"Patterns and Determinants of Carbon Emissions Embodied in Trade"

Master thesis

Jan Siegmeier

Technische Universität Berlin

December 23rd, 2011

Masterarbeit für den Studiengang:

Master of Industrial and Network Economics

Matrikelnummer: 328435

Betreuer:

Prof. Dr. Ottmar Edenhofer, Dr. Robert Marschinski

Zusammenfassung

Basierend auf jüngsten Forschungsergebnissen zur großen und wachsenden Bedeutung des Welthandels bzw. der für den Handel produzierten Güter für die Emission von Treibhausgasen, widmet sich diese Arbeit der Analyse von Mustern des CO2-Fussabdrucks verschiedener Handelsströme sowie möglicher Ursachen. Es werden die gleichen Daten verwendet wie von Peters et al. (2011). Die Analyse erfolgt in zwei Schritten:

Erstens werden zur genaueren Bestimmung der Muster die Veränderungen des im Handel „enthaltenen" CO2 zwischen 1992 und 2006 zerlegt, und zwar in diejenigen Anteile, die auf Veränderungen bilateraler Handelsvolumina, länder- und sektorspezifischer CO2-Intensitäten und der Sektorzusammensetzung einzelner bilateraler Handelsströme zurückzuführen sind. Der Skaleneffekt ist der dominierende, insbesondere für Exporte von Ländern ohne Emissionsreduktionsziel unter Annex B des Kyoto-Protokolls, gefolgt von Intensitätsveränderungen, die für Exporte aus Annex B - Ländern besonders wichtig sind. Zusammen führen die beiden Nettoeffekte in einem Größenverhältnis von etwa 4 zu 1 zu einer Verdreifachung der Netto-CO2-Importe der Annex-B-Länder. Die Sektorzusammensetzung spielt nur eine geringe Rolle.

Im zweiten Schritt werden ökonometrische Schätzverfahren angewendet, um die Mechanismen der den CO2-Bilanzen zugrunde liegenden Handelsströme mit einem Gravitationsmodell zu analysieren. Dabei werden die CO2-Intensitäten der Handelsströme als exogen angenommen. Es wird der Versuch gemacht, die Ergebnisse von Feenstra et al. (2001) zu reproduzieren. Diese nutzen jahres- und sektorweise Schätzungen für die Elastizitäten des Handels bezüglich des BIP des Exporteurs und Importeurs, um auf verschiedene mit der Struktur des Gravitationsmodells konsistente Theorien zu schließen. Sie argumentieren außerdem, dass die unterschiedlichen Ergebnisse für Schätzungen mit dem Gravitationsmodell in der Literatur nicht von verschiedenen Ländercharakteristiken abhängen, sondern von der Art der Güter, die sie handeln. Die Ergebnisse in Analogie zu Feenstra et al. werden bezüglich ihrer Robustheit bei Verwendung anderer Schätzverfahren überprüft.

Theoretische und methodische Aspekte der ökonometrischen Modellierung werden diskutiert und eine Reihe von Tests durchgeführt, um das am besten geeignete Modell auszuwählen. Es werden drei Arten von Schätzungen vorgenommen: Erstens für auf drei Güterarten aggregierte Handelsdaten mit OLS, analog zu Feenstra et al. (2001); zweitens mit Länderpaar-spezifischen Achsenabschnitten und der einfachen OLS-Methode überlegenen Panelschätzverfahren, und drittens mit Panelverfahren auf Sektorebene.

Unsere Ergebnisse stützen diejenigen von Feenstra et al. (2001) für differenzierte Güter. In diesem Fall weist eine höhere Elastizität des Handels bezüglich des BIP des Exportlandes auf ein Modell des monopolistischen Wettbewerbs mit einem "Heim-Markt-Effekt" hin, statt auf ein alternatives

Modell, das Güter auf Länderebene unterscheidet. Dieses Ergebnis ist robust bei den meisten Schätzverfahren (abgesehen von einem Panelansatz, der Differenzen zwischen aufeinanderfolgenden Jahren bildet).

Für den Handel mit homogenen Gütern werden die Thesen von Feenstra et al. allerdings nicht gestützt. Zwar finden wir eine Tendenz zu niederigerem (höherem) Einfluss des BIPs des Exporteurs (Importeurs), im Vergleich zu Handel mit differenzierten Gütern. Aber die Elastizität bezüglich des Importeurs bleibt weiterhin oft kleiner als die des Exporteurs, oder ist nur um einen insignifikanten Betrag größer. Daher ergibt sich kein Hinweis für Marktzugangsbeschränkungen in einem reziproken "Dumping"-Modell, wie sie von Feenstra et al. berichtet werden, und auch nicht für den entgegengesetzten Fall mit freiem Zugang für Firmen.

Die Schätzungen für homogene Güter liefern im Gegenteil derart schwankende Ergebnisse und oft unplausible negative Elastizitäten und insignifikante Parameterschätzungen, dass bezweifelt werden muss, ob das Gravitationsmodell in seiner hier verwendeten Form zur Beschreibung solcher Handelsströme überhaupt geeignet ist. Die schlechte Prognose des Handels mit homogenen Gütern kann als ein Hauptgrund dafür gesehen werden, dass unsere Schätzungen die beobachteten CO2-Muster des Handels insbesondere für die Exporte der „nicht-Annex-B-Länder" nicht vollständig reproduzieren können.

Eidesstattliche Versicherung

Die selbstständige und eigenhändige Anfertigung versichere ich an Eides statt.

Berlin, den 23.12.2011

…………………………………………………….

(Jan Siegmeier)

Danksagung

Ich möchte mich ganz herzlich und offiziell bei meinen Betreuern und Kollegen, meinen Freunden und meiner Familie für alle Formen der Unterstützung während der Anfertigung dieser Arbeit bedanken. Insbesondere danke ich auch Glen Peters und Jan Minx für die freundliche Bereitstellung der Daten und die Beantwortung meiner Fragen, sowie meiner Frau für ihre große Geduld!

Jan Siegmeier, Berlin am 23.12.2011

Diese Arbeit enthält alle wesentlichen Ergebnisse.

Ihr liegt dennoch eine CD mit ergänzenden Ergebnissen bei, auf die im Text mehrfach Bezug genommen wird.

Contents

1 Introduction

With reports of ever higher global carbon emissions (IEA 2011), concerns about climate change and policy instruments to limit emissions of greenhouse gases (GHG), in particular CO_2, are growing. In the current debates about a potential successor to the Kyoto Protocol, patterns of trade and emissions embodied in trade play an important role. Recent research has highlighted that countries with a binding target in Annex B to the Kyoto Protocol ("B countries")[1] on the whole are net importers of carbon emissions, and that the imbalance as well as total trade-related emissions have been growing over the last 20 years. This thesis will identify patterns of emissions embodied in trade (EET) by a simple decomposition and try to analyze some of their determinants by econometric estimations based on trade theory.

GHG emissions from production can either be attributed to the region where the production process or a particular step takes place (production-based accounting), or to the region where the final good is consumed (consumption-based accounting). For a closed economy or the global total emissions, both measures yield the same result; with trade, they may differ by the amount of emissions related to production of the traded goods or services. A country may import more or different goods than it exports, or may use less carbon-intensive technologies than its trade partners. This is a mere accounting exercise, but depending on how these measures are used in international agreements to limit emissions, they might themselves influence patterns of emissions and trade.

An environmental input-output analysis for 87 countries for the year 2001 found that 5.7 Gt or 23.0% of global carbon emissions are related to trade (Peters and Hertwich 2008). Since there are significant imbalances regarding the EET, the production-side accounts of territorial emissions differ markedly from consumption-based accounts for many countries: Carbon imports of B countries from non-Annex B ("non-B") countries, called weak carbon leakage by the authors, amount to 10.8% of their domestic carbon emissions. On balance, B countries consume 5.6% more carbon than they produce. Some European countries import more than 30% of the emissions embodied in the goods they consume (Davis and Caldeira 2010).

Trade, EET and related imbalances are highly dynamic. Figure 1 shows the index of global aggregates of trade value and of EET against the base year 1990, based on Peters et al. (2011) who analyzed the EET for a panel of 57 sectors in 113 regions between 1990 and 2008. We see that total trade increased by 207% or a factor of 3.07, while total EET increased by only 81%. In other words, the

[1] See subsection 3.3.2 for a detailed definition of "(non-)Annex B countries". Here, Annex B countries include the USA.

carbon intensity of total trade decreases (1.81/3.07=0.59 or -41%)[2]. Trade only between B and non- B countries develops similarly, with dollar volume +197% and EET +77% (intensity -40%).

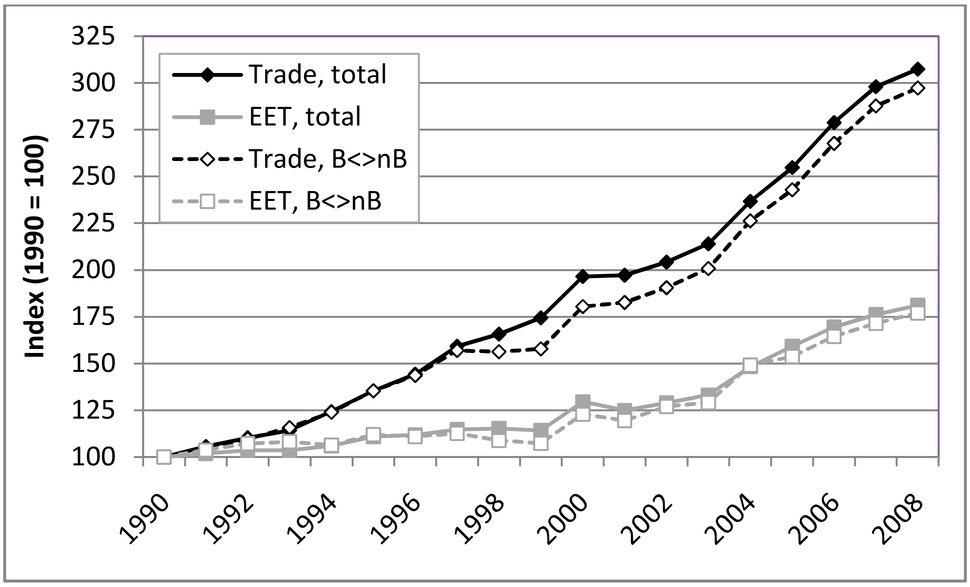

Figure 1: Trade value and EET (index base year 1990) for total international trade and trade among Annex B and non-Annex B countries. See section 3.3 for data sources.

At a more disaggregate level, Figure 2 shows the directional trade flows and embodied emissions from B to non-B countries and vice versa in absolute terms. As for global trade we observe that the embodied emissions grew less than the respective trade volume (+33% vs. +162% for B to non-B and +134% vs. +256% for non-B to B). B countries are net importers in terms of EET across the whole period; net EET imports quadrupled from 383 Mt CO_2 to 1608 Mt CO_2 due to the stronger growth of EET for non-B exports. However, in real (2005) dollar terms the group of B countries had a small export surplus until 1998 and only became net importers from 1999 onwards, with a growing imbalance amounting to 854 billion dollar in 2006.

A first research question is to which extent these developments of EET are attributable to changes in the scale of bilateral trade flows (stronger trade growth for countries with relatively clean exports), to changes in the sector composition of trade (growing share of clean sectors), or to changes in country- and sector-specific carbon intensities. Peters et al. (2011) report trends in EET for seven aggregate sectors and six regions, but they do not analyze in detail the interaction of country shares of world trade and country-sector-intensities. Using their data, we will use a decomposition analysis to complement their results. We find that scale (country composition) and intensity effects are much

[2] For comparison, from figure 1 in Peters et al. (2011), we see that total global emissions grew by about 40%, while world GDP increased by roughly 70%, so the "CO_2 intensity of world GDP" decreased by -18% and much less than the CO_2 intensity of trade. If trade (+207%) had the same composition as total output and thus the same carbon intensity, EET would have increased by 150%. The difference is due to the composition of total output compared to trade, non-traded goods, potentially domestic structural change, etc. which we don't analyze here in detail.

more important for the observed changes in EET than changes in bilateral trade flows' sector composition.

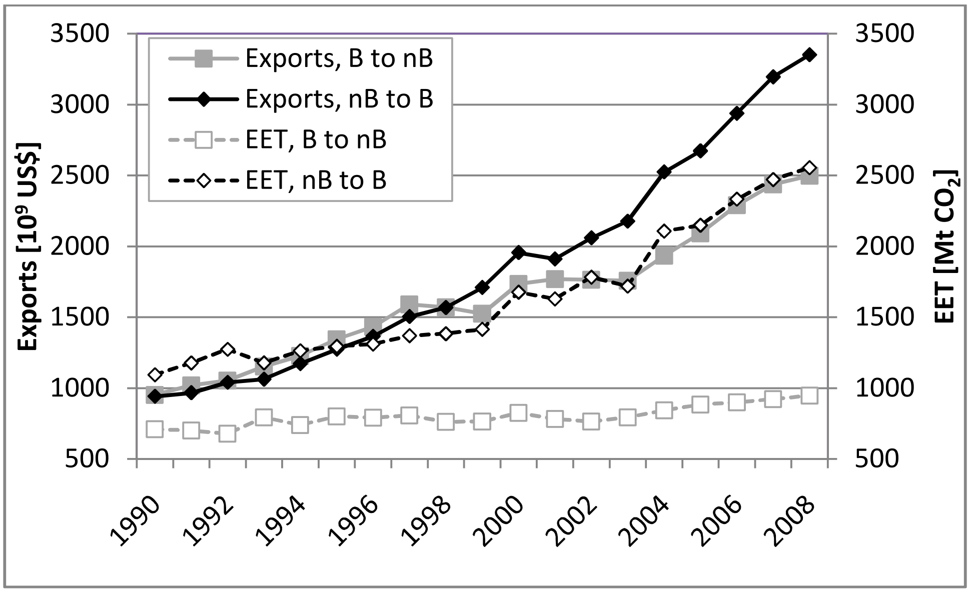

Figure 2: Volume [10^9US$, 2005] and EET [Mt CO_2] of exports from Annex B- to non-Annex B countries and vice versa. See section 3.3 for data sources.

Decompositions are descriptions and not directly linked to a theoretical explanation of the determinants behind the observed patterns. However, they may guide the choice of a suitable approach for further econometric analysis. In our case, we can first choose between looking into country- and sector-specific decarbonisation processes leading to intensity changes, or analysis of trade patterns underlying the scale and composition effects. We opt for the latter and treat intensity changes as given. The second choice is between different theories explaining trade, based on comparative advantage or specialization. We choose the specialization approach, leading to a "gravity model" where the product of country sizes affects bilateral trade. This seems to be the most promising model to explain the dominance of scale- over composition effects. There are also "hybrid" models, for example linking trade patterns to energy intensity (Gerlagh and Mathys 2010) or environmental policy (Antweiler et al. 2004)[3]. We take this into account in a very simplified way by including a "Kyoto variable" into our model.

One motivation of the gravity model is the explanation of intra-industry trade between similarly developed countries, and its applicability for example to trade between OECD- and non-OECD countries is controversial (Hummels and Levinsohn 1995, Debaere 2005). Feenstra, Markusen and Rose (2001, "FMR") argue that it is more useful to distinguish trade in different goods types rather than country types. They show that different theories of specialization in trade predict different

[3] Other alternative approaches include models of endogenous learning affecting competitiveness and carbon intensity, the analysis of international production networks and trade in intermediates (Hummels et al. 2001), or beyond the trade focus, models of structural change (Krüger 2008).

relative sizes for the elasticities of trade with respect to the exporter's and the importer's GDP. Using goods type- and year-wise estimations of a gravity model, they conclude that trade in differentiated goods[4] is consistent with a model of monopolistic competition, while estimates for homogenous goods suggest a model of reciprocal dumping with restricted entry.

Our second research question is if FMR's empirical results can be reproduced with a similar model for the data of Peters et al. (2001), and if they are robust to other estimation strategies. We will see that we can only partly confirm their findings, since we obtaining similar results only for trade in differentiated goods and when using their estimation approach (cross-section OLS). Employing panel estimation techniques to capture unobserved heterogeneity leads to inconclusive results due to large errors and depending on the specific estimator. The "within" estimator yields a better model fit than the "first-difference" estimator in particular for differentiated goods, but only insignificantly small differences between elasticities, preventing the decision between alternative models underlying the gravity equation. We do not find a consistent effect of the Kyoto variable on trade.

While we test FMR's prediction on trade flows of three aggregate goods types, we extend our panel estimations to a more detailed level with 41 sectors matching our decomposition. The results are satisfactory for differentiated goods but relatively poor for homogenous goods. For illustration, we use the bilateral trade predicted by one of our sector-wise estimations, weight them with the original carbon intensities to obtain "predicted EET" and repeat our initial decomposition. While the qualitative patterns and the relative size of the effects are reproduced, we find that the absolute size of the trade and thus EET flows is underestimated, in particular for non-B exports. This suggests that the gravity equation, at least in our simple form, may be inappropriate for a general application to trade of all countries and goods types, rather than differentiated goods trade between developed countries.

This thesis is structured as follows: Chapter 2 briefly reviews some relevant theory, with a focus on gravity models of trade. Chapter 3 summarizes our decomposition method, econometric models and estimation approach, as well as variable definitions and data sources. Chapter 4 reports the results of our decomposition and estimations. Chapter 5 discusses the results and chapter 6 concludes.

[4] See subsection 2.2.2 for a definition of goods types.

2 Review of relevant theories

This chapter reviews classical trade theories based on Heckscher-Ohlin and Ricardo models, before turning to the gravity model for bilateral trade and its theoretical underpinnings are introduced in detail. We consider models focusing on the effect of environmental regulation (e.g. related to the Kyoto protocol) and trade. Finally we comment on a selected alternative model and theories of structural change.

2.1 Models of fully competitive trade

Classical models of trade assume constant returns to scale and thus perfect competition. Rather than by access to a larger market, trade is motivated by comparative advantage: A country will export goods where production costs, expressed in terms of the value of foregone production of other goods, are lower than in competing countries (the exporter's relative price for the good is lower under autarky; Feenstra 2004, p.2).

In a Ricardian model of trade, technology (the productivity of a factor that is assumed to be immobile between countries) determines comparative advantage. Countries export from sectors in which their factor productivity relative to other countries' factor productivity is higher than in other sectors. Typically labor is used as the input factor. In the context of embodied carbon emissions, energy resources could be an important input factor and the efficiency of energy use would determine EET patterns: Countries with high relative energy productivity in energy-intensive sectors would specialize in and export from these sectors.

Alternatively in Heckscher-Ohlin models all countries use the same technologies and countries' endowments with input factors (at least two, for instance labor and capital) determine comparative advantage. With more than two goods actual trade flows cannot be predicted, but only the implied factor contents (Bernhofen 2010): Since it is also assumed that consumers in all countries have identical, homothetic preferences, they use implicit factors in proportion to their share in world GDP and the difference to their country's endowment will be imported or exported. For embodied emissions we would thus expect countries rich in fossil fuel reserves to export carbon-intensive goods – if fossil fuels were immobile. Otherwise if fuels are traded but capital can be assumed to be immobile and if capital intensity of production is correlated with emissions, capital endowment could serve as an approximation. However, Heckscher-Ohlin models without different technologies have been an empirical failure, so modified hybrid models have been developed (Feenstra 2004).

2.2 "New" trade theory and the gravity equation

Ruled out by the trade models above, it is plausible that positive scale and network effects play a role for patterns of trade. Corresponding models provided the first explanation of the gravity equation, although it was empirically successful and popular much earlier it is now known to be consistent with variants of the above models which do not feature increasing returns to scale as well.

2.2.1 Overview

Inspired by gravity in physics, the basic gravity model of trade as proposed by Tinbergen (1962) states that exports X_{ij} from country i to country j is proportional to the trade partners' sizes, for instance total output[5] Y_i and Y_j, and inversely proportional to their distance d_{ij} or more generally barriers to trade, although proportionality may not be direct if exponents α, β and γ differ from one:

$$X_{ij} = c\frac{{Y_i}^{\alpha}\,{Y_j}^{\beta}}{{d_{ij}}^{\gamma}} \qquad (2.1.1)$$

The basic model and its variants proved to be very successful empirically. It has been used in many different applications examining the determinants of bilateral trade flows, in particular the effects to trade barriers like borders and tariffs, but also of bilateral foreign direct investment and migration flows (see Anderson and van Wincoop (2004), Anderson (2011) or Bergstrand and Egger (2011) for recent reviews).

The gravity model is consistent with empirical observations of significant trade flows between technologically similar countries and intra-industry trade of goods with similar factor intensities[6]. But these are not directly predicted by classical trade theories, like the Ricardo or the two-sector, two-factor Heckscher-Ohlin model, so the gravity equation at first lacked a theoretical foundation. Meanwhile various models based on the specialization of countries into different goods (varieties) have filled this gap. In the simplest case (see Rauch, 1999 or Feenstra 2004, p.145), a gravity specification similar to equation 2.1.1 can be motivated as follows: Apart from countries' specialization in varieties of a final good, assume free, balanced trade and thus identical prices (normalized to one) as well as identical, homothetic demand across countries. Then, an amount y_{ik} of variety k produced in country i is consumed by other countries j according to their share s_j of global GDP, where $s_j = Y_j / Y_w$ and $Y_w = \sum_j Y_j$. Exports of variety k from i to j are $X_{ijk} = s_j\, y_{ik}$, and summing over all varieties, we obtain for total exports from i to j:

$$X_{ij} = s_j \sum_k y_{ik} = s_j Y_i = s_j s_i Y_w = \frac{Y_j Y_i}{Y_w} \qquad (2.1.2)$$

When prices are allowed to differ across countries, more elaborate general equilibrium models with explicit utility functions are necessary (Feenstra 2004, p.152), e.g. for the analysis of obstacles to trade like borders or tariffs, but also the assessment of environmental policies affecting prices. In their survey, Bergstrand and Egger (2011) group "conditional" and "unconditional general

[5] The original model in the book by Tinbergen (1962) uses GNP.
[6] Helpman and Krugman (1985) use this characterization of intra-industry trade; Davis (1995) expands on it.

equilibrium models". Conditional general equilibrium models based on Anderson (1979) separate and neglect first-stage production and consumption decisions to focus on a second stage, where countries have an exogenously given "endowment" of goods to trade and only the trade partners are determined. Bergstrand (1985) and Anderson and van Wincoop (2003) are two prominent contributions that derive import demand by maximizing utility functions discriminating goods by origin (Armington assumption) and introduce price indices reflecting different trade costs.

On the other hand, "unconditional general equilibrium models" feature explicit production functions and endogenize the production decisions. An example is the number of varieties produced by a firm in the monopolistic competition approach based on increasing returns to scale by Krugman (1979) and Helpman and Krugman (1985). More recently, Eaton and Kortum (2002) derived a gravity equation from a Ricardian approach with a continuum of goods and a stochastic distribution of technologies among countries[7]. Davis (1995) shows how specialization may also occur with constant returns to scale in what he calls a Heckscher-Ohlin-Ricardo model.

2.2.2 Applicability of gravity models to country and commodity types

Despite its general empirical success the gravity equation might not apply to all kinds of trade flows. The gravity prediction is theoretically motivated through specialization in differentiated goods, allowing the explanation of trade of similar factor content between developed countries with similar technology. It would not be expected to hold for trade between developed and "poor" countries or among the latter.

Specifically, Helpman (1987) assumes that countries are specialized, have identical homothetic tastes and trade is frictionless and uses a monopolistic competition setting based on Helpman and Krugman (1985). He then predicts that the total volume of bilateral trade (VT_A) within a group of countries A, relative to their total size as a group in terms of GDP ($Y_A = \sum_{i \in A} Y_i$), increases with their share in world GDP (Y_w) and if their individual sizes are similar. This can be written as

$$\frac{VT_A}{Y_A} = \frac{Y_A}{Y_w}\left[1 - \sum_{i \in A}\left(\frac{Y_i}{Y_A}\right)^2\right],$$

where the second term on the right-hand side is a "similarity index".

Surprisingly, Hummels and Levinsohn (1995) found support for the Helpman's prediction not only for trade among OECD countries, but also for trade among non-OECD countries. Debaere (2005) argues

[7] Bernhofen (2010) notes that in the Ricardo framework by Eaton and Kortum (2002), the same good is produced by multiple producers, but exported by those with a comparative advantage, and lower trade costs lead to higher volumes via a larger set of traded goods. This is different from the concept of specialization in varieties used in models with product differentiation (increasing returns to scale), "where a decrease in trade costs […] induces consumers to spend more on each imported variety." (Bernhofen 2010, p.9)

that this is due to a misspecification of their econometric model: Their dependent variable is trade (not trade over GDP, as above), and the sum of GDPs is included on the right-hand side, so it is the correlation between the two which dominates the result, rather than the similarity of GDPs.

Carefully implementing the formula above as an econometric model with country-pair fixed effects to deal with unobserved heterogeneity between country pairs, for instance different trade costs, Debaere confirms Helpman's prediction for OECD countries. But for non-OECD countries, he finds no clear relation between the trade to GDP ratio and the similarity index (no support for the Helpman prediction), but a positive coefficient for the share in the world economy. However, Debaere (2005) acknowledges missing support for a model based on monopolistic competition at the aggregate level of trade among non-OECD countries does not entirely "disqualify" the gravity equation for the study of these trade flows. First, it could still be used for analysis at a more disaggregate level; second, we have seen that there are alternative theories yielding gravity models.

In fact, Feenstra, Markusen and Rose (2001; "FMR" in the following) show that a gravity equation not only emerges for specialization in "differentiated goods"[8], where it can be derived in a monopolistic competition model or from the Armington assumption, but also for trade in "homogenous goods" if competition is imperfect and markets are segmented ("reciprocal dumping" models with free or restricted entry). They derive different predictions regarding the relative importance of the exporter's and importer's GDPs for trade flows, which enables them to distinguish between the four variants using the coefficients from estimations of a gravity model of the form

$$\ln\left(X_{ij}^{g}\right) = \beta_0^g + \beta_X^g \ln(Y_i) + \beta_M^g \ln\left(Y_j\right) - \beta_3^g \ln\left(D_{ij}\right) + \beta_4^g CONT_{ij} + \beta_5^g LANG_{ij} + \beta_6^g FTA_{ij} + \beta_7^g REM_{ij} + \epsilon_{ij}^g \qquad (2.1.3)$$

where $X^g{}_{ij}$ are exports of goods type g from country i to country j, Y_i and Y_j are the countries' real GDPs, $\beta^g{}_X$ and $\beta^g{}_M$ are the elasticity of exports with respect to exporter's and importer's GDP, respectively, and $\beta^g{}_0$ is a constant (capturing the effect of world GDP in their model, among other things). Then, a number of ad-hoc auxiliary country-pair variables known to affect bilateral trade follow: geographical distance D_{ij}, dummies for contiguity $CONT_{ij}$, common language $LANG_{ij}$ and common membership in a free trade agreement FTA_{ij}, and finally a "remoteness" measure REM_{ij} (see below). $\in^g{}_{ij}$ is an orthogonal error term. FMR predict the following parameter relations for different goods types and models of trade:

For trade in a *differentiated* good, a *monopolistic competition* model leads to a "home market effect" (Krugman 1980): Since trade costs put imports at a disadvantage, more firms (each producing one variety, but the same amount of the differentiated good) enter in large countries, where most of the

[8] See below and section 3.3.2 for details of the classification.

demand is located, until profits are zero. This bias towards large countries is more than proportionate to country size, overcompensating higher demand, so larger countries are net exporters of the differentiated good. FMR show that the impact of the exporter's GDP on exports will then be larger than that of the importer's GDP ($\beta_X > \beta_M$).

