Retrieved from `www.un.org/en/development/desa/policy/wesp/wesp{_}current/2014wesp{_}country{_}classification.pdf[accessed:09.08.2014]`

Worldbank. (2014). *World Development Indicators [Dataset].* Retrieved from `data.worldbank.org/data-catalog/world-development-indicators[accessed:27.11.2014]`

Appendices

A Information to data

Product category	*HS02*
Durable consumer goods	162
Consumption goods	1265
Capital goods	674
Intermediate goods	3145
Total	5222

Table 4: Nr. of products in HS classification

Data	**Variables**	**Source**
Percentage. of products/markets	$mRATE_i$= Pc. of destinations $mRATE_j$= Pc. of origins nm_i^{dcg} =Nr. of dcg destinations $mRATE_j^{low}$ =Pc. of low inc. origins $mRATE_i^{dcg\|high}$ =Percent of high inc. destinations served $pRATE_{ij}$ = Pc. of products $np_{ij}^{consumption}$ = Pc. of consumption products	Gaulier & Zignago (2010) for underlying prepared U.N. Comtrade data, disaggregated to HS6-digit level
Income data	GDP, GDP p.c., Gini coefficients, Income shares (quintiles)	Worldbank (2014)
Income groups	*Low income* = < $1'035 *Lower middle income* = $1'036 < x < $4'085 *Upper middle income*= $4'086< x < $12'615 *High income*= >n $12'615 by GNI per capita 2012	United Nations (2014)
Bilateral gravity	distw = weighted distance colony = colonial relationship rta = regional trade agreement comleg= common legal origin comcur= common currency	Head & Mayer (2013)
Unilateral gravity	International distance, landlocked dummy	Mayer & Zignago (2011)
Trade costs	Next to gravity data: Nr. of import/export documents cost to import/export ($ per container) cost to start up business (index in %) time to import/export (days)	Worldbank (2014)
BEC	Categories: Total, consumption, durable consumer goods, capital goods, intermediate goods	United Nations (2002)

Table 5: Variables explained in detail

Country	iso3	Country	iso3
Angola	AGO	Kenya	KEN
Albania	ALB	Kyrgyz Republic	KGZ
Argentina	ARG	Cambodia	KHM
Armenia	ARM	Korea, Rep.	KOR
Australia	AUS	Liberia	LBR
Austria	AUT	Sri Lanka	LKA
Azerbaijan	AZE	Lithuania	LTU
Burundi	BDI	Latvia	LVA
Belgium	BEL	Morocco	MAR
Benin	BEN	Moldova	MDA
Burkina Faso	BFA	Madagascar	MDG
Bangladesh	BGD	Mexico	MEX
Bulgaria	BGR	Mali	MLI
Bosnia and Herzegovina	BIH	Malta	MLT
Belarus	BLR	Mozambique	MOZ
Bolivia	BOL	Mauritania	MRT
Brazil	BRA	Malawi	MWI
Central African Republic	CAF	Malaysia	MYS
Canada	CAN	Niger	NER
Switzerland	CHE	Nigeria	NGA
Chile	CHL	Nicaragua	NIC
China	CHN	Netherlands	NLD
Cote d'Ivoire	CIV	Norway	NOR
Cameroon	CMR	Nepal	NPL
Congo, Rep.	COG	New Zealand	NZL
Colombia	COL	Pakistan	PAK
Costa Rica	CRI	Peru	PER
Czech Republic	CZE	Philippines	PHL
Germany	DEU	Papua New Guinea	PNG
Denmark	DNK	Poland	POL
Dominican Republic	DOM	Portugal	PRT
Ecuador	ECU	Paraguay	PRY
Egypt, Arab Rep.	EGY	Romania	ROM
Spain	ESP	Russian Federation	RUS
Estonia	EST	Rwanda	RWA
Ethiopia	ETH	Senegal	SEN
Finland	FIN	Singapore	SGP
France	FRA	Sierra Leone	SLE
Gabon	GAB	El Salvador	SLV
United Kingdom	GBR	Slovak Republic	SVK
Georgia	GEO	Slovenia	SVN
Ghana	GHA	Sweden	SWE
Guinea	GIN	Syrian Arab Republic	SYR
Gambia, The	GMB	Chad	TCD
Guinea-Bissau	GNB	Togo	TGO
Greece	GRC	Thailand	THA
Guatemala	GTM	Tajikistan	TJK
Honduras	HND	Turkmenistan	TKM
Croatia	HRV	Tunisia	TUN
Hungary	HUN	Turkey	TUR
Indonesia	IDN	Tanzania	TZA
India	IND	Uganda	UGA
Ireland	IRL	Ukraine	UKR
Iran, Islamic Rep.	IRN	Uruguay	URY
Iraq	IRQ	United States	USA
Iceland	ISL	Uzbekistan	UZB
Israel	ISR	Venezuela, RB	VEN
Italy	ITA	Vietnam	VNM
Jamaica	JAM	Yemen, Rep.	YEM
Jordan	JOR	South Africa	ZAF
Japan	JPN	Congo, Dem. Rep.	ZAR
Kazakhstan	KAZ	Zambia	ZMB
		Zimbabwe	ZWE

Table 6: Countries covered in detail (125 countries in total)

B Descriptive statistics

	mean	sd	min	max
nm_i	105.008	20.99558	34	124
nm_i^{dcg}	63.48	39.24892	0	124
$nm_i^{consumption}$	92.952	30.1006	13	124
$nm_i^{capital}$	81.528	36.05139	7	124
$nm_i^{intermediate}$	98.304	25.29	25	124
$mRATE_i$	84.68387	16.93192	27.41936	100
$mRATE_i^{dcg}$	51.19355	31.65235	0	100
$mRATE_i^{consumption}$	74.96129	24.27468	10.48387	100
$mRATE_i^{capital}$	65.74839	29.0737	5.645161	100
$mRATE_i^{intermediate}$	79.27742	20.39516	20.16129	100
nm_j	105.008	17.42194	55	124
nm_j^{dcg}	63.48	22.24446	19	115
$nm_j^{consumption}$	92.952	19.25502	44	123
$nm_j^{capital}$	81.528	20.21002	32	124
$nm_j^{intermediate}$	98.304	20.31654	42	124
$mRATE_j$	84.68387	14.04996	44.35484	100
$mRATE_j^{dcg}$	51.19355	17.93908	15.32258	92.74194
$mRATE_j^{consumption}$	74.96129	15.52824	35.48387	99.19355
$mRATE_j^{capital}$	65.74839	16.2984	25.80645	100
$mRATE_j^{intermediate}$	79.27742	16.3843	33.87097	100
N	125			

