

TI 2017-121/VI
Tinbergen Institute Discussion Paper



Untangling Real Gravity

Franç Klaassen¹
Rutger Teulings¹

¹ University of Amsterdam; Tinbergen Institute, The Netherlands

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

Untangling Real Gravity

FRANC KLAASSEN

University of Amsterdam and Tinbergen Institute

and

RUTGER TEULINGS *

University of Amsterdam and Tinbergen Institute

December 29, 2017

Abstract

We derive a real gravity equation and gain several new insights that were hidden in the nominal specification used so far. Most importantly, the real effective exchange rate (REER) of the exporter and, via the importer's terms of trade, also the importer's REER matter, and we can identify the elasticity of substitution. We estimate real gravity for 18 OECD countries. Therefore, we extend the untangling normalization method from an *it* to an *ijt* panel data model and use it to exploit all variables proposed by theory, despite a broad set of fixed effects (FE). We find that both REERs are important and estimate an elasticity of substitution of 1.5. If we assume homogeneous parameters, as is common, the remaining unexplained exporter-time and importer-time deviations are still substantial, relaxing this assumption improves this. We now explain 64 and 70% of the exporter-time and importer-time deviations, respectively, and thus the majority of the exporter and importer multilateral resistances. Untangling normalization helps to get a better view of what is still unexplained by theory.

Key words: real gravity equation, exchange rate, multilateral trade resistance, untangling normalization.

JEL classification: C18; F10; F14.

*Corresponding author. Address: Amsterdam School of Economics, PO Box 15867, 1001 NJ Amsterdam, The Netherlands; tel. +31-20-5254157; e-mail: r.m.teulings@uva.nl.

1 Introduction

The gravity model is the workhorse model for explaining international trade flows, as is explained by [Head & Mayer \(2014\)](#). In this paper we contribute in two ways to the existing gravity literature. First, we derive a real gravity equation and we show that it has several advantages over the existing nominal gravity equation. Furthermore, the real gravity equation gives us some new insights. For example, it shows how the terms of trade (ToT) and the real effective exchange rate matter for exports. Second, we log-linearize the real gravity equation and we apply it to explain bilateral export flows between 18 OECD countries in the time period 1965-2011. We use untangling normalization, as proposed by [Klaassen & Teulings \(2015\)](#), to investigate the exporter and importer multilateral resistance (MR). We show what the model can explain, but also what is still not accounted for in the model.

For our first contribution, we take the standard nominal gravity equation by [Anderson & van Wincoop \(2003\)](#) as a starting point to derive a real gravity equation. We carefully ensure that we account for the different variable dimensions in the nominal gravity equation improving the standard gravity equation.

To properly account for the different currency dimensions in the nominal gravity equation we introduce the exchange rate into the equation, because it is important in explaining exports, see for example [Baldwin & Krugman \(1989\)](#) and [Campa & Goldberg \(2005\)](#). Still, it is typically omitted in the gravity equation. [Bergstrand \(1985\)](#) is an exception.

Now we can derive the real gravity equation. So we have an equation that explains quantities of exports instead of the value of exports. The real equation is consistent with the nominal model; so we use the same underlying model. One advantage is that we can fully account for endogenous price changes in nominal output that in the nominal equation is typically assumed constant. We only impose exogeneity on real output, so less stringent.

Real gravity gives us four important new insights. The first insight is that we need ToT to determine the purchasing power of the importing country. In the nominal equation the ToT of the importer is hidden in the endogenous nominal GDP. If its exporter price index increases it earns more from exporting and can therefore import more. However, if its imports become more expensive relative to its own good it can afford less import. Substituting out all endogenous variables introduces the importer multilateral resistance (MR) from the supply side. The intuition behind this is that, if the importer MR from the supply side goes up its own good becomes more expensive for

the rest of the world (RoW), reducing demand for its good and therefore its purchasing power.

The second insight is that the importer MR in the form of the consumer price index (CPI) no longer plays a role, because the CPI is absorbed in order to deflate nominal variables. In the nominal equation the importer MR affects exports through the demand side. Instead in the real gravity equation a new importer MR is introduced from the supply side, as explained above.

The third insight, is that the real gravity equation not only depends directly on the real exchange rate, but also indirectly through the exporter and importer MR because they capture the real effective exchange rate (REER) of the exporter and importer, respectively. The nominal gravity equation also depends indirectly on the exchange rate through the MRs, only now the latter capture the nominal effective exchange rate (NEER). Real gravity shows that if the REER depreciates vis-à-vis the RoW, the exporting country becomes more competitive leading to higher demand by the RoW for its good reducing bilateral export flows with the importing country. If the REER of the importer depreciates it becomes more competitive leading to higher purchasing power for the importing country, so it will import more.

Finally, the fourth insight is that we can identify the elasticity of substitution directly. while in the nominal gravity equation the elasticity of substitution can only be estimated indirectly. We find, using the exchange rate, a significant estimate of 1.51, close to the median estimate for this identification channel of 2.38 by [Head & Mayer \(2014\)](#) in their literature review.

We bring both the nominal and the real gravity equation to the data and show that three restrictions (exporter and world GDP have the same absolute impact and exchange rates do not matter for exports) implied by the nominal equation are not supported by the data. Real gravity relaxes two of these assumptions. However, the restriction on the exchange rate of the exporter is still rejected, offering a direction for potential future research.

For our second contribution we need to exploit the explanatory power of all variables that follow from the real gravity equation. Therefore, we use untangling normalization by [Klaassen & Teulings \(2015\)](#). This method is developed for a two dimensional *it*-panel data model. We extend their approach to a three dimensional *ijt*-panel data model for a broad range of FE-types and constant regressors, that is regressors that are constant in one or more dimensions and vary in the remaining. Untangling normalization uses orthogonality conditions as normalizations, untangling different effects in the data and let them be captured by the different fixed effect (FE) types in the model. For example,

it lets the constant capture the overall intercept, the trend the overall trend, the country FE the country specific deviation from the overall intercept and the country trend FE the country specific deviations from the overall trend. This improves interpretation of the FE estimates allowing for new insights into potentially ignored variables. Untangling normalization also allows us to exploit the explanatory power and possibly identify the impact of the constant regressors that are otherwise multicollinear with the FE. [Klaassen & Teulings \(2015\)](#) develop a test to examine whether we actually identify the true value or the pseudo-true value of the impact of the constant regressors.

By applying untangling normalization we can visualize the exporter-time and importer-time variation in our bilateral export equation, an *ijt*-panel data model, and see how much is explained by the variables proposed by real gravity. The remaining variation is captured by the country time FE. In the standard model we explain 9% and 54% of the exporter-time and importer-time variation, respectively, by the constant regressors. Allowing for heterogeneous parameters, this improves to 64 and 70%, respectively. This indicates that there is still a substantial amount that the gravity model cannot explain in bilateral export flows. However, by using untangling normalization we get a better view of the remaining unexplained variation. This offers new starting points for future research.

The remainder of the paper is structured as follows. In [Section 2](#) we extend the nominal gravity equation with the exchange rate. Next, we transform the nominal gravity equation into a real gravity equation in [Section 3](#). In [Section 4](#) we introduce our empirical model equation based on the theoretical real gravity equation and describe the data and our estimation method, giving special attention to the extension of the untangling normalization method from a two to a three dimensional panel data model. We present the results in [Section 5](#). In [Section 6](#) we investigate and untangle the exporter and importer multilateral resistance. Finally, we conclude in [Section 7](#).

2 The exchange rate in the nominal gravity model

In this section we will introduce the nominal gravity equation, based upon the model of [Anderson & van Wincoop \(2003\)](#), henceforth abbreviated as AvW, extended with the nominal exchange rate, such that we can relax the implicit assumption of single currency and ensure dimensions match. We derive this extended nominal model step by step so that in the real derivation the differences with the nominal part comes out clearly.

2.1 Export demand

We use the structural gravity model by AvW as a starting point. They start with several basic assumptions. On the supply side they assume that each country is endowed with only one distinguishable good and its real endowment is fixed. On the demand side they assume that the consumers in all countries have identical and homothetic preferences given by a CES utility function, all countries import for consumption use only and trade is balanced. Like AvW, we define country i (Home) to be the country that exports goods to country j (Foreign). In total we have J countries.

However, where AvW implicitly assume that there is one global currency, we allow each country to have its own currency. Therefore, each nominal variable has its own currency denomination. Country-specific variables, such as income, are denoted in the currency of their country. Country-pair-specific variables, such as exports from Home to Foreign, can be in Home or Foreign currency. So, we use an i or j in the superscript to indicate that a variable is in either Home (i) or Foreign (j) currency, respectively. As tractable example we use Home as a euro country (€) and Foreign as Japan (¥) throughout this section.

2.1.1 CES-structure

Given the above assumptions, each consumer in country j wants to maximize his utility subjective to the budget constraint:

$$\max_{x_{ij}} c_j = \left(\sum_i \lambda_i^{\frac{1}{\sigma}} x_{ij}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad (1)$$

$$s.t. \sum_i p_{ij}^j x_{ij} = E_j, \quad (2)$$

where $c_j = E_j/P_j$ is the aggregate consumption index, x_{ij} is the quantity that country j consumes of the good of country i (note that x_{jj} is the quantity country j consumes of its own good), σ is the elasticity of substitution between goods and is assumed to be larger than one (so all goods are gross substitutes), λ_i is a taste parameter for good i (where $\sum_i \lambda_i = 1$ must hold and is independent of j due to identical preferences), E_j is the nominal expenditure of country j and p_{ij}^j is the price country j has to pay to buy good i in j currency.

We maximize the consumer's utility (1) with respect to the budget constraint (2) for each j and obtain the exports demand for good i by country j in nominal terms,

using that the nominal value of exports in j currency is given by $X_{ij}^j = x_{ij}p_{ij}^j$,

$$X_{ij}^j = \lambda_i \left(\frac{p_{ij}^j}{P_j} \right)^{1-\sigma} E_j, \quad (3)$$

where P_j is give by

$$P_j = \left[\sum_i \left(p_{ij}^j \right)^{1-\sigma} \lambda_i \right]^{\frac{1}{1-\sigma}}. \quad (4)$$

From (4) it follows that P_j is the price index of country j , considering that the right-hand-side (RHS) is a weighted summation over all good prices country j faces, including its own good, and the taste parameter λ_i functions as weight. Hence, one unit of the aggregate consumption index c_j has price P_j .

2.1.2 Introducing the exchange rate

So far we did not further specify p_{ij}^j . Now we introduce the exchange rate. If Foreign (j) imports goods from Home (i) it needs to exchange its income from Foreign into Home currency to be able to pay for the goods of Home. So we need an exchange rate. The exchange rate has a strong effect on exports. If, say, Home currency depreciates with respect to Foreign currency, Home becomes cheaper for Foreign but Foreign becomes more expensive for Home. We define the nominal exchange rate, S_{ij} , to be the Home currency price for one unit of Foreign currency, euro per yen ($\text{€}/\text{¥}$) in our example. So an increase in S_{ij} implies a deprecation of Home currency.

The price country j pays for good i in j currency, p_{ij}^j , depends on the exporters factory gate price, p_i , the bilateral trade barrier factor for exporting from country i to j , b_{ij} , and the nominal exchange rate to change the denomination from i currency to j currency, or in our example from € to ¥ . So we get

$$p_{ij}^j = p_{ij}^i / S_{ij} = p_i b_{ij} / S_{ij}, \quad (5)$$

where the trade barrier b_{ij} is defined as

$$b_{ij} = t_{ij} \left(\frac{S_{ij}}{S_{ij}^b} \right)^\xi. \quad (6)$$

The bilateral trade barrier consists of two parts. The first part, t_{ij} , is the bilateral trade cost factor as in AvW. This factor captures everything from the bilateral distance between two countries to whether the two countries signed an FTA. The second part,

$\left(S_{ij}/S_{ij}^b\right)^\xi$, captures how much a change in the nominal exchange rate with respect to some base level S_{ij}^b affects the price p_{ij}^i , such that $\left(S_{ij}/S_{ij}^b\right)^\xi 1/S_{ij}$ captures the exchange rate pass-through. This way of modeling the exchange rate pass-through is not new and can for example be seen in [Sutherland \(2005\)](#).¹

The parameter ξ measures the flexibility in the price level. If $\xi = 0$ a change in the nominal exchange rate does not affect p_{ij}^i and is completely passed-through to the importer price p_{ij}^j . This is the baseline value in the remainder of the paper. If $\xi = 1$, a change in the exchange rate is completely incorporated into p_{ij}^i and not passed-through to p_{ij}^j . In this case the price p_{ij}^i increases (decreases) with respect to the factory gate price in case of a depreciation (appreciation). For every value of $0 < \xi < 1$ the change in the exchange rate is partially absorbed by p_{ij}^i and partially passed-through to p_{ij}^j . Perverse exchange rate pass-through, i.e. when $\xi > 1$ or $\xi < 0$, is possible and in this case p_{ij}^i is higher or lower than the actual change in the nominal exchange rate (cf. [Kenneth A. Froot \(1989\)](#)).