On the other hand, in a model of trade in a differentiated good using the *Armington* assumption, commodities are differentiated at the country level, the number of varieties per countries is fixed (usually at n=1, implying restricted entry, in contrast to the monopolistic competition case above), and larger countries are net importers. This leads to the opposite prediction regarding the impact of trade partners' GDP ($\beta_X < \beta_M$)[9].

For goods that are more *homogeneous*, up to the point of being traded on organized exchanges, "reciprocal dumping models" with Cournot oligopoly pricing and segmented markets leads to similar predictions, depending on the openness to new firms' entry assumed in the model: With *free entry*, there will again be disproportionately more firms in the larger country (with lower Cournot prices), exporting to the smaller one, since otherwise the zero-profit condition would be violated in one or both countries. In this case, we again expect $\beta_X > \beta_M$. With *no entry* (only one firm per country), it can be shown that "at constant marginal cost, each firm must have an equal market share in the other firm's market, implying that the small country is the net exporter." (Feenstra et al. 2001, p. 439) Thus, the prediction is $\beta_X < \beta_M$.

FMR test their model on five worldwide cross-sections of bilateral trade data between 1970 and 1990; we present their results for the first and last cross-section in Table 1. Separate estimations are performed for the three different categories of goods as classified by Rauch (1999), treating goods as "homogenous" when they are traded in an organized exchange, as "reference priced" when they are similar enough for a published price quote to exist, but not traded in volumes and frequencies high enough for a central market place, and as "differentiated" if not even a reference price is available. Their results indicate that trade in differentiated goods and those with a reference price is best described by a monopolistic competition model (with β_X around 1.1 and 0.9, respectively, β_M around 0.65 and R^2 about 0.5), while the estimated coefficients for trade in homogeneous goods match a reciprocal dumping model with restricted entry (with β_X around 0.5, β_M around 0.8 and R^2 about 0.35). They report that estimations for $\beta_{X/M}$ are robust to restrictions in the set of countries (only among OECD countries, where R^2 values are even higher, or between OPEC and non-OPEC states,

[9] Quoting Feenstra et al. (2001), p.436: "In the Armington formulation, by contrast, *aggregate* production and consumption are what matters. If the price of country *i* 's variety is the same as the price of country *j* 's variety (they produce *x* in proportion to size), each country will demand the domestic variety and the foreign variety in the same ratio. But this cannot be consistent with total production's being in proportion to income in each country, since there would be an excess demand for the small country's good. With the two *x* goods being symmetric but imperfect substitutes, the small country must be the net exporter of *x*."

with slightly lower R^2). Also, including the share of homogeneous, resource-intensive goods (minerals and fuels) in GDP as an additional explanatory variable – as one might suspect that for countries where these industries play a particular role, a higher GDP would have less of an effect on exports – leaves the results unchanged[10]. This means that differing estimates of the gravity equation coefficients depend on the nature of the goods, rather than the characteristics of the countries in the sample.

Selected results from Feenstra et al. (2001): Aggregate bilateral trade, pooled over all regions						
Dependent variable: ln(Xgijt)						
	Cross-section (OLS)					
	homogenous goods		**reference-priced goods**		**differentiated goods**	
	1970	**1990**	**1970**	**1990**	**1970**	**1990**
$\ln(Y_X)$	0,44 (0,02)	0,54 (0,02)	0,94 (0,02)	0,91 (0,02)	1,11 (0,02)	1,12 (0,02)
$\ln(Y_M)$	0,85 (0,02)	0,81 (0,02)	0,69 (0,02)	0,74 (0,01)	0,62 (0,02)	0,72 (0,02)
ln(DIST)	-0,75 (0,04)	-0,89 (0,04)	-1,06 (0,04)	-1,15 (0,04)	-1,11 (0,04)	-1,1 (0,04)
REM_{XM}	227 (67)	384 (72)	523 (66)	719 (63)	493 (81)	794 (62)
FTA	0,77 (0,22)	1,06 (7,5)	1,73 (0,15)	1,38 (0,11)	2,2 (0,12)	1,73 (0,11)
CONT	-0,07 (0,40)	0,26 (0,16)	0,06 (0,16)	-0,01 (0,15)	0,02 (0,16)	-0,0 (0,16)
LANG	0,91 (0,10)	0,61 (0,09)	0,66 (0,16)	0,55 (0,08)	0,94 (0,08)	30,69 (0,08)
R^2	0,35	0,4	0,47	0,56	0,49	0,57
N	5505	5095	5381	5439	6498	6367

Table 1: Selected results from table 2 in Feenstra et al. (2001), p.442

2.3 Environmental regulation and trade

Another attempt to explain the observed trade patterns and EET net imports of B countries would be to attribute them to climate policy, e.g. policies by countries to fulfill their commitment under Annex B of the Kyoto Protocol. The impact of trade on the environment (with or without environmental policy), and in the other direction the effect of environmental regulation on trade, have been widely discussed even before carbon emissions became an issue. We summarize the findings of this literature and apply them to the case of CO_2.

In the one direction, trade may affect emissions and thus the environment via production changes, as countries increase or decrease their production due to exports or imports, or via income changes: For many pollutants, an inverse-U-shaped relationship termed “environmental Kuznets curve” (EKC; Grossman and Krueger, 1993) is found between income and polluting emissions. Possible explanations for the EKC include shifts from capital- and pollution-intensive to human-capital-intensive production, increasing demand for environmental quality, political “threshold” effects or increasing returns to abatement (Copeland and Taylor, 2004). If trade affects income, it would thus

[10] Other robustness checks in Feenstra et al. (2001) which leave the results virtually unchanged are the inclusion of GDP per capita and using Tobit instead of OLS estimation.

also affect emissions, and pollution "embodied" in trade. Similarly, in the other direction, trade patterns could be determined not only by factor endowments, technology, etc., but also by domestic policies (Tobey 1990), which in turn depend on income.

In the model of Antweiler et al. (2001), trade is motivated by differences in endowments with classical input factors, capital and labor, and pollution policy, which is determined endogenously by a demand and supply model for pollution. For SO_2 emissions, they find an overall positive effect of trade on the environment, since the increasing scale of "dirty" production, in wealthy countries with high capital endowment, is offset by the use of cleaner technology, induced by more stringent policies (confirmed e.g. by Ederington and Minier, 2003; Frankel and Rose, 2005). Their findings suggest that a "weak pollution-haven" effect exists, where – all else equal – countries with weaker environmental regulation produce and export more dirty goods. The "strong" version labeled pollution haven hypothesis (PHH), where the effect dominates trade patterns, is rejected (Copeland and Taylor, 2004; Cole et al., 2005; Spatareanu, 2007).

For the case of CO_2, an important point is that most of the literature reporting an EKC for local pollutants does not find one for carbon emissions[11]: Shafik and Bandyopadhay (1992) report an EKC for example for SO_2 from fossil electricity generation, but increasing CO_2 emissions with income. This is attributed to free-riding at the expense of the global public and future generations, while pollutants like SO_2 have an immediate local effect, creating public pressure for policy measures. Matching this, the authors find higher indebtedness (pointing to high discount rates) to be positively correlated with carbon emissions, and negatively with SO_2 emissions (see also Holtz-Eakin et al., 1995; Arrow et al., 1995). This has two consequences here: First, if a "strong" pollution haven effect has not been found for immediate, local pollutants and related policies, it is unlikely to be found for carbon emissions as well. Second, for the choice of an appropriate model for our purposes, if higher income doesn't increase public pressure for carbon policy because CO_2 is not an immediate local pollutant, the applicability and added value of a direct pollution demand-and-supply model linking endogenous policy and trade as in Antweiler et al. (2001) to the case of CO_2 is questionable[12].

Alternatively, instead of a direct link between income, regulations targeting CO_2, and trade, we could imagine an indirect link: CO_2 emissions are complementary to other pollution facing stricter regulation, so CO_2 patterns of trade for example just reflect SO_2 or NO_X patterns (as speculated e.g. by Cole and Elliott, 2003). Higher income inducing regulations that target complements of CO_2, e.g.

[11] All explanations for the EKC stated above are consistent with finding an EKC for local, immediate pollutants and simultaneously no EKC for CO_2, as shown in Siegmeier (2010), mainly due to the different possibility of actors to internalize the benefits of abatement and technological decoupling of abatement for different pollutants.

[12] Gerlagh and Mathys (2010) note that for the effect of environmental policy on trade, income is empirically less important than the other way around.

SO_2, would then also affect CO_2. However, such an approach would be complicated, data intensive and prone to measurement and specification errors.

Since we focus on trade theory, we use a simpler approach to control for the impact of climate policy on trade and just include a “Kyoto” variable in our gravity specification described below (compare Aichele and Felbermayr, 2010).

2.4 Application to the analysis of EET

The basic Heckscher-Ohlin and Ricardo models are theoretically suitable for the description of the commodity composition (or factor content) of trade; however, in our decomposition analysis, sector composition will turn out to be of minor importance for the observed EET changes. Also, these models do not describe bilateral trade in a multi-country setting[13], which is what we are mainly interested in here.

In contrast, we saw that the gravity model may be a suitable description for bilateral trade not only between similar (developed) countries, but also for trade between countries as different as Annex-B and non-Annex-B countries, as long as different types of goods are treated separately. Moreover, it is possible to draw conclusions about the specific underlying theory if the estimated elasticities of trade with respect to importers’ and exporters’ GDP are different.

Finally, conventional models integrating pollution reduction policy as an additional motive for trade are designed for pollution emissions with an immediate and local effect and not easily applicable to the case of carbon emissions.

Together this indicates that instead of directly analyzing the multi-dimensional problem of EET, we should use a somewhat simpler approach and analyze the underlying trade pattern. A sector-wise gravity model of trade has the best chance of explaining and reproducing our decomposition results regarding sector composition and scale effects. The intensity effect in this case is treated as exogenous.

Specifically, our second research question after the decomposition is if FMR’s empirical results can be reproduced with a model similar to equation 2.1.3 for the data of Peters et al. (2001). That is, we will test whether the observed trade in differentiated goods underlying our EET data can best be explained by monopolistic competition or national product differentiation, or reciprocal dumping with or without free entry in the case of homogenous goods. We will thus use an approach and model similar to FMR, first to verify FMR’s findings at the aggregate goods-type level for our dataset

[13] With multiple countries, Ricardo and HO models can only be used to model “pooled” world trade (imports and outputs per country, aggregate over all trade partners) and cannot be directly applied to bilateral trade (since the rest of the world needs to be taken into account).

(subsection 4.2.1) and to compare different estimation approaches as described in section 3.2. We will proceed to estimations on a more detailed sector level in subsection 4.2.2.

Because we are more interested in high predictive power for comparison to our decomposition results than for example in identifying the effect of a certain trade barrier, we prefer the simpler model of FMR to a more elaborate approach like the general equilibrium model by Anderson and van Wincoop, 2003. However, equation 2.1.3 augments the simple model of equation 2.1.2 by ad-hoc variables not formally derived from theory, in an attempt to capture some of the more complex effects – for example, "remoteness" mimics the general equilibrium price effects. We return to this methodological issue in section 3.2.7. Also, motivated by endogenous policy models, we will control for GDP per capita in our estimations by adding population as additional explanatory variables.

Finally, we include a "bilateral Kyoto dummy variable" in our gravity estimations, which serves as a proxy for the actual (asymmetric) introduction of various policies in order to comply with a ratified Annex B target (which potentially affect costs and thus trade), such as emission standards or trading schemes. The variable is defined as follows:

$$KYOTO_{ijt} = \begin{Bmatrix} 1 & \text{if i has a binding Annex B target, while j does not} \\ 0 & \text{if both i and j have the same status} \\ -1 & \text{if j has a binding Annex B target, while i does not} \end{Bmatrix} \ldots at\ time\ t \quad (2.4.1)$$

Here, i is the exporter and j the importer. A "binding Annex B target" means that the country has been assigned an emissions target in the Kyoto protocol's Annex B and has actually *ratified* the protocol in the respective year or earlier (so the USA are treated as "non-B" in this case).

2.5 Selected alternative approaches

While we will focus mainly on "new" trade theory, we would like to highlight two alternative approaches that we find important in the context of EET.

One is the analysis of Gerlagh and Mathys (2010), applying a concept developed by Romalis (2004), who integrates a many-country Heckscher-Ohlin model with a monopolistic competition model with trade costs. Trade is then determined by the interaction of sector-wise factor intensities with country-wise factor abundance. Gerlagh and Mathys use energy as a factor and use inverse energy prices as a measure of abundance. With their approach, patterns of EET could be explained as energy content of trade, weighted by each country's carbon intensity of energy.

The other important approach, or rather field of research, is structural change (for a recent review, see Krüger, 2008). The trade models described above are static and assume technologies, factor endowments, etc. as exogenous and in most cases constant. At least in the long run, domestic structural change might play a role, so these variables would have to be treated as endogenous in appropriate dynamic models, which would also allow us to control for path-dependencies. However, theories of structural change require consideration of detailed domestic economic processes beyond trade which is out of scope for this thesis.

3 Methods and data

The objective of this thesis is two-fold: One aim is to clarify the extent to which changes in carbon intensity of production, in composition of trade and in volume of bilateral trade flows respectively drive the observed developments of emissions embodied in trade in general and the balance between Annex-B and non-Annex-B-countries in particular. Based on these results, the second objective is to identify possible determinants behind the directly trade-related drivers, composition and scale, while intensity is treated as exogenous. We thus describe two methods of analysis in this section.

First, we choose an additive decomposition using Laspeyres indexes. Decomposition analysis is a simple and transparent tool to identify the relative contribution of several changing variables to an overall change in the aggregate quantity of interest, and is a well-established method for example in industrial energy demand analysis (Ang and Zhang 2000). However, it is highly dependent on the choice of variables and based on an identity, so it cannot be used to refute any theory. From the growth rates shown in the introduction, changes of trade volumes can be expected to dominate this decomposition, but the relative importance of sector composition and carbon intensities is still of interest, also for the choice of the theory that we are modeling and testing in the next step.

Second, we try to explain the observed pattern with an econometric model motivated by trade theory. Gravity models are suitable for the description of the volume of bilateral trade and empirically successful, as widely documented in the literature. Breaking down total bilateral trade into separate goods types, as in Feenstra et al. (2001), allows a distinction between different theories consistent with the general gravity specification. We perform similar estimations at the goods type level with our data, with different model specifications and estimators, to reproduce and test the robustness of the results of Feenstra et al. Then, we extend this to the sector level, which matches the level of detail that we used for our decomposition: A gravity model predicting different changes in trade for each sector is potentially able to explain both scale and composition effects.

3.1 Decomposition methods

This section outlines basic decompositions using emission intensity, trade composition and trade volume, similar to methods originally developed to isolate the effect of structural shifts in production on industrial energy use (see, for example, Schäfer 2005)[14]. Ang and Zhang (2000) comprehensively review such decompositions. We selectively summarize their exposition in the following and apply it to our circumstances.

[14] This threefold split is also reminiscent of the analysis of the "environmental Kuznets curve" pioneered by Grossman and Krueger (1993), who distinguish effects of technology (determining the pollution intensity of production), composition and scale.

We are interested in changes in the total emissions embodied in international trade and in the net balance of emissions embodied in trade between Annex-B and non-Annex-B countries. Both can be explained by changes at a more granular level, namely the emissions embodied in bilateral trade flows per sector, which can then be aggregated over sectors and regions at will to give the total and balance figures. We thus start with the decomposition of emissions embodied in exports in sector *s* from country i to country j (EEX_{sij}), where s, i and j are running over all sectors and regions in our data set, and later sum over sets of countries and/or sectors. The basic identity is

$$EEX_{sij} = \frac{EEX_{sij}}{X_{sij}} \frac{X_{sij}}{X_{ij}} X_{ij} = I_{si} C_{sij} X_{ij} \tag{3.1.1}$$

where X_{sij} and X_{ij} denote the sectoral and total export volumes from i to j, respectively, the first factor is the emission intensity of exports per sector (I_{si} , which is independent of the importer j), and the second is the sector's share in total exports from i to j (C_{sij})[15].

Ang and Zhang (2000) distinguish additive and multiplicative concepts of decomposition; both lead to the same conclusions. We follow the authors' suggestion and use additive decomposition when comparing only two periods, as we will do for the quantities EEX_{sij} and aggregates thereof (EEX and $BEET_{BnB}$, see below), because interpretation is more intuitive.

An additive decomposition of changes in EEX_{sij} reads

$$\Delta EEX_{sij} = EEX_{sij}^{1} - EEX_{sij}^{0} = \Delta_{int} EEX_{sij} + \Delta_{comp} EEX_{sij} + \Delta_{vol} EEX_{sij} \tag{3.1.2}$$

where superscripts 1 and 0 signify values from periods t_1 and t_0 , and Δ_{int} , Δ_{comp} and Δ_{vol} denote changes due to varying I_{si}, C_{sij} and X_{sij}.

The calculation of the respective contributions $\Delta<.>$ depends on the employed decomposition method. For our additive decomposition, we use the residual-free method of refined Laspeyres indexes: In the simple Laspeyres index method, each contribution is calculated by multiplying the change between two periods in one variable in the basic identity (here, equ. 3.1.1) by all other variables' base period value. For example, the effect of a change in X_{ij} describes how much EET_{sij} would have changed if only bilateral trade volume had been scaled up (thus the name "direct scale effect"), while the sector shares defining the composition of bilateral trade (C_{sij}) and the sector-wise intensities I_{sij} remained constant. Since the individual factors do not change in isolation, but simultaneously, interaction terms of changes in one variable with changes in the other variables arise. In the simple method, these are lumped into a residual term, which might be of substantial size and render the decomposition meaningless. In the refined method, the interactions are split evenly

[15] Additional time indexes on all variables are omitted in this and the next two equations for better readability.

between the variables ("jointly created and equally distributed principle", Ang and Zhang 2000, p. 1164-5). For example, in our case the contribution of intensity changes to changes in EEX_{sij} is:

$$\Delta_{int}EEX_{sij} = \Delta I_{si}\, C_{sij}^0\, X_{ij}^0 + \tfrac{1}{2}\Delta I_{si}\, \Delta C_{sij}\, X_{ij}^0 + \tfrac{1}{2}\Delta I_{si}\, C_{sij}^0\, \Delta X_{ij} + \tfrac{1}{3}\Delta I_{si}\, \Delta C_{sij}\, \Delta X_{ij}\,, \qquad (3.1.3)$$

with $\Delta I_{si} \equiv \left(I_{si}^1 - I_{si}^0\right)$, $\Delta C_{sij} \equiv \left(C_{sij}^1 - C_{sij}^0\right)$ and $\Delta X_{ij} \equiv \left(X_{ij}^1 - X_{ij}^0\right)$

and similarly for $\Delta_{comp}EEX_{sij}$ and $\Delta_{vol}EEX_{sij}$. This can be extended with more interaction terms if the basic identity contains more factors. The simple Laspeyres method, in contrast, uses only the first term, while the interaction of intensity changes with composition and volume changes contribute to a residual term, additionally to the three terms in equation 3.1.2. Since the interaction terms are of higher order in ΔI, ΔC and ΔX, they are of relatively less importance the smaller the changes between the two periods are.

Based on the decomposition for each country-pair and sector, we can now decompose the changes in total emissions embodied in exports ($\sum_{sij} EEX_{sij}$) between two periods (*0* and *1*) as

$$\Delta EEX = \sum_{sij}\left[EEX_{sij}^1 - EEX_{sij}^0\right] = \Delta_{int}EEX + \Delta_{comp}EEX + \Delta_{vol}EEX = \sum_{sij}\Delta EEX_{sij} =$$
$$= \sum_{sij}\Delta_{int}EEX_{sij} + \sum_{sij}\Delta_{comp}EEX_{sij} + \sum_{sij}\Delta_{vol}EEX_{sij} \quad (3.1.4)$$

The decomposition of changes in the balance of emissions embodied in trade ($\sum_{sij} EEX_{sij} - EEX_{sji}$) between Annex-B and non-Annex-B countries (BnB) becomes

$$\Delta BEET_{BnB} = BEET_{BnB}^1 - BEET_{BnB}^0 = \sum_{s,i\in B,j\in nB}\left[\left(EEX_{sij}^1 - EEX_{sji}^1\right) - \left(EEX_{sij}^0 - EEX_{sji}^0\right)\right] =$$
$$= \sum_{s,i\in B,j\in nB}\Delta EEX_{sij} - \sum_{s,i\in B,j\in nB}\Delta EEX_{sji} \qquad (3.1.5)$$

We will decompose changes in EET between 1992 and 2006 (since we only extrapolate the bilateral and sector split to '90 and '08, respectively; see section 3.3), but we have data for all years in between. Instead of doing one decomposition with rather large changes in the variables, we can also decompose the changes year by year and add the results. As we saw in equation 3.1.3 this is preferable because it reduces the influence of the interaction terms.

A more detailed decomposition analysis would compare the results of the Laspeyres index with base year weights to those of a Paasche index with terminal year weights, or a Marshall–Edgeworth index using the mean of base and terminal year weights (Hoekstra et al. 2003). We skip this step here due to space restrictions and attach more importance to the statistical methods. Moreover, the size differences between the scale, composition and intensity effects obtained by the Laspeyres index and presented in section 4.1 are so large that it seems unlikely that a different index method would affect our qualitative results, given that we use a year-wise decomposition.

3.2 Gravity model and panel estimation

Econometric methods provide rigid checks for the relevance of explanatory variable, but can also be misleading: In most cases, they at best show consistency of a model with the data rather than causality, and results can again only be interpreted within the specific model. Potential sources of error include misspecification regarding omitted or endogenous variables and functional form, and incorrect assumptions about the error term e.g. with respect to heteroskedasticity or auto-correlation. We will thus discuss theoretical aspects of model specification as well as several empirical tests.

3.2.1 Aggregation level

The standard theories underlying the gravity model that were discussed above are based on countries' specialization in one or several substitutable variants of one good. For an estimation of the parameters of the gravity equation, this suggests full aggregation of all kinds of trade from one country to another into a single flow. FMR predict different parameter values depending on the type of traded goods, but since different goods types are traded simultaneously, they break down total trade flows into three parts for separate estimations, without further analysis of how dropping the assumption that all goods are substitutable to the same degree might affect the functional form and expected parameter values in the gravity model. Verifying this is beyond the scope of this thesis, but if we accept their implicit assumption that separate estimations for three goods types are consistent with the model, we might as well perform separate estimations for more detailed trade data[16]. Thus, we proceed in two stages: First, we use goods-type-wise aggregated data for estimations and compare them to the results of FMR, since we build on their theoretical results. Second, we want to explore the explanatory power of the gravity model regarding the scale and composition effects found in our decomposition, which is based on a dataset with goods trade in 41 sectors, so we use sector-wise estimations matching our level of analysis.

Smaller subsamples at the sector level potentially impair the significance of our estimates and tests; with our large data set, this is no major problem. However, there are three other potential econometric problems, some of them related to the use of sector subsamples:

First, tightly linked to the question of aggregation level we have problems to deal with zero bilateral trade. Their share is larger in sector-wise trade data. If we do not use more elaborate estimation methods but restrict our estimations to the intensive margin, we have to be aware of the potential selection bias (subsection 3.2.2).

[16] Hallak (2010) explicitly argues for a sector-wise version of the gravity model.

Second, it is not clear if we can assume trade to be homogeneous in the sense that the coefficients of our basic model (stated in subsection 3.2.3) apply across all sectors, years and country-pairs, or if we should rather perform separate estimations for several subsamples (subsection 3.2.4). This includes the discussion of general time-dependent effects that apply to all country-pairs in a cross-section.

Third, the functional form of our basic model might omit country-pair specific variables that are correlated with the included variables, leading to biased estimates. Since we use the same functional form for all sectors, or groups thereof, we might omit potentially different variables for each sector. The omitted variable bias is probably even stronger for sector-wise than for pooled estimations, where the different biases might offset each other to some extent. As long as the omitted variables are constant for country-pairs, specifications with individual fixed effects will take care of this (subsection 3.2.5).

3.2.2 Handling of "zero trade"

Even if we make no attempt to theoretically explain occurrences of zero bilateral trade for a sector or overall, zeros are still a problem for estimating a logarithmic gravity specification. We observe that we have a larger share of zeros at the detailed sector level (Table 9). Basically, there are four options to handle this issue:

The first is to omit all cases of zero trade. This implies that we are effectively only estimating the "intensive margin"[17] of bilateral trade, that is, the effect of the explanatory variables on trade *volume* in a specific sector *given that two countries already trade in that sector*; for example, positive elasticities of trade with respect to trade partners' GDPs mean that larger countries trade more goods (or more variants) within a sector. At the "extensive sector margin", they would trade in a larger *number of sectors*. Effects at the "extensive country margin" would affect the *number of country-pairs* trading at all. Using a higher level of aggregation, like the three goods types in FMR, captures more of the two extensive margin effects than estimations for 41 separate GTAP sectors. The importance of this effect becomes obvious when we compare the share of zeros for separate sectors and goods types (Figure 3, see also Table 9 in the appendix): For example, for the 14 sectors of the "homogenous" goods type under the conservative classification, on average only around 5% of the data points are non-zero and 13% of the country-pairs have non-zero data for at least one period, but if we aggregate this data, we have 32% of non-zero values and 56% of country-pairs entering the sample[18].

[17] For a detailed analysis of the relationship between the "intensive" and "extensive" margins of trade, see Felbermayr and Kohler (2006).

[18] If in contrast, we pooled rather than aggregated across sectors, we would obtain significantly smaller estimates for β_X and β_M, as test runs confirmed.