Table 7: Summary nm_i, nm_j, $mRATE_i$ & $mRATE_j$

	mean	sd	min	max
GDP p.c.	10594.26	15463.2	153.1431	66825.02
GDP	4.29e+11	1.45e+12	7.35e+08	1.41e+13
Gini coefficient in %	38.09963	8.13116	24.82	65.02
N	125			

Table 8: Summary income

	mean	sd	min	max
Documents to import	7.379032	2.858713	2	17
Cost to import (US$ per container)	1926.476	1593.299	420	9800
Time to import (days)	25.37097	20.15568	4	112
Cost of business (% of GNI per capita)	33.74919	48.62487	0	284.7
Documents to export	6.185484	2.253694	2	13
Cost to export (US$ per container)	1605.944	1218.264	435	8450
Time to export (days)	22.56452	16.34747	6	81
N	125			

Table 9: Summary fixed trade costs

	mean	sd	min	max
np_{ij}	362.6551	749.8811	0	4647
np_{ij}^{dcg}	12.52006	26.4055	0	150
$np_{ij}^{consumption}$	100.3017	200.1608	0	1166
$np_{ij}^{capital}$	61.65652	120.9782	0	637
$np_{ij}^{intermediate}$	198.522	435.1949	0	2877
$pRATE_{ij}$	6.944755	14.36004	0	88.98889
$pRATE_{ij}^{dcg}$	7.728435	16.29969	0	92.59259
$pRATE_{ij}^{consumption}$	7.928991	15.82298	0	92.17391
$pRATE_{ij}^{capital}$	9.147851	17.94928	0	94.51038
$pRATE_{ij}^{intermediate}$	6.312305	13.83768	0	91.47854
N	15500			

Table 10: Summary np_{ij} & $pRATE_{ij}$

	mean	sd	min	max
Distance (weighted, pop-wt, km)	7239.909	4243.49	114.6373	19650.13
Regional trade agrement	.1254821	.3312762	0	1
Contiguity	.0252066	.1567576	0	1
Common legal origin	.3115702	.4631512	0	1
Common currency	.0172176	.1300859	0	1
Common official language	.1008264	.3011091	0	1
N	14520			

Table 11: Summary variable trade costs

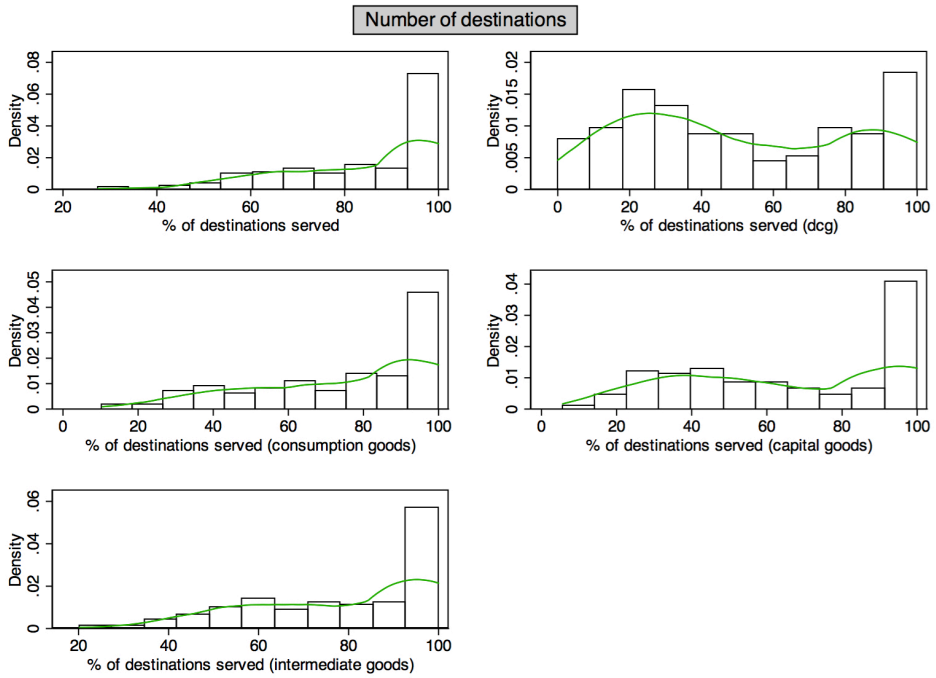

Figure 22: Histogram percentage of destinations served ($mRATE_i$)

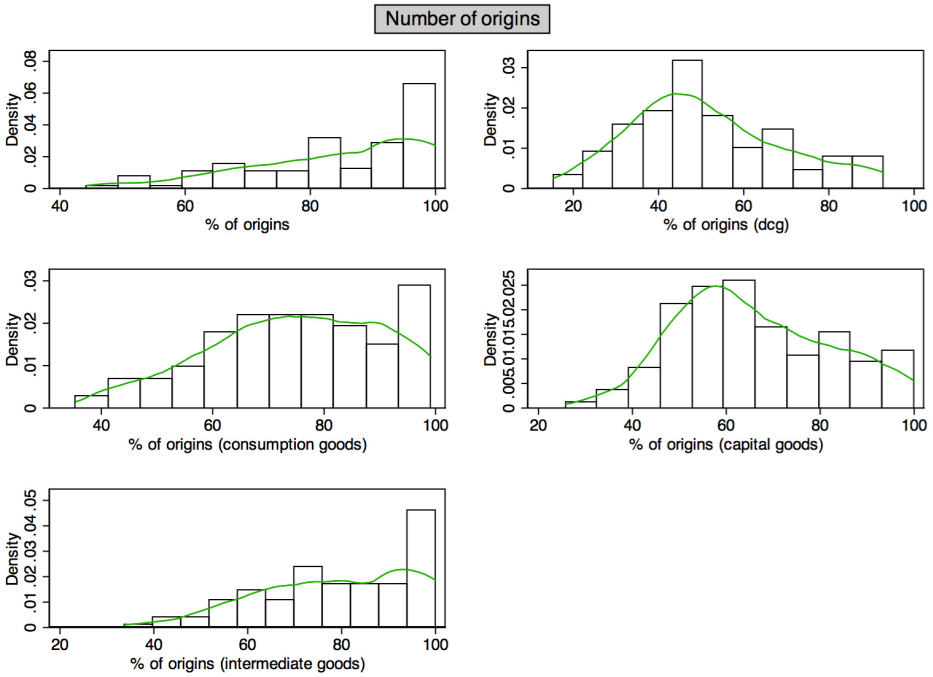

Figure 23: Histogram percentage of origins ($mRATE_j$)