We now substitute (5) into (3) introducing the exchange rate into the gravity model

$$X_{ij}^j = \lambda_i \left(\frac{p_i b_{ij} / S_{ij}}{P_j} \right)^{1-\sigma} E_j. \quad (7)$$

Note, that if we set $S_{ij} = 1$ for all country pairs, we are back in the standard AvW setting.

2.2 Equilibrium prices

In equilibrium all markets clear, implying supply equals demand by all countries j for good i as there is no capital in this model. So, the market clearance condition for good i becomes

$$Y_i = \sum_j X_{ij}^i, \quad (8)$$

where $X_{ij}^i = X_{ij}^j S_{ij}$. We substitute (7) into the market clearing condition and solve for equilibrium prices p_i

$$p_i = \left(\frac{w_i^y}{\lambda_i} \right)^{\frac{1}{1-\sigma}} \frac{1}{\Pi_i}, \quad (9)$$

¹ [Anderson et al. \(2016\)](#) introduce a similar ratio as $\left(S_{ij}/S_{ij}^b\right)^\xi$ into the gravity model via the bilateral trade cost factor. However, they do not explicitly account for all the different currency dimensions in the gravity model and still assume a single currency. We allow countries to have different currencies and ensure that dimensions match.

where Π_i is given by

$$\Pi_i = \left[\sum_j \left(\frac{b_{ij}/S_{ij}}{P_j} \right)^{1-\sigma} w_j^e \right]^{\frac{1}{1-\sigma}} \quad (10)$$

and we define $w_i^y = Y_i/Y_W^i$ as the output share of country i with respect to the total world income in i currency, Y_W^i , and $w_j^e = E_j/Y_W^j$ as the expenditure share of country j with respect to the total world output in j currency.² Both w_j^e and w_i^y are dimensionless.

We replace p_i , after using (5), in (4) by (9) and obtain

$$P_j = \left[\sum_i \left(\frac{b_{ij}/S_{ij}}{\Pi_i} \right)^{1-\sigma} w_i^y \right]^{\frac{1}{1-\sigma}}. \quad (11)$$

Let us now check the dimensions in the nominal gravity model. We start by checking the dimensions of P_j and Π_i . The RHS of (4) sums over all prices in $\mathbb{Y}$. So, given that the powers cancel out, P_j has the dimension $\mathbb{Y}$. This is in line with the interpretation of P_j being a price index. Using this result, the RHS in (10) has the dimension $1/\mathbb{E}$, so we know the dimension of Π_i on the left-hand-side (LHS). To check this result, we compare the RHS of (11) with P_j on the LHS, both have the dimension $\mathbb{Y}$ so the dimensions match. Finally, we check the dimensions of (7). The fraction on the RHS is dimensionless, using that P_j has dimension $\mathbb{Y}$. Since E_j is the only variable with a currency dimension, the RHS must have dimension $\mathbb{Y}$. This coincides with the dimension of X_{ij}^j on the LHS.³

In the literature Π_i and P_j are usually interpreted as the multilateral resistance (MR) of the exporter and importer, respectively, because both depend positively on the trade cost barrier. However, they capture more than just the resistance to export and import, they capture supply side shocks, changes in the supply of good i affecting w_i^y (see (11)), and demand side shocks, changes in the preference parameter λ_i (see (4)). Hence, Π_i is the exporter MR from the supply side and P_j the importer MR from the demand side, where we already established for the latter that it is the consumer price index of country j . Only if the distinction is necessary we will use this refinement.

The exporter MR in (10) is a reciprocal measure of competitiveness of country i . The higher Π_i , the higher the resistance to sell good i and the lower the competitiveness of country i . This is in line with the currency dimension of Π_i ; it measures the resistance

²Note that total world expenditure equals total world output.

³Note that AvW consider the solution $\Pi_i = P_i$ when they solve (10) and (11) for P_i and Π_i . However, the dimensions of Π_i show that this solution, although mathematically valid, cannot be true in economic terms. If we use $\Pi_i = P_i$ and assume a common currency, say, $\mathbb{E}$, then the currency dimensions on the LHS of the standard nominal gravity equation becomes $\mathbb{E}$ and on the RHS $\mathbb{E}^{2\sigma-1}$. For the dimensions to match we need $\sigma = 1$, while we typically assume $\sigma > 1$.

to supply good i for every euro spend. It increases if the bilateral barrier b_{ij} increases because country i becomes less competitive. If currency i depreciates good i becomes cheaper for country j and country i becomes more competitive. Finally, if P_j increases, country i becomes relatively cheaper for country j compared to the rest of the world (RoW) making country i more competitive, decreasing Π_i . The expenditure shares serve as weight.

The importer MR in (11) increases with the trade barrier b_{ij} because good i becomes more expensive. If currency i depreciates, good i becomes cheaper. Finally, if Π_i of country i increases, demand for good i by the RoW will decrease making good i cheaper, also for country j , so that P_j decreases. The output shares serve as weight in the price index.

Now we can also better understand the equilibrium price expression (9) for good i . It increases with the preference parameter for good i , increasing demand. It decreases if the output share of country i increases because of a higher supply of good i . Finally, it decreases if Π_i increases because of a decreases in the competitiveness of country i , leading to less demand for good i .

2.3 Balanced trade

So far we left the relation between the endogenous determined expenditure and output of a country unspecified. But to be able to determine expenditure and solve the gravity model we introduce balanced trade

$$E_j = Y_j, \tag{12}$$

although this is in reality an unrealistic assumption because trade is almost never balanced, see for example the prolonged current account deficit of the US. However, in this model balanced trade follows from the fact that this is a cross-sectional model without capital such that there is no possibility for saving or borrowing and therefore no possibility for inter-temporal trade. Therefore, countries can only spend what they earn and trade is automatically balanced.⁴

In the literature Y_j is often assumed as constant, as is noted by [Head & Mayer \(2014\)](#), both in theory as well as in empirics. A general equilibrium approach, such as

⁴[Anderson et al. \(2015\)](#) introduce the possibility of imbalanced trade through an exogenous parameter, such that the balanced trade condition becomes $E_j = \phi Y_j$. This can be introduced both in our nominal and our real model. However, this is still an exogenously determined trade imbalance. If we want an endogenous trade imbalance we need to introduce dynamics into the model, something outside the scope of this paper.

AvW, allows Y_j to be endogenous and it is updated by p_j . Typically one assumes Y_j is p_j times the real endowment of country j .

However, the trade barrier b_{ji} actually contributes to the income of country j , see also AvW, because according to the market clearance condition (8) we sum over all nominal exports $X_{ji}^j = x_{ji}p_{ji}^j$ and (5) shows that p_{ji}^j depends on b_{ji} ; so Y_j depends on the export-weighted trade barrier. So, we should take the trade barrier into account when calculating Y_j to remain consistent with the model. We can interpret this as an additional income for the exporting country, for example a transport sector who transports the good to other countries or a government that collects the export tariffs. This is different from some of the interpretations in the literature that identify the trade cost barrier as iceberg trade cost, implying that the number of goods ‘melts’ with a factor b_{ji} when exported from country j to i . However, this invalidates the balanced trade condition.

We have now derived the complete system of equations underlying the nominal gravity model, i.e. (7) subject to (9), (10), (11) and (12).

2.4 The nominal gravity equation

To finally obtain the nominal gravity equation we take the exports demand equation (7) and substitute in the equilibrium prices (9) for the price ratio and balanced trade (12) for E_j arriving at

$$X_{ij}^j = w_i^y \left(\frac{b_{ij}/S_{ij}}{\Pi_i P_j} \right)^{1-\sigma} Y_j, \quad (13)$$

subject to (10) and (11). Again we can return to AvW’s version of the structural gravity model by assuming a common currency, $S_{ij} = 1$.

Now we check the dimensions. Earlier we derived that P_j has the dimension € and Π_i has the dimension $1/\text{€}$. Consequently, (13) is in € on both sides of the equality sign. So, all dimensions are correct in the new structural gravity equation.

Like AvW, we distinguish the same three different driving forces of bilateral exports in (13). The first driving force is output of country j . If Y_j increases it has more money to spend on all goods including good i leading to more bilateral exports. Note that this is the only term that is not relative in the gravity equation and is responsible for the scale of exports.

The second driving force is the relative trade cost between country i and j , $\left(\frac{b_{ij}/S_{ij}}{\Pi_i P_j} \right)$. If it increases bilateral exports decreases. On its turn this term consists of four different factors. The first two are the exporter and importer MRs. If either increases, the trade cost term decreases and bilateral exports increase. To understand this a priori

counterintuitive result, consider an increase in Π_i of country i . Given b_{ij}/S_{ij} , this increase is due to an increase in the resistance between i and the RoW. The RoW demands less of good i decreasing its price, see (9). Hence, country j shifts its demand towards the now relatively cheaper good i ; so exports increase. A similar reasoning holds for P_j .

The last two variables in the trade cost term are b_{ij} and S_{ij} . They both have a direct and indirect effect on bilateral exports. The direct effect of b_{ij} causes a decrease in bilateral exports because it makes good i more expensive shifting the consumption of country j towards cheaper goods. If S_{ij} increases, a depreciation of currency i , the direct effect causes bilateral exports to go up because country i becomes cheaper for country j . The indirect effects work through Π_i and P_j and the indirect effects of b_{ij} and S_{ij} have a positive and negative effect on bilateral exports, respectively, as discussed in Section 2.2. So, the direct and indirect effects work in opposite directions. Together, Π_i and P_j are homogeneous of degree one in b_{ij} and S_{ij} ,⁵ so the indirect effects together offset the direct effect in case b_{ij} or S_{ij} is increased with the same amount for all country-pairs. So, b_{ij} and S_{ij} are both homogeneous of degree zero in (13). Hence, it is not an absolute rise in all b_{ij} or S_{ij} that affects bilateral exports, but a relative rise with respect to the RoW. So, if the trade cost between all country-pairs fall with the same amount due to, say, the introduction of the internet, bilateral exports will not be affected. However, if country i signs an FTA with a group of countries bilateral exports is affected. These countries start trading more with each other and less with non-member countries because the bilateral trade cost are now relatively lower between members than with non-members. Similarly, non-members will trade more with each other and less with member countries.

The third and final driving force is the output share of country i . If w_i^y increases, bilateral exports increase because an increase in the output share of country i will lower the price of good i relative to other goods and country j will substitute its consumption towards good i ; so exports increase.

3 Real gravity

We will now develop a real structural gravity equation from the nominal gravity model in Section 2.4. Focusing on real exports has several motivations (except for the new

⁵This can be seen by taking (10) and divide both sides by Π_i such that we get $1 = \left[\sum_j \left(\frac{b_{ij}/S_{ij}}{\Pi_i P_j} \right)^{1-\sigma} \theta_j \right]^{\frac{1}{1-\sigma}}$. Now we see that, for the RHS to be homogeneous of degree zero in the bilateral trade cost or the exchange rate, like the LHS, $\Pi_i P_j$ must be homogeneous of degree one in the bilateral trade cost.

insights it will deliver later on). First, in the nominal model it is common to take Y_j as constant, as discussed in Section 2.3. Instead the real gravity equation takes real endowment y_j as fixed and lets Y_j automatically adjust in a model-consistent way, that is, by price and trade-barrier changes, as motivated in Section 2.3.

Second, as the empirical focus has shifted from cross-sectional data to time-series data, it is important to adjust for inflation, as is argued by Anderson & Yotov (2010). They show that it is necessary to convert to a real currency, in their paper to Alberta 1992 dollars. Our real gravity equation does this automatically and there is no longer the need to choose an arbitrary deflator.

3.1 Export demand

The main idea to make the gravity model real is to focus on the quantity of bilateral exports, x_{ij} , instead of the value of bilateral exports, X_{ij}^j . Therefore, we use that $x_{ij} = X_{ij}^j/p_{ij}^j$, where p_{ij}^j is given by (5).

We start from the demand equation (7). To make the nominal gravity model real, we divide both the LHS and RHS by $p_i b_{ij}/S_{ij}$ and replace the nominal by the real exchange rate (RER) $s_{ij} = S_{ij}P_j/P_i$

$$x_{ij} = \lambda_i \left(\frac{\frac{p_i}{P_i} b_{ij}}{s_{ij}} \right)^{-\sigma} c_j. \quad (14)$$

So, in the real demand equation the quantity of bilateral exports depends on relative prices and the aggregate consumption index.