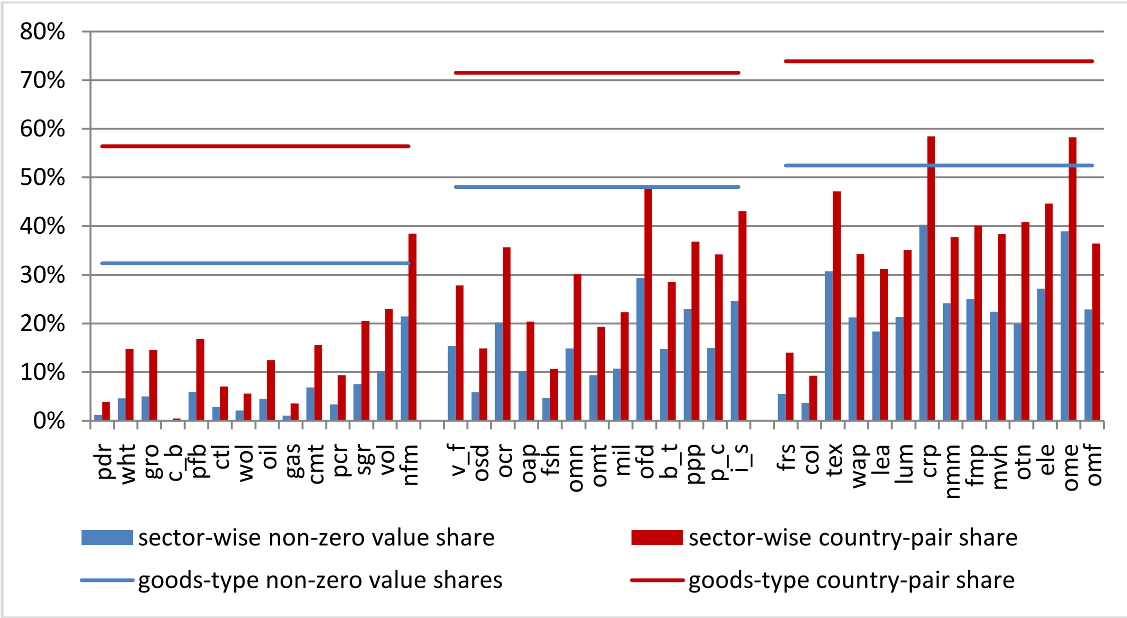

Figure 3: Shares of non-zero values in sector samples (blue bar) and share of country-pairs entering the respective sample (red bar), grouped by goods-type (from left: homogenous, reference-priced, differentiated, using the "conservative" classification) and compared to respective values for aggregate data

The second option is to replace all cases of zero trade by a small value, which is sometimes motivated by assuming that some tiny trade flows might not have been detected or included in the statistics (Debaere 2005). However, at least at the detailed sector level, where the problem is more severe, this assumption is not plausible (e.g. Oceania does probably not trade bi-directionally with Poland in all 41 sectors). With our data, test runs gave large numbers of negative estimates for β_1 and β_2, pointing to a large negative bias, so we do not use this approach.

The third and fourth alternatives are to capture both intensive and extensive effects with a 2-stage estimation procedure, motivated by a model using an underlying latent (unobservable) variable (Eaton and Tamura 1994, Rauch 1999), or to use a Poisson estimator (Santos Silva and Tenreyro 2006, Bergstrand and Egger 2011, p.30f). Comparing these methods is beyond the scope of this thesis.

We choose the first option here and omit zero trade values, so our estimations are only for the effect of the explanatory variables on the intensive margin.

Note that conventional OLS estimation also uses "one-off trade" of country-pairs with data for only one period (for a sector or goods type). The alternative within or first-difference estimators described below use the variation in trade and explanatory variables over at least two periods of an existing bilateral trade relation, matching the definition of the "intensive margin". This is a first reason why these estimators are more appropriate.

3.2.3 Basic model specification

Our gravity model is similar to the one in Feenstra et al. (2001). Their specification, based on equation 2.1.3, reads

$$\ln(X_{sijt}) = \alpha_0^{st} + \beta_1^{st}\ln(Y_{it}) + \beta_2^{st}\ln(Y_{jt}) + \beta_3^{st}\ln(D_{ij}) + \beta_4^{st}CONT_{ij} + \beta_5^{st}LANG_{ij} + \beta_6^{st}FTA_{ijt} + \beta_7^{st}REM_{ijt} + \epsilon_{sijt} \qquad (3.2.1)$$

where in a period t, X_{sijt} denotes exports from country i to country j in sector s (FMR use aggregate goods types instead of our more detailed sectors, see below), Y_{it} and Y_{jt} are the countries' real GDPs, ${\beta^{st}}_1$ and ${\beta^{st}}_2$ are the elasticity of exports with respect to exporter's and importer's GDP, respectively, and ${\alpha^{st}}_0$ is constant across country-pairs (capturing world GDP in their model). They include a number of ad-hoc auxiliary country-pair variables known to affect bilateral trade: geographical distance D_{ij}, dummies for contiguity $CONT_{ij}$, common language $LANG_{ij}$ and common membership in a free trade agreement FTA_{ijt} , and a "remoteness" measure REM_{ijt}. Finally, $\in_{sijt}$ is an orthogonal error term.

We use the same auxiliary variables (see section 3.2.7 for a detailed discussion and section 3.3 for exact definitions and sources), except for the remoteness measure: FMR define REM_{ijt} as "remoteness of *j*, given *i*, equal to the GDP-weighted negative of distance" (Feenstra et al. 2001, p.441). As it is not clear what that means, remoteness of country j is defined here as

$$REM_j^t = \left(\sum_k \frac{Y_k^t}{D_{ij}}\right)^{-1} \qquad (3.2.2)$$

as in Head (2003), where inclusion of this variable is theoretically suggested only for the importer, but we include it for the exporter as well as a control. Additionally, we include the trade partners' populations, P_{it} and P_{jt} (Cheng and Wall 2005; Aichele and Felbermayr 2011), and a variable for Kyoto Annex B status of the trade partners, K_{ijt} (see sections 2.4 and 3.3.2), obtaining as a basic model:

$$\ln(X_{sijt}) = \alpha_0 + \beta_X\ln(Y_{it}) + \beta_M\ln(Y_{jt}) + \beta_3\ln(P_{it}) + \beta_4\ln(P_{jt}) + \beta_5\ln(D_{ij}) + \beta_6 CONT_{ij} + \beta_7 LANG_{ij} + \beta_8 FTA_{ijt} + \beta_9 REM_{it} + \beta_{10} REM_{jt} + \beta_{11} K_{ijt} + \epsilon_{sijt} \qquad (3.2.3)$$

This is the most restricted formulation, with constant parameters across all sectors, periods and country-pairs, which we call the "full pooling" (FP) model. Our "default" estimation is by OLS, and we routinely provide heteroskedasticity-robust standard errors[19].

[19] We use estimates of the covariance matrix with diagonal elements weighted according to the "HC1" scheme introduced by MacKinnon and White (1985) and described in Zeileis (2004); Long and Ervin (2000) report the best small-sample performance for the "HC3" weighting scheme, but this is computationally expensive, and we don't have to deal with very small samples in our OLS estimations.

Our initial dataset has three "dimensions", country-pairs (113*112), time (15 periods), and goods types (3) or sectors (41), and even if we omit occurrences of zero trade, it is very large. This raises the question if the same model is appropriate for the entire dataset. Even if the same functional form is imposed, as we do here, coefficients may vary between sectors or over time, which we will check in the next section. Parameter variations over country-pairs will be discussed separately, as this is related to other estimation approaches.

3.2.4 Pooling or heterogeneity across the sector and time dimension

If the same variant of the gravity model applied to trade in all sectors, estimations of equation 3.2.3 using the full sample or any subset of sector-wise trade flows would always yield the same coefficients in the limit (apart from the intercept, which in our logarithmic specification scales the starting point volume of trade for each sector). If trade followed a gravity model for all sectors, but with a different underlying theory depending on the sector or goods type (e.g. monopolistic competition or reciprocal dumping, as in FMR), the basic functional form of equation 3.2.3 would still hold for all subsamples, but some coefficients would vary, depending on the sector (in particular, we expect the relative size of β_X and β_M for the GDP's to vary). While this already provides a theoretical reason for goods-type/sector-wise estimation, and all of the main results that we discuss in section 4 will be for separate goods types or sectors, we also need to check that such structural breaks are statistically significant (compare Cheng and Wall (2005) for a systematic comparison of gravity models with different restrictions along the time and country-pair dimension).

Relaxing the restriction implied by equation 3.2.3 along the sector dimension (but maintaining the pooling over periods for the moment) leads to the "goods-type pooling" (GP) model, where we allow parameters to vary between three groups of sectors, according to their goods-type classification. The simple "pooling" (P) model is even less restricted and parameters are estimated separately for each individual sector (at the goods-type aggregation level, only the full and the simple pooling model exist). Separate estimation is equivalent to pooled OLS estimation using a more general specification of equation 3.2.3 in which the *k*-th coefficient is written as a base value β_k^0 and a sum of components specific to the sector subsample, β_k^s, multiplied by subsample dummies *ds* for all but one subsample, e.g. $\beta_3 = \beta_3^0 + d2*\beta_3^2 + d3*\beta_3^3 + \ldots + d41*\beta_3^{42}$ if the subsamples are simply the sectors. We then test the null hypothesis that $\beta_k^s = 0$ for all *k* and *s* simultaneously. The corresponding test statistic[20] is

$$f = \frac{(RSS_{restr} - RSS_{unrestr}) \,/\, ((L-1)K)}{RSS_{unrestr} \,/\, (N-LK)} \;\sim\; F_{N-LK}^{(L-1)K}, \qquad (3.2.4)$$

[20] This kind of F-test is also called a Chow test (e.g. Wooldridge, 2000) for a structural break between two or more pre-defined subsamples. It compares the residual sums of squares of a pooled and the separate OLS estimations, adjusted for the respective degrees of freedom.

where RSS_{restr} and $RSS_{unrestr}$ are the residual sums of squares (RSS) from OLS estimation of the model in equation 3.2.3 with the relevant pooling restriction across sectors and for separate estimations (summed), respectively. N is the size of the full sample, L the number of sector subsamples (separate estimations), and K the number of explanatory variables in the restricted model, including the intercept. This statistic follows an F-distribution with $(L\text{-}1)K$ and $N\text{-}LK$ degrees of freedom.

At the goods type aggregation level, we test the fully restricted FP model against the P model (which has three separate estimations). At the sector level, we first test the FP model against the GP model with its three separate estimations for pools of 14, 13 and 14 sector. Second, we test the restriction of the GP model versus the P model, for homogenous, reference-priced and differentiated goods. In all cases, we obtain values of the f-statistic larger than 400 and p-values below numerical accuracy (see Excel file). Since all restrictions along the sector dimension are rejected and the P model has been identified as the most suitable specification up to now, we will only use goods-type- and sector-wise estimations in the following.

Now, coefficients could also change over time due to exogenous factors affecting some or all sectors, like new production and transport technologies, economic or political crises, etc. We relax the restrictions along the time dimension, at first only for the intercept:

$$\ln(X_{sijt}) = \alpha_0^s + \alpha_t^s + \beta_X^s \ln(Y_{it}) + \beta_M^s \ln(Y_{jt}) + \beta_3^{st} \ln(P_{it}) + \beta_4^{st} \ln(P_{jt}) + \beta_5^s \ln(D_{ij}) + \beta_6^s CONT_{ij} + \beta_7^s LANG_{ij} + \beta_8^s FTA_{ijt} + \beta_9^s REM_{it} + \beta_{10}^s REM_{jt} + \beta_{11}^{st} K_{ijt} + \epsilon_{ij}^{st} \qquad (3.2.5)$$

Here, $\alpha^s{}_t$ are time-specific intercept components (set to zero for the first period) [21], which exploit the panel structure of our data by absorbing unobserved shocks that affect all country-pairs likewise. We call this a "time fixed effect" (time FE) model[22] in the following, emphasizing the similarity to individual fixed effects introduced below. Note that since we estimate for each sector separately, the time FE might be a different one for each sector (denoted by the superscript s, also for all other parameters). As before, we test the time FE model against the more restricted model, in this case the P model, which sets $(T-1)$ of the intercept components to zero. The test statistic is

$$f = \frac{(RSS_P - RSS_{timeFE}) \,/\, (T-1)}{RSS_{timeFE} \,/\, (N-(K+T-1))} \sim F_{N-(K+T-1)}^{T-1} \qquad (3.2.6)$$

The results of this test are presented in Table 12 (columns on the left) in the appendix. The restriction of constant intercepts across time is rejected for all goods types and most sectors (34 out

[21] Time-dependent intercepts are implemented numerically using dummy variables $d\tau^t$ (1 if $t = \tau$, 0 otherwise).

[22] Cheng and Wall (2005) classify both our time-FE and the pooling model as "pooled cross-section models" (PCS), testing only the case with time-dependent intercept.

of 41, with five sectors without significant time FE classified as homogenous). The time FE model of equation 3.2.5 is the first model for which we will discuss detailed results in section 4.2.[23]

Apart from the intercept, the time FE model still imposes time-independent parameters. If we relax this final restriction, we obtain

$$\ln(X_{sijt}) = \alpha_0^s + \alpha_t^s + \beta_X^{st}\ln(Y_{it}) + \beta_M^{st}\ln(Y_{jt}) + \beta_3^{st}\ln(P_{it}) + \beta_4^{st}\ln(P_{jt}) + \beta_5^{st}\ln(D_{ij}) + \beta_6^{st}CONT_{ij} + \beta_7^{st}LANG_{ij} + \beta_8^{st}FTA_{ijt} + \beta_9^{st}REM_{it} + \beta_{10}^{st}REM_{jt} + \beta_{11}^{st}K_{ijt} + \epsilon_{ij}^{st} \qquad (3.2.7)$$

This is the standard sector-wise cross-sectional (CS) model, similar to that used by FMR[24] (compare equ. 3.2.1). Again, we can test the null hypothesis that the βs are not time-dependent - similar to the first test for sector heterogeneity, we can think of all coefficients in equation 3.2.7 as sums of a base value and year-wise components, with the null hypothesis that all year-wise components are zero (only the base values remain, except for the intercept). We obtain the test statistic

$$f = \frac{(RSS_{timeFE} - RSS_{CS}) \,/\, ((T-1)(K-1))}{RSS_{CS} \,/\, (N-TK)} \sim F_{N-TK}^{(T-1)(K-1)} \qquad (3.2.8)$$

RSS_{timeFE} is from the time-wise pooled estimation of equation 3.2.5 for the respective sector (or goods type), and RSS_{CS} is the sum of RSS from as many sector-wise estimations of equation 3.2.7 as there are periods (T). As before, N is the full sample size and K is the number of explanatory variables including the intercept. In the unrestricted model, TK is the number of coefficients including the time FE, $(T\text{-}1)(K\text{-}1)$ of which are set to zero under the null hypothesis (which still allows for time FE, so $(K+T\text{-}1)$ coefficients remain). The results of the test based on equation 3.2.8 for aggregate goods trade will be discussed along with the estimation results for the CS model in section 4.2. We do not consider a CS model at the sector level, where data is too sparse for many sectors and years to estimate this number of parameters.

3.2.5 Heterogeneity across country-pairs and fixed effects

While potential heterogeneity along the sector and time dimensions can be accommodated with the above models, they all impose homogeneity across country pairs. Similar to the approach for *sectors*, we could also estimate separately for subsamples of *country-pairs*. For example, FMR repeat their cross-sectional, goods-type-wise estimations for the subsample of trade among OECD countries as a qualitative robustness check. In our case, a grouping according to the importer's and exporter's status in Annex B to the Kyoto Protocol would be an obvious choice (see section 3.3.2); we could distinguish trade among B-countries, among non-B-countries, from B- to non-B -countries and vice

[23] Since we pool over periods for estimation the time FE model, and later compare it to fixed effects models, we report the same heteroskedasticity- and autocorrelation-consistent error terms as for the within and FD estimators (see footnote 30)
[24] While the sector aggregation level in FMR is motivated by theory, the choice of year-wise estimations by OLS is made ad hoc, and only qualitative consistency of all estimations with theory is discussed.

versa. We will only briefly discuss the results of such subsample estimations because we prefer a different approach to heterogeneity across country-pairs that is robust against an important source of omitted variable bias:

Although trade partners' GDP and distance have high predictive power for bilateral trade, it is clear that other geographic, historical, cultural and economic factors will also have a significant effect. The specifications above attempt to control for this by including auxiliary variables like contiguity or common language dummies. But they still impose constant effects of these variables across country-pairs; furthermore, unbiased estimates require that the set of variables be complete for all country-pairs and sectors, which is a rather strong assumption. For example, depending on the sector, additional variables like resource endowment, a skilled workforce or infrastructure may play a role.

Since we are not confined to cross-sectional data, but have a panel data set, we can allow for unobserved heterogeneity between country-pairs by including country-pair-specific intercept components (fixed effects $\alpha^{s}{}_{ij}$). Along with unobserved effects, these absorb the effects of all time-invariant country-pair variables used before: distance, contiguity, common language, and membership in free-trade agreements for country-pairs where this did not change over the covered period. This also conveniently rids us of problems related to the definition and uniform treatment of "contiguity" and "distance" (Cheng and Wall 2005), which could be measured e.g. between capitals or economic centers, and could be of different importance depending on land and sea transport options (see Head and Mayer, 2001). "Individual" country-pair FE are commonly included in models where slope coefficients are constant across time and country-pairs, while time FE may be retained (see Cheng and Wall 2005). The following specification thus builds on equation 3.2.5, rather than the CS model of equation 3.2.7, and is called a "two-way FE" model:

$$\ln(X_{sijt}) = \alpha_0^s + \alpha_t^s + \alpha_{ij}^s + \beta_1^s \ln(Y_{it}) + \beta_2^s \ln(Y_{jt}) + \beta_3^{st} \ln(P_{it}) + \beta_4^{st} \ln(P_{jt}) + \beta_8^s FTA_{ijt} + \beta_9^s REM_{it} + \beta_{10}^s REM_{jt} + \beta_{11}^{st} K_{ijt} + \epsilon_{ij}^{st} \qquad (3.2.9)$$

Results of an F-test of the two-way FE model against the time FE model, to test the joint significance of individual effects, are discussed in the results section[25].

Alternatively, we could include separate FE for the exporter and importer. Cheng and Wall (2005) observe that this is equivalent to cross-pair restrictions on the country-pair FE, but reject them in favor of the less restricted model (see also Egger and Pfaffermayr, 2003; Bergstrand and Egger, 2011). The restriction of symmetric country-pair FE ($\alpha_{ij} = \alpha_{ji}$) is also rejected. We do not include these additional specifications and tests here.

[25] This test is implemented in the *plm* package for *R* as *pFtest()*.

In fixed effects models, the time- or individual-specific contributions to the intercept are interpreted as additional coefficients which could in principle be estimated, e.g. by explicitly including dummy variables – although their value might be of little interest by itself. If the latter is the case, the slope coefficients are commonly estimated by pooled ordinary least squares (OLS) after some transformation eliminated the FE (see below). Correlation of the individual effects with the other explanatory variables does not lead to inconsistency. If, on the other hand, the time- or individual effects can be assumed to be randomly distributed and independent of the other explanatory variables, this is called a "random effects" (RE) model, estimated e.g. with a generalized least squares estimator (Verbeek 2004). However, the assumptions underlying the RE model are not very plausible in our case, where "individuals" are country-pairs and each country "one of a kind" (Verbeek 2004, p.351) with non-random characteristics potentially correlated to our variables. While there is a formal test available for the null hypothesis that the individual effects and the explanatory variables are uncorrelated (Hausman test, see e.g. Verbeek 2004), we skip this here for brevity and directly use the FE approach, as most of the literature on gravity estimations does (Bergstrand and Egger 2011).

3.2.6 Estimation of FE models using "within" and first-difference estimators

For the time FE, we explicitly include dummy variables as described above, since the econometrics package that we use (see below) is inefficient with two-way FE in unbalanced panels (Millo 2009). Individual FE are not estimated directly[26], but eliminated before estimation with one of two common transformation: With "time demeaning", all variables in the model are restated as the difference from their respective average over all available periods (e.g. $\ddot{Y}_{ijt} = Y_{ijt} - \sum_t Y_{ijt} / T$). Pooled OLS estimation of the transformed model is then called "within" (or fixed effects) estimation. Alternatively, for each variable, we can subtract adjacent variable values (the value in the first available period is subtracted from the second value, the second value from the third, and so on), before pooled OLS estimation. This procedure is called a "first difference" (FD) estimator.

Note that although *pooled* OLS imposes that a change in a variable between two years or two individuals has the same effect, estimations with the *transformed* data are driven by "within-country-pair" variations, rather than variation between country-pairs (Verbeek 2004, p.347).

The within and FD estimators are similar in the sense that both are consistent and unbiased under similar assumptions (see Wooldridge 2000, pp.447-448), most importantly strict exogeneity of the explanatory variables[27]. Subtle differences arise with respect to serial correlation in the error term

[26] The estimator explicitly estimating the FE is called the Least Squares Dummy Variables estimator, and gives the same results as the "within" estimator (unless time FE are also included, which makes a small difference since an additional restriction is required, e.g. setting the first time FE to zero, see Cheng and Wall 2005)

[27] Strict exogeneity requires that in any period t, the expected value of the error term ε_{it}, conditional on the explanatory variables in all T periods and the fixed effect, is zero: $E(\varepsilon_{it} \mid x_{i1}, \ldots x_{iT}, a_i) = 0$

(Croissant and Millo 2008, pp.27-28): If the original error terms $\varepsilon_{ij}{}^{t}$ are serially uncorrelated the transformed errors of the first-differenced model, $\Delta\varepsilon_{ij}{}^{t} = \varepsilon_{ij}{}^{t} - \varepsilon_{ij}{}^{t-1}$, are correlated ($cor(\Delta\varepsilon_{ij}{}^{t};\ \Delta\varepsilon_{ij}{}^{t-1}) = -0.5$), and the within estimator is more efficient[28]. The FD estimator, on the other hand, is more efficient when there is no serial correlation in the first-differenced errors, which means that the original error terms are following a random walk. In our case, this could be explained by trade relations and volumes changing slowly due to the underlying international production networks and capital stocks.

The presence of either type of serial correlation can be tested using a first-difference based test by Wooldridge (2002)[29]. If one of the two null hypothesis, "No serial correlation in first-differenced errors" or "No serial correlation in original errors", is not rejected, while the other is rejected, the estimator corresponding to the Null that is not rejected should be used. If both are rejected, both estimators will suffer from serial correlation, and we have to use an autocorrelation-robust covariance matrix for inference. For our within and first-difference estimations, we only report heteroskedasticity- and autocorrelation-robust standard errors in this work, computed using the Arellano method (Croissant and Millo 2008)[30].

3.2.7 General equilibrium effects, "remoteness" and endogenous variables

Anderson and van Wincoop (2003) model implicit price indices to capture multilateral equilibrium effects, in particular the evolution of other trade options, e.g. changes of third countries' size and bilateral trade frictions ("multilateral resistance"). They argue that this more appropriate than ad-hoc "remoteness" measures based only on GDP-weighted geographical distance, like in our specifications above, and report a significant effect e.g. on estimations of border effects. However, their approach requires a solution not only to a gravity equation, but simultaneously for the unobservable implied prices; custom programming of a nonlinear solver for the structural general equilibrium model is beyond the scope of this thesis.

Bergstrand (1985) and later Baier and Bergstrand (2009) approximated the price indices by GDP deflators, but this may not fully reflect all costs of trade (Feenstra 2004).

Another simple approach employs country fixed effects, which inter alia absorb all country-specific constant trade costs (see Feenstra 2004 for a summary), or country-time effects to reflect the change of trade barriers over time (Bergstrand and Egger 2011, pp.28-29). As noted above, country fixed effects are a restricted form of country-pair fixed effects, which we already in our model, while the

[28] This efficiency comparison requires homoskedasticity.
[29] Implemented in the *plm* package as function *pwfdtest*.
[30] The Arellano method is suitable in a panel setting where the original White estimator for heteroskedasticity-consistent covariance matrix is inconsistent for fixed T and growing n (Croissant and Millo 2008, p.31)

restriction is rejected by Cheng and Wall (2005); country-time effects, on the other hand, preclude the separate estimation of GDP effects, which we need to draw conclusions about the underlying mechanism.

We thus keep on using the ad-hoc “remoteness” measure in all models, and country-pair FE in the panel estimations, accepting a potential bias of the parameter estimates for variables related to trade barriers (distance, contiguity, language in the OLS estimations, and membership in free-trade agreements), since they are not in our focus anyway. However, we have to keep in mind that the parameter estimate for the Kyoto variable may also be affected.

Finally we note that the Kyoto variable may be endogenous, like other environmental policies discussed in section 2.3, and also the dummy we use for mutual membership in free agreements. This could be addresses by instrumental variables and two-stage least squares (2SLS) estimation or the or country-year effects that we already discussed above (Bergstrand and Egger 2011, Aichele and Felbermayer 2010).

3.2.8 Numerical implementation

All data manipulations and estimations were performed in *R*, version 2.12.1, using the packages *filehash* to store large datasets on the hard drive rather than RAM, *plm* for panel estimations (Croissant and Millo 2008) and *sandwich* (Zeileis 2004) for computation of robust standard errors. The latter had to be slightly modified to deal with unbalanced panels under first-differencing, since a previous debug to that end did not take into account cases where the first period is missing.

3.3 Data sources and issues

3.3.1 Exports and emissions embodied in trade

We used the dataset of Peters et al. (2011) for the emissions embodied in trade (EET). For the decomposition and the gravity model, we also need bilateral trade data consistent with the EET data in construction, resolution and coverage (sectors, regions and time). Along with the EET data, we obtained the underlying GTAP trade data time series described in the supporting information of Peters et al. (2011b) from the authors and performed the manipulations outlined therein to arrive at a matching data set.

The starting point is a four-dimensional data set of the exports among 113 regions in 41 goods sectors (there is no data for a 42nd sector, raw milk), for the years 1992 to 2006 (for region and sector definitions, see Table 8 in the appendix or GTAP 2011). This is only used for the *split* of each region’s annual *goods* exports into trade partners and sectors, while the total export volume is taken from the United Nations Statistics Division (UNSD 2011). The latter reports cumulated exports in goods and services, while our trade time series does not include services, so we take regions' service shares in exports[31] for the years 1997, 2001 and 2004 from GTAP versions 5, 6 and 7, and extrapolate them to the periods from 1992 to 1998, 1999 to 2002, and 2003 to 2006, respectively. Since GTAP 5 and 6 use only 66 and 87 regions, the service share is also assumed to be the same for all "members" of a region that are only reported in detail in GTAP 7. From the detailed goods exports data and the service shares, we construct the complete exports split with respect to goods and service sectors and trade partners (normalized to one for each exporting region) and apply it to the UNSD exports data. Since EET are given in physical units (Mt of CO_2), we choose the UNSD exports time series in constant rather than current terms[32], converted to US$ for the year 2005 using market exchange rates (without adjustments for purchasing power, which are important sometimes in the domestic context, but less so for international trade).