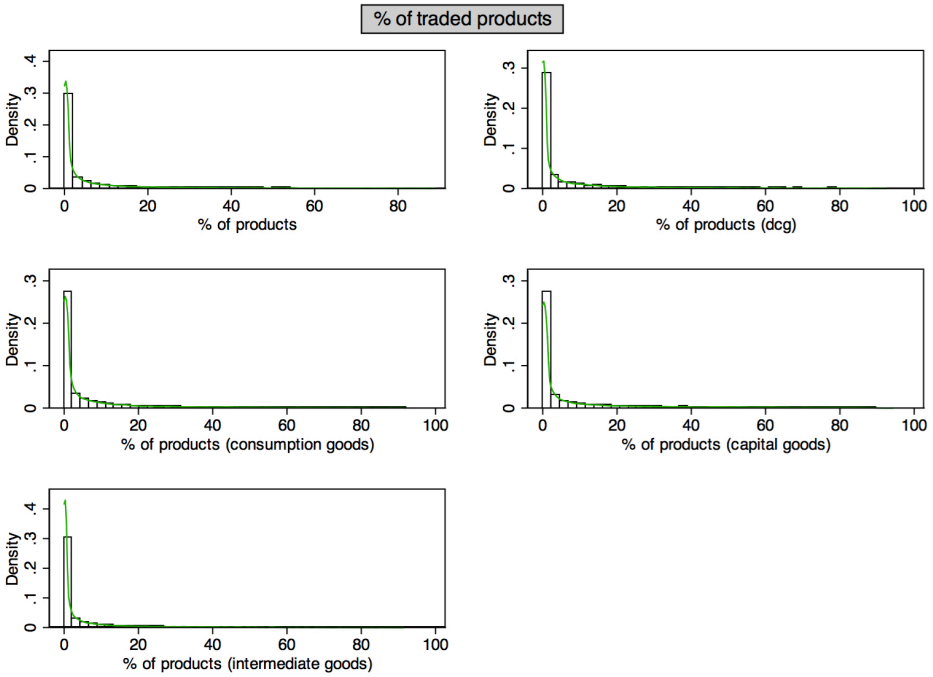

Figure 24: Histogram percentage of exported products ($pRATE_{ij}$)

C Additional figures

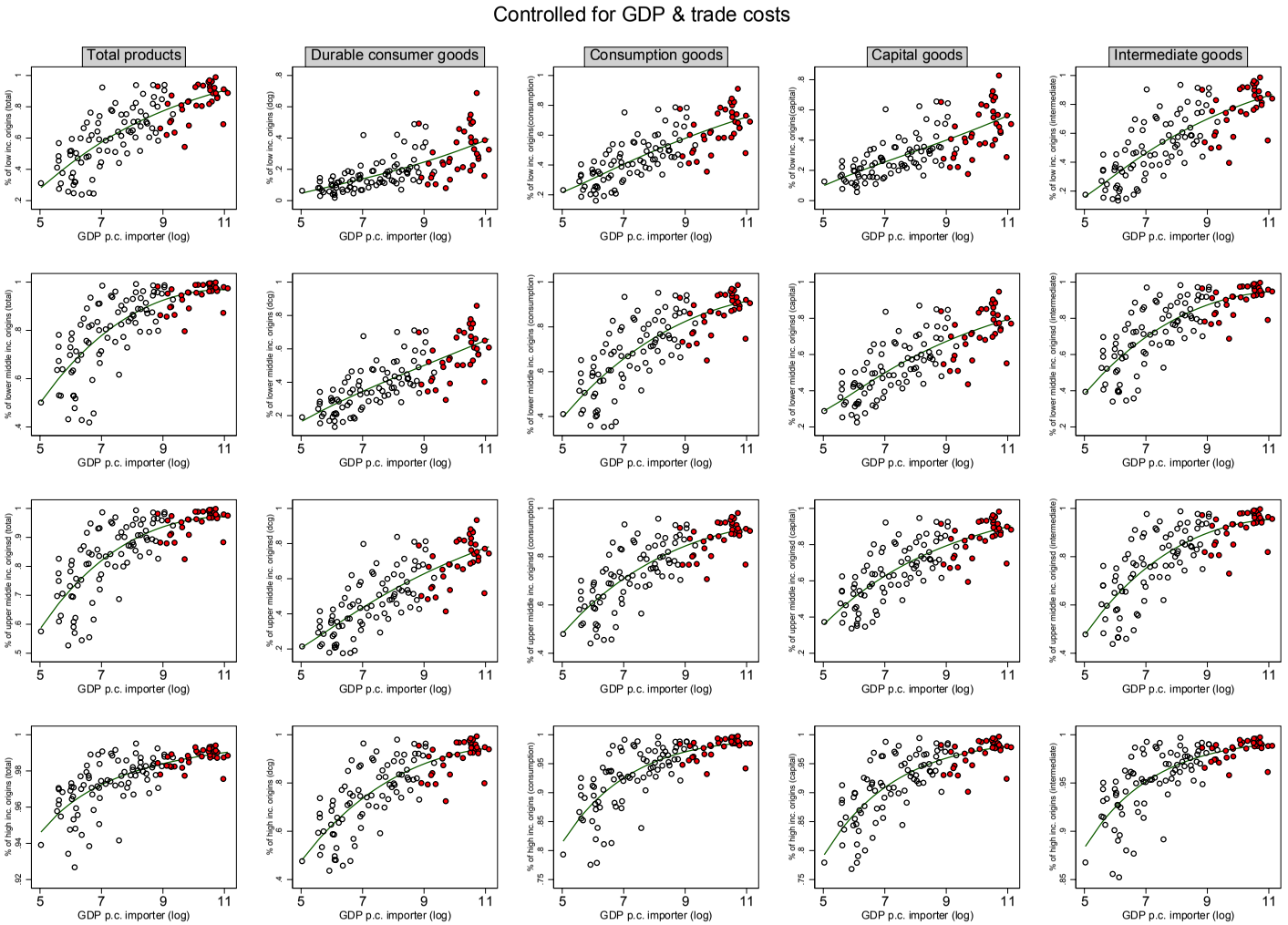

Figure 25: Percentage of origins across all product categories (BEC)