3.2 Equilibrium prices

Next, we use the market clearance condition (8) to obtain the relative equilibrium price

$$\frac{p_i}{P_i} = \left(\frac{w_i^y}{\lambda_i} \right)^{\frac{1}{1-\sigma}} \frac{1}{\pi_i}, \quad (15)$$

where

$$\pi_i = \left[\sum_j \left(\frac{b_{ij}}{s_{ij}} \right)^{1-\sigma} w_j^e \right]^{\frac{1}{1-\sigma}} \quad (16)$$

and $\pi_i = \Pi_i P_i$ is the real exporter MR. This is consistent with the dimension of Π_i ($1/\text{€}$).⁶

⁶We can also start from the consumer problem (1)-(2) in quantities. This leads to the same real demand equation. Subsequently, the market clearance condition (8) in quantities leads to the same equilibrium relative price condition if we properly account for the difference between world output and

3.3 Balanced trade

In Section 2.3 we use balanced trade to solve the gravity model. So, we take nominal balanced trade (12) and rewrite it into a real version. We start by rewriting Y_j using (8), (5) and $y_j = \sum_i x_{ji}$ so that $Y_j = P_j^x y_j$, where P_j^x is the exporter price index which is given by

$$P_j^x = p_j \psi_j, \quad (17)$$

where ψ_j is the average barrier to export for country j

$$\psi_j = \sum_i b_{ji} w_{ji}^x, \quad (18)$$

and $w_{ji}^x = x_{ji} / \sum_i x_{ji}$ is the bilateral export share of good j to country i with respect to y_j .

Next we use that nominal expenditure is linked to real expenditure by $E_j = P_j c_j$, so real balanced trade becomes

$$c_j = \frac{P_j^x}{P_j} y_j, \quad (19)$$

where the ratio P_j^x / P_j is the ToT of country j (including the prices of exports to and imports from itself). Hence, real consumption increases after a ToT improvement.

3.4 The real gravity equation

We have now derived the complete system of equations underlying the real gravity model. To obtain the real gravity equation we take the real export demand equation (14) and substitute in both the real equilibrium prices (15) for the price ratio and real balanced trade (19) for c_j , where P_j^x is given by (17), arriving at

$$x_{ij} = \left(\frac{w_i^y \lambda_j}{\lambda_i w_j^y} \right)^{\frac{-1}{1-\sigma}} w_i^y \left(\frac{b_{ij}}{s_{ij} \pi_i} \right)^{-\sigma} \frac{\psi_j}{\pi_j} y_j, \quad (20)$$

subject to (16) and (18). Note that π_j is the new importer MR, but now from the supply side. The importer MR from the demand side P_j in the trade cost term is absorbed by the real exchange rate. The exporter and importer MR no longer depend on one and other, like in the nominal equation. The MRs are untangled. This is another advantage of the real gravity equation.

The real gravity equation consists of six driving forces of which three are new compared to the nominal gravity equation (13). The first three driving forces, that is,

world consumption.

y_j , $b_{ij}/(s_{ij}\pi_i)$ and w_i^y are the same as in the nominal gravity equation. The first has the same role as E_j in (13), as discussed in Section 2.4. Again it is the only variable that is in levels, in quantities this time, while all other variables are relative. The latter two are discussed in Section 2.4.

The fourth driving force, w_i^y/λ_i , reflects changes in the relative equilibrium price for good i p_i/P_i , see (15). If the output share of country i increases the relative price for good i decreases, so bilateral exports increase. If λ_i increases, the taste for good i increases and this leads to an increase in demand for good i and consequently a rise in its price, so bilateral exports decrease.

Finally, the fifth and sixth driving force, that is, ψ_j/π_j and λ_j/w_j^y capture the ToT effect of country j $P_j^x/P_j = (w_j^y/\lambda_j)^{\frac{1}{1-\sigma}} \psi_j/\pi_j$, where we consecutively substitute in (17) and (15) to obtain the RHS expression. If the ToT increases, the purchasing power of country j goes up, so bilateral exports increase. To see the individual contribution of each variable to the ToT, we first consider ψ_j/π_j . If ψ_j increases, the barriers to export increase and this leads to an increase in the exporter price index of country j , so the ToT increases and bilateral export increases. Contrary, if π_j increases it becomes harder to export for country j and demand for good j falls resulting in a lower ToT, so bilateral exports decrease. Next, consider λ_j/w_j^y . If λ_j increases, the price of good j goes up, so the ToT increases and bilateral exports increase. If output share of country j increases, the price for good j goes down, so the ToT decreases and bilateral exports decrease.

4 Empirical model specification and estimation method

In Section 1 we argued that it is better to use the real gravity equation for an empirical model. Therefore, we will introduce our empirical model specification based on our theoretical real structural gravity equation from Section 3. First, we derive the log-linear representation and use it to obtain an empirical model specification, where we add a rich set of fixed effects (FE). Next, we extend the untangling normalization method, as proposed by Klaassen & Teulings (2015), such that we can apply it to our *ijt* panel data model and untangle the FE. Finally, we discuss the data.

4.1 Loglinear real gravity and the real effective exchange rate

To be able to estimate the real gravity equation with OLS, we take the log of (20).⁷ This yields the log of π_i , π_j , and ψ_j , given by (16) and (18). Log-linearizing the former results in

$$\log(\pi_i) = \sum_j (\log(b_{ij}) - \log(s_{ij})) w_j^e, \quad (21)$$

where we assume that for each i the linearization is around a point that does not depend on the importing countries j . For the countries in our sample b_{ij} and s_{ij} are relatively stable and therefore the approximation error is most likely reasonably small. We do not log-linearize around the weighting variable w_j^e . In a similar way we approximate π_j and ψ_j .

We obtain two new insight from (21): $\log(\pi_i)$ depends on both the weighted average of $\log(b_{ij})$ and $\log(s_{ij})$, where w_j^e serves as weight. The weighted average of $\log(s_{ij})$ is the log of the real effective exchange rate (REER) of country i , where an increase implies that the currency of country i depreciates vis-à-vis the RoW. If the REER of country i increases it becomes cheaper for the RoW to buy good i , increasing its competitiveness; so $\log(\pi_i)$ decreases. Similarly, the weighted average of $\log(b_{ij})$ is the average log trade barrier of country i . If it increases vis-à-vis the RoW, country i becomes less competitive and $\log(\pi_i)$ increases.⁸

Next, we replace $\log(\pi_i)$, $\log(\pi_j)$ and $\log(\psi_j)$ by their log-linear approximations and we arrive at

$$\begin{aligned} \log(x_{ij}) = & \frac{-1}{1-\sigma} \left((\log(w_i^y) - \log(\lambda_i)) + (\log(\lambda_j) - \log(w_j^y)) \right) + \log(w_i^y) \\ & - \sigma \left(\log(b_{ij}) - \log(s_{ij}) - \sum_j (\log(b_{ij}) - \log(s_{ij})) w_j^e \right) \\ & + \sum_i (\log(b_{ji}) w_{ji}^x - \log(b_{ji}) w_i^e + \log(s_{ji}) w_i^e) + \log(y_j). \end{aligned} \quad (22)$$

The nominal gravity equation can be log-linearized in a similar way.

⁷Takings logs does not allow for zero trade flows. However, our data does not exhibit zero trade flows and therefore this worry can be ignored.

⁸In a similar way we can approximate the MR Π_i and P_j in (13) using log-linearization. We find that Π_i and P_j depend on the average trade cost barrier and the nominal effective exchange rate and the weighted global consumer price indexes P_j and the weighted global competitive measure Π_i , respectively. Substituting the log-linearized expression of P_j into Π_i and vice versa, results in a demeaned approximation of Π_i and P_j independent from each other.

4.2 Model specification

We use (22) to come up with an empirical model. We consider an ijt -panel with N country-pairs and T time periods. This results into

$$\begin{aligned} exp_{ijt} = & \beta_1 FTA_{ijt} + \gamma_1 \overline{FTA}_{it}^e + \gamma_2 (\overline{FTA}_{jt}^x - \overline{FTA}_{jt}^e) + \\ & \gamma_3 gdp_{it} + \gamma_4 gdp_{jt} + \gamma_5 gdp_{wt} + \gamma_6 reer_{it} + \gamma_7 reer_{jt} + \\ & \alpha + \alpha_i^x + \alpha_j^m + \alpha_{ij} + \tau \cdot t + \tau_i^x \cdot t + \tau_j^m \cdot t + \tau_{ij} \cdot t + \theta_t + \theta_{it}^x + \theta_{jt}^m + \varepsilon_{ijt}. \end{aligned} \quad (23)$$

where normalization of the FE will be discussed, in Section 4.4. The dependent variable exp_{ijt} is (the logarithm of) the real bilateral exports from country i to country j at time t .

We use an FTA dummy variable to capture the trade barrier $\log(b_{ij})$. The $\overline{FTA}_{it}^e$ and $\overline{FTA}_{jt}^x$ represent the geometrically weighted average of $\log(b_{ij})$ using expenditure and export weights, respectively. We split up the w_i^y , w_j^y and y_j into gdp_{it} , gdp_{jt} and gdp_{wt} , where all are in logs. Hence, $\gamma_3 = -\sigma/(1-\sigma)$, $\gamma_4 = (2-\sigma)/(1-\sigma)$ and $\gamma_5 = 1$. Finally, $reer_{it}$ and $reer_{jt}$ capture the exporter and importer specific REER in logs, respectively. The RER s_{ij} is also captured by these two variables, because triangular arbitrage implies that $\log(s_{ij}) = reer_i - reer_j$. Therefore, both REER variables capture both the direct and indirect RER effect through the MR. So $\gamma_6 = 0$ and $\gamma_7 = 1 - \sigma$.

We add an extensive set of FE. In total we have three FE-families, α , τ , and θ . Each FE-family targets a specific dimension of unobserved heterogeneity. On its turn each FE-family consists of different FE-types, ranging from homogeneous FE-type, say, α to completely heterogeneous FE-type, say, α_{ij} .

The α -family captures the overall, exporter, importer, and country-pair specific effects by α , α_i^x , α_j^m , and α_{ij} , respectively, where we list the FE-types from most homogeneous to most heterogeneous. These FE capture intercept differences, driven by selection into trading effects, due to, for example, a country being landlocked or cultural differences between a country pair.

The τ -family captures the overall, exporter, importer, and country-pair trends by $\tau \cdot t$, $\tau_i^x \cdot t$, $\tau_j^m \cdot t$, and $\tau_{ij} \cdot t$, respectively. Not including these FE-types might lead to an omitted variable bias as shown by [Bun & Klaassen \(2007\)](#).

The θ -family captures time, exporter-time, and importer-time specific effects by θ_t , θ_{it}^x and θ_{jt}^m , respectively. In this way we capture all global-, exporter-, and importer-specific developments like oil shocks and country-specific economic crises.

Finally, we assume that the error term ε_{ijt} has a zero mean conditional on the regressors at all times. We assume it is not cross-sectionally correlated and allow it to

be heteroscedastic and serially correlated. We estimate the empirical model by OLS.

4.3 Identifying the elasticity of substitution

The elasticity of substitution σ is important for deriving the welfare effects of different trade policies, as is shown by [Arkolakis *et al.* \(2012\)](#). However, it is difficult to directly estimate σ from the nominal gravity equation. Potentially σ can be identified through the bilateral trade cost b_{ij} or the exporter or importer MR. However, b_{ij} consist of multiple variables, where each has a separate coefficient as well. So, we only identify a combination of the elasticity with other coefficients. It is also not possible to use both MRs to identify σ , since we do not have any variable that proxies the MRs reliably. Different identification strategies exist in the literature, although most are not straightforward or rely on data that is difficult to obtain; see [Anderson & van Wincoop \(2004\)](#) and [Head & Mayer \(2014\)](#) for an overview of different strategies and σ estimates in the literature. The gravity literature usually finds that $\sigma > 1$.

However, with our real gravity approach it is possible to estimate the elasticity of substitution through three different channels in (23). All three channels rely on widely available variables.

The first channel is through gdp_{it} . This implies that we expect to find $\gamma_3 = -\sigma/(1 - \sigma) > 1$, if $\sigma > 1$. Similarly, the second channel is through gdp_{jt} . This implies that $\gamma_4 = (2 - \sigma)(1 - \sigma) < 1$, if $\sigma > 1$. Finally, the last channel is through $reer_{jt}$. This implies that $\gamma_7 = 1 - \sigma < 0$, if $\sigma > 1$.

4.4 Untangling normalization

In (23) we have perfect multicollinearity between different FE-types and between (country-)time FE and “constant” regressors, that is, regressors that are constant in one or more dimensions and vary in the remaining. We have different types of constant regressors: three constant it -regressors, three constant jt -regressors and one constant t -regressor. These constant regressors are perfect multicollinear with the FE-types θ_{it}^x , θ_{jt}^m and θ_t , respectively. In (23) we have one regressor that varies over all three dimension, so called ijt -regressor. To allow us to identify the FE and the impact of the constant regressors we need normalizations.