For our motivational graphs in section 1, we extended the data set to 1990 and 2008, using the sector-region-splits of 1992 and 2006, respectively, and also included service trade in the aggregates. For our main analysis, we only use the 1992-2006 *goods* trade data, to avoid bias due to an over-emphasis of the split data for 1992 and 2006[33].

The resulting data panel is unbalanced, since trade and EET data is missing for some regions and years[34]. Although this is no major obstacle for estimation and methods for unbalanced panels are available, numeric efficiency is impaired for some methods of the *plm* library for *R*, namely fixed effects estimation with individual and time effects (Millo 2009). Instead of specifying "two-way effects" in the convenient *plm()* command, we had to use a model with explicit time dummies in combination with the "individual effects" option in *plm()* in these cases.

3.3.2 Classification of trade in sectors into goods types

Rauch's (1999) classification (see section 2.2.2) maps commodity trade at the 4- and 5-digig SITC, revision 2 levels to three goods types. Since our trade data is at the much coarser GTAP sector level, we could not directly use this mapping but had to construct our own. We obtained Rauch's classification as a spreadsheet from Haveman (2011) together with a verbal description of the SITC levels. For additional information, we used the mapping from SITC revision 2 to ISIC revision 3 by Affendy et al. (2010) and a verbal description of ISIC rev. 3 by Hutcheson (2006). We then matched

[31] Service exports are actually reported separately again for 15 sectors in GTAP, but since we only analyze goods trade here, we don't need this level of detail.

[32] The GTAP data are in current US$, but this is no problem as we only use it for year-wise sector and region splits.

[33] Note that this will affect our definition of Annex-B and non-Annex-B countries below.

[34] TSTRD and EET data is missing for imports and exports of the Czech Republic, Slovakia and Ethiopia ('90-'92), Luxembourg ('90-'98), Botswana and "Rest of South African Customs Union" ('90-'99, GTAP-Code XSC). For Norway, although EET data is reported by Peters et al. (2011), neither EET nor TSTRD data for Norwegian *exports* is contained in the data set I received, while *import* data is included.

SITC commodities to GTAP sectors (2011) by hand and classified a GTAP sector according to the majority of classifications of the corresponding commodities (see Excel file).

Rauch (1999) provides two classifications: *"Because ambiguities arose that were sometimes sufficiently important to affect the classification at the three- or four-digit level, both 'conservative' and 'liberal' classifications were made, with the former minimizing the number of three- and four-digit commodities that are classified as either organized exchange or reference priced and the latter maximizing those numbers."* (Rauch 1999, p.15) We only report and discuss the results for the conservative classification here. The liberal classification gave similar results which are included in the accompanying Excel results file.

3.3.3 Climate policy, Kyoto membership and emission intensity of exports

We use the category of "having an emissions reduction target in Annex B of the Kyoto Protocol" in two ways here: For groups of countries that we talk about or which we pool for analysis, and for a specific country-pair's relative Kyoto status, captured by the Kyoto variable in equation 2.4.1.

The latter is time-dependent, and we count ratification of the Kyoto protocol (commitment to a target) only to the next year if the ratification date is in the second half of a calendar year.

Whenever we use *groups* of countries, a country is labeled a "B country" if it was assigned an emission reduction target in Annex B of the Kyoto protocol, irrespective of the actual status and timing of ratification of the protocol by that country. In particular, this includes the USA as the only country with a target in Annex B that did not ratify the protocol to date, and countries that ratified the protocol after 2006, the time horizon of our analysis. We use the distinction between B- and non-B countries because it is in keeping with Peters at al. (2011), which motivated our analysis and serves as a reference for our results, and because it is currently the most relevant categorization in the international climate policy context. Alternatively, we motivate country groupings normatively by countries' responsibility for emission reduction due to current or historical emissions, economic capacity for mitigation, equity and growth concerns, which suggests criteria like emission intensity, cumulative past emissions or per capita income. The analysis of the determinants of EET patterns would motivate criteria like a country's productivity or resource endowments. However, having an Annex B target, while subject to political negotiations and thus some arbitrariness, can be expected to be correlated with many of these alternative criteria and to lead to similar outcomes[35]; also, we explicitly control for some of the other criteria in our estimations (per capita income).

Specifically, in our detailed decomposition results in Table 11, we explicitly distinguish between B- and non-B countries and the USA, so that we can compare similarly developed (groups of) regions

[35] At least, a grouping Annex B / non-Annex B is not more arbitrary than one based on, say, OECD membership.

with and without broad climate policy measures. In some of our estimations in section 4.2, we groups of trade flows between non-B and B regions (and vice versa and within these groups) with “B” now including the USA, which reflects the mentioned assumption that these countries share important (but unobserved) socio-economic characteristics that might affect trade and thus the parameter estimates. This is similar to one of the sensitivity tests by FMR, who re-run their cross-section OLS estimations on the subsample of trade among OECD countries to check if their results are really due different goods types and not country characteristics. Potential differences in the estimates are thus not to be interpreted as effects of implemented emission reductions policies.

Some countries only ratified the Kyoto Protocol in 2007 and later, which is beyond the time horizon of our analysis: Australia ratified in December 2007 and Croatia in May 2007. While we treat them as “B” countries whenever we pool or aggregate data for decomposition, statistical analysis or plots, their ratification is too late to affect variation in the “Kyoto variable” of our estimations.

The same applies to the compound region “Rest of North America” used GTAP7 (coded as XNA), which comprises Greenland, which is represented by Denmark in the Kyoto protocol, Bermuda, and Saint Pierre and Miquelon. Bermuda has a GDP and population significantly higher than the other two, so we define the Kyoto status of "XNA" according to Bermuda's status. Bermuda is a UK overseas territory and represented by the UK with respect to the Kyoto Protocol. The UK ratified the protocol in 2002, with an extension to Bermuda in 2007 (UNFCCC 2011a, footnote 7), so we treat XNA as a “B” region throughout (unlike Peters et al., 2011), but this doesn’t affect the Kyoto variable.

GTAP7 definitions in two more cases lump countries of different target status into one region: “Rest of European Free Trade Agreement” (XEF) consists of Liechtenstein, which ratified in 2005, and Iceland, which ratified in 2002. Since the latter’s GDP is more than double the size, and Iceland can be assumed to be more relevant for goods production trade than Liechtenstein, we list XEF as a Kyoto Annex B country from 2002 onwards. “Rest of Europe” (XER) lumps non-Annex B countries such as Kosovo, Bosnia, Montenegro, Macedonia, Serbia, Andorra and San Marino, and Monaco as the only Annex-B country (ratification in 2006), which we neglect, so XER is treated as a non-B country here.

3.3.4 Other explanatory variables in the gravity model

Market size in gravity models could be measured in terms of GDP or population; many models include both as explanatory variables (Bergstrand and Egger 2011). This implicitly also controls for per capita GDP, which is sometimes argued to be a proxy for consumption patterns. Our GDP data is from the UNSD (2011) and population data from the Worldbank (2011). A table with distances measured as great circle distances between capitals, contiguity and common language dummies are from Mayer and Zignago (2006). For our dummy variable FTA_{ijt} for common membership in a regional free trade agreement, we consider the EU, NAFTA, ASEAN, Mercosur and Caricom with dates taken from Wikipedia (2011), and Anzcerta (2011). We do not use additional dummies for former colonial ties.

3.3.5 Descriptive statistics of the trade dataset

Some descriptive statistics on the trade dataset we constructed are given in Table 9 in the appendix, including the shares of zeros for goods types and sectors, which was discussed above and plotted in Figure 3. We also included the number of of country-pairs trading per sector, that is, the distribution of the number of available periods and data on the concentration of trade into the hands of a few exporters and importers.

Additionally, we obtained the number of data points per sector and year, which rises from 1992 to 2006 for all sectors (except "wol"), by more than 60% on average.

Under the conservative classification, differentiated goods make up 66.3% of trade volume in 1992, reference-priced goods 18.4% and homogenous goods 15.3%. World trade grows by 168% for differentiated, 135% for reference-priced and 143% for homogenous goods, so the respective shares are relatively stable until 2006.

The shares of each goods types in interregional trade flows are given in Table 2. Three quarters of the exports from B countries to both B and non-B countries consist of differentiated goods, while the share of homogenous goods falls below 10%. Non-B countries initially export 50% and later over 60% differentiated goods, while the share of homogenous goods falls from one third to one fifth.

1992	**B**		**non-B**	
B	hom	10.9%	hom	8.1%
	ref	19.0%	ref	15.9%
	diff	70.1%	diff	75.9%
nB	hom	33.0%	hom	29.6%
	ref	18.0%	ref	21.3%
	diff	49.0%	diff	49.0%

2006	**B**		**non-B**	
B	hom	7.9%	hom	7.8%
	ref	16.7%	ref	15.3%
	diff	75.5%	diff	76.9%
nB	hom	21.2%	hom	21.5%
	ref	13.8%	ref	16.7%
	diff	64.9%	diff	61.8%

B: Annex B countries incl. USA, nB: non-Annex B countries
hom: homogenous goods, ref: reference-priced goods, diff: differentiated goods

Table 2: Share of goods types (conservative classification) in interregional trade volumes

4 Results

4.1 Decomposition analysis

The results of the decomposition of the change in total emissions embodied in trade between 1992 and 2006, according to the method described in section 3.1, are summarized in Table 3.

Totals in first and last period		**B to nB**	**nB to B**	**B to B**	**nB to nB**	**Total EET**	**Total B<>nB**	**Net B to nB**
EET ‘92	abs. [Mt CO_2]	587	1011	1584	475	3656	1598	-424
EET ‘06	abs. [Mt CO_2]	781	1996	1778	1794	6349	2777	-1215
Trade ‘92	[10^9 US$, 2005]	687	883	2437	403	4410	1570	-196
Trade ‘06	[10^9 US$, 2005]	1832	2322	5269	1955	11379	4154	-490
CO_2 intensity ‘92	[kgCO_2 / $]	0,85	1,14	0,65	1,18	0,83	1,02	-
CO_2 intensity ‘06	[kgCO_2 / $]	0,43	0,86	0,34	0,92	0,56	0,67	-
Decomposition summary								
ΔEET	**abs. [Mt CO_2]**	**194**	**985**	**194**	**1320**	**2693**	**1179**	**-791**
	vs. '92 [%]	33%	97%	12%	278%	74%	74%	187%
thereof, due to…								
- scale change	**abs. [Mt CO_2]**	**592**	**1607**	**1291**	**1540**	**5030**	**2199**	**-1015**
	direct effect	561	1620	1281	1461	4923	2181	-1060
	vs. '92 [%]	101%	159%	82%	324%	138%	138%	239%
- composition change	**abs. [Mt CO_2]**	**61**	**89**	**198**	**66**	**413**	**150**	**-28**
	direct effect	-101	-98	24	-239	-413	-199	-4
	vs. scale eff. [%]	10%	6%	15%	4%	8%	7%	3%
	vs. '92 [%]	10%	9%	12%	14%	11%	9%	7%
- intensity change	**abs. [Mt CO_2]**	**-459**	**-710**	**-1295**	**-286**	**-2750**	**-1169**	**252**
	direct effect	-590	-828	-1465	-619	-3503	-1418	237
	vs. scale eff. [%]	-77%	-44%	-100%	-19%	-55%	-53%	-25%
	vs. '92 [%]	-78%	-70%	-82%	-60%	-75%	-73%	-59%

Cumulated yearwise decomposition with refined Laspeyres index method (interaction terms distributed symmetrically). B = Annex B countries (incl. USA), nB = non-Annex B countries, EET=emissions embodied in trade.

Table 3: Decomposed changes in CO_2 emissions embodied in goods trade, 1992-2006

The total changes of emissions embodied in trade are positive for all flows and thus also for the total EET and total emissions embodied in bidirectional trade between Annex-B countries and non-Annex-B countries, which makes up 44% of total EET in both '92 and '06. But the increase is higher for trade flows originating in non-B countries, e.g. the growth of embodied emissions for non-B to B trade (nB2B) was roughly five times the increase for B to non-B trade (B2nB). In relative terms, EET for B2nB increased by 33% with respect to their 1992 levels, while in the opposite direction, they increased by 97% (the increase within the non-B group was almost three times higher again). Consequently, the negative balance of EET between B and non-B countries almost tripled.

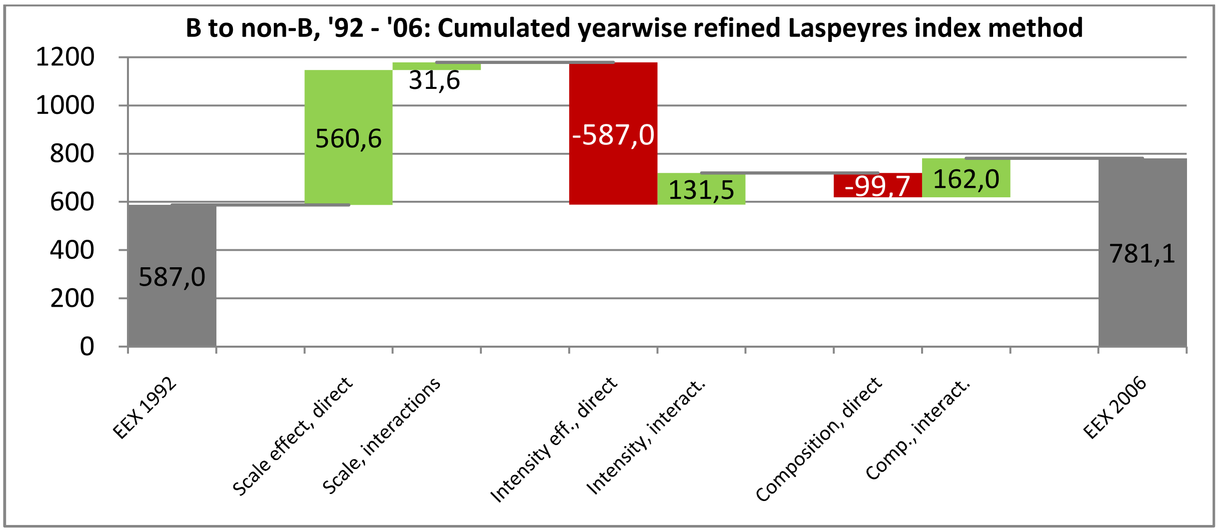

Figure 4: Decomposition of changes in CO_2 emissions embodied in goods trade [Mt CO_2], from Annex B countries (incl. USA) to non-Annex B countries

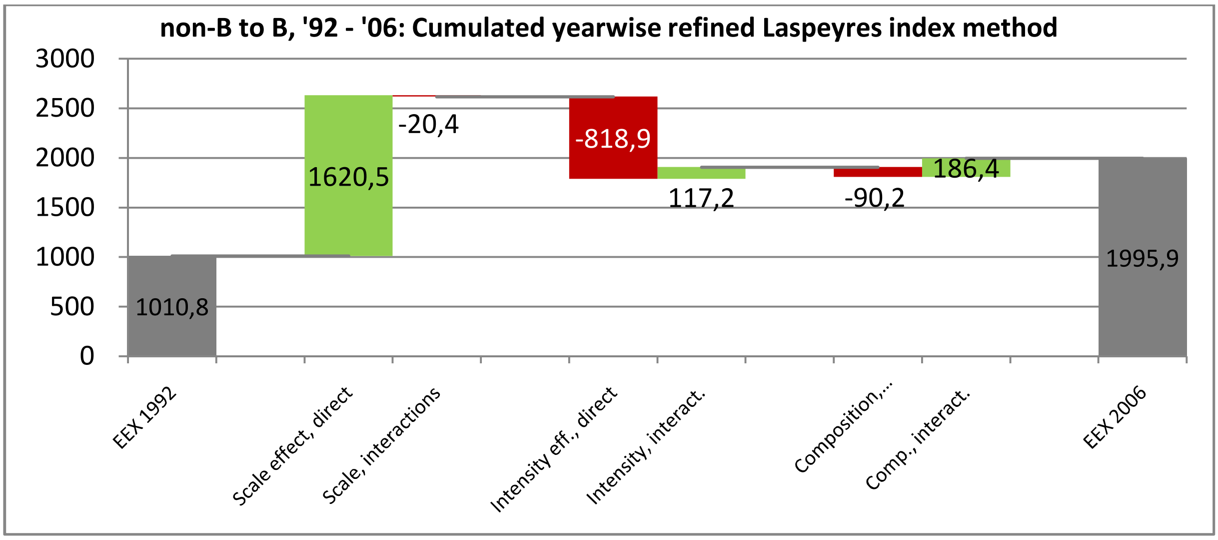

Figure 5: Decomposition of changes in CO_2 emissions embodied in goods trade [Mt CO_2], from non-Annex B countries to Annex B countries (incl. USA)

For trade between B and non-B countries, summarized in Figure 3 and Figure 4, scale changes are by far the largest contributors: While the positive scale effect for B2nB exports is of the same size as the B2nB total EET level in the year 1992, around 590 Mt CO_2, the scale effect in the reverse direction makes up 1607 Mt CO_2, the 2.7-fold, which by itself would have deepened the imbalance between B and non-B countries to 3.4 times the level of 1992. This is partly offset by negative intensity changes, which are 1.5 times larger for nB2B (-710 Mt CO_2) than for B2nB (-459 Mt CO_2), so the net intensity effect amounts to 60% the size of the imbalance in 1992 (with the opposite sign). Note that this is due to the larger volume of nB2B trade, but doesn't mean that nB2B trade has become "cleaner" at a faster rate than vice versa - the opposite is the case, since the intensity decrease makes up only 44% of the scale effect for nB2B, but 77% for B2nB trade. This is also apparent if we take into account the value of traded goods : While the B2nB average intensity fell from 0.85 kg CO_2/$ to 0.43 kg CO_2/$, it only fell from 1.14 kg CO_2/$ to 0.86 kg CO_2/$ for nB2B, so the gap actually widened. The composition effects are small compared to the other effects, making up 9-14% of the respective trade flow's 1992 level, and positive; the net impact is slightly adding to the imbalance. Without the interaction terms, which are dominated by the much larger changes in scale and intensity, the individual terms would have been of roughly the same magnitude, but with a negative sign (and still a negative impact on the balance).

Within the B and non-B blocks, while the scale effect is large in both cases, these gains are entirely offset by intensity decreases among the B countries, so the overall effect is of the same size as the composition effect, which is only 1/7th of each of the other effects. At the other extreme, in the case of intra-non-B trade, intensity improvements still leave 81% of the scale effect, so the total change in EET is the highest of all four aggregate trade flows.

We see that the direct effects are large and the interaction terms' impacts small, generally significantly below the difference between the nB2B and B2nB effects, so using a decomposition method with a different allocation of the residuals would not have change the qualitative results.

More detailed decompositions separating five aggregate goods sectors (as in Peters et al., 2011, see Table 8) , which are given in Table 10 the appendix, show that the results above are mainly driven by "manufacturing" and "energy-intensive manufacturing", followed at some distance by "mining", while "agriculture" and "food" are much less important. We note that the three main aggregate sectors all show the same growing imbalance of EET between n and non-B countries that we already saw for the total flows, but for different "reasons": For manufacturing, both the scale and the intensity effects for nB2B embodied emissions are five times larger than those for B2nB, and similarly for mining. The intensity effects are roughly half as large as the corresponding scale effects for both directions, so the offset due to decarbonisation in B2nB countries is quite symmetric (as far as the flows between the two country groups are concerned). For energy-intensive manufacturing, the nB2B scale effect is only 50% larger, but the intensity effect is 50% smaller than for B2nB: Decarbonisation in the exporting countries compensates 90% of the volume increase in B2nB exports, but only 30% for nB2B exports. This gap is even more pronounced for trade among B-countries and among non-B countries, where the intensity effect is -109% and -16% of the respective scale effect for energy-intensive manufactures. The large difference in decarbonisation effects for trade within the country groups extends to manufacturing and mining (and also the aggregate food sector).

We summarize that, first, EET (total and between non-B and B) grew over 70% mainly due to increases in trade volumes (in particular from non-B countries) , which were almost double the size of intensity improvements; second, that disproportionately high growth of non-B to B embodied emissions, which were already 70% larger than B to non-B embodied emissions in 1992, enlarged the net EET deficit of B countries; third, that relatively higher intensity improvements for B exports, especially for energy-intensive manufactures and thus probably largely due to decarbonisaton in the energy sector, were not sufficient in absolute terms to offset this deficit; and fourth, that composition effects did not play a major role for the EET changes, although the export structure in 1992 is of course important as it defines the starting point.

In the next section, we want to analyze potential determinants of these developments with an econometric model. We found that scale effects are most important, while composition effects play a minor role. We will thus choose a gravity model of trade, which has been empirically successful and is suitable to explain scale effects via the direct link to GDP growth. Sector-wise estimations will lead to a different impact of the explanatory variables in each sector, so the model can also reproduce composition changes. For a final illustration of the estimation results, we will use them to predict trade, weight the results by the observed sector-wise export emission intensities to obtain predicted EET, and repeat the decomposition (section 0).

The gravity model that we are going to use also contains a Kyoto dummy as a control for potential effects of climate policy on trade. Before discussing the more reliable statistical results, we could attempt a preliminary check using a modified decomposition: Since the USA have not ratified the Kyoto protocol, but can be assumed to have an economic structure similar to other B countries, they can serve as a "control group" to some extent. Table 11 in the appendix presents the a decomposition with three regions ("USA", "B*" denoting Annex B countries without the USA, and "non-B").

If climate policy with a regulation of production-related emissions was the main determinant behind trade patterns, we would naively expect that exports from B* to non-B countries show a stronger carbon intensity decrease than US exports to non-B, which is indeed the case (0.91 to 0.41 vs. 0.72 to 0.47). Then one would also expect that B* countries now cover a larger share of their dirty goods consumption by imports, because domestic production became more expensive. This would imply that the decarbonisation of total B* imports (from non-B and the USA) should be lower than the decarbonization of total US imports. But we observe that B* countries import 941 (1577) $MtCO_2$ and 880 (1981) billion US$ worth of goods in 1992(2006), implying an intensity decrease of over 34%, while the USA import 487 (986) MtCO2 and 674 (1742) US$, which means that intensity decreased by less than 28%. Thus, it is not immediately evident how the Kyoto Protocol would have dominated trade patterns. However, these results are difficult to interpret due to equilibrium effects, as it is not clear how changes of domestic production costs will affect prices, which in turn affect consumption patterns and trade patterns.

4.2 Estimation of gravity model of trade

In this section we discuss estimation results obtained at the goods type level, compare them to the results reported by FMR and select the most suitable estimation approach. Then we report results for sector-wise estimations matching the decomposition.

4.2.1 Estimations at the aggregate goods type - level

The results reported by FMR (Table 1) were obtained by separate estimation for three goods types, for five cross-sections each. Several pooling tests in the methods subsection 3.2.4 showed that there is sufficient heterogeneity in our dataset to justify separate estimation for goods types, and also to allow for time-dependent intercepts. Moreover, the restriction of constant slope coefficients imposed by the time FE model (equ. 3.2.5) is rejected by an F-test evaluated for our goods-type aggregate data, which gives a p-value of 0.002 for homogenous goods and lower for differentiated and reference-priced goods (see Excel results file). While this is not surprising – given the size of our sample, most homogeneity restrictions would be rejected – it in principle supports the choice of a CS model over the time FE model. However, the qualitative results of the two approaches are not very different. The five columns on the left of Table 4, Table 5 and Table 6 give the results for differentiated, reference-priced and differentiated goods, respectively, for four selected cross-sections[36] and the time FE model.

A first inspection shows that the parameters of $log(Y_X)$, $log(Y_M)$ and $log(DIST)$ are significant with the expected signs and magnitudes in all cases (for distance, we get values around -0.76). The common language, contiguity and FTA dummies also have the expected positive impact, are mostly significant and of similar magnitude. For these six variables, estimates tend to increase when going from homogenous, to reference-priced, to differentiated goods. The same holds for the model fit (adjusted R^2), with values just below 0.4, 0.6 and 0.7, respectively. Note that the time FE model achieves a fit similar to the respective cross-section estimations.

The picture is less clear for the population and remoteness variables: While the parameter of $log(POP_X)$ has small negative values, and is mostly significant except for CS estimations for homogenous goods, for $log(POP_M)$ it changes sign and is frequently insignificant for CS estimations for reference-priced and homogenous goods. For $log(REM_X)$, we obtain relatively large positive estimates for homogenous and reference-priced goods, but significant (though small) negative values

[36] Results for all 15 cross-sections can be found in the accompanying Excel file.

for differentiated goods. The parameter of $log(REM_M)$ is mostly insignificant, except for differentiated goods, where it has small positive values.

The intercept estimates are mostly significant, positive for homogenous and negative for differentiated goods; this reflects the respective magnitude of trade for each goods type. Inspection of the year fixed effects in the time FE model (see Excel file) reveals that their size is mostly below 1% of the base intercept, so these are only minor adjustments.

Finally, for the Kyoto dummy, no clear pattern emerges: For the CS estimations after 2001, which is the first year with a ratification of a B country, almost half of the estimates are larger than zero (seven out of 18, for reference-priced and homogenous goods), and just as many are insignificant (two of the positive and five of the negative estimates, four of which are in the first two years). A test run using a CS model with the Kyoto-variable weighted by the carbon emissions intensity of the respective trade flow gave even fewer negative estimates, especially for differentiated goods, and no improvement regarding significance[37]. The time FE model gives a very small, but significant negative estimate for homogenous goods, but also a slightly larger, significantly positive estimate for reference-priced goods, and is insignificant for differentiated goods.

We now take a closer look at the parameter estimates for $log(Y_X)$ and $log(Y_M)$, β_X and β_M, visualized in Figure 6, to see how they compare to the findings of FMR. The results for the time FE model lie right among their respective CS counterparts, on which we thus focus in the first step.

For differentiated goods (Table 4), the parameter for $log(Y_X)$, β_X , is larger than β_M for $log(Y_M)$ in all periods for the CS model. β_X rises from 0.90 in 1992 to 1.06 in 2006, with some minor disruptions of the upward trend; the β_M value is initially around 0.1 lower (0.81 in 1992), but the gap widens to 0.3 as β_M falls to 0.75 in 2005 and is significant at the 5% level[38] throughout. This confirms the findings of FMR, although their absolute β_X and β_M values are slightly higher and lower, respectively.

For reference-priced goods (Table 5), β_X is in the interval [0.73; 0.81], with an upward trend from '93 to '06 after an initial decrease and minimum values in '99 and '00. β_M values are about 0.1 lower and in the interval [0.61; 0.69], with no dominant pattern over time. The gap between the elasticities widens from 0.05 in 1993 to 1.4 in 2006. The somewhat lower absolute values, compared to differentiated goods, and the smaller gap between the elasticities is in line with FMR.