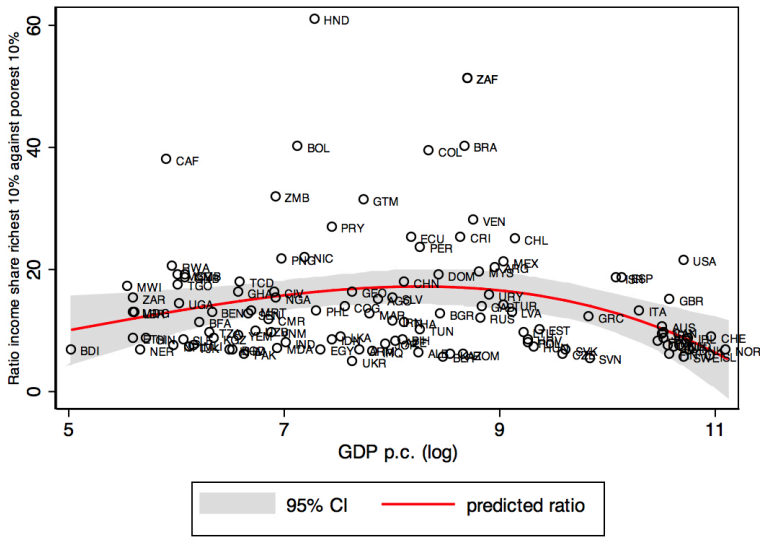

Figure 26: Income share ratio highest decile/lowest decile and GDP p.c.

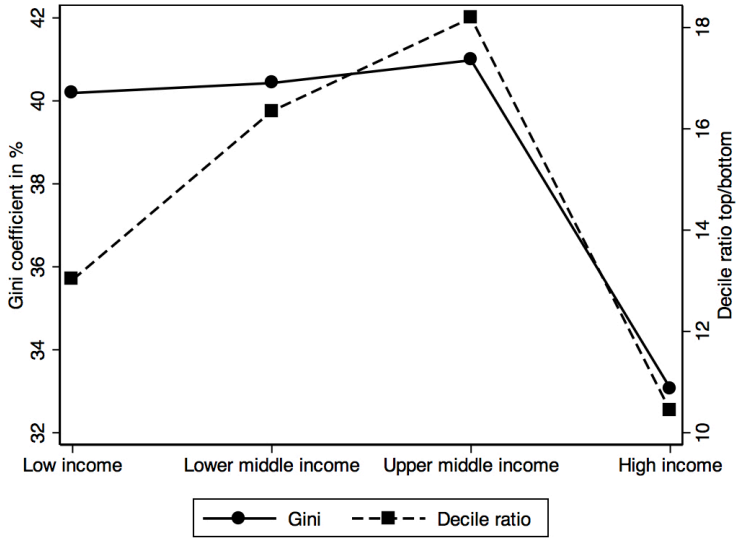

Figure 27: Comparison decile ratio and Gini coefficient

When comparing the two inequality measures, it is crucial to understand the difference. The Gini coefficient is derived from the Lorenz-Curve measuring the deviation from a perfect/fair income distribution to the actual income distribution using normalised income shares to model the distribution. The normalisation leads to a more insensitive measure for extreme values or in this case, extreme inequality. A ratio using the income shares in % from the top decile and the bottom decile takes a more unequal income distribution into account. We can see the result in figure 26 observing significant more "outliers" like Brazil, Colombia, Chile etc. known for their severe inequality. Also in figure 27 a sharp increase of the decile ratio is evident for upper middle income countries. In comparison, the Gini coefficient does not seem to capture these effects.

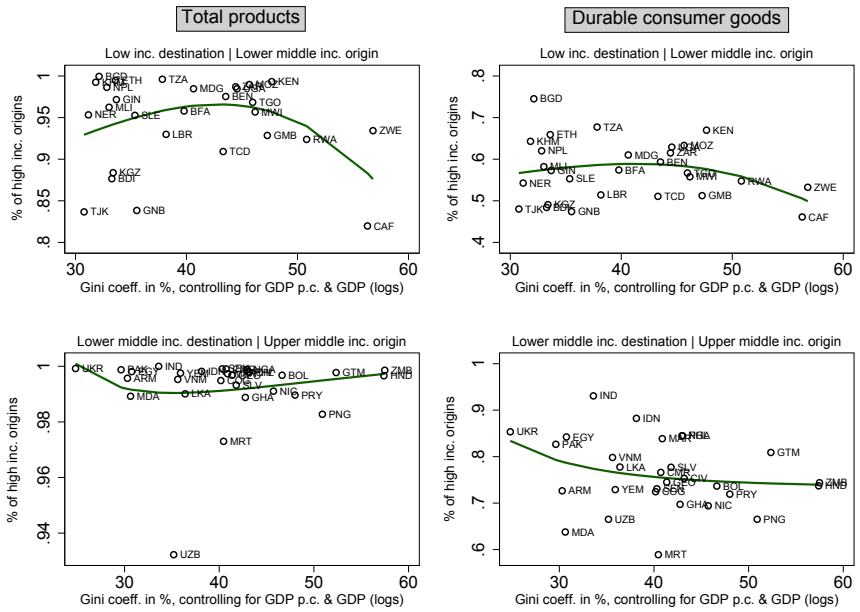

Figure 28: Percentage of high inc. origins and inequality in low & lower middle income countries