Untangling normalization disentangles FE dummies and constant regressors. [Klaassen & Teulings \(2015\)](#) develop this method for an it panel data model, see their paper for a more extensive discussion of untangling normalization. Here we will shortly discuss the general idea of untangling normalization. Next, we will show how to extend it to an ijt panel data model. We denote untangled parameters with a u superscript and

constant regressors with the common variable z and common parameter γ .

4.4.1 General idea

Untangling normalization uses orthogonality conditions as normalizations to prevent multicollinearity between two sets of variables. This allows us to obtain parameter estimates for both set of variables, that would otherwise be impossible due to multicollinearity. For example, the vector of ones and exporter dummies are multicollinear, preventing us from estimating α and α_i^x simultaneously. Untangling normalization orthogonalizes the set of variables that are most heterogeneous, the exporter dummies, to the set of variables that is most homogeneous, the vector of ones, implying that we normalize the mean of α_i^{xu} to zero. So the parameter α^u captures the overall intercept, while the parameters α_i^{xu} capture the exporter deviation on the overall intercepts.

Untangling normalization has two advantages over the commonly used zero normalization, where one normalizes some parameters to zero to prevent multicollinearity. First, it improves interpretation by assigning to each FE-type an effect orthogonal to the other effects; the effects are untangled. To clarify this, consider the example above, where we assign the overall mean to α^u and the exporter deviation to α_i^{xu} . If we use zero normalization instead, we are not able to separate these two different effects and assign them to different parameters. For example, if we normalize α to zero, the remaining α_i^x parameters capture the exporter deviation plus the overall mean making the effects ‘tangled’ and the interpretation more difficult.

Second, untangling normalization allows us to exploit constant regressors. For example, we have a country specific effect that consists of $\alpha_i^x + z_i'\gamma$; so the i -regressors is multicollinear with the exporter dummies. Untangling normalization allows us to exploit the i -regressors, by orthogonalizing z_i with respect to the exporter dummies. We prevent multicollinearity and obtain parameter estimates for the constant regressors. The α_i^{xu} capture the exporter deviation that is unexplained by the constant regressors. This is not possible if we apply zero normalization.

The untangled constant-regressor parameter of the i -regressors are not the true but the pseudo-true value, due to the normalizations. This is because we can estimate $\alpha_i^x + z_i'\gamma$, but we cannot identify γ from this. However, if $\alpha_i^{xu} = 0$ for all i , then $\alpha_i^{xu} + z_i'\gamma^u = z_i'\gamma^u$ and $\alpha_i^x + z_i'\gamma$ is completely explained by $z_i\gamma^u$, so $\gamma^u = \gamma$ and we identify the true value.

4.4.2 Untangling normalization in an ijt panel data model

Now we show how we can extend untangling normalization to a three-dimensional ijt panel data model, like (23). We do not allow for all possible FE-configurations in our three-dimensional model, but only for the ‘sensible’ FE-configurations. A sensible FE-configurations satisfies the general rule that if we include, say, α_i^x , all more homogeneous FE-types are included as well, so we also include α , since the former encompasses the latter. This allows for the complete untangling of all different effects in the model and assigning them to separate parameters. For example, we assign the overall intercept to α and the exporter deviations on the overall intercept to α_i^x . If we do not include both, we hamper the interpretation of α_i^x because they capture both the exporter and overall effects. So way we apply the general idea of untangling normalization to assign each effect to a separate FE-type, see Section 4.4.1. This rule is straightforwardly applied to the α - and τ -families, because there is no multicollinearity possible between both FE-families.

However, the θ -family varies both over countries and over time, resulting in interdependence with the α - and the τ -families, respectively. For example, θ_t can capture both the overall intercept and trend. So this implies that if we include, say, θ_t into the model, we also must include α , $\tau \cdot t$ into the model.

In Table A.1 in the Appendix we summarize all possible untangling normalizations in an ijt panel data model. It shows for each FE-type in the leftmost column which normalization this FE-type enforces on other FE-types, if they are present in the FE-configuration as well. For example, consider the top row of our table. If we include α into our model equation, we have to normalize the sum of α_i^x over i , the sum of α_j^m over j , and the sum of θ_t over t to zero, if they are present in the FE-configuration as well.

Note that, if we have a FE-type that varies over two dimensions, say, θ_{it}^x and perform a normalization on one of the two dimensions, for example we normalize the sum over i to zero, we do not perform one but T normalizations. If we now perform a second normalization on θ_{it}^x , for example we normalize the sum over t to zero, we lose one degree of freedom, so we have $T + J - 1$ normalizations. For every additional normalization on θ_{it}^x we lose another degree of freedom.

4.5 Data

Our data set consists of 18 countries and therefore $N = 306$ country-pairs. We include all EU-15 countries, except for Belgium and Luxembourg⁹, Canada (CN), Japan (JP), Norway (NW), the United States (US) and Switzerland (SW). We use annual data from 1965-2011 ($T = 47$) resulting into 14382 observations.

Real exports is measured in exporter baskets of good i valued in 2010 dollars. To compute this we start with the monthly nominal bilateral export data from the IMF Direction of Trade Statistics (DOTS). The export data is converted in Home (i) currency using the exchange rate from the International Financial Statistics (IFS) of the IMF. We sum to get yearly aggregates and convert it from nominal to real using the exporter price index of country i from the AMECO database of the European Commission. Finally, we divide by the purchasing power parity (PPP) exchange rate of i currency versus US dollars in the baseyear 2010, obtained from the OECD Economic Outlook. This ensures a common scaling of all bilateral export flows.

The FTA dummy variable is similar to the one used by [Baier & Bergstrand \(2007\)](#). Table A.2 in Appendix A shows the different FTAs we take into account and all entry and exit dates of its members.

The average FTA variables are constructed using geometric weighted averages, where the weights are constructed using expenditure and export. We construct expenditure weights using nominal expenditure from AMECO. Nominal expenditure is transformed into national currency with the AMECO exchange rate. We divide by the AMECO expenditure deflator to obtain real Expenditure. Export weights are constructed using above constructed annual bilateral exports.

Nominal yearly GDP data from AMECO is transformed to the national currency, using the AMECO exchange rate. Next, we construct real GDP using the AMECO GDP deflator. Similar to real export, GDP is scaled by dividing by the PPP of the US dollar in i currency. We impute German GDP of before 1991 from West-German GDP.

The monthly REER, based on the consumer-price index, is taken from the Bank for International Settlements (BIS) and averaged to construct yearly data and inverted.

⁹ Austria (OE), Denmark (DK), Finland (FN), France (FR), Germany (BD), Greece (GR), Ireland (IR), Italy (IT), the Netherlands (NL), Portugal (PT), Spain (ES), Sweden (SD), United Kingdom (UK).

5 Estimating the real gravity equation

5.1 Real versus nominal

Before we go to the estimation results for the real gravity equation, we will show why we consider a real and not a nominal gravity equation. To compare the performance of a real versus nominal gravity equation we first derive an estimation equation for the nominal gravity equation from (13). We do this in the same way as with the real equation in Section 4.1, that is, we derive the log-linear representation, and add the FE

$$\begin{aligned} Exp_{ijt}^j = & \beta_1 FTA_{ijt} + \gamma_1 \overline{FTA}_{it}^e + \gamma_2 \overline{FTA}_{jt}^e + \\ & \gamma_3 GDP_{it} + \gamma_4 GDP_{jt} + \gamma_5 GDP_{Wt} + \gamma_6 NEER_{it} + \gamma_7 NEER_{jt} + \\ & \alpha + \alpha_i^x + \alpha_j^m + \alpha_{ij} + \tau \cdot t + \tau_i^x \cdot t + \tau_j^m \cdot t + \tau_{ij} \cdot t + \theta_t + \theta_{it}^x + \theta_{jt}^m + \varepsilon_{ijt}, \end{aligned} \quad (24)$$

where Exp_{ijt}^j is nominal exports in country j currency in logs¹⁰, GDP_{it} , GDP_{jt} and GDP_{Wt} are the nominal GDP of country i , j and the world in logs, respectively, and finally $NEER_{it}$ and $NEER_{jt}$ are the nominal effective exchange rates (NEER) of country i and j in logs, respectively. All variables are in logs.

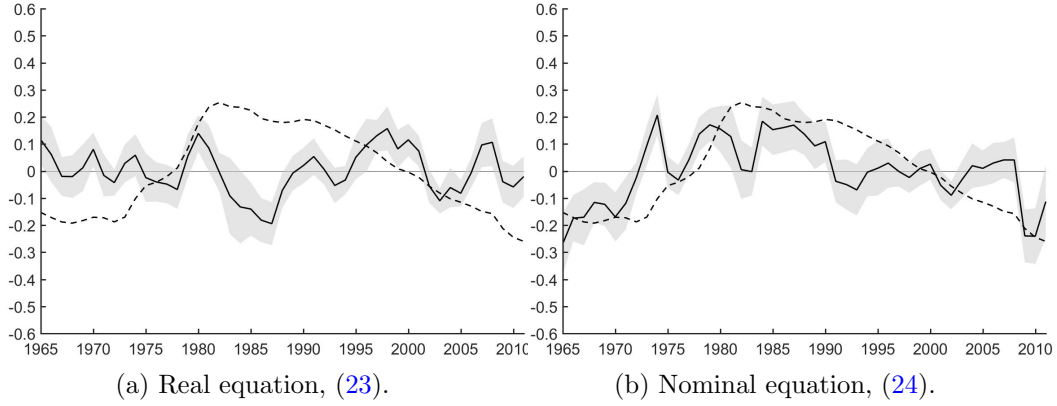
The nominal gravity equation (13) induces three restrictions on (24). First, $\gamma_3 = -\gamma_5$, because together they form w_i^y in the nominal gravity equation. While in the real gravity equation there are three GDP variables w_i^y , w_j^y and y_j . So the nominal equation loses one degree of freedom compared to the real equation. We clearly reject this restriction, with a t-statistic of 156.

The second and third restrictions are $\gamma_6 = \gamma_7 = 0$, since triangular arbitrage implies that $S_{ijt} = NEER_{it} - NEER_{jt}$. For each estimate this restriction is rejected, with a t-statistic of 9 and 14, respectively. So the nominal equation is too restrictive.

We estimate both the real and nominal equation using (23) and (24), respectively, where we restrict $\gamma_3 = -\gamma_5$ in the nominal model, and plot $\theta_t^u + \theta_{jt}^{mu}$. In Figure 1 we show the results for the US, which are indicative for the general picture. Substituting θ_{jt}^{mu} for θ_{it}^{xu} does not change the result. The dashed line represents $CPI_{US,t}$. We detrend and demean $CPI_{US,t}$ and multiply it with the ratio of the standard deviation of $\theta_t^u + \theta_{US,t}^{mu}$, estimated with the nominal equation, over the standard deviation of $CPI_{US,t}$. The latter is for ease of comparison. If we compare the real estimates in Figure 1a with the nominal

¹⁰In the literature it is common to denominate all variables in current US dollars. If we do this only γ_7 will change because it will capture $S_{US,jt}$. We deviate from this practice because we want to stay as close to our theoretical nominal gravity model (13) as possible.

Figure 1: Untangled ($\theta_t^u + \theta_{US,t}^{mu}$) of the US in a real and nominal gravity equation.



The dashed line represents detrended and demeaned $CPI_{US,t}$. For ease of comparison it has been scaled, that is, the ratio of the standard deviation of $(\theta_t^u + \theta_{US,t}^{mu})$, estimated with the nominal equation, over that of the standard deviation of $CPI_{US,t}$.

estimates in Figure 1b, we see that the nominal estimates display great similarity with the US inflation pattern while the real equation does not show any inflation pattern. This is not a problem if we are only interested in the impact of FTA_{ijt} or any other ijt -regressor, because then the FE are nuisance parameters. However, when we want to better understand the MRs and want to open the black box of the FE, we need to get rid of the inflation effects because they dominate all other effects. This is what we want to do in the remainder of this paper and therefore we will use the real gravity equation.

Baldwin & Taglioni (2006) warn for the so called bronze medal mistake. They show that it is hard to find the right deflators for different variables like export and real GDP. The wrong deflators will cause measurement errors and therefore a bias in the estimates. So they advice to always use nominal variables because the (country-)time FE will capture inflation. However, the restriction imposed by the nominal model is rejected and we want to get a better understanding of the MRs and of what the FE actually capture. Letting the (country-) time FE capture inflation will hamper this effort.

5.2 Empirical results versus theory

Now we will present the estimation results of estimating (23). We will compare our empirical estimation results from Table 1 with the theoretical implications of the real gravity equation from Section 3. We will pay separate attention to the identification of the elasticity of substitution σ . We use both a static, Model 1 and 2, and dynamic model, Model 3.