[37] See accompanying Excel file for detailed results.

[38] In the following, "significant" always means significance at the 5% level.

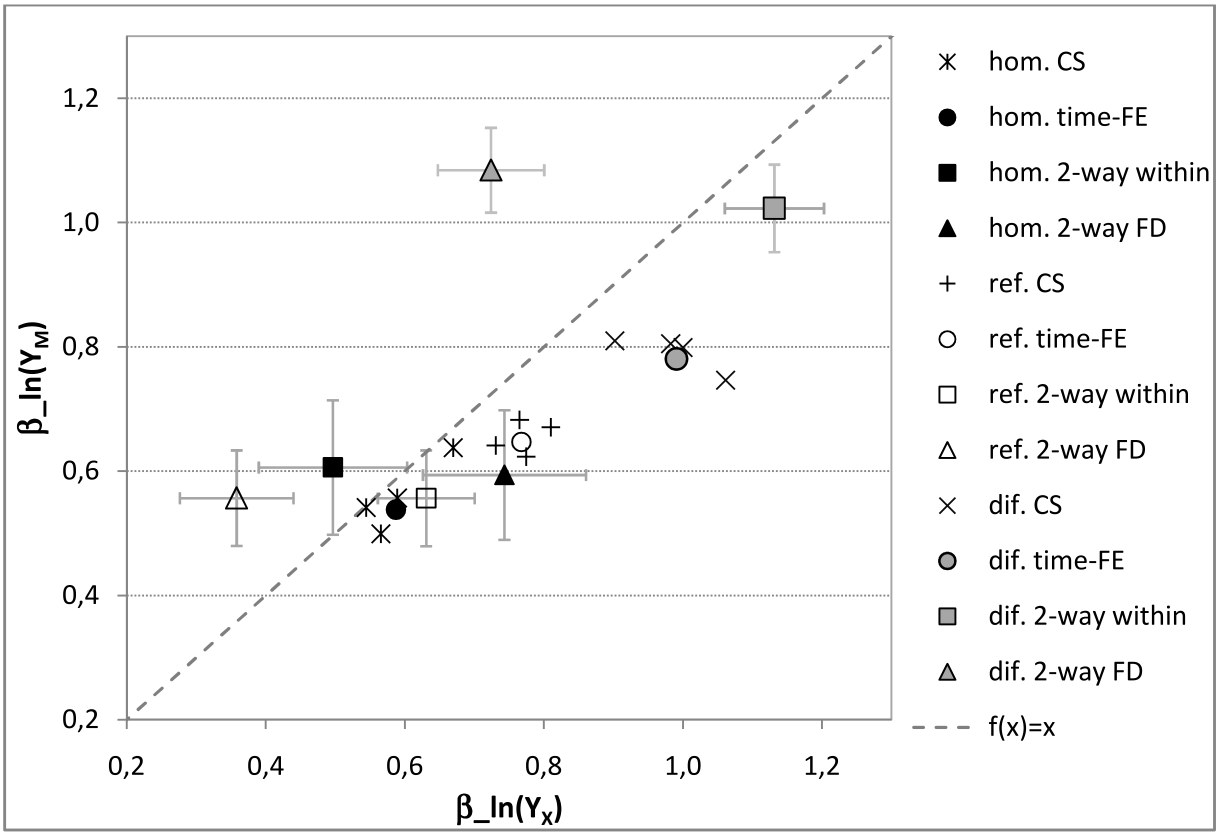

Figure 6: Elasticities of trade with respect to exporter and importer GDP per goods type, for selected estimations (based on Table 4 to Table 6; short error bars for CS and time FE omitted for readability).

For homogenous goods (Table 6), both elasticities are close together and show a similar, if not parallel, increase, with β_X always slightly larger than β_M ([0.52; 0.67] vs. [0.48; 0.64]). However, looking at the standard errors reveals that the gap is insignificant for 9 out of 15 periods. FMR found β_M to be significantly larger than β_X for homogenous goods, with values around 0.8 and 0.5, respectively.

Differentiated goods. Dependent variable: ln(X_{gijt}) (X: aggregate bilateral trade, regions pooled)							
	Cross-section (OLS)				**Pooled over 15 periods**		
	1992	**2000**	**2001**	**2006**	**time eff.**	**2-way FE (within est.)**	**2-way FE (FD est.)**
Intercept	-14,78 (1,54)	-15,98 (1,36)	-15,61 (1,36)	-21,07 (1,39)	-	-	-
ln(YX)	0,90 (0,02)	0,98 (0,02)	1,00 (0,02)	1,06 (0,02)	0,99 (0,01)	1,13 (0,07)	0,72 (0,08)
ln(YM)	0,81 (0,02)	0,80 (0,02)	0,80 (0,02)	0,75 (0,02)	0,78 (0,01)	1,02 (0,07)	1,08 (0,07)
ln(Dist)	-0,62 (0,03)	-0,81 (0,02)	-0,86 (0,02)	-0,80 (0,03)	-0,79 (0,02)	-	-
ln(PopX)	-0,16 (0,02)	-0,15 (0,02)	-0,15 (0,02)	-0,12 (0,02)	-0,14 (0,02)	-0,35 (0,16)	*0,22 (0,19)*
ln(PopM)	-0,11 (0,02)	-0,08 (0,02)	-0,06 (0,02)	*0,00 (0,02)*	-0,06 (0,01)	-1,33 (0,13)	-1,00 (0,16)
ln(RemX)	-0,14 (0,04)	-0,10 (0,04)	-0,08 (0,04)	-0,19 (0,04)	-0,11 (0,03)	*-0,10 (0,25)*	0,48 (0,24)
ln(RemM)	0,21 (0,05)	0,17 (0,04)	0,17 (0,04)	0,15 (0,04)	0,17 (0,04)	*-0,03 (0,28)*	-1,18 (0,26)
FTA	0,87 (0,09)	0,62 (0,08)	0,50 (0,09)	0,60 (0,06)	0,57 (0,06)	0,19 (0,03)	*-0,01 (0,02)*
CONT	0,83 (0,11)	0,81 (0,08)	0,78 (0,09)	1,11 (0,09)	0,87 (0,08)	-	-
LANG	0,65 (0,06)	0,70 (0,05)	0,67 (0,05)	0,65 (0,05)	0,70 (0,04)	-	-
KYOTO	-	-	*-0,06 (0,09)*	-0,08 (0,04)	*-0,02 (0,02)*	*0,04 (0,01)*	-0,04 (0,01)
parameters	11	11	12	12	26	8566	8566
N	5016	6873	6959	7748	99590	98787	98787
adj. R^2	0,63	0,69	0,69	0,69	0,68	0,18**	0,01**

Heteroskedasticity- (and with FE, autocorrelation-)robust standard errors in brackets, *insignificant values at the 5% level in italics.*
**R-squared for 2-way FE models not directly comparable to OLS and time FE models.

Table 4: Estimation results for aggregate differentiated goods

Reference-priced goods. Dependent variable: ln(X_{gijt}) (X: aggregate bilateral trade, regions pooled)							
	Cross-section (OLS)				**Pooled over 15 periods**		
	1992	**2000**	**2001**	**2006**	**time eff.**	**2-way FE (within est.)**	**2-way FE (FD est.)**
Intercept	3,25 (1,53)	4,01 (1,32)	5,52 (1,34)	*-1,62 (1,38)*	-	-	-
ln(YX)	0,77 (0,02)	0,73 (0,02)	0,77 (0,02)	0,81 (0,02)	0,77 (0,02)	0,63 (0,07)	0,36 (0,08)
ln(YM)	0,68 (0,02)	0,64 (0,02)	0,62 (0,02)	0,67 (0,02)	0,65 (0,01)	0,56 (0,08)	0,56 (0,08)
ln(Dist)	-0,66 (0,03)	-0,77 (0,02)	-0,79 (0,02)	-0,88 (0,03)	-0,78 (0,02)	-	-
ln(PopX)	-0,26 (0,02)	-0,17 (0,02)	-0,21 (0,02)	-0,12 (0,02)	-0,18 (0,02)	-0,47 (0,16)	0,51 (0,20)
ln(PopM)	*-0,03 (0,02)*	*0,02 (0,02)*	0,04 (0,02)	0,05 (0,02)	0,02 (0,01)	*0,26 (0,14)*	*-0,07 (0,18)*
ln(RemX)	0,54 (0,05)	0,51 (0,04)	0,57 (0,04)	0,55 (0,04)	0,55 (0,03)	*0,03 (0,29)*	1,23 (0,30)
ln(RemM)	*0,03 (0,05)*	*0,07 (0,04)*	0,08 (0,04)	*-0,08 (0,04)*	*0,01 (0,03)*	*-0,26 (0,29)*	-1,16 (0,29)
FTA	0,84 (0,09)	0,71 (0,08)	0,65 (0,08)	0,18 (0,07)	0,48 (0,06)	0,16 (0,03)	0,08 (0,02)
CONT	0,49 (0,11)	0,76 (0,09)	0,86 (0,09)	1,03 (0,09)	0,82 (0,08)	-	-
LANG	0,55 (0,06)	0,66 (0,05)	0,64 (0,05)	0,72 (0,05)	0,65 (0,04)	-	-
KYOTO	-	-	*-0,04 (0,10)*	0,16 (0,04)	0,11 (0,02)	*0,02 (0,01)*	*-0,01 (0,01)*
parameters	11	11	12	12	26	8229	8229
N	4614	6278	6301	7179	91250	90411	90411
adj. R^2	0,54	0,58	0,58	0,60	0,58	0,08**	0,01**

Heteroskedasticity- (and with FE, autocorrelation-)robust standard errors in brackets, *insignificant values at the 5% level in italics.*
**R-squared for 2-way FE models not directly comparable to OLS and time FE models.

Table 5: Estimation results for aggregate reference-priced goods

Homogenous goods. Dependent variable: ln(X_{gijt}) (X: aggregate bilateral trade, regions pooled)							
	Cross-section (OLS)				**Pooled over 15 periods**		
	1992	**2000**	**2001**	**2006**	**time eff. (OLS)**	**2-way FE (within est.)**	**2-way FE (FD est.)**
Intercept	11,48 (2,05)	13,33 (1,89)	14,75 (1,91)	13,31 (2,02)	-	-	-
ln(Y_X)	0,57 (0,03)	0,54 (0,03)	0,59 (0,03)	0,67 (0,03)	0,59 (0,02)	0,50 (0,11)	0,74 (0,12)
ln(Y_M)	0,50 (0,03)	0,54 (0,02)	0,56 (0,02)	0,64 (0,03)	0,54 (0,02)	0,61 (0,11)	0,59 (0,10)
ln(DIST)	-0,70 (0,04)	-0,71 (0,04)	-0,74 (0,04)	-0,85 (0,04)	-0,70 (0,03)	-	-
ln(POP_X)	-0,07 (0,03)	-0,02 (0,03)	-0,06 (0,03)	-0,03 (0,03)	*-0,04 (0,02)*	*-0,36 (0,25)*	*0,15 (0,31)*
ln(POP_M)	0,07 (0,03)	0,05 (0,03)	0,07 (0,03)	0,09 (0,03)	0,08 (0,02)	0,57 (0,20)	0,94 (0,24)
ln(REM_X)	0,79 (0,06)	0,74 (0,06)	0,89 (0,05)	0,97 (0,06)	0,83 (0,04)	*-0,14 (0,41)*	1,28 (0,42)
ln(REM_M)	*-0,07 (0,07)*	*0,11 (0,06)*	*0,06 (0,06)*	*0,08 (0,07)*	*0,04 (0,05)*	-1,23 (0,41)	-1,39 (0,40)
FTA	0,68 (0,13)	0,46 (0,11)	0,38 (0,11)	*0,10 (0,09)*	0,33 (0,08)	0,22 (0,05)	*0,04 (0,04)*
CONT	*0,23 (0,14)*	0,56 (0,11)	0,43 (0,12)	0,57 (0,11)	0,45 (0,09)	-	-
LANG	0,54 (0,08)	0,52 (0,07)	0,57 (0,07)	0,69 (0,07)	0,64 (0,06)	-	-
KYOTO	-	-	*-0,24 (0,16)*	*0,02 (0,06)*	-0,06 (0,03)	-0,11 (0,02)	-0,07 (0,02)
parameters	11	11	12	12	26	6108	6108
N	3231	4181	4167	4856	61421	60373	60373
adj. R^2	0,35	0,36	0,38	0,42	0,37	0,05**	0,01**
Heteroskedasticity- (and with FE, autocorrelation-)robust standard errors in brackets, *insignificant values at the 5% level in italics.* **R-squared for 2-way FE models not directly comparable to OLS and time FE models.							

Table 6: Estimation results for aggregate homogenous goods

FMR repeat their estimations on subsamples of country-pairs (within OECD and between OPEC and non-OPEC countries) to check the robustness of their argument that the different impact of exporters' and importers' GDP is attributable to goods types rather than "country types". We compare CS estimations for four different subsamples: trade among B countries (B2B), among non-B countries (nB2nB), from B countries to non-B-countries (B2nB) and vice versa (B2nB). The results for the GDP elasticities are plotted in Figure 9 in the appendix, the full results can be found in the Excel results file. We find that regional subsample CS estimations for nB2B and nB2nB are similar to those we obtained by pooled CS above for all goods types and in all years. For B2B homogenous goods trade and B2nB homogenous and reference-priced goods, we have a systematic deviation and obtain $\beta_M > \beta_X$. The fact that we obtain homogenous goods estimates consistent with FMR for B2B and B2nB trade, but not for our full sample suggests that those trade flows might have higher relative representation in FMR's sample than in ours.

Overall, our time FE and CS estimations confirm FMR in the sense that we find a trend of decreasing impact of Y_X and increasing impact of Y_M as we go from differentiated to reference-priced, to homogenous goods, even though this trend is less pronounced in our estimations than in those of FMR. For differentiated and reference-priced goods, where β_X is significantly larger than β_M, this suggests a model of monopolistic competition is most appropriate, as argued by FMR. However, in contrast to FMR, our estimates of β_X remain larger than those of β_M for homogenous goods as well, so we do not find support for *restricted entry* in a reciprocal dumping model, as FMR do. Since the difference between β_X and β_M for homogenous goods is insignificantly small in a majority of the periods, no clear case can be made for a reciprocal dumping model with *free entry* either. The latter point is confirmed by subsample estimations using only trade between or within B and non-B country groups.

The models so far may be biased due to omitted variables, which can be captured by fixed effects (FE) in the case of constant variables. An F-test confirms the joint significance of individual effects[39] for all goods types (see Table 12). The introduction of such country-pair FE changes the picture substantially.

Focus on the β_X and β_M results first, for within estimation: Trade in differentiated goods would be more susceptible to changes in both exporters' and importers' GDP than in the time FE model, with β_X estimated as 1.13 instead of 0.99, and β_M at 1.02 from 0.78. Although β_X remains larger, the gap to β_M is now insignificant, since the robust standard errors are also larger (0.07).[40] The same holds for for reference-priced goods, where both elasticities are lower than in the time FE model (β_X changes from 0.77 to 0.63, and β_M from 0.65 to 0.56). Finally, for homogenous goods, β_X (from 0.59 to 0.50) is now *smaller* than β_M (from 0.54 to 0.61), as in FMR – but again, the gap is too small to be significant. Overall, within estimation gives qualitative results similar to FMR, but the differences in elasticities are now too small and the errors too large to decide between the theories behind the gravity model.

Second, using the FD estimator yields qualitatively very different, almost reversed results: β_X (0.72) is now *significantly smaller* than β_M (1.08) for *differentiated* goods, while for homogenous goods, the relatively higher importance of changes in Y_X compared to those of Y_M is even more pronounced than under time FE: β_X for homogenous goods (0.74) is even slightly larger than the first-difference estimate for differentiated goods. For reference-priced goods, β_X (0.36) is now quite far below β_M (0.56), so homogenous and reference-priced goods under FD estimation basically switch positions compared to within estimation[41]. The qualitative pattern is thus reversed between differentiated and homogenous goods and using the theoretical arguments in FMR, different conclusions concerning the model at work would have to be drawn (at least for those FD estimates where the difference between β_X and β_M is significant).

---

[39] The F-test compares a two-way FE model under within estimation to a time FE model under pooled OLS.

[40] It will turn out below that we have to use heteroskedasticity- and autocorrelation-consistent standard errors, which we thus report for our fixed effects and time FE estimations throughout. Even with non-robust errors, which are about half as large, the gap between β_X and β_M is insignificant in most cases.

[41] Using the liberal classification emphasizes the patterns for differentiated and reference-priced goods even more; for homogenous goods, β_X is already higher than in the conservative case for time-FE, and increases even more, also relative to β_M, when we use fixed effects under either estimator (see Excel results file).

Regarding the other variables, the impact of FTAs remains significantly positive under within estimation, but it is smaller than CS and time-FE estimates. It practically vanishes under FD estimation, while the impact of populations and remoteness is as heterogeneous as before, with changing signs, size and significance.

The within estimate of the Kyoto variable is significantly negative (but small) for homogenous goods, but significantly positive for differentiated goods and insignificant for reference-priced goods. FD estimates are similar for homogenous and reference-priced goods, but we get a small negative value for differentiated goods.

To determine if the within or the FD estimates are more reliable, we first evaluate the first-difference based test for serial correlation described in the methods section (3.2.6.). Both the null hypothesis of no serial correlation in the original errors and of no serial correlation in the differenced errors are rejected with very high Chi-squared values, ranging from 1284 to 1581 in the former and 2572 to 3332 in the latter case (see Table 13), and corresponding p-values below the level of numerical accuracy provided by R. For this case, Croissant and Millo (2005, p.28 and p.31) suggest using heteroskedasticity- and autocorrelation-robust errors calculated by the Arellano method, which we already reported above. Unfortunately, this does not help us to choose between the two estimators. On the contrary it suggests that a dynamic model might be more suitable than ours.

A second criterion would be the goodness of fit. This is not routinely evaluated for panel models; several possible definitions for R^2 exist, e.g. treating individual FE as “explaining” or just “capturing” some part of the variation (Verbeek 2004, p.352). This is also the reason why the R^2 values given in Table 4 to Table 6 for the models with individual FE are not directly comparable e.g. to those of the time FE model, and other criteria and tests become more important. But R^2 values, adjusted for degrees of freedom, are defined consistently between the within and FD estimation, and are much larger for within estimation in all cases. This is of particular importance as we want to reproduce the decomposition results with the predicted results in the final step.

Overall, we prefer two-way fixed effects models to CS on theoretical grounds (unobserved heterogeneity) and to time FE models also due to the results of the Hausman test for significant individual effects. Among the panel methods we have a tendency towards “within” rather than “first-difference” estimation based on the serial correlation tests and the fit measures, although the source of the difference between the two remains unknown.

Using the within estimator, we perform one more robustness check. As for CS estimations above, we repeat the estimation for subsamples of intra- and interregional trade flows for the B and non-B country groups. Figure 7 displays the key results, the GDP elasticities of aggregate trade in differentiated and homogenous goods. The full results are contained in the Excel file. The results of the estimation pooled over all regions described above are included for comparison. With the exception of goods exports from B to non-B countries, all estimates for differentiated goods trade lie relatively close together with $\beta_X > \beta_M$. For homogenous goods trade, on the other hand, we obtain widely varying results far from the pooled estimate, with $\beta_X < 0$ for B to B and B to non-B trade and $\beta_M < 0$ for non-B to B trade. This provides more evidence that our models work better for differentiated goods trade (and between similar countries) than for homogenous goods trade.

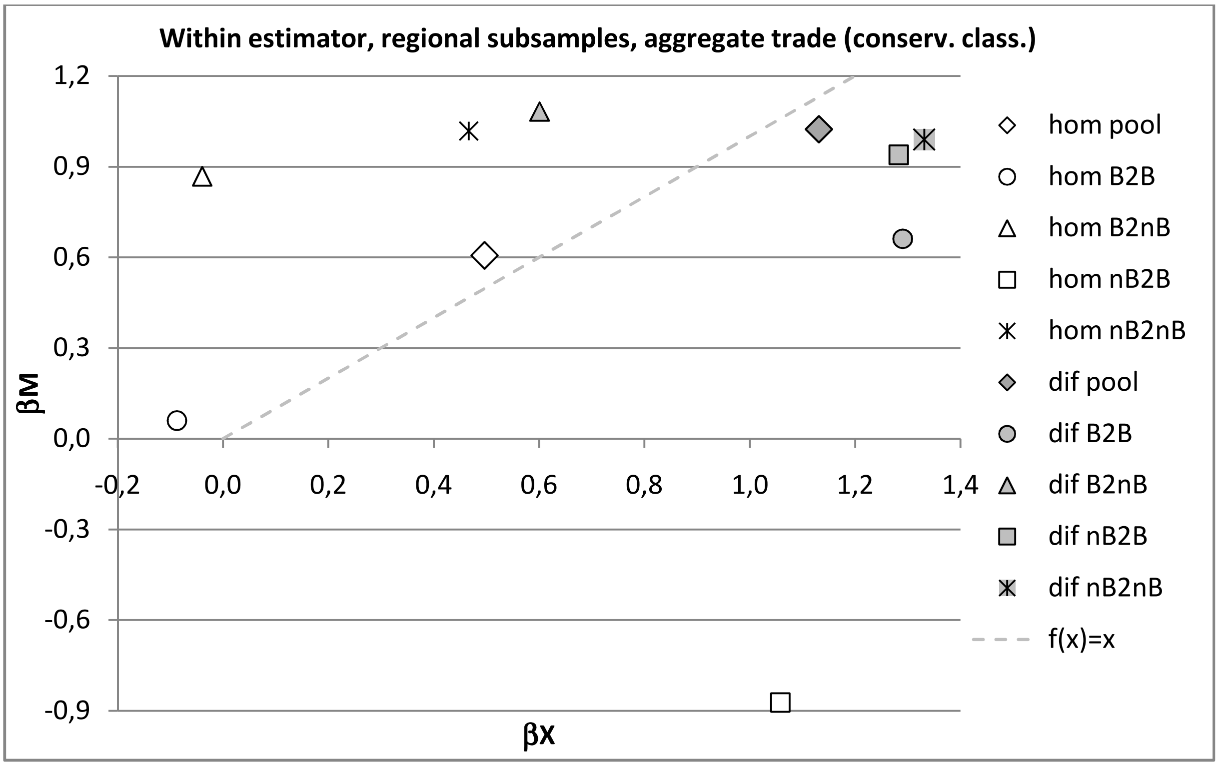

Figure 7: GDP elasticities of goods trade for regional subsamples and within estimation

4.2.2 Estimations at the sector level

We now turn to our estimations at the detailed sector level. As for aggregate goods, tests in subsection 3.2.4 justified separate estimation for each sector and time-dependent intercepts (sector-wise time FE model). Due to data restrictions, a CS model is not feasible at the sector level. We maintain the restriction of constant slope coefficients throughout and only compare the time FE model to two-way FE models under within and FD estimation. The joint significance of country-pair fixed effects is confirmed across all sectors with an F-test, while significance of time FE is rejected for a few sectors (mainly of the homogenous goods type; see Table 12). We nevertheless use both effects in all estimations for ease of comparison. Table 14, Table 15 and Table 16 in the appendix give selected results for differentiated, reference-priced and differentiated goods, respectively[42]. We first

[42] Full results, for all variables and estimators, are included in the accompanying Excel results file.

discuss the parameters of $ln(Y_X)$ and $ln(Y_M)$, β_X and β_M, which are plotted for each sector in Figure 8, grouped by goods types.

For the 14 sectors classified as differentiated goods, within estimation gives β_X values that are mostly higher than for the aggregate case (up to 2.3), and β_M values that are often similar to the estimate for aggregate trade in differentiated goods, although there are some up- and downward deviations. β_X is significantly larger than β_M, with two noteworthy exceptions: Coal ("col") and forestry ("frs") have larger elasticities regarding the importer's GDP, with the latter even showing a negative (but insignificant) estimate for β_X. These two sectors are actually "reference-priced" under the liberal classification. They have low R^2 values, and all of our other explanatory variables are found to be insignificant (apart from $ln(POP_X)$ for forestry, which has a large negative impact). Additionally, the sample for these sectors are smaller than for other differentiated goods, and Table 9 shows that exports and imports are relatively concentrated, with a few countries dominating trade. Both sectors would be candidates for the use of additional explanatory variables such as surface area or natural resource endowment in a time FE or CS[43] model. If these variables are slow-moving, their effect would be captured by fixed effects. It is plausible nevertheless that the impact of exporter's GDP is smaller in these cases and that the gravity model does not work well for them, also because the exporter's income may be endogenously affected by trade.

Contrary to within estimation, sector-wise estimation with time effects yields lower values for both elasticities than for aggregate data of trade in differentiated goods. The decrease is particularly high for β_X, which falls as low as 0.18 and significantly below β_M in some cases, e.g. wearing apparel (wap), leather (lea) and textiles (tex).

First-difference estimation gives β_M values between 0.7 and 1.2, which is similar to within estimation, and β_X between 0.5 and 1, which is much lower (and significantly below β_M). There are three "outliers" – coal and forestry, as before, and transport equipment (otn), with $\beta_X < \beta_M$ at a low level.

The sector-wise estimates for 13 reference-priced goods, while being somewhat smaller in total, have roughly the same position relative to the estimates for the aggregate data that we described for the differentiated goods: Mostly higher β_X for the within estimator, although β_M here also increases and is now sometimes larger than β_X ; FD estimates with more variation in estimated β_M than in β_X, but now also grouped around the diagonal where $\beta_M = \beta_X$; and both elasticities lower (and roughly balanced) than in the aggregate case for time effects. The time, FD and within estimations, with average elasticities rising in this order, also maintain their relative positions as for differentiated

[43] Feenstra et al. (2001) use the share of minerals and fuel production in the exporter's and importer's GDP as additional variables for a robustness check. They find that they are significant, but have little impact on GDP elasticities.

goods. Note that a number of estimates are negative and/or insignificant: four out of 13 sectors are affected for within and six for FD estimation, namely fishing (fsh), meat products (omt), dairy products (mil) and beverages and tobacco (b_t) for both estimators plus oil seeds (osd) and ferrous metals (i_s) under FD estimation.

The estimates for homogenous goods show large variations between the 14 sectors. As before, the model with time FE gives the lowest absolute values, lower than for the estimation with aggregate trade. These estimates are relatively close together across sectors, with no discernable pattern regarding the relative sizes of β_X and β_M. Four sectors have one insignificant elasticity estimate: gas, sugar (sgr), sugar cane and beet (c_b) and rice (pdr). The FD estimator yields higher absolute values in a large interval, with β_X from -0.57 to 1.4 and β_M between -1.05 and 1.24. Robust standard errors are large and only four sectors have two significant elasticity estimates: Grains (gro), plant-based fibers (pfb) and metals (nfm) have no significant gap between β_X and β_M though, and cement (cmt) even has a negative β_X. Within estimation gives even more extreme and equally incoherent estimates ranging from -0.93 to 1.25 (β_X) and -4.42 to 1.28 (β_M). Sectors with two significant elasticities are again grains, fibers and metals, and additionally oil.