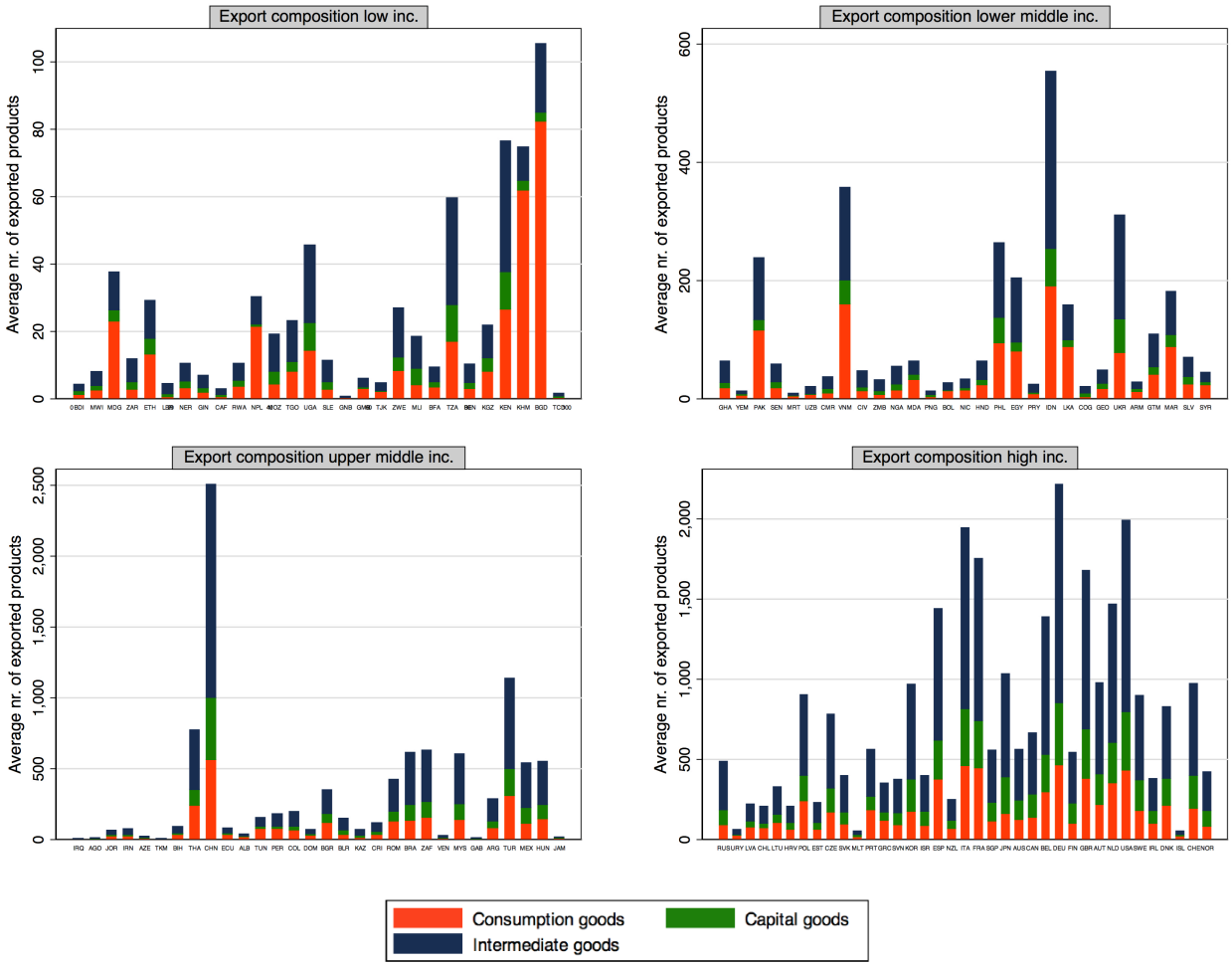

Figure 29: Composition of exports: Average np_{ij} per sectors and income group

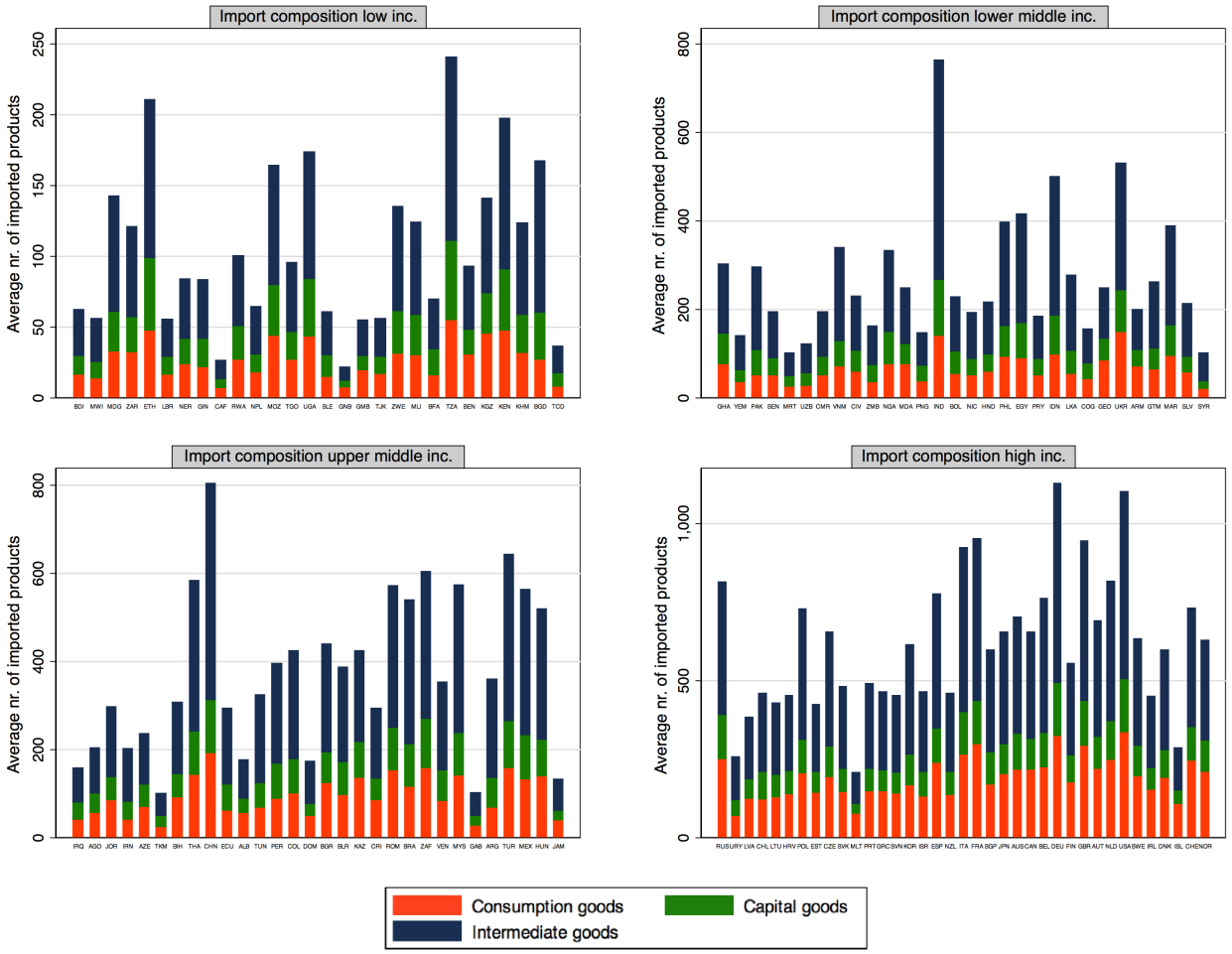