Table 1: Estimating the real gravity equation for exp_{ijt} based on (23)

Model Description	1 No z	2 Static	3 DL	4 DL, Hetero.
FTA_{ijt}	0.13 * (0.03)	0.13 * (0.03)	0.29 * (0.04)	0.29 * (0.03)
$\overline{FTA}_{it}^e$		-0.05 (0.08)	-0.09 (0.08)	-0.39 * (0.09)
$\overline{FTA}_{jt}^x - \overline{FTA}_{jt}^e$		0.18 * (0.05)	0.13 * (0.05)	0.09 (0.06)
gdp_{it}		0.73 * (0.07)	0.62 * (0.07)	-0.34 * (0.11)
gdp_{jt}		1.07 * (0.06)	1.01 * (0.06)	1.01 * (0.10)
gdp_{Wt}		-0.93 * (0.12)	-1.18 * (0.14)	-0.65 * (0.20)
$reer_{it}$		0.29 * (0.04)	0.35 * (0.04)	1.06 * (0.06)
$reer_{jt}$		-0.40 * (0.04)	-0.51 * (0.04)	-0.58 * (0.06)
Wald tests				
$\theta_{it}^x = 0$	5.58 * [0.00]	5.03 * [0.00]	5.13 * [0.00]	3.19 * [0.00]
$\theta_{jt}^m = 0$	4.42 * [0.00]	2.64 * [0.00]	2.47 * [0.00]	2.13 * [0.00]
$R_{\theta^x}^2$	—	0.09	0.09	0.64
$R_{\theta^m}^2$	—	0.40	0.54	0.70

For the DL model we included 10 phase-in lags for FTA_{ijt} and 2 lags for every other variable. The heterogeneous parameter model is a DL model where we allow for heterogeneity over both the exporter and importer for the level variables, except for FTA_{ijt} , $\overline{FTA}_{it}^e$ and $\overline{FTA}_{jt}^x - \overline{FTA}_{jt}^e$ which have homogeneous parameters. We only report the mean effect. The reported Wald statistics are divided by the number of restrictions and we relate them to an F-distribution. $R_{\theta^x}^2$ ($R_{\theta^m}^2$) is the fraction of the variance of the untangled exporter-time (importer-time) FE from a model without it -regressors (jt -regressors) that is explained once detrended it -regressors (jt -regressors) are included, respectively. Standard errors are between brackets and they are based on Newey & West (1987, 1994), which gives three lags. * indicates significance at 5% level, that we use throughout the paper. P-values are between square brackets.

We have included one ijt -regressor. The FTA_{ijt} dummy captures the bilateral trade barrier and its impact is not affected by the normalization of the FE and the inclusion of constant regressors. In Model 1 we estimate the impact of FTA_{ijt} without exploiting any constant regressors (z). We find that export increases with $([\exp(0.13) - 1] * 100 =) 14\%$. However, the estimated effect is substantially smaller than the 58% increase in export found by [Baier & Bergstrand \(2007\)](#). [Teulings \(2017\)](#) shows that this difference is mostly explained by the inclusion of $\tau_{ij} \cdot t$ into our model, this dampens the impact of FTA_{ijt} .

In Model 3 we allow for dynamics in the FTA_{ijt} effect, as is done by among others [Baier & Bergstrand \(2007\)](#), to capture the gradual process of the implementation of an FTA, so called phase-in. We include 10 lagged first differences, ensuring that the level variable captures the overall effect. We find an overall effect of 0.29 resulting in a 32% increase in exports. In a similar model, [Baier & Bergstrand \(2007\)](#) find an increase in export of 114%. Again the difference between their and our estimate is the inclusion of $\tau_{ij} \cdot t$, as is shown by [Teulings \(2017\)](#). He finds a 44% increase in exports, similar to our estimate, and that the $\tau_{ij} \cdot t$ are needed for a proper model specification, see also Section 6.3. In the remainder of this section our main focus is on Model 3. We find similar results for Model 2.

Model 3 is a distributed lag (DL) model where we include lags in the form of lagged first differences, such that the level variable captures the overall impact and is comparable to the static model. In Table 1 we only report the level effects. For all it - and jt -regressors we include two lags. Dynamics in trade are important, even though the standard gravity model does not account for it. For example, [Klaassen & Teulings \(2015\)](#) show that allowing for distributed lags helps to better explain exports from 17 OECD countries to the US.

We include it - and jt -regressors while keeping the exporter-time and importer-time FE at the same time, using untangling normalization to avoid multicollinearity. As explained in Section 4.4, we do not necessarily estimate the true values of the impacts of the constant regressors. Only if $\theta_{it}^{xu} = 0$ and $\theta_{jt}^{mu} = 0$ separately, we identify the true value of the impact of the it - and jt -regressors, respectively.

Testing the first restriction, the reported Wald test for Model 3 in Table 1 rejects the null hypotheses, compared to a critical value of 1.11 in an F-distribution with 722 degrees of freedom in the numerator and 11642 in the denominator. We can only explain a small part of θ_{it}^{xu} , as is quantified by the $R_{\theta_x}^2 = 9\%$. The Wald-statistic reduces substantially if we allow for heterogeneous parameters in Model 4, although we still reject the null hypothesis (see Section 6.2 for an more extensive discussion of Model

4). So we cannot conclude that the estimates of the it -regressors reflect the true value. However, the large Wald-statistic seems mainly driven by large outliers, considering that only 146 out of the 810 θ_{it}^{xu} estimates are significant and only 65 have t-statistic larger than 3. Furthermore, the $R_{\theta_x}^2$ increases substantially in Model 4, explaining 64% of the variation in θ_{it}^{xu} of Model 1. This let us tentatively conclude that the pseudo-true estimates are quite close to their true estimates in both Model 4 and 3, since the latter has similar estimates as the former. See also [Klaassen & Teulings \(2015\)](#) for a more extensive discussion on the identification of the impact of constant regressors.

Testing the second restriction, the reported Wald test for Model 3 in Table 1 rejects the null hypothesis. Although the inclusion of the jt -regressors leads to a strong decrease in the Wald-statistic and an increase in the $R_{\theta_m}^2 = 54\%$, so that importer constant regressors explain quite some of the importer-time FE. If we allow for heterogeneous parameters in Model 4, the Wald-statistic reduces further, although the reduction is minor. Still we reject the null hypothesis. So, again we cannot conclude that the estimates of the jt -regressors reflect the true value. However, the estimated θ_{jt}^{mu} are even closer to zero than the estimated θ_{it}^{xu} , so the rejection of the null-hypothesis is driven by large outliers. Only 88 out of the 810 θ_{jt}^{mu} estimates are significant and only 22 have a t-statistic larger than 3. Furthermore, the $R_{\theta_x}^2 = 70\%$. This let us tentatively conclude that the pseudo-true estimates are quite close to their true values in Model 4 and 3, since the latter are similar to the former.

Apart from the direct FTA effect, our real gravity approach motivates two new FTA-based regressors. The indirect FTA effect $\overline{FTA}_{it}^e$ stems from the exporter MR π_i . Theory predicts a negative impact on export because when country i signs an FTA with the RoW demand for good i by the RoW increases causing a price increase and crowding out import of good i by country j . However, in Model 3 we find an insignificant estimate.

The $\overline{FTA}_{jt}^x - \overline{FTA}_{jt}^e$ variable is a contraction of the exporter weights from the average export barrier ψ_j and the MR π_j , respectively. Together they measure the change in the ToT when country i signs an FTA with the RoW. Theory shows that there are two opposing effects. Consider country j signs an FTA with the RoW. First, the competitiveness of country j increases, increasing $\overline{FTA}_{jt}^e$, and this leads to an increase in its exports and therefore its purchasing power increases. On the other hand, trade cost revenue for country j decreases, decreasing $\overline{FTA}_{jt}^x$, and this erodes its purchasing power. We find a significant estimate of 0.13, this implies a 14% increase in exports. Our positive estimate shows that the former dominates the latter suggesting that the ToT of country j improves allowing country j to import more of good i .

The estimated impacts of gdp_{it} , gdp_{jt} and gdp_{wt} are all significant and the signs are consistent with theory. The estimated impact of the latter does not deviate significantly from its theoretical value of -1.

Finally, we discuss the exchange rate effects. The estimated impact of $reer_{it}$ is significantly positive and a 1% real depreciation leads to a 0.35% increase in exports. This estimate captures two effects, the direct effect through $s_{ij} = reer_i - reer_j$, and the indirect effect through the MR π_i . Consider an increase in $reer_{it}$. First, a depreciation of currency i vis-à-vis currency j (lowering the exchange rate) makes good i cheaper for j increasing exports from i to j . Second, a depreciation of currency i vis-à-vis the RoW makes good i cheaper for the RoW, increasing demand and price of i , which crowds out imports of good i by country j . Theory predicts that these two effects cancel each other out. However, we find that the former dominates the later. This suggests that there is room for improvement in real gravity, just as there were three improvements possible in the nominal approach. Apparently, real gravity has resolved two of those three by accounting for the endogeneity of Y_j .

The $reer_{jt}$ estimate is significantly negative and a depreciation of 1% lowers exports by 0.51%. Like with $reer_{it}$, the estimated impact captures both the direct effect and indirect exchange rate effect. First, a depreciation of currency j vis-à-vis currency i makes good i more expensive for j decreasing exports from i to j . Second, a depreciation vis-à-vis the RoW through the MR π_j makes good j cheaper for the RoW and generates more income for country j , so it can buy more of good i . Theory predicts that the former dominates the latter. This is in line with our result.

5.2.1 Estimating the elasticity of substitution

[Anderson & van Wincoop \(2004\)](#) discuss the different estimation results for σ in the literature and conclude it is somewhere between five and ten. More recently [Head & Mayer \(2014\)](#) review a wide range of studies and they find for studies who use the exchange rate to identify σ a median estimate of 2.12 and for all papers that use a structural gravity approach a median σ estimate of 4.78.

In Section 4.3 we discussed three new channels through which we can identify σ in the real gravity equation. The first two channels to identify σ are through gdp_{it} and gdp_{jt} . The theoretical model in Section 3 predicts an impact of $\gamma_3 = -\sigma/(1 - \sigma)$ and $\gamma_4 = (2 - \sigma)/(1 - \sigma)$, respectively. So, the results from Table 1 leads to $\hat{\sigma} = -1.64$ (0.51) and $\hat{\sigma} = -184$ (2165), respectively, where the numbers in brackets are the standard errors based on the delta method. The first is significantly negative and goes against the common result in the literature that $\sigma > 1$, the latter is insignificant and the large

standard error suggests weak identification of σ via gdp_{jt} . If we allow for heterogeneous parameters in Model 4 we find $\hat{\sigma} = 0.25$ (0.06) through the gdp_{it} channel. Although this estimate is positive, it still contradicts the common result in the literature that $\sigma > 1$ and furthermore the impact of gdp_{it} is negative contrary to the theoretical predictions. Finally, the reported standard error is that of the mean effect γ_3 but does not take into account that we allow for heterogeneity (see for more information Section 6.2) over both the exporter γ_{3i}^x and importer γ_{3j}^m . To account for this we calculate the standard errors for all 306 country pairs and take the average. We find a standard error of 0.22, so, if we account for the correct standard error, $\hat{\sigma}$ is insignificant. It seems that the GDP variables are not so useful to identify σ .

The last channel is through $reer_{jt}$. The theoretical model predicts an impact of $\gamma_7 = 1 - \sigma$. Hence, we find $\hat{\sigma} = 1.51$ (0.04). This estimate is in line with [Head & Mayer \(2014\)](#).

6 Capturing the multilateral resistance terms

AvW stress the importance of the MRs. Our empirical model (23) fully controls for them by exporter-time and importer-time FE. The exporter-time FE have been normalized by untangling normalization, as explained in Section 4.4. The resulting θ_{it}^{xu} are orthogonal to other FE and the constant regressors, here all it -regressors. So they capture in a clean way what the model does not explain in the it -dimension. Likewise, the the untangled importer-time FE θ_{jt}^{mu} capture the unexplained part in the jt -dimension of our model.

Now, we will take a closer look to the untangled θ_{it}^{xu} and θ_{jt}^{mu} estimates and investigate how much is explained by the constant regressors implied by gravity theory. Next, we will show that we can better explain these untangled FE by allowing for parameter heterogeneity. Finally, we will investigate the generally accepted conclusion that we need to include exporter- and importer-time FE when we estimate a gravity equation. Is it really needed to include that many FE, or is it sufficient to impose linearity on their development over time?

6.1 Untangled country-time fixed effects

In the sub-figures in the left column of Figure 2 and 3 we display the estimated θ_{it}^{xu} and θ_{jt}^{mu} for Model 1, respectively. We group countries that show similar patterns over time and ignore confidence bands to improve the visibility of the figures. To give an idea about the size of the standard errors we report the median of the standard errors

Figure 2: Untangled exporter-time FE θ_{it}^{xu} for Model 1 and 4 in Table 1.