The Kyoto dummy has small parameter estimates which are often insignificant, even for differentiated goods sectors under within or time FE estimation where β_X and β_M are significant, and more frequently for reference-priced and homogenous goods. Positive estimates occur and make up one third to one half of the significant values. We thus find no consistent effect of unilateral Kyoto membership on trade flows at the individual sector level.

The parameter estimates for the other variables are reported in the accompanying Excel file: The common language and contiguity dummies and $ln(DIST)$ are only estimated in the time FE model. They have the expected signs but are only roughly half as large as for the aggregate case (e.g. for distance, the simple average of all coefficients is -0.36), with numerous insignificant values in particular for homogenous goods sectors and for CONT and LANG. The coefficient of the FTA is also smaller and insignificant for most FD and many within estimations. For the other auxiliary variables ($ln(POP_X)$, $ln(POP_M)$, $log(REM_X)$, $log(REM_M)$), from time FE to within to FD estimation, we have decreasing numbers of significant estimates with a larger magnitude and switching signs more often, again especially for homogenous goods.

As a general trend we obtain fewer significant estimates of β_X and β_M and the other parameters from differentiated, to reference-priced, to homogenous goods sectors. Simultaneously, the fit (R^2) of the time FE and within estimators decreases, while it is uniformly low for FD estimation even compared to within estimation. These patterns are similar to the aggregate case[44], but at the sector level aggravated by smaller samples dominated by smaller numbers of country-pairs (Table 9) for homogenous and to some extent reference-priced sectors.

As for the goods-type aggregate trade estimations the results of the within estimator can be assumed to be more reliable than time FE due to elimination of unobserved heterogeneity, and more trustworthy than FD estimation due to the better model fit. The null hypothesis of no serial correlation in both the original and differenced errors is rejected again, for all sectors (Table 13), so we continue to use autocorrelation-robust errors.

[44] Although the fit for the sector-wise time FE model seem to be lower, on average, than in the aggregate goods type case, this is due to the definition of adjusted R^2 in terms of deviations from the respective (sub-)sample's average. The overall fit of 41 separate estimations is of course better. In models with individual FE, the number of parameters is much higher and we achieve a similar R^2 for the subsamples and the total sample.

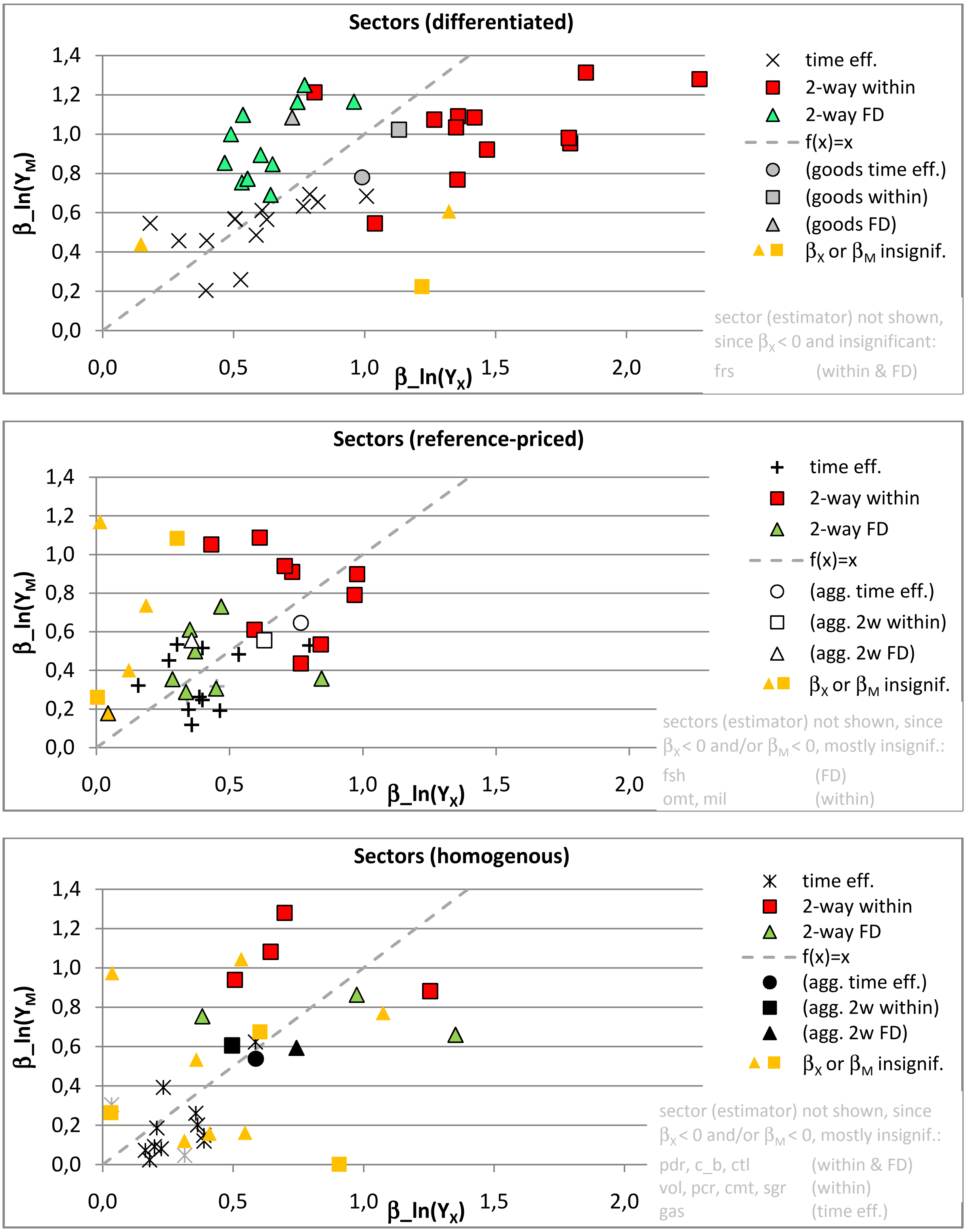

Figure 8: Elasticities of trade with respect to exporter and importer GDP per sector, for selected estimations (based on Table 14 to Table **16**; different symbol for time eff. than in Figure 6 and negative values omitted for readability).

4.2.3 Decomposition of predicted trade

To further illustrate the results of the sector-wise estimation, we use the predicted trade volumes from one of the sector-level models to predict EET and repeat our initial decomposition. This is not an additional statistical test, but will show to which extent the results of the decomposition can be reproduced by the gravity model.

Specifically, we choose the two-way FE model under within estimation[45] to predict bilateral trade in dollar terms. In our basic decomposition identity for EET_{sij} (equ. 3.1.1), this gives the bilateral scale and composition factors X_{ij} and C_{sij}. Intensity as the third factor I_{sij} is equal to the originally observed carbon intensities for each sector and country-pair, which we made no attempt to explain and must thus treat as exogenous. We can now calculate the predicted EET and perform the decomposition defined by equations 3.1.3 and 3.1.4. Note that although our prediction of trade volumes affects only scale and composition but not intensity for any single period, they weight the respective intensities in a decomposition of changes in EET over time and thus also have an impact on the intensity effect. The results are given in Table 7, with selected results from our first decomposition in brackets.

In terms of the relative size of the flows and effects, the general patterns that we found in Table 3 are reproduced well: EET for total trade and between B and non-B countries grows strongly at 41% and 52% (vs. 74% in the original data), driven by the scale effects particularly for trade originating in non-B countries, which make up around 60% of total trade scale effects in the predicted and the observed case. Scale effects are again much larger than intensity improvements, and relatively higher intensity improvements for B exports cannot make up for this. EET from non-B to B are already 60% larger than B to non-B embodied emissions in 1992 (original data: 70%), and they grow four times (three times) faster, so the net EET deficit of B countries increases by 170% (186%). The composition effects predicted by within estimation are even smaller than in the original data.

However, these qualitative results cannot conceal that the fit is poor in absolute terms. The "within model" systematically underestimates the volumes of trade and thus of EET. The mismatch grows over time, e.g. B to non-B trade (EET) is underestimated by 7% (3%) in 1992, but 21% (14%) in 2006. Predictions for intra-non-B trade (EET) are particularly low, for example by -17% (-21%) in 1992 and -38% (-41%) in 2006. This is reflected by too low predictions for regional and total ΔEET, small scale effects and an underestimated net EET balance.

One possible explanation is that the within estimator eliminates fixed effects by a "demeaning" transformation and only uses within-individual variation of the variables for estimation, thus

[45] Within estimation cannot take into account country-pairs with data for only period. Since these cases make up only 1% of total trade and EET over all sectors and periods we assume that their omission does not strongly affect the results.

discarding their total values. The initial offset must have been positive in our case, which would have lead to higher absolute values. This leaves us with the insufficient prediction of changes over time. Note that the relative underestimation of trade volume is small for trade among B countries and changes the least, with -7% in 1992 and -9% in 2006. This block makes up 55% of total trade volume in 1992 and 46% in 2006 and also has the largest number of data points. This might have led to a large impact on parameter estimates and a good fit for the trade dynamics of this block, and worse fits for the other regional trade flows. The latter also have different composition in terms of goods types (see Table 2) with a larger share of homogenous goods in particular for non-B exports. We saw that the estimates for these goods are relatively poor. This is consistent with the argument that only trade between relatively similar countries and in differentiated goods is described well by a gravity model.

Predicted totals in first and last period		**B to nB**	**nB to B**	**B to B**	**nB to nB**	**Total EET**	**Total B<>nB**	**Net B to nB**
EET '92	abs. [Mt CO2]	570	905	1571	375	3420	1474	-335
	(orig. data)	(587)	(1011)	(1584)	(475)	(3656)	(1598)	(-424)
EET '06	abs. [Mt CO2]	668	**1572**	1541	**1047**	4829	2241	-904
	(orig. data)	(781)	**(1996)**	(1778)	**(1794)**	(6349)	(2777)	(-1215)
Trade '92	[10^9 US$, 2005]	641	849	2268	335	4094	1491	-208
	(orig. data)	(687)	(883)	(2437)	(403)	(4410)	(1570)	(-196)
Trade '06	[10^9 US$, 2005]	**1444**	**1941**	4799	**1211**	9395	3385	-497
	(orig. data)	**(1832)**	**(2322)**	(5269)	**(1955)**	(11379)	(4154)	(-490)
CO_2 intensity '92	[kg CO2 / $]	0,89	1,07	0,69	1,12	0,84	0,99	-
	(orig. data)	(0,85)	(1,14)	(0,65)	(1,18)	(0,83)	(1,02)	-
CO_2 intensity '06	[kg CO2 / $]	0,46	0,81	0,32	0,86	0,51	0,66	-
	(orig. data)	(0,43)	(0,86)	(0,34)	(0,92)	(0,56)	(0,67)	-
Decomposition summary								
ΔEET	**abs. [Mt CO2]**	**99**	**668**	**-29**	**672**	**1409**	**766**	**-569**
	(orig. data)	(194)	(985)	(194)	(1320)	(2693)	(1179)	(-791)
	vs. '92 [%]	17%	74%	17%	179%	41%	52%	170%
thereof, due to...								
- scale change	**abs. [Mt CO2]**	**489**	**1334**	**1070**	**974**	**3867**	**1823**	**-845**
	(orig. decomp.)	(592)	(1607)	(1291)	(1540)	(5030)	(2199)	(-1015)
	vs. '92 [%]	86%	147%	68%	260%	113%	124%	252%
- composition change	**abs. [Mt CO2]**	**-16**	**6**	**43**	**-27**	**6**	**-10**	**-23**
	(orig. decomp.)	(61)	(89)	(198)	(66)	(413)	(150)	(-28)
	vs. scale eff.	-3%	0%	4%	-3%	0%	-1%	3%
	vs. '92 [%]	-3%	1%	3%	-7%	0%	-1%	7%
- intensity change	**abs. [Mt CO2]**	**-374**	**-673**	**-1143**	**-275**	**-2465**	**-1046**	**299**
	(orig. decomp.)	(-459)	(-710)	(-1295)	(-286)	(-2750)	(-1169)	(252)
	vs. scale eff.	-76%	-50%	-107%	-28%	-64%	-57%	-35%
	vs. '92 [%]	-66%	-74%	-73%	-73%	-72%	-71%	-89%

Cumulated yearwise decomposition, refined Laspeyres index method (interaction terms distributed symmetrically).
B = Annex B countries (incl. USA), nB = non-Annex B countries, EET=emissions embodied in trade.

Table 7: Decomposition of changes in CO_2 emissions embodied in goods trade, based on trade predicted by within estimation of two-way FE model and exogenous observed intensities.

5 Discussion

Our initial observation was that total EET and EET between B and non-B countries grew strongly between 1990 and 2008, and that growth of trade volumes even exceeded this. Decomposing the data analyzed by Peters et al. (2011) with a refined Laspeyres index (Ang and Zhang 2000) year-wise for the period from 1992 to 2006, we found that growth of specific bilateral trade relations and smaller intensity decreases are responsible for the observed changes in EET. Changes in bilateral trade flows' sector composition have a relatively small effect. Scale effects of exports from non-B countries were particularly large, while intensity improvements offset a major part of the already smaller scale effects for B countries' exports. Hence the net EET deficit of B countries almost tripled, with a 1:4 ratio of the contributions of the net intensity and net scale effects. Manufacturing, energy-intensive manufacturing and mining are the most important aggregate sectors behind these results.

We did not check the effect of using other than basic year weights (Hoekstra et al. 2003), but believe that the effect on our results would be small, because the relative size difference of the effects is large and based on the sum of year-wise decompositions rather than just one decomposition using the first and last year.

A modification of interest would be to introduce a fourth factor into the identity for EET, splitting the scale factor Xij into $Xij/Xtot$ and $Xtot$, where $Xtot$ denotes the total world trade volume. This would allow us to separate changes in country-pairs' shares of world trade from changes of trade volume, which is not possible in our specification. We suspect that intensity changes would then have the largest "single" effect on many trade flows.

We now summarize and interpret our estimation results, compare them to the literature and discuss some econometric issues.

We first reported results for the aggregate goods type level, starting with cross-sectional estimations for comparison to the results by Feenstra et al. (2001). We presented selected results for alternative specifications using time FE or additional country-pair FE, estimated by within and first-difference estimators. The cross-section and time FE estimations yield roughly similar results, with the elasticity of trade in *differentiated goods* with respect to the exporter's GDP (β_X) *larger* than that for the importer's GDP (β_M). Although in our estimations the difference is less pronounced, the models without FE would thus confirm FMR's conclusion that a home market effect is at work and a monopolistic competition is the relevant model in this case. These simple estimations also confirm the trend of decreasing (increasing) relative importance of the exporter's (importer's) GDP as we go from differentiated, to reference-priced, to homogenous goods, consistent with FMR. However, in our case, the exporter's GDP is more important even for homogenous goods, so there is weak evidence for a reciprocal dumping model with free rather than restricted entry, which FMR favored ("reverse home-market effect"). Moreover, a robust check with different subsamples of inter- and intra-regional trade flows raised doubts about the applicability of the same gravity models for all kinds of trade flows, in terms of goods types and countries.

The estimate of β_M for homogenous goods only becomes larger than β_X when we use FE for country-pairs and within estimation, and we still have $\beta_X > \beta_M$ for differentiated goods. But in both cases heteroskedasticity- and autocorrelation-robust errors are too large for the gap to be significant, so we cannot identify a home-market effect and decide between theories anymore. First-differencing leads to markedly different results, with a significant difference between elasticities at least for differentiated goods, but $\beta_X < \beta_M$, contrary to the findings by FMR and suggesting national product differentiation as the model of choice for this goods type.

We argued that cross-sectional estimations as used by FMR are potentially unreliable due to unobserved heterogeneity, and that panel estimators are preferable. This is in line with the findings summarized in the review by Bergstrand and Egger (2011, p.28). They additionally argue that instead of country-pair fixed effects, country- year effects should be used to fully capture general equilibrium effects, which we avoided here since would have precluded comparisons to FMR. Cheng and Wall (2005) similarly disapprove of cross-section and time FE estimations. They specifically test and reject the restrictions imposed by the FD estimator in favor of the within estimator, although their parameter estimates are similar for both estimators. Although we did not perform such a test, we found a poor fit of the FD- compared to the within estimator, and thus prefer the latter. We could not reject serial correlation in the original or differenced errors. Although this does not help with the choice of an estimator, it suggests that a dynamic model with lagged variables might be necessary.

Second, we applied the time FE model and both two-way effects models to sector-wise data. With the within estimator, for practically all sectors classified as differentiated goods, we found a relatively good fit and $\beta_X > \beta_M$ as in the aggregate case. But now estimates for β_X are mostly substantially larger, making the gap significant and supporting monopolistic competition as the source of the gravity mode. For reference-priced sectors, within estimates are clustered at higher values than for the aggregate case, but with varying relative sizes and thus no conclusive support for either theory. This is also true for homogenous goods, where estimates are widely scattered, frequently negative or insignificant, and the model fit is low. The deviations of the time FE and FD estimations from the within estimations in terms of parameter values, significance and fit are similar to the aggregate goods type case.

Overall, regarding our second research question we find that results are not robust to the choice of estimation approach for all goods types. For differentiated goods, the aggregate CS and time FE models and the preferable within estimation at the aggregate and sector level gave elasticity of trade that were larger with respect to the exporter's GDP than for the importer's GDP (although the difference is insignificant for aggregate within estimation due to large errors). This provided some support of FMR's claim that this is governed by a home-market effect in a monopolistic competition model. First-difference estimation, which we find less convincing due to poor fit, lead to the opposite relative sizes of the elasticities as the Armington assumption would predict. For other goods types, however, we did not find evidence in support of FMR, as our estimated GDP elasticities are inconclusive at the aggregate and sector level. In the case of homogenous goods, where results are often insignificant and negative, we would even argue that the gravity model as we specified it might not be appropriate at all. This is underlined by a robustness check using within estimation for subsamples of intra- and interregional trade, which gave very different results for homogenous goods depending on the regions sample. We provided additional illustration by using the sector-wise trade data and exogenous emission intensities to "predict" EET and repeating our initial decomposition on this data. This showed that, apart from a general underestimation of trade flows which is probably due to the short time frame of our panel, predictions for non-B exports are particularly too low. Either trade originating in non-B countries is underrepresented in our data, since for example B to B trade is predicted well, or again homogenous goods trade, which makes up a larger share for non-B exports, is not modeled well.

This contrasts FMR's finding that "the differing estimates of the gravity equation pertain to types of *goods*, rather than being features of *countries* with differing factor endowments" (Feenstra et al. 2001, p.444). Without explicitly separating goods types for estimation, Debaere (2005) rejects a monopolistic competition model for non-OECD trade flows as reported by Hummels and Levinsohn (1995). He argues that this could be due to their lower share of differentiated goods, consistent with our findings. Cheng and Wall (2005) restrict their sample to trade between upper-middle or high income countries right from the start and exclude high-income oil exporters.

Before discussing additional findings regarding the impact of Kyoto membership on trade, we comment on some issues that might affect the robustness of our results. These issues are inter alia exposed by the large variation in parameter results at the sector level, their deviation from the aggregate goods case and their frequent insignificance.

The first concern is the aggregation of sectors into goods types. While we only reported results for aggregate goods types based on the "conservative" commodity classification scheme by Rauch (1999), where in ambiguous cases commodities are included in the relatively less homogenous class, he also provides a "liberal" classification which maximizes the number of homogenous and reference-priced goods. This shifts three sectors from the class of differentiated to reference-priced goods and one to homogenous goods, and six sectors from reference-priced to homogenous. At the sector level, this just affects the grouping of our results for presentation, and the results are also independent of potential errors that we made when mapping Rauch's original SITC classification on GTAP sectors. The aggregate estimations yield slightly different results (see Excel results file), but our qualitative conclusions are unaffected.

The second issue is the handling of instances of zero trade, which we simply omitted, similar for example to Cheng and Wall (2005) and FMR. FMR additionally report that their results are not changed by an alternative Tobit estimation, while Rauch (1999) compares both cases and prefers the Tobit approach. As discussed in subsection 3.2.2, omission of zero trade implies that we are only estimating the effect of our explanatory variables on the volume of trade in a sector between country-pairs that are assumed to have already traded in this sector before ("intensive margin"). The question when country-pairs trade at all, or in which sectors (we called this the "extensive country and sector margins") is neglected. Trade data aggregated into three goods types rather than 41 sectors has a lower share of zeros because it captures more of the effects at the extensive margin, which in turn affects the estimates, although it is not clear into which direction.

Third, data quality might have impaired our results. Our trade dataset is relatively sparse for some sectors and country pairs (Table 9). Lower performance of the model for homogenous goods and sectors may be linked to these smaller samples and the larger contribution of a few exporters and importers to our samples. We cannot rule out that this is due to some selection bias. More concerns relate to our construction of the *trade data*: We use countries' total export volumes reported by the UNSD (2011) whose global sum differs from that of the imports, so there must be some measurement or reporting error. There is also uncertainty regarding the conversion of data in current local currencies into constant 2005 US$ (which we use), undertaken by the UNSD. Then, since we are only analyzing commodity trade flows, we extrapolate the goods-service-share from three years of GTAP data (1997, 2001 and 2004) to the other periods. Regarding our *explanatory variables*, GDP data (UNSD 2011) is subject to the same constant dollar conversion problems and the best measure of "distance" for our CS and time FE model is controversial (Head and Mayer 2001).

Fourth, apart from the dynamic models we mentioned and potential misspecification of the "remoteness" variable (subsection 3.2.7), alternative model specifications could be relevant. One example is the treatment of the Kyoto variable, which we discuss next.

None of our estimations found a consistently negative effect the Kyoto variable across goods types and sectors. One explanation could be a poor fit of the model, but this is implausible here since we also found insignificant or even positive estimates for differentiated goods estimations where the model worked well. Secondly, one might expect that pollution-intensive exports are more affected by policies introduced in response to ratification of an Annex B target than "clean" exports[46]. To rule out that the uniform treatment of all countries affects the estimates, we explored a specification where we weighted the Kyoto variable with the respective trade flow's carbon intensity, which did not affect the qualitative result. This is an ad-hoc solution; Gerlagh and Mathys (2010) build a more careful model of the effect of countries' energy abundance and sectors' energy intensity (which is related to emissions), but only apply it to 14 high-income countries to find a positive effect of energy abundance on net energy embodied in exports. Third, commitment to a Kyoto Annex B target might be a bad proxy for the actually implemented policies, since targets and efforts to achieve them vary by country, sectors are not treated equally and sometimes exempted, or policies might be ineffective. This does not impair tests for the impact of Kyoto membership as such, though. Fourth, as Aichele and Felbermayr (2010) argue, the Kyoto variable could be endogenous[47]. To capture all factors that might lead a country to join the Kyoto protocol, they include country FE and their interaction with time FE into their gravity model. These also absorb the GDP variable, which is necessary in our case since we are mainly concerned with FMR's predictions, but is of no interest if the focus is on the effect of "Kyoto". Aichele and Felbermayr (2010, p.27) report that "carbon imports of a committed country from a non-Kyoto exporter are about 10% higher than if the country had no commitments. This carbon leakage is strongest in the most carbon-intensive sectors." While the reason that we cannot confirm this finding might be our different focus and thus different modeling, we suggest that a reliable test of the effect of the Kyoto protocol on trade and EET requires a full general equilibrium model with the corresponding estimation techniques suggested by Anderson and van Wincoop (2003).

[46] It is not entirely clear if absolute pollution intensity is important, or intensity relative to the same exporter's emission intensity in other sectors. Under a global, binding climate policy regime, the intensity relative to competing exporters in the same sector would also be an alternative measure.
[47] See Antweiler et al. (2001) for another example of an endogenous environmental policy model, discussed in section 2.3.

6 Conclusion

Motivated by recent research highlighting the large and growing importance of trade-related carbon emissions and the imbalance of net embodied carbon flows for certain regions and countries, this thesis analyzed the patterns and some potential determinants of emissions embodied in trade, using the same dataset as Peters et al. (2011). We proceeded in two steps:

First, to identify the patterns more clearly, we decomposed the changes in emissions embodied in bilateral trade among 113 regions for 41 sectors between 1992 and 2006 into the effects of changes in the scale of bilateral trade volumes, in country- and sector-wise carbon intensities and in sector compositions of bilateral trade flows. We found the effect of scale increases to dominate, in particular for non-Annex B countries' exports, followed by intensity decreases, which are relatively important for Annex B exports. Contributing with a 4:1 ratio, the two net effects lead to a tripled net EET deficit for B countries. Composition effects as they are defined here are small throughout.

Second, we treated carbon intensities as given and chose a gravity model for an econometric analysis of the determinants of trade flows underlying the EET data. Specifically, we attempted to reproduce the results by Feenstra et al. (2001), who performed year-wise estimations for separate goods types to distinguish between different theoretical underpinnings of the gravity equation from the relative size of the estimated elasticities of trade with respect to exporters' and importers' GDP. They argue that varying results of gravity estimations in the literature are not driven by differing country characteristics, for example between OECD and non-OECD countries, but the type of goods dominating their trade. Moreover, we checked the results' robustness to the use of alternative estimation approaches.

We discussed theoretic and methodological issues regarding model specification and performed a series of tests to identify the most suitable pooling level. We then obtained three sets of estimation results: The first was for aggregate bilateral trade data for three goods types, estimated by pooled OLS with time effects or cross-section OLS as in Feenstra et al. (2001). A second set was obtained for the aggregate goods level as well, but with country-pair fixed effects and panel estimation techniques, namely within or first-difference estimation, which are preferable due to their robustness against unobserved heterogeneity bias. The third set was an application of these panel techniques to sector-level data.

Our results provide support for the findings of Feenstra et al. (2001) for differentiated goods, where elasticity with respect to exporters' GDP is larger than for the importers' GDP, implying that a monopolistic competition model with firm-level specialization and a home market effect is more appropriate in this case than a model with national product differentiation. Time effect, cross-section and within estimation yield this result at the goods and sector level and also for different region subsamples. Only the first difference estimates diverge, but these have a poor overall fit of the model.

However, we cannot confirm the results of Feenstra et al. for homogenous goods. Although we find a trend towards lower (higher) estimates of the exporter's (importer's) GDP compared to the differentiated goods case for all but the first-difference estimators, the importer elasticity is often still smaller than the exporter elasticity or only insignificantly larger. There is thus no support for restricted firm entry in a reciprocal dumping model, as argued by Feenstra et al., and neither for the opposite (free entry). The estimates for reference-priced goods, which were an intermediate case in Feenstra et al. (2001) similarly allow no conclusion regarding the appropriate model.