Figure 30: Composition of imports: Average np_{ij} per sectors and income group

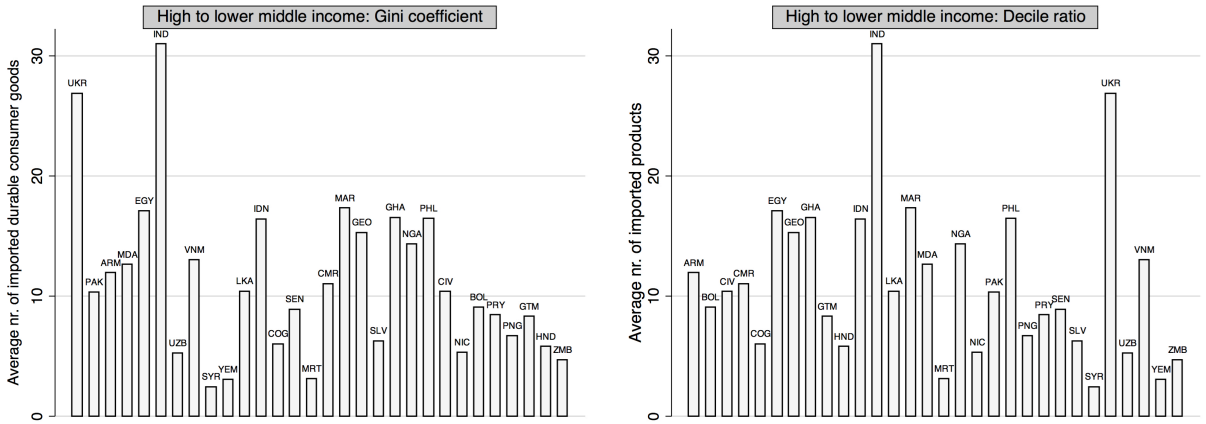

Figure 31: The number of imported durable consumer goods and inequality, lower middle inc. example