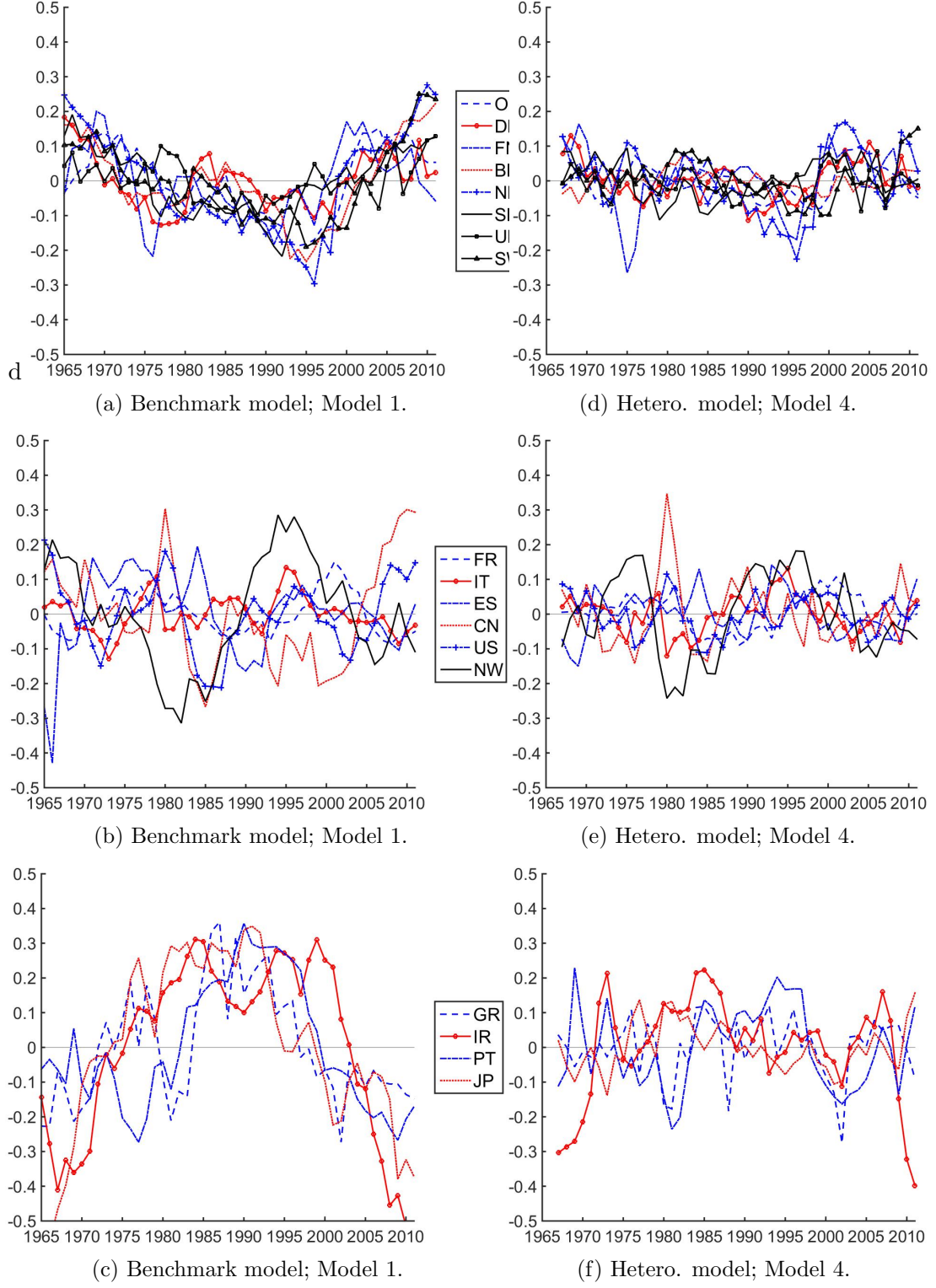
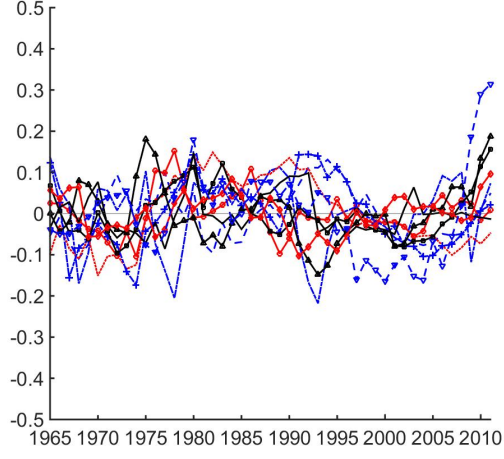
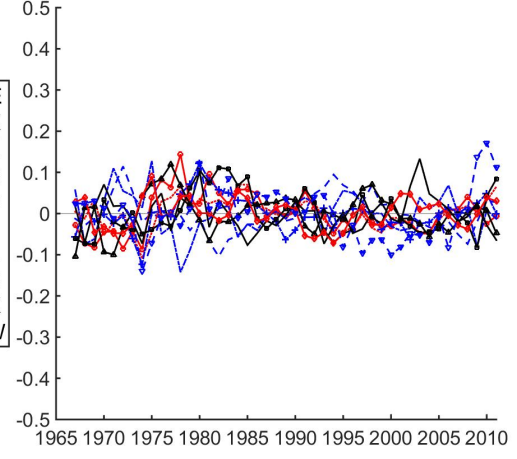


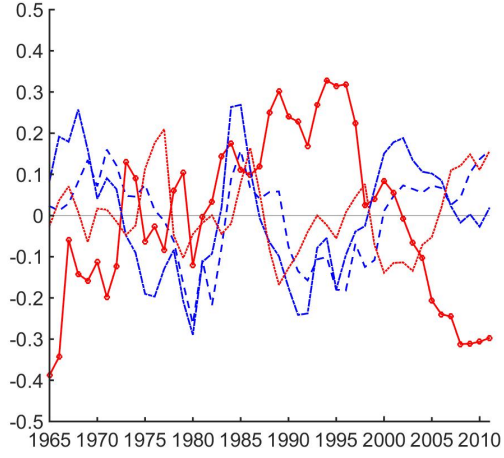
Figure 3: Untangled importer-time FE θ_{jt}^{mu} for Model 1 and 4 in Table 1.



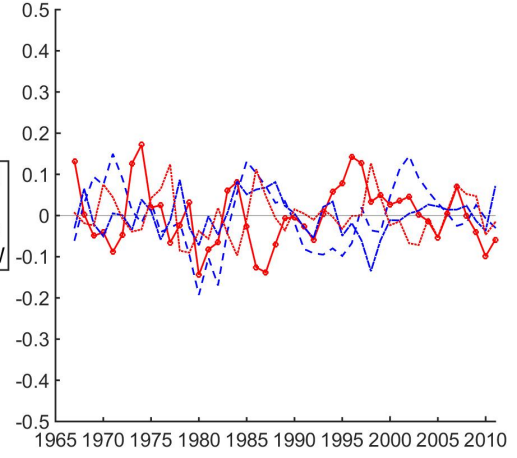
(a) Benchmark model; Model 1.



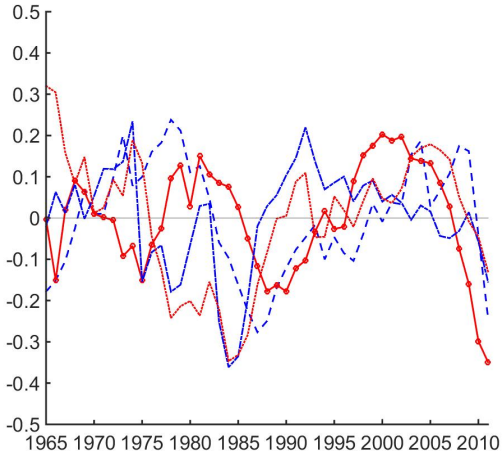
(d) Hetero. model; Model 4.



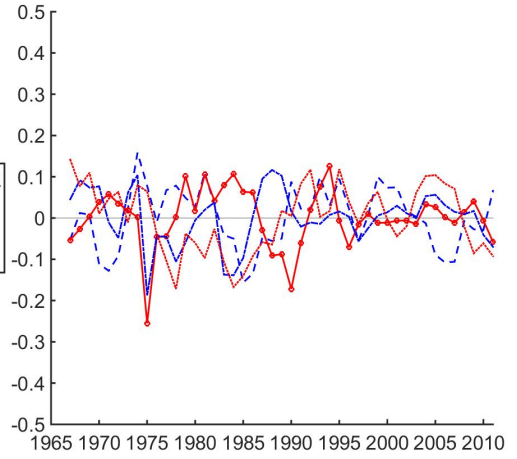
(b) Benchmark model; Model 1.



(e) Hetero. model; Model 4.



(c) Benchmark model; Model 1.



(f) Hetero. model; Model 4.

for each country in Table A.3 of Appendix A.

Figure 2a shows that the estimated θ_{it}^{xu} display a v-shape for a group of Northern- and Western-European and Scandinavian countries, except for Ireland and Norway. This shape implies that they export less than based on all other model fundamentals from about the mid seventies to the end of the nineties. This turns around in the beginning of the 2000s.

In Figure 2b we display the estimated θ_{it}^{xu} for a disperse group of countries. For France and Italy the estimates fluctuate closely around zero, while the remaining four countries are varying much more widely. Canada and the US are moving relatively jointly probably because they are both geographically and economically very close, explaining similar estimates for θ_{it}^{xu} .

In Figure 2c we display four countries with a distinctive hump-shape. This shape indicates that these countries export less than average in the first half of the sample, more than average from the end of the seventies to the end of the nineties and finally again less than average in the 2000s. This hump-shape is the reverse of the v-shape in the first group, although it is way more pronounced. For Greece, Ireland and Portugal a possible explanation for this shape might be that entering the EU for these countries was a serious boost to export, but when they changed their own currency for the Euro they lost their competitiveness. An other explanation lies in the resemblance of this hump-shape with the worldwide inflation pattern. It is possible that exports of these four countries suffered more from worldwide inflation than other countries. Further research is needed. The estimated θ_{it}^{xu} of Model 2 and 3 are close to that of Model 1.

In Figure 3a we can see a group of ten countries, all EU countries except Switzerland, that have θ_{jt}^{mu} estimates fluctuating closely around zero.

In Figure 3b we display the estimated θ_{jt}^{mu} for a disperse group of four non-EU countries that widely vary around zero. Again Canada and the US are moving relatively closely together.

In Figure 3c we display the four PIGS countries that display the same distinct pattern, especially from the end of the eighties they all increase, importing more than based on the other fundamentals, until they sharply plummet at the start of the financial crisis; so they are hit harder by the financial crisis than other countries. Again the PIGS countries might see a sharp additional increase in import due to better access to other European markets.

So both θ_{it}^{xu} and θ_{jt}^{mu} have substantial unexplained country-time deviations, adding *it*- and *jt*-regressors explains these deviations partially.

6.2 Heterogeneous parameters

A possible further improvement in explaining the remaining country-time deviation might be by the introduction of heterogeneous parameters. [Chen & Novy \(2011\)](#) show that there is substantial heterogeneity across industries. [Herwartz & Weber \(2013\)](#) show that the euro effect on export and the MR terms are both subject to cross-sectional heterogeneity both across sectors and country pairs. Time specific parameter heterogeneity seems not as important as heterogeneity over countries.

We allow for parameter heterogeneity over both the exporter and importer and let one homogeneous parameter capture the mean effect. That is, we substitute $z'\gamma$ by $z'\gamma + \sum_i z'_i \gamma_i^x + \sum_j z'_j \gamma_j^m$ and we normalize $\sum_i \gamma_i^x = 0$ and $\sum_j \gamma_j^m = 0$ to ensure that γ captures the mean effect and that there is no multicollinearity.

In Model 5 in Table 1 we estimate a DL model, similar to Model 3, where we allow all level variables to have heterogeneous parameters except for the parameters of FTA_{ijt} , $\overline{FTA}_{it}^e$ and $\overline{FTA}_{jt}^x - \overline{FTA}_{jt}^e$. The FTA_{ijt} is a dummy variable and therefore does not have enough exporter or importer specific variation for all heterogeneous parameters to be identified, so we decide not to allow for parameter heterogeneity. In Table 1 we only report the mean level effects. In A.4 in Appendix A we report all exporter and importer specific heterogeneous parameter estimates. The θ_{it}^{xu} and θ_{jt}^{mu} estimates of Model 4 are plotted in the sub-figures on the right in Figure 2 and 3, respectively.

Two of the three variables with homogeneous level parameters in Model 4 hardly change compared to Model 3: FTA_{ijt} and $\overline{FTA}_{jt}^x - \overline{FTA}_{jt}^e$. Although the latter becomes insignificant. However, the estimated $\overline{FTA}_{it}^e$ effect is substantially larger and even becomes significant, leading to a 32% decrease in export. This estimate is in line with the predictions by theory, see Section 5.2.