Panel estimations for homogenous goods trade for aggregate or separate sectors, with the full dataset or subsamples, invariably yields widely varying, often negative or insignificant parameter results and a low fit. Although we cannot rule out that this is due to flaws in our data, we doubt that our type of gravity specification is suitable at all to describe trade in homogenous commodities. Apart from the limited time dimension of our panel, this is the main reason why our trade model does not fully reproduce the observed EET patterns, as we illustrated by repeating the decomposition with "predicted EET": Homogenous commodities make up a large part of non-Annex B exports, which in turn yield the largest scale effects and are the main driver behind the net EET imports of Annex B countries, so these trade flows would have to be explained well.

To do this, it seems useful to distinguish and separately estimate different goods types, but either different models would have to be used for each, or a general model would have to take into account other determinants of bilateral trade volume than GDP and exogenous trade frictions, for example resource endowments, endogenous price and policy effects, or even structural change. Another option is a dynamic model, since we could not reject serial correlation in our panel estimations.

As a by-product, we obtained estimates for the impact of Kyoto membership on trade by using a dummy variable in our gravity specification. While we find no consistent evidence of a negative impact of Kyoto membership, this simple approach does not replace a full analysis of this issue, which has to take into account inter alia potential endogeneity of climate policy and general equilibrium effects.

7 References

Affendy, A.M., Sim Yee, L. and M.Satoru (2010): "*Commodity-industry classification proxy: A correspondence table between SITC revision 2 and ISIC revision 3",* Journal of International Economic Studies, 24: 185-202.

Aichele, R. and G.Felbermayr (2010), "*Kyoto and the carbon content of trade*", FZID Discussion Papers, Universität Hohenheim.

Anderson, J.E. (1979), *"A Theoretical Foundation for the Gravity Equation",* American Economic Review, 69(1): 106-116.

Anderson, J. E. and E. van Wincoop (2003), *"Gravity with Gravitas", American Economic Review, 93(1): 170-192.*

Anderson, J. E. and E. van Wincoop (2004), *"Trade Costs",* Journal of Economic Literature, 42: 691-751.

Anderson, J.E. (2011), "*The Gravity Model*", Annual Review of Economics, vol. 3.

Angn B.W. and F.Q. Zhang (2000), "*A survey of index decomposition analysis in energy and environmental studies"*, Energy (25): 1149–1176.

Antweiler, W., B.R. Copeland and M.S. Taylor, (2001) "*Is free trade good for the environment?",* American Economic Review, 91(4): 877–908.

Anzcerta (2011), *data on ANZCERTA membership*, retrieved on August 15th 2011 from http://203.6.168.90/fta/anzcerta/anzcerta_history.html

Arrow, K., B. Bolin, R. Costanza, P. Dasgupta, C. Folke, C.S. Holling, B.-O. Jansson, S. Levin, K.G. Maler and C. Perrings (1995) "*Economic Growth, Carrying Capacity, and the Environment"*, Science, 268: 520-21.

Bergstrand, J.H. (1985), *"The Gravity Equation in International Trade: some Microeconomic Foundations and Empirical Evidence"*, Review of Economics and Statistics, 67(3): 474-481.

Baier, S.L. and J. H. Bergstrand (2009), *"Bonus Vetus OLS: A Simple Method for Approximating International Trade-Cost Effects using the Gravity Equation",* Journal of International Economics, 77: 77-85.

Bergstrand, J. H. and P. Egger (2011), *"Gravity Equations and Economic Frictions in the World Economy",* in D. Bernhofen, R. Falvey, D. Greenaway and U. Krieckemeier (eds.), *Palgrave Handbook of International Trade*, Palgrave-Macmillan Press, in press; retrieved as a working paper from http://www.nd.edu/~jbergstr/Working_Papers/

Cheng, I-H. and H. Wall (2005), *"Controlling for Heterogeneity in Gravity Models of Trade and Integration"*, Federal Reserve Bank of St. Louis Review, 87(1): 49-63.

Cole, M.A. and R.J.R. Elliott (2003), *"Determining the trade–environment composition effect: the role of capital, labor and environmental regulations"*, Journal of Environmental Economics and Management, 46: 363–383.

Cole, M. A., Elliott, R. J. R. and K. Shimamoto (2005), *"Why the grass is not always greener: the competing effects of environmental regulations and factor intensities on US specialization"*, Ecological Economics, 54:95–109.

Copeland, B.R. and M.S. Taylor (2004), *"Trade, Growth, and the Environment"*, Journal of Economic Literature, 42(1): 7-71

Croissant, Y. and G. Millo (2008), *"Panel Data Econometrics in R: The plm Package "*, Journal of Statistical Software, 27(2).

Davis, D.R. (1995), "*Intra-industry trade: A Heckscher-Ohlin-Ricardo approach*", Journal of International Economics, 39: 201-226

Davis, S. J. and K. Caldeira (2010), *"Consumption-based accounting of CO_2 emissions"*, PNAS, 107(12): 5687–5692

Deardorff, A. V. (1998), "*Determinants of bilateral trade: does gravity work in a neoclassical world?",* in J. A. Frankel (ed.), *The Regionalization of the World Economy*, University of Chicago Press, Chicago.

Debaere, P. (2005), *"Monopolistic competition and trade, revisited: testing the model without testing for gravity"*, Journal of International Economics, 66: 249–266

Eaton, J. and S. Kortum (2002), *"Technology, Geography, and Trade"*, Econometrica, 70: 1741-1780.

Eaton, J. and A. Tamura (1994), *"Bilateralism and regionalism in Japanese and U.S. trade and direct foreign investment patterns"*, Journal of the Japanese and International Economies, 8:478–510.

Ederington, J. and J. Minier (2003), *"Is environmental policy a secondary trade barrier? An empirical analysis"*, Canadian Journal of Economics, 36:137-154

Egger, P. and M. Pfaffermayr (2003), *"The Proper Panel Econometric Specification of the Gravity Equation: A Three-Way Model with Bilateral Interaction Effects",* Empirical Economics , 28: 571-580

Feenstra, R.C., Markusen, J.R. and A.K.Rose (2001), *"Using the gravity equation to differentiate among alternative theories of trade"*, Canadian Journal of Economics, 34(2): 430-447

Feenstra, R.C. (2004), "*Advanced International Trade",* Princeton University Press, Princeton, New Jersey, USA.

Felbermayr, G. and Kohler, W. (2006), "*Exploring the Extensive and Intensive Margins of World Trade*", Review of World Economics, 142(4): 642-674

Frankel, J.A. and A.K. Rose (2005), *"Is trade bad or good for the environment: Sorting out the causality"*, Review of Economics and Statistics, 87(1): 85-91

Gerlagh, R. and N.A. Mathys (2010), *"Energy abundance, trade and industry location"*, working paper, Fondazione Eni Enrico Mattei.

Grossman, G.M. and A.B. Krueger (1993): "*Environmental Impacts of a North American Free Trade Agreement"*, in The U.S.-Mexico Free Trade Agreement. Peter M. Garber, ed. Cambridge, MA: MIT Press, pp. 13-56.

GTAP Global Trade Analysis Project (2011), *GTAP sector description*, retrieved on July 1st 2011 from https://www.gtap.agecon.purdue.edu/databases/contribute/detailedsector.asp

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

Haveman (2011), "*Rauch's classification of commodity trade"*, retrieved on July 7th 2011 from http://www.macalester.edu/research/economics/page/haveman/trade.resources/tradedata.html

Head, K. (2003), "*Gravity for beginners*", retrieved on April 10th 2011 from http://economics.ca/keith/gravity.pdf

Head, K. and T. Mayer (2001): "*Illusory Border Effects: How Far Is an Economy from Itself?"*, Working Paper, University of British Columbia

Helpman, E. (1987), *"Imperfect competition and international trade: evidence from fourteen industrial countries",* Journal of the Japanese and International Economics, 1: 62–81

Helpman, E. and P. Krugman (1985), "*Market Structure and Foreign Trade",* MIT Press, Cambridge, Massachusetts, USA.

Hoekstra, R. and J.J.C.J.M. van der Bergh (2003), *"Comparing structural and index decomposition analysis"*, Energy Economics, 25:39–64

Holtz-Eakin, D. and T.M. Selden (1995): "*Stoking the Fires? C02 Emissions and Economic Growth"*, Journal of Public Economices, 57(1): 85-101.

Hummels, D., Ishii, J. and K.M. Yi (2001), *"The nature and growth of vertical specialization in world trade"*, Journal of International Economics, 54: 75–96

Hummels, D. and J. Levinsohn (1995), *"Monopolistic competition and international trade: Reconsidering the evidence"*, Quarterly Journal of Economics 110, 799–836.

Hutcheson, T. (2006), *"HS2002-CPC 1.1-ISIC, Rev3-GTAP Concordance "*, GTAP Resource Center, https://www.gtap.agecon.purdue.edu/resources/res_display.asp?RecordID=1916

IEA, International Energy Agency (2011), *"World Energy Outlook 2011"*, Paris.

Krüger, J.J. (2008), *"Productivity and structural change: A review of the literature"*, Journal of Economic Surveys, Vol. 22(2): 330–363

Krugman, P. R. (1979), "*Increasing Returns, Monopolistic Competition, and International Trade"*, *Journal of International Economics*, 9: 469-479.

Krugman, P. R. (1980), *"Scale economies, product differentiation, and the pattern of trade*", American Economic Review, 70: 950–59

Long, J.S. and L.H. Ervin (2000), *"Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model"*, The American Statistician, 54:217–224.

MacKinnon, J.G. and H. White (1985), *"Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties"*, Journal of Econometrics, 29:305–325.

Mayer and Zignago (2006), *"Notes on CEPII's distances measures"*, retrieved together with a table of distance measures, contiguity and language dummies on August 10th 2011 from http://www.cepii.fr/anglaisgraph/bdd/distances.htm

Millo, G. (2009), post in the help forum of the R project on August 28th 2009, http://www.mail-archive.com/r-help@r-project.org/msg67156.html

Peters, G. and E. Hertwich (2008), "*CO2 Embodied in International Trade with Implications for Global Climate Change Policy*", Environmental Science & Technology, 42(5): 1401-1407

Peters, G., Minx, J., Weber, C.L. and O. Edenhofer (2011), *"Growth in emission transfers via international trade from 1990 to 2008"*, PNAS early edition, March 2011.

Rauch, J. (1999), *"Networks versus markets in international trade",* Journal of International Economics, 48: 7–37.

Romalis, J. (2004), *"Factor Proportions and the Structure of Commodity Trade"*, American Economic Review, 94(1): 67–97.

Santos Silva, J.M.C. and S. Tenreyro (2006), *"The log of gravity"*, The Review of Economics and Statistics, 88:641-658.

Schäfer, A. (2005), *"Structural change in energy use",* Energy Policy, 33: 429–437

Shafik, N. and S. Banyopadhyay (1992): *"Economic Growth and Environmental Quality: Time Series and Cross-Sectional Evidence"*, World Bank working paper 904.

Spatareanu, M. (2007): *"Searching for pollution havens: The impact of environmental regulations on foreign direct investment"*, Journal of Environment Development, 16 (2): 161–182

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

UNFCCC (2011a), ratification dates of the Kyoto Protocol, retrieved on August 12th 2011 from http://unfccc.int/kyoto_protocol/status_of_ratification/items/2613.php

UNFCCC (2011b), Annex B targets of the Kyoto Protocol, retrieved on August 12th 2011 from http://unfccc.int/kyoto_protocol/items/3145.php

UNSD United Nations Statistics Division (2011), "*System of National Accounts*", export and GDP data retrieved on August 20th from http://data.un.org/

Wikipedia (2011), data on membership in the EU, NAFTA, ASEAN, Mercosur and Caricom, retrieved on August 15th 2011 from http://en.wikipedia.org/

Wooldridge, J.M. (2000): "Introductory Econometrics – A Modern Approach", ...

Wooldridge, J.M. (2002): *"Econometric Analysis of Cross-Section and Panel Data"*, MIT Press.

Worldbank (2009), *"World Development index 2009"*, population data retrieved on August 20th from http://databank.worldbank.org/ddp/home.do

Zeileis, J. (2004), *"Econometric Computing with HC and HAC Covariance Matrix Estimators"*, Journal of Statistical Software, 11(10).

8 Appendix

8.1 Sectors and goods types

GTAP no.	GTAP code	Sector name	aggregate sector (Peters et al. 2011)	goods type conservative	goods type liberal
1	pdr	Paddy rice	Agriculture	hom	hom
2	wht	Wheat	Agriculture	hom	hom
3	gro	Cereal grains nec	Agriculture	hom	hom
4	v_f	Vegetables, fruit, nuts	Agriculture	ref	ref
5	osd	Oil seeds	Agriculture	ref	**hom**
6	c_b	Sugar cane, sugar beet	Agriculture	hom	hom
7	pfb	Plant-based fibers	Agriculture	hom	hom
8	ocr	Crops nec	Agriculture	ref	**hom**
9	ctl	Bovine cattle, sheep, goats, horses	Agriculture	hom	hom
10	oap	Animal products nec	Agriculture	ref	**hom**
11	rmk	Raw milk	Agriculture	-	-
12	wol	Wool, silk-worm cocoons	Agriculture	hom	hom
13	frs	Forestry	Agriculture	dif	**ref**
14	fsh	Fishing	Agriculture	ref	ref
15	col	Coal	Mining	dif	**ref**
16	oil	Oil	Mining	hom	hom
17	gas	Gas	Mining	hom	hom
18	omn	Minerals nec	Mining	ref	**hom**
19	cmt	Bovine meat products	Food	hom	hom
20	omt	Meat products nec	Food	ref	**hom**
21	vol	Vegetable oils and fats	Food	hom	hom
22	mil	Dairy products	Food	ref	**hom**
23	pcr	Processed rice	Food	hom	hom
24	sgr	Sugar	Food	hom	hom
25	ofd	Food products nec	Food	ref	ref
26	b_t	Beverages and tobacco products	Food	ref	ref
27	tex	Textiles	Manufacturing	dif	dif
28	wap	Wearing apparel	Manufacturing	dif	dif
29	lea	Leather products	Manufacturing	dif	dif
30	lum	Wood products	Manufacturing	dif	**hom**
31	ppp	Paper products, publishing	Energy-int. manufact.	ref	ref
32	p_c	Petroleum, coal products	Energy-int. manufact.	ref	**hom**
33	crp	Chemical, rubber, plastic products	Energy-int. manufact.	dif	dif
34	nmm	Mineral products nec	Energy-int. manufact.	dif	**ref**
35	i_s	Ferrous metals	Energy-int. manufact.	ref	ref
36	nfm	Metals nec	Energy-int. manufact.	hom	hom
37	fmp	Metal products	Manufacturing	dif	dif
38	mvh	Motor vehicles and parts	Manufacturing	dif	dif
39	otn	Transport equipment nec	Manufacturing	dif	dif
40	ele	Electronic equipment	Manufacturing	dif	dif
41	ome	Machinery and equipment nec	Manufacturing	dif	dif
42	omf	Manufactures nec	Manufacturing	dif	dif

Table 8: GTAP sectors (GTAP 2011) and their aggregation as in Peters et al. (2011) and using Rauch's (1999) of goods types (see section 3.3.2)

8.2 Sample statistics

		Totals			no. of countries exporting >= X% of tot. vol.			no. of countries importing >= X% of tot. vol.			no. of country-pairs with ... available periods								share of country-pairs with >= 1 periods	share of country-pairs with >= 2 periods
		no. of aggreg. sectors	N (w/o zeros, NAs, interreg. trade)	non-zero share	25%	50%	75%	25%	50%	75%	0 per.	1 per.	2 per.	3-5 per.	6-8 per.	9-11 per.	12-14 per.	15 per.		
Goods cons. classif.	hom.	14	61421	32%	2	7	17	2	7	17	5522	1048	555	1103	788	694	931	2015	56%	48%
	ref.-p.	13	91250	48%	4	11	25	3	8	21	3610	839	468	1146	940	807	1405	3441	71%	65%
	diff.	14	99590	52%	3	7	16	2	8	19	3309	803	467	916	898	741	1400	4122	74%	68%
Goods liberal classif.	hom.	22	84624	45%	3	9	22	2	7	18	3928	916	561	1131	909	829	1261	3121	69%	62%
	ref.-p.	10	94073	50%	3	8	19	3	8	21	3726	748	465	967	827	768	1378	3777	71%	65%
	diff.	9	90915	48%	3	6	15	2	7	18	3834	852	486	968	855	694	1279	3688	70%	63%
Homogenous goods sectros (conservative)	pdr	-	2237	1%	1	4	9	4	8	16	12164	178	70	105	39	35	28	37	4%	2%
	wht	-	8754	5%	1	3	6	5	12	25	10781	593	286	426	207	135	99	129	15%	10%
	gro	-	9532	5%	1	2	6	2	7	18	10807	525	302	374	216	144	122	166	15%	10%
	c_b	-	175	0%	1	1	1	1	1	2	12601	22	14	9	5	3	2	0	0%	0%
	pfb	-	11186	6%	2	3	9	3	8	17	10526	577	279	502	254	210	183	125	17%	12%
	ctl	-	5321	3%	2	5	10	2	3	9	11762	240	99	198	86	75	90	106	7%	5%
	wol	-	4005	2%	1	1	2	2	4	9	11950	180	89	152	88	83	67	47	6%	4%
	oil	-	8521	4%	1	3	7	2	5	13	11083	483	215	309	158	119	108	181	12%	9%
	gas	-	2034	1%	2	3	7	2	4	9	12203	169	65	82	56	27	27	27	4%	2%
	cmt	-	12920	7%	2	4	9	3	6	12	10685	438	215	390	264	178	177	309	16%	12%
	vol	-	19403	10%	2	6	12	4	11	27	9749	572	303	646	365	311	304	406	23%	18%
	pcr	-	6371	3%	1	3	7	3	10	20	11477	339	154	255	133	83	102	113	9%	7%
	sgr	-	14232	7%	2	5	15	3	9	22	10063	663	337	581	348	255	228	181	20%	15%
	nfm	-	40669	21%	4	9	19	3	7	15	7791	767	407	793	534	406	601	1357	38%	32%
Reference-priced goods sectors (cons.)	v_f	-	29231	15%	3	8	19	2	6	13	9133	535	292	577	428	328	467	896	28%	24%
	osd	-	11156	6%	1	2	6	2	5	14	10776	486	234	381	223	156	191	209	15%	11%
	ocr	-	38305	20%	3	9	22	2	6	16	8151	683	351	676	511	498	609	1177	36%	30%
	oap	-	19396	10%	2	6	15	3	7	15	10074	437	255	503	303	256	333	495	20%	17%
	fsh	-	8872	5%	4	10	22	2	5	10	11311	262	145	363	118	103	135	219	11%	9%
	omn	-	28264	15%	3	8	17	2	7	14	8848	670	383	766	437	326	451	775	30%	25%
	omt	-	17777	9%	2	5	11	2	5	13	10214	440	231	485	287	306	267	426	19%	16%
	mil	-	20301	11%	2	4	9	3	7	17	9835	531	310	543	330	256	312	539	22%	18%
	ofd	-	55672	29%	3	10	23	3	7	19	6584	721	412	911	712	532	868	1916	48%	42%
	b_t	-	27981	15%	2	5	12	3	6	17	9047	610	318	676	439	335	411	820	29%	24%
	ppp	-	43596	23%	3	6	12	3	8	19	7995	593	316	649	467	382	605	1649	37%	32%
	p_c	-	28464	15%	3	7	16	2	9	23	8329	940	465	907	479	426	614	496	34%	27%
	i_s	-	46801	25%	4	9	19	4	10	22	7204	796	479	811	573	530	771	1492	43%	37%
Differentiated goods sectors (conservative)	frs	-	10305	5%	2	4	12	2	5	12	10887	394	212	476	204	151	148	184	14%	11%
	col	-	6981	4%	1	3	7	1	6	14	11485	351	146	214	98	82	98	182	9%	6%
	tex	-	58288	31%	3	8	18	3	8	23	6688	660	375	775	562	530	834	2232	47%	42%
	wap	-	40307	21%	2	8	21	1	4	11	8319	504	313	601	473	409	612	1425	34%	30%
	lea	-	34871	18%	2	5	14	2	5	15	8711	532	289	603	455	358	548	1160	31%	27%
	lum	-	40579	21%	3	7	17	1	4	13	8217	531	311	676	463	429	668	1361	35%	31%
	crp	-	76475	40%	3	7	14	3	8	21	5258	638	423	829	712	661	1105	3030	58%	53%
	nmm	-	45780	24%	3	7	16	3	8	21	7877	601	263	594	493	482	674	1672	38%	33%
	fmp	-	47602	25%	3	7	15	2	8	20	7582	639	370	652	486	492	680	1755	40%	35%
	mvh	-	42497	22%	2	4	9	2	5	12	7797	682	378	785	492	438	608	1476	38%	33%
	otn	-	37748	20%	1	4	9	3	8	21	7489	1037	543	907	576	476	563	1065	41%	33%
	ele	-	51582	27%	3	6	11	2	6	15	7011	729	406	781	629	510	771	1819	45%	39%
	ome	-	73920	39%	2	5	12	3	8	20	5289	826	439	841	685	571	1029	2976	58%	52%
	omf	-	43440	23%	3	6	13	1	5	11	8047	617	310	595	413	408	592	1674	36%	32%
-	rmk	-	0	0%	0	0	0	0	0	0	12656	0	0	0	0	0	0	0	0%	0%

Table 9: Sample statistics for trade dataset (X, not EET)

8.3 More detailed decomposition results

		Total effect due to scale change		Total effect due to composition change		Total effect due to intensity change		Total change in EET	
		B	nB	B	nB	B	nB	B	nB
Manufacturing	B	387,43	180,35	-58,58	-16,21	-283,85	-89,61	45,00	74,53
	nB	943,26	607,16	-6,58	46,65	-465,18	-137,03	471,50	516,78
Energy-intensive manufacturing	B	649,91	340,35	76,12	63,21	-706,02	-306,07	20,01	97,49
	nB	479,46	721,94	55,56	26,03	-142,17	-114,27	392,84	633,70
Mining	B	170,25	23,58	245,57	23,18	-192,80	-12,75	223,02	34,01
	nB	162,78	152,23	82,63	27,18	-82,07	-23,12	163,34	156,29
Food	B	52,75	18,49	-18,16	-0,04	-58,47	-17,79	-23,88	0,65
	nB	49,30	50,30	-15,80	-12,68	-16,86	-5,88	16,65	31,74
Agriculture	B	30,67	29,38	-47,23	-9,30	-53,41	-32,64	-69,97	-12,56
	nB	-27,96	8,37	-27,01	-21,17	-4,21	-6,06	-59,18	-18,85

Cumulated yearwise decomposition with refined Laspeyres index method (interaction terms distributed symmetrically). Exporters in rows, importers in columns (B = Annex B countries (incl. USA), nB = non-Annex B countries).

Table 10: Decomposed changes in CO_2 emissions embodied in goods trade between non-Annex B and Annex B countries [$MtCO_2$], 1992-2006, for aggregate sectors (see Table 8)

Totals in first and last period		B* to nB	US to nB	nB to B*	nB to US	B* to B*	B* to US	US to B*	Net B* to nB	Net US to nB
EET '92	abs. [Mt CO_2]	438	149	716	295	1167	192	225	-278	-146
EET '06	abs. [Mt CO_2]	546	235	1310	686	1210	300	268	-764	-451
Trade '92	[10^9 US$, 2005]	480	207	571	312	1766	363	308	-91	-104
Trade '06	[10^9 US$, 2005]	1328	504	1409	912	3868	830	572	-82	-408
CO_2 intensity '92	[kg CO_2 / $]	0,91	0,72	1,25	0,95	0,66	0,53	0,73	-	-
CO_2 intensity '06	[kg CO_2 / $]	0,41	0,47	0,93	0,75	0,31	0,36	0,47	-	-
Decomposition summary										
ΔEEX	**abs. [Mt CO_2]**	**108**	**86**	**594**	**391**	**44**	**108**	**42**	**-486**	**-305**
	vs. '92 [%]	25%	58%	83%	133%	4%	56%	19%	175%	209%
thereof, due to…										
- scale change	**abs. [Mt CO_2]**	**419**	**173**	**965**	**642**	**838**	**260**	**192**	**-546**	**-469**
	direct effect	382	179	953	668	815	265	201	-571	-489
	vs. '92 [%]	96%	116%	135%	218%	72%	136%	85%	197%	321%
- composition change	**abs. [Mt CO_2]**	**61**	**0**	**56**	**33**	**173**	**50**	**-25**	**5**	**-32**
	direct effect	-99	-2	-123	26	3	41	-20	24	-28
	vs. scale eff.	15%	0%	6%	5%	21%	19%	-13%	-1%	7%
	vs. '92 [%]	14%	0%	8%	11%	15%	26%	-11%	-2%	22%
- intensity change	**abs. [Mt CO_2]**	**-372**	**-87**	**-427**	**-283**	**-967**	**-202**	**-125**	**55**	**196**
	direct effect	-500	-91	-575	-252	-1149	-195	-122	75	162
	vs. scale eff.	-89%	-50%	-44%	-44%	-115%	-78%	-65%	-10%	-42%
	vs. '92 [%]	-85%	-58%	-60%	-96%	-83%	-106%	-55%	-20%	-134%

Cumulated yearwise decomposition with refined Laspeyres index method (interaction terms distributed symmetrically).
B = Annex B countries without USA), nB = non-Annex B countries, EET=emissions embodied in trade.*

Table 11: Decomposed changes in CO2 emissions embodied in goods trade, 1992-2006, separating the USA from other Annex B countries.

8.4 Detailed estimation and test results

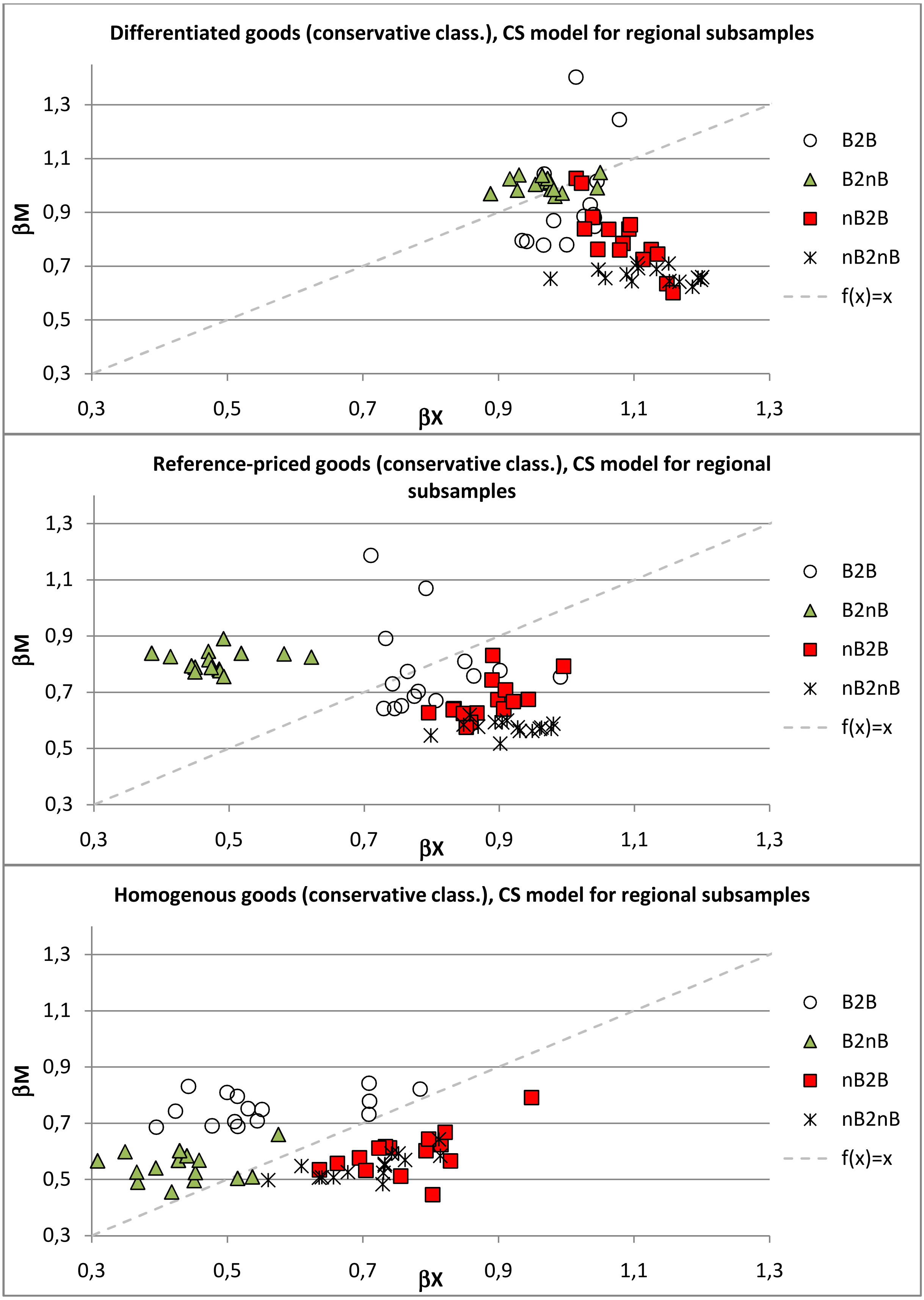

Figure 9: GDP elasticities from cross-section estimations for regional subsamples, aggregate goods types

Null hypothesis: Joint insignificance of FE; alternative: significant effects									
		time FE (vs. pooling)				Two-way FE (vs time FE)			
Goods type / sector		**F-statistic**	**p-value**	**df1**	**df2**	**F-statistic**	**p-value**	**df1**	**df2**
Homogenous goods		4,28	1,23E-07	14	60347	28,42	<1,0E-300	6082	54265
Reference-priced g.		3,69	3,26E-06	14	90385	30,99	<1,0E-300	8203	82182
Differentiated g.		13,34	2,88E-32	14	98761	38,66	<1,0E-300	8540	90221
Homogenous goods sectors	pdr	*0,68*	*7,98E-01*	*14*	*2033*	9,56	1,85E-219	310	1723
	wht	2,13	8,31E-03	14	8135	10,61	<1,0E-300	1278	6857
	gro	*1,45*	*1,23E-01*	*14*	*8981*	14,84	<1,0E-300	1320	7661
	c_b	2,42	4,80E-03	14	127	4,38	2,03E-08	29	98
	pfb	4,80	6,51E-09	14	10583	14,33	<1,0E-300	1549	9034
	ctl	5,26	4,98E-10	14	5055	21,05	<1,0E-300	650	4405
	wol	2,53	1,35E-03	14	3799	16,93	<1,0E-300	522	3277
	oil	7,33	1,96E-15	14	8012	23,29	<1,0E-300	1086	6926
	gas	*1,14*	*3,19E-01*	*14*	*1839*	25,90	<1,0E-300	280	1559
	cmt	4,17	2,37E-07	14	12456	27,56	<1,0E-300	1529	10927
	vol	4,01	5,74E-07	14	18805	17,93	<1,0E-300	2331	16474
	pcr	*1,26*	*2,24E-01*	*14*	*6006*	21,42	<1,0E-300	836	5170
	sgr	*1,59*	*7,44E-02*	*14*	*13543*	15,78	<1,0E-300	1926	11617
	nfm	4,98	2,23E-09	14	39876	26,01	<1,0E-300	4094	35782
Reference-priced goods sectors	v_f	5,61	5,64E-11	14	28670	42,24	<1,0E-300	2984	25686
	osd	2,13	8,27E-03	14	10644	21,87	<1,0E-300	1390	9254
	ocr	48,85	1,40E-135	14	37596	35,54	<1,0E-300	3818	33778
	oap	9,49	2,29E-21	14	18933	26,65	<1,0E-300	2141	16792
	fsh	*0,71*	*7,68E-01*	*14*	*8584*	30,57	<1,0E-300	1079	7505
	omn	2,05	1,16E-02	14	27568	25,03	<1,0E-300	3134	24434
	omt	2,09	9,75E-03	14	17311	26,92	<1,0E-300	1998	15313
	mil	8,48	1,29E-18	14	19744	29,29	<1,0E-300	2286	17458
	ofd	8,36	2,46E-18	14	54925	39,46	<1,0E-300	5347	49578
	b_t	4,96	2,50E-09	14	27345	28,58	<1,0E-300	2995	24350
	ppp	11,19	3,73E-26	14	42977	41,89	<1,0E-300	4064	38913
	p_c	34,73	5,82E-94	14	27498	15,28	<1,0E-300	3383	24115
	i_s	3,22	4,02E-05	14	45979	25,28	<1,0E-300	4652	41327
Differentiated goods sectors	frs	3,59	5,81E-06	14	9885	29,53	<1,0E-300	1371	8514
	col	*1,14*	*3,13E-01*	*14*	*6604*	24,36	<1,0E-300	816	5788
	tex	20,65	3,04E-53	14	57602	44,87	<1,0E-300	5304	52298
	wap	26,85	3,35E-71	14	39777	52,47	<1,0E-300	3829	35948
	lea	16,18	2,72E-40	14	34313	37,20	<1,0E-300	3409	30904
	lum	10,71	8,55E-25	14	40022	48,07	<1,0E-300	3904	36118
	crp	12,37	1,71E-29	14	75811	40,27	<1,0E-300	6756	69055
	nmm	66,40	2,00E-187	14	45153	30,49	<1,0E-300	4174	40979
	fmp	3,61	4,91E-06	14	46937	30,49	<1,0E-300	4431	42506
	mvh	8,87	1,03E-19	14	41789	27,30	<1,0E-300	4173	37616
	otn	3,30	2,69E-05	14	36685	12,72	<1,0E-300	4126	32559
	ele	16,64	1,20E-41	14	50827	39,70	<1,0E-300	4912	45915
	ome	15,87	1,77E-39	14	73068	37,54	<1,0E-300	6537	66531
	omf	45,51	6,23E-126	14	42797	48,79	<1,0E-300	3988	38809
-	rmk	-	-	-	-	-	-	-	-

Table 12: Results of F-test for time and country-pair FE, for aggregate goods types and sectors

Null hypothesis if NOT rejected:		no serial correl. in DIFFERENCED errors use FD estimator		no serial correl. in ORIGINAL errors use within estimator	
		Chi-squared statistic	p-value	Chi-squared statistic	p-value
Homogenous goods		2572,36	<1,00E-300	1284,33	2,88E-281
Reference-priced goods		3332,23	<1,00E-300	1581,37	<1,00E-300
Differentiated goods		2867,44	<1,00E-300	1561,86	<1,00E-300
Homogenous goods sectors	pdr	139,56	3,32E-032	8,88	2,89E-003
	wht	924,86	3,87E-203	99,63	1,84E-023
	gro	718,31	3,11E-158	187,93	9,01E-043
	c_b	12,16	4,87E-004	13,06	3,02E-004
	pfb	462,63	1,29E-102	381,28	6,53E-085
	ctl	115,01	7,85E-027	187,18	1,31E-042
	wol	49,86	1,65E-012	77,26	1,50E-018
	oil	323,39	2,64E-072	144,51	2,75E-033
	gas	21,02	4,54E-006	115,02	7,77E-027
	cmt	176,13	3,40E-040	489,32	2,00E-108
	vol	837,98	2,99E-184	568,52	1,18E-125
	pcr	323,66	2,31E-072	111,63	4,30E-026
	sgr	949,02	2,16E-208	167,24	2,97E-038
	nfm	1160,72	2,10E-254	1152,59	1,23E-252
Reference-priced goods sectors	v_f	701,05	1,77E-154	500,28	8,26E-111
	osd	683,75	1,02E-150	144,38	2,94E-033
	ocr	1275,22	2,74E-279	587,37	9,37E-130
	oap	474,07	4,16E-105	706,84	9,74E-156
	fsh	153,12	3,61E-035	240,47	3,10E-054
	omn	784,92	1,03E-172	573,85	8,15E-127
	omt	258,30	4,04E-058	608,48	2,39E-134
	mil	440,14	1,01E-097	318,63	2,88E-071
	ofd	1248,46	1,80E-273	1629,82	<1,00E-300
	b_t	284,30	8,68E-064	716,47	7,83E-158
	ppp	1154,34	5,11E-253	1071,75	4,56E-235
	p_c	2190,25	<1,00E-300	596,13	1,16E-131
	i_s	1731,08	<1,00E-300	1297,30	4,38E-284
Differentiated goods sectors	frs	180,19	4,40E-041	355,18	3,15E-079
	col	221,29	4,73E-050	180,07	4,67E-041
	tex	872,32	1,02E-191	2402,33	<1,00E-300
	wap	383,64	2,00E-085	1253,55	1,40E-274
	lea	342,91	1,49E-076	1254,48	8,82E-275
	lum	330,81	6,40E-074	1283,45	4,46E-281
	crp	2204,00	<1,00E-300	1658,24	<1,00E-300
	nmm	958,35	2,03E-210	1448,56	<1,00E-300
	fmp	1564,70	<1,00E-300	867,96	9,07E-191
	mvh	620,19	6,80E-137	1650,67	<1,00E-300
	otn	3081,91	<1,00E-300	338,82	1,15E-075
	ele	1349,49	1,99E-295	1429,27	<1,00E-300
	ome	2531,11	<1,00E-300	1535,78	<1,00E-300
	omf	1189,36	1,25E-260	909,65	7,82E-200
-	rmk	-	-	-	-

Table 13: Result of Wooldrigde's first-difference based test for autocorrelation

Differentiated goods - dependent variable: ln(Xsijt) (X: bilateral trade, regions pooled)

	Sector	frs	col	tex	wap	lea	lum	crp	nmm	fmp	mvh	otn	ele	ome	omf
Time effects, pooled OLS	**ln(YX)**	0,39	0,53	0,40	0,18	0,29	0,61	0,82	0,51	0,63	0,77	0,59	0,79	1,01	0,51
		(0,06)	(0,05)	(0,02)	(0,03)	(0,02)	(0,03)	(0,02)	(0,02)	(0,02)	(0,02)	(0,02)	(0,03)	(0,02)	(0,02)
	ln(YM)	0,21	0,26	0,46	0,55	0,46	0,61	0,65	0,57	0,57	0,63	0,49	0,69	0,68	0,57
		(0,05)	(0,06)	(0,02)	(0,02)	(0,02)	(0,02)	(0,01)	(0,02)	(0,02)	(0,02)	(0,02)	(0,02)	(0,01)	(0,02)
	Kyoto	*-0,04*	*-0,06*	-0,12	*-0,02*	*-0,01*	0,20	-0,07	0,10	-0,05	*-0,02*	-0,12	-0,25	-0,07	-0,05
		(0,06)	*(0,08)*	(0,02)	*(0,03)*	*(0,03)*	(0,03)	(0,02)	(0,02)	(0,02)	*(0,03)*	(0,03)	(0,03)	(0,02)	(0,03)
	param.	25	25	25	25	25	25	25	25	25	25	25	25	25	25
	N	10305	6981	58288	40307	34871	40579	76475	45780	47602	42497	37748	51582	73920	43440
	adj. R^2	0,167	0,243	0,444	0,377	0,335	0,441	0,638	0,548	0,568	0,525	0,346	0,45	0,658	0,444
2-way eff., within estimator	**ln(YX)**	*-0,30*	0,81	1,35	1,79	1,85	1,36	1,04	1,42	1,35	1,78	1,22	2,28	1,27	1,47
		(0,25)	(0,26)	(0,11)	(0,13)	(0,15)	(0,12)	(0,08)	(0,11)	(0,11)	(0,18)	(0,16)	(0,13)	(0,09)	(0,12)
	ln(YM)	1,60	1,21	0,77	0,95	1,31	1,09	0,55	1,08	1,03	0,98	*0,23*	1,28	1,07	0,92
		(0,32)	(0,38)	(0,10)	(0,12)	(0,14)	(0,13)	(0,07)	(0,10)	(0,09)	(0,12)	*(0,14)*	(0,12)	(0,08)	(0,10)
	Kyoto	*-0,04*	*0,01*	0,03	*0,04*	*0,02*	-0,10	0,05	*-0,02*	-0,05	-0,15	-0,08	-0,11	*-0,02*	*0,01*
		(0,05)	*(0,06)*	(0,02)	*(0,02)*	*(0,02)*	(0,02)	(0,01)	*(0,02)*	(0,02)	(0,02)	(0,03)	(0,02)	*(0,01)*	*(0,02)*
	param.	1396	841	5329	3854	3434	3929	6781	4199	4456	4198	4151	4937	6562	4013
	N	9911	6630	57628	39803	34339	40048	75837	45179	46963	41815	36711	50853	73094	42823
	adj. R^2	0,036	0,068	0,090	0,100	0,077	0,183	0,220	0,103	0,186	0,221	0,055	0,227	0,211	0,086
2-way effects, FD estimator	**ln(YX)**	*-0,10*	1,32	0,65	0,60	0,74	0,53	0,64	0,49	0,47	0,77	*0,15*	0,96	0,54	0,55
		(0,24)	(0,25)	(0,09)	(0,11)	(0,11)	(0,10)	(0,08)	(0,10)	(0,11)	(0,16)	*(0,19)*	(0,12)	(0,09)	(0,12)
	ln(YM)	*0,36*	*0,61*	0,85	0,89	1,16	0,75	0,69	1,00	0,85	1,25	0,44	1,17	1,10	0,77
		(0,32)	*(0,38)*	(0,08)	(0,10)	(0,11)	(0,11)	(0,06)	(0,09)	(0,09)	(0,11)	(0,19)	(0,09)	(0,07)	(0,09)
	Kyoto	*-0,03*	*-0,02*	-0,04	-0,03	-0,07	-0,05	-0,04	-0,04	-0,03	-0,05	*-0,02*	-0,07	-0,05	-0,09
		(0,03)	*(0,05)*	(0,01)	(0,01)	(0,01)	(0,01)	(0,01)	(0,01)	(0,01)	(0,01)	*(0,03)*	(0,01)	(0,01)	(0,01)
	param.	1396	841	5329	3854	3434	3929	6781	4199	4456	4198	4151	4937	6562	4013
	N	9911	6630	57628	39803	34339	40048	75837	45179	46963	41815	36711	50853	73094	42823
	adj. R^2	0,014	0,035	0,020	0,020	0,021	0,026	0,013	0,054	0,014	0,021	0,004	0,022	0,019	0,029

Heteroskedasticity- and autocorrelation-consistent standard errors in brackets, *insignificant values at the 5% level in italics.*
**R-squared for two-way effect models not directly comparable and time FE models.

Table 14: Selected estimation results at sector level, for differentiated goods.

Reference-priced goods - dependent variable: ln(Xsijt) (X: bilateral trade, regions pooled)														
		Sector												
		v_f	**osd**	**ocr**	**oap**	**fsh**	**omn**	**omt**	**mil**	**ofd**	**b_t**	**ppp**	**p_c**	**i_s**
Time effects, pooled OLS	**ln(YX)**	0,45	0,36	0,16	0,34	0,30	0,40	0,35	0,46	0,53	0,39	0,80	0,27	0,40
		(0,03)	(0,05)	(0,02)	(0,03)	(0,06)	(0,03)	(0,03)	(0,03)	(0,02)	(0,03)	(0,02)	(0,03)	(0,02)
	ln(YM)	0,32	0,12	0,32	0,26	0,53	0,25	0,20	0,19	0,48	0,26	0,53	0,45	0,52
		(0,03)	(0,04)	(0,02)	(0,02)	(0,05)	(0,03)	(0,03)	(0,03)	(0,02)	(0,02)	(0,02)	(0,03)	(0,02)
	Kyoto	-0,11	*-0,01*	*0,05*	*0,07*	*0,03*	-0,08	*-0,08*	0,17	*0,00*	*0,02*	*0,17*	*0,03*	*0,18*
		(0,03)	*(0,06)*	*(0,03)*	*(0,04)*	*(0,07)*	(0,04)	*(0,05)*	(0,04)	*(0,02)*	*(0,03)*	*(0,02)*	*(0,04)*	*(0,03)*
	param.	25	25	25	25	25	25	25	25	25	25	25	25	25
	N	29231	11156	38305	19396	8872	28264	17777	20301	55672	27981	43596	28464	46801
	adj. R^2	0,280	0,161	0,273	0,274	0,244	0,248	0,288	0,290	0,411	0,310	0,511	0,231	0,403
2-way eff., within estimator	**ln(YX)**	0,74	0,43	0,84	0,59	*0,30*	0,61	*-0,26*	-0,47	0,77	*0,00*	0,97	0,98	0,71
		(0,13)	(0,20)	(0,11)	(0,16)	*(0,31)*	(0,16)	*(0,23)*	(0,20)	(0,10)	*(0,17)*	(0,12)	(0,17)	(0,12)
	ln(YM)	0,91	1,05	0,54	0,61	1,09	1,09	*0,18*	0,47	0,44	*0,26*	0,79	0,90	0,94
		(0,15)	(0,41)	(0,14)	(0,20)	(0,34)	(0,17)	*(0,20)*	(0,16)	(0,09)	*(0,14)*	(0,11)	(0,19)	(0,12)
	Kyoto	*0,00*	*0,04*	*0,05*	*0,05*	*0,00*	*0,00*	-0,12	-0,19	*0,00*	*-0,04*	*0,02*	*-0,05*	-0,06
		(0,02)	*(0,04)*	*(0,02)*	*(0,03)*	*(0,05)*	*(0,03)*	(0,04)	(0,03)	*(0,02)*	*(0,02)*	*(0,02)*	*(0,04)*	(0,02)
	param.	3009	1415	3843	2166	1104	3159	2023	2311	5372	3020	4089	3408	4677
	N	28696	10670	37622	18959	8610	27594	17337	19770	54951	27371	43003	27524	46005
	adj. R^2	0,052	0,016	0,052	0,025	0,067	0,064	0,076	0,060	0,082	0,067	0,149	0,107	0,097
2-way eff., FD estimator	**ln(YX)**	0,37	*0,19*	0,34	0,35	*-0,35*	0,45	*-0,00*	*0,12*	0,29	*0,04*	0,47	0,85	*0,01*
		(0,11)	*(0,21)*	(0,09)	(0,14)	*(0,23)*	(0,15)	*(0,20)*	*(0,16)*	(0,10)	*(0,15)*	(0,13)	(0,19)	*(0,12)*
	ln(YM)	0,50	0,74	0,29	0,61	*0,34*	0,31	0,35	0,40	0,35	*0,18*	0,73	0,36	1,17
		(0,14)	(0,25)	(0,11)	(0,13)	*(0,25)*	(0,15)	(0,15)	(0,13)	(0,09)	*(0,13)*	(0,09)	(0,18)	(0,12)
	Kyoto	*-0,03*	*-0,03*	*-0,01*	*0,02*	*0,03*	-0,05	-0,04	-0,05	-0,06	-0,05	-0,04	*-0,03*	*0,01*
		(0,01)	*(0,03)*	*(0,01)*	*(0,02)*	*(0,03)*	(0,02)	(0,02)	(0,02)	(0,01)	(0,02)	(0,01)	*(0,04)*	*(0,01)*
	param.	3009	1415	3843	2166	1104	3159	2023	2311	5372	3020	4089	3408	4677
	N	28696	10670	37622	18959	8610	27594	17337	19770	54951	27371	43003	27524	46005
	adj. R^2	0,010	0,011	0,027	0,015	0,023	0,009	0,010	0,012	0,011	0,007	0,023	0,056	0,020

Heteroskedasticity- and autocorrelation-consistent standard errors in brackets, *insignificant values at the 5% level in italics.*
**R-squared for two-way effect models not directly comparable and time FE models.

Table 15: Selected estimation results at sector level, for reference-priced goods.

Homogenous goods - dependent variable: ln(Xsijt) (X: bilateral trade, regions pooled)		**Sector**													
		pdr	**wht**	**gro**	**c_b**	**pfb**	**ctl**	**wol**	**oil**	**gas**	**cmt**	**vol**	**pcr**	**sgr**	**nfm**
Time effects, pooled OLS	**ln(YX)**	0,31	0,39	0,36	*0,03*	0,16	0,21	0,39	0,23	*-0,03*	0,36	0,22	0,20	0,18	0,59
		(0,05)	(0,04)	(0,05)	*(0,18)*	(0,04)	(0,08)	(0,06)	(0,05)	*(0,13)*	(0,04)	(0,03)	(0,05)	(0,03)	(0,03)
	ln(YM)	*0,05*	0,15	0,20	0,31	0,07	0,19	0,12	0,39	0,44	0,26	0,08	0,09	*0,02*	0,62
		(0,06)	(0,04)	(0,04)	(0,14)	(0,04)	(0,06)	(0,06)	(0,06)	(0,12)	(0,04)	(0,02)	(0,05)	*(0,03)*	(0,02)
	Kyoto	*-0,21*	*0,14*	*0,06*	*-0,08*	*-0,03*	*0,11*	-0,18	-0,25	*-0,13*	*0,01*	-0,13	*-0,08*	*0,05*	*0,05*
		(0,12)	*(0,06)*	*(0,06)*	*(0,33)*	*(0,06)*	*(0,09)*	(0,09)	(0,10)	*(0,19)*	*(0,06)*	(0,04)	*(0,08)*	*(0,04)*	*(0,03)*
	param.	25	25	25	25	25	25	25	25	25	25	25	25	25	25
	N	2237	8754	9532	175	11186	5321	4005	8521	2034	12920	19403	6371	14232	40669
	adj. R^2	0,176	0,229	0,17	0,369	0,06	0,187	0,238	0,158	0,242	0,266	0,155	0,146	0,109	0,375
2-way eff., witin estimator	**ln(YX)**	*-0,28*	*0,60*	0,51	*1,06*	1,25	*-0,57*	0,91	0,70	*0,03*	-0,93	*0,01*	*0,57*	*-0,05*	0,64
		(0,70)	*(0,35)*	(0,25)	*(1,54)*	(0,22)	*(0,38)*	(0,44)	(0,26)	*(0,56)*	(0,28)	*(0,23)*	*(0,32)*	(0,23)	(0,14)
	ln(YM)	-1,79	0,68	0,94	*-4,42*	0,88	*0,71*	*0,00*	1,28	*0,26*	*0,41*	*-0,20*	*-0,25*	0,20	1,08
		(0,72)	(0,28)	(0,31)	*(2,77)*	(0,31)	*(0,46)*	*(0,42)*	(0,51)	*(0,85)*	*(0,23)*	*(0,20)*	*(0,28)*	(0,20)	(0,15)
	Kyoto	*-0,02*	-0,11	-0,10	*0,61*	*0,01*	*0,12*	-0,17	*-0,02*	*-0,16*	*-0,04*	-0,24	*0,02*	*0,04*	*0,01*
		(0,12)	(0,05)	(0,05)	*(0,39)*	*(0,05)*	*(0,07)*	(0,09)	*(0,07)*	*(0,11)*	*(0,05)*	(0,04)	*(0,06)*	*(0,04)*	*(0,02)*
	param.	335	1303	1345	54	1574	675	547	1111	305	1554	2356	861	1951	4119
	N	2059	8161	9007	153	10609	5081	3825	8038	1865	12482	18831	6032	13569	39902
	adj. R^2	0,039	0,019	0,018	0,314	0,075	0,026	0,124	0,086	0,117	0,023	0,096	0,023	0,023	0,092
2-way eff., FD estimator	**ln(YX)**	*-0,04*	*0,04*	0,97	*1,40*	1,35	*-0,57*	*0,36*	1,08	*0,53*	-0,41	*0,41*	*0,31*	0,55	0,38
		(0,50)	*(0,38)*	(0,31)	*(1,39)*	(0,19)	*(0,34)*	*(0,29)*	(0,28)	*(0,64)*	(0,22)	*(0,22)*	*(0,35)*	(0,22)	(0,14)
	ln(YM)	*-1,05*	0,97	0,86	*-0,10*	0,66	1,24	*0,53*	*0,77*	*1,05*	0,65	*0,16*	*0,12*	*0,16*	0,75
		(0,61)	(0,32)	(0,27)	*(2,21)*	(0,26)	(0,47)	*(0,31)*	*(0,44)*	*(0,66)*	(0,17)	*(0,16)*	*(0,34)*	*(0,22)*	(0,13)
	Kyoto	-0,22	*-0,04*	*-0,05*	*0,49*	-0,07	*0,01*	-0,14	-0,14	-0,22	*-0,01*	*-0,05*	*0,02*	*-0,04*	*-0,02*
		(0,08)	*(0,05)*	*(0,04)*	*(0,26)*	(0,03)	*(0,05)*	(0,06)	(0,05)	(0,09)	*(0,03)*	*(0,03)*	*(0,05)*	*(0,03)*	*(0,02)*
	param.	335	1303	1345	54	1574	675	547	1111	305	1554	2356	861	1951	4119
	N	2059	8161	9007	153	10609	5081	3825	8038	1865	12482	18831	6032	13569	39902
	adj. R^2	0,020	0,010	0,010	0,259	0,024	0,016	0,058	0,029	0,040	0,014	0,030	0,010	0,030	0,020

Heteroskedasticity- and autocorrelation-consistent standard errors in brackets, *insignificant values at the 5% level in italics.*
**R-squared for two-way effect models not directly comparable and time FE models.

Table 16: Selected estimation results at sector level, for homogenous goods.