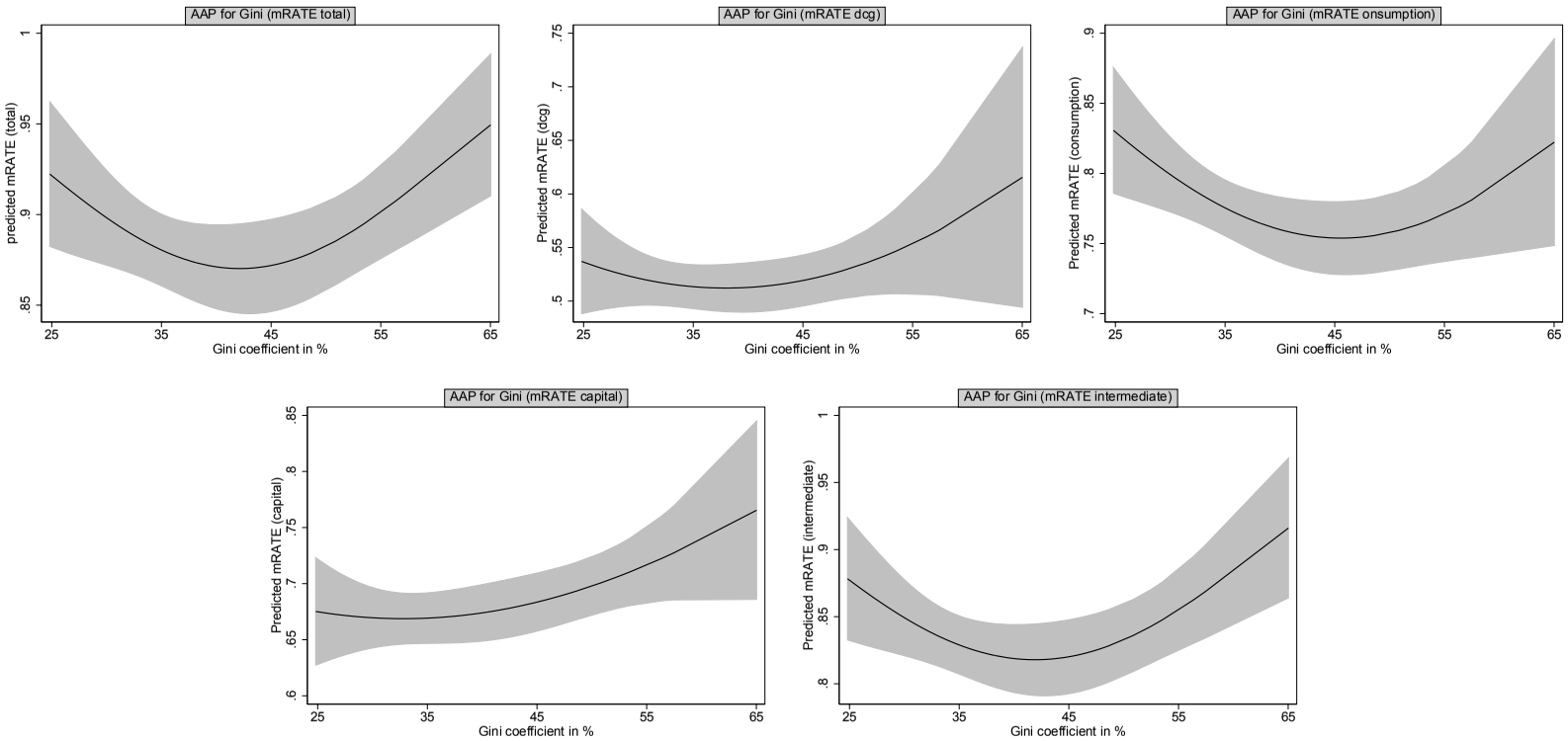

Figure 32: Effect on $mRATE_j$: AAPs for the Gini coefficients

D Robustness checks

D.1 Number of markets

Table 12: Robustness mRATE_j with linear gini-effect

	(1) Total	(2) DCG	(3) Consumption	(4) Capital	(5) Intermediate
GDP p.c.$_j$ (log)	0.00876	0.0351***	0.0213***	0.0132	0.00631
	(0.00819)	(0.00793)	(0.00820)	(0.00885)	(0.00808)
GDP$_j$ (log)	0.0438***	0.0538***	0.0438***	0.0587***	0.0596***
	(0.00591)	(0.00553)	(0.00592)	(0.00591)	(0.00573)
Gini_j^2	-0.000136	0.00105	-0.00128	0.00168*	-0.0000830
	(0.000983)	(0.00115)	(0.00105)	(0.000982)	(0.00106)
Landlocked	0.0115	0.0295	0.00605	0.0322	0.0167
	(0.0260)	(0.0228)	(0.0251)	(0.0261)	(0.0281)
Cost to import (log)	-0.0387**	-0.0135	-0.0234	-0.0206	-0.0416**
	(0.0193)	(0.0163)	(0.0171)	(0.0188)	(0.0211)
Observations	123	123	123	123	123

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(i) Average partial effects (APE) $\frac{dy}{dx}$ of fractional probit response model

(ii) Robust standard errors

Table 13: Robustness mRATE_j with non-linear decile ratio

	(1) Total	(2) DCG	(3) Consumption	(4) Capital	(5) Intermediate
GDP p.c.$_j$ (log)	0.00909	0.0329***	0.0240***	0.0102	0.00641
	(0.00834)	(0.00807)	(0.00832)	(0.00911)	(0.00820)
GDP$_j$ (log)	0.0427***	0.0543***	0.0415***	0.0591***	0.0583***
	(0.00610)	(0.00587)	(0.00601)	(0.00624)	(0.00591)
Decile ratio^2	-0.00154	-0.000492	-0.00303**	0.000325	-0.00183
	(0.00126)	(0.00121)	(0.00123)	(0.00127)	(0.00131)
Landlocked	0.00480	0.0268	0.000393	0.0235	0.00893
	(0.0264)	(0.0235)	(0.0256)	(0.0269)	(0.0286)
Cost to import (log)	-0.0371*	-0.0148	-0.0226	-0.0169	-0.0394*
	(0.0197)	(0.0167)	(0.0176)	(0.0195)	(0.0217)
Observations	116	116	116	116	116

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(i) Average partial effects (APE) $\frac{dy}{dx}$ of fractional probit response model

(ii) Robust standard errors

(iii) Decile ratio is originally quadratic, but due to APE now depicted to be linear

Table 14: Extended regression of analysis $mRATE_j$

	(1) mR_j^{total}	(2) $mR_j^{total\|high}$	(3) mR_j^{DCG}	(4) $mR_j^{DCG\|high}$	(5) $mR_j^{con.}$	(6) $mR_j^{con.\|high}$	(7) $mR_j^{cap.}$	(8) $mR_j^{cap.\|high}$	(9) $mR_j^{int.}$	(10) $mR_j^{int.\|high}$
GDP p.c.$_j$ (log)	0.00349	-0.00761**	0.0315***	0.0350***	0.0165*	0.00352	0.0107	0.00209	0.000471	-0.00461
	(0.38)	(-2.55)	(3.66)	(3.82)	(1.84)	(0.75)	(1.13)	(0.51)	(0.05)	(-1.36)
GDP$_j$ (log)	0.0447***	0.00807**	0.0545***	0.0399***	0.0445***	0.0161***	0.0592***	0.0233***	0.0605***	0.0138***
	(7.34)	(2.30)	(9.75)	(6.48)	(7.42)	(3.50)	(9.89)	(5.16)	(10.27)	(3.25)
Gini$_j$	-0.00111	-0.0000366	0.0000962	-0.00187*	-0.00233**	-0.000865	0.00105	0.000852	-0.00122	0.0000778
	(-1.01)	(-0.10)	(0.09)	(-1.88)	(-2.10)	(-1.56)	(0.96)	(1.55)	(-1.05)	(0.17)
Landlocked	0.00368	-0.0110	0.0251	0.00913	-0.000355	-0.0189	0.0289	0.000534	0.00837	-0.0113
	(0.14)	(-1.30)	(1.06)	(0.44)	(-0.01)	(-1.51)	(1.09)	(0.04)	(0.29)	(-0.97)
logcosttoimport	-0.0416**	-0.0123*	-0.0163	-0.0347**	-0.0268	-0.0180**	-0.0224	-0.0208**	-0.0451**	-0.0188**
	(-2.08)	(-1.86)	(-0.98)	(-2.45)	(-1.51)	(-2.07)	(-1.15)	(-2.01)	(-2.06)	(-2.52)
Observations	123	123	123	123	123	123	123	123	123	123

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Average partial effects (APE) $\frac{dy}{dx}$ of fractional probit response model.

*T*he count of high income origins across sectors is included as well in order

*T*o get a more pronounced inequality-effect.