Heterogeneous parameters influence mostly the estimated mean effects of the it -regressors and less so that of the jt -regressors. Possibly because the jt -regressors Model 3 already explain a large part of θ_{jt}^{mu} but the it -regressors not so much of θ_{it}^{xu} .

The estimated mean effect of gdp_{it} becomes negative, opposite from the theoretical prediction. We find positive and significant estimates for France, Portugal and Japan, such that the total negative impact of gdp_{it} becomes positive. The importer specific estimates vary less widely and less are significant. Ireland, the UK and Japan all have positive significant estimates, such that the total impact of gdp_{it} becomes positive.

The overall estimated impact of gdp_{jt} does not change, although we find substantial deviations from the overall estimate. The exporter specific significant estimates are large and widely varying, although for none of the countries the total impact of gdp_{jt} becomes negative. Only for the UK does the negative importer estimate lead to a

negative total impact of gdp_{jt} .

The estimated impact of gdp_{wt} is almost halved and seems to have a widely varying effect on trade. The exporter specific significant estimates are large and widely varying. Germany, the Netherlands, Sweden, Canada, the US and Switzerland have positive significant estimates such that the total impact of gdp_{wt} for these countries is positive. A possible explanation might be that these countries are very open to trade, so an improvement of the world economy will boost their exports, even though this is against the predictions of the gravity model. The importer specific estimates vary less widely. We find positive significant estimates for Finland, Italy and Norway making the total impact of gdp_{wt} positive.

In gdp_{it} and gdp_{wt} we see a grouping of countries for the exporter specific parameters. Germany, the Netherlands, Sweden, Canada, the US and Switzerland, have negative estimates for gdp_{it} and positive for gdp_{wt} . The most of these countries have similar θ_{it}^{xu} estimates, see Figure 2a, except for Canada and the US. They form a separate group with comparable estimates. The second group consists of Finland, Ireland, Italy Portugal, Spain, Japan and Norway, having opposite signs for gdp_{it} and gdp_{wt} compared to the first group. Their θ_{it}^{xu} estimates do not have a clear similar pattern in Model 1.

The estimated impact of $reer_{it}$ becomes three times as large and a 1% depreciation leads to 1.06% increase in export. The total effect for each country is always positive. Thus, the direct exchange rate effect always dominates the indirect effect.

The estimated impact of $reer_{jt}$ becomes slightly stronger and a 1% depreciation leads to a 0.58% decrease in export. The total effects are always negative, except for Spain.

In Figure 2d-2f we plot the estimated θ_{it}^{xu} for Model 4. The v-shape from Figure 2a is mostly gone. In Figure 2e most variations disappear, except for Canada and Norway. This might be explained by their dependence on natural resources. The pronounced hump in Figure 2c disappears in Figure 2f. However, we keep seeing a strong variation around zero.

In Figure 3d-3f the estimated θ_{jt}^{mu} are closer to zero than their counterparts in the sub-figures in the left column. We no longer see distinct patterns in the estimates. So heterogeneous parameter reduce the fluctuations substantially. The plots of the estimated θ_{jt}^{mu} in Model 2 and 3 are similar to that of Model 4.

So allowing for heterogeneous parameters allows us to capture more exporter-time and importer-time specific variation. However, are the elasticities actually heterogeneous or do the heterogeneous parameters capture the effect of unobserved variables.

Especially with the exporter-time variation we see some distinct patterns for different country groups, which we cannot capture by variables suggested by the gravity literature. Better understanding and explaining these patterns is left for future research.

6.3 The importance of trend FE

Table 2: The effect of omitting trend FE, provided we have omitted θ_{it}^x and θ_{jt}^m

Model	2	5	6	7
θ_{it}^{xu} & θ_{jt}^{mu}	Yes	No	No	No
τ_{ij}^u	Yes	Yes	No	No
τ_i^{xu} & τ_j^{mu}	Yes	Yes	Yes	No
FTA_{ijt}	0.13 * (0.03)	0.14 * (0.03)	0.36 * (0.03)	0.51 * (0.03)
$\overline{FTA}_{it}^e$	-0.05 (0.08)	-0.06 (0.08)	-0.36 * (0.10)	-0.35 * (0.08)
$\overline{FTA}_{jt}^x - \overline{FTA}_{jt}^e$	0.18 * (0.05)	0.16 * (0.06)	-0.06 (0.07)	0.02 (0.06)
gdp_{it}	0.73 * (0.07)	0.73 * (0.09)	0.71 * (0.11)	1.87 * (0.07)
gdp_{jt}	1.07 * 0.06	1.07 * (0.08)	1.09 * (0.10)	0.95 * (0.06)
gdp_{wt}	-0.93 * (0.12)	-0.93 * (0.14)	-0.87 * (0.20)	-1.42 * (0.21)
$reer_{it}$	0.29 * (0.04)	0.29 * (0.05)	0.28 * (0.06)	-0.04 (0.05)
$reer_{jt}$	-0.40 * (0.04)	-0.40 * (0.04)	-0.41 * (0.06)	-0.54 * (0.05)

The dependent variable is exp_{ijt} . The note to Table 1 provides further details.

So far we focused on explaining exporter- and importer-specific deviation captured by θ_{it}^{xu} and θ_{jt}^{mu} , respectively. However, we will now investigate the common view in the literature that omitting θ_{it}^x and θ_{jt}^m will seriously bias estimates of trade cost variables such as FTA_{ijt} . Is this bias really caused by unobserved country-time specific deviations or is it actually caused by country specific trends? Therefore, we investigate the importance of trend FE.

Country-time dummies overlap with country trends; they are multicollinear. In the literature $\tau_i^x \cdot t$ and $\tau_j^m \cdot t$ are normalized to zero and one just adds $\tilde{\theta}_{it}^x$ and $\tilde{\theta}_{jt}^m$, as defined in the literature, such that $\tilde{\theta}_{it}^x$ captures both $\tau_i^x \cdot t$ and θ_{it}^x and similarly for $\tilde{\theta}_{jt}^m$. We use this notation to distinguish our untangled θ_{it}^{xu} and θ_{jt}^{mu} from the ones commonly used in the literature.

In Table 2 we show the estimation results for different static model specifications with homogeneous parameters to see the impact of $\tau_i^{xu} \cdot t$, $\tau_j^{mu} \cdot t$ and $\tau_{ij}^u \cdot t$. In all model specifications we exclude θ_{it}^{xu} and θ_{jt}^{mu} . Using the DL model yields similar conclusions.

Model 5 show the parameter estimates if we only omit θ_{it}^{xu} and θ_{jt}^{mu} . We compare

these with the “baseline” estimates of Model 2, repeated here in the first column, and see that the estimates are very similar. This suggests that the exporter-time and importer-time specific deviation in export is not correlated with FTA_{ijt} and does not cause an omitted variable bias in the corresponding estimate.¹¹ Hence, opposite to what is suggested in the literature, we do not find any evidence for the so called Gold medal mistake, failing to account for θ_{it}^{xu} and θ_{jt}^{mu} (see Baldwin & Taglioni (2006)). Still, some other source of unobserved heterogeneity may cause a severe omitted variable bias in economic integration estimates that lets the literature conclude that it is necessary to include $\tilde{\theta}_{it}^x$ and $\tilde{\theta}_{jt}^m$.

In Model 6 we omit $\tau_{ij}^u \cdot t$. The estimated FTA_{ijt} parameter increases almost three-fold indicating that the estimate has a severe bias due to omitting $\tau_{ij}^u \cdot t$. This result is in line with Bun & Klaassen (2007), Baier *et al.* (2014) and Bergstrand *et al.* (2015), who all show the importance of including $\tau_{ij}^u \cdot t$ into the model. Still, this does not explain why the literature finds a substantial omitted variable bias when they omit $\tilde{\theta}_{it}^x$ and $\tilde{\theta}_{jt}^m$, because $\tau_{ij}^u \cdot t$ does not overlap with $\tilde{\theta}_{it}^x$ and $\tilde{\theta}_{jt}^m$ and captures a different part of unobserved heterogeneity. This explains why the estimates of the constant regressors remain relatively unchanged except for $\overline{FTA}_{it}^e$ and $\overline{FTA}_{jt}^x - \overline{FTA}_{jt}^e$, who are affected by the bias in FTA_{ijt} due to the high correlation between these variables.

If we omit $\tau_i^{xu} \cdot t$ and $\tau_j^{mu} \cdot t$ in Model 7, we find a substantial additional bias in FTA_{ijt} . The constant regressors are also affected. The three most heavily affected variables are gdp_{it} , gdp_{wt} , increasing substantially, and $reer_{it}$, becoming negative and insignificant.

So, it appears that the bias in economic integration estimates is not caused by the Gold medal mistake, but failing to account for $\tau_i^{xu} \cdot t$ and $\tau_j^{mu} \cdot t$. This explains why the literature finds that $\tilde{\theta}_{it}^x$ and $\tilde{\theta}_{jt}^m$ are so important. But it actually is the linear-trend part of the $\tilde{\theta}_{it}^x$ and $\tilde{\theta}_{jt}^m$ that matters, not so much the non-linear deviations from that trend. So we add $2 \cdot 18 \cdot 47 = 1692$ parameters where just $2 \cdot 18 = 36$ parameters suffice. If it is indeed the case that $\tau_i^{xu} \cdot t$ and $\tau_j^{mu} \cdot t$ are important causes for the omitted variable bias instead of θ_{it}^{xu} and θ_{jt}^{mu} , then we need a better theoretical understanding why we need trends within the gravity framework. This is left for future research.¹²

¹¹Note that it is not informative to look at the estimated parameters of the constant regressors, since the untangling normalization already orthogonalizes the exporter-time and importer-time specific FE regarding these regressors.

¹²Note that we do not argue that θ_{it}^{xu} and θ_{jt}^{mu} are redundant. In Section 6 we show that there is country-time unobserved heterogeneity, however we argue that these unobserved country-time deviations do not matter for the estimated impact of FTA variable. They are a kind of random effects.

7 Conclusion

We derive a real gravity equation based on a standard nominal gravity equation. This presents us with four new insights. First, the importer MR in the form of the CPI no longer plays a role, as it is absorbed in order to deflate nominal variables. Instead in the real gravity equation a new importer MR is introduced from the supply side. The higher this importer MR, the less the importing country earns from exports, so it can import less. Second, the ToT is introduced into the real gravity equation to capture the purchasing power of the importing country. Third, the real gravity equation needs the RER. We show that the real gravity equation depends both directly on the RER as well as indirectly through the exporter and importer MRs. The latter results in weighted aggregates of the exchange rate and we show that they are a proxy for the REER of the exporter and importer, respectively. Finally, it is possible to identify the elasticity of substitution through three different channels and the most promising route is by the combination of the direct and indirect exchange rate effect of the importer.

We use the real gravity equation to estimate bilateral exports between 18 OECD countries in the time period 1965-2011. To allow us to exploit the constant regressors proposed by the real gravity equation we use the untangling normalization method. Therefore, we extend this method from an *it*- to *ijt*-panel data model. We show that the nominal model imposes three restrictions (GDP_{it} and GDP_{Wt} have the same absolute impact and exchange rates do not matter for exports), which are all rejected by the data. Real gravity takes away two of those restrictions except for $reer_{it}$, which is still rejected by the data. So there is room for further improvement. The other estimates of the real model are in line with the theoretical predictions and we find an elasticity of substitution estimate of 1.51.

We also investigate how much the proposed MR variables can explain from the country-time specific deviations. Using untangling normalization, we exploit these constant regressors and let the country-time FE capture the remaining unexplained country-time deviation. This unexplained part is quite substantial. If we allow for heterogeneous parameters we eventually explain 64 and 70% of the exporter-time and importer-time deviation, respectively. The remaining unexplained country-time deviation still shows some distinct patterns, indicating that further gains can be made in the gravity literature.

Finally, we investigate the consequences of omitting country-time FE for the estimated impact of the free trade agreement (FTA), an *ijt*-regressor. We find that this does not bias the estimate, even though the exporter-time and importer-time FE esti-

mates are non-zero. So the exporter-time and import-time deviations are not important for explaining the impact of FTA. Instead we find that a substantial bias arises if we omit the country-pair trend and the exporter trend and importer trend FE. So trends are important, but there is not yet a theoretical justification for adding trend variables into the gravity model. This would be an important addition to the gravity model.

References

- Anderson, James E., & van Wincoop, Eric. 2003. Gravity with Gravitas: A Solution to the Border Puzzle. *The American Economic Review*, **93**(1), pp. 170–192.
- Anderson, James E., & van Wincoop, Eric. 2004. Trade Costs. *Journal of Economic Literature*, **42**(3), pp. 691–751.
- Anderson, James E., & Yotov, Yoto V. 2010. The Changing Incidence of Geography. *The American Economic Review*, **100**(5), pp. 2157–2186.
- Anderson, James E., Larch, Mario, & Yotov, Yoto V. 2015. *Estimating General Equilibrium Trade Policy Effects: GE PPML*. Tech. rept.
- Anderson, James E., Vesselovsky, Mykyta, & Yotov, Yoto V. 2016. Gravity with scale effects. *Journal of International Economics*, **100**, pp. 174–193.
- Arkolakis, Costas, Costinot, Arnaud, & Rodriguez-Clare, Andres. 2012. New Trade Models, Same Old Gains? *American Economic Review*, **102**(1), pp. 94–130.
- Baier, Scott L., & Bergstrand, Jeffrey H. 2007. Do free trade agreements actually increase members’ international trade? *Journal of International Economics*, **71**(1), pp. 72–95.
- Baier, Scott L., Bergstrand, Jeffrey H., & Feng, Michael. 2014. Economic integration agreements and the margins of international trade. *Journal of International Economics*, **93**(2), pp. 339–350.
- Baldwin, Richard, & Krugman, Paul. 1989. Persistent Trade Effects of Large Exchange Rate Shocks. *The Quarterly Journal of Economics*, **104**(4), pp. 635–654.
- Baldwin, Richard, & Taglioni, Daria. 2006. *Gravity for Dummies and Dummies for Gravity Equations*. NBER Working Papers 12516. National Bureau of Economic Research, Inc.
- Bergstrand, Jeffrey H. 1985. The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence. *The Review of Economics and Statistics*, **67**(3), pp. 474–481.
- Bergstrand, Jeffrey H., Larch, Mario, & Yotov, Yoto V. 2015. Economic integration agreements, border effects, and distance elasticities in the gravity equation. *European Economic Review*, **78**, pp. 307–327.

- Bun, Maurice J. G., & Klaassen, Franc J. G. M. 2007. The Euro Effect on Trade is not as Large as Commonly Thought. *Oxford Bulletin of Economics and Statistics*, **69**(4), pp. 473–496.
- Campa, José M., & Goldberg, Linda S. 2005. Exchange Rate Pass-Through into Import Prices. *The Review of Economics and Statistics*, **87**(4), pp. 679–690.
- Chen, Natalie, & Novy, Dennis. 2011. Gravity, trade integration, and heterogeneity across industries. *Journal of International Economics*, **85**(2), pp. 206 – 221.
- Head, Keith, & Mayer, Thierry. 2014. *Gravity Equations: Workhorse, Toolkit, and Cookbook*. Handbook of International Economics, vol. 4. Elsevier. Chap. 3, pages pp. 131–195.
- Herwartz, Helmut, & Weber, Henning. 2013. The role of cross-sectional heterogeneity for magnitude and timing of the euro’s trade effect. *Journal of International Money and Finance*, **37**(October), pp. 48–74.
- Kenneth A. Froot, Paul D. Klemperer. 1989. Exchange Rate Pass-Through When Market Share Matters. *The American Economic Review*, **79**(4), 637–654.
- Klaassen, Franc J. G. M., & Teulings, Rutger. 2015 (December 24). *Untangling Fixed Effects and Constant Regressors*. Tech. rept. 15-137/VI. Tinbergen Institute Discussion Paper.
- Newey, Whitney K., & West, Kenneth D. 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, **55**(3), pp. 703–708.
- Newey, Whitney K., & West, Kenneth D. 1994. Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies*, **61**(4), pp. 631–653.
- Sutherland, Alan. 2005. Incomplete pass-through and the welfare effects of exchange rate variability. *Journal of International Economics*, **65**(2), pp. 375–399.
- Teulings, Rutger. 2017 (Aug.). *Brexit and The Impact of Gradual Economic Integration on Export*. Tinbergen Institute Discussion Papers 17-075/VI. Tinbergen Institute.

A Additional tables

Table A.1: Normalization decision rule per FE-type in an ijt panel.

FE to be normalized															
α-family					τ-family			θ-family			const. regressors (z)				
α	α _i ^x	α _j ^m	α _{ij}	τ	τ _i ^x	τ _j ^m	τ _{ij}	θ _t	θ _{it} ^x	θ _{jt} ^m	z _i	z _j	z _t	z _{it}	z _{jt}
α	∑ _i	∑ _j						∑ _t							
α _i ^x			∑ _j						∑ _t						
α _j ^m			∑ _i							∑ _t					
α _{ij}															
τ					∑ _i	∑ _j		∑ _t							
τ _i ^x							∑ _j		∑ _t						
τ _j ^m							∑ _i			∑ _t					
τ _{ij}															
θ _t									∑ _i	∑ _j					
θ _{it} ^x															
θ _{jt} ^m															
z _i	∑ _i														
z _j		∑ _j													
z _{ij}			∑ _{i,j}												
z _t								∑ _t							
z _{it}									∑ _{i,t}						
z _{jt}														∑ _{j,t}	

This table shows for each FE-type (row) what sum of other FE-types (columns) it restricts to zero, provided that they are present in the FE-configuration. For example, if τ is present we have to normalize the sum of τ_i^x over i , of τ_j^m over j , and of $t \cdot \theta_t$ over t to zero, provided that they are present in the FE-configuration.

Table A.2: FTA starting dates and participating countries to define FTA_{ijt}

FTA name	Start	Participating countries
European Union (EU)	1958	Austria (1995), Denmark (1973), Finland (1995), France, Germany, Greece (1981), Ireland (1973), Italy, Netherlands, Portugal (1986), Spain (1986), Sweden (1995), United Kingdom (1973)
European Free Trade Association (EFTA)	1960	Austria (until 1995), Denmark (until 1973), Finland (1986-1995), Portugal (until 1986), Sweden (until 1995), United Kingdom (Until 1973), Norway, Switzerland
EU-EFTA	1973	Agreement between all EU and EFTA members Finland enters in 1974 as associate EFTA member
CUSFTA	1989	United States and Canada
European Economic Area (EEA)	1994	New agreement between all EU and EFTA members
North American Free Trade Agreement (NAFTA)	1994	United States, Canada

Number between brackets are entry or exit dates of individual countries.

Table A.3: Median of the standard errors for Figures 2 and 3

Median std. err.	θ_{it}^{xu}		θ_{jt}^{mu}	
	1	5	1	5
OE	0.03	0.03	0.04	0.03
DK	0.04	0.03	0.04	0.03
FN	0.04	0.04	0.04	0.04
FR	0.03	0.03	0.03	0.02
BD	0.02	0.02	0.03	0.03
GR	0.06	0.05	0.06	0.05
IR	0.05	0.05	0.05	0.04
IT	0.03	0.02	0.03	0.03
NL	0.03	0.03	0.03	0.03
PT	0.04	0.04	0.05	0.04
ES	0.04	0.04	0.04	0.03
SD	0.03	0.03	0.03	0.03
UK	0.03	0.03	0.04	0.03
CN	0.06	0.05	0.05	0.04
JP	0.05	0.04	0.04	0.04
US	0.04	0.03	0.04	0.03
NW	0.06	0.05	0.05	0.04
SW	0.03	0.02	0.04	0.04

Table A.4: Heterogeneous parameter estimates of model 4 in Table 1

Variable	gdp_{it}		gdp_{jt}		gdp_{wt}		$reer_{it}$		$reer_{jt}$	
Hetero. over	i	j	i	j	i	j	i	j	i	j
Mean	-0.34 * (0.11)		1.01 * (0.10)		-0.65 * (0.20)		1.06 * (0.06)		-0.58 * (0.06)	
OE	0.21 (0.25)	0.30 (0.28)	0.10 (0.16)	-0.18 (0.34)	0.64 (0.37)	0.29 (0.45)	1.06 * (0.22)	-0.21 (0.14)	-0.22 * (0.09)	-0.18 (0.32)
DK	-0.87 * (0.33)	0.37 (0.27)	0.62 * (0.20)	-0.26 (0.35)	-0.41 (0.43)	-0.09 (0.43)	0.07 (0.23)	0.08 (0.15)	-0.16 (0.14)	-0.10 (0.25)
FN	0.18 (0.24)	-0.31 (0.26)	0.16 (0.21)	-0.15 (0.24)	-1.63 * (0.44)	1.08 * (0.48)	-0.10 (0.15)	0.07 (0.12)	-0.34 * (0.14)	0.41 * (0.15)
FR	0.47 * (0.21)	0.08 (0.13)	0.03 (0.14)	0.99 * (0.19)	0.01 (0.29)	-1.63 * (0.29)	0.18 (0.18)	0.08 (0.08)	0.20 * (0.08)	-0.34 (0.20)
BD	-0.98 * (0.17)	0.11 (0.17)	0.00 (0.14)	0.10 (0.19)	2.44 * (0.27)	-1.26 * (0.35)	0.23 * (0.11)	-0.04 (0.09)	0.06 (0.08)	0.12 (0.15)
GR	-0.47 (0.25)	-1.01 * (0.27)	-0.66 * (0.30)	0.78 * (0.19)	1.10 (0.88)	0.11 (0.72)	1.14 * (0.22)	0.39 * (0.14)	-0.22 (0.16)	0.12 (0.20)
IR	-0.40 * (0.19)	0.94 * (0.31)	0.42 (0.32)	0.21 (0.16)	-4.19 * (0.62)	-1.72 * (0.36)	1.45 * (0.22)	-0.24 (0.17)	0.15 (0.19)	0.48 * (0.17)
IT	0.04 (0.16)	-0.61 * (0.18)	0.57 * (0.16)	0.10 (0.17)	-1.64 * (0.35)	1.80 * (0.38)	-0.72 * (0.10)	0.21 (0.11)	0.14 (0.08)	0.29 * (0.12)
NL	-0.56 * (0.25)	-0.16 (0.22)	-0.44 * (0.16)	-0.62 * (0.30)	4.12 * (0.35)	-0.68 (0.45)	-0.12 (0.14)	0.13 (0.12)	0.24 * (0.11)	-0.46 * (0.17)
PT	1.60 * (0.18)	-0.59 (0.33)	-0.68 * (0.25)	0.15 (0.19)	-3.68 * (0.58)	0.13 (0.57)	-0.49 * (0.15)	-0.46 * (0.21)	-0.14 (0.17)	-0.44 * (0.17)
ES	0.24 (0.30)	-0.16 (0.20)	0.46 (0.26)	0.33 (0.26)	-1.05 * (0.48)	-0.40 (0.54)	-0.34 * (0.14)	-0.07 (0.12)	0.12 (0.13)	0.65 * (0.15)
SD	0.05 (0.28)	-0.20 (0.19)	-0.44 * (0.19)	0.16 (0.25)	2.06 * (0.46)	0.12 (0.49)	-0.03 (0.13)	-0.25 * (0.10)	-0.12 (0.11)	0.16 (0.15)
UK	0.30 (0.31)	0.66 * (0.23)	-0.49 * (0.20)	-1.53 * (0.42)	0.31 (0.41)	1.03 (0.53)	-0.74 * (0.11)	-0.30 * (0.14)	0.05 (0.11)	0.28 * (0.14)
CN	-1.48 * (0.46)	0.27 (0.23)	0.36 (0.25)	-0.20 (0.35)	2.50 * (0.72)	-0.34 (0.55)	-1.14 * (0.18)	0.28 * (0.13)	0.37 * (0.17)	0.02 (0.15)
JP	2.31 * (0.19)	0.69 * (0.28)	1.07 * (0.31)	0.00 (0.17)	-1.89 * (0.52)	-1.13 * (0.54)	-0.49 * (0.10)	0.08 (0.13)	-0.20 (0.13)	0.10 (0.10)
US	0.17 (0.33)	0.57 (0.29)	-0.64 * (0.21)	-0.03 (0.42)	0.87 (0.38)	0.15 (0.41)	-0.66 * (0.12)	-0.06 (0.12)	0.00 (0.12)	-0.53 * (0.14)
NW	-0.09 (0.45)	-1.12 * (0.29)	-0.23 (0.36)	-0.06 (0.28)	-1.50 * (0.67)	2.04 * (0.59)	0.91 * (0.34)	0.06 (0.16)	0.02 (0.23)	-0.20 (0.27)
SW	-0.70 * (0.26)	0.17 (0.29)	-0.21 (0.15)	0.21 (0.41)	1.93 * (0.36)	0.50 (0.61)	-0.22 * (0.11)	0.25 (0.14)	0.06 (0.08)	-0.36 (0.19)

The dependent variable is exp_{ijt} . The heterogeneous parameter model is a DL model where we allow for heterogeneity over both the exporter and importer for the long term variables, except for FTA_{ijt} , FTA_{it} and $FTA_{jt}^x - FTA_{jt}$, which have homogeneous parameters. Here we only report the heterogeneous parameter estimates and their mean counterparts. For the other table notes see Table 1.