D.2 Number of product categories

Flexible pseudo-maximum likelihood estimation

Table 15: Robustness checks for $pRATE_{ij}$: Flexible specification

	(1) pR_{ij}^{Total}	(2) pR_{ij}^{DCG}	(3) $pR_{ij}^{Consumption}$	(4) $pR_{ij}^{Capital}$	(5) $pR_{ij}^{Intermediate}$
GDP p.c.$_i$ (log)	-0.000618	0.000750	-0.00629***	0.00820***	-0.000332
	(-0.83)	(0.77)	(-7.30)	(7.98)	(-0.47)
GDP p.c.$_j$ (log)	0.00122	0.00767***	0.00996***	-0.00153	-0.00171
	(0.90)	(4.30)	(6.10)	(-0.86)	(-1.26)
GDP$_i$ (log)	0.0359***	0.0401***	0.0384***	0.0444***	0.0344***
	(67.46)	(61.28)	(60.43)	(62.26)	(66.15)
GDP$_j$ (log)	0.0171***	0.0156***	0.0159***	0.0214***	0.0173***
	(18.97)	(14.09)	(13.90)	(18.14)	(19.25)
Gini$_j$	0.000221	0.000295	0.0000528	0.000513	0.000259
	(0.83)	(0.91)	(0.15)	(1.48)	(1.09)
Distance (log)	-0.0437***	-0.0452***	-0.0487***	-0.0516***	-0.0417***
	(-25.64)	(-21.16)	(-22.68)	(-22.51)	(-27.16)
RTA	0.0133***	0.0104***	0.0177***	0.0142***	0.0121***
	(4.83)	(3.11)	(5.22)	(3.53)	(4.79)
Common language	0.0214***	0.0287***	0.0246***	0.0282***	0.0190***
	(7.51)	(7.75)	(7.16)	(6.62)	(7.33)
Colonial relationship	0.0339***	0.0330***	0.0509***	0.0456***	0.0263***
	(5.12)	(4.61)	(5.94)	(4.51)	(4.73)
Cost to export (log)	-0.0367***	-0.0512***	-0.0487***	-0.0415***	-0.0319***
	(-34.33)	(-36.10)	(-33.62)	(-29.40)	(-31.79)
Cost to import (log)	-0.00420	-0.00148	0.000289	-0.00251	-0.00644***
	(-1.49)	(-0.44)	(0.07)	(-0.65)	(-2.65)
MTR$_i$	2.72e-12***	-7.39e-11***	2.62e-11***	-1.75e-11***	-6.87e-13***
	(19.41)	(-15.36)	(16.67)	(-18.58)	(-19.52)
MTR$_j$	3.45e-10***	3.34e-10***	3.44e-10***	4.30e-10***	3.43e-10***
	(376.70)	(166.54)	(334.19)	(355.15)	(373.59)
Observations	13806	13806	13806	13806	13806

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(i) Average partial effects (APE) $\frac{dy}{dx}$ of FLEX-model proposed by Santos Silva et. al (2014))

(ii) Variance estimator clustered on importer (j) to account for intragroup correlation. Countries itself are independent observations, but within the countries the bilateral trade flows are correlated

(iii) Gini is originally quadratic, but due to APE now linear

The flexible Bernoulli (maximum) likelihood estimation is maximised by the objective function:

$$L(\beta, w_m) = w_m y_m ln\{G(x'_m\beta)\} + w_m(1 - y_m)ln\{1 - G(x'_m\beta)\} \tag{10}$$

The maximised weights allow the distribution to be either symmetric ($w = 1$), left-skewed ($w < 1$) or right-skewed ($w > 1$). This specification, as the distribution is dictated by the data rather than assumption, allows for more flexibility.

OLS estimation

Table 16: Robustness Checks for $pRATE_{ij}$: OLS estimation

DV in logs	(1) np_{ij}^{Total}	(2) np_{ij}^{DCG}	(3) $np_{ij}^{Consumption}$	(4) $np_{ij}^{Capital}$	(5) $np_{ij}^{Intermediate}$
GDP p.c.$_i$ (log)	0.0604	-0.00483	-0.0822*	0.196***	0.0843
	(1.19)	(-0.10)	(-1.68)	(3.28)	(1.52)
GDP p.c.$_j$ (log)	0.0540	0.113***	0.190***	-0.0150	-0.0310
	(1.46)	(4.31)	(5.36)	(-0.55)	(-0.90)
GDP$_i$ (log)	0.743***	0.564***	0.664***	0.636***	0.747***
	(18.44)	(17.77)	(18.03)	(14.90)	(17.28)
GDP$_j$ (log)	0.411***	0.221***	0.292***	0.337***	0.443***
	(17.25)	(12.74)	(12.57)	(19.14)	(19.73)
Gini$_j$	-0.00323	0.000837	-0.00475	0.00484	0.0000507
	(-0.56)	(0.18)	(-0.76)	(1.03)	(0.01)
Distance (log)	-0.921***	-0.658***	-0.859***	-0.778***	-0.925***
	(-16.73)	(-14.20)	(-16.55)	(-15.54)	(-16.87)
RTA	0.475***	0.275***	0.480***	0.392***	0.495***
	(4.57)	(3.20)	(5.14)	(4.50)	(4.96)
Common language	0.650***	0.325***	0.496***	0.498***	0.579***
	(7.93)	(4.03)	(6.04)	(6.25)	(6.72)
Colonial relationship	0.584***	0.498***	0.739***	0.559***	0.551***
	(4.07)	(4.77)	(5.27)	(4.00)	(4.05)
Cost to export (log)	-0.867***	-0.623***	-0.937***	-0.509***	-0.690***
	(-7.62)	(-5.65)	(-8.83)	(-4.32)	(-6.26)
Cost to import (log)	-0.148**	0.0164	-0.0630	-0.0304	-0.199***
	(-2.19)	(0.29)	(-0.85)	(-0.51)	(-3.25)
MTR$_i$	1.93e-09	-7.52e-10	3.39e-09**	2.29e-10	1.34e-09
	(1.28)	(-0.58)	(2.24)	(0.15)	(0.93)
MTR$_j$	5.81e-09***	4.72e-09***	5.62e-09***	5.42e-09***	6.04e-09***
	(5.24)	(4.96)	(5.24)	(5.30)	(5.44)
Observations	11810	7168	10456	9212	11080

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes:

(i) Average partial effects (APE) $\frac{dy}{dx}$ of OLS-model only for Gini relevant

(ii) Gini is originally quadratic, but due to APE now linear

(iiI) Variance estimator two-way clustered on importer (j) and exporter (i) to account for intragroup correlation. Countries itself are independent observations, but within the countries the bilateral trade flows are correlated

Declaration of authorship

I hereby declare

- that I have written this thesis without any help from others and without the use of documents and aids other than those stated above,
- that I have mentioned all the sources used and that I have cited them correctly according to established academic citation rules,
- that the topic or parts of it are not already the object of any work or examination of another course unless this has been explicitly agreed on with the faculty member in advance,
- that I will not pass on copies of this work to third parties or publish them without the University's written consent if a direct connection can be established with the University of St. Gallen or its faculty members,
- that I am aware that my work can be electronically checked for plagiarism and that I hereby grant the University of St.Gallen copyright in accordance with the Examination Regulations in so far as this is required for administrative action.

Date and Signature

Brunnen, 12.11.2015 

